DGK Deutsche Geodätische Kommission
der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 766

Clémence Dubois

Interferometric Synthetic Aperture RADAR and Radargrammetry towards the Categorization of Building Changes

München 2016

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission beim Verlag C. H. Beck

ISSN 0065-5325 ISBN 978-3-7696-5178-2

DGK Deutsche Geodätische Kommission
der Bayerischen Akademie der Wissenschaften

Reihe C Dissertationen Heft Nr. 766

Interferometric Synthetic Aperture RADAR and Radargrammetry towards the Categorization of Building Changes

Zur Erlangung des akademischen Grades eines
Doktor-Ingenieurs (Dr.-Ing.)
von der Fakultät für Bauingenieur-, Geo- und Umweltwissenschaften
des Karlsruher Instituts für Technologie (KIT)
genehmigte Dissertation

von

Dipl.-Ing. Clémence Dubois
aus Saint-Lô (Frankreich)

München 2016

Verlag der Bayerischen Akademie der Wissenschaften
in Kommission beim Verlag C. H. Beck

ISSN 0065-5325 ISBN 978-3-7696-5178-2

Adresse der Deutschen Geodätischen Kommission:

Deutsche Geodätische Kommission
Alfons-Goppel-Straße 11 ● D – 80 539 München
Telefon +49 – 89 – 23 031 1113 ● Telefax +49 – 89 – 23 031 - 1283 / - 1100
e-mail hornik@dgfi.badw.de ● http://www.dgk.badw.de

Referent:	Prof. Dr.-Ing. habil. Stefan Hinz Institut für Photogrammetrie und Fernerkundung (IPF), Karlsruher Institut für Technologie (KIT)
Korreferent:	Prof. Dr.-Ing. Uwe Sörgel Institut für Geodäsie, Technische Universität Darmstadt

Tag der mündlichen Prüfung: 11.11.2015

Diese Dissertation ist auf dem Server der Deutschen Geodätischen Kommission unter <http://dgk.badw.de/>
sowie auf dem Server des Karlsruher Instituts für Technologie unter
<http://digbib.ubka.uni-karlsruhe.de/volltexte/1000051736> elektronisch publiziert

ISSN 0065-5325

ISBN 978-3-7696-5178-2

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt, die wörtlich oder inhaltlich übernommenen Stellen als solche kenntlich gemacht und die Satzung des KIT zur Sicherung guter wissenschaftlicher Praxis in der jeweils gültigen Fassung beachtet habe.

Karlsruhe, den 15. Oktober 2015

Dubois

Abstract

Change detection in urban areas plays an ever-growing role, be it for construction monitoring in urbanization projects, or for rapid damage assessment at building level after a natural disaster. Remote sensing SAR data are particularly popular for this purpose, as they are deployable independently of daylight and weather.

They present a particular advantage after an earthquake or a storm, where on-site visit is often limited or impossible.

The current generation of satellite SAR platforms such as TerraSAR-X, TanDEM-X and COSMO-SkyMed provides resolutions up to one meter, permitting the detailed analysis of urban areas, whilst covering large zones. Their specific orbital constellations enables, besides the acquisition of single images, the application of other SAR techniques relying on multiple image acquisition, such as SAR interferometry (InSAR) and radargrammetry. Compared to single image analysis, those methods offer the possibility of three-dimensional scene reconstruction, which is of high interest for the analysis of urban areas. InSAR uses the phase difference between acquisitions taken from similar incidence angles, whereas radargrammetry relies on the amplitude disparity between images taken under different incidence angles.

The purpose of this work is the investigation of such techniques for fully automatic and rapid change detection analysis at building level. In particular, the benefits and limitations of a complementary use of InSAR and radargrammetry in an emergency context are examined in term of quickness, globality and accuracy.

First, the potential of InSAR phase data for building detection and reconstruction is inspected. Especially, focus is put on building layovers, arising due to the specific sensor geometry. Their particular appearance in interferometric SAR images permits the development of two different detectors, whose combined utilization in order to segment building candidates is investigated. As sometimes more than one facade is visible for one particular building, a tool for differencing adjoined facades is developed. Based on the extracted facade segments, an algorithm is presented that permits their reconstruction into geometrical shapes, from which building parameters are deduced straightforwardly.

Second, the suitability of radargrammetric data for building reconstruction and change detection at building level is examined. To achieve this goal, data fusion is performed between interferometric and radargrammetric datasets. It aims at the identification of corresponding buildings and the transmission of building parameters for improving the robustness of the radargrammetric method. The developed radargrammetric approach intends to preserve linear structures, as found at building locations. For this, modifications of standard image matching approaches are proposed and analyzed. The resulting appearance of buildings in the radargrammetric disparity map is examined, leading to the recognition of specific geometrical shapes. The consideration of the statistical information and matching parameters permits the extraction of those shapes and the subsequent determination of the building parameters.

Based on the extracted building parameters from both techniques, a change detection approach is developed that aims not only at detecting where changes occurred, but also at giving an indication about its nature. Accordingly, several change categories and change parameters are introduced, and their relevance is investigated.

In this work, the analysis is restricted to isolated, rectangular shaped, middle-rise buildings with flat-roof, as found in many suburbs of metropoles. The efficiency of the proposed methodology is evaluated quantitatively and qualitatively for very high resolution TerraSAR-X and TanDEM-X data of a construction site situated in the north-east of Paris.

Kurzfassung

Die Änderungsdetektion im städtischen Gebiet spielt eine immer größer werdende Rolle, sei es zum Monitoring des Fortschritts bei Neubauten oder Abrissen, für stadtplanerische Zwecke oder zur Schadensanalyse auf Gebäudeebene nach einer Katastrophe. Hierfür sind satellitengestützte SAR Daten besonders geeignet, da sie zu jeder Zeit und bei jedem Wetter eingesetzt werden können. Sie sind z.B. nach einem Erdbeben oder einem Sturm von Vorteil, da eine terrestrische Analyse durch Ortsbegehung oft nur sehr begrenzt möglich ist.

Neben der Möglichkeit einer großflächigen Aufnahme erreicht die derzeitige Generation satellitengestützter Plattformen, wie z.B. TerraSAR-X, TanDEM-X und COSMO-SkyMed, Auflösungen bis zu einem Meter, was die detaillierte Analyse städtischer Gebiete erleichtert. Neben Einzelbildaufnahmen bieten ihre spezifischen Konstellationen die Anwendung weiterer SAR-Techniken, die auf mehrfacher Szenenaufzeichnung beruhen, wie z.B. SAR Interferometrie (InSAR) und Radargrammetrie. Im Vergleich zur Einzelbildanalyse ermöglichen diese Methoden eine dreidimensionale Szenenrekonstruktion, was für die Analyse städtischer Gebiete von besonderem Interesse ist. Bei InSAR wird vor allem die Phasendifferenz zwischen zwei Aufnahmen mit ähnlichem Einfallswinkel genutzt, während bei der Radargrammetrie der Amplitudenversatz zwischen zwei Aufnahmen unterschiedlicher Einfallswinkel analysiert wird.

Ziel dieser Arbeit ist die Erforschung solcher Techniken für eine vollautomatische und schnelle Änderungsdetektion auf Gebäudeebene. Insbesondere die Vorteile und Einschränkungen einer kombinierten Anwendung von InSAR und Radargrammetrie in einer Notsituation werden hinsichtlich Schnelligkeit, Globalität und Genauigkeit untersucht.

Zuerst wird das Potenzial von InSAR-Phasen zur Gebäudedetektion und -rekonstruktion dargelegt. Insbesondere Gebäude-Layover, die durch die spezifische Sensorgeometrie entstehen, werden analysiert. Ihr besonderes Erscheinungsbild im interferometrischen Phasenbild ermöglicht die Entwicklung zweier unterschiedlicher Detektoren, deren kombinierte Verwendung zur Segmentierung von Gebäudehypothesen untersucht wird. Da zum Teil mehr als eine Fassade für ein bestimmtes Gebäude erkennbar ist, wird zusätzlich eine Methode zur Differenzierung angrenzender Fassaden vorgestellt. Basierend auf den extrahierten Fassadensegmenten wird ein Algorithmus entwickelt, der ihre Rekonstruktion in bekannte geometrische Formen durchführt, von denen wiederum Gebäudeparameter abgeleitet werden können.

Darüber hinaus wird die Eignung radargrammetrischer Daten für die Gebäuderekonstruktion und Änderungsdetektion auf Gebäudeebene analysiert. Hierfür erfolgt eine Fusion zwischen interferometrischem und radargrammetrischem Datensätzen, mit dem Ziel der Identifikation von Gebäudekorrespondenzen zwischen den Datensätzen und der Übertragung der Gebäudeparameter zur Verbesserung der Robustheit der radargrammetrischen Methode sowie der späteren

Veränderungsanalyse. Der entwickelte radargrammetrische Ansatz beabsichtigt die Erhaltung linearer Strukturen, die an Gebäuden gefunden werden. Folglich werden übliche Bildmatchingsverfahren modifiziert und ihr Beitrag analysiert. Das resultierende Erscheinungsbild von Gebäuden im radargrammetrischen Versatzbild wird interpretiert, was zur Erkennung spezifischer geometrischer Formen führt. Die Berücksichtigung von statistischer Information und abgeleiteten Matching-Parametern ermöglicht die Extraktion dieser Formen und die darauffolgende Ermittlung der Gebäudeparameter.

Basierend auf den durch beide Techniken extrahierten Gebäudeparametern wird ein Ansatz zur Änderungsdetektion entwickelt, der nicht nur die einfache Detektion von Veränderungen anstrebt, sondern auch deren Art und Umfang beschreibt. Hierfür werden mehrere Veränderungsklassen und -parameter eingeführt, und ihre Relevanz wird untersucht.

In dieser Arbeit beschränkt sich die Analyse auf freistehende, rechteckige Gebäude mit einem Flachdach, die eine mittlere Höhe von ungefähr zehn Stockwerken aufweisen und sich vermehrt in Vororten von Großstädten befinden. Die Effizienz der vorgestellten Methodik wird quantitativ und qualitativ für sehr hochaufgelöste TerraSAR-X und TanDEM-X Daten von einer Großbaustelle im Nord-Osten von Paris bewertet.

Résumé

Que ce soit à des fins d'auscultation de projets d'aménagement urbain, ou bien d'évaluation rapide des dégâts à l'échelle des bâtiments après une catastrophe naturelle, la détection de changements en milieu urbain joue un rôle de plus en plus important dans notre société. Les données de télédétection RSO sont particulièrement appréciées pour cette tâche, car elles peuvent être employées indépendamment de la luminosité et du temps qu'il fait. Elles présentent un avantage notable à la suite d'un tremblement de terre ou d'une tempête, lorsque l'accès aux lieux est souvent limité, voire impossible.

Tout en permettant un grand champ de couverture, l'actuelle génération de plateformes satellites, telles que TerraSAR-X, TanDEM-X ou COSMO-SkyMed, atteint des résolutions allant jusqu'à un mètre, ce qui permet l'analyse détaillée des milieux urbains. Outre l'acquisition d'images simples, leurs constellations orbitales respectives favorisent l'application d'autres techniques de RSO utilisant l'acquisition d'images multiples, telles que l'interférométrie RSO (InSAR) ou la radargrammétrie. Par rapport à l'analyse d'images simples, ces méthodes offrent la possibilité de reconstruire des scènes en trois dimensions, ce qui représente un grand intérêt pour l'analyse des zones urbaines. L'InSAR utilise principalement la différence de phase entre deux images obtenues avec des angles de vue similaires, tandis que la radargrammétrie exploite la disparité entre deux images d'amplitudes obtenues avec des angles de vue différents.

Le but de ce travail est l'investigation de telles techniques en vue d'une détection de changement rapide et totalement automatique à l'échelle des bâtiments. En particulier, les bénéfices et limitations d'un usage combiné de l'INSAR et de la radargrammétrie dans un contexte d'urgence sont analysés en termes de rapidité, globalité et précision.

Dans un premier temps, le potentiel des images de phases InSAR pour détecter et reconstruire les bâtiments est examiné. En particulier, l'attention est portée sur les repliements de façade provenant de la géométrie spécifique propre aux capteurs RSO. Leur apparence spécifique dans les images de phase interférométriques incite au développement de deux détecteurs distincts, dont l'utilisation combinée est étudiée, afin de segmenter des hypothèses de bâtiments. Comme parfois plus d'une façade est reconnaissable pour un bâtiment, un outil est développé permettant la différenciation de façades jouxtantes. Puis, un algorithme est implémenté permettant la reconstruction des segments en formes géométriques, à partir desquelles les paramètres des bâtiments peuvent être déduits de manière directe.

Dans un second temps, l'aptitude des données radargrammétriques à reconstruire les bâtiments et détecter les changements est analysée. Une fusion des données interférométriques et radargrammétriques intervient, ayant pour but l'identification des buildings correspondants entre les deux jeux de données et la transmission des paramètres de bâtiments. Cette dernière vise à la fois l'amélioration de la robustesse de la méthode radargrammétrique et la détection ultérieure de changements. L'approche radargrammétrique développée prévoit la préservation des structures linéaires, comme on peut en trouver au niveau des bâtiments. Pour cela, diverses modifications sont apportées à des approches usuelles de mise en correspondance d'images, et leurs bénéfices examinés. L'apparence résultante des bâtiments dans l'image de disparité radargrammétrique est analysée, conduisant à la reconnaissance de formes géométriques spécifiques. La considération des statistiques de l'image de disparité ainsi que des paramètres de mise en correspondance permet l'extraction de ces formes et la détermination consécutive des paramètres de bâtiments. A partir des paramètres de bâtiments déterminés à l'aide des deux méthodes, une approche pour la détection de changement est développée, visant non seulement à détecter où les changements sont apparus, mais aussi à indiquer le type et un ordre de dimension du changement. Pour cela, différentes catégories et paramètres de changement sont introduits, et leur pertinence est explorée.

Pour cette thèse, l'analyse est limitée à des bâtiments isolés, de forme rectangulaire et à toit plat, ayant une hauteur moyenne d'environ dix étages. On trouve ce type de bâtiments aux périphéries de grandes villes. L'efficacité de la méthodologie proposée est évaluée quantitativement et qualitativement pour des données TerraSAR-X et TanDEM-X très haute résolution, pour un projet d'aménagement urbain situé dans la banlieue Nord-Est de Paris.

Contents

1

Introduction

In recent years, the number of spaceborne Synthetic Aperture RADAR (SAR) systems increased significantly. Nowadays, most of the developed countries possess their own SAR satellite and developing countries cooperate in order to create evermore performing platforms. Concurrently, the technical developments undergo a rapid evolution. Each new SAR satellite or satellite constellation addresses new requirements in terms of resolution, scale of acquisition, configuration, autonomy and/or revisit time.

Especially due to their weather and daylight independence, spaceborne SAR systems permit scene acquisition under most difficult conditions, where optical or airborne systems reach their limits (e.g., cloud coverage, night, stormy weather). Such characteristics make them very attractive in cases of natural disasters, where rapid response is required. Different available resolutions and coverages additionally permit a wide range of applications. Low resolution and large coverage is preferred for flood analysis in rural areas, where the knowledge of the total extent of the damage is more important that the exact boundaries of the water line. Higher resolution is required as soon as detailed and local description is necessary. Current spaceborne missions achieve very high resolutions under 1 m. Obviously, with such products, the study of man-made structures is possible. In particular, a detailed analysis of urban areas after an earthquake or a cyclone provides useful information for the local authorities and international emergency organizations. Knowledge about the current state of buildings is not only mandatory for finding possible victims, but also necessary for estimating which roads are still passable or which buildings could be used as transitional shelter.

The topic of rapid mapping and change detection in urban areas by use of SAR data has been addressed already by several approaches. Many of them deal with change detection at large scale, e.g. at city level, using lower resolution data. The availability of high resolution data since a few years pushes the development of approaches for finer scale analysis, e.g. at building level. Yet, one main challenge those approaches are confronted to is their global applicability, without any external data. Indeed, often no preliminary (pre-event) data exist, and fast change analysis is performed with a single, post-event, SAR image, combined to pre-event ancillary information. This work fits into the topic of rapid mapping and change detection in urban areas using exclusively SAR data and aims at overcoming the lack of data or information for the scene of interest whilst permitting a global applicability of the developed approach and a fine scale analysis of the buildings.

1.1 Problem Statement

One crucial element when speaking about rapid mapping is the time component. After a natural disaster (later referred to as *event*) occurred, emergency response is required. Information about the extent and the kind of the damage, as well as logistic information, should be available rapidly. Current spaceborne SAR missions allow the fast observation and scene acquisition of every point on Earth in a short time interval, independently from daytime and weather condition. Yet, depending on the applied technique and used sensors, the response time and accuracy can vary significantly. It is therefore important to choose the most suitable system.
Particularly sensitive and challenging are urban areas, which present high people density and should provide rapid logistic solutions for finding and offering shelter to potential victims. An important migration of population from rural to urban areas is observable since over half of a century, principally for work, but also for health, cultural, social or economical reasons. This leads to a significant urbanization, especially in the city outskirts, where buildings capable of lodging hundreds of families grow very rapidly since the 1970s. Such areas are particularly sensitive in case of a natural disaster. Namely, the buildings are often old and not built for countering such catastrophes. Furthermore, many families are living in such buildings, inducing high fatalities in case of collapse. Finally, they are usually located near main access roads to the city centers, and their collapse could induce a total obstruction of the city entrances for rescue teams. Hence, changes should preferably be modeled at building level, in order to facilitate the logistic decision making. In this work, a methodology is developed that allows to distinguish individual buildings and determine their state of damage.

To this end, three main questions arise:

- Which SAR system(s) and technique(s) enable(s) a rapid and global application?
  The main limitation of change detection analysis is the poor availability of pre-event data, due to the unpredictable occurrence of an event. This work addresses accurate pre-event techniques that can be applied globally and post-event techniques that permit a fast response.

- How do buildings look like, using the chosen technique(s)?
  Considering the chosen techniques, it is necessary to analyze and understand the appearance of buildings in the resulting SAR images. This enables to estimate the accuracy to which the building analysis is possible, and to extract and compare individual buildings in order to perform change detection.

- Is it possible to detect changes at building level with respect to the applied technique(s)?
  Depending on the accuracy of the chosen techniques and the precision of the subsequent building recognition and extraction, change detection analysis can be performed to a specific level of detail. This work addresses the capacity of the chosen techniques and applied

methodology to detect and categorize changes at building level, e.g. if it is still standing or if part of it is destructed and, for the later case, to which approximate amount.

This work addresses and answers those questions, discussing qualitatively and quantitatively the advantages of the chosen techniques and methodology, and their limitations.

1.2 State-of-the-Art

SAR image interpretation is highly related to the desired application. Depending on the chosen mode of acquisition, the appearance of SAR images varies a lot and multiple complementary information can be extracted. Figure 1.1 shows a sample of SAR images acquired under different illumination conditions and acquisition techniques. Amplitude-based image interpretation relies on the statistical analysis of the backscattering intensity. This allows to define specific texture characteristics for different objects in the scene, leading to subsequent classification or feature extraction. The result highly depends on the sensor resolution. A similar interpretation can be performed using polarimetric data, where object characteristics are defined with respect to their interaction with the emitted signal (Figure 1.1a). In order to fulfill the information and enhance feature extraction, images taken from different directions of acquisition (multi-aspect data) can be used (Figure 1.1b). The directions of acquisition are marked with the yellow arrows. In other cases, additional information coming from the combination of two images taken under the same viewing geometry can be used to enhance or replace the amplitude-based interpretation (Figure 1.1c). Using the coherence between two images, similar analysis of the coherence values can be performed as with the intensity, leading to the classification of temporally stable or unstable objects. Moreover, the analysis of the phase difference between two images taken under the same illumination conditions give additional information about the topography of the scene, permitting its 3D interpretation.

Different approaches have been developed in the last decade that permit automatic change detection in urban areas using the SAR technology. One of the main problem is the sufficiency of pre-event data, as most of the natural disasters are not predictable. Therefore, several approaches deal with the fusion of pre-event optical or GIS data and post-event SAR data. Due to the permanent development of the SAR systems, low resolution SAR images are available for the pre-event analysis of most areas since a few years. The post-event analysis is mostly performed using a single SAR image, in order to furnish a rapid response.

The existing approaches not only differ in the kind of data used for the change analysis, but also in the level of detail of the damage assessment. Whereas many methods produce maps of entire areas or cities with an indication of the districts that underwent the most damages, only few methods perform analysis at building level.

An exhaustive review of existing approaches for damage detection in urban areas using different kinds of data is presented in (Dong & Shan 2013). In the following, the existing approaches using SAR data are categorized, considering if they rely only on SAR data or use additional information.

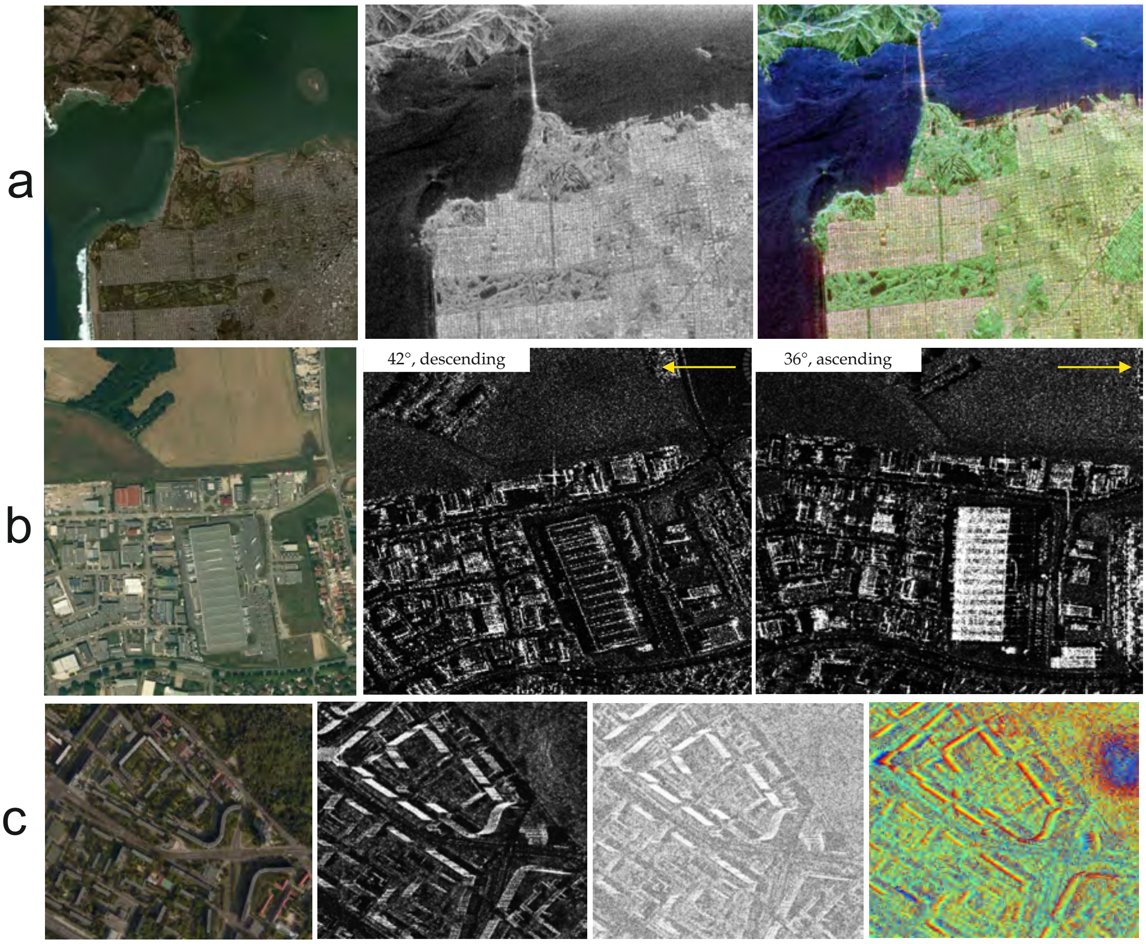


Figure 1.1: Aspect of SAR data under different acquisition modes; (a) optical (source: Bing Maps), amplitude and color-coded polarimetric data (source: JPL-NASA, AIRSAR); (b) optical image (source: Bing Maps) and amplitude images of multi-aspect data; (c) optical image (source: Bing Maps), interferometric (InSAR) intensity, coherence and phase

Damage assessment using external ancillary data

In (Mansouri et al. 2007), a GIS layer containing building footprints and heights is used for analyzing four multi-aspect post-event Envisat ASAR images of Tehran, Iran. After georeferencing, different statistical features of the SAR images, such as coherence and backscattering intensity, are analyzed at the building edges and corner locations, permitting to detect changes at parcel level due to the low resolution of the data. In (Liu et al. 2012), building footprints are manually extracted from GIS data and changes of backscattering intensity are detected within the extracted footprints. In (Trianni & Gamba 2008), pre- and post-event ALOS/PALSAR data are acquired and change detection performed by classification of extracted features such as coherence loss and intensity changes. The pixel-based damage assessment is merged in order to provide a parcel level damage information, using parcel border information from GIS data. Such approaches provide satisfying results. However, due to the resolution of the data, they often do not allow to detect changes at building level. Additionally, they rely on the availability of GIS information for the area of interest, which is not always ensured, particularly in remote areas where natural disasters occur.

To overcome this lack of data, some authors propose the combined use of different sensors for assessing changes. In (Butenuth et al. 2011), the authors propose a combined use of optical, SAR and DEM information into a general damage assessment system, detecting road infrastructures after flooding. (Chini et al. 2013) use optical, thermal and SAR data for analyzing the flooding after the 2011 Japan Tsunami. The adaptability of those approaches for buildings damage detection after destruction still has to be proved, as the occlusion effects are very different. In (Chini et al. 2009), the authors use multiple sensors in order to give damage information at several levels. Change detection at building level is performed using pre- and post-event very high resolution optical imagery whereas change detection at district level is carried out with lower resolution SAR imagery from Envisat ASAR. Furthermore, the authors analyze the behavior of a change criterion defined with the pre- and post-event SAR images using a building mask created with optical data. High correlation between the different building classes and the variation of the criterion is observed, proving the complementarity of optical and SAR data. In (Dell'Acqua et al. 2011), the authors also use optical data for building damage assessment at pixel level and SAR data for damage assessment at district level. For the latter, the area is divided into several districts, and a ‚damaged area ratio' is created and thresholded by analyzing several texture parameters derivable from amplitude SAR data, such as variance, mean and entropy. Similarly, (Upret & Yamazak 2012) use two TerraSAR-X images of the pre- and post-event, taken under same viewing geometry, and analyze the difference of backscattering coefficient and the loss of correlation between both images. Based on the different values, thresholds corresponding to different building states are defined. Those are assessed using building footprints digitalized from a post-event optical image. In (Dong et al. 2011), the authors draw building footprints on pre-event Quickbird images, and analyze the behavior of the RADAR backscattering coefficient of a post-event TerraSAR-X image within the defined footprints. Three classes are defined a posteriori: undamaged, semi collapsed and totally collapsed. Thresholding the backscattering coefficient on other areas or using further images enables the determination of potentially damaged buildings. In (Stramondo et al. 2006), the authors propose a complementary use of optical and SAR data for district level damage assessment. After a coregistration and resampling of the optical images to the resolution of the SAR images, different feature vectors are defined, using SAR complex coherence and intensity as well as optical indexes as NDVI and pixel-to-pixel difference. Several combinations of the determined indexes are analyzed, showing that the combination of optical and SAR data provide better change detection accuracy than their separate use. In (Brunner et al. 2010), a method is proposed that fuses pre-event very high resolution optical imagery with post-event high resolution SAR imagery in order to detect changes at building level. First, relevant building parameters are extracted manually from very high resolution optical images. Those parameters are then used for simulating the building appearance in SAR images. Finally, the simulated building is compared to real post-event SAR imagery and a similarity criterion is calculated using mutual information. The change decision is performed using Bayes' rule for distinguishing two classes: damaged and undamaged. This approach shows good results, yet the framework could be improved using automatic building extraction methods. A similar approach has been developed in (Wang & Jin 2012), defining three building classes: collapsed, subsided and deformed.

It is obvious that the use of complementary information in the form of optical and SAR imagery is feasible for obtaining detailed damage information whilst permitting a rapid application. Until less than a decade ago, the resolution of spaceborne sensors did not permit a detailed analysis at building level using only SAR data. Optical images allowed to understand the backscattering behavior of SAR images at building location, and derive some thresholds for damage classification. The approaches presented previously reveal yet two main drawbacks. First, if pre- and post-event SAR data are available, both must be acquired under the same incidence angle in order to analyze the correlation and difference of intensities. Second, only two or three damage classes are derivable, all giving only a rough, scattered-based, information about the building state. However, the use of new spaceborne sensors like TerraSAR-X or COSMO-Skymed, which allow very high resolution, permits a more detailed analysis of the changes. It should be possible to define the amount of the change, and give some information about the planimetric and altimetric extent of the change. For this, a simple analysis of the backscattering coefficient is not enough, and feature based approaches should be envisaged.

The presented approaches using simulated data go in this direction. However, more advanced simulation approaches can be applied that take into account speckle noise and surroundings, as presented in (Auer et al. 2011b). Such a method is used in (Tao et al. 2011), whereby both optical and SAR simulations are performed based on a LIDAR derived DEM in order to facilitate the visual interpretation of SAR data. A further development leads to the simulation of whole city areas under different incidence angles (Tao et al. 2013). This approach relies on the principle that the available SAR images before and directly after an event may not have been acquired under the same incidence angle. In order to detect changes, the simulation has to represent buildings as they appear in the post-event SAR image. Based on the LIDAR DEM, the scene is simulated and different masks can be created. This approach focuses only on the analysis of layover areas, created by building walls. A pixel-based analysis between simulated and real post-event layover areas enables the detection of changes at building level. The originality of this approach is that it is not necessary to have SAR images acquired under the same incidence angle for pre- and post-event analysis, allowing a faster response. Yet, it requires the availability of an accurate DEM of the scene of interest, which is, in most of the cases, not available.

Damage assessment using only SAR data

Approaches relying only on SAR data can be applied in remote areas, where no other kind of information, or bad optical information is available.

In (Balz & Liao 2010), the authors analyze the appearance of damaged buildings in multiple post-event high resolution SAR amplitude images. Different levels of damage are analyzed theoretically, simulated and detected manually on diverse real amplitude images. The relevance of using multiple viewing geometries to recognize the damages is shown. This approach provides a good description of the appearance of damaged buildings in amplitude SAR data, but still has to be automatized. Similar works are presented in (Shinozuka et al. 2000, Wang et al. 2009). In (Kuny & Schulz 2014), the authors present an extension of the previous approach by simulating heaps of debris. Simulated SAR data are trained considering several texture features, and heaps of debris are subsequently detected automatically in real SAR data. This method shows good results in terms of completeness of the extraction, however, many other objects like trees are also classified as heaps of debris due to their similar signature. Further texture features should be defined in order to diminish the false alarm rate. The drawback of methods based on simulations is the quantity of possible appearances a building and a damaged building can take, which should all be modeled in order to detect every possible change.

The methods using only SAR data and no simulation can be grouped into amplitude driven methods and phase driven methods, whereby some approaches use the full information content of the SAR or InSAR images.

The amplitude driven methods typically use the backscattering coefficent and intensity correlation between one pre-event and one post-event SAR image, both taken under the same viewing geometry. In damaged areas, the intensity correlation becomes lower and the difference of backscattering coefficient between pre- and post-event increases (Matsuoka & Yamazaki 2004). At building level, additional features can be considered. In (Guida et al. 2010) and (Brett & Guida 2013), the authors propose to extract double-bounces corresponding to building footprints in one pre-event SAR amplitude image. A double-bounce mask is created. Additionally, the ratio of the pre-event with one post-event amplitude image taken under the same viewing conditions is built. The ratio image is multiplied with the double-bounce mask. Double-bounce lines showing a high ratio are classified as damaged buildings. In (Polli et al. 2010), the authors attempt to classify damaged urban areas using only post-event amplitude imagery of COSMO-Skymed. There, different textures are defined and analyzed in groups of pixels, trying to differentiate several damage levels. Better performance of this method is shown using block analysis at lower resolution SAR images (Cossu et al. 2012), as it suffers less of influence of small objects.

The phase driven methods exploit the loss of phase coherence between interferometric image pairs of pre- and post-event. In (Ito et al. 2000), multi-temporal interferograms are used, using SAR images acquired before and after the event. Several combinations of the images are used, providing different coherence information. Image pairs acquired before the event provide a pre-event coherence, and co-seismic coherence is estimated using one pre-event and one post-event image. Damaged areas corresponding to a coherence loss between the pre-event and co-seismic interferograms are classified using neural networks. A similar damage assessment is performed in (Hoffmann 2007). In (Bignami et al. 2004), a coherence based damage classification of urban areas is performed, showing high correlation with damage assessment derived from optical data.

A few methods use both amplitude and phase information. In (Matsuoka & Yamazaki 2000), the intensity difference, the intensity correlation and the coherence difference between pre- and post-event interferograms are analyzed, showing that the coherence is suitable for categorizing different degrees of damages. A similar approach is presented in (Yonezawa & Takeuchi 2001), showing that damages at building level influence the backscattering characteristics of the signal and induce a decorrelation between pre- and post-event data. In general, the amplitude information is used for determining areas where potential changes occurred, and phase coherence information is used to assess and specify the changes (Arciniegas et al. 2007, Yonezawa et al. 2002).

Another approach relying on the analysis of polarisation changes between pre- and post-event is presented in (Singh et al. 2013). It shows good results using low resolution SAR images for detecting damaged areas, and should be tested on higher resolution images for detecting single damaged buildings. However, due to the higher level of detail at higher resolution, the robustness of the classifier towards other objects and infrastructures should be tested.

Limitations of current approaches

In general, approaches relying only on SAR data for change detection assess the damages at a block or district level, classifying the backscattering characteristics and features calculated from the image statistics. Partly due to the lower resolution of the imagery provided by the chosen sensors, only few approaches perform damage assessment at building level. All presented approaches rely on pixel-based methods, or use external information in order to simulate and quantify the change in terms of building extent and dimension. However, no approach specifies the amount of the change in terms of building height and planimetric dimension.

In this work, it is attempted to use a feature based method, permitting to quantify planimetric and height changes at building level. Besides, no external information is used or simulation performed, but the building parameters are estimated using pre- and post-event SAR data only. For this, not only a rapid technique for post-even analysis is looked for, but also a method permitting a global and accurate estimation of the pre-event state is addressed.

1.3 Contribution

The existing approaches use either external (e.g. optical, LIDAR) or only locally available (e.g. single SAR image, GIS data) pre-event data. Such data are often not up-to-date and sometimes show lower resolutions or bad quality. Therefore, they are difficult to interpret for building level analysis. The approach developed in this work attends to overcome this lack by using global SAR data provided by the current tandem mission of the German Aerospace Center (DLR). This mission aims at creating a global DEM of the Earth using the principle of SAR interferometry with the two satellites TanDEM-X and TerraSAR-X flying on parallel orbits. To this goal, interferometric data are acquired worldwide, providing a global coverage.

For post-event analysis with SAR data, most approaches use either a single SAR amplitude image or an interferometric phase image. The use of a single SAR image presents the advantage of a very short time of acquisition after an event. However, it suffers from the poor information content due to the single aspect. Occlusions occur, which hinder a good building recognition and reconstruction. Furthermore, the height information can only be derived based on the 2D measurement of characteristic image shapes, which in turn can suffer from occlusion effects. The approaches using interferometry seem to be a good compromise, as the elevation information can be deduced directly using the phase values of the interferogram. Such approaches rely mainly on the change of coherence in order to detect areas where changes occurred. However, due to the loss of coherence after an event, the phase information cannot be evaluated anymore, and elevation retrieval is usually impossible. The tandem mission may allow a more coherent interferometric acquisition due to a single-pass of the satellites above the scene of interest, but when this mission will finish, the only possibility for spaceborne interferometry will be repeat-pass acquisitions of the same sensor, inducing additional noise due to time decorrelation. In this work, a new solution for post-event analysis is investigated, allowing to retrieve height information under non coherent conditions, whilst using acquisitions taken within a short time span. Radargrammetric data, i.e. at least two images taken from different incidence angles, are used, and their amplitude information is treated similar to the optical stereoscopy in order to retrieve height information. With such data, incoherent 3D change detection can be performed.

In this work, pre-event interferometric and post-event radargrammetric SAR data are used in order to detect and extract buildings. The determined building parameters from both methods are then compared in order to detect and quantify the changes.
Considering the pre-event data, an automatic processing chain is developed in order to detect and extract buildings from interferometric SAR data. Based on existing building detectors using the phase image, a new detector is created. A new algorithm for building extraction is subsequently introduced, combining information coming from the building detection step and geometrical shape considerations. The building detection rate and correctness of the extracted building parameter are analyzed thoroughly, and the robustness of the extraction is shown.
For the post-event analysis, a new fully automatic processing chain for radargrammetric data is developed. Contrary to most existing feature-based approaches for building extraction by radargrammetry, this work relies on a pixel-based approach, combining recent knowledge on SAR image matching and methods coming from the optical stereoscopy. A recent approach for SAR image matching is used for automatic image coregistration, and methods for improving the image matching such as a pyramidal approach and backmatching are adapted to SAR image processing. Based on the developed algorithm, a thorough analysis of the building appearance in radargrammetric SAR data is conducted, permitting to derive a methodology for extracting building and change parameters. Furthermore, a new relative height calculation is introduced, taking into account the angle difference on the ground between two SAR acquisitions. A detailed analysis of the performance and limitations of the developed methodology considering different acquisition configurations is performed, defining requirements for future SAR missions.

Finally, a new fusion strategy involving building parameters deduced from interferometric and radargrammetric SAR data is developed in this work. The data fusion occurs at different levels of the methodology. First, information deduced from the interferometric processing is transmitted directly to the radargrammetric processing. It introduces some constraints that allow to improve the radargrammetric processing in terms of accuracy and time. Second, the deduced building parameters from both processing chains are brought together in a rule system permitting to evaluate the probability and the amount of a change.

In conclusion, the main contributions of this work can be resumed in the following points:

- an investigation of the combined use of pre-event interferometric and post-event radargrammetric SAR data,
- a new algorithm for building detection and extraction using interferometric SAR data,
- a new methodology for building detection and extraction using radargrammetric SAR data,
- a thorough analysis of the different acquisition parameters and configurations, allowing to define requirements for future SAR missions, and
- an original change detection strategy, allowing fast response under non coherent conditions whilst preserving height information.

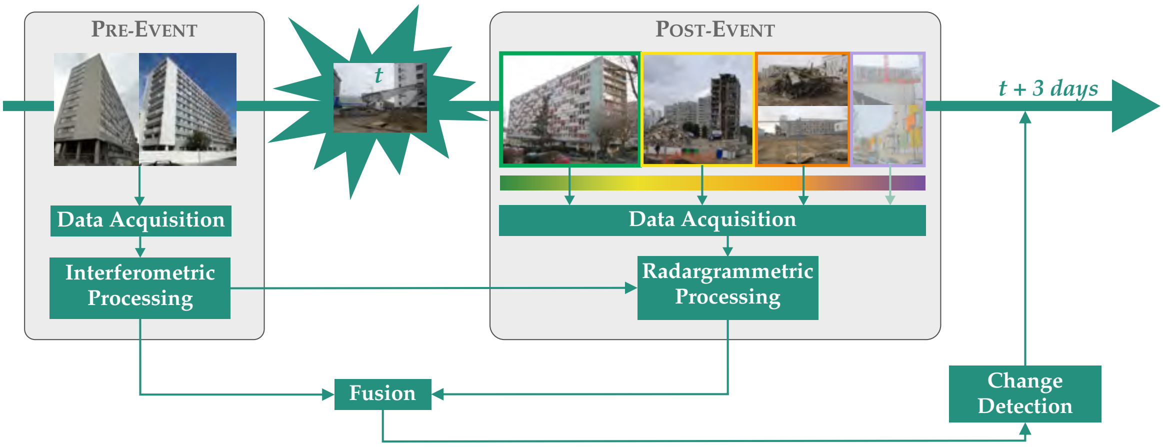


Figure 1.2: Concept of this work: pre-event interferometric data and post-event radargrammetric data are fused in order to detect changes in urban areas within a short timespan (maximum 3 days) after an event

In this work, the focus is put on middle-rise rectangular shaped buildings with flat roofs, mostly isolated from the surroundings, as can be found in many city outskirts. A few examples of such buildings are shown Figure 1.2.

1.4 Structure of the Thesis

The concept of this work is depicted in Figure 1.2, showing the overall structure of the thesis. Building parameters are extracted for the pre-event analysis from interferometric data. These are used for the radargrammetric processing in order to improve the robustness and limit processing time for the post-event analysis. New building parameters are deduced from the radargrammetric processing. Both sets of parameters are then fused in a rule system in order to detect changes.
Chapter 2 recalls the principles of the synthetic aperture RADAR (SAR) and highlights in more details the two SAR techniques employed in this work: SAR interferometry and SAR radargrammetry.
Chapter 3 describes the pre-event interferometric analysis. The building appearance in interferometric SAR data is reviewed, leading to the description of the overall interferometric workflow for building detection and extraction developed in this work. The new building detector is introduced, and the different processing steps are specified. The determination of the pre-event building parameters follows subsequently.
Chapter 4 is the equivalent of Chapter 3 for post-event radargrammetric data. First, the different building appearances in the amplitude images resulting from the different incidence angles are analyzed. Then, the new methodology for performing stereoscopy on SAR images is described step-by-step. It is explained at which step the pre-event building parameters can be considered in order to restrain the processing time and improve the robustness of the algorithm. Finally, new building parameters are extracted, corresponding to the post-event state.
Chapter 5 comprises the fusion of interferometric and radargrammetric data. First, the geometrical aspect is explained, describing how the pre-event building parameters are transformed for being used as constraint during radargrammetric processing. Second, the change detection

approach, using extracted pre- and post-event building parameters, is described, finally leading to the different considered changes.

Chapter 6 shows the suitability of the developed methodology by means of an application example. The test site and acquired data are described, and results of the different processing steps are shown. A thorough evaluation using different acquisition parameters is performed, and the obtained results are discussed and compared to reference data.

A final discussion and outlook of this work are given in Chapter 7.

2

SAR Principles

The wide field of remote sensing can be divided into two main sensing strategies: passive and active. Passive systems record the radiation reflected or emitted by an object. In optical imagery, the object is illuminated by the sun or other illumination sources. The optical sensor records the reflected light on the object surface. In thermal imagery, the object specific thermal emission is recorded. Active systems, on the contrary, have their own illumination source. They emit radiations that are reflected by an object. The system records the part of the reflection that is returned towards the sensor. For example, active sonar systems emit a sound whose return time indicates the distance of a detected object. In airborne and spaceborne remote sensing, most sensors use electromagnetic radiation as emission source. For example, LIDAR emits laser pulses that are reflected by the object. The resulting product represents the distance between sensor and object, determined by the speed of light and the time needed by the emitted pulse to return to the sensor. Whereas LIDAR systems use the optical domain of the electromagnetic spectrum, RADAR systems work in the microwave domain, with wavelengths varying from a few millimeters to about one meter.

Due to their active mode of measurement, RADAR systems are very attractive for night measurements. Indeed, they need no sun or other illumination source. Moreover, they can acquire data even by dense cloud coverage, as their long wavelength is affected only to a limited extent by the atmospheric conditions.

In this chapter, basics on RADAR are first recalled (Section 2.1). Then, the SAR principle is introduced and the different acquisition modes are presented (Section 2.2). Afterwards, the INSAR principle is explained and the current spaceborne constellation TanDEM-X is introduced (Section 2.3). The last section gives an exhaustive presentation of the radargrammetric principle (Section 2.4).

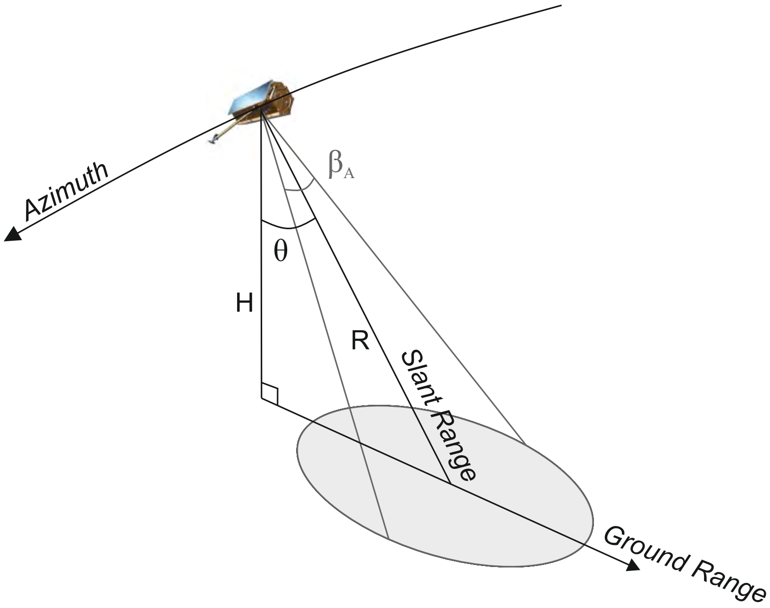


Figure 2.1: Geometry of a Real Aperture RADAR - RAR

2.1 Real Aperture RADAR (RAR)

RADAR is an acronym for *Radio Detection And Ranging* and is a range measuring method based on the emission and the time-delayed reception of an electromagnetic signal in the microwave domain. In the following, real aperture RADAR (RAR) is explained, before presenting the advantages of a synthetic aperture RADAR (SAR) in Section 2.2.

2.1.1 RAR Principles

Figure 2.1 shows the principle of acquisition of a real aperture RADAR. The ‚azimuth direction‘ corresponds to the flight path of the sensor. The perpendicular direction corresponds to the beam direction of the sensor. It is called ‚range direction‘ in the following, as it is the direction of measurement of the sensor-object distance. It has to be distinguished between slant range direction, which corresponds to the line-of-sight of the sensor, and ground range direction, which is its projection on a reference surface. First, the sensor emits a signal towards the scene. Second, the signal is reflected by the illuminated objects. Depending on the reflection characteristics of the material and the orientation of the considered object, the signal is either partially reflected towards the sensor (diffuse reflection) or totally reflected (specular reflection). The latter induces either that no signal comes back to the sensor (e.g., on very smooth surfaces, like calm water or oil) or that all the signal comes back to the sensor (e.g., specific material or succession of specular reflections bringing the signal back to its origin). The received signal is recorded, processed, and transformed into an image. The runtime of the signal between emission and reception is analyzed, leading to the distance between the illuminated object and the RAR system. An exhaustive explanation of the signal processing is given in Section 2.2.2, for SAR images. The resulting 2D image is complex and each pixel has two values, representing the real (in-phase) and imaginary (quadrature) part of the signal. Using angular functions, they can be related to the amplitude of the reflection and to the phase of the received radiation. The latter contains information about the sensor-object distance with an ambiguity corresponding to the number of full wavelengths between sensor and object. The run-time of the signal for two objects situated at the same distance of the sensor is the same. Thus, both objects are mapped at the same position in the resulting 2D image and it is impossible to discern them.

With perspective nadir geometry, as for optical sensors, it would be impossible to represent separately objects situated at the same distance but on different sides of the sensor. The side-looking geometry of RAR systems avoids this problem, as the beam is oriented in slant direction. In the final 2D images, objects are represented from near range to far range, in slant geometry.

The geometrical resolution corresponds to the minimal distance between two objects, such that they can still be distinguished from each other. For RAR systems, the geometrical resolution is usually different in range and in azimuth direction. The range resolution depends on the duration τ of the transmitted pulse, as well as on its speed, i.e. the speed of light c (Klausing & Holpp 2000):

$$\rho_{SR} = \frac{c\tau}{2}, \text{ whereby } \tau \approx \frac{1}{B}. \tag{2.1}$$

The pulse duration τ can be expressed in function of the bandwidth B of the emitted frequency modulated signal, referred in Section 2.2.2 as *chirp*. In ground range direction, it additionally depends on the incidence angle θ:

$$\rho_{GR} = \frac{c\tau}{2\sin\theta}. \tag{2.2}$$

Using shorter τ improves the range resolution. However, using too short pulses would yield a too weak signal power and deteriorate the signal-to-noise ratio. Yet, using pulse compression and matched filtering (Section 2.2.2) allows to obtain high resolution by maintaining low noise level. By reducing the incidence angle θ, the ground range resolution becomes large. The critical point is achieved when $\theta = 0$, i.e. by nadir illumination. The ground range resolution goes infinite, and all illuminated objects within the beam would be mapped within the same range cell. This is why the main characteristic of a RAR system is its side-looking geometry.

The azimuth resolution $\rho_{A_{RAR}}$ depends on the angular spread β_A of the signal - $\beta_A = \frac{\lambda}{L_A}$-, whereby λ is the signal wavelength and L_A is the length of the antenna in azimuth direction. It is expressed as:

$$\rho_{A_{RAR}} = R\beta_A = \frac{R\lambda}{L_A}, \tag{2.3}$$

whereby R is the slant distance between sensor and object. Two ground objects situated at the same distance R from the sensor can only be distinguished if they are not within the same beam. Considering Equation (2.3), it is obvious that the azimuth resolution ρ_A depends highly on the range R between sensor and object. Increased ranges lead to poorer resolution. In order to enhance the resolution, shorter wavelengths λ could be used. However, they are more sensitive to atmospheric attenuation. The use of a longer antennas (i.e. increased L_A) would also improve $\rho_{A_{RAR}}$, but it is subject to technical and material limitations. Longer antennas need very high transmission power, and are very heavy to deploy on a satellite platform. Instead of dealing with a real, long antenna, the creation of a synthetic longer antenna is possible, and led to the development of SAR sensors, which permit the achievement of a better azimuth resolution.

Table 2.1: RAR Wavelengths and Applications

Frequency Band	Frequency Range	Wavelength	Application Example
VHF	300KHz-300MHz	10m-1m	Biomass analysis, Foliage and ground penetration
P-Band	300MHz-1GHz	1m-30cm	Soil moisture, Ground penetration (Archaeology, Geology)
L-Band	1GHz-2GHz	30cm-15cm	Land-use analysis (Forestry), Soil moisture
S-Band	2GHz-4GHz	15cm-7.5cm	Land-use analysis
C-Band	4GHz-8GHz	7.5cm-3.75cm	Land-use analysis (Agriculture), Oceanography
X-Band	8GHz-12GHz	3.75cm-2.50cm	Object detection (Roads, buildings)
Ku-Band	14GHz-18GHz	2.1cm-1.7cm	Mapping of snow cover
Ka-Band	27GHz-47GHz	1.1cm-6.4mm	Target Tracking

2.1.2 RAR Wavelength

Commonly used frequency bands are presented in Table 2.1. In spaceborne applications, mostly X-, C- and L-Band, and sometimes P-Band are used, as they are a good compromise between low generation power and low atmospheric extinction. Besides, using these four bands, a wide range of applications can be covered. Due to their long wavelengths, P-Band and L-Band can penetrate the ground, allowing the analysis of ground structure, which is of high interest for archaeological and geological applications. Furthermore, they can penetrate through dense foliage, allowing ground mapping instead of forest cover. On the contrary, C-Band RADAR penetrates to a lower extent, permitting mapping of forest structure and biomass evaluation. Finally, X-Band wavelengths are too short to penetrate dense foliage. Reflections happen at canopy height, allowing classification of forest types by considering intensity of the reflection and volume decorrelation.

2.2 Synthetic Aperture RADAR (SAR)

The principle of Synthetic Aperture RADAR is to create an artificial very long antenna. Practically, instead of one single sensor position, the whole scene is illuminated at several consecutive sensor locations, creating a very long antenna. In the following, a straight flight path is assumed. The derived formulas can be found in more details in (Klausing & Holpp 2000).

2.2.1 SAR Azimuth Resolution

The angular spread β_A of the antenna at each sensor position is the same. An object in the scene is illuminated for each beam at a different position. This corresponds to different ranges to the sensor, as shown in Figure 2.2. Therefore, a Doppler effect occurs, with increasing signal frequency when the sensor comes closer to the object and decreasing signal frequency when the sensor moves away. The signal frequency f_0 at the object is expressed as:

$$f_0 = \left(1 + \frac{v}{c}\right) f_s, \tag{2.4}$$

whereby v is the component of the sensor velocity in range direction and f_s the frequency emitted by the sensor. The frequency shift between sensor and object is $f_0 - f_s$. Furthermore, the returned echo is shifted by the same amount, leading to a total Doppler frequency shift f_d at the object of:

$$f_d = 2(f_0 - f_s) = \frac{2v}{c} f_s. \tag{2.5}$$

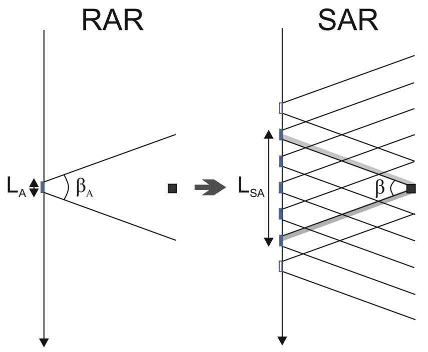


Figure 2.2: SAR Principle

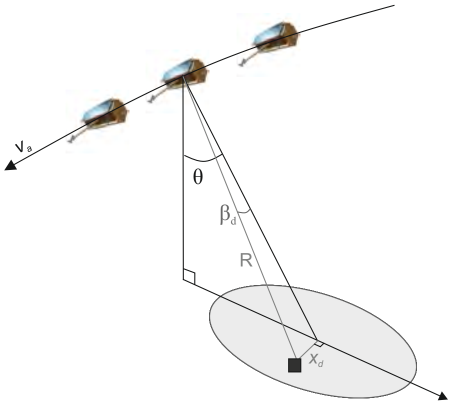


Figure 2.3: Geometry of a Synthetic Aperture RADAR - SAR

This expression can be rewritten as function of the velocity v_a of the sensor along its flight path:

$$f_d = \frac{2v_a \sin\beta_d}{c} \cdot \frac{c}{\lambda} = \frac{2v_a x_d}{\lambda R}. \tag{2.6}$$

whereby β_d is the angular difference between the object and the range direction and x_d is the azimuth coordinate of the object in a coordinate system with the sensor as origin, as represented in Figure 2.3.

The shift between emitted and returned frequency yields the correct object location, by extracting x_d.
Considering these observations, two objects can be distinguished if they have different Doppler frequency shifts. Thus, the azimuth resolution depends on the resolution ρ_{f_d} of the Doppler frequency shift:

$$\rho_{A_{SAR}} = \left(\frac{\lambda R}{2v_a}\right)\rho_{f_d}. \tag{2.7}$$

Assuming that the Doppler frequency shift is constant during the illumination time Δt of the object, ρ_{f_d} can be expressed as:

$$\rho_{f_d} = \left(\frac{1}{\Delta t}\right), \text{ whereby } \Delta t = \frac{R\beta}{v_a} = \frac{R\lambda}{L_A v_a}. \tag{2.8}$$

$R\beta$ is the length between the start and the last position of the sensor from which the object is observed (cf. Figure 2.2). It corresponds to the length of the synthetic aperture $L_{SA} = R\beta = \frac{R\lambda}{L_A}$. From Figure 2.2, it becomes clear that the angular aperture β at the object is equal to the angular spread of the signal β_A of a single antenna position ($\beta = \beta_A$). Combining Equations (2.7) and (2.8) leads to the azimuth resolution of SAR systems:

$$\rho_{A_{SAR}} = \left(\frac{L_A}{2}\right). \tag{2.9}$$

Contrary to RAR systems, the azimuth resolution of SAR systems is independent of the range of the sensor platform. This result should be interpreted cautiously, as the assumption that the Doppler frequency shift stays constant during the illumination of the object is only true if the antenna is steadily perpendicular to the flight path and has constant velocity during the acquisition (cf. StripMap mode, Section 2.2.5). In cases where the antenna direction does not stay perpendicular to the flight path during the acquisition, or rotate around the flight direction, this assumption is not validated (e.g., ScanSAR or Spotlight data, Section 2.2.5). In general, it can be retained that by increasing the range R between sensor and object, an increasing synthetic aperture L_{SA} has to be used in order to maintain the azimuth resolution.

2.2.2 Pulse compression

This previous formula corresponds to the azimuth resolution of compressed data, i.e. the multiple echos due to the multiple antenna positions are compressed in order to recover the original object position. This step is called ‚pulse compression‘ and is enlightened in the following. To this goal, it is necessary to recall that RADAR systems in general, and SAR systems in particular, are coherent. The phase of the transmitted signal is well defined and can be retrieved with high accuracy. Only under this condition, it is possible to analyze the phase of the returned echo and deduce the path of the signal.

As for RAR systems, yet at each sensor position, the antenna emits a pulsed signal. Mostly, transmitted pulses are modulated, so that the signal can be transmitted and returned back to the sensor without loss of the phase information. The frequency modulated emitted signal, or *chirp*, was already mentioned in Section 2.1.1. For every sensor position, the signal received by the sensor is first demodulated, converted into a digital signal, and mapped in a 2D matrix. At this point, it is necessary to notice the difference between resolution and pixel sampling (Raney & Wessels 1988). In order to preserve all signal information during the analog to digital conversion, the Nyquist criterion has to be respected, i.e. the sampling frequency f_{ns} must be greater than twice the signal bandwidth B. Only under this condition, a perfect, unambiguous signal can be reconstructed, without aliasing. This oversampling of the signal leads consequently to a pixel spacing in the resulting 2D matrix smaller than the resolution.

Rows of the matrix correspond to the different sensor positions in azimuth direction, and columns to the range of the illuminated objects. The raw matrix has to be processed in order to acquire the final image data. Indeed, the representation of point targets is spread in both azimuth and range directions, respectively due to the Doppler effect and the pulse dispersion. As a scene contains more than one object of interest, spread signals of several objects mostly overlap, and the raw data is hardly directly interpretable (cf. Figure 2.4).

The processing happens in two main steps: range compression on each RADAR echo, followed by azimuth compression on each column of the range compressed data. For both of them, so-called 2D matched filters are employed (Chan & Koo 2008). Such filters consist of a reference function corresponding to the transmitted *chirp* for the range compression and to the Doppler effect for the azimuth compression. By convolving these reference functions with the 2D signal, peaks corresponding to the responses objects are extracted, and superimposed signals of different objects are therefore separated. Usually, those convolutions in the spatial domain are performed as correlation in the frequency domain, using the Fourier transform (Moreira et al. 2013). Near the main peak (i.e. main lobe) of the filter response, side-lobes are present that blur the response of the object. For an object presenting a strong reflection (i.e. a strong backscatterer), its side-lobes could interfere with main lobes of other objects, presenting lower backscattering properties. Specific weighting functions are applied in order to smooth those side-lobes (Harris 1978).

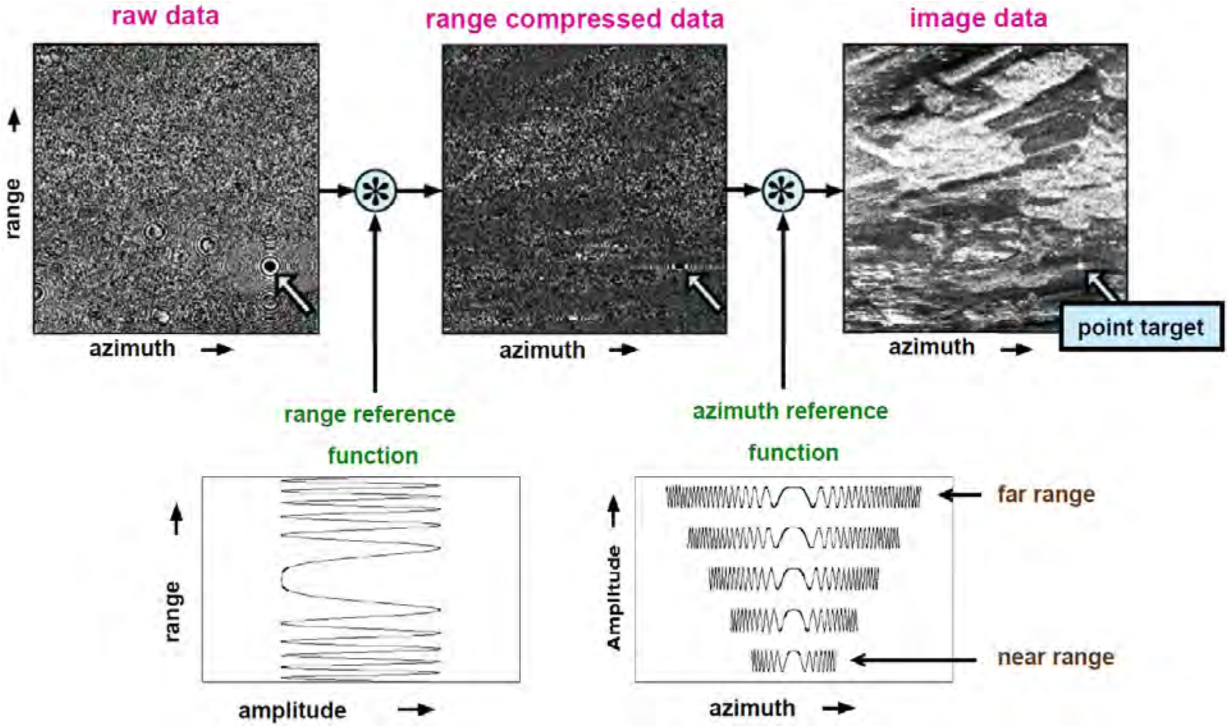


Figure 2.4: Pulse compression (source: (Moreira 2000))

2.2.3 SAR Image Data

The resulting image data is a complex matrix, containing in-phase I_p and quadrature Q_p components (i.e. respectively real and imaginary part). The complex data can also be expressed in amplitude and phase information, whereby the amplitude A is the backscattering coefficient of the signal, and the phase ϕ is the phase of the returned signal. The relations between complex (I_p, Q_p) and polar (A, ϕ) coordinates are recalled in the following equations:

$$I_p = A\cos\theta, \quad Q_p = A\sin\theta \tag{2.10}$$

$$A = \sqrt{I_p^2 + Q_p^2}, \quad I = A^2 \tag{2.11}$$

$$C = I_p + iQ_p = A\exp(i\phi) \tag{2.12}$$

I is the intensity of the signal. C is the complex image.

The amplitude image has a grainy appearance, even on homogeneous areas. This ‚salt-and-pepper‘ pattern is called *speckle* and is due to the superimposition of the backscattering of different targets within the same resolution cell (Goodman 1976). Due to the sensor resolution, single scatterers cannot be resolved individually. Within every resolution cell, echoes of several scatterers interfere. Their coherent summation is either constructive or destructive, depending on their relative phase. This results in brighter or darker pixels, producing the ‚salt-and-pepper‘ appearance (Moreira et al. 2013).

Especially for land-use classification or object detection, speckle makes SAR image interpretation more difficult and has to be reduced. Several methods exist for reducing, or filtering, the speckle. They mostly assume a multiplicative speckle model, i.e. the speckle is directly proportional to the intensity, or grey level, of the area. It implies that its variance increases with the

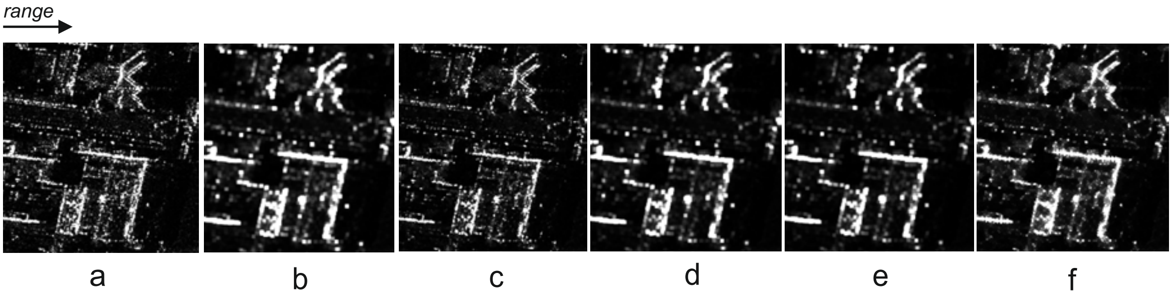


Figure 2.5: Example of standard Speckle Filtering: (a) original intensity image; (b) multilook 5x5; (c) Frost filter 5x5, damping=2; (d) Gamma MAP 5x5, ENL=1; (e) Lee 5x5; (f) Refined Lee 5x5, edge threshold=0.5

intensity. The probability density function (pdf) of the intensity of the received signal under fully developed speckle follows the Rayleigh Speckle model (Lee et al. 1994):

$$pdf(I) = \frac{1}{\sigma^2} exp^{\left(\frac{-I}{\sigma^2}\right)} \tag{2.13}$$

where σ is the mean RADAR cross section (RCS) of the considered area: $\mu(I) = \sigma^2$.
Speckle filters can be separated into two main categories: local and non-local filters. The most common local filter is a *multilook*, practically an averaging of the intensity image within a specified neighborhood. This spatial averaging induces a degradation of the image resolution. Other local filter methods permit to reduce the speckle by preserving the spatial resolution. Adaptive local filters, such as the Frost (Frost et al. 1982), Kuan (Kuan et al. 1985), Lee (Lee 1980) and Gamma MAP filter (Lopes et al. 1993), take the local image statistics into account for weighting the filter coefficients. Such filters are nowadays state-of-the-art and implemented in most SAR softwares. Particularly interesting are the refined Lee filter (Lee 1981) and the enhanced Lee filter (Lopes et al. 1990). The first uses the local gradient information, considering eight mask directions. Fluctuations along the edges are removed but edges are preserved. The second separates the image into three different regions: homogeneous, heterogeneous with structures, and point targets. The weighting coefficients are different for these three regions, allowing the preservation of structure information, without blurring it. Based on this principle, non-local filters have been developed lately, as in (Deledalle et al. 2010). They permit to complete the local information with information of similar areas, situated at other locations in the image. A similarity criterion has to be defined for finding analogous areas. Such filters are very efficient for preserving structural information, such as roads, in the image, by guarantying a very good smoothing of homogeneous areas. However, due to their non-local approach, the results of such filters are still noisy at the boundaries between two different areas, i.e. around edges. Figure 2.5 shows the results of some speckle filtering methods for the same window size.

2.2.4 Geometrical Image Distortion

Due to the side-looking configuration of RADAR acquisitions, not only simple reflections, but also typical geometric distortions appear in the SAR image, caused by the existing relief. These phenomena are imaged schematically in Figure 2.6 and recalled briefly here. The different grey levels represent the intensity of the corresponding effect in the SAR image, bright grey representing high and dark very low intensity.

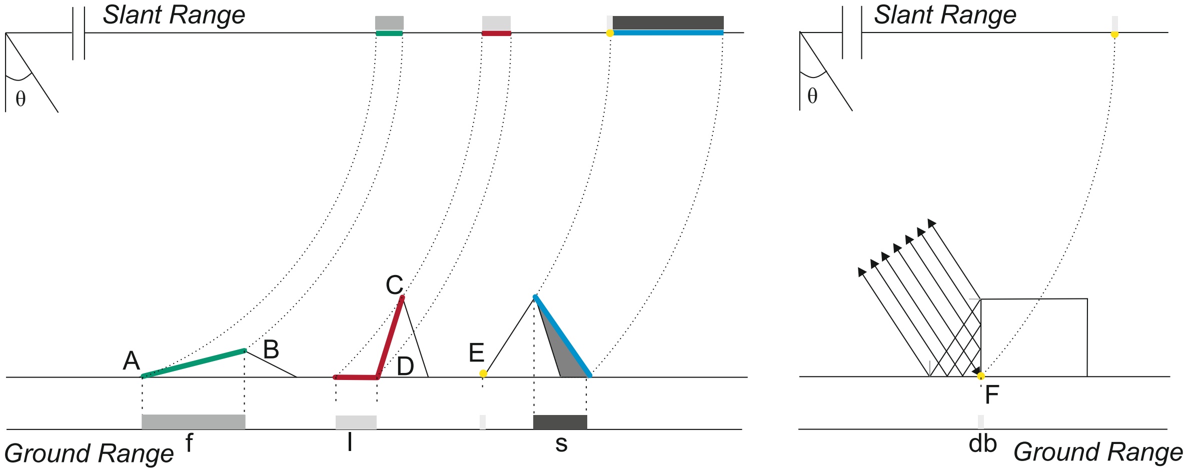


Figure 2.6: Schematic representation of side-look induced geometrical image distortions

Foreshortening (f)

This phenomenon appears principally in mountainous areas, when the illuminated terrain shows a slope smaller than the incidence angle θ. Considering the points A and B in Figure 2.6, situated at the foot and on top of the hill, respectively, their representation in slant range direction appears compressed (green lines). This terrain compression in the slant range representation yields brighter areas. Maximum foreshortening occurs when the slope increases until it becomes equal to the incidence angle. Indeed, it is then perpendicular to the RADAR beam, so that the whole slope has about the same range to the sensor. Its representation is then compressed into a few range pixels, leading to a very bright line (yellow point E).

Layover (l)

When the slope continues increasing and becomes greater than the incidence angle, layover occurs. There, the RADAR beam reaches the top of the feature before it reaches its foot. In the resulting image, the foot of the hill (D) is mapped at a farther range than the top (C). Layover induces therefore the summation of several contributions within one range cell. Slope points are situated at the same distance to the sensor as ground points (red lines) and therefore represented at the same position, leading to brighter areas.

Double-bounce (db)

This phenomenon appears principally on man-made objects, as it occurs mostly on vertical structures. There, part of the signal coming on the vertical wall undergoes a specular reflection that redirects the signal to the ground. Here again, the same phenomenon occurs, redirecting the signal towards the sensor (Franceschetti et al. 2002, Dong et al. 1997). This phenomenon occurs also in the opposite direction, i.e. from ground to wall. This is represented on the right of Figure 2.6, whereby the RADAR beam is reduced to a parallel illumination front. The path of the signal that undergoes a so-called double-bounce reflection has the same length for every point on the vertical structure. It corresponds to the range between sensor and foot of the vertical feature. Therefore, every scatterer of the ground and of the wall contributing to the double-bounce is

represented at the same range, corresponding to the distance to the foot of the structure (yellow point F in Figure 2.6). This leads to very bright lines in the resulting image.
Such phenomenon is not rare on man-made structures, and under the same principle, triple bounces or even higher reflections can be observed (Auer & Gernhardt 2014, Auer et al. 2011a).

Shadow (s)
The last phenomenon observable in SAR images are shadow areas. They result from foreshortening and layover. Hill sides looking back from the sensor are not illuminated (area marked grey under blue line in Figure 2.6). Therefore, no information returns to the sensor. Those areas are very dark in the SAR image.

2.2.5 SAR Acquisition Modes

Apart from the different utilizable wavelengths, SAR systems allow several imaging modes, which also influence the spatial resolution. The choice of the best suited acquisition mode depends on the application. SAR systems destined for land-use applications do not need the same spatial resolution as SAR systems designed for finer object analysis. In the following, focus is laid on modes enabled by the German satellite TerraSAR-X (Fritz et al. 2008), launched in June 2007, whose data are exploited in this work. Other SAR systems propose, at least, one of those imaging modes, with similar specifications.

StripMap Mode
The StripMap mode is the most basic imaging mode, following the SAR principle described previously. It consists in illuminating the scene with the same incidence angle and fixed range direction during an arbitrary time. This results in a single stripe of arbitrary azimuth length. The standard TerraSAR-X products have a coverage of 30 km in range direction and 50 km in azimuth direction. Depending on the incidence angle, the ground range resolution goes from 1.7 m up to 3.5 m. The obtained azimuth resolution comes to 3.3 m. StripMap data are mostly used for mapping and monitoring of land use, but can also be employed for object recognition such as ship detection.

ScanSAR Mode
The ScanSAR mode enables a very large range coverage, by switching slightly the incidence angle of the antenna. This results in multiple stripes along the range direction. During the change of incidence angle, transmitter and receiver are off. Therefore, each stripe is illuminated for a shorter time as for the StripMap mode, leading to a degradation of the azimuth resolution. For TerraSAR-X, the usual coverage with ScanSAR is 100 km in range direction and 150 km in azimuth direction. The ground range resolution varies between 1.7 m and 3.5 m, depending on the incidence angle. The azimuth resolution is about 17 m. A new wide ScanSAR mode is available since 2013 that enables a coverage of 270 km x 200 km with a resolution of 40 m. The ScanSAR mode finds application for large-area monitoring and mapping of wide disasters, as oil slick, sea ice or glacier monitoring, or mapping of devastated areas by a forest fire. Using interferometry (cf. Section 2.3), topography mapping and deformation monitoring is also possible.

Spotlight Mode
In the Spotlight mode, the incidence angle of the antenna stays fixed, but the antenna is steered in azimuth direction, permitting a longer illumination of the scene. The rotation center of the antenna steering is situated behind the scene at far range. The azimuth resolution is highly enhanced compared to the previous modes, at the detriment of the coverage. For high resolution Spotlight products, the scene coverage is 10 km in range direction and 5 km in azimuth direction. The ground range resolution is about the same as for the previous modes, varying from 1.5 m to 3.5 m depending on the looking angle. However, the achieved resolution in azimuth direction is 1.1 m. Due to their high resolution in both directions, it is possible to distinguish small objects. Therefore, Spotlight data are used for the detection and recognition of man-made structures and objects, such as roads, buildings and even vehicles.

Staring Spotlight Mode
The Staring Spotlight mode is available since 2013 on the TerraSAR-X platform (Mittermayer et al. 2014). It is similar to the Spotlight mode, with the exception that the rotation center for the antenna steering is situated at nearer range position, within the illuminated scene. For a comparable acquisition length, the normal Spotlight mode has a smaller antenna steering. The achieved coverage for the Staring Spotlight mode is about 5 km in range direction and 2.4 - 3.3 km in azimuth direction, for an azimuth resolution up to 0.25 m. With such a resolution, not only single objects can be detected, but also their structure and shape can be analyzed.
Further enhancements of the resolution are planned for the future, in particular within the TerraSAR-X Next Generation mission, that should continue to provide and enhance very high resolution modes up to 0.25 m beyond 2025 (Janoth et al. 2014).

Due to the available modes at the beginning of this work, but also due to the necessary and sufficient resolution for building analysis, the following work focuses on the processing and analysis of Very High Resolution Spotlight data.

2.2.6 Special SAR Techniques

Due to the coherent imaging capability of SAR sytems, and considering the polarization of the transmitted electromagnetic wave, several techniques have been developed in order to analyze terrain topography and enhance object detection. This subsection briefly presents the most common SAR techniques. A good overview is given in (Moreira et al. 2013) and (Toutin & Gray 2000). Sections 2.3 and 2.4 explain two of these techniques in more details, as they are used in this work.

Polarimetry
The technique of polarimetry relies on the analysis of the polarization of the signal. The emitted electromagnetic signal can be described by two components: one in horizontal and one in vertical direction. The targeted emission or reception of those wave planes allows the analysis of the polarimetric properties of specific targets. Depending on the roughness and backscattering properties of the surface of the illuminated object, they appear brighter or darker depending on

the analyzed polarization state. This allows their distinction and enables the creation of physical objects models. Polarimetry is therefore a very convenient method for classification. Current research on polarimetry includes land-use and soil moisture classification (Gosselin et al. 2014), as well as the detection of man-made objects (Bhattacharya & Touzi 2012).

Interferometry (InSAR)

The technique of SAR interferometry analyzes the phase difference between two SAR images taken from sensor positions characterized by slightly different incidence angles. The sensor-object range differs between both sensor positions, leading to different phases of the returned signal. The phase difference is directly related to the range difference. Thus, the surface topography can be deduced with high accuracy for the whole scene. The wide research field of SAR interferometry is yet not limited to the creation of Digital Elevation Models (DEM). Considering different baselines, i.e. different distances between both sensor positions, increasing the time between both acquisitions, or changing the configuration of the sensors with respect to the flight direction, provides a lot more fields of applications. Among them, surface displacement (Hu et al. 2012) and glacier flow analysis (Li et al. 2014), atmosphere estimation (Gong et al. 2011), as well as monitoring of land subsidence (Chaussard et al. 2014) or detection of moving targets (Budillon et al. 2013) are just a few examples. InSAR is explained in more details in Section 2.3.

Radargrammetry

The first definition of radargrammetry was given by (Leberl 1990): ‚Radargrammetry is the technology of extracting geometric information from RADAR images‘. It contains all techniques permitting to retrieve geometric information based on measurements on SAR images. Shape-from-shadow analysis from a single SAR image as in (Bolter & Leberl 2000*b*) is an example of radargrammetric analysis. Lately, the term ‚radargrammetry‘ has evolved to the description of the RADAR equivalence to the optical stereoscopy. It is under this definition that it is used in this work. It requires at least two images of the same scene. Compared to InSAR, two main differences exist. First, the incidence angles of both radargrammetric acquisitions are very different. Considering the provided incidence angles from TerraSAR-X, the difference of incidence angles reaches from about 10° up to more than 20°. Second, instead of phase differences, disparities between amplitude images are calculated. Indeed, such a large difference between both incidence angles yields important distortions between both represented scenes. Resolution cells of both images contain different contributing scatterers, making the phase difference analysis useless. Radargrammetry is therefore an incoherent method, whereas InSAR uses the coherence of the signal between both acquisitions. Radargrammetry is mostly used for the creation of DEM (Toutin et al. 2013), and is also employed in combination with InSAR for resolving problems in terms of phase ambiguities (Crosetto & Pérez Aragues 2000, Lachaise et al. 2014). This technique is presented in more details in Section 2.4.

Tomography

The technique of SAR tomography is a relatively new technique, whose applications in forested and urban areas are booming since about five years. Instead of using a one-dimensional synthetic aperture, tomography exploits a bi-dimensional aperture. Additionally to the synthetic aperture

Table 2.2: SAR Systems and their Characteristics

SAR System	Launch	Band	Resolution (Range/ Azimuth)	Repeat-Cycle
RadarSAT-II (Canada)	2007	C-Band	1.6m/0.8m	24 days
TerraSAR-X (Germany)	2007	X-Band	1m/0.24m	11 days
TanDEM-X (Germany)	2010	X-Band	1m/0.24m	11 days
COSMO-Skymed (Italy)	2007-2010	X-Band	1m/1m	1/16 days
Kompsat-5 (Korea)	2013	X-Band	1m/1m	28 days
HJ-1C-SAR (China)	2013	S-Band	5m	4 days
RISAT-1 (India)	2012	C-Band	0.67m/1m	25 days
SENTINEL-1a/b (Europe)	2014/2015	C-Band	5m/5m	12 days
PAZ (Spain)	2014	X-Band	1m/1m	4-7 days
ALOS-2 (Japan)	2014	L-band	3m/1m	14 days

in flight direction, another aperture is created, yet in elevation direction, by multiple passes of the sensor at slightly different heights. Such a configuration allows to bypass the main problem of the side-looking SAR, namely that targets situated at the same range from the sensor are represented in the same resolution cell and cannot be separated. By analyzing the scatterers' distribution at different heights, tomography allows to resolve this ambiguity and obtain accurate height profiles of vertical objects (Frey et al. 2014).

2.2.7 SAR Systems

In this short subsection, an overview of current civil spaceborne SAR missions is given, with respect to their frequency band, their best achievable resolution and their interferometric repeat-cycle (Table 2.2). The last one characterizes the time needed between two acquisitions under the same illumination conditions, which is important for repeat-pass interferometry (cf. Section 2.3.5).

Considering repeat-cycle and resolution, PAZ, Kompsat-5, COSMO-Skymed, TerraSAR-X and TanDEM-X meet the requirements for building analysis. They provide very high resolution, and interferometry is possible within a suitable time span, thus avoiding strong effects of temporal decorrelation. Moreover, the tandem configuration of TerraSAR-X and TanDEM-X satellites allows simultaneous interferometric acquisitions, removing all temporal effects (cf. Section 2.3.6). BIOMASS (Heliere et al. 2009), TanDEM-L (Moreira et al. 2009), and the future constellation of RadarSAT (Seguin et al. 2014) systems are planned in order to extend the global Earth coverage to several wavelengths, ensuring a wide spectrum of applications.

2.3 Interferometric SAR (InSAR)

As stated in the previous section, interferometry evaluates the phase difference between two acquisitions taken from slightly different incidence angles. In this section, the interferometric image formation process is explained and the importance of the acquisition configuration is motivated. Besides, the tandem satellite mission of the German satellites TerraSAR-X and TanDEM-X is reviewed, as it is used in the following of this work.

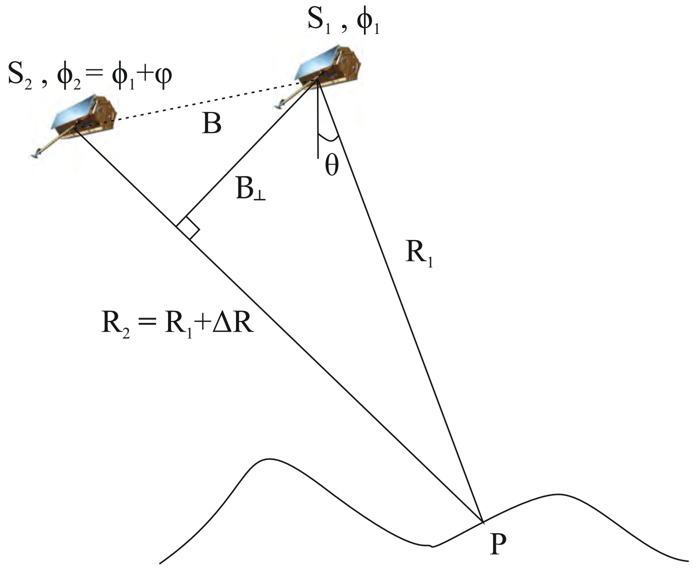


Figure 2.7: Interferometric acquisition geometry

2.3.1 InSAR Principle

A schematic representation of an InSAR configuration is represented Figure 2.7. S_1 and S_2 are the satellite positions of the two acquisitions. They are situated at distance B from another, called *baseline*, whose orthogonal component to the looking direction is the *effective baseline* $B_\perp$. Considering one single SAR image, the phase ϕ_s of the returned signal for a specific point target is proportional to its range R:

$$\phi_s = \frac{2\pi}{\lambda} R + \phi_{scatt}, \tag{2.14}$$

whereby the factor of 2 corresponds to the two ways from and back to the sensor, and ϕ_{scatt} corresponds to a possible phase shift due to the roughness of the scatterer itself. As mentioned previously, the interferogram builds the phase difference between received signals at both sensor positions. This phase difference is expressed as:

$$\varphi = \phi_{s1} - \phi_{s2} = p_f \frac{2\pi}{\lambda}(R_1 - R_2) = p_f \frac{2\pi}{\lambda} \Delta R. \tag{2.15}$$

Here, ϕ_{scatt} is considered to be the same for both images. The factor p_f is either 1 or 2, respectively if the signal is transmitted only at one position and received on both (standard mode) or transmitted at both sensor positions, where each position receives its own echo (‚ping-pong‘ mode) (Rosen et al. 2000). ΔR is the range difference between both sensor positions and the target.

In practice, the interferogram formation consists in performing a complex multiplication between both single SAR images as described in Equation (2.12):

$$C = C_1 C_2^* = A_1 \exp(i\phi_{s1}) A_2 \exp(-i\phi_{s2}) = |A_1||A_2| \exp(i\varphi) \tag{2.16}$$

As a result, two images are obtained: the intensity image and the phase image, corresponding to $|A_1||A_2|$ and φ, respectively.

2.3.2 InSAR Phase

Several components contribute to the phase φ (Ferretti et al. 2007):

$$\varphi = \varphi_{topo} + \varphi_{flat} + \varphi_{def} + \varphi_{atm} + \varphi_{noise} \tag{2.17}$$

φ_{topo} is the main contribution to the final phase φ and represents the terrain relief. φ_{flat} arises from the side-looking configuration. Indeed, a point target situated at far range shows a different phase contribution as a point target situated nearer to the sensor, even if both points have the same height above ground. φ_{def} and φ_{atm} are caused by temporal decorrelation, if both acquisitions are taken at different times: φ_{def} arises if the terrain suffers some deformation between both acquisitions, and φ_{atm} is induced by possible changes of the atmospheric conditions. Finally, several factors contribute to φ_{noise}: thermal noise inherent to the sensor, orbit errors and noise due to volume, temporal, or geometric decorrelation. The latter is caused by the baseline B. Namely, in order to have an interpretable phase φ, it is necessary that the same scatterers, or at least the same strong scatterers are present in the same resolution cells of both images. Increasing the baseline B induces a poorer concordance of the scatterers in both images. The critical baseline is reached when the object spectra do not overlap anymore, i.e. the distributions of the backscattered signal for each object within the resolution cell do not superimpose anymore between both acquisitions.

As for single SAR images, the phases of different scatterers sum up within the same resolution cell, and some scatterers can appear or disappear from one image to the other. As a consequence, interferometric SAR images are also affected by speckle. Several methods exist for filtering the interferometric phase. They can be separated into two main groups: boxcar and oriented filters. Among the first group, simple filters such as median and mean (multilooking) of phase values within a specific neighborhood can be found. Also, more sophisticated filters using the frequency domain can be applied (Goldstein & Werner 1998, Baran et al. 2003). These filters can separate high frequencies corresponding to noise from low frequencies, allowing total preservation of the phase information. Furthermore, region growing approaches exist, which create asymmetrical filter windows of similar image statistics. Such filters are mainly employed on the interferometric coherence image (Vasile et al. 2004, Vasile et al. 2006). Finally, as for single SAR amplitude images, non-local approaches can be deployed for finding regions of similar statistics within the image (Deledalle et al. 2014). The second group of filters consists of oriented filter masks. In (Rejichi et al. 2010), the local phase gradient is analyzed in order to define masks parallel to the local orientation of constant phase values by morphological operations. The most effective oriented filter is probably the adaptation to phase images of the enhanced Lee filter, presented in Section 2.2.3. It defines sixteen different oriented masks at each pixel position, with uniformly distributed orientations within $[0, \pi]$ (Lee et al. 1998). After analyzing the local statistics, i.e. the variance within each mask, the mask showing the lowest variance is considered for the filtering. The filtering is performed in the edge direction, allowing a good preservation of image structures. A slightly modified approach was presented in (Bo et al. 1999), considering twenty-eight masks instead of sixteen, and combining oriented masks with boxcar filtering on homogeneous areas. Recently, an approach relying on adaptive dynamic masks has been developed for smoothing fringe patterns along given feature directions (Dubois et al. 2012).

2.3.3 InSAR Coherence

The quality of the interferogram is usually evaluated by the coherence γ, which represents a measure of the correlation between two complex SAR images:

$$\gamma \approx \frac{|\sum_{W} C_1[i,k] C_2^*[i,k]|}{\sqrt{\sum_{W} |C_1[i,k]|^2 \sum_{W} |C_2[i,k]|^2}} \tag{2.18}$$

Usually, the coherence is estimated within a small neighborhood W for each pixel, $[i,k] \in W$. C_1 and C_2 represent two complex acquisitions. The coherence γ takes values between 0 and 1, 0 standing for no coherence and 1 for a perfect correlation. For example, in shadow areas, where no signal arrives, both acquisitions are completely decorrelated and show very low coherence. Depending on the backscattering properties of the illuminated surfaces, the coherence varies. Therefore, the coherence image is often used for land-use classification (Martinez et al. 1998, Strozzi et al. 2000). Furthermore, temporal decorrelation between both acquisitions may cause important changes in the backscattering properties of a surface, inducing a poor coherence. Coherence estimation can thus be used for change detection, for example over forested areas (Askne & Hagberg 1993).

2.3.4 InSAR Height

The interferometric phase φ is principally defined by its topographic phase contribution φ_{topo}. A small interferometric phase difference $\partial\varphi$ between two point targets on the ground corresponds to a small height difference ∂h of these two points (Bamler & Hartl 1998). The so-called phase-to-height sensibility is expressed as:

$$\frac{\partial\varphi}{\partial h} = p_f \frac{2\pi}{\lambda} \frac{B_\perp}{R \sin\theta} \ , \tag{2.19}$$

where θ represents the incidence angle of the master sensor position. The interferometric phase is defined with a 2π ambiguity, which corresponds to the range ambiguity of the transmitted signal. Indeed, the phase shift of the signal coming back to the sensor can be measured precisely. However, the total number of full waves between sensor and object is unknown. This ambiguity yields a fringe pattern within the interferometric phase image (Figure 2.8b). One fringe corresponds to one 2π cycle, given in the interval $]-\pi,\pi]$. A key parameter is hence *the height of ambiguity* h_a, which is the height corresponding to a phase change of 2π:

$$h_a = \frac{\lambda}{p_f} \frac{R \sin\theta}{B_\perp}. \tag{2.20}$$

The retrieval of the correct terrain topography, i.e. the creation of a phase field proportional to the local topography, involves phase unwrapping (Figure 2.8(d,e)). The main existing phase unwrapping algorithms are all based on the same principle: phase variations are assumed to be less than π in the direct pixel neighborhood, and the phase field should be continuous. This means, that the sum of the wrapped phase differences in a closed path should be equal to zero. This assumption is verified on most terrains (Ferretti et al. 2007). Therefore, if this

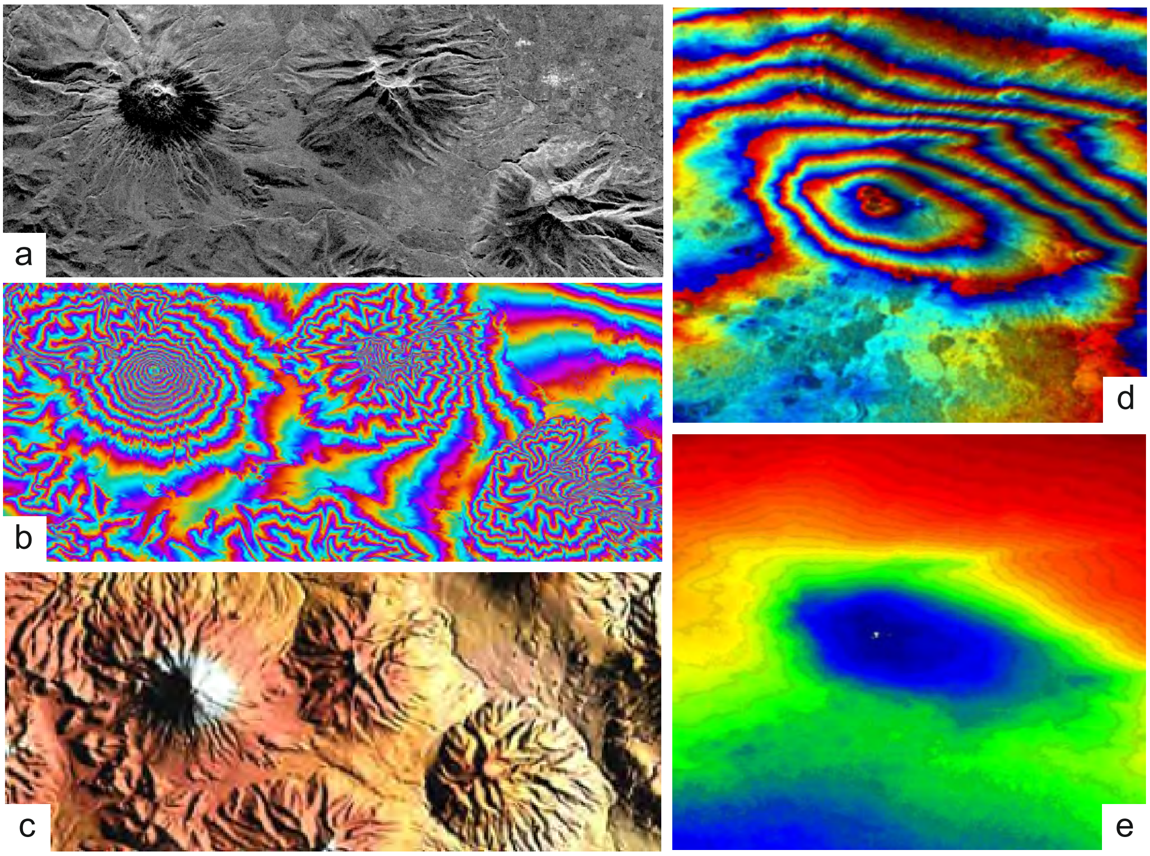


Figure 2.8: first column: Cotopaxi Volcano, Ecuador: (a) amplitude image; (b) phase image (fringes); (c) DEM; second column: Mount Etna, Italy: (d) wrapped phase image; (e) unwrapped phase image (source: JPL-NASA, SIR-C)

path integration is not verified, phase discontinuities, also called *phase residues*, occur, and phase unwrapping has to be performed. The most commonly used algorithm was developed by Goldstein (Goldstein et al. 1988). It is based on a branch-cut method, i.e. residues are selected and connected to each other in order to lead the phase unwrapping along consistent paths. These branch cuts must not be crossed by the unwrapping process. Other methods based on least square matching have been developed, whose goal is to minimize a specific cost function (Ghiglia & Pritt 1998). Further advances in phase unwrapping techniques, as considering multiple baselines, can be found in (Ferretti et al. 2007).

2.3.5 InSAR Configurations

Depending on the geometric and temporal configuration of the InSAR acquisitions, the interferometric phase at a certain location may vary a lot. In this subsection, a general overview of the possible acquisition configurations is presented, and their influence on the interferometric phase is emphasized.

Along-Track and Across-Track Configurations

Depending on the desired analysis, two different configurations of the sensors with respect to the flight direction are possible.

In along-track interferometry, both antennas are aligned in the flight direction. Both acquisitions are taken at the same range within a very short time lapse, so that the resulting interferometric phase is equal to zero except of the inherent sensor noise. However, if the illuminated target moves between both acquisitions, a relative phase shift between both acquisitions appears at this target. Its interpretation permits the estimation of its velocity. Along-track interferometry is therefore used for ocean surface movement estimation (Romeiser et al. 2010), and moving target detection (Hinz et al. 2008).

In across-track interferometry, the baseline between both antenna positions has a cross-track component. This is the basic configuration presented in Section 2.3.1. Due to the small range difference of both acquisitions towards the target, the phase difference contains information about the surface topography φ_{topo}.

In this work, static objects are considered, and their height is looked for. Thus, across-track configurations are considered.

Single-Pass and Repeat-Pass Data

Another parameter that has to be looked at when using interferometric data is the time between both acquisitions.

By single-pass interferometry, both images are acquired simultaneously. Therefore, considering Equation (2.17), the contributions φ_{def} and φ_{atm} can be neglected. Indeed, no surface deformation and no atmospheric change can be observed. Furthermore, the part of φ_{noise} corresponding to temporal decorrelation can be neglected as well, as the scene does not change between both acquisitions. The thermal noise subsists though. Consequently, single-pass interferometry is very useful for DEM generation, as the interferometric phase contains mostly the topographic contribution φ_{topo} and the remaining component of φ_{noise}. Single-pass interferometry can be achieved either by two antennas mounted on the same platform, or by two different platforms flying close to each other. Considering spaceborne InSAR, the first configuration was deployed during the SRTM mission (Shuttle RADAR Topography Mission), whereby the first antenna was situated on the space shuttle and the second mounted at the end of a deployable mast. This 60 m long mast allowed to obtain a sufficient interferometric baseline. With this mission, a first global DEM in a 90 m raster was created. It has an absolute height accuracy of 16 m and relative height accuracy of 6 m (Rabus et al. 2003). The second configuration is currently achieved by the tandem mission of TerraSAR-X and TanDEM-X, explained exhaustively in Section 2.3.6. With two satellites flying in a tandem configuration, the main goal of this mission is the production of a second global DEM, with a 12 m raster, an absolute height accuracy of 10 m and a relative height accuracy of 2 m (Moreira et al. 2004).

By repeat-pass interferometry, both acquisitions are taken at different times. All contributions of the phase presented in Equation (2.17) are present. Repeat-pass data is affected by temporal decorrelation, yielding more noisy data. However, repeat-pass interferometry can be used for detecting changes that happened between both acquisitions, as change in the forest cover (Askne et al. 1997), long-time terrain deformation (Bürgmann et al. 2000), or glacier displacement (Atwood et al. 2010). For those applications, the evaluation of the coherence is of high importance, since areas showing low coherence are more probable to have changed. Furthermore, analysis and mapping of the atmospheric effects is possible (Li et al. 2004). Repeat-pass interferometry is easier to deploy, as a single antenna is sufficient, which has to fly twice over the region of interest. Current spaceborne missions are also characterized by their repeat-cycle, i.e. the time needed from the sensor or sensor constellation to perform a second acquisition under the same geometry as the first. Currently, the Italian sensor constellation COSMO-Skymed allows the shortest repeat-cycle (one day), using four satellites that circle the Earth in the same orbit. The German satellite TerraSAR-X has a repeat-cycle of eleven days. In Table 2.2, an overview of spaceborne SAR and InSAR missions is given.

With repeat-pass interferometry, differential interferometry (DInSAR) (Ferretti et al. 2007) and persistent scatterer interferometry (PSI) are also possible. The first one consists in isolating the contribution φ_{def} of the interferometric phase by substracting a DEM, i.e. φ_{topo} from the interferogram. The DEM can come from an external source, or have been created by interferometric acquisitions before the deformation occurred. Differential interferometry is often used for monitoring seismic deformations and volcano activities (Massonnet et al. 1995). PSI is usually employed for land subsidence analysis (Ferretti et al. 2000), and single building deformation analysis in urban areas (Gernhardt & Bamler 2012). Based on the same principle as DInSAR, long time series of the same scene are acquired, whereby the topographic contribution of the phase can be removed. However, contrary to DInSAR, PSI processing focuses on very stable radiometric targets, called persistent scatterers, whose deformation is analysed.

In this work, both single-pass and repeat-pass data were used, depending on the available acquisitions, as later explained in Section 6.1.2.

Ascending and Descending Orbits

SAR satellites are placed on sun-synchronous polar orbits in order to provide global coverage. The orbit inclination is around 97°, depending on the considered spaceborne system. Each satellite has a specific looking direction, even if this can be changed for some satellites and for specific applications. Considering a nominally right looking antenna, a scene is illuminated from west to east when the sensor flies from south to north, i.e. on ascending orbit. On the contrary, the same scene is illuminated from east to west when the sensor travels the orbit from north to south, i.e. in descending orbit. Considering the side-looking geometry and the surface relief, the geometric distortions mentioned in Section 2.2.4 are present. For DEM creation, such distortions hinder a good retrieval of the topography, particularly in shadow areas, where the interferometric phase shows only noise. Shadow masks are often used to hide these areas during DEM creation. Combining ascending and descending orbits contributes to compensate the missing information and enhance the DEM, as phase information from both mountain sides is considered (Carrasco et al. 1997).

Effective Baseline

As already mentioned in this section, the baseline between both satellites in across-track configuration plays an important role for the retrieval of terrain topography. The interferometric - or *effective* baseline $B_\perp$ may differ from the spatial baseline B. $B_\perp$ is the perpendicular projection of satellite position S_1 into the range direction R_2, and may be shorter as the spatial distance B between both satellite positions. Considering Equation (2.20), the longer the effective baseline $B_\perp$, the smaller the height of ambiguity. Long baselines provide thus higher fringe frequency and a phase unwrapping is often necessary. On the contrary, small baselines induce longer heights of ambiguity, i.e. more height values are contained in the 2π phase interval. Therefore, height estimation is not as accurate as with long baselines. However, for flat areas, phase unwrapping may not be necessary. Consequently, the accuracy of the height estimation and the necessity of phase unwrapping depends on the effective baseline, and a compromise has to be found. In order to facilitate the phase unwrapping whilst providing an accurate height estimation, multi-baseline approaches exist (Ferretti et al. 1997). Such methods consider information coming from several interferograms built with different baselines. Usually, the interferogram created with the smallest baseline is first used to retrieve a coarse unwrapped height information, and a finer height estimate is then assessed using interferograms acquired with longer baselines, by performing phase unwrapping.

2.3.6 The TanDEM-X System

The TanDEM-X system is composed of two satellites flying in close helix formation: TerraSAR-X and its twin satellite TanDEM-X. Their formation allows single-pass interferometry, with baselines between 250 m and 500 m (Gonzalez et al. 2009). The principal goal of this mission is to provide a worldwide DEM, with the accuracy requirements as stated in Section 2.3.5.

The helix formation is achieved by horizontal and vertical orbital displacement of both satellites relative to an original orbit, resulting in a helix-like movement of the satellites relative to each other. Hence, both satellites have orbits that never cross. Small displacements of both satellites around their orbit are allowed. As they are twin satellites (Krieger et al. 2007), similar acquisition configurations can be obtained for the complete Earth surface, considering that ascending configurations in one hemisphere have the same geometry as descending configurations in the other hemisphere, and vice-versa. The helix formation enables very diverse acquisition configurations and baselines, permitting a wide field of applications in addition to its original purpose that is the global DEM formation.

Four operating modes are enabled: bistatic (standard), pursuit monostatic, alternating bistatic (also called ping-pong) and simultaneous transmit, all allowing Spotlight, StripMap and ScanSAR imaging modes (Krieger et al. 2007). For the global DEM formation, bistatic StripMap mode is used, i.e. one of the two satellites transmits and both receive the signal simultaneously. In pursuit monostatic mode, both satellites emit and receive their own signal. In this mode, both satellites do not need to be synchronized as for the bistatic mode. The alternating bistatic mode is a combination of both previous modes, as the transmitter is switched in a pulse-to-pulse basis, and both satellites receive. This provides for every pass two monostatic and two bistatic images. This difference of signal path is considered by the factor p_f of Equations (2.15) and (2.20). In bistatic mode, $p_f = 2$ and the height of ambiguity is halved. By combining the different modes, interferograms with different phase-to-height sensitivities can be built, which facilitates the phase unwrapping process. The fourth operation mode makes an improvement in terms of signal processing, allowing pulse transmission simultaneously on both satellites.

In this work, standard bistatic data are used, yet acquired in Spotlight mode, for allowing the analysis of single buildings. Several baselines are acquired, permitting the estimation of a preferable configuration for the analysis of urban areas.

2.4 Radargrammetry

In this section, the principle of radargrammetry is described, similarly to the previous section for interferometry. In particular, image matching and height retrieval are explained, and the influence of the acquisition configuration is emphasized.

2.4.1 SAR Stereo Principle

First spaceborne radargrammetric investigations started about forty years ago, with the 3D mapping of the moon surface by the Apollo Lunar Sounder Experiment (ALSE) (Leberl 1976). Twenty years later, radargrammetry was used anew for mapping the surface of planet Venus during the Magellan mission (NASA) (Leberl & Maurice 1993). Further developments of spaceborne sensors allow to broaden the field of applications, leading to the creation of DEM over mountainous areas (Fayard et al. 2007) or glacier regions (Toutin et al. 2013), or to canopy height determination (Perko et al. 2011). Since about ten years, interest has grown for 3D city models, leading to further developments. An overview of radargrammetric contributions for the mapping of urban areas is presented in Section 4.2.

As for interferometry, radargrammetry requires at least two SAR acquisitions. However, both images are acquired with different incidence angles, θ_m and θ_s. A schematic representation of the radargrammetric principle is given in Figure 2.9.

The baseline B between both acquisitions is longer than the critical interferometric baseline mentioned in Section 2.3.2. Consequently, radargrammetric processing does not consider phase differences between both images for retrieving the height information. Instead, the parallax p arising between the representations in both amplitude images of the same point target P is determined and evaluated (Figure 2.9). This parallax, or disparity, is induced by the different incidence angles, and its determination follows the same principle as optical stereoscopy, yet for RADAR geometry. Homologous points in both images are found during a matching process (cf. Section 2.4.2) that determines their disparity. Depending on the acquisition configurations, the higher the target, the larger the parallax (cf. Section 2.4.3). Using the RADAR stereo model or geometric considerations permits the retrieval of the absolute or relative height of different targets (cf. Section 2.4.4).

2.4.2 Matching

Image matching consists in finding homologous points in both acquired images. Due to the difference of incidence angles, both images present different geometry and radiometry, which complicates the matching process. Two main strategies exist for matching: grey-level area based, and feature based.

Area based matching consists in area correlation. For a specific area in the master image, the slave image is searched in order to find a corresponding area, showing the best similarity with the master area. An overview of existing similarity criteria is later presented in Section 4.5.1.

Feature based matching strategies rely on the principle that specific structures have similar appearance in both images. Principally, points and lines are extracted in both images and

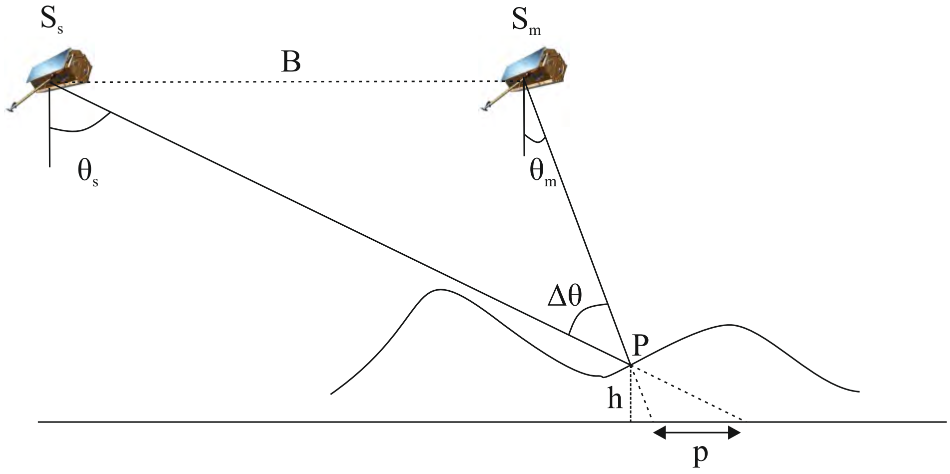


Figure 2.9: Radargrammetric acquisition geometry

matched using sophisticated schemes. Such methods are well developed for optical imagery, where standard feature extraction methods can be employed for extracting man-made structures (San A et al. 2007). However, the main drawback of feature matching is its sparseness. Only single features are matched instead of all image pixels, leading to a sparse terrain reconstruction. In optical stereoscopy, hybrid approaches have been developed, matching features in a first step and pixel points in a second step, as refinement and information enhancement (Baillard 1997). In (Paar & Pölzleitner 1992), an approach is presented that matches texture feature vectors. For each pixel in the template and search image, a set of texture features is defined, calculated using the statistics of its neighborhood. For example, mean, variance, but also gradient and edge properties are defined. Corresponding points in both images present the maximum similarity of their feature vectors.

An overview of existing feature based and area based approaches for applications in urban areas for SAR images is given in Section 4.2.

The result of the matching is stored in a so-called disparity map, where, for each pixel of the master image, the parallax to its homologous point in the slave image is recorded.

The different matching techniques principally consist in defining a similarity function for finding homologous points of both images. Yet, matching can be improved using further methods that focus on simple rules and image geometry. In (Fua 1991), master and slave are inverted after first matching and a second matching is performed. Only points showing the same (opposite) disparity value in both results are considered reliable. This forward-backward matching, or *backmatching*, was also presented in (Paar & Pölzleitner 1992) for optical imagery. Hierarchical approaches as in (Fayard et al. 2007) consist in building image pyramids by reducing image resolution. The similarity function is evaluated at each pyramid level, starting from the lowest resolution, and results of one level are used as input for the disparity calculation of the next higher level. Such a strategy allows to deal with large images but particularly to reduce wrong matches and noise by using reduced resolution. In (Schubert et al. 2002), a multiresolution wavelet matching is performed.

In order to restrain the search area for matching, i.e. not considering the whole slave image, several constraints can additionally be introduced. This ensures to spare computation time and reduce the risk of wrong matching. The constraints can be separated into local and global

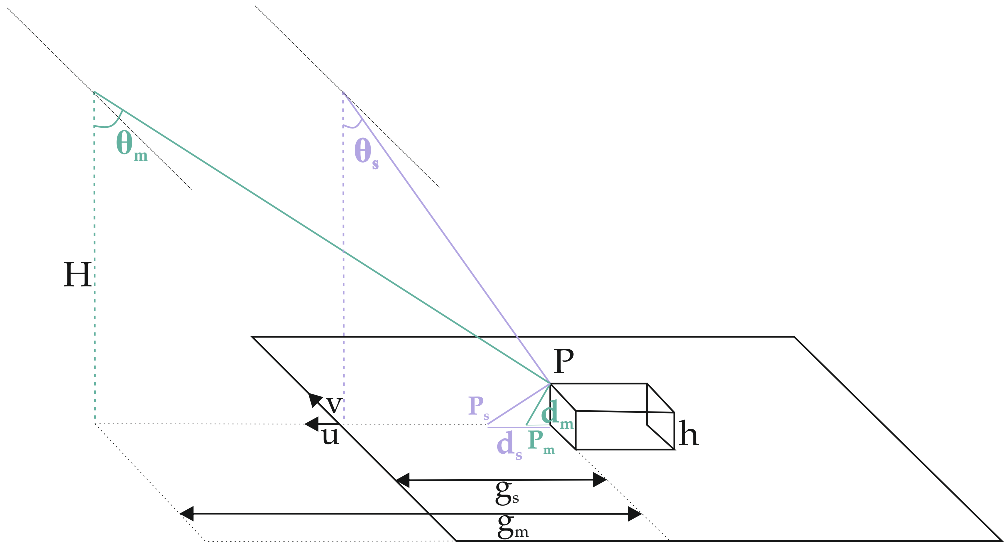


Figure 2.10: Principle of epipolar geometry adapted to radargrammetry along parallel flight path

constraints (Simonetto 2002). The main local constraint is the epipolar constraint. In (Sörgel et al. 2009), the general case of a flight path forming an angle β is considered, and the epipolar lines are represented for $\beta = [0, 90°]$. In Figure 2.10 the epipolar geometry is depicted for parallel flight paths of same height H. Due to layover effects, a point P situated at height h above the ground is represented in P_m and P_s in the ground geometry of the co-registered master and slave image, respectively. The distances d_m and d_s represent the respective displacement of P_m and P_s to the projected ground position of P. The local system (u, v) is defined by u, parallel to the range direction of master and slave, and v parallel to flight direction, u and v being orthogonal. The displacements d_m and d_s are expressed in this system as follows:

$$d_m = \begin{pmatrix} h/\tan\theta_m \\ 0 \end{pmatrix} \cdot \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} h \cdot H/g_m \\ 0 \end{pmatrix} \cdot \begin{pmatrix} u \\ v \end{pmatrix} \text{ and } d_s = \begin{pmatrix} h/\tan\theta_s \\ 0 \end{pmatrix} \cdot \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} h \cdot H/g_s \\ 0 \end{pmatrix} \cdot \begin{pmatrix} u \\ v \end{pmatrix} \tag{2.21}$$

The parameters g_m and g_s are the ground distances between the ground projections of P and the sensor positions. Considering Equation (2.21), for finding the match between both representations of point P in the co-registered ground images, a simple translation along the range direction u should be applied. The search area for a match thus corresponds to a single range line. It is equivalent to the epipolar lines defined in optical stereoscopy. The radargrammetric epipolar lines are therefore here all parallel, and correspond to the range lines. Consequently, the search area in azimuth direction can be reduced to a line, i.e. to a thin strip of one pixel width. In practice, both radargrammetric acquisitions are not exactly parallel, but differ from a small heading angle ζ. The search area is therefore restrained to a slightly larger strip (about three pixels wide) (Méric et al. 2009). A specific study of these effects in slant range geometry for layover areas is shown in Section 4.

Three global constraints are defined in (Simonetto 2002) for a terrain with low relief. First, it can be assumed that points are in the same order in both images. Second, each point has a single correspondence in the other image. Finally, a continuous disparity is assumed, i.e. the gradient of the disparity should not exceed a specific value. These assumptions are verified for almost flat terrains, with low relief changes. However, considering mountainous or urban areas, those assumptions are not verified. In optical imagery, a very efficient and robust method for getting rid of disparity gaps at object borders has been presented in (Hirschmüller 2005). It consists in using templates oriented in different directions, always looking for disparities along the epipolar line. The final disparity map is created by considering for each pixel the disparity for which the matching criterion was the highest along all considered directions. This approach provides very sharp object borders. An overview of the domain of validity of the constraints, depending on the terrain slope, is given in (Nocera 1996). On terrain affected by foreshortening, the uniqueness assumption is no more valid as the compression effect is not the same on both acquisitions. Also, if layover occurs only in one image, or if the acquisitions are taken from opposite directions, the points' order is not preserved (cf. Section 2.4.3). Finally, considering man-made structures, as given in urban areas, where abrupt elevation changes occur, the assumption of continuous disparity is no longer verified. Indeed, high disparity changes occur at the border between layovers and flat terrain. This aspect is discussed in Section 4.5.3.

An original approach has been proposed in (Nascetti 2013) for reducing the search area. It is based on the radargrammetric model presented in Section 2.4.4. Instead of limiting the search for the best match in a 2D space, a 3D space is used. Each acquisition is reprojected on ground for several specified terrain heights, leading to an image stack. Voxels corresponding to different height levels are defined for both images and the search for the best match is first carried out in the height direction, before performing fine search within the best height level in the 2D planar direction. Such an approach allows to perform the matching directly in the object space.

2.4.3 Acquisition Configurations

As in optical stereoscopy, the radargrammetric configuration of acquisition plays an important role for the correct retrieval of the 3D information. In this chapter, three main configuration parameters are depicted. As this works focuses on building analysis, the parallax at building location is considered in the explanation. The parallax p between both representations of a roof corner P is represented in ground geometry in Figure 2.11.

Acquisition Geometry

There are two acquisition geometries: same-side and opposite-side (Figure 2.11, I and II, respectively). By same-side radargrammetry, both acquisitions are taken from the same direction, i.e. the scene is illuminated from the same side twice. For spaceborne radargrammetry, possible same-side configurations are descending/descending and ascending/ascending, considering the flight direction of the satellite along its orbit. On the contrary, by opposite-side configuration, both acquisitions are taken from different sides of the illuminated scene. This corresponds to descending/ascending and ascending/descending configurations. With same-side configurations, the radiometry of both images is similar and the matching process is facilitated. However, only one side of the scene is illuminated, and 3D information may not be totally retrieved, particularly in case of high terrain relief inducing shadow areas. On the contrary, by opposite configurations, almost the whole scene is illuminated, preserving all the terrain information. The radiometry of both acquisitions differs significantly. Especially in case of high terrain relief and man-made structures, the typical SAR phenomena of foreshortening, layover, shadows and

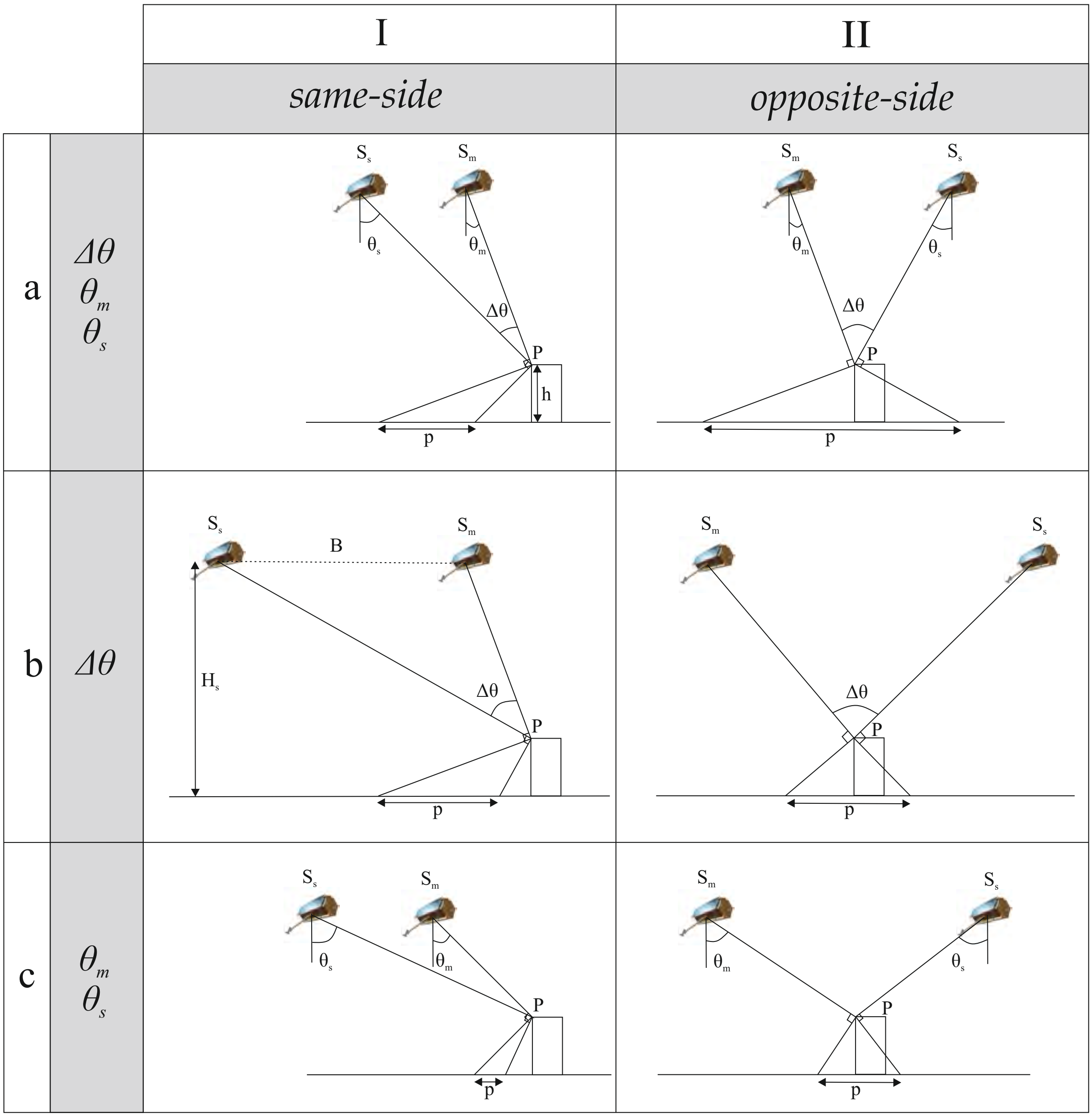


Figure 2.11: Radargrammetric properties in (I) same-side and (II) opposite-side acquisition; (a) acquisition geometry and parameters; (b) influence of intersection angle compared to (a); (c) influence of incidence angle compared to (a)

possible double-bounce occur. Those phenomena are inverted in both images, as the shadow of one side corresponds to the layover or foreshortening of the other side. The matching process is therefore more difficult for opposite-side configurations, in case of high relief disparities. Some approaches deal with this problem by inverting the radiometry of one image, to make it similar to the other. In (Fullerton et al. 1986), three zones are defined: bright, dark and intermediary. Bright and dark zones are inverted with respect to the intermediary zone. However, such approaches have to be considered carefully, as points with similar radiometry in opposite directions do not represent the same points in reality. This is explained in more details in Section 4.1.

Intersection Angle

Another important parameter is the intersection angle $\Delta\theta$, formed by both acquisitions. A schematic representation is given in Figure 2.11(a, b) for different cases. For same-side configurations, the larger the intersection angle, the longer the parallax between both images (Figure 2.11Ib). Making a parallel with the optical stereoscopy, the base-to-height ratio $\frac{B}{H_s}$ is used to quantify the quality of the intersection angle. The larger this factor, the larger the parallax, and the better the 3D intersection. Indeed, the intersection of both RADAR rays is better defined (Leberl 1990). However, due to the larger incidence angle difference, imaged objects appear differently distorted. Large intersection angles produce more radiometric disparities between both images. Particularly in dense urban areas, phenomena as layover and shadow vary significantly, and overlapping phenomena in one image may not be present in the other image, leading to a completely different radiometry. Finding homologous points is therefore more difficult. The relation of the parallax to the intersection angle is inverted considering opposite-side configurations. Indeed, for such configurations, a smaller intersection angle yields a larger parallax (Figure 2.11IIa). Yet, the intersection angle is highly related to the inclination of the incidence angles, discussed in the following paragraph.

Incidence Angles

It has to be distinguished between steep and shallow incidence angles. Considering Figure 2.11, it is obvious that steep incidence angles (a) provide larger parallaxes than shallow angles (c) (Toutin & Gray 2000), for both acquisition geometries. They are thus preferable for the 3D calculation. However, similar observations can be formulated as for the other criteria. Whereas shallow angles produce longer shadows, steep angles in turn produce larger layover areas. A compromise has to be made in order to have sufficient radiometric information for matching, without too much overlapping effects between objects, which could deteriorate the radiometry.

In conclusion, for a correct retrieval of the 3D information, large parallaxes p are preferred, as they offer a better geometry (Leberl 1990). However, they are often achieved at the expense of good radiometry for matching. It is obvious from the previous explanations that the acquisition configuration, the intersection and incidence angles are highly related to each other. A compromise has thus to be made when considering radargrammetric acquisitions, permitting a good geometry for 3D retrieval without hindering the matching process through a worsened radio-

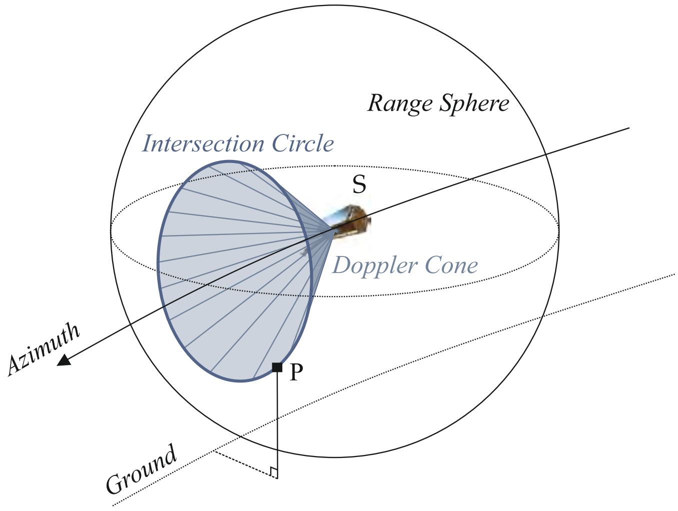


Figure 2.12: Radargrammetric Range-Doppler Model

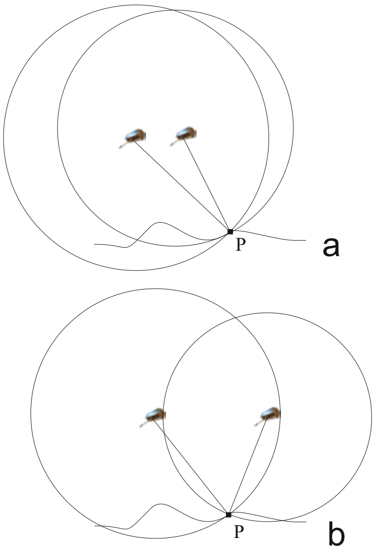


Figure 2.13: Radargrammetric intersections; (a) same-side; (b) opposite-side

metry. In Sections 6.1.3 and 6.3.3, the different acquisition configurations considered in this work are listed and discussed, with a focus on SAR typical geometric distortions at building location.

2.4.4 RADAR Stereo Model

In this section, the RADAR stereo model for extracting the 3D information is presented. First, the rigorous model used for retrieving absolute height and positions in a specific reference system is explained. Second, a more practicable but approximated model using simple geometric considerations is described. This model permits a relative height estimation based on the disparity map.

Absolute Height Determination

Figures 2.12 and 2.13 show the geometry of the RADAR stereo model. Lets first consider a single SAR acquisition, as depicted in Figure 2.12. The position of a point target P is defined by its range r_p and its Doppler frequency f_p. Considering the sensor position S as origin, the point P is situated on a sphere of radius r_p. Besides, the site of equal Doppler frequency can be represented as a cone with the sensor as origin and oriented towards the azimuth axis, defined by:

$$f_d = \frac{2}{\lambda} \cdot \frac{\langle \vec{\dot{S}}, \vec{SP} \rangle}{\vec{SP}} \tag{2.22}$$

whereby $\vec{\dot{S}}$ is the velocity of the RADAR at position S (Méric et al. 2009). This equation is similar to Equation (2.6), yet with new denominations. The squint angle ξ corresponds to the angular difference β_d (cf. Equation 2.6) between range direction and point P and can be expressed as:

$$\sin \xi = \frac{\lambda f_d}{2} \tag{2.23}$$

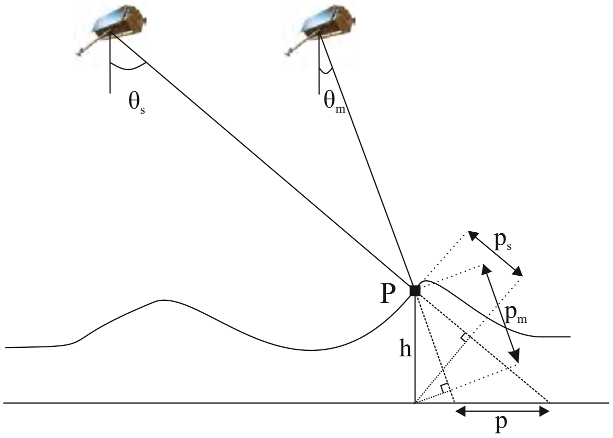


Figure 2.14: Geometric considerations for relative height determination by radargrammetry

The position of P can therefore be reduced to the intersection circle between the range sphere and the Doppler cone, leading after some simplifications to the equation system:

$$\begin{cases} |P-S| &= r_p \\ \dot{S}\cdot(P-S) &= \sin\xi\cdot|\dot{S}|\cdot r_p \end{cases} \tag{2.24}$$

However, for SAR sensors, a zero-Doppler frequency, i.e. a squint angle ξ equal to zero, is assumed after processing (cf. Section 2.2.2) (Leberl 1990). Therefore, the equation system (2.24) reduces to:

$$\begin{cases} |P-S| &= r_p \\ \dot{S}\cdot(P-S) &= 0 \end{cases} \tag{2.25}$$

Considering now radargrammetric acquisitions, the point P is at the intersection of the two circles defined by both acquisitions (cf. Figure 2.13). It is obvious that two solutions exist, whereby only one makes sense. The previous equation system (2.25) yields four equations for only three unknowns, which are the coordinates of the point $P(x_p, y_p, z_p)$. This non-linear system is overdeterminated and solved by a least-squares method (Leberl 1990).
Usually, in order to refine the parameters of the stereo model (sensor position S and velocity $\dot{S}$), ground control points (GCPs) are required (Toutin & Gray 2000). Recently, an algorithm allowing an accurate orientation of the radargrammetric model without GCP has been proposed (Capaldo et al. 2011).

Relative Height Determination

Instead of considering 3D intersection in space for absolute height determination, the disparity map provides relative terrain heights. Indeed, the parallax, or disparity, is directly related to the object height h by simple geometric considerations (cf. Figure 2.14). Making the assumption that both acquisitions are taken at the same sensor height H, on parallel flight tracks, with the baseline perpendicular to the azimuth direction, the following expression can be derived, in ground range geometry:

$$h = \frac{p}{\cot\theta_m \pm \cot\theta_s} \tag{2.26}$$

whereby ‚$-$' applies for same-side configurations and ‚$+$' for opposite-side configurations. Considering slant range geometry, this expression can be rewritten as follows (Fayard et al. 2007):

$$h = \frac{p_m - p_s}{\cos\theta_m \pm \cos\theta_s} \tag{2.27}$$

This last expression has to be handled carefully as p_m and p_s represent the distances between the top and the bottom of the object in master and slave image geometries, respectively. They do not correspond to the disparities between homologous points of master and slave image. The differentiation between ground and slant geometry is of high relevance. In ground range geometry, both images are first reprojected on the ground and the disparity is calculated then. In slant range geometry, the slave image is reprojected on the master in order to derive the disparities that are expressed in the slant range geometry of the master image. This aspect was first pointed out in (Goel & Adam 2012), and is explained in more details in Section 4.6.3. In this work, the slant range geometry of the master is used. Moreover, considering spaceborne TerraSAR-X data, the effect of the heading angle between both images is introduced. Namely, the absolute parallelism between both acquisitions is not validated. This is discussed in Section 4, where an exhaustive disparity analysis in layover areas is given, and a new height calculation is derived.

3

Extraction of Building Parameters from InSAR Data

In this chapter, the developed approach for the extraction of building parameters from InSAR data is presented. The overall appearance of buildings in InSAR data is shortly recalled first, focusing on rectangular shaped buildings with flat roof and predominant layover (Section 3.1). Considering this appearance, existing approaches for building detection in InSAR data are presented in a second part (Section 3.2), showing the pertinence of the method developed in this work, whose overall workflow is depicted in Section 3.3. Finally, Sections 3.4 to 3.6 present in more details the several steps of the developed approach.

3.1 Appearance of Buildings in InSAR Data

The general appearance of buildings in InSAR data is presented thoroughly in (Sörgel 2003), (Thiele et al. 2007a) and (Thiele et al. 2007b). A schematic representation is given in Figure 3.1 for amplitude A (a) and phase φ (b) behaviors. Two different building forms are represented (black and green). At the sensor looking building facade, layover occurs, as signals from ground, wall and roof overlap. As a result, the amplitude image shows high values at layover locations. They are characterized by bright areas, as represented schematically in Figure 3.1, from A to C (1). Considering the phase image, the layover is characterized by a descending phase ramp from near to far range. Equation (2.19) shows that the phase values are correlated to the building height. Alongside the signals coming from ground and roof, the begin of the layover in near range contains backscattered signals coming from the top of the facade, whereas the end of the layover at far range contains signals coming from the bottom of the building, therefore at lower height. The descending height values along the building facade produce the phase ramp. Depending on the building dimensions (height h and width w) and on the acquisition geometry (Sörgel 2003), the three signals of ground, wall and roof overlap, or only two signals from ground and wall. In the latter case, the layover in the amplitude image is slightly darker, as only two signals are superimposed (Figure 3.1, B to C (2)). Also the phase ramp is marked by a phase jump, as the contribution from the roof disappears. However, the slope stays the same, as a constant phase contribution is removed. Directly after the layover, at the foot of the building, the previously mentioned double-bounce effect occurs (C), leading to a very bright line in the amplitude image.

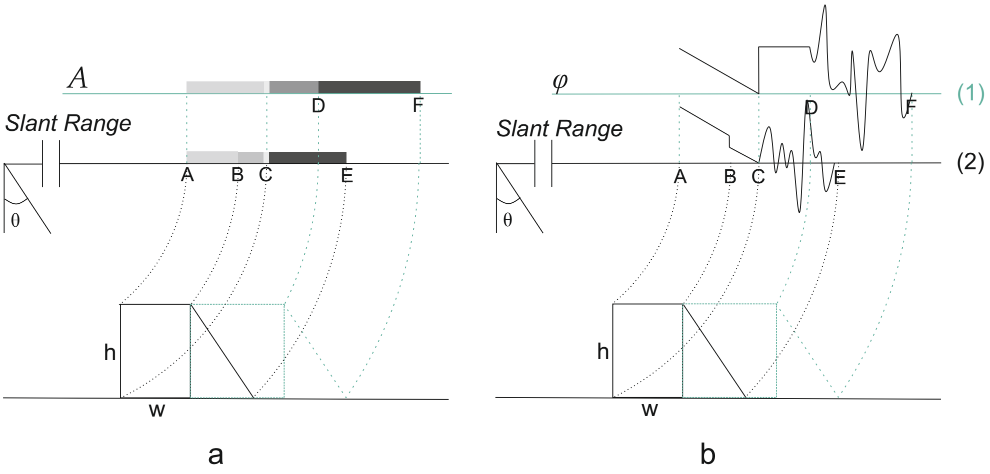


Figure 3.1: Schematic representation of building appearance in InSAR data for two different building widths: (1) visible roof part; (2) roof entirely contained in the layover; (a) amplitude image; (b) phase image

In the phase image, the phase value of the double-bounce line corresponds to the phase value of the ground level. After the double-bounce comes usually the roof area, where single backscattered signals coming on the roof are directly sent back to the sensor (Figure 3.1, C to D (1)). This is here only observable for the building delimited in green color (1), as for the black building (2) the entire roof is contained in the layover. Depending on the material properties of the roof, this area is more or less bright in the amplitude image. The roof is characterized in the phase image by constant phase values corresponding to the building height. Finally, the building side facing away from sensor, as well as the area situated in its ‚shadow', are not illuminated, leading to a dark area in the amplitude image, and noise in the phase image.

Figure 3.2 depicts three buildings in InSAR data, meeting the rectangular shape and flat roof description. The corresponding optical images show that the two first buildings have a regular structure, the first one consisting mainly in concrete material and the second one in glass and metal. The third building, for its part, shows high irregularities in the facade, as the concrete walls are interspersed by irregular window patterns, whereby shutters and french balconies alternate in an arbitrary way.

Considering the corresponding amplitude images, it is observable that the layover areas of all buildings differ a lot, even if they all show brighter intensities. The layover of the concrete building is a very homogeneous area (a), whereby the building consisting in glass and metal presents a very regular line pattern (b), where every line corresponds to a window floor. Changes in the facade orientation imply changes in the line direction. Depending of the considered incidence angle, the third building presents as well a linear pattern (d, e), or shows more heterogeneity (c, f). Layover areas in amplitude InSAR images appear to have very varying appearances, depending on material and acquisition configurations. Amplitude-based approaches for extracting buildings based on layover areas suffer from this heterogeneity. Likewise, the double-bounce lines (cf. Section 2.2.4) corresponding to the building footprint cannot always be distinguished in the two dimensional image, either because of trees situated in front of the building and changing the signal path, or because they do not show much higher intensities as the layover area, as for the first building (a). Furthermore, regarding dimensions and acquisition configuration, each building of Figure 3.2 represents the case (2) of Figure 3.1, whereby the entire roof area is con-

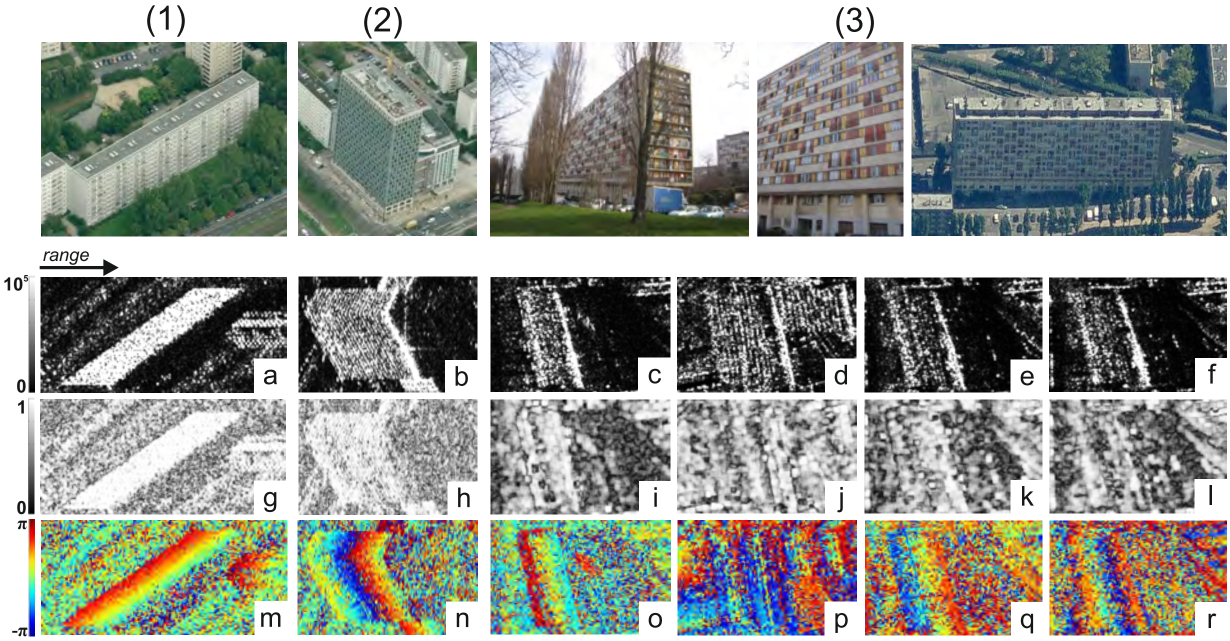


Figure 3.2: Appearance of three different buildings (1-3) in InSAR data (source for optical images (1-2): Bing Maps); (a-f) amplitude image; (g-l) coherence image; (m-r) phase image; Different acquisition configurations for building (3): (c, i, o) $\theta = 56°$, $B_\perp = 172$ m, $h_a = 65$ m; (d, j, p) $\theta = 21°$, $B_\perp = 186$ m, $h_a = 17$ m; (e, k, q) $\theta = 47°$, $B_\perp = 151$ m, $h_a = 56$ m; (f, l, r) $\theta = 47°$, $B_\perp = 234$ m, $h_a = 36$ m

tained in the layover. Its extraction would require first the layover detection and suffers from the same drawbacks. Besides, no significant variation of the layover brightness is observable for the buildings of Figure 3.2, making roof extraction nearly impossible. Finally, the shadow areas are not always well distinguishable from the surroundings. In dense urban areas, layover from other buildings occur that diminish the shadow (a, d), but also some processing driven artifacts or phenomena as urban canyons (Mokadem et al. 2013) may appear that cause the shadow areas not to be homogeneously dark (b, c, e, f).

Similar observations can be made for the coherence images as for the amplitude image (g to l), except that double-bounce lines are even more difficult to distinguish from the layover, as they show very similar coherence values.

The interferometric phase image contains other information as the amplitude image. For every building presented in Figure 3.2 (m to r), the characteristic fringe pattern, i.e. the descending phase ramp, is recognizable in the layover. Contrary to the amplitude image, the appearance of layover areas in phase image is consistent, independent of the building material. When two facades are visible, the orientation of the fringe pattern changes, the fringes being parallel to the orientation of each facade. Double-bounce lines have phase values corresponding to the phase values of the ground level. Their extraction presents two drawbacks: first, not only double-bounce lines, but whole ground areas would be extracted at the same time. Second, due to the 2π ambiguity of the phase, phase values corresponding of ground level may correspond in reality to wrapped phase values of features situated higher than the ground level. Therefore, extraction of double-bounce lines in phase images is not trivial and external information is necessary. As mentioned previously, for the buildings analyzed in this work, the roof areas are entirely contained in the layover. Due to the acquisitions' configurations, two different layover areas should be distinguished and a phase jump should appear. However, none of these observations can be made for the presented buildings, showing that the contribution of the roof is negligible. In fact, is has been proven that the facade contribution is the most important to the layover (Rossi et al. 2014a). Finally, shadow areas in the phase image show only noise, and cannot be distinguished easily from the surroundings.

As a conclusion, the most reliable features for the detection of middle to high-rise buildings, as those considered in this work, are layovers in the phase image, as those areas are very consistent, independent of building material. Due to the side-looking geometry, each layover area has the shape of a parallelogram, with two sides parallel to the range direction, and the two other sides parallel to the facade orientation, as is the fringe pattern.

Besides, their appearance using different acquisition geometries is predictable. Considering different configurations, important observations can be made: as stated from Equation (2.20), both incidence angle θ and effective baseline $B_{\perp}$ play an important role in the fringe appearance. This is shown in Figure 3.2 for building (3), which represents the same building taken with different baselines, and different incidence angles. The shorter the baseline, the longer the height of ambiguity, and vice versa. For very small baselines, the height of ambiguity is very large, leading to very long 2π ambiguities. Therefore, the fringe pattern is not so obvious as for longer baselines. Layover areas have a smaller ramp inclination, which complicates the recognition of layover areas from the surroundings. Similar conclusions can be drawn for the incidence angle, whereby a larger incidence angle induces a larger height of ambiguity. In Figure 3.2, almost two fringes, i.e. two 2π cycles, are observable for the considered building with an incidence angle of 21° (p) whereas only one fringe is noticeable for an incidence angle of 56° (o). Also, for the same incidence angle of 47°, the fringe repetition is observable within the layover by an effective baseline of 234 m (r), which makes phase unwrapping mandatory. On the contrary, less than one fringe is observed for an effective baseline of 150 m (q), where phase unwrapping would be superfluous. A compromise has to be found between unwrapping issues and good recognition of the typical layover phase ramp. These conclusions are discussed in more details in Section 6.2.1.

3.2 State-of-the-Art

The state-of-the-art in building reconstruction with spaceborne high resolution InSAR data can be separated into two main trends.

The first considers sparse reconstruction by analyzing a stack of interferograms. A very efficient and popular method therefore is Persistent Scatterer Interferometry (PSI), which consists in extracting stable points in a stack of several interferograms and retrieving their height based on their phase values. After unwrapping and georeferencing, 3D scatter points along building facades are retrieved (Adam et al. 2008, Gernhardt et al. 2010). Such approaches provide very precise height information which, combined with a time analysis, permit to evaluate deformations in millimeter accuracy (Gernhardt & Bamler 2012). Tomographic approaches, using multiple SAR images acquired from slightly different looking angles along a vertical baseline, permit to separate multiple scatterers situated in the same resolution cell (e.g., separation from ground and wall scatterers) (Zhu & Bamler 2010). The fusion of PSI and TomoSAR methods allows thus to overcome the superposition problem of layover in most urban scenarios and reconstruct separately ground, facade and roof (Frey et al. 2014). Despite of the high geometric accuracy, such methods have two main drawbacks. First, they allow only a sparse reconstruction of the building facade, so that the extracted point clouds have to be further processed in order to determine which buildings are represented and retrieve their respective shape, i.e. retrieve the objects (Zhu & Shahzad 2014). Second, such a processing suffers of the huge quantity and requirements on the acquisitions, which makes it at this moment unsuitable for global application.

The second trend concentrates more on the global appearance of buildings in InSAR data, trying to extract specific features that distinguish them from the surroundings. The determination of building parameters from the extracted features is then straightforward. Here, it has to be distinguished between multi-aspect and single-aspect InSAR data. In (Bolter & Leberl 2000*a*), four interferograms (from each building side) are acquired and processed. The building footprint is delimited on each side by the shadow, and elevation is retrieved by taking the maximum InSAR height of all four views. Coherence is used in order to distinguish buildings from other high objects. In (Sörgel 2003), building primitives as edges and lines are extracted from the magnitude data and building hypotheses are made by fusion with connected components of the elevation data. Using multi-aspect data in order to suppress layover and shadow effects, this approach performs iteratively in order to retrieve the correct building shape. In (Thiele 2014), primitives are extracted from magnitude and interferometric phase data and fused in the same geometry. Based on the created building hypotheses, InSAR phase signatures (especially layover and shadow) of buildings are simulated and compared to the real data in order to assess the extracted parameters. Those multi-aspect approaches have the advantage of using information from several building sides and allow a better elimination of trees and neighboring effects. However, the recognition of building features in single-aspect data is of paramount importance for each of these approaches. In addition, as for natural disaster management, only using single-aspect data is preferable, due to the time pressure. Finally, depending on the mapped region, sometimes only single-aspect data are available or analyzable. In (Dell'Acqua et al. 2001), the authors propose to detect nearby buildings with different elevations by a joint segmentation and region growing of magnitude and elevation data of a single-aspect image. As pre-processing steps in order to separate building candidates from their surrounding, the density of edge pixels is exploited using the elevation data. This method performs well for the kind of buildings found in industrial areas, i.e. low rise and long flat roof, but shows limitations if applied on buildings with large layover. In (Gamba et al. 2000), a similar segmentation is performed in the elevation data and combined with a plane fitting strategy for ground and roof. It leads to an underestimated building footprint, as layover areas and shadows are not considered. On the contrary, in (Tison et al. 2004), the authors consider the shadow areas for detecting buildings. After the segmentation of the shadow areas in the magnitude image, the building footprint is estimated by considering both magnitude and interferometric phases, minimizing an energy function. This method performs only for isolated buildings, where the shadow is not occluded by surrounding objects. Only very few authors consider layover areas in a single interferometric phase image. In (Petit et al. 2000), the authors exploit the spectral shift between the interferometric image pair in order to separate the vertical signal from the horizontal signals. In (Ferraioli 2010), edges are detected in the phase image at both layover borders. However, the method does not utilize the whole layover part, which could become critical if one edge is hidden from the surroundings. In

(Wegner et al. 2014), double-bounce lines are first extracted from the amplitude image. Then, a line is searched for in the phase image, parallel to the double-bounce line towards near range, presenting maximum phase values. A parallelogram is created consequently, connecting horizontally the extracted parallel lines. Yet, this method involves the extraction of double-bounce lines, not always recognizable, and the subsequent use of both amplitude and phase images. With the TanDEM-X mission, started 2010, whose main purpose is the creation of a global World DEM at a 12 m x 12 m resolution, new approaches have been developed. From the already created TanDEM DEM, some rough estimation of local mean building heights in square grids has been undertaken (Marconcini et al. 2014). A fusion of the TanDEM DEM with a PSI generated DEM has also been performed in order to enhance both geolocation and building height accuracy of the original DEM (Rossi & Gernhardt 2013). Besides, due to the capability of single-pass interferometry, single interferograms resulting from such a bistatic configuration show no time decorrelation (Zink et al. 2008). This is optimal for urban processing. Based on such interferometric data, new methods are currently being developed in order to detect buildings and retrieve their shape. In (Guo & Zhu 2014), the authors first propose a segmentation of building layover candidates based on amplitude and coherence image, before determining their orientation using the fringe frequency of the interferometric phases. Yet, mainly the magnitude image is used for layover detection and shape determination. This has the drawback of being sensitive to changes in building appearance due to incidence angle or material changes, for example. In (Rossi et al. 2014b), a layover map is created by counting the occurrence of a single SAR pixel in several geocoded cells, using the phase discontinuity between master and slave image. The retrieval of the building parameters of single buildings from the created map is not obvious though, as many areas are overlapping.

In this work, for the pre-event analysis, single-aspect InSAR data are considered, as this algorithm should be applicable in regions where often few interferometric data and only a coarse DEM are available. The particularity of this pre-event algorithm is that it relies only on the analysis of interferometric phase images for detecting, extracting and analyzing the buildings. It uses the advantage that buildings have similar appearance in phase images, contrary to their appearance in the amplitude images that can differ depending on the construction material. In particular, focus is put on layover areas, as they are preponderant for the buildings studied in this work. Furthermore, the developed algorithm is able to distinguish single buildings and even single facades from each other, and reconstruct their geometrical shape, so that building parameters can be deduced straightforwardly from the extracted shapes. It is tested on both single-pass TanDEM-X and repeat-pass TerraSAR-X data.

3.3 Overall InSAR Workflow for Building Extraction

Figure 3.3 depicts the workflow of building detection and parameter estimation by InSAR. It can be divided into three main steps, after interferogram calculation: building detection, facade recognition, and reconstruction.

The first step consists in separately applying two detectors on the interferometric phase image. Their aim is to find the aforementioned interferometric fringes in layover areas. They are de-

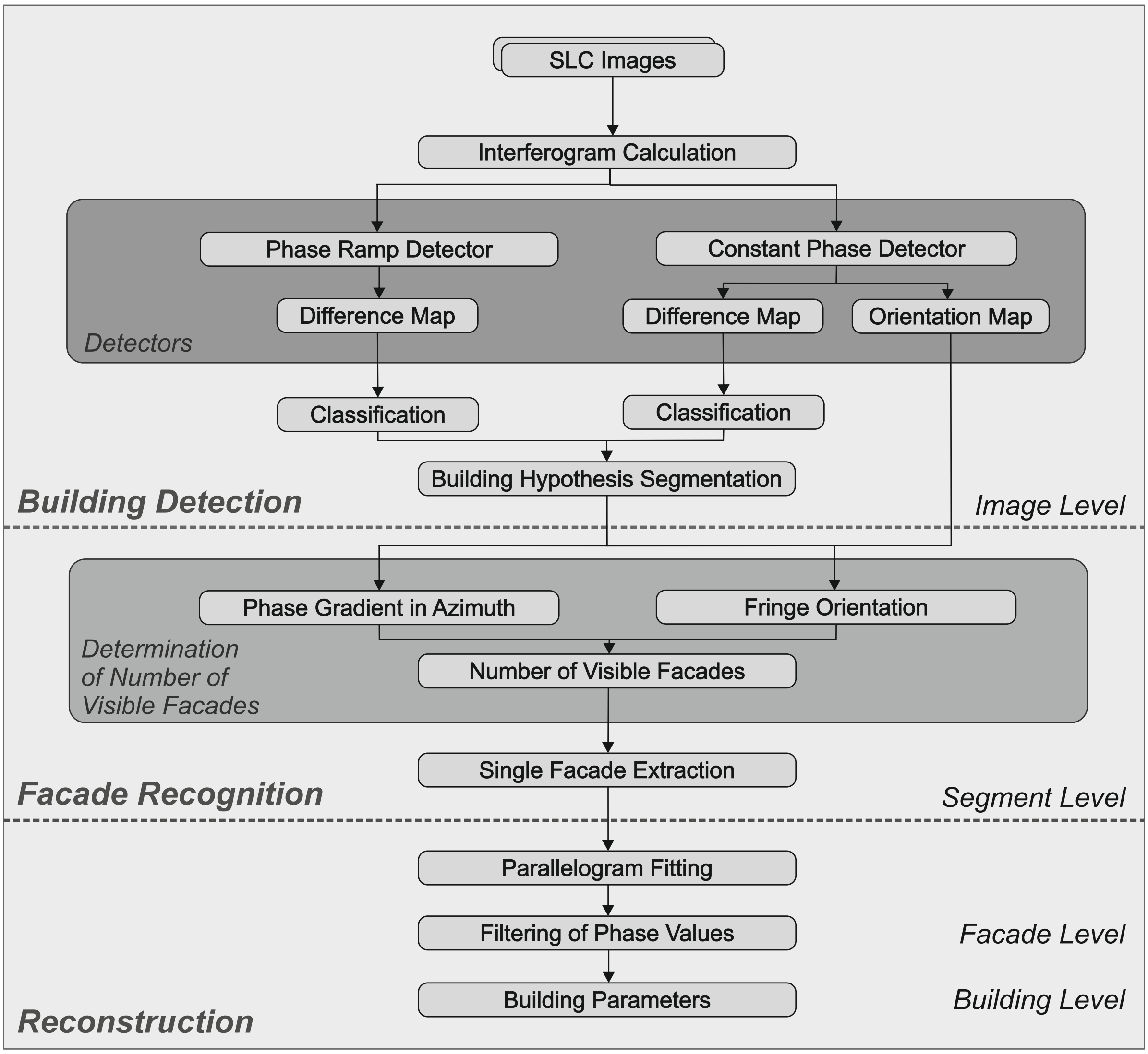


Figure 3.3: Overall InSAR workflow

scribed in Section 3.4. Their outputs are then analyzed, classified and fused, in order to segment building candidates that correspond to building layovers.

Each building hypothesis is then further analyzed individually during the facade recognition step, whereby the number of visible facades is determined in two ways. First, phase values within each segment are considered, and changes in the phase behavior are detected. Indeed, changes in the direction of the phase gradient along the azimuth direction indicate potential change in the facade orientation. Simultaneously, the orientation map, by-product of one of the detectors, is searched for orientation changes within the extracted segments. Considering rectangular buildings, the change of fringe orientation between two consecutive facades should come to 90°. The dimensions of the considered segment decide which of the two methods prevails for the determination of the number of visible facades. At the end of this step, building hypotheses are separated into patches containing single building facades. More details about facade recognition are given in Section 3.5.

Finally, during the reconstruction step, parallelogram shapes are fitted for each facade individually, permitting to reconstruct the building layover. Special cases are also considered (cf.

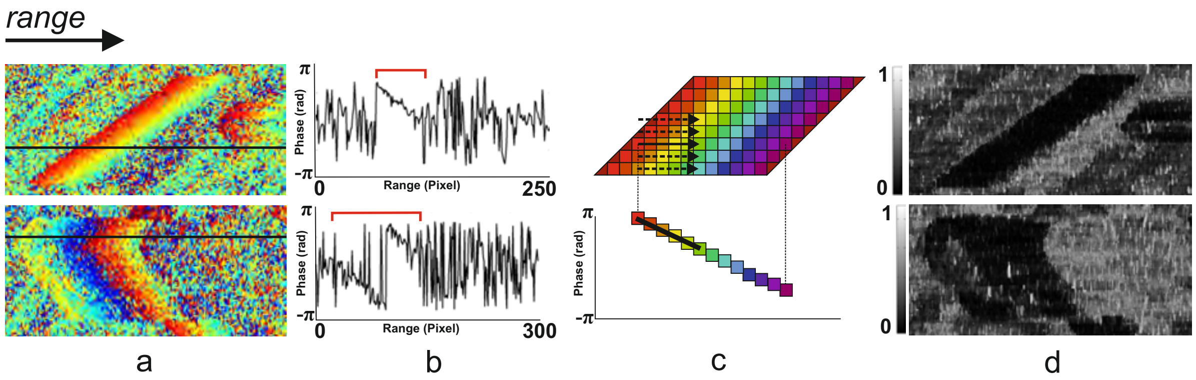


Figure 3.4: Phase ramp detector for one- (first row) and two- (second row) facades buildings: (a) phase image; (b) phase profile along the black line marked in (a) - layover areas are indicated by the red segments; (c) schematic representation of real (colored squares) and synthetic (black line) phase values; (d) resulting difference map

Section 3.6). Within the fitted parallelograms, a filtering is performed in order to smooth the underlying phase values and improve the later height determination. The building parameters are determined subsequently, based on the dimensions of the fitted parallelograms and on the filtered phase values.

The following sections give a detailed overview of the different steps, with the assumption of rectangular shaped, middle-rise buildings with flat roof and predominant layover.

3.4 Building Detection

Building detection focuses on the layover areas, characterized by a descending phase ramp along the range direction, and constant phase values parallel to building edges. The following subsections present the two detectors that are employed in this work, as well as their parameter setting.

3.4.1 Phase Ramp Detector

This detector was first presented in (Thiele et al. 2013) for TanDEM-X data. The principle is recalled here shortly and shown in Figure 3.4. Considering a building situated on flat terrain with vertical walls, the theoretical phase value at each pixel position within the layover is expressed as:

$$\varphi_S(k) = k \cdot \frac{h_a \cos\theta}{d_r} \tag{3.1}$$

whereby d_r is the pixel spacing, h_a the height of ambiguity, θ the incidence angle at the pixel position, and k the pixel position along the descending ramp. The main idea of the phase ramp detector is to compare the expected theoretical phase values with the real data. Starting at a specific pixel position (r, c) within the image, r being the row and c the column value, the detector defines a synthetic phase ramp of specific length L_{SR} along the range direction, corresponding to expected phase values at these positions. Then, the difference between the real phase values $\varphi_r(k)$ (colored squares in Figure 3.4c) and the created synthetic phases $\varphi_S(k)$ (black line in

Figure 3.4c) is built for each position k , and the mean difference is assigned to the pixel position (r, c):

$$D_{PR}(r,c) = \frac{-1}{L_{SR}} \sum_{k=1}^{L_{SR}} |W(\varphi_r(c+k) - \varphi_r(c)) + \varphi_S(k)| \tag{3.2}$$

W is an unwrapping function, as phase jumps due to short height of ambiguity or high-rise buildings are taken into account. The output for the whole image is a difference map D_{PR}, whereby areas with small difference values are good candidates for building layovers. This detector is run twice, in both directions (from near to far range and vice versa), in order to avoid smoothing effects at layover edges. The minimum difference value of both runs is taken at each pixel position as final difference. In Figure 3.4d, the difference map for both presented buildings is given, showing that the phase jump due to the 2π phase ambiguity for the building of the second row has no effect on the result. From Equation (3.1), this detector is defined only for detecting the phase ramp caused by the layover of vertical elements on a flat terrain. Besides, as the incidence angle θ is considered at each pixel position, the synthetic ramp is robustly defined for the whole image, i.e. as well in near as far range.
Considering only differences along one single row tends to produce higher differences, as noisy phase values have an important influence on the difference calculation. Therefore, difference calculations on several rows are averaged at a specific pixel position, as represented in Figure 3.4c with the dotted arrows. Besides, the phase values are weighted by their coherence, so that phase values showing high coherence have more influence on the result as noisy phase values showing bad coherence. On layover borders, the averaging along different rows produces a smoothing of the difference values, as a part of the considered pixels does not belong to the layover. That is why small averaging window sizes should be used. In this work, five rows are used for averaging.
It is obvious from Equation (3.2) that the length L_{SR} of the synthetic phase ramp plays an important role for the detection, as do the height of ambiguity and the incidence angle (cf. Equation (3.1)). Long L_{SR} tend to smooth the difference values on layover borders, whereas short L_{SR} are more influenced by noise. A thorough analysis of these parameters has been carried out in this work for determining their optimal value. This is presented in Section 6.2.1.

In (Thiele et al. 2013), the authors perform a region growing in the resulting difference map, in order to retain segments corresponding to layover areas. The phase image is used in each segment for determining the number of visible facades. Afterwards, the amplitude image is used for extracting the layover borders by Hough transform. This has two main drawbacks. First, simple region growing approaches can be stuck at locations with high intensity changes, which implies that only a part of the layover may be recovered. Second, the Hough transform performed in the amplitude image may be influenced by other narrow objects, leading to a wrong orientation of building edges. The approach developed in this work also uses the phase image for determining the number of visible facades. Changes compared to the approach of (Thiele et al. 2013) are presented in Section 3.5.

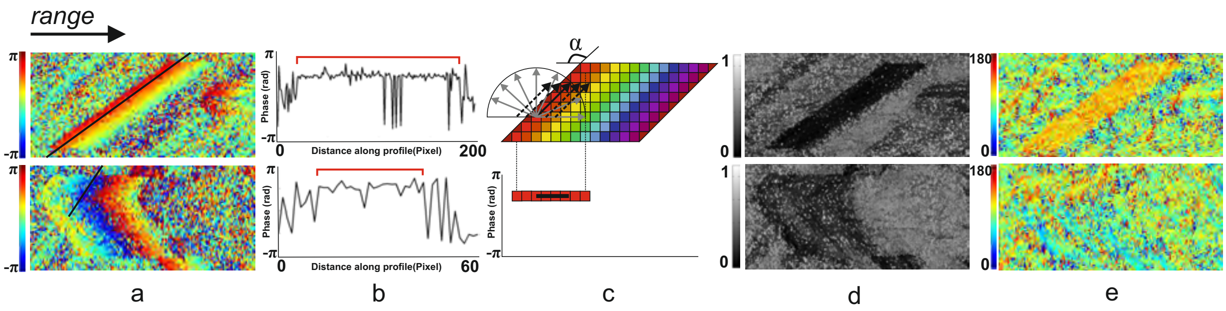


Figure 3.5: Constant phase detector for one- (first row) and two- (second row) facades buildings: (a) phase image; (b) phase profile; (c) schematic representation of real (colored squares) and synthetic (black line) phase values; (d) resulting difference map; (e) resulting orientation map

3.4.2 Constant Phase Detector

Figures 3.4a and 3.5a show the two phenomena observable in building layover. Due to the constant fringe orientation along one facade, building layover in the phase image is marked not only by a phase ramp in range direction, but also by constant phase values parallel to the building edges. The corresponding areas are delimited with the red lines in the profiles of Figures 3.4b and 3.5b, respectively. Constant phase values correspond to constant height values in reality (e.g. floor levels). Developed during this work for completing the phase ramp detector, the idea of the constant phase detector is to detect segments of constant phase values and their directions. Therefore, this detector not only allows to detect the layover in a complementary way to the first detector, but also to give an indication about the direction of the building facade. Moreover, phase ramps could occur on other high-rise objects than buildings, typically on every object producing a layover, e.g. trees. Considering constant phase values and combining both types of information allows therefore to filter potentially false building hypotheses from the phase ramp detector.

Figure 3.5 depicts the principle of the constant phase detector. As for the previous detector, a phase segment is defined from a specific pixel position (r, c). However, the synthetic phase values of this segment are all the same, equal to the phase value of the original pixel $\varphi(r, c)$. Considering a phase profile, this corresponds to a horizontal line, depicted in black in Figure 3.5c. Also for this detector, the mean difference between real phases (red squares in Figure 3.5c) and created synthetic phases along the segment is built. Moreover, several directions are defined for the segment, as the orientation α of the building edge is not known:

$$D_{CP}(r,c) = \min\left(\frac{1}{L_{SR}} \sum_{k=1}^{\sin\alpha\cdot L_{SR}} \sum_{l=1}^{\cos\alpha\cdot L_{SR}} |W(\varphi(r+k,c+l)-\varphi(r,c)|, \alpha\in[0,\pi[\right) \tag{3.3}$$

For better homogeneity of the results and for saving computation time, α is given in $[0, \pi[$ in 5° steps. Indeed, it is not necessary to consider the whole interval $[0, 2\pi[$, as only the orientation of the building edge is of high relevance, and not its direction.

This detector has two outputs: a difference map D_{CP}, similar to the first detector, containing the minimum mean differences between real and synthetic segments for each pixel position, and an orientation map, containing for each pixel the orientation α of the segment for which the minimal mean difference is achieved. Both outputs are represented for the two buildings of Figure 3.5,

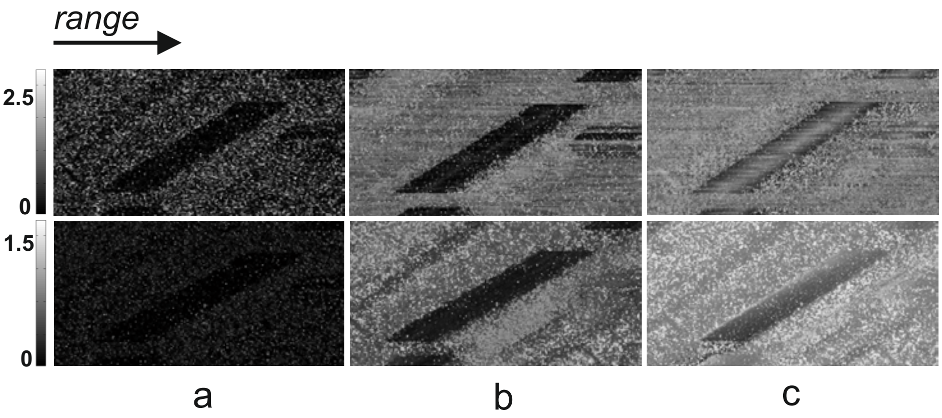


Figure 3.6: Difference maps of phase ramp (first row) and constant phase (second row) detectors: (a) $L_{SR} = 2$; (b) $L_{SR} = 32$; (c) $L_{SR} = 80$

respectively in d and e. For the second building, where two facades are visible, a change of the facade orientation values is observable in the orientation map. Here, the first facade shows homogeneous orientation values around 130° (yellow) whereas the second facade shows orientations around 80° (blue). The orientation map is very useful, as it gives an indication about the local fringe orientation.

As for the phase ramp detector, the phase values are weighted by their coherence values. Moreover, an average on parallel segments is also performed in order to filter outliers. Contrary to the phase ramp detector, where the averaging happens along several rows, the averaging happens here along several segments of the specified orientation α, as represented with the dotted arrows for the correct orientation on Figure 3.5c. For each pixel position along the defined segment, the same amount of pixels is taken left and right of the segment. Every pixel value of this new, wider area is then subtracted from the original phase value and the mean difference is built. Indeed, if the segment has the orientation of the building edge, subtracting the area situated on its left by the same area situated on its right gives the original phase value. The phase behavior along the fringes is assumed linear for flat vertical building facades. The layover borders suffer of such an averaging: if only the left side of the segment is contained in the layover, or vice-versa, the calculated difference becomes larger. In order to reduce this impact, the averaging windows are kept small. In this work, two pixels are taken on each side, in order to produce an averaging window of five pixel width, as for the phase ramp detector in azimuth direction.

Also for this detector, the length L_{SR} of the defined segment is of high importance for the quality of the detection. This is discussed in the next subsection and analyzed in Section 6.2.1.

3.4.3 Parameter Setting

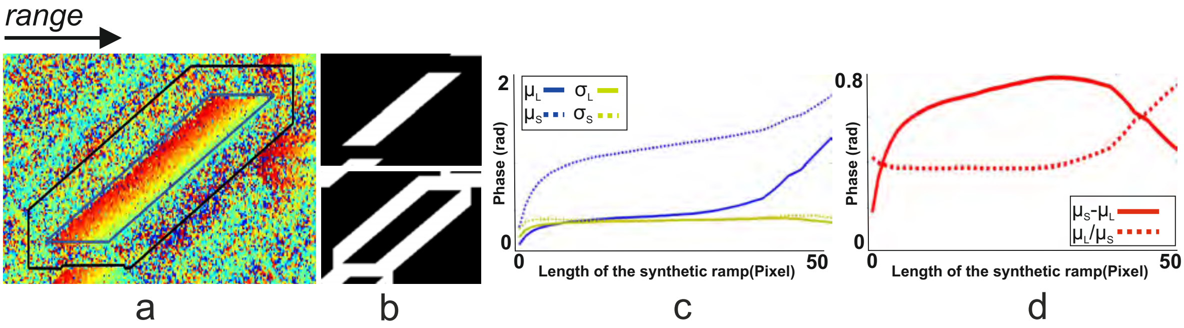
The parametrization of both detectors plays an important role for the subsequent extraction of the layover areas from the difference map. Depending on the length of the synthetic ramps L_{SR}, on the height of ambiguity h_a, i.e. of the effective baseline $B_\perp$, but also on the incidence angle θ, the resulting difference maps have very different appearances. An example is given in Figure 3.6 for both detectors, whereby, for the same acquisition configuration, the length of the synthetic ramps varies. For small lengths L_{SR} (Figure 3.6a), the difference between the real and the synthetic phases is calculated only along a few pixels, and the noise has a high effect. Differences calculated within and outside the layover are similar. For very long L_{SR} (Figure 3.6c), not only the layover, but also the surroundings are taken for the difference calculation, leading to higher differences. It is obvious that the layover is easier to extract with a middle length L_{SR} as shown in Figure 3.6b, as the contrast between layover areas and surroundings is higher.

A strategy for evaluating the best parameter settings was developed during this work and is presented in (Dubois et al. 2015). Training data from different test scenes were considered. For each layover area of a specific test scene, two classes are defined: layover segments, containing only the layovers (area marked in blue in Figure 3.7a), and surrounding segments, corresponding to the direct surroundings of the layover areas (area marked in black in Figure 3.7a). It is ensured that no other building part or disturbing object is contained in the surroundings. The results of both segments are distinctly isolated by multiplying the difference map with these segments separately (Figure 3.7b). As shown previously, a good contrast, i.e. a maximum difference between layover and surrounding is preferred. Several tests involving the mean and standard deviation of the difference values within both classes have been performed. For example, the mean difference within the layover is expected smaller as the mean difference within the buffer. Regarding single assessment parameters separately does not give much insight: small ramps L_{SR} yield low mean values, in both layover and surroundings.

Similarly, small standard deviations are obtained in both layover and surroundings for small ramps and larger standard deviations for larger ramps. These conclusions are shown in Figure 3.7c for the phase ramp detector where a constant increase of mean and standard deviation is observable for both layover and surroundings by increasing the synthetic ramp length L_{SR}. However, considering respective behaviors of layover and surroundings simultaneously permits to analyze their contrast. Both, ratio $\frac{\mu_L}{\mu_S}$ and difference $\mu_S - \mu_L$ of the mean values within the layover μ_L and the surroundings μ_S are built. They are represented in Figure 3.7d, where it is observable that their behavior is opposite. Namely, considering the ratio, the best parameter setting is reached by lower ratio values, i.e. when mean difference values within the layover are small and mean differences in the surroundings are high. In that case, the contrast between layover and surrounding is high. Considering the difference of the mean $\mu_S - \mu_L$, the higher the difference, the better the parameter setting, as layover and surrounding should be as different as possible. With the difference calculation, more distinct optimal lengths L_{SR} are achieved as with the ratio, as shown in Figure 3.7d. Therefore, $\mu_S - \mu_L$ is retained for the setting of the other parameters as well, as shown in Section 6.2.1.

3.4.4 Fusion of Detectors

In this section, the extraction of building hypotheses from the results of both detectors is presented. Both detector results are first analyzed in order to determine the best fusion strategy. Second, specific decision criteria are discussed, which leads to the extraction of building segments.



a b c d

Figure 3.7: Strategy for parameter setting: (a) phase image with layover (blue) and buffer (black) areas; (b) binary layover and buffer areas; (c) results of mean and standard deviation calculation for layover and surroundings, for the phase ramp detector; (d) results of mean difference and mean ratio calculation for the phase ramp detector

Statistical Analysis

The difference maps produced by the detectors vary a lot from one dataset to another, due to different amount of noise and data quality. A simple thresholding of the detector results for extracting building candidates, retaining lower differences values, is not optimal, as the threshold value would change for every dataset and each detector. In the following, an approach is presented that allows to combine both detector results in order to extract building hypotheses, without the use of any manually set threshold.

The combination of both detectors cannot succeed without the analysis of their respective statistics. In this work, one-dimensional histograms are built for each detector result separately, as shown in Figure 3.8 for different datasets. In order to facilitate the comparison between both detectors, the difference maps are normalized, so that all differences are contained in the interval $[0, 1]$. Regarding the created one-dimensional histograms, both detectors present similar distributions for the same considered dataset. In Figure 3.8, the best fitting normal distribution (Hazewinkel 2001) is represented for each histogram (blue lines), and corresponding parameters μ and σ are given. Other distributions, as for example inverse Gaussian (Chhikara 1988), or more typical SAR distributions such as the Rayleigh distribution (Kuruoğlu & Zerubia 2004) could be fitted. However, tests not shown here proved that the distribution that best fits the data is different for each histogram. Therefore, only the normal distribution is represented here. In order to combine both detectors, not only one-dimensional histograms are interesting, but rather the behavior of one detector compared to the other. Therefore, two-dimensional histograms are further built, in order to visualize the joint distribution of both detector results. The intensity of a pixel within the created two-dimensional space represents the number of pixels in the image for which the detectors' outputs are a certain combination (D_{PR}, D_{CP}). If both detector results are correlated, the two-dimensional histogram consists principally of a line with a slope of about 45°, as for every pixel location, and the results of both detectors are the same. Two-dimensional histograms are thus a good indicator for correlation between both detectors. Figure 3.8c shows the two-dimensional histograms of both datasets. Also, their correlation was calculated numerically by:

$$\nu = \frac{\sum_{r=1}^{n}\sum_{c=1}^{m}(X_1(r,c)-\mu_1)(X_2(r,c)-\mu_2)}{\sqrt{\sum_{r=1}^{n}\sum_{c=1}^{m}(X_1(r,c)-\mu_1)^2\sum_{r=1}^{n}\sum_{c=1}^{m}(X_2(r,c)-\mu_2)^2}} \tag{3.4}$$

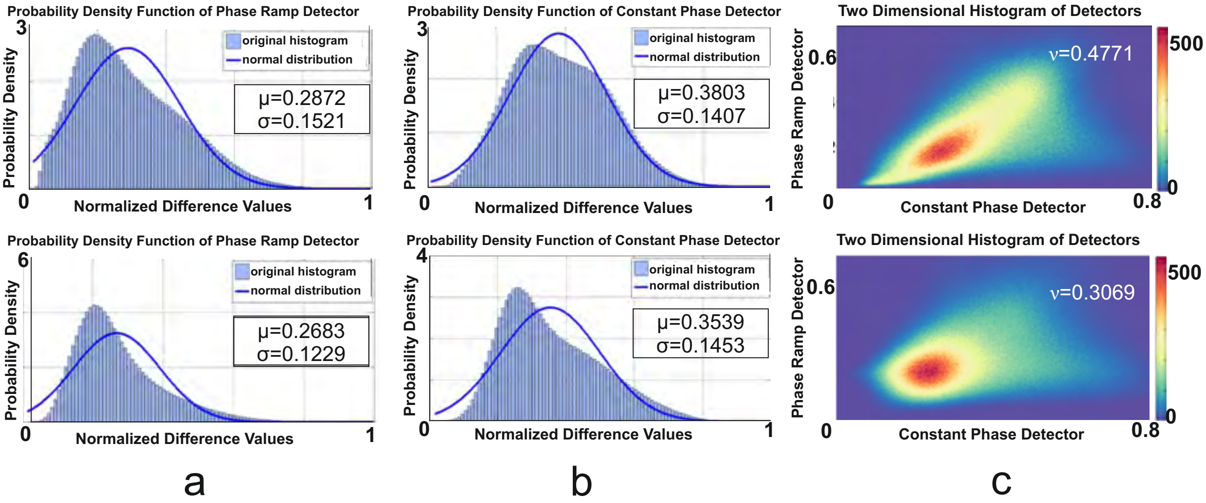


Figure 3.8: Statistical analysis of the detector results for two different scenes (first and second row): (a) 1D-histogram and best-fit normal distribution for the phase ramp detector; (b) 1D-histogram and best-fit normal distribution for the constant phase detector; (c) 2D-histogram

whereby X_1 and X_2 are the difference values of the first and second detector, respectively, and μ_1 and μ_2 their average. n and m are the total number of rows and columns of the image, whereby (r, c) are row and column indices, respectively. There is no correlation between both results when $\nu = 0$ and the highest correlation is given by $\nu = 1$. For the first dataset, the characteristic diagonal shape, representative of correlation between both detectors, is slightly recognizable. The calculated correlation value is however low, explaining the wide spread of values around the diagonal. For the second dataset, the correlation is even worse, which is confirmed by the appearance of the two-dimensional histogram, where no diagonal shape arises, but a disk shape. This disk, or ellipsoidal shape, also shows the concentration of both detector results in a small interval of values, as confirmed by the corresponding one-dimensional histograms of Figure 3.8a and b.

As a conclusion, even if both detectors show similar distributions in the one-dimensional histograms, the combination pairs (D_{PR}, D_{CP}) of detector values at each pixel location are different. As a consequence, both results are independent and complementary, showing the importance of using both detectors.

Considering first each single detector result separately for classification of building layovers and then combining both building classes allows to preserve all relevant information of the single detectors. In the following, the classification of each detector result is explained, as well as the validation of the created classes. Later, the combination of the extracted building classes from both detectors is presented.

Classification and Validation Index

As mentioned previously, every histogram has a different best-fit distribution. The classification of the result image based on its distribution is therefore limited to the considered dataset and detector. This would need supervised classification, as specific values of the distribution parameters should be defined for every class. Due to the supervision by the operator and the large amount of possible distributions, computation time would be highly increased.

In this work, a more generic approach, based on k-means (Hartigan & Wong 1979), is used. k-means aims to partition a set of observations $X = (x_1, x_2, x_3, ..., x_m)$ into k clusters $C = (C_1, C_2, ..., C_k)$ whereby $k < m$, by minimizing the within-cluster sum of squares. An initialization step is required to set the center of the clusters, that are updated each time a new observation comes into the cluster. Considering the detector results in city areas, the optimal number of classes is not known. At least two classes exist, namely ‚building layover‘ and ‚surroundings‘. Particularly for the phase ramp detector, at least three different classes can be visually distinguished, which could be assimilated to ‚building layover‘, ‚roads and other objects‘ and ‚shadows (noise)‘. Therefore, the k-means clustering is performed for $k = [2, 5]$: in urban areas, a limitation of the number of classes to five seems a good compromise, in order to maintain homogeneity of the results.

For assessing the classification results, several validation indices, or criteria, exist. The final optimal number of classes is estimated by analysis of those criteria. They can be separated into external, internal and relative validation indices. Whereas external indices are used to quantify the quality of the classification with respect to externally given class labels, internal indices evaluate the classification without using external information. Relative indices are used to compare two different classes or classifications, and are therefore preferred in this work, as the goal is to distinguish the building layover class from the others. Different validation criteria are used: silhouette, weighted inter-intra ratio, separation and cohesion:

- The silhouette coefficient quantifies the similarity of an observation x_l to the other observations within its cluster, compared to observations in the other clusters (Kaufman & Rousseeuw 2009). It can be written as:

  $$S_l = \frac{b_l - a_l}{\max(a_l, b_l)} \tag{3.5}$$

  whereby a_l is the mean distance from observation x_l to the other observations of its cluster, and b_l the mean distance of x_l to observations in the statistically ‚nearest‘ cluster. This silhouette coefficient gives values in the interval $[-1, 1]$, whereby 1 stands for homogeneous, well separated classes. For evaluating the optimal number of classes, a mean silhouette coefficient is built over all observations.

- The weighted inter-intra ratio (Strehl 2002) is defined by

  $$W(X, \lambda) = 1 - \frac{\sum_{C_i=1}^{k} \sum_{C_j \in 1,...,C_i-1,C_i+1,...,k} n_{C_j} \cdot \mathrm{inter}(X, \lambda, C_i, C_j)}{\sum_{C_i=1}^{k} n_{C_i} \cdot \mathrm{intra}(X, \lambda, C_i)} \tag{3.6}$$

  whereby

  $$\mathrm{inter}(X, \lambda, C_i, C_j) = \frac{1}{n_{C_i} n_{C_j}} \cdot \sum_{\lambda_a = C_i, \lambda_b = C_j} s(x_a, x_b) \tag{3.7}$$

  and

  $$\mathrm{intra}(X, \lambda, C_i) = \frac{2}{(n_{C_i} - 1) n_{C_i}} \cdot \sum_{\lambda_a = \lambda_b = C_i, b > a} s(x_a, x_b) \tag{3.8}$$

  X is a vector containing the observations, λ is a vector containing the cluster value for each observation, C_i and C_j are different cluster indices, k is the number of clusters, n the

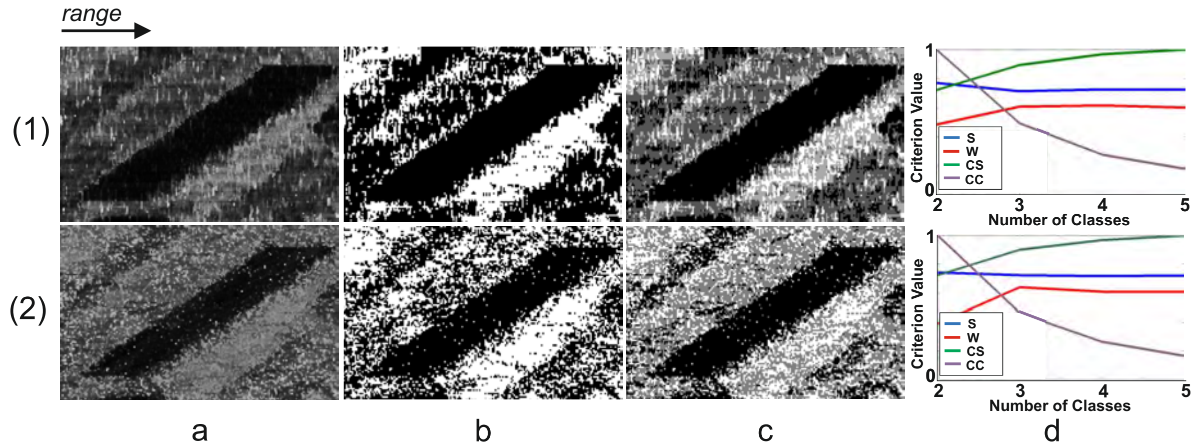


Figure 3.9: Choice of the number of classes by the validation indices for phase ramp detector (1) and constant phase detector (2): (a) difference map; (b) classification result with k-means for $k = 2$; (c) classification result with k-means for $k = \max(W)$; (d) value of the validation indices for $k = [2, 5]$

number of elements within the corresponding cluster and s is a similarity function. In this work, the Euclidean distance is used as similarity function. W is contained in the interval $[0, 1]$ whereby $W = 1$ indicates that samples from different clusters have a zero similarity, and that at least two samples from the same cluster have a non-zero similarity, which is preferred. In other terms, the higher W, the higher the intra cluster similarity and the smaller the inter cluster similarity.

- Both silhouette and weighted inter-intra ratio are measures that attend to compare the homogeneity, or cohesion, of the clusters with their separation. Therefore, the last evaluated indices are cluster cohesion CC and cluster separation CS, defined as follows:

$$CC = \frac{1}{k} \sum_{C_i=1}^{k} \sum_{l=1}^{n_{C_i}} (x_l - \mu_{C_i})^2, \text{ and } CS = \frac{1}{k} \sum_{C_i=1}^{k} |n_{C_i}| (\mu_{C_i} - \mu)^2 \qquad (3.9)$$

whereby μ is the mean value of all observations x_l. A low CC is preferred, as it means that each observation x_l within a cluster is statistically ‚near‘ the cluster mean value. On the contrary, high CS values are searched, as this corresponds to better separation of the cluster centers.

Considering the previous conclusions and Figure 3.9, looking only at cohesion and separation would lead to chose $k = 5$ as optimum number of classes. Lower values of CC are reached with an increasing number of classes (grey lines). Indeed, the more classes, the better the homogeneity within each class. Consequently, the separation capacity increases by increasing the number of classes, whereby it seems to converge to a maximum value for a higher number of classes. However, considering the silhouette parameter, better results are obtained for $k = 2$, for both detectors. Considering the corresponding classified images in Figure 3.9(1,2)b, it is obvious that a large amount of the surroundings is in the same class as the building layover (dark areas). The last criterion, W (red lines in Figure 3.9d) shows a maximum at $k = 4$ for the phase ramp detector and at $k = 3$ for the constant phase detector. Corresponding classified images

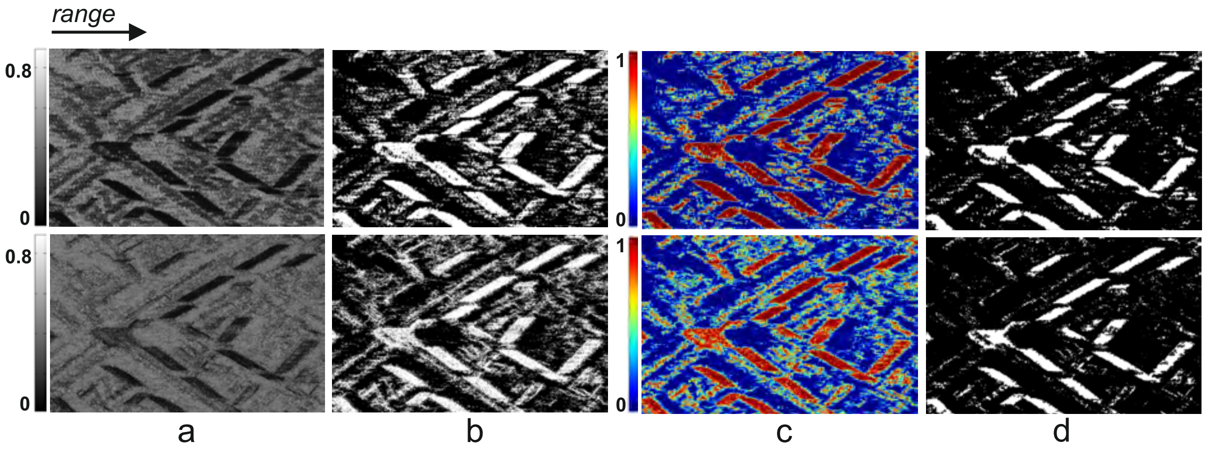


Figure 3.10: Separated post-processing for phase ramp detector (first row) and constant phase detector (second row): (a) difference map; (b) building class after k-means; (c) density map; (d) building class after density filtering

are presented in Figure 3.9(1)c and (2)c, respectively. Here, the building layover (dark areas) is well separated from its surroundings (dark grey to white areas). The ‚shadows/noisy‘ areas are separated into two classes, which are not relevant, as only building layover are of interest. However, this shows that $k = 4$, respectively $k = 3$ as given by the criterion weighted inter-intra ratio is a good compromise. Therefore, in this work, the weighted inter-intra ratio is the index that is used for determining the optimal number of classes for the classification of the detector results using k-means, i.e. $k = \max(W)$ is used.

The classification of each difference map is performed with the defined optimal k for each detector, and the building class is extracted. This later step is straightforward, as from the detector definition, the detector values should be minimal at the building location. The building class has thus the smallest cluster center. In Figure 3.10b, the extracted building class for both detectors is represented in black. The building layovers are well recognizable, yet noise subsists. For extracting only building hypotheses, post-processing is needed.

Segmentation of Single Building Hypothesis

The post-processing of the building class happens in two steps: post-processing for both detectors separately (Figure 3.10), and post-processing after fusion of both results (Figure 3.11).

In order to filter the persisting noise from the surrounding, a density map is first produced for each detector (Figure 3.10c). For each pixel, the pixel density is calculated, which corresponds to the percentage of building pixels in a 3x3 surrounding patch. Only pixels with density strictly higher than 0.5 are retained, i.e. more than half of the patch contains pixels belonging to the building class. This step is necessary before fusing both detector results, as summing up before would produce very dense surroundings that could not be filtered as easily. The final building class is shown in Figure 3.10d for both detectors, whereby holes within closed facades are filled. Here, for both detectors, the main part of the surrounding noise has been removed.

The fusion of both detectors happens then by adding both final building classes, relying on their complementarity. The fusion result is given in Figure 3.11b. Considering the original phase image, shown in Figure 3.11a, it is obvious that all well defined building layover are retrieved. Having a closer look at the still noisy areas, roads can be distinguished, principally due to their

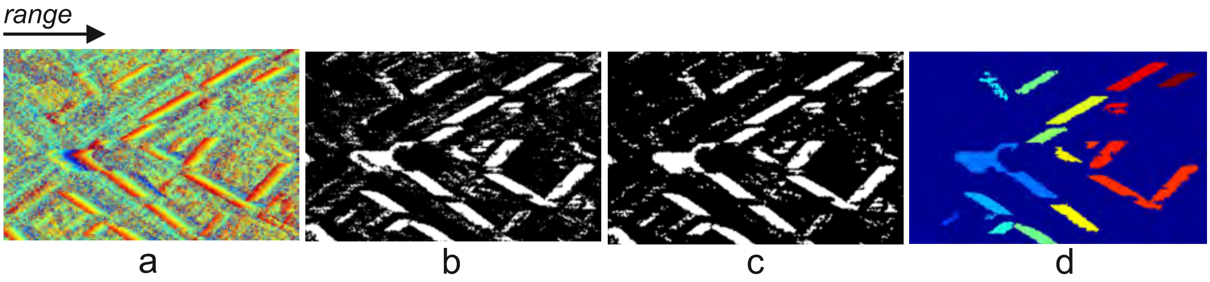


Figure 3.11: Fusion and extraction of building hypotheses; (a) original phase image; (b) fused building class from both detectors; (c) fused building class after morphological filtering; (d) final building hypotheses

constant phase, detected by the second detector. Furthermore, some holes in building structures are visible, as for the high-rise building situated on the middle-left part of Figure 3.11b.

The next post-processing step consists in performing a filtering of the fused image (Figure 3.11b) by using morphological operators. For this, two different structure elements are used, consisting in a vertical and then a horizontal line of two pixel width. With TerraSAR-X Very High Resolution Spotlight, the structure elements correspond to about 1.6 m in azimuth and 1 m in range direction. Both are applied subsequently, using opening operation. Their small size allows to keep most information on layover borders and to sharpen the layover boundaries, while eliminating small noise. Subsequently, remaining holes are filled, in order to obtain plain building facades. The result is shown in Figure 3.11c, whereby layovers are more dense and have sharp border lines. Moreover, noise of the roads mainly disappeared. Finally, based on this result, connected components of a specific size are searched for. Even if the searched building parameters are not known, a minimal expected layover size can be defined. Knowing the area of interest, land-use maps are often available that give an indication about the kind of building situated in the search area (single residential houses, multiple dwelling, office buildings, financial district...). From this indication, coarse expected building dimensions are established. The final extraction result is shown in Figure 3.11d, whereby each color corresponds to a different component. Here, not directly connected components, but connected components within a certain neighborhood, are searched for. Components having a small gap between each other can be combined to a single component. In this work, a neighborhood of 7x3 (7 pixel in azimuth direction and 3 pixel in range direction) is considered for all datasets. For high resolution Spotlight TerraSAR-X data, this corresponds to about 6 m in azimuth direction and 1.5 m in range direction. Thus, possible layover gaps are taken into account. Such gaps are principally caused in azimuth direction by trees situated in front of the building or by specific building structures in the facade that could alter the phase fringes. This parameter is of high importance for the datasets shown in Section 6. In suburban areas, a distance of about 6 m between buildings is reasonable. If smaller distances occur, it is often due to joint building blocks, handled as a single building.

Some conclusions can be drawn at this point, considering the final image (Figure 3.11d). All the noise has been removed, and the extracted connected components all correspond to building layovers. Even if most connected components represent only single facades or buildings, some connected components seem to contain more than one building, as for the middle blue and red ones. Those buildings are very close to each other and were already connected after the building

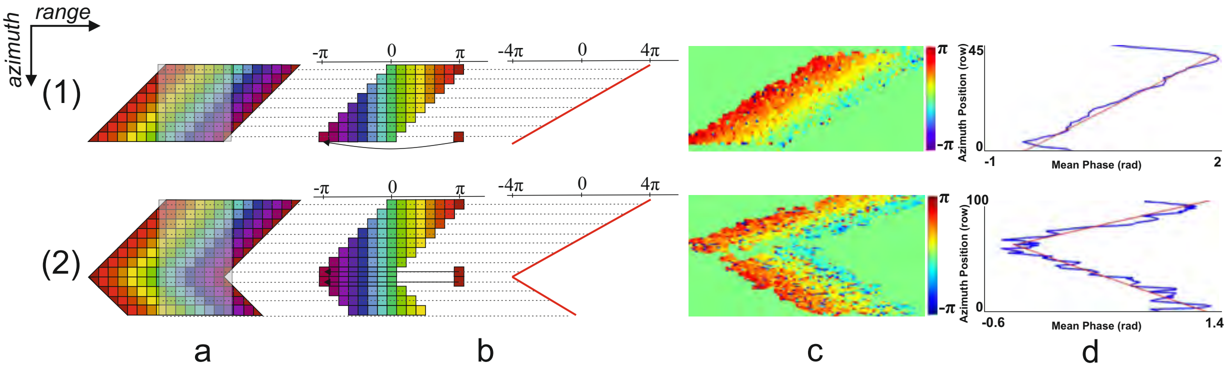


Figure 3.12: Strategy of ridge detection: (1) one-facade; (2) two-facades; (a) schematic representation of phase values with area of interest marked in grey; (b) phase values within the area of interest and average; (c) real phase values; (d) real average of phase values (blue) and fitted lines (red)

classification of both detectors. They are taken into account in the further processing for the extraction of single facade parameters (cf. Section 3.6.3).

3.5 Facade Recognition

The second step of building detection and reconstruction with InSAR performs facade recognition on the extracted building segments. Depending on the acquisition configuration and on the building orientation towards the sensor, one or two facades may be visible. In order to prevent false building reconstruction and complete the building information, it is necessary to determine the number of visible facades for each building segment. In the case of two visible facades, the building can be entirely reconstructed, as a rectangular shape is assumed for each building. The strategy of facade recognition is explained in the following, whereby two different analyses are performed and combined. The first consists in ridge detection in the phase image, and the second in fringe orientation analysis, using the orientation map produced by the constant phase detector. Depending on the shape and dimension of the building hypothesis, the outcomes are combined differently in order to extract individual facades.

3.5.1 Ridge Detection

The ridge detection is explained schematically in Figure 3.12 and was partly described in (Dubois et al. 2014). Considering a single column within the area marked in grey, phase values are descending from one building edge to the other, regarding only a single facade, as represented in Figure 3.12(1). However, by two-facades buildings, the phase values increase again within the second facade, forming a ridge along the azimuth direction. Its position corresponds to the building edge between both facades. The ridge detection step aims at detecting this valley, by analyzing the phase gradient parallel to the azimuth direction.
Practically, the approach consists first in creating a bounding box for each building hypothesis. Within this bounding box, only phase values that belong to the building layover are taken into account for further processing (Figure 3.12c). Therefore phase values and noise from the surroundings have no influence on the result. First, the necessity of performing a phase unwrapping

is tested. For this, the phase values along the azimuth direction are summed up and averaged in each column. The minimal and maximal value of the obtained profile are then analyzed. A difference higher than π means that a phase jump occurs and unwrapping is necessary. Unwrapping is performed with the Goldstein method (Goldstein et al. 1988), yielding continuous phase values. In a second step, ridge detection is performed. Detecting the ridge considering only one column could be influenced by noise, thus the phase values are summed up and averaged along the range direction within a specific area, marked in grey in Figure 3.12. This area is centered on the centroid of the building hypothesis. Its dimension corresponds to the distance between centroid and the nearest border, in near or far range. This summation is repeated for each row of the bounding box (Figure 3.12b), so that the ridge detector is described as:

$$M_{\varphi} = \sum_{k=1}^{r} M_{\varphi,k} \tag{3.10}$$

with

$$M_{\varphi,k} = \frac{1}{w \cdot \sum_{l=c_m-\frac{w}{2}}^{c_m+\frac{w}{2}} \gamma(k,l)} \sum_{l=c_m-\frac{w}{2}}^{c_m+\frac{w}{2}} \varphi(k,l) \cdot \gamma(k,l) \tag{3.11}$$

for a specific row k. c_m is the column position of the center of mass of the image, and $w = \min(c_m, c - c_m)$. The row and column position of the considered pixel are k and l, respectively, and (r, c) is the total number of rows, respectively columns, in the bounding box. φ represents the phase values. $\gamma(k,l)$ represents the coherence at the considered pixel position, showing that each phase value is weighted by its coherence during summation, reducing the influence of noisy phase values.

The mean of the phase value for each row k, i.e. each azimuth position, is represented in red in Figure 3.12b for schematic and in blue in Figure 3.12d for real data. The ridge can be well recognized for the two-facades building, whereby the one-facade building is characterized by a linear decreasing phase behavior. The decision if one or two facades are visible is made by fitting lines to the curves of mean phase values. For the line fitting, linear regression is used. For each position k along the vector of mean phase values, two polynomials of first degree $y_1 = a_1x_1 + b_1$ and $y_2 = a_2x_2 + b_2$ are fitted in a least squares sense. The first line has the first vector element as start position and the considered vector position k as end position $(x_1 = [1 : k])$. The second line has the considered position k as start position and the last vector element as end position $(x_2 = [k : r])$. For each azimuth position, the root mean square error ($RMSE$) between the fitted lines y_1 and y_2 and the mean phase values M_{φ} is evaluated. The position k showing the minimal $RMSE$ is retained as possible building edge between both facades. The final decision about the number of facades depends on the position of the minimal $RMSE$, as well as of the signs of the coefficients a_1 and a_2. A building hypothesis is considered containing two facades if both following conditions are fulfilled:

- The position k of the minimal $RMSE$ should be at least 6 pixels away from the borders of the bounding box, in azimuth direction. Indeed, a facade can only be retrieved as such if it has a minimal size. 6 pixels in azimuth direction corresponds for HR Spotlight TerraSAR-X data to about 5 m in azimuth direction, and is reasonable for a minimal facade length, without considering its orientation and thus its extension in range direction.

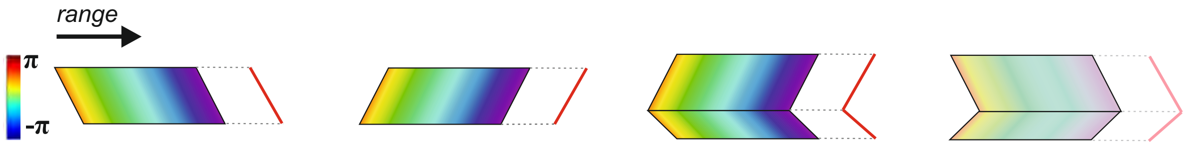


Figure 3.13: Different possible appearances of buildings in InSAR data

- Both fitted lines should have coefficients a_1 and a_2 of opposite signs, as the facades orientations change. Considering the appearance of buildings resulting from the side-looking geometry, further restriction can be made, as represented in Figure 3.13. Only three possible behaviors of the mean phase values exist. The first two represent one-facade buildings with different orientations. The behavior of the summed phase values is linear decreasing or increasing, depending on the facade orientation. The third represents the two-facades building, whereby first a decrease and then an increase of the phase values occur. The last appearance, with an increasing behavior followed by a decreasing behavior could occur for example at inner courtyards or for adjacent or L-shaped buildings. This case was not considered in this work, as only rectangular shaped buildings were analyzed. Therefore, for two-facades buildings, not only both coefficients a_1 and a_2 should have opposite signs, but the first coefficient should be negative ($a_1 < 0$) and the second positive ($a_2 > 0$), taking the azimuth direction as axis. In every other case, the building is considered having only one visible facade.

The separation into segments containing individual facades is then straightforward: the building edge between both facades corresponds to the position of the ridge, i.e to the position of the minimal $RMSE$. Also for buildings with only one visible facade, a refinement is performed, whereby the position of the minimal $RMSE$ is taken as new border for the building hypothesis. This allows to eliminate noise at the upper or lower borders, as can be seen in Figure 3.12(1) for the one-facade building.

3.5.2 Fringe Orientation Analysis

A schematic representation of the fringe orientation analysis for detecting multiple facades is given in Figure 3.14 with corresponding real data. For this analysis, the orientation map, created during the detection of constant phases, is used. The change in orientation of the fringe pattern is analyzed. For vertical, straight facades, each facade is characterized by a specific orientation. It is well recognizable in Figure 3.14(a,b), where each color represents a specific orientation. The one-facade building shows a dominant yellow-orange color, corresponding to an orientation of about 130°. In comparison, the two-facades building shows two dominant colors, depending of the considered facade: the upper facade shows similar orientation as the previous building, whereas the second facade shows dominant blue-cyan color, corresponding to an orientation of about 80°. Averaging all orientation values for each row leads to the profiles represented schematically in red and for the real data in blue in Figure 3.14(a,d), respectively. For the one-facade building, the profile is continuously vertical, whereas an orientation jump is observed for the building with two facades. This discontinuity is situated at the edge between both facades. The goal of the fringe orientation analysis is to determine how many different main orientations

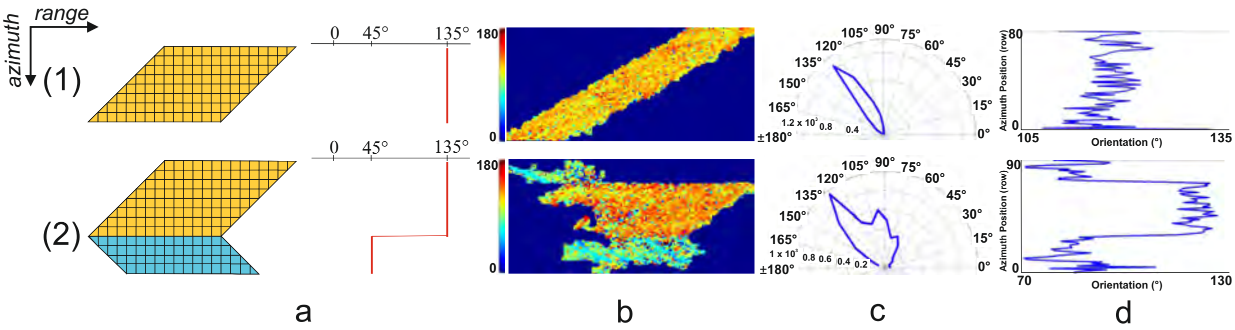


Figure 3.14: Strategy of fringe orientation analysis: (1) one-facade; (2) two-facades; (a) schematic representation of orientation values and average; (b) real orientation values; (c) histogram of orientations; (d) real average of orientation values

exist in the building hypothesis, and find the location of the potential change of orientation. Practically, as for the ridge detection, only orientations within the detected layover area are used (Figure 3.14b). However, instead of considering only a reduced area (grey area of Figure 3.12a), the whole building hypothesis is taken into account. First, a histogram of the orientation values is built, whereby each orientation is weighted by its corresponding coherence value, so that orientations with low coherence do not have high influence on the histogram result. The polar histograms are represented Figure 3.14c for the two building types. For the one-facade building, only one peak is visible, around 130°. For the two-facades building, several peaks can be distinguished. For each created histogram, the peaks are analyzed. Up to three peaks are searched for, and the main peak is considered corresponding to the orientation of the main facade. If only one peak is detected, the building hypothesis is considered representing only one facade, whose orientation is the one of the detected peak. If more than one peak is detected, the angles between the side peaks and the main peak are calculated. The orientation given in the histogram corresponds to slant range orientations. The angle difference is therefore in slant geometry. In order to determine the angular difference between two peaks on the ground, it is necessary to consider some geometric relations, depicted in Figure 3.15. Considering the orientation difference $\Delta\alpha''$ formed by the angles α''_1 and α''_2 in slant range ($\Delta\alpha'' = \alpha''_1 + \alpha''_2$), the following equations can be written:

$$X' = X'' \cdot dr = \cos\alpha''_i \cdot dr, \text{ and } Y' = Y'' \cdot da = \sin\alpha''_i \cdot da \tag{3.12}$$

whereby dr and da are pixel spacing in range and azimuth direction, respectively. i takes values of 1 or 2 depending on the considered angle. X' and Y' are the values in meter of X'' and Y'' in slant geometry.

Projecting now X' and Y' on the ground (Figure 3.15b), the following equations are obtained:

$$X = \frac{X'}{\sin\theta} = \frac{\cos\alpha''_i \cdot dr}{\sin\theta}, \text{ and } Y = Y' = \sin\alpha''_i \cdot da \tag{3.13}$$

whereby X and Y are the representations of X' and Y' in ground geometry, and θ represents the incidence angle.

Finally, the angles α_i on the ground (Figure 3.15c) can be expressed as:

$$\alpha_i = \arctan\frac{Y}{X} = \arctan\frac{\sin\alpha''_i \cdot \sin\theta \cdot da}{\cos\alpha''_i \cdot dr} \tag{3.14}$$

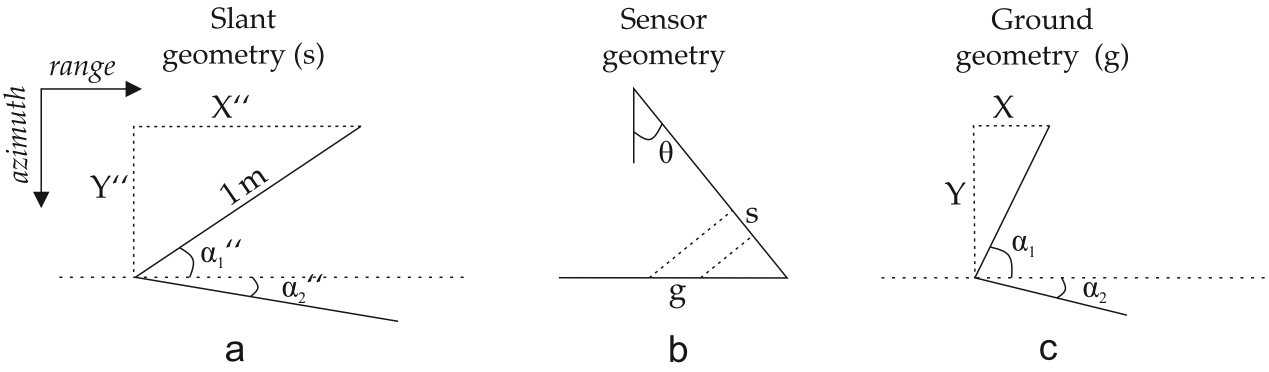


Figure 3.15: Angle difference between slant and ground range geometry: (a) slant range geometry; (b): sensor geometry; (c) ground range geometry

Therefore, the ground angle difference $\Delta\alpha$ can be expressed as:

$$\Delta\alpha = \alpha_1 + \alpha_2 = \arctan(\tan\alpha_1'' \cdot k) + \arctan(\tan\alpha_2'' \cdot k) \tag{3.15}$$

whereby

$$k = \frac{\sin\theta \cdot da}{dr} \tag{3.16}$$

Considering these equations, it is straightforward to determine the ground angle difference between the main peak of the histogram and its side peaks. In this work, as rectangular shaped buildings are considered, ground angle differences of 90° are looked for. However, as the side facades are often thinner than the main facades and their orientation values are more noisy (cf. Figure 3.14b), an offset of ±10° is allowed, to the original difference of 90°. If one of the side peaks fulfills this condition with the main peak, the building hypothesis is considered representing two facades.

For two-facades buildings, the position of the building edge is determined using the profile of the averaged orientation values (Figure 3.14(a,d)). The mean orientation calculated for a row r is subtracted from the mean orientation of its direct neighbor $r+1$. Within the difference profile, peaks are searched for. As for the ridge detection, the possibilities presented in Figure 3.13 are considered. Expected phase jumps are positive, as the upper facade has an orientation value higher than the lower facade (Figure 3.14b). The position of the highest positive orientation difference gives the position of the building edge. Considering the example in Figure 3.14(2)d, the first orientation jump is not considered as possible building edge, because the orientation difference is negative. It is due to the surroundings belonging to the building hypothesis. However, such a jump is detected and used for refining the bounding box of the layover, eliminating the corresponding surroundings.

3.5.3 Combination

Due to phase and orientation noise, both methods do not always yield the same conclusions about the number of visible facades for a specific building. In order to provide reliable single facade segments, rules are defined for combining both methods. For this purpose, the dimension of the bounding box of each building hypothesis is used. The height h_{BB} of the bounding box represents its extent from near to far range, and its width w_{BB} its extent in the perpendicular azimuth direction. If both methods yield the same conclusion, the combination is trivial. However, in case both methods disagree and give contradictory information, two possibilities are evaluated:

- $w_{BB} > h_{BB}$: in this case, the result of the ridge detection is considered. Indeed, the ridge detection is more reliable if the lines are fitted on a larger set of values, only possible if w_{BB} is large. Moreover, only a thin area is used for averaging the phase values in range direction (grey area in Figure 3.12), therefore h_{BB} can be small.
- $h_{BB} > w_{BB}$: in this case, the fringe orientation analysis is taken into account. Indeed, the detection of the orientation jumps relies on the good averaging of the orientation values along each row, which is more reliable with a large amount of orientation values.

The choice of the method to apply hence directly depends on the dimensions of the building hypothesis. Subsequently to the determination of potential adjacent facade and position of the building edge, single facade segments are extracted, and new surrounding bounding boxes are defined for each single facade (w_{BB_f}, h_{BB_f}).

3.6 Building Reconstruction

The last step of the InSAR processing chain consists in reconstructing the geometry of the facades, in order to extract building parameters: orientation, height, length and potential width. The different stages, developed in the following, consist in fitting parallelogram shapes to single facades and filtering the phase values within the parallelograms. Then, robust building parameters can be deduced.

3.6.1 Parallelogram Fitting

The building analysis of Section 3.1 outlined the parallelogram shape of each building facade, assuming rectangular shaped buildings and side-looking geometry. The horizontal edges, parallel to the range direction, correspond to the layover length, and give an approximate building height h. The complementary edges are parallel to the building orientation and give its length l or width w, depending on the considered building. For two-facades buildings, the relative size of the complementary edges from one facade compared to the other will decide which is the main facade of length l and which is the side facade of width w. The given facade orientation corresponds to the angle α formed by the edges parallel to the building facade with the range direction. All parameters are represented in Figure 3.16(b,d).
Starting from the single facade segments extracted in the previous step, parallelogram shapes are fitted by performing an affine transform of a rectangular template that has the size of the facade bounding box. In this case, three parameters are used for transforming the rectangle into the parallelogram:

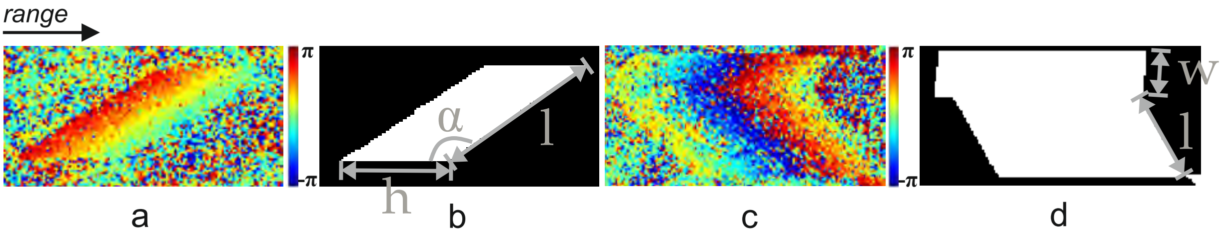


Figure 3.16: Parallelogram fitting and building parameters: (a) one-facade building in phase image; (b) fitted parallelogram for one-facade building with corresponding building parameters; (c) two-facades building; (d) fitted parallelograms with remaining parameters

- α_x: this is the main parameter, allowing shearing of the rectangle into a parallelogram. This shearing is applied only in the horizontal (range) direction, as two edges stay parallel to their original range direction.
- s_x: this is a scale factor, in horizontal direction, allowing the reduction of the size of the template, so that it fits better the building segment.
- t_x: the last parameter is a translation parameter, necessary to adjust the parallelogram position within the bounding box in order to reduce the influence of residual noise at the border of the segments.

The affine transform can be written as:

$$\begin{pmatrix} x_p \\ y_p \\ 1 \end{pmatrix} = \begin{pmatrix} s_x & \alpha_x & t_x \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \tag{3.17}$$

whereby (x_p, y_p) are the transformed coordinates of original pixel position (x, y).

The best fit is obtained by least squares matching (LSM), whereby approximate values for all parameters are given as follows:

- α_0: as approximate orientation value, the mean facade orientation deduced from the histogram analysis of the orientation map is taken.
- s_{x0}: for the determination of the approximate scale factor, the number of segment pixels n_r in each row k of the facade bounding box is first determined. The scale factor s_{x0} is then determined by

  $$s_{x0} = \frac{1}{w_{BB_f} \cdot h_{BB_f}} \sum_{k=1}^{w_{BB_f}} n_r(k) \tag{3.18}$$

  whereby w_{BB_f} and h_{BB_f} are the dimensions of the facade bounding box.
- t_{x0}: the translation is set to 0, as the original facade segment is expected to fit exactly into the bounding box.

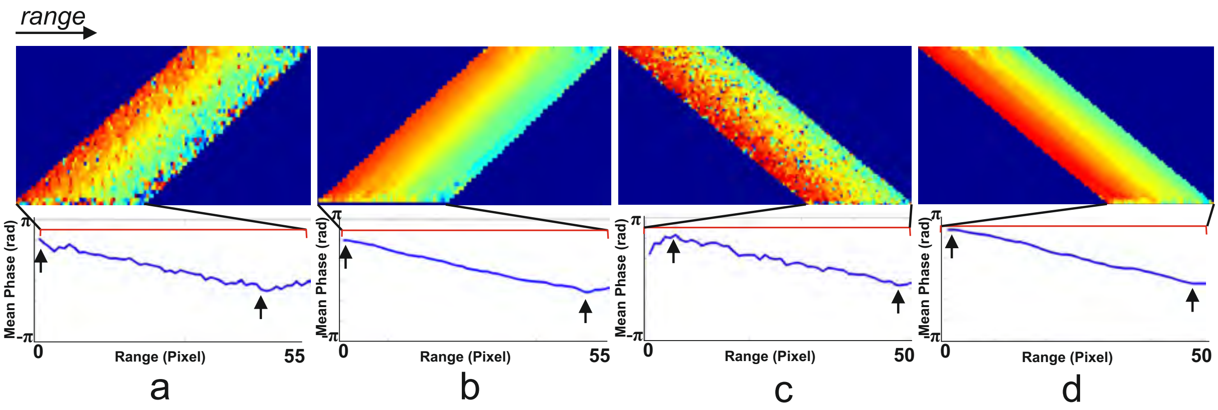


Figure 3.17: Filtering of phase values within the extracted parallelograms, with corresponding average of phase values along the building orientation - the positions of minimal and maximal average phases are marked with arrows: (a,c): real phase values; (b,d): filtered phase values

Figure 3.16(b, d) shows the fitted parallelograms for the two building types. Considering only the geometry, it is possible to deduce all building parameters. However, the parallelogram shapes tend to encompass the segments, leading to higher buildings. The analysis of the phase values within the extracted parallelograms allows to obtain supplementary information that make the height estimation more robust.

3.6.2 Phase Analysis and Filtering

Figure 3.17 shows the phase values within the extracted parallelograms. As stated in Section 2.3.4, the calculation of the height from the phase values is possible. The phase difference $\Delta\varphi$ between building foot and roof provides their height difference, and thus the building height. Practically, the phase values within the fitted parallelograms are summed up and averaged along the building orientation α. Corresponding unwrapped mean phase profiles are shown in Figure 3.17(a, c). At the building foot, in far range, the phase value is minimal, whereas at the roof, in near range, the phase value is maximal. The minimum and maximum phases are marked by the black arrows in the profiles of Figure 3.17. It is already observable from the profiles that using only the parallelogram dimensions would lead to higher values than the real building height, as phase values not belonging to the building layover appear at near and far range. The detection of the minimal and maximal mean phase value is trivial, which leads to the phase difference $\Delta\varphi$, from which the building height can be estimated:

$$h = \frac{\Delta\varphi \cdot h_a}{2\pi} \tag{3.19}$$

However, in order to reliably estimate minimal and maximal phase values, noisy phases should be filtered first. In this work, adaptive filtering as presented in (Dubois et al. 2012) is used within the extracted parallelogram. Thin dynamical filter masks are created parallel to the building edge, whose vertical dimension is adapted depending on the pixel position relative to the bounding range lines. Due to their parallelism to the building edges, only similar phase values are present in each mask, allowing to preserve the fringe pattern of the layover. For each

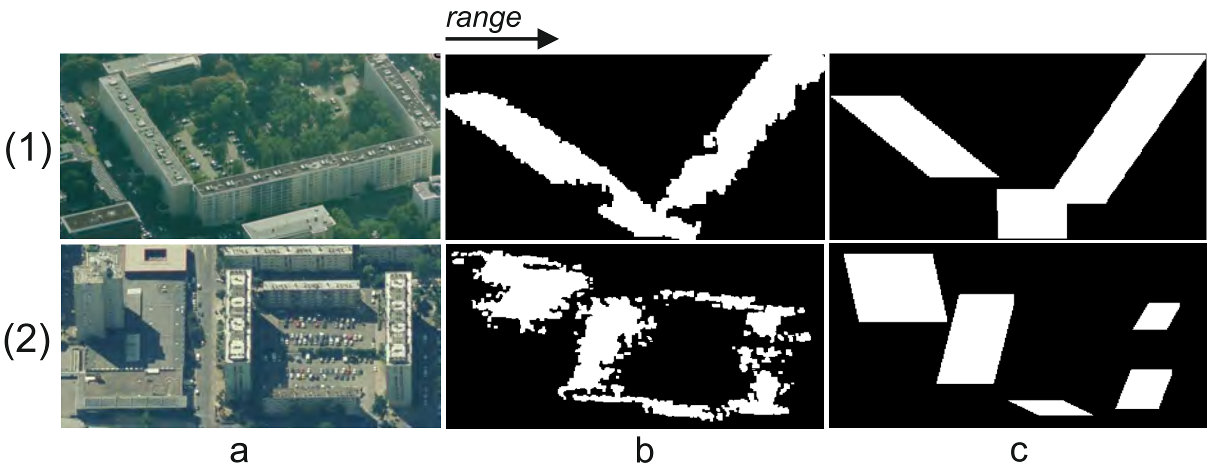


Figure 3.18: Special cases due to building blocks and close neighborhood: (a) optical image (source: Bing Maps); (b) extracted connected component after first processing; (c) extracted parallelograms after refinement

pixel position, the phase values, weighted by their coherence, are averaged. Figure 3.17(b, d) shows the consequently filtered parallelograms and mean phase profiles.
The subsequent building parameter estimation goes straightforward, using the parallelogram dimensions for determining orientation α, length l and width w, and the filtered phase values for determining the height h. Results are presented in Section 6.2.3.

3.6.3 Special Cases

Figure 3.18 shows two connected components representing more than one building, as can be observed in the corresponding optical imagery. However, they are considered by this approach as one single building, which leads to wrong parallelogram estimation and parameter extraction. In order to avoid such problems, an iterative approach is developed, whose goal is to separate these rather big connected components into smaller ones, containing each only one building. In order to recognize such cases, the length h_{BB} of each bounding box is first analyzed. Given a maximal possible building height h_{max} in the considered area, the expected maximal layover length l_{max} can be calculated in slant range by:

$$l_{max} = \frac{h_{max} \cos\theta}{dr} \tag{3.20}$$

In this work, $3 \cdot l_{max}$ is considered as maximal allowed h_{BB}, in order to take into account building orientation and length. Bounding boxes whose length is greater than this value are treated as follows: the segment is iteratively eroded, using a square structure element with a size of 2x2 pixels. For each iteration, the number of directly connected components is tested, and the steps from the facade recognition up to the parallelogram fitting are performed for each detected connected component. The residuals after parallelogram fitting are evaluated. This erosion process is performed until no segment is left. The iteration for which the residuals after parallelogram fitting are minimal is retained, whereby the extracted parallelograms are considered as the effective facades, from which building parameters are derived. The result of this process is shown in Figure 3.18c, whereby it is obvious that dense segments deliver better results. Segments that already present missing information tend to be oversplit. This will be further discussed in Section 6.2.3.

3.7 Conclusion

In this chapter, the development of the InSAR processing chain has been described, from the sole phase image to the extraction of the building parameters of single facades. After the detection of the characteristic building layover in the phase image, layover segments are analyzed in order to extract single facades, on which parallelograms are fitted. From the extracted parallelograms, building parameters are retrieved. Results about the completeness and accuracy of the extraction are given in Section 6.2.3. The building orientations α, their lengths l, and heights h are relevant for the following change detection, and therefore used as input for the radargrammetric processing, as constraint parameters. Namely, if no change occurs, the building dimensions estimated by the radargrammetric processing should be the same as the ones after the InSAR processing. The following chapter focuses on the radargrammetric processing, and in Section 5, the fusion of both approaches for change detection is enlightened. Numerical results of this chapter are shown with more examples in Section 6.2.

4

Radargrammetric Building Detection and Extraction

In this chapter, the radargrammetric approach for building detection and parameter extraction is presented. It is the equivalent to the previous chapter, considering radargrammetry instead of InSAR. As taken from the overall workflow for change detection presented in Section 1.4, InSAR is employed for pre-event and radargrammetry for post-event analysis. Results obtained during previous InSAR processing are used as input for the radargrammetric processing. Considering a specific building, its orientation but also its expected height and facade length are known from the InSAR processing. These parameters enable the definition of some processing parameters during the radargrammetric processing and facilitate the building detection. Changes in the building appearance are also detected more easily. This is presented in this chapter. First, the appearance of buildings considering different incidence angles is recalled as well as the implications in terms of parallax and matching (Section 4.1). Second, existing radargrammetric approaches for the analysis of urban areas are shortly recalled (Section 4.3). Subsequently, the approach developed in this work is presented (Section 4.3), and its different steps are explained in more detail (Sections 4.4 to 4.6). The fusion step between InSAR and radargrammetry is developed in Chapter 5, where the georeferencing of both methods for finding building correspondences, but also fusion and change detection rules are explained in more detail.

4.1 Appearance of Buildings in Radargrammetric Data

The influence of the radargrammetric acquisition configuration on the observed parallax at building location has already been presented in Section 2.4.3, for different incidence and intersection angles, as well as different acquisition geometries. In Section 3.1, the overall appearance of buildings in SAR images has been recalled briefly, insisting particularly on the layover that is observable for the buildings considered in this work. Figure 4.1 depicts different SAR acquisitions, representing the scene shown in the optical image. The layover of each building is recognizable, even if its appearance depends on the considered acquisition. Whereas double-bounce lines are visible for almost all buildings, roof areas are entirely contained in the layover for these acquisitions. Shadow areas are hardly distinguishable from the surroundings for most buildings,

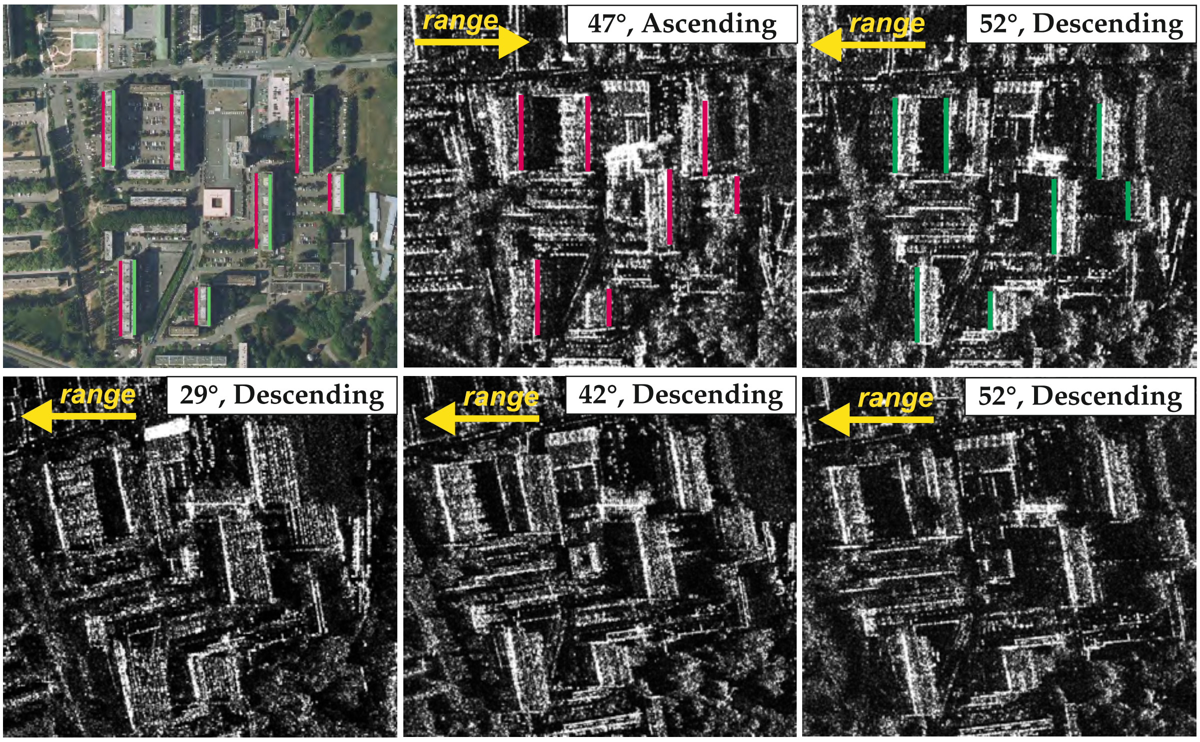


Figure 4.1: Appearance of buildings in radargrammetric data for different acquisition configurations; first row: geocoded images; second row: slant range geometry (source of the optical image: Bing Maps)

especially because of layover or multiple reflections occurring on other objects, mapped within these areas.

On the first row of Figure 4.1, opposite-side images are represented, geocoded for better interpretation with the optical image. Even if symmetrically opposed, the radiometry of both images is very similar, as they have almost the same incidence angles. However, inverting bright and dark zones for radargrammetric processing as proposed in (Fullerton et al. 1986) could prove difficult, as shadow areas are barely identifiable. Besides, matching double-bounce lines and layover areas between both images could even yield wrong results. Indeed, for a specific building, the double-bounce line of one image does not represent the same building foot as the one of the other image. Both double-bounce lines are represented in green and magenta on the optical and on both opposite images (Figure 4.1). It is obvious that they are situated on opposite building sides, whereby only one is visible in each SAR image, the other one being in the building shadow. The same phenomenon occurs for layover areas, whereby facades situated at opposite building sides are mapped in each image. If the considered building shows perfect symmetry of both facades, but also on the roof and on the ground on each side, the matching of both layovers is theoretically possible. The real parallax could then be retrieved, assuming the building width is known. However, this is often not the case, as objects situated on the ground on both building sides vary and facades may be different, due to balconies or other structures. The use of opposite-side configuration at building location is therefore not possible, as points that seem radiometrically similar do not correspond to the same points in reality.

The second row of Figure 4.1 presents three SAR acquisitions taken from the same-side, but with different incidence angles. Layover and double-bounce lines are well recognizable in all images, although the layover appearance varies a lot between all images. Especially, considering the steep incidence angle of 29°, the layover is characterized by regular bright lines corresponding to the window front on the different floors. The two other images present shorter layover areas, whereby the line pattern is replaced by more homogeneous bright areas. The strong texture tends to melt together by shallower incidence angles, as more information is mapped in a shorter area. Even if the combination of the steep incidence angle with a shallower incidence angle produce better geometry, due to the larger intersection angle, it is obvious that homologous points are more difficult to find, as the radiometry is completely different. Besides, more overlapping with surrounding objects occurs within the layover area of the steep incidence angle (29°), making matching even more complex. In extreme cases of very large intersection angles or other building shapes, it may even be possible that the roof area appears only in one of both images, creating an area with new radiometry that is not present in the other image. On the contrary, the two images taken from shallower incidence angles (42° and 52°, Figure 4.1), i.e. forming a smaller intersection angle, have very similar radiometry, facilitating the search for homologous points. In this latter case, the geometry for 3D reconstruction is less optimal.

As a conclusion, as for InSAR processing, shadow areas are useless for radargrammetric processing in urban areas. Not only are they barely identifiable, undergoing some overlapping with surrounding objects, but also using their boundaries could yield wrong matches, as the points situated at a shadow boundary in one image do not correspond to the points situated at the shadow boundary of the other image. This assertion is true for both opposite- and same-side configurations. The consideration of double-bounce lines and layover areas is preferable, whereby it has to be handled carefully, depending on the acquisition configuration. Considering opposite-side configurations in urban areas may turn out dangerous, as points with similar radiometry in both images do not correspond to the same points in reality. On the contrary, the search for homologous points is facilitated with same-side configurations, whereby some radiometric differences arising due to incidence and intersection angles may hinder the matching. Finally, matching of layover areas stays a difficult task, as each pixel contains two or three different contributions from ground, facade, and potentially roof. For each pixel in the master image, at least two homologous points would have to be found in the slave image. However, as the facade contribution prevails (Rossi et al. 2014a), this work focuses on finding homologous facade points. This will be further enlightened in Section 4.5.

Different radargrammetric same-side configurations are considered in this work. Resulting extracted building parameters allow their comparison and the estimation of preferable acquisition parameters, as shown in Section 6.3.3.

4.2 State-of-the-Art

Depending on the chosen matching method (feature/area-based), on the method of improvement (e.g., backmatching, pyramid), on the chosen matching criterion, on the acquisition configuration (same-side, opposite-side, orthogonal flight paths), or on the method chosen for determining the height (absolute/relative), the state-of-the-art in radargrammetry offers different possibilities for building recognition and reconstruction. In order to keep a sorted overview, the existing approaches are presented with respect to the matching method.

Several approaches exist that employ feature-based matching. In (Michaelsen et al. 2008, Sörgel et al. 2009), salient lines and points belonging to buildings are detected in images acquired on orthogonal flight paths. A production system is then used for merging them into 2D-structures, producing building corners with specific angle and symmetry. Within the defined production system, line and angle correspondences are found between the orthogonal acquisitions, and 3D information is retrieved. This approach provides promising results, particularly due to the orthogonal flight paths, which enable to find corner information, at least for one corner. With such information, the building can be better modeled. With spaceborne data, only ascending and descending path are possible, representing either only one facade or opposite building sides, which are more complicated to associate. In (Goel & Adam 2012), point scatterers (PS) situated on building facades in same-side amplitude images are detected and their disparity and corresponding absolute height is assessed using Bayesian networks. To this goal, three configurations using different incidence angles are used. First, PS are extracted for each configuration using a stack of images taken under the same incidence angle. Second, the PS positions extracted from the different configurations are scaled with respect to the chosen master image. Their relative shift between the different configurations is assessed using a Bayesian network, which permits to find PS correspondences by evaluating the shift direction and derived height. This approach shows good results in terms of PS location and absolute height, which is of great advantage if combined with PS interferometry, where only relative height can be determined. Yet, in order to extract reliable PS, a large image stack has to be available for each incidence angle, which makes its utilization for rapid demand useless. Feature-based approaches are bounded to the extraction of specific points or lines, and do not consider all the information contained in the building signature of the intensity image.

On the contrary, area-based approaches consider the full image information in order to extract disparities. In (Dai et al. 2008), the authors extract buildings using polarimetric classification in opposite-side images, and determine building height using disparities calculated within pixel location profiles of the classified images. Each image is classified first, and the pixels belonging to one building are summed up in their respective range directions. The disparities calculated between both resulting profiles provide the building height. The classification allows to process information of the whole buildings without suffering from the influence of their surroundings. Nevertheless, the opposite-side configuration implies that the same part of the building (or at least building roof) are visible on both images, which is generally not the case. In (Oriot & Cantalloube 2008), several acquisitions taken under the same incidence angle, but under a circular trajectory (i.e. different azimuth angles), are considered. Normalized cross-correlation calculation between overlapping images, followed by geocoding of all resulting height maps, permits to retrieve the DEM. In order to reduce false detection, a threshold is set on the cross correlation coefficient. This approach shows good results, however, the layover of neighbor images are quite similar, facilitating the point matching and disparity calculation.Recently in (Capaldo et al. 2012), the authors analyzed the quality of an urban DSM created by radargrammetry, performing coarse matching first at a 3D scale, and fine matching at a 2D scale, on the correct height level. The results show yet discrepancies to the correct building height.

An hybrid approach is presented in (Simonetto 2002, Simonetto et al. 2005), whereby building features are extracted first, and the height deduced from an area-based approach is allocated to the extracted features in a second step. In this approach, a classification is first performed in order to extract bright lines in the magnitude images, that are then merged to lines extracted using a ratio based method. A binary mask is created in which crossing points presenting at least two bright arms are extracted for further processing. Bright segments in the prolongation of the cross arms are connected, in order to reconstruct the building borders. False alarms are eliminated by retaining only the crosses and segments present in both stereoscopic images. Such segments may represent either the building footprint or the top of the roof. Area-based matching is performed in order to determine the ground height, leaving a buffer around the building footprint, so that the building height and layover do not influence the calculation. A pyramidal approach with backmatching is employed. The final building height is estimated by using either stereoscopic or monoscopic measurement on the extracted segments. Only connected segments showing a height difference to the ground height are considered belonging to the roof and used for the final height computation. This approach represents the interest of combining the wide area height information of area-based matching and structure information of feature extraction. However, it relies on the principle that roof edges appear distinctly as brighter lines in the SAR image. If not, the whole layover is considered as a thick segment and height information has to be retrieved using monoscopic measurement of the layover thickness.

It has been shown in Section 4.1 that buildings can take very different appearances considering different incidence angles. Especially, the extraction of linear features in the layover (i.e. window lines) could be possible in one image, but not in the other, leading to a poor matching between images and a very sparse reconstruction. The choice of area based matching for this work is straightforward, as wrong matches can be compensated by the quantity of information. A matching criterion taking into account the image statistics is considered, in order to improve the matching localization of geometric patterns, i.e. to take into account the feature information. Furthermore, a pyramidal approach combined with a filtering is used in order to compensate different building appearance caused by different incidence angles. Finally, backmatching is performed in order to enhance the matching reliability.

4.3 Overall Radargrammetric Workflow

Figure 4.2 depicts the overall workflow of building detection and extraction by radargrammetry. It consists in three main steps: pre-processing, matching, and building extraction. The first step is performed at image level, and the two last steps only at patch and facade level, using input from the previous InSAR processing.

During pre-processing, the images are calibrated, in order to minimize object independent radiometric differences between both images. The calibrated images are then coregistered. In this work, the slave image is transformed into the slant-range geometry of the master image. A SAR-SIFT like algorithm (Dellinger et al. 2015) is used for finding corresponding points in both images, followed by a projective transformation and resampling of the slave image. This step is described in Section 4.4.

After coregistration, both images are in epipolar geometry, and matching, i.e. search of homologous points, is performed. The developed matching approach is presented in Section 4.5, taking into account geometrical as well as radiometrical aspects. Particular attention is paid to layover

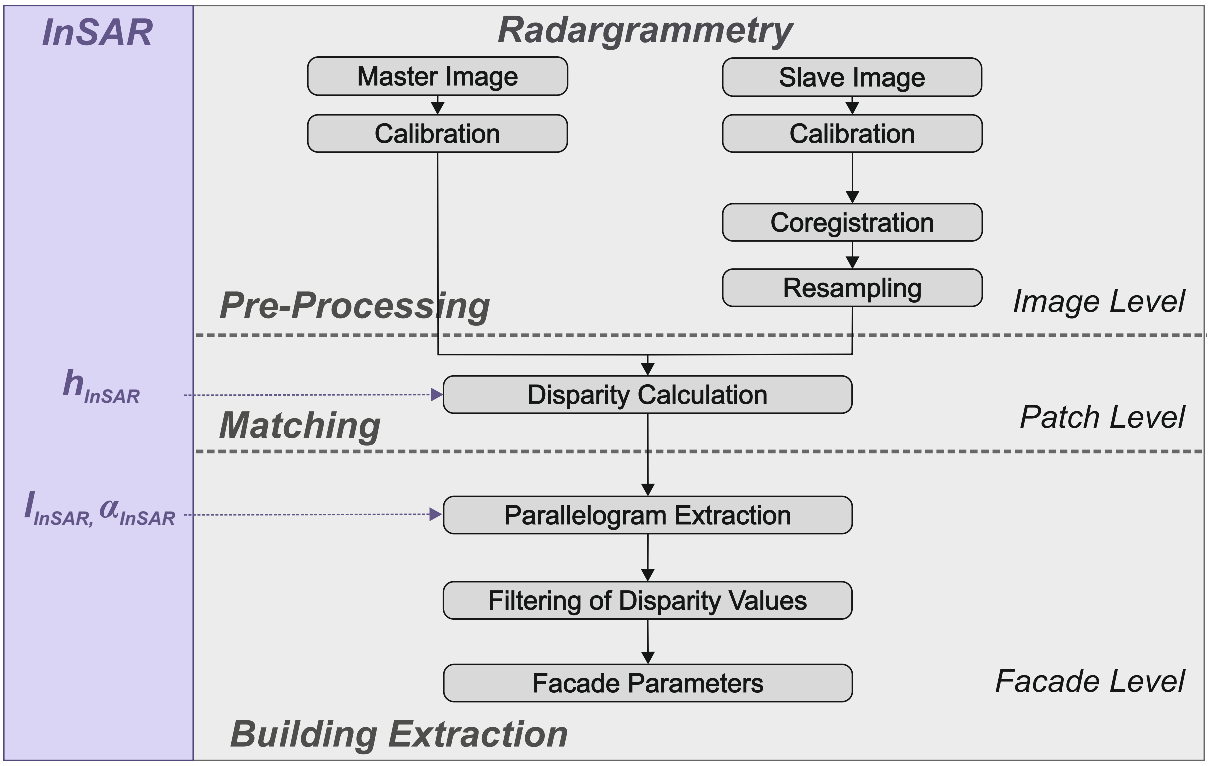


Figure 4.2: Overall radargrammetric workflow

areas, in order to define new constrains for reducing the search area for matching. Considerations on the acquisition configuration allow to restrain the search area in azimuth direction whereas considerations on building heights obtained during InSAR processing constrain the search in range direction.

The disparity map resulting from the matching step is analyzed for each facade separately, using the results obtained during InSAR processing. The appearance of rectangular shaped buildings in the disparity map is analyzed, revealing for each facade two parallelograms, whose extraction is performed using both radargrammetric processing parameters and known building information from the InSAR processing. In order to determine possible changes in the building appearance and deduce post-event building parameters, a further analysis of the parallelogram shapes and disparity values is performed. Relying on these new results, parallelogram dimensions are adapted, leading to the post-event facade length. Within the final parallelograms, a filtering of the disparity values is performed. As disparities are directly related to height values, smoothing of noisy disparities permits to improve the height determination. A post-event relative height estimate is finally determined for each facade by converting the filtered disparity values using geometric considerations. This step is explained exhaustively in Section 4.6.

4.4 Radargrammetric Pre-Processing

Pre-processing aims at bringing both radargrammetric acquisitions into the same radiometric range and geometry for facilitating the later matching. Both pre-processing steps are explained in the following.

4.4.1 Calibration

Changes of incidence angles or of orbit orientation produce radiometric differences between images that should be minimized for better comparison. The aim of image calibration is to reduce and eliminate in each image all radiometric contributions that are not directly due to the target characteristics.

Two different calibration coefficients can be computed: beta-naught β^0 and sigma-naught σ^0. Beta-naught corresponds to the RADAR brightness and is expressed as:

$$\beta^0 = k_s \cdot |A|^2, \text{ whereby } A = \sqrt{I_p^2 + Q_p^2} \tag{4.1}$$

A is the amplitude of the signal at pixel position p. The calibration factor k_s is given in the product files, and is sensor and product dependent (Fritz & Werninghaus 2007). The RADAR brightness β^0 only takes into account radiometric differences due to changes of the sensor or of the mode of acquisition (e.g., StripMap, Spotlight). It does not consider the incidence angle, which is mandatory for radargammetric analysis. The second coefficient sigma-naught σ^0 takes into account the orientation and distance of each resolution cell towards the sensor. The influence of the local incidence angle θ_{loc} is therefore considered:

$$\sigma^0 = (k_s \cdot |A|^2 - NEBN) \cdot \sin\theta_{loc} \tag{4.2}$$

$NEBN$ is the Noise Equivalent Beta Naught, and represents the influence of several noise contributions to the signal. It is a more precise estimation of β^0, computed as a summation of polynomials of different degrees, weighted by the calibration factor k_s. The polynomials' coefficients represent the influence of different noise factors such as transmitted power and antenna noise, and are given in the product file.

In this work, the second coefficient sigma-naught σ^0 is used, in order to obtain images of similar radiometry, independently of the incidence angle of the acquisition. Master and slave images are both calibrated, producing σ_m^0 and σ_s^0, respectively. Those resulting images are used for further processing.

4.4.2 Coregistration

Both calibrated images σ_m^0 and σ_s^0 are then coregistered. In this work, the slave image is coregistered on the master image, so that points situated on the ground are at the same position. This ensures that the double-bounce lines of the buildings have same position in both images after coregistration.

Coregistration may be performed in three ways. The first consists in georefecing both images, using a specific reference system. This usually leads to distortion of image features, as the reference system may not be a plane. In urban areas, double-bounce reflections at building location are characterized by lines and layover areas have parallelogram shapes. Distortion of such linear features complicates feature recognition and extraction for later building parameter estimation. Besides, relief implied image distortion would hinder the matching. The second possibility is to project both images into ground geometry, using a reference surface. An ellipsoid would yield the same drawbacks as georeferencing, but a common plane surface allows the preservation of the image geometric shapes. Yet, both images still undergo a transformation, implying interpolation errors and changes in their radiometry. The third possibility for image coregistration implies the slave image to be reprojected into the slant geometry of the master image. Here, only one image is interpolated and modified. Besides, the projection surface is still a plane, preserving the shape of geometric features. This last option is chosen in this work, as conservation of geometric features is mandatory for disparity calculation and later fusion.

Coregistration of radargrammetric SAR images often combines the different presented ways. The goal is to preserve the geometric and radiometric information of both images by allowing a facilitated matching, respecting the epipolar constraint (Méric et al. 2009). In (Simonetto 2002), an approach similar to optical stereoscopy is used, whereby points of the master image are projected to several ground heights and then projected into the slave image, creating epipolar curves in the slave image, and vice-versa. Homologous points are searched along the direction of the epipolar lines. This approach applies multiple transformations, as each image point is projected into several heights. Similar to this approach, (Nascetti 2013) proposes to project both master and slave images into ground planes of several heights. Each image provides an image stack, or voxel grid, and matches are first searched in the height direction between both stacks. This approach, as the previous one, combines the two steps of image coregistration and matching. In (Perko et al. 2011), a regular grid of points is first defined in the master image and projected onto a coarse DSM. This created DSM grid is then projected into the slave image geometry, permitting the definition of an affine transformation between master and slave image, in a least-squares sense. The slave image is then registered to the master image using this affine transformation. As a result, the direction of the disparity is aligned with one image dimension. Considering parallel flight tracks, across-track geometry of both acquisitions and constant altitude of the sensor, as for spaceborne imagery, the search for matches can then be reduced to the range direction. This quasi epipolar geometry is also used by (Méric et al. 2009), whereby the search in azimuth direction is extended to a small stripe (3 pixels) in order to balance the estimation error of the coregistration. All those approaches necessitate a coarse DSM, or an approximate terrain height of the area of interest. Besides, they suffer from interpolation errors, as at least two projections have to be performed.

An original approach is proposed in (Schubert 2004), where the slave image is simply resampled into the geometry of the master image, by eliminating regularly complete columns of the slave in range direction. Yet, it involves information loss.

In this work, a new approach is used, not necessitating any external information and requiring only one image reprojection. It is based on the SAR-SIFT algorithm presented in (Dellinger et al. 2015). The main change compared to the original SIFT algorithm (Lowe 1999) is the introduction of a gradient computation called gradient by ratio (GR). It is based on the logarithm of the ratio of exponentially weighted averages (ROEWA), defined in (Fjortoft et al. 1998). During key-point detection, a multiscale representation of the original image is used. Contrary to the original SIFT scale space, which is defined using Laplacian of Gaussian (LoG), the SAR-SIFT uses a multiscale SAR-Harris function based on the GR. The scale parameter is replaced by the smoothing parameter of the GR calculation. It considers adaptive smoothing of the image without changing its scale. Namely, both SAR images have the same scale. A comparison with key-points extracted with the original SIFT detector shows that the extracted SAR key-points

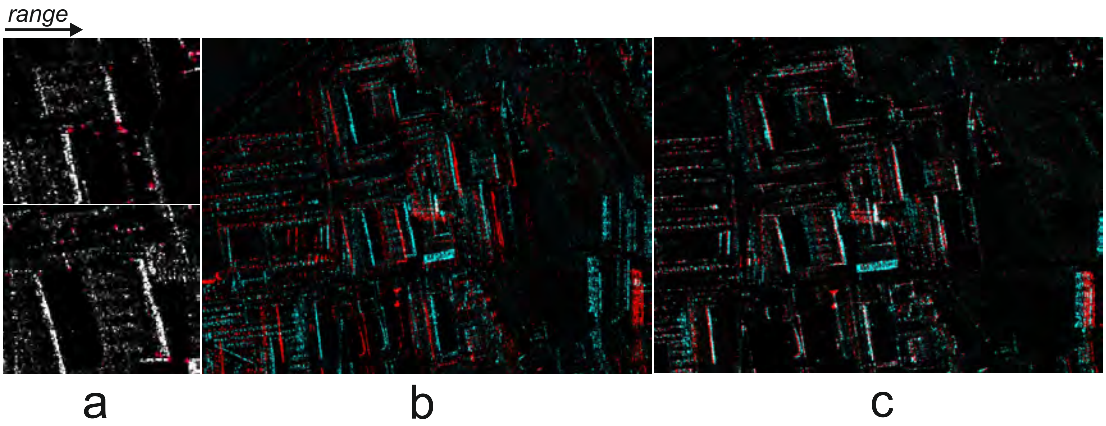


Figure 4.3: SAR-SIFT: (a) extracted key-points using gradient by ratio; (b) master (red) and slave (blue) images before SAR-SIFT coregistration; (c) master (red) and slave (blue) images after SAR-SIFT coregistration

are situated on distinct features, as corners and edges, and that their extraction suffers less from inherent multiplicative speckle. Figure 4.3a shows key-points extracted with the new method. During orientation assignment and descriptor extraction, the GR is used again for computing the histograms of gradient orientation. In (Dellinger et al. 2015), a circular neighborhood, separated in polar sectors, is used for computing the histograms of gradient orientations. In this work, a square neighborhood as in the original SIFT is used. Indeed, in urban areas, key-points are situated on double-bounce lines and linear edges. Square windows allow to preserve the rectilinear geometry of urban structures. Key-points of both master and slave images are then matched by computing nearest neighbor analysis on both sets of descriptors. For coregistration of master and slave images, a transformation has to be fitted to the corresponding points of both images. In (Dellinger et al. 2015), an affine transformation is defined. However, both images are in their respective slant geometry. Considering spaceborne acquisitions on parallel flight-tracks in across-track geometry, this is a plane-to-plane transformation, corresponding to a homography. Thus, in this work, instead of an affine transformation, a homography is defined between the two sets of points. This transformation is defined by eight parameters, requiring at least four pairs of corresponding points between both images. In order to robustly estimate the transformation parameters, outlier correspondences are filtered by employing a RANSAC (RANdom SAmple Consensus (Bolles & Fischler 1981)) scheme. Figure 4.3(b,c) shows in false-color representation both images before and after SAR-SIFT coregistration. It is observable that the double-bounce lines of all buildings are well aligned after coregistration. Parameter settings of this approach, as well as subsequent accuracy tests are presented in Section 6.3.1. A comparison to a standard coregistration method is shown.

Using SAR-SIFT allows to directly project the slave image into the slant geometry of the master image, without using external information. Compared to the methods presented previously, only one transformation occurs, reducing interpolation errors, and only one image -the slave image- is modified. During transformation, resampling ensures that the coregistered slave image has the same sampling as the master image. Calibrated amplitude values of the coregistered slave image are determined using bilinear interpolation. Furthermore, the final coregistered images as represented in Figure 4.3c are in epipolar geometry, whereby epipolar lines correspond to the range direction. The search for homologous points for disparity calculation can thus be performed along horizontal direction for each pixel. Yet, due to slight differences of heading angles between both SAR acquisitions, a slightly wider search window has to be defined. This is explained in Section 4.5.3, considering layover areas.

Finally, SAR-SIFT is a very attractive method for the coregistration of two images taken from different incidence angles. It allows to find enough matches between both images for defining a projective transform. Having a closer look at the extracted key-points, it is obvious that they are mainly situated on the double-bounce lines, corresponding to ground level, where the disparity between both coregistrated images is zero. Modifying the parameters in order to obtain key-points corresponding to other building parts (e.g. layover) is possible. This would allow to calculate disparities between both images. However, only sparse disparity calculation would be allowed with such an approach. Besides, the difference of incidence angles induces very different appearance of the layover areas, as mentioned in Section 4.1. Finding corresponding key-points in layover areas may prove nearly impossible. Hence in this work, a different approach for disparity calculation is used, and described in Section 4.5.

4.5 Matching

Matching is the most important step of the radargrammetric processing cahin, as it allows to determine the disparity, or parallax, between both acquisitions. From this disparity, the object height can be deduced. Section 2.4.2 detailed two main aspects of matching: the matching method itself and the method for improving the matching. Whereas the matching method consists in defining a similarity function for finding radiometrically homologous points in both images, the improved method aims at ameliorating geometrically the similarity calculation in order to make the determined disparity values more reliable. The strategies used in this work for both aspects are explained in more details in the following.

4.5.1 Matching Method

As explained in Section 4.2, area based matching is performed in this work. The usual similarity criterion is the normalized cross correlation ρ, making an equivalent to the coherence, defined for the InSAR processing in Equation (2.18):

$$\rho = \frac{E(X_m X_s) - \mu_m \mu_s}{\sigma_m \sigma_s} \tag{4.3}$$

X_m and X_s are two variables, containing the amplitude values of the pixels within the specified areas, or windows, of the master and slave image, respectively. E is the expectation, μ_i and σ_i are respectively mean and standard deviation of the values within the considered windows. In the following, the defined area in the master image is called template window, and the scanned area for similarity in the slave image is called search window. The template window is always smaller than the search window.

Other similarity criteria such as sum of normalized absolute and squared differences (Leberl et al. 1994), wavelet decomposition (Schubert et al. 2013), or coefficient of variation (Tupin & Nicolas 2002) were investigated for SAR images.

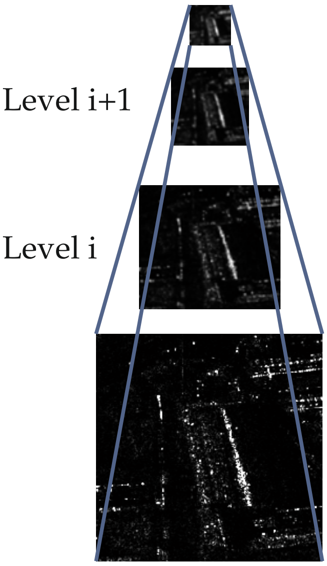


Figure 4.4: Example of an image pyramid

In this work, the criterion υ based on the coefficient of variation is applied (Tupin & Nicolas 2002). It is preferred for the matching of structured areas. It considers radiometric ratios instead of direct radiometric differences between both images:

$$\upsilon = \frac{c_{\upsilon_m} c_{\upsilon_s}}{E[(\frac{X_m}{\mu_s} - \frac{X_m}{\mu_s})^2]} \tag{4.4}$$

c_{υ_i} is the coefficient of variation, defined as $c_{\upsilon_i} = \frac{\sigma_i}{\mu_i}$. This criterion is more robust to structure changes, as edges, and is preferable for use in urban areas. Indeed, it allows to cope with remaining radiometric changes between both calibrated images, among others due to speckle. Areas of similar ratio are searched for matching, privileging areas containing strong features, and reducing the impact of noise. In urban areas, particularly the matching of double-bounce lines and layover edges is interesting in order to determine disparities at building location and deduce building height. Such areas present high ratio values, compared to areas only affected by noise, and are matched more easily with this criterion.

4.5.2 Improved Matching

In this work, a specific method for improving the matching is developed. Similar to (Fayard et al. 2007), an image pyramid is built (Figure 4.4), yet using an edge preserving method. The images are first matched at low resolution and the output is used for matching at higher resolution. Several methods exist for creating an image pyramid. The most commonly used method consists in using Gaussian filtering from one level to the next one, whereby the image size is reduced at each level by a factor of two. Considering G_i as the original image, the down-sampled image at the next level G_{i+1} is expressed as:

$$G_{i+1} = downsample(G_i) \tag{4.5}$$

downsample is a down-sampling operation, performed by convolution of a low-pass filter with the image G_i. The filter mask is a kernel k_{Gr} defined by:

$$k_{Gr} = \begin{pmatrix} \frac{1}{4} - \frac{a}{2} & \frac{1}{4} & a & \frac{1}{4} & \frac{1}{4} - \frac{a}{2} \end{pmatrix} \tag{4.6}$$

It defines a 5x5 pattern of weights. The weighting parameter a usually takes values between 0.3 and 0.6 (Burt & Adelson 1983). In order to obtain a symmetric Gaussian kernel k_G, the parameter a is set to 0.375. The final Gaussian kernel k_G is the convolution of the horizontal component k_{Gr} with the vertical component k_{Gc}:

$$k_G = k_{Gr} * k_{Gc} = k_{Gr} * k_{Gr}^T \tag{4.7}$$

During convolution, the center pixels obtain higher weights than neighbor pixels. Using a Gaussian kernel has however the drawback of smoothing edges, which complicates the matching. In (Fayard et al. 2007), a similar method is used whereby, instead of a Gaussian kernel, the grey values of down-sampled images are defined by averaging the grey values of four pixels in the previous image. However, as the image is not prefiltered, speckle has a high impact on the average grey value of the new level, and matching could prove difficult, as speckle is not the same in both images.

In this work, Laplacian pyramids, i.e. difference of Gaussians (DoG), were used instead of a Gaussian pyramid. Few works already use Laplacian pyramids for speckle reduction in SAR images (Jin et al. 2005). One level of the Laplacian pyramid corresponds to the difference between two levels of the Gaussian pyramid. Figure 4.5 shows this principle. The image L_i at a specified pyramid level is expressed as:

$$L_i = G_i - upsample(G_{i+1}) \tag{4.8}$$

upsample is an upsampling operation that expands G_{i+1} to the size of G_i using a specified filter kernel (Paris et al. 2011). As Gaussian pyramids perform a low-pass filtering, the difference between both levels represents the higher frequencies, which characterize the details of the images, i.e. noise and edges. In order to eliminate speckle, Refined Lee speckle filtering (Lee 1981) is used on the original image P_1. As explained in Section 2.2.3, this speckle filtering method allows to reduce speckle without smoothing edge information, by taking into account image statistics. Hence, the resulting levels of the Laplacian pyramids show edges, without being affected by speckle. Finally, the Gaussian filter kernel defined in Equation (4.6) is replaced in this work by a 9/7 wavelet filter (Daubechies et al. 1992), described in (Khazaei & Georgiadis n.d.) as more efficient for higher frequencies. The final image P_i at each pyramid level, containing high frequency edge information as well as lower frequency information, is reconstructed using:

$$P_i = L_i + upsample(P_{i+1}), \text{ whereby } L_i = P_i - upsample(P_{i+1}) \tag{4.9}$$

In this work, four pyramid levels are defined. Compared to the original Gaussian pyramid G_i represented in Figure 4.5, it is obvious that edges are better preserved in the highest pyramid levels (smaller image size) of the reconstructed Laplacian pyramids P_i (Figure 4.5).

Several indexes allow to give numerical validation of this conclusion. They are presented in (Sheng & Xia 1996), and (Iqbal et al. 2013), and originally served for quality evaluation of speckle filters, and their availability to preserve features. The indexes used in this work for evaluating the edge preserving capability of the image pyramid are resumed in the following:

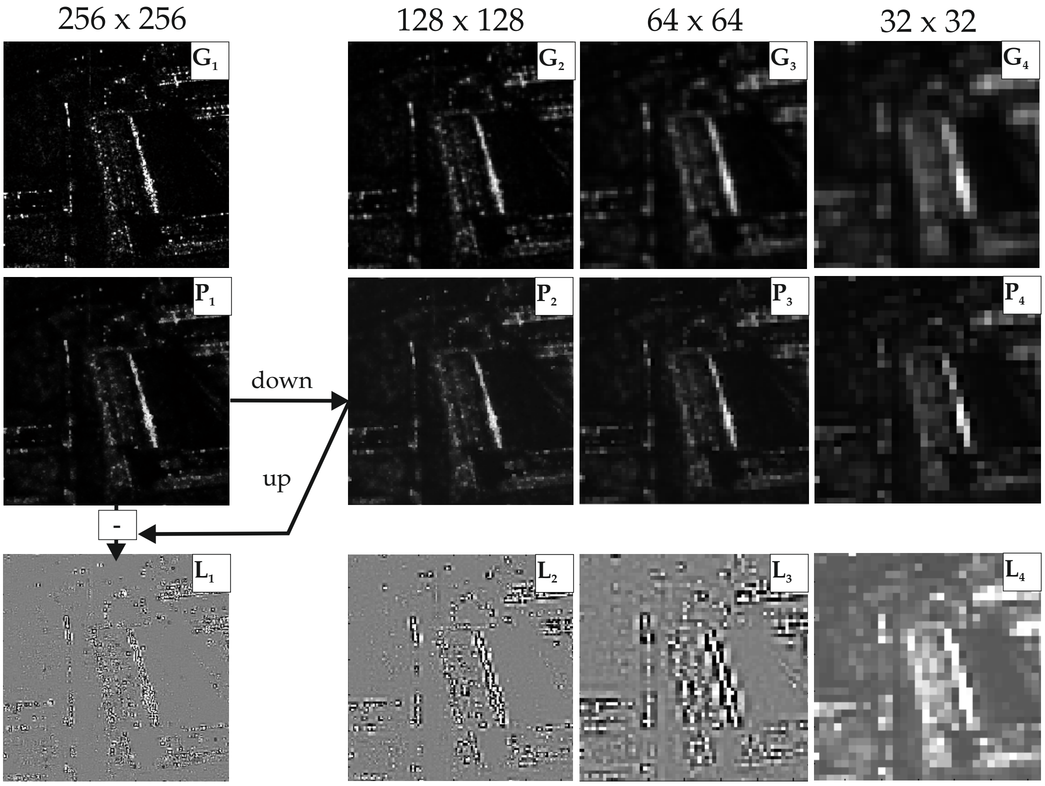


Figure 4.5: Gaussian (G_i), Laplacian (L_i) and reconstructed Laplacian (P_i) image pyramids

- Equivalent Number of Look (ENL): this index is the square of the ratio of the mean image value above its standard deviation and represents the ability of the method of flattening, or smoothing the image.

  $$ENL = \left(\frac{\mu_{P_i}}{\sigma_{P_i}}\right)^2 \tag{4.10}$$

  Sometimes, the Flatness Index (FI), $FI = \sqrt{ENL}$ is used instead of ENL. The larger the index, the higher the image smoothing.

- Speckle Suppression Index (SSI): is the ratio of the coefficient of variation of filtered and original image. In this work, as there is no original image except for the first pyramid level, the result G_i of the Gaussian pyramid is considered as the original image for each level, and the image P_i provided by the Laplacian Pyramid on the speckle filtered image is considered as the filtered image.

  $$SSI = \frac{c_{v_{P_i}}}{c_{v_{G_i}}} = \frac{\sigma_{P_i}}{\mu_{P_i}} \frac{\mu_{G_i}}{\sigma_{G_i}} \tag{4.11}$$

  Smaller values of SSI mean that P_i shows better speckle suppression as G_i.

- Speckle suppression and Mean Preservation Index ($SMPI$): this index allows to take into account a potential overestimation of the image mean induced by filtering, which is not taken into account by ENL and SSI.

  $$SMPI = Q \cdot \frac{\sigma_{P_i}}{\sigma_{G_i}}, \text{ whereby } Q = 1 + |\mu_{G_i} - \mu_{P_i}| \tag{4.12}$$

  Here again, lower values of $SMPI$ indicate better speckle suppression.

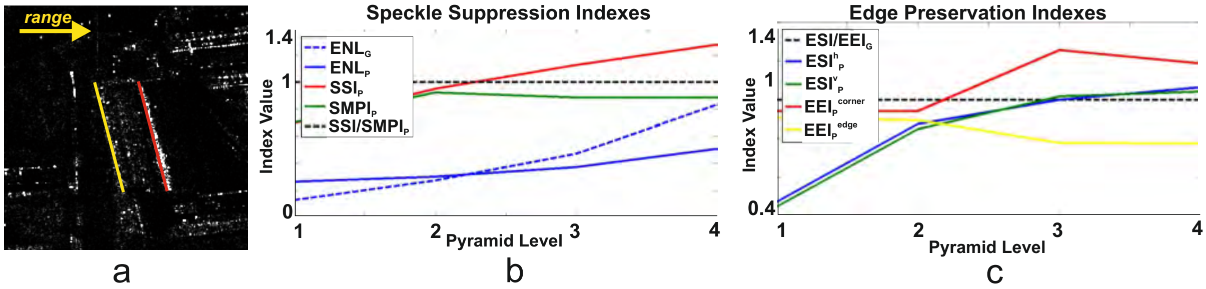


Figure 4.6: Comparison of speckle filtering and edge preserving indexes for Gaussian and reconstructed Laplacian image pyramid: (a) considered building with defined double-bounce (red) and layover border (yellow) lines; (b) speckle filtering indexes; (c) edge preserving indexes

The three previous indexes permit the evaluation of speckle suppression. However, when speckle is reduced, so are the image structures. In order to evaluate the edge preserving capability of the pyramid, complementary indexes have to be used. The better method is the one presenting a good compromise between high speckle suppression and good edge preservation. The used edge preserving indexes are presented in the following:

- Edge Saving Index (ESI): this index measures the capability of the method of preserving edges. It has a horizontal component ESI^h and a vertical component ESI^v, permitting to take into account several edge directions.

$$ESI^h = \frac{\sum_{k=1}^{m}\sum_{l=1}^{n-1}|P_i(k,l+1) - P_i(k,l)|}{\sum_{k=1}^{m}\sum_{l=1}^{n-1}|G_i(k,l+1) - G_i(k,l)|} \tag{4.13}$$

$$ESI^v = \frac{\sum_{k=1}^{m-1}\sum_{l=1}^{n}|P_i(k+1,l) - P_i(k,l)|}{\sum_{k=1}^{m-1}\sum_{l=1}^{n}|G_i(k+1,l) - G_i(k,l)|} \tag{4.14}$$

m and n are respectively the number of columns and rows in the image at pyramid level i. The higher ESI^h and ESI^v, the better the edge preservation ability. Usually, ESI^h and ESI^v are smaller than 1. However in this work, better edge preservation is expected for P_i as for G_i, which is smoother. Therefore, values of ESI^h and ESI^v greater than 1 prove better edge conservation for P_i than for G_i.

- Edge Enhancing Index (EEI): similar to the ESI, the EEI considers only the neighborhood of edges instead of the whole image. There, pixels of either side of an edge (L and R) are selected and substracted from each other, in both filtered and original images. The total number of selected pixels in each image is N.

$$EEI = \frac{\sum_N |P_{iL} - P_{iR}|}{\sum_N |G_{iL} - G_{iR}|} \tag{4.15}$$

In this work, a specified buffer zone has been defined at each pyramid level i for selecting the pixels. Besides, EEI was evaluated on a line (double-bounce, red line in Figure 4.6a) and on an edge (layover border in near range, yellow line in Figure 4.6a). Here also, the higher EEI, the better the edges are preserved.

Figure 4.6 presents the results of the different indexes for the several pyramid levels of both Gaussian (G_i) and reconstructed Laplacian pyramids (P_i), taking as example the image of the building presented in Figure 4.6a. Whereas ENL and SSI show a better result for the two higher levels of the Gaussian pyramid, $SMPI$ is better for the Laplacian pyramid, at all levels. Concerning the indexes showing the quality of the edge preservation, both ESI^h and ESI^v show better results for the Laplacian pyramid at the two highest levels. Having a look at Figure 4.5, edges are less smoothed than using the Gaussian pyramid. Similar observations can be made for EEI calculated for the double-bounce line. However, for the layover edge, EEI is higher for the Gaussian pyramid, at all levels. For matching, it is important that the start images, i.e. G_4 or P_4, at low resolution, have good quality, as their result is used as input for the next levels. A matching error at the highest level would propagate down to the lowest level, i.e. the full resolution image. Both speckle filtering and edge preserving indexes indicate that the use of Laplacian pyramid outperforms the use of Gaussian pyramid for the two highest levels, therefore the Laplacian pyramid is employed in this work. The lowest level G_1 of the Gaussian pyramid represents the original, unfiltered image. In comparison, the lowest level P_1 of the reconstructed Laplacian pyramid represents the speckle filtered image, which explains the poorer results of edge preservation indexes of the Laplacian at the lower levels.

Starting from the highest level (P_4 in Figure 4.5), disparities are calculated using the defined matching criterion. The disparity map of one level (P_{i+1}) is used as input for the disparity calculation of the next lower level (P_i). Therefore, the disparity map of the highest level is upsampled by a factor of 2. Before upsampling, backmatching is performed, i.e. master and slave images are switched. Only pixels presenting the same -opposite- disparity values are transmitted to the next level. At the last level, at full resolution, the disparity map is filtered in order to smooth the final disparity map. Indeed, the disparity values between neighbor pixels are expected to be homogeneous. Therefore, for each pixel in the created disparity map, the most occurring disparity value within the pixel's direct neighborhood (3x3pixels) is retained as its final disparity value.

4.5.3 Constraints

In order to restrict the search area for matching, and eliminate wrong disparities, several constraints are defined. Among the constraints presented in Section 2.4.2, one particular needs an adaptation for urban areas: the epipolar constraint. This constraint permits to define, case dependent, the dimension of the search area w_s and l_s, as shown in Figure 4.8.

After coregistration, both images are in epipolar geometry and the search for matches could be limited along a single horizontal line, i.e. along the range direction. This assumption holds, considering strictly parallel flight paths of both acquisitions. However, in reality, both acquisitions have slightly different heading angle, leading to a small rotation between both images, represented as ζ, in Figure 4.7b. In this work, a thorough analysis of the influence of this angle on the epipolar constraint in layover areas is conducted. Figure 4.7(a,b) shows a schematic representation of a building layover area after coregistration, in both images, in ground and slant range, respectively. The layover in the master image is represented in bright grey and the one in the slave image in blue. Figure 4.7b shows that both layovers are not perfectly aligned in horizontal direction after coregistration. A point in the layover area is the summation of most often

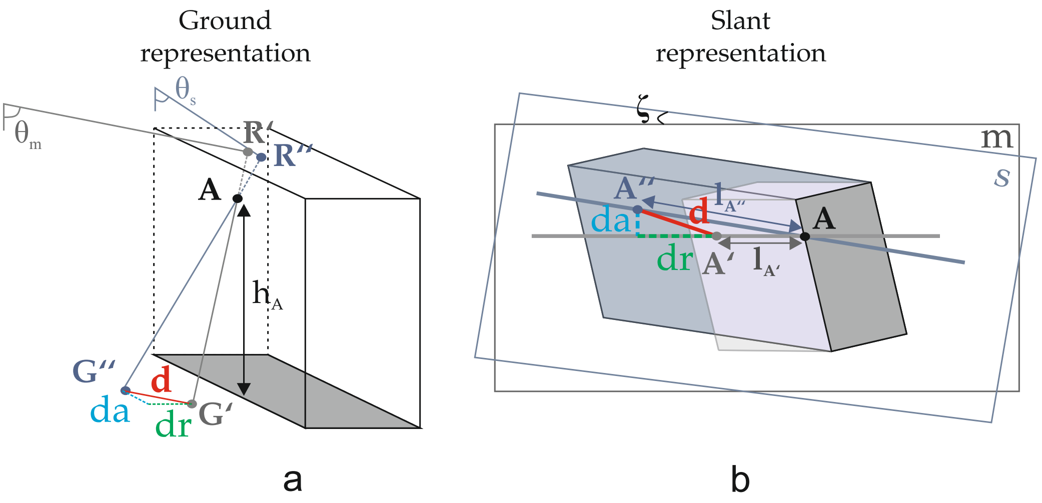


Figure 4.7: Epipolar constraint for building layover areas: (a) ground geometry; (b) slant range geometry of master image m

three contributions, which are aligned in range direction: considering the master image m, those points are represented as G' on the ground, A on the facade and R' on the roof (Figure 4.7a). In the slant range image, they are all represented at the same position A', whereby $A' = G' = R'$ (Figure 4.7b). Due to the different heading and incidence angles of the master image m and slave image s, the facade point A is imaged in A'' in image s (Figure 4.7b). A' and A'', representations of the same facade point A, are not on the same azimuth line, and correspond to two distinct ground points, $G' = A'$ and $G'' = A''$ (Figure 4.7a). In the same way, the roof information contained in A' and A'' comes from two distinct roof points, represented by R' and R'' in Figure 4.7a, respectively. The distance d between A' and A'' is the disparity to determine. It can be split into two parts: dr, due to the difference of incidence angles of both images, and da, due to the difference of heading angles ζ of both images. Estimating dr and da allows reducing the search area for matches along the range and the azimuth direction, respectively. Due to layover, matching of the two facade's contributions A' and A'' involves matching different ground and roof contributions, which do not represent the same scatterers. However, the facade contribution is the most important contribution of layover, and the most important one for retrieving disparity at facade points and consequently estimating building height. Hence in this work, the focus is put on matching facade points. Determining the maximum dr and da arising in layover areas gives an idea about the maximal matching error induced to ground and roof points, but also allows to determine the necessary dimensions of the search area for matching. Regarding some geometric considerations in Figure 4.7, the layover length $l_{A'}$ at point A' in slant range for the master image m is expressed as:

$$l_{A'} = h_A \cos\theta_m \tag{4.16}$$

whereby h_A is the height of facade point A, and θ_m the incidence angle of master image m. A similar equation can be written for the layover length $l_{A''s}$ in the slave image:

$$l_{A''s} = h_A \cos\theta_s \tag{4.17}$$

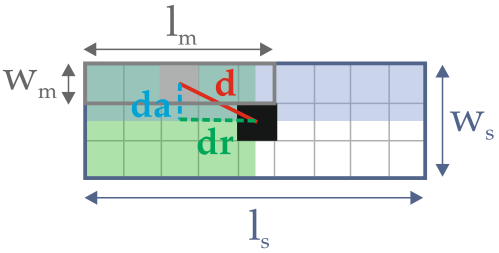


Figure 4.8: Dimensions of template (w_m, l_m) and search (w_s, l_s) windows for matching

However, the slave image s is reprojected in the slant geometry of image m during coregistration, inducing a scaling effect on $l_{A''s}$ described in (Goel & Adam 2012). The coregistered layover length $l_{A''}$ at point A'' for the slave image corresponds therefore to:

$$l_{A''} = \frac{\sin\theta_m}{\sin\theta_s} l_{A''s} \tag{4.18}$$

Additionally considering the difference of heading angle ζ between both images, the necessary width da of the search window in azimuth direction for point A is expressed as:

$$da = l_{A''}\sin\zeta = \frac{\sin\theta_m}{\sin\theta_s} h_A \cos\theta_s \sin\zeta \tag{4.19}$$

In this work, the building height h, and therefore facade points' heights h_A are searched for. However, average building heights in a specific scene are assumed to be known. Depending of the considered area, land use maps are available and permit to evaluate a maximum building height for the area. For example, in financial districts, high-rise buildings are frequent, whereby in residential districts, low-rise buildings are predominant. Besides, prior information about the building heights are available from the InSAR processing (cf. Section 3.6). Fixing h_A to h_{max}, which is the expected building height known from InSAR, yields the determination of da_{max} from Equation (4.19), i.e. to the maximum expected disparity in azimuth direction. da_{max} is therefore the epipolar constraint for matching in layover areas: the search window in the slave image has the dimension da_{max} in azimuth direction. For sake of the implementation, a search window of dimension $w_s = 2da_{max} + 1$ in azimuth dimension is defined in order to obtain symmetrical windows, centered on the point of interest, but the area of search for the maximum value of the matching criterion is then restrained to the correct half, as represented in Figure 4.8 in blue. In Section 6.3.2, a table depicts the values of da_{max} depending on the available acquisitions and corresponding difference of heading angles ζ, taking into account the expected maximum height value in the scene of interest.

Besides da_{max} in azimuth direction, the maximum expected disparity between both images at layover location in range direction, dr_{max}, is determined. Regarding Figure 4.7, dr is expressed as:

$$dr = \sqrt{l_{A''}^2 - da^2} - l_{A'} = \sqrt{l_{A''}^2(1 - \sin^2\zeta)} - l_{A'} \tag{4.20}$$

Replacing h_A by h_{max} in Equations (4.16) and (4.17), the maximum expected disparity value in range direction dr_{max} is defined. Those values are also given in Section 6.3.2 for different configurations. dr_{max} allows to determine the necessary search area for matching in range direction. Considering Figure 4.8, the second dimension l_s of the search area has to follow the rule:

$$dr \leq \frac{l_s}{2} - \frac{l_m}{2}, \text{ i.e. } l_{s_{min}} = 2dr_{max} + l_m \tag{4.21}$$

$l_{s_{min}}$ is the minimum length of the search window so that dr_{max} can be retrieved, using a template window in the master image of size l_m. Knowing the incidence angles of both master and slave, it is straightforward to deduce the direction of the disparity within the layover. Considering the symmetry of the search window with respect to the point of interest, the search for the correct match can also be restrained to half of the search window, represented in green in Figure 4.8. Additionally to the previous half defined from the azimuth direction, this leads to a privileged quarter (upper left side in Figure 4.8) where layover matches can be found. Practically, a match is retained in this area if its criterion value is at least 95% of the highest criterion value of the whole search window. This allows to privilege matching of layover areas without completely neglecting the building surroundings, that could lead to displacements in the other direction.

As a conclusion, in areas affected by important layover, a consideration of the epipolar constraint is mandatory for setting precisely the parameters for matching, which ensure to spare computation time and avoid wrong matching. Another constraint mentioned in Section 2.4.2 is not fulfilled in layover areas: the assumption of continuous disparity. At the border of layover areas, at near range, the disparity values are very high, corresponding to the building height. The neighboring pixels, situated at nearer range, belong however to the ground, and show smaller disparities. Therefore, disparity jumps occur around layover areas, yielding discontinuous disparities.

4.6 Building Extraction

This step focuses on the recognition and extraction of building facades from the created disparity map. It is performed at facade level, whereby areas surrounding each facade are extracted using results of the previous InSAR processing. This fusion procedure between InSAR results and the disparity map is explained in Chapter 5. In this section, the signature of buildings in the disparity map is first analyzed, and corresponding characteristic shapes are extracted subsequently. Possible modifications of the building parameters between InSAR data and radargrammetric data due to changes are considered. The extracted shapes are eventually filtered, in order to estimate final building parameters.

4.6.1 Facade Signature in the Disparity Map

Figure 4.9d shows the disparity map of a single facade, calculated using the method described in Section 4.5 on speckle simulated data. The simulation is performed by adding uniformely distributed speckle noise to the image, considering only double-bounce, layover and surroundings. This is not exactly the Rayleigh distribution defined in Section 2.2.3, but is sufficient for the aim of the simulation. Both master and slave simulated images are represented Figure 4.9(a,b), whereby the double-bounce line is situated at far range. Figure 4.9c shows the corresponding criterion values. The footprint of the layover of the master image is represented in blue in the slave image and in the disparity map, for better interpretation. Two distinct parallelograms, each presenting homogeneous disparity values, are observable, on each border of the layover of the

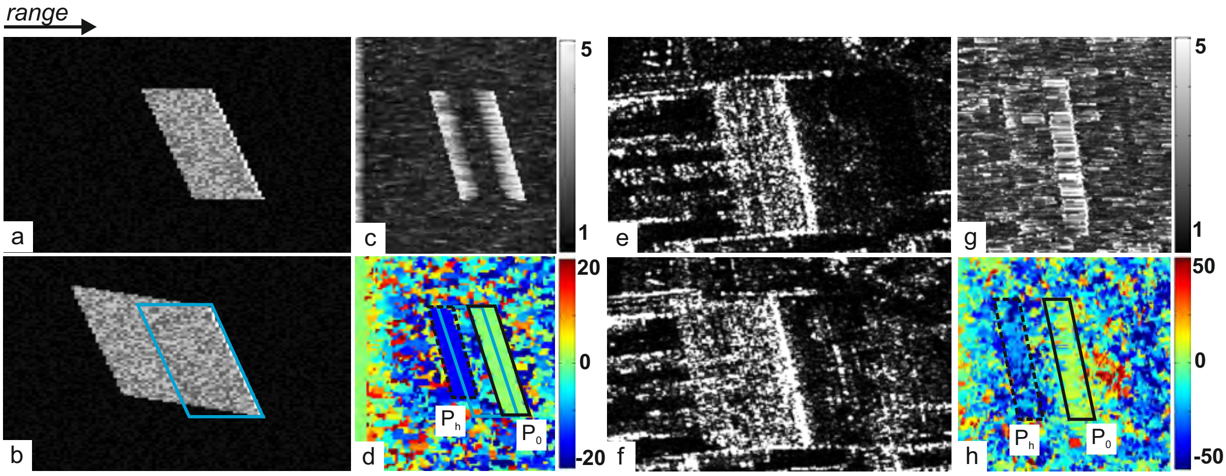


Figure 4.9: Building signature in the disparity map, in simulated and real data: (a) speckle simulated master image; (b) speckle simulated slave image; (c) matching criterion values ν for the simulated data; (d) disparity map for the simulated data; (e) real master image; (f) real slave image; (g) matching criterion values ν for the real data; (h) disparity map for the real data

master image. The parallelogram P_0 is situated around the double-bounce line, and shows homogeneous disparity values around zero (green). As slave is coregistered on master, the building footprints are situated at the same position and show zero disparity. The second parallelogram, P_h, is situated around the border of the layover of the master image. It shows homogeneous disparity values, different from zero (blue). Those values correspond to the difference $d = l_{A'} - l_{A''}$ (cf. Section 4.5), whereby the facade point A would be at the highest facade position. This corresponds to the difference of layover length between both coregistered images. In this example, the negative disparities are due to the respective size of the layover in the master and in the slave image, due to their respective incidence angles. Inverting the incidence angles of master and slave, i.e. choosing $\theta_m < \theta_s$ would provide positive disparities. The transformation of the disparity values into height values is explained at the end of this section.
Next to the simulated data are corresponding real data results. The parallelogram P_0 is easier to distinguish than P_h, and seems less noisy. This is due to the higher intensity values of the double-bounce line, which make its matching between both images unequivocal. Moreover, the presence of parking lots in front of the building (horizontal bright patterns), influences the estimation of matching criterion at the near layover edge, leading to more noisy estimated disparity values within P_h. On the simulated data, P_h is smaller than P_0. It is due to the difference of heading angle ζ between both acquisitions, exaggerated in the simulated data. The length difference corresponds to the previously defined da. However, in the real data, this difference is barely visible, hence two parallelograms of the same dimension are assumed in the following.

4.6.2 Facade Extraction

In order to retrieve the building parameters, both parallelograms have to be extracted. Their extraction can be facilitated considering previous results from InSAR processing and matching parameters.

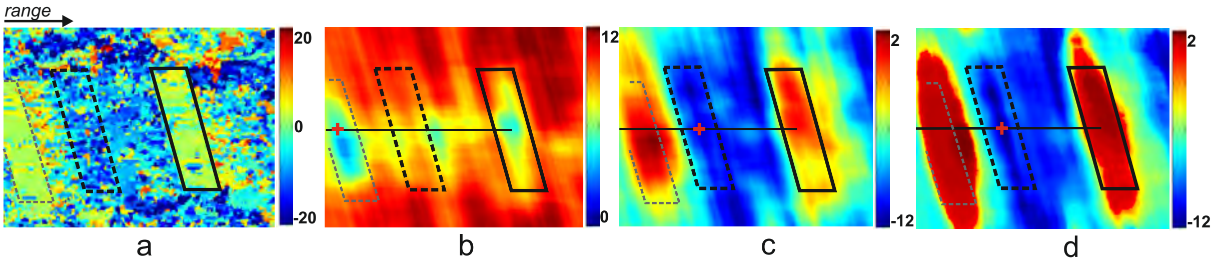


Figure 4.10: Strategy for parallelogram extraction in the disparity map: (a) disparity map; (b) standard deviation map, calculated for each pixel position within the parallelogram footprint represented in black; (c) mean map; (d) median map

The widths w_0 and w_h of both parallelograms are defined by:

$$w_0 = w_h = l_m - 1, \text{ if } l_m - 1 < L_m \tag{4.22}$$

whereby L_m is the length of the layover in the master image. This length can be determined using the building height $h = h_{InSAR}$ determined during the previous InSAR processing: $L_m = h_{InSAR} \cdot \cos\theta_m$. The length of the template window l_m should be long enough in order to contain enough reliable radiometric information for matching, but it should not exceed L_m, in order to reduce mismatching due to the other layover border. Particularly for buildings where the double-bounce line is not well recognizable compared to the layover, both layover borders have similar radiometric ratios. This is also the reason why master images usually have larger incidence angle as the slave images: within the search area, it should not be possible to find an area of the size of the template window containing both layover edges. This avoids mismatching. The direct proportionality existing between the size of the chosen template window l_m and the widths w_0 and w_h of the parallelograms is caused by the chosen matching criterion, which emphasizes the matching of edges. As long as an edge is present within the template, the matching criterion will favor this position. Sliding the template along the range direction would imply sliding the best match possibility within the search window of the same amount, therefore the disparities stay the same as long as an edge is contained in the template, i.e. for $l_m - 1$. The area between both parallelograms is characterized by $w_m = L_m - (l_m - 1)$.

The lengths l_0 and l_h of the parallelograms are defined using the facade length of the previous InSAR processing: $l = l_{InSAR}$. Using the determined facade orientation $\alpha = \alpha_{InSAR}$ from the InSAR processing as well, the parallelogram shapes can be constructed and used for further extraction. The transmission of the InSAR length and orientation into the geometry of the master image is explained in Section 5.1.

Using a different parameter setting for l_m with respect to L_m, other parallelogram widths and building signatures would be expected.

Knowing the expected parallelogram dimensions $w_{0,h}$, $l_{0,h}$ and $\alpha_{0,h}$ as well as the position of the double-bounce line (cf. Section 5.1), their extraction is straightforward. The extraction of P_0 consists in extracting the defined parallelogram shape around the double-bounce line, whose position is known from InSAR processing (Section 3.6), the double-bounce line being exactly centered within the parallelogram. Practically, due to geocoding issues, only an approximative position of the double-bounce line is given. The correct position of P_0 is adjusted looking for the minimal absolute mean within the defined parallelogram around the coarse double-bounce position (Figure 4.10c). The use of minimal standard deviation would be possible, yet in cases where P_0 is more noisy, this could lead to wrong extraction results. As the disparity map is in slant range geometry, the parallelogram P_h is centered on the same azimuth line as P_0. The search for this parallelogram is reduced along a horizontal direction, from far to near range. However, the extraction of P_h requires a further analysis of the disparity values within the defined parallelogram shape. Standard deviation, mean and median disparity values within the parallelogram are considered. The final position of parallelogram P_h is defined as being the one for which the median value within the parallelogram is the highest, along the azimuth line. Having a look at Figure 4.10, the choice of the median value is obvious: first, it is not influenced by noisy disparity values within the parallelogram, as could be the mean calculation. Second, the parallelogram P_h is more noisy than P_0, and the calculated standard deviation within the parallelogram at its correct position could be higher than the determined standard deviation over a neighboring, flat area.

Figure 4.10 shows the standard deviation, mean and median maps produced by determining the corresponding criteria within the defined parallelogram shapes at every pixel position in the image. The azimuth line for search and the correct position of P_h are drawn in black. The maximum criterion values (minimal standard deviation, maximum absolute mean and median) along the azimuth line are represented as the red cross. The median value proves to provide a better, unambiguous estimation of the correct position of P_h. The highest absolute median value is searched for, as the disparity is expected to be the highest at the layover border, as a flat area is assumed for the direct building surroundings. The distance w_m between both parallelograms is used for restricting the search for P_h along the range direction so that the final position of P_h is not higher that h_{InSAR}.

The extraction algorithm developed in this work leaves scope for detecting possible changes in the height and shape of the building, starting from the determined InSAR parameters. Changes of the building height can be deduced by analyzing the disparity values within P_h. A lower building height will produce smaller disparities. Besides, a subsequent lower median value within the extracted parallelogram P_h would be calculated, and its estimated position would be closer to P_0. Aside from changes in the building height, changes in the building shape, i.e. facade length, are determined. For this, adaptive parallelogram lengths are defined. First, changes in P_0 are analyzed. Starting from the theoretical facade length $l_0 = l_{InSAR}$, the parallelogram length is reduced until $l_0 = 1$. In this work, only a diminution of the building extend was allowed, as changes due to collapse or demolition are considered, and no new construction. For every value of l_0, the standard deviation within the redefined parallelogram shape is calculated, at every possible pixel position k along the double-bounce line. The criterion of standard deviation is chosen here, as P_0 shows low noise in the disparity map, and high standard deviation values imply a higher amount of noise. This subsequently indicates a possible reduction of the parallelogram length, i.e. a possible change. For every l_0, the pixel position k_{l_0} is retained for which the smallest standard deviation within the parallelogram is achieved. The value of the standard deviation is normalized by the number of pixels within the parallelogram P_0 for the length l_0. The final parallelogram length $l_{0_{radarg}}$ corresponds to the length l_0 for which the smallest normalized standard deviation is achieved. Consequently, the parallelogram position corresponds

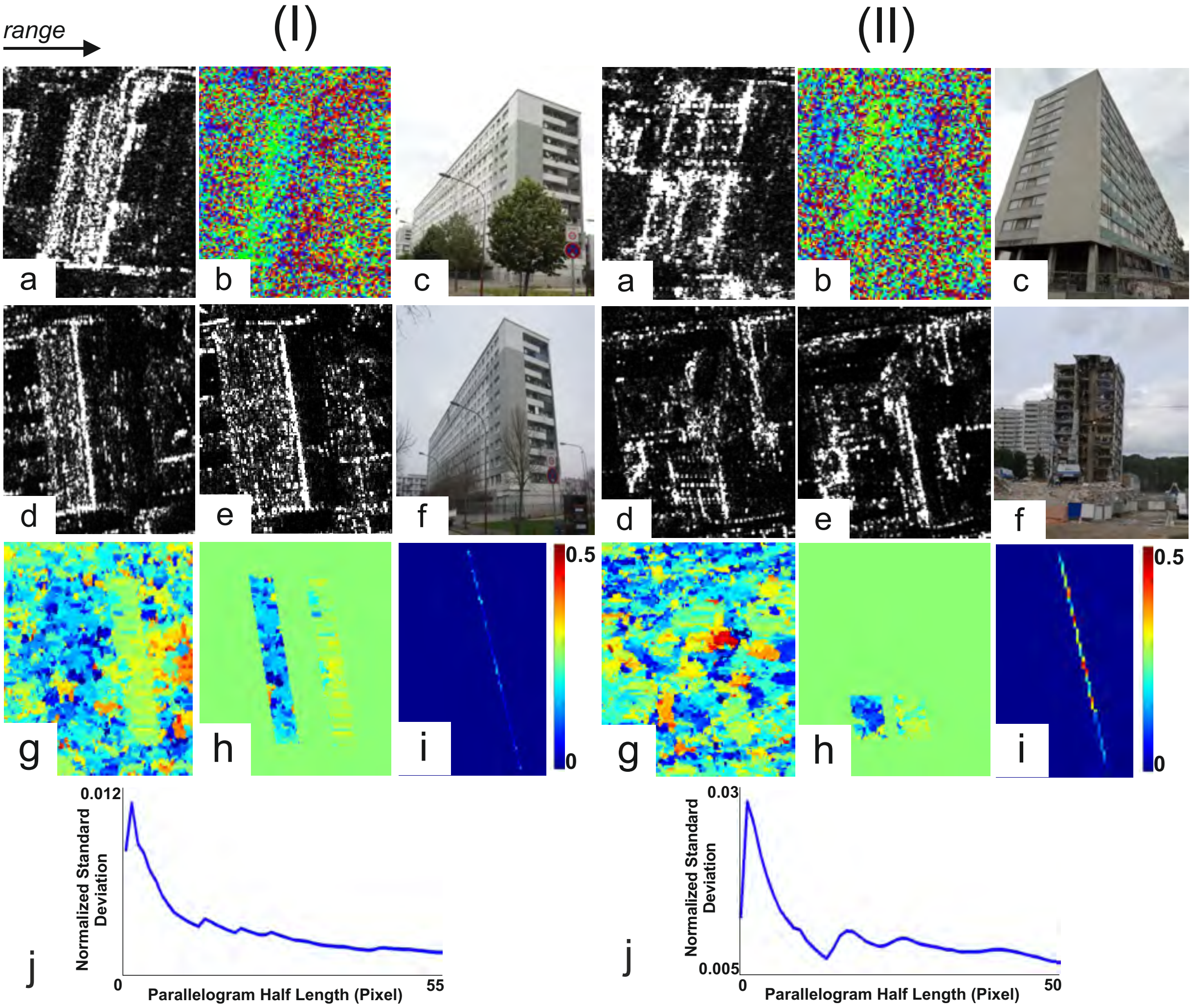


Figure 4.11: Parallelogram extraction for two differently affected buildings (I) and (II): (a) InSAR amplitude image (before change); (b) InSAR phase image (before change); (c) optical image (before change); (d) radargrammetric master image (after change); (e) radargrammetric slave image (after change); (f) optical image (after change); (g) disparity map; (h) extracted parallelograms; (i) values of the standard deviation for every pixel along the double-bounce line, for $l = 1$; (j) minimal determined standard deviation with P_0 for $l = [1, l_{InSAR}]$, normalized by the number of pixels within the created parallelogram

to the pixel position $k_{l_0} = k_{l_{0_{radarg}}}$ for which the minimal standard deviation is achieved for the specified length $l_{0_{radarg}}$. The relation between P_0 and P_h is assumed to be injective in this work, i.e. changes in P_0 imply changes in P_h, but changes in P_h can occur without changes in P_0. Indeed, a building can be only partly destructed, implying that its height and possibly the length at its top would change, but its foot stays the same. However, it is not possible that changes occur at the foot of the building without the top being affected. Moreover, changes at the top are assumed to be at least of the same amount as those on the foot. Changes in the second parallelogram P_h are subsequently determined using $l_{0_{radarg}}$ and $k_{l_{0_{radarg}}}$. P_h is searched along the azimuth line corresponding to the pixel position $k_{l_{0_{radarg}}}$ of the center of parallelogram P_0 of length $l_{0_{radarg}}$. Here also, the position is searched for which the highest median value within the defined parallelogram shape of length $l_{0_{radarg}}$ is achieved. Then, the same process is performed as for P_0, along the parallelogram orientation, in order to determine further possible

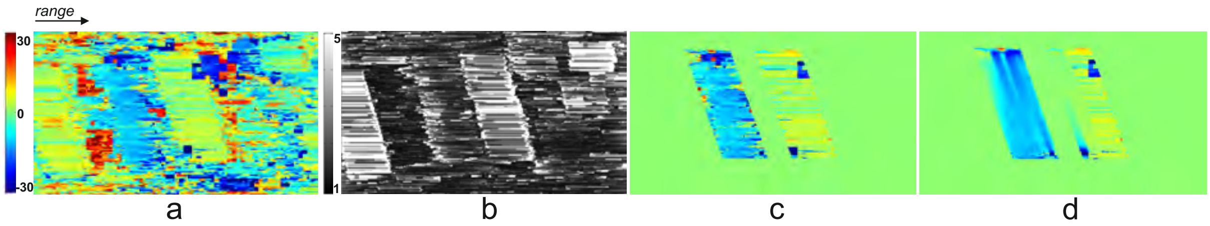


Figure 4.12: Filtering of parallelogram P_h: (a) height map; (b) criterion map; (c) extracted parallelograms; (d) filtered height map

changes of the length of the building top, yet considering l_h starting from $l_h = l_{0_{radarg}}$ to $l_h = 1$. Figure 4.11 depicts this strategy for two different buildings, differently affected: the first building did not change between InSAR and radargrammetric acquisition, and the length of the second building changed between both acquisitions. Figure 4.11(I)i shows the standard deviation of the disparity values within the parallelogram for $l = 1$, for P_0 of the first building. The minimal standard deviation, normalized over all possible lengths, is obtained for $l_{radarg} = 106$, i.e $\frac{l_{radarg}}{2} = 53$. This is also the original length l_{InSAR}. For the second building, the minimum standard deviation is achieved for $l_{0_{radarg}} = 28$, i.e. $\frac{l_{0_{radarg}}}{2} = 14$, as can be observed in Figure 4.11(II)j. The standard deviation itself is a factor of ten higher than the first building. For this building, the minimal standard deviation on the top is observed for $l_{h_{radarg}} = l_{0_{radarg}} = 28$, meaning the building was affected by the same amount on the top and the foot, as visible in Figure 4.11(II)f. The extracted parallelograms are shown in Figure 4.11(I)h and (II)h, respectively. The utilization of several configurations can yield different results and is therefore useful for avoiding false change detection.

4.6.3 Filtering and Estimation of Facade Parameter

Similar to the InSAR processing, a filtering is performed within the extracted P_h, in order to smooth the disparity values and improve building height determination. The dynamic filtering presented in (Dubois et al. 2012) and used for the InSAR processing in Section 3.6.2, is employed, considering disparity values instead of phase values. Moreover, the coherence weighting is replaced by a weighting with the values of the matching criterion at the corresponding pixel positions. Only filtering within P_h is performed, as it is the parallelogram from which the building height is retrieved. Figure 4.12 shows the result of filtering and the input disparity and criterion map.

The facade parameters can be retrieved subsequently. The facade length determined by radargrammetry is the final length $l_{(0,h)_{radarg}}$ determined for the extracted parallelograms. Here, the length of the top $l_{h_{radarg}}$ can differ from the length of the building foot $l_{0_{radarg}}$. The orientation is assumed not to change. The final building height is determined by considering the filtered disparity values within the final filtered parallelogram P_h. In this work, the relative building height is determined (cf. Section 2.4.4). The parallelogram P_0 shows zero disparities and therefore a zero height value. Consequently, the conversion of the disparity values of P_h into height values directly provides the building height. Considering the scaling of the layover presented in (Goel & Adam 2012) and the newly introduced influence of heading angle ζ (cf. Equation (4.19)), the final disparity d depicted in Figure 4.7 ($d^2 = da^2 + dr^2$) is expressed using trigonometric relations in triangle $(A''A'A)$ as:

$$d^2 = l_{A''}^2 + l_{A'}^2 - 2 l_{A''} l_{A'} \cos\zeta \tag{4.23}$$

Expressing $l_{A''}$ and $l_{A'}$ as function of building height h_{radarg}, using Equations (4.16), (4.17) and (4.18) yields (Dubois et al. 2013):

$$h_{radarg} = \sqrt{\frac{d^2}{D}}, \text{ whereby } D = \sin\theta_m^2 \cot\theta_s^2 + \cos\theta_m^2 - 2\sin\theta_m \cos\theta_m \cot\theta_s \cos\zeta \tag{4.24}$$

Another formula was derived in (Desai 1997). There, different formulas for the height estimation are defined, considering different flight paths. The heading angle difference ζ is taken into account, but not the scaling of the slave image onto the master image (Goel & Adam 2012). Equation (4.24) considers both effects. It has been shown previously in this section that the disparities within P_h are homogeneous and correspond to the displacement of the layover border between both images. Considering Figure 4.7, these are the disparities at point A, whereby all points A are situated at the top of the facade. After conversion, the resulting height values are homogeneous and correspond to the final building height h_{radarg}. This conversion is performed on the filtered disparity map. The final building height h_{radarg} is determined by averaging the filtered height values within P_h.
Results regarding different configurations of acquisitions are presented and discussed in Section 6.3.3.

4.7 Conclusion

In this chapter, the radargrammetric processing chain has been presented. After the calibration of both images, a SAR-SIFT like algorithm is used for coregistration, differing from the usual coregistration methods used for radargarmmetry. Through this pre-processing, image geometry and radiometry are better preserved. Further focus was put on the matching method. The new method for improving the matching relies on a Laplacian pyramid, a backmatching and a filtering of the disparity map at the last pyramid level for obtaining homogeneous disparities. A thorough analysis of the building appearance in the disparity map has been performed, revealing for the chosen pixel-based matching method two parallelograms of the same size and homogeneous disparity values. The extraction algorithm has been presented, allowing possible changes in the building appearance with respect to previous InSAR processing. A strict analysis of the epipolar constraint at building location, considering difference of heading angle ζ, permitted to refine and restrict processing parameters and finally yielded a new calculation of the relative building height. Based on the final building parameters, change detection can be performed, by comparing the radargrammetric results with the InSAR results. This is shown in Section 5, whereby the fusion step between InSAR and radargrammetry is explained in details.

5

Fusion of InSAR and Radargrammetry

In the previous chapters, both interferometric and radargrammetric approaches for building detection and parameter extraction have been described exhaustively. In addition, it has been shown, how the building parameters deduced from the interferometric processing are embedded into the radargrammetric processing. It reduces first the search area for matching, and consequently the risk of wrong matching and the computation time. Second, the building detection itself is facilitated in the radargrammetric disparity map, knowing original building length and orientation from InSAR processing. For this purpose, the building position in the disparity map, hence in the radargrammetric image pair, has to be known. Its position in the InSAR dataset has to be transferred into the geometry of the radargrammetric master image. Furthermore, the building parameters used for the parallelogram extraction (l_{InSAR}, α_{InSAR}) need to be adapted to the new geometry ($l0_{radarg}$, α_{radarg}). Finally, new building parameters inferred during radargrammetric processing ($l0_{radarg}$, $l_{r_{radarg}}$, h_{radarg}) have to be investigated in order to determine a possible change of the building.

In this chapter, the fusion step of pre-event InSAR data and post-event radargrammetric data is enlightened. In a first part, the approach used for determining the building positions and extent in the radargrammetric dataset starting from the detected buildings in the InSAR data is explained (Section 5.1). Particularly, the transfer of the building parameters is addressed. In a second part, the change detection approach used in this work is outlined, relying on the different changes appearing on buildings in the test scene (Section 5.2). The test area as well as results of both processing chains and change detection are given in Chapter 6, leading to a qualitative and quantitative discussion.

5.1 Geometrical Matching

The determination of the buildings' positions in the radargrammetric image pair is mandatory for the radargrammetric processing. In this work, it is performed using the previous location information from the detected buildings in the InSAR data. For each building detected in the InSAR processing chain, its position in the radargrammetric master image has to be found. This allows to define regions of interest in both coregistered radargrammetric images around

the building location and perform the radargrammetric processing in this reduced area, in order to save computation time. For this purpose, the corner coordinates of the building footprint are used. After extracting patches in both coregistered images, the disparity between master and slave is calculated. To this goal, an approximate building height, h_{InSAR} is needed in order to reduce the search area for matching. Finally, the extraction of both parallelograms, corresponding to the building signature in the disparity map, as well as the determination of new building parameters, requires knowledge about the original building orientation α_{InSAR} and length l_{InSAR}.
The height h_{InSAR} can be directly transferred from the InSAR to the radargrammetric dataset, as it is not influenced by the acquisition parameters. α_{InSAR} is the relative orientation of the building in the InSAR dataset, being measured from the range direction of the InSAR data. As the acquisition parameters change between InSAR and radargrammetry, its value has to be adapted for the radargrammetric dataset, as the range direction of the radargrammetric master image may differ from the range direction of the InSAR dataset. The length l_{InSAR} has been determined by measuring the length of the parallelogram edge situated at far range on the double-bounce line, in the InSAR dataset. Even if its value in meter unit stays the same, due to the different acquisition geometry, the extent of the double-bounce line in the radargrammetric image is different, as observed from another sensor position. A corresponding length in pixels, considering the new azimuth and range directions of the radargrammetric master image, has to be deduced from l_{InSAR}. Finally, the building position itself is defined by the corner coordinates of the double-bounce line, for each facade. Those coordinates have to be transferred to the new geometry of the radargrammetric dataset, i.e in the geometry of the radargrammetric master image.
To this goal, a projection of the extracted features, especially of the corner coordinates of the double-bounce line, has to be performed from the InSAR to the radargrammetric dataset. This projection is explained in the following, as well as the subsequent adaptation of the original building parameters to the new geometry.

5.1.1 From Interferometric to Radargrammetric Geometry

A direct projection of the InSAR geometry into the radargrammetric geometry calls for an estimation of the transformation between the image coordinates of the InSAR and the radargrammetric image respectively. Theoretically, as both images are taken independently, from different sensor positions, this is not possible directly. Practically, the parameters of the transformation could be determined using the presented SAR-SIFT algorithm on the master images of both InSAR and radargrammetric dataset. This algorithm shows its limits for a large difference of incidence angles, or for different orbit geometries (i.e. ascending, descending), which restricts its usage as soon as totally different acquisition configurations between InSAR and radargrammetry occur. This will be shown in more details in Section 6.3.1. However, using the respective sensor positions from which both InSAR and radargrammetric master images were acquired, an indirect transformation of the InSAR coordinates into the radargrammetric image coordinates is possible. The sensor position, as well as the necessary components of its velocity vector, are contained in the product data file of each acquisition. The indirect transformation consists first in geocoding

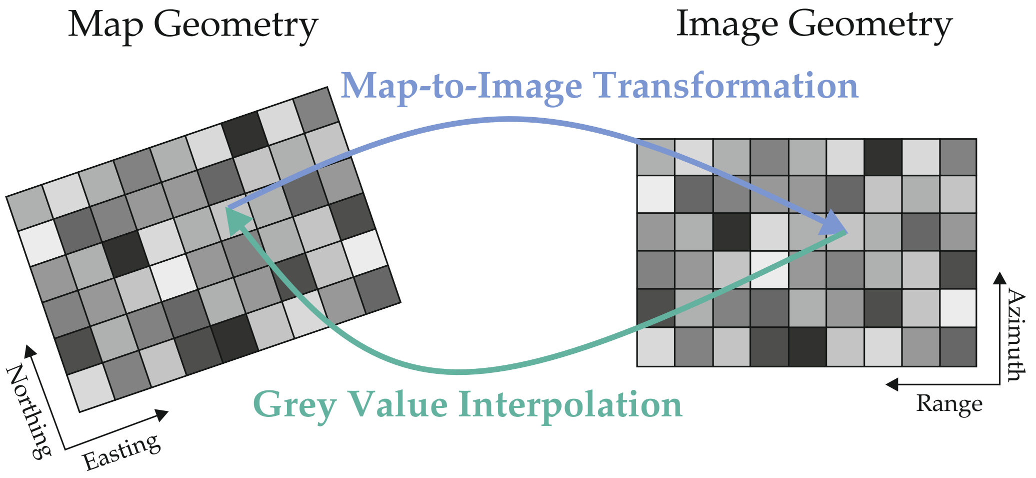


Figure 5.1: Schematic representation of backward geocoding

the InSAR image (*forward geocoding*), and second in projecting the geocoded image into the geometry of the radargrammetric master image (*backward geocoding*).

Forward Geocoding

The forward geocoding consists in determining the map coordinates (E, N) of each image point (x_{InSAR}, y_{InSAR}) in a specified map projection system. Thus, the 3D points have to be calculated in a world coordinate system $(\varphi_w, \lambda_w, H_w)$ first, followed by a projection on a reference plane defined by (E, N). For this, a reference surface for the Earth has to be defined. In some cases, a simple ellipsoid is considered, which leads to *ellipsoid corrected* images. More often, a more exact representation of the Earth surface is considered, by using Digital Elevation Models (DEM). The image is then *terrain corrected* (Schreier 1993). Usually, the DEM created by the SRTM mission is used, as freely available for the whole Earth surface. Most software solutions download the necessary region automatically.

In practice, geocoding is performed using the backward technique (Small & Schubert 2008), i.e. for this work, two subsequent backward geocodings are performed.

Backward Geocoding

As its name implies backward geocoding is the reverse operation of the forward geocoding. Starting from a projected map point (E, N), its corresponding position $(x_{InSAR/radarg}, y_{InSAR/radarg})$ (in range and azimuth, respectively) is retrieved in the slant range SAR image. A schematic representation is given in Figure 5.1. It consists in two main steps: map-to-image coordinate transformation, and grey value interpolation (Small & Schubert 2008, Schreier 1993). The second step is performed at the end of the geocoding, and consists only in interpolating the resulting grey value in the map geometry from neighbor pixels of a non-geocoded SAR image, using one of the usual criteria (nearest neighbor, bilinear or bicubic interpolation). The map-to-image transformation consists in the following steps (Schreier 1993):

- transformation of projected map coordinates into world geographic coordinates:

$$(E, N) \rightarrow (\varphi_w, \lambda_w) \tag{5.1}$$

- transformation of world geographic coordinates and topographic height into local cartesian system coordinates
$$(\varphi_w, \lambda_w), H_w \rightarrow (X_l, Y_l, Z_l) \tag{5.2}$$
- transformation of local cartesian system coordinates into geo-centered cartesian coordinates through Helmert transformation
$$(X_l, Y_l, Z_l) \rightarrow (X, Y, Z) \tag{5.3}$$
- transformation of geo-centered cartesian coordinates (X, Y, Z) into image coordinates
$$(X, Y, Z) \rightarrow (x_{InSAR/radarg}, y_{InSAR/radarg}) \tag{5.4}$$

The only transformation that differs from the usual geocoding procedure is the last one, as it takes into account the SAR specific side-looking geometry. An overview of the other transformations can be found in (Schreier 1993). The transformation of geo-centered cartesian coordinates into image coordinates starts from the Equations (2.22) and (2.25) of the RADAR model, defined in Section 2.4.4, whereby the Doppler frequency shift f_d is not zero. Both equations can be rewritten for an arbitrary image point (x, y) corresponding to a real point P as:

$$\begin{cases} F_1(x,y) = f_d - \frac{2}{\lambda} \cdot \frac{\langle \vec{S}, \vec{SP} \rangle}{\vec{SP}} = 0 \\ F_2(x,y) = r_0 - m_r x - |\vec{P} - \vec{S}| \end{cases} \tag{5.5}$$

r_0 and m_r are two constants, respectively image and sensor specific. S is the sensor position, dependent of the azimuth position y. Both f_d and S are approximated by polynomials:

$$\begin{cases} f_d = d_0 + d_1 x + d_2 x^2 + d_3 y + d_4 y^2 \\ \vec{S} = \begin{cases} x_s = a_0 + a_1 y + a_2 y^2 + a_3 y^3 \\ y_s = b_0 + b_1 y + b_2 y^2 + b_3 y^3 \\ z_s = c_0 + c_1 y + c_2 y^2 + c_3 y^3 \end{cases} \end{cases} \tag{5.6}$$

All polynomial coefficients (a_k, b_k, c_k, d_k) are given in the product files.
Both Equations (5.6) are non-linear in x and y, and must be solved iteratively, using an approximate solution (x_0, y_0) for x and y. The geo-centered cartesian coordinates of the scene center $(X_{P_0}, Y_{P_0}, Z_{P_0})$ can be deduced from the geographical coordinates (φ_0, λ_0) given in the product file and an additional DEM, and corresponding (x_0, y_0) is determined straightforwardly, thus this point is used as first approximated value. The Equations (5.6) can be solved in $(\Delta x, \Delta y)$ using the following partial derivatives:

$$\begin{cases} F_1(x_0, y_0) + \frac{\partial F_1}{\partial x} \Delta x + \frac{\partial F_1}{\partial y} \Delta y = 0 \\ F_2(x_0, y_0) + \frac{\partial F_2}{\partial x} \Delta x + \frac{\partial F_2}{\partial y} \Delta y = 0 \end{cases} \tag{5.7}$$

After the improvement of the image scene center to $(x_0 + \Delta x, y_0 + \Delta y)$, this result is used as starting value for a neighbor point of the geolocation grid defined using the geographical coordinates of the scene corners and center.
A complete description of the backward geocoding approach, considering also datum transformation, can be found in (Schreier 1993).

Practical Application

Backward geocoding is implemented in numerous software solutions. For this work, the Open Source software NEST (ESA) is used, whose geocoding method follows the method explained previously (Small & Schubert 2008). The DEM created by the SRTM mission is used as reference for the Earth surface. The output of the geocoding procedure is a geocoded map, whose pixel coordinates correspond to map coordinates in a chosen projection system. With this freeware, it is not possible to reproject a geocoded image into a given geometry, as for example the geometry of the radargrammetric master image. Of course, it is possible to perform a geocoding of both InSAR and radargrammetric master images, and extract a patch of the geocoded radargrammetric image around the desired building area. This has the drawback, that the radargrammetric images are transformed, and not in slant geometry anymore. Straight lines such as building double-bounces or layover edges appear distorted, and the radargrammetric processing, valid for images in slant range geometry, cannot be transferred straightforwardly in the new geometry. The building coordinates have to be determined in the slant range geometry of the radargrammetric master image. For this purpose, meshgrids containing the pixel coordinates in the InSAR and in the radargrammetric master images are created. Those meshgrids are geocoded using the product metadata of the corresponding amplitude images. Knowing the image coordinates (x_{InSAR}, y_{InSAR}) of a point (e.g. building corner) in the InSAR data, it is straightforward to look for the pixel showing these values in the geocoded meshgrids. This pixel is characterized by (E, N). The pixel coordinates (x_{radarg}, y_{radarg}) corresponding to (x_{InSAR}, y_{InSAR}) in the radargrammetric master slant range geometry are given at the position (E, N) in the geocoded meshgrids of the radargrammetric master image.

As both InSAR and radargrammetric datasets do not have the same extent, the freeware QGIS is complementary used in this work, in order to obtain images of same dimension and pixel size. Due to the accuracy of the chosen DEM, the reprojected building corners in the radargrammetric master image may differ from the actual position of the corners in the image. However, this accuracy is sufficient for the radargrammetric processing. In a first step, it is only necessary to have an approximate building position in order to extract patches in the radargrammetric images for further processing. Also, for the later parallelogram extraction, the approximate position of the corner is sufficient, as the optimal building position is searched using the position of the minimal absolute average of the disparity values within a specified parallelogram. Therefore, this method can also be used for fusing ascending InSAR data with descending radargrammetric data, or vice-versa. This is explained in more details in the next subsection, together with the propagation of the building parameters.

5.1.2 Transformation of the Building Parameters

The transfer of the building parameters from the InSAR to the radargrammetric dataset relies on the propagated position of the building corners. The positions (x_{radarg}, y_{radarg}) in the radargrammetric master image of the facade corners after geocoding allow the straightforward determination of facade length $l_{0_{radarg}}$ and orientation α_{radarg}.

The length $l_{0_{radarg}}$ corresponds to the norm between the two facade's corners. In meter unit, it should correspond to l_{InSAR}. However, in pixel unit, this length may differ from l_{InSAR} as the

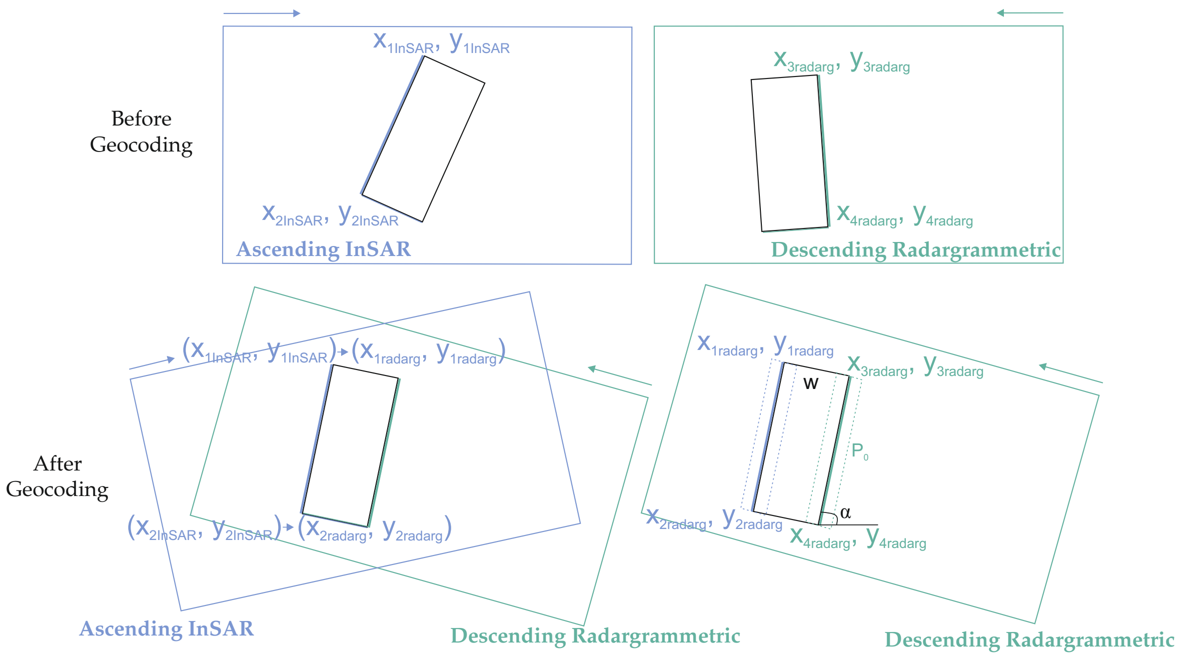


Figure 5.2: Estimation of corner coordinates in case of opposite-side acquisitions for geocoding

building may be oriented differently in the image due to the acquisition configuration (heading angle, ascending/descending path). The pixel spacing in azimuth and range direction may also differ between InSAR and radargrammetric dataset.

The parameter α_{radarg} is determined by calculating the arctangent of the angle formed by the line between both facade corners (double-bounce line) and the range direction.

The building height h_{radarg} is simply transferred as metadata from one image to the other, as it is independent of the heading angle. However, it is used for determining w_m (cf. Section 4.6.2) for the extraction of the parallelograms.

In the case of fusion of same-side configuration, e.g. ascending InSAR and ascending radargrammetric dataset, the extracted facade corners from the InSAR dataset correspond to the determined facade corners in the radargrammetric dataset. If their position differs, it is just about a few pixels due to inaccuracies induced by the geocoding. The real position of the facade footprint is found by searching the position for which the defined radargrammetric parallelogram P_0 shows the lowest absolute mean disparity.

In the case where InSAR and radargrammetric datasets are acquired from opposite directions, the facade corners extracted by the InSAR processing may not correspond to the facade corners represented in the radargrammetric dataset. Figure 5.2 depicts this issue, where only the visible corners in the respective acquisition are labeled before geocoding. In other words, the determined reprojected facade corners $(x_{(1,2)radarg}, y_{(1,2)radarg})$ from the InSAR coordinates $(x_{(1,2)InSAR}, y_{(1,2)InSAR})$ may not have any corresponding distinctive feature in the radargrammetric dataset. Indeed, they may be situated in the building shadow. However, the length $l_{0_{radarg}}$ and orientation α_{radarg} determined with these coordinates are correct, as the building hypothesis is rectangular. The determination of the position of the represented facade in the radargrammetric dataset delimited by $(x_{(3,4)radarg}, y_{(3,4)radarg})$ is relevant for the extraction of the parallelograms P_0 and P_h. It corresponds to the position for which parallelogram P_0 shows the lowest absolute mean disparity, searching from the determined corner coordinates $(x_{(1,2)radarg}, y_{(1,2)radarg})$ towards near range.

After every parameter has been transferred, the radargrammetric processing allows to extract the building facades and possible new building parameters. The comparison between building parameters derived by interferometry and radargrammetry leads to the change decision.

5.2 Change Detection Rules

After the building parameters have been determined by both InSAR and radargrammetric processing, change detection can be conducted, by comparing pre-event with post-event parameters. In principle, if the parameters calculated during InSAR and radargrammetric processing are the same, no change occurred. On the contrary, if the outcomes from InSAR and radargrammetric processing are different for a specific facade, the occurrence of a change in the building facade is very probable. For example, considering only the facade length, if $l_{InSAR} = l_{radarg}$, there exists a good chance that no change occurred. If additionally $h_{InSAR} = h_{radarg}$, the chance that no change occurred becomes even higher. However, if $l_{InSAR} = 90$ m and $l_{radarg} = 50$ m, a change is more probable to have occurred, especially if h_{InSAR} is different from h_{radarg}.Considering a complementary post-event analysis of the facade with a radargrammetric image pair acquired with a different configuration, different outcomes may arise, for example $l_{radarg} = 87$ m. In this case, did the building change or was there in the first configuration an obstacle that hindered the complete building recognition? And does a length difference of 3 m between pre-event and post-event indicate a change or is it only due to the accuracy of the respective processing? It is also possible that the building parameter extracted by radargrammetry are more correct than those extracted by interferommetry. Both extracted pre-event and post-event data have different degrees of uncertainty, that make the change detection approach very challenging. As the outcomes may differ considering different acquisition configurations and different parameters, it is necessary, in order to determine with high reliability the occurrence of a change, to define decision rules that consider all extracted facade parameters and their extent in terms of precision. Furthermore, as it is possible to consider several pre-event but also post-event configurations taken in a short time span, such a decision strategy should enable to have more than one value for a specific parameter.

Several decision strategies are described in the literature. Three main rule systems can be distinguished: fuzzy, Dempster-Shafer's and Bayes' theory (Hedman 2010, Hommel 2010). The fuzzy theory deals with approximate reasoning. The occurrence of an event can not only be true or false, but can be for example true to a certain degree situated in the range $[0, 1]$, where 0 and 1 are false and true, respectively. The variables are not sorted into two classes ‚true‘ and ‚false‘, but they are assigned a degree of affinity situated in the range $[0, 1]$. By fuzzy logic, rules have to be defined by the operator, e.g. ‚If‘. . . ,and‘. . . ,then‘. Those are linguistic rules, which are interpreted into numerical values during a process called *inference*, leading to the final degree of affinity for all combined variables. The fuzzy logic was used by (Hommel 2010) for the classification of building states using LIDAR data. In (Dell'Acqua & Gamba 2001), fuzzy logic was used on SAR data for detecting urban structures, in particular road networks. As rules have to be defined by the operator before classification, fuzzy logic presents two drawbacks: first, the quantity of rules to be defined can vary with respect to the application, and may become tremendously time consuming. Second, once the rules are defined, they are fix and cannot evolve, for example new features cannot be taken into account, without having to define new rules. The Dempster Shafer's theory allows to overcome these problems, while leaving the possibility of uncertain or incomplete data. It builds upon the framework of the evidence theory, and is a generalization of the probability theory for the imprecise nature of an event. For classification applications, Dempster Shafer's theory is able to deal with not well defined or union of classes. Mass functions are defined for each class, and degrees of belief for new variables are given within an interval of probabilities, making the decision imprecise. In (Poulain et al. 2011), the authors use Dempster Shafer's theory in order to fuse features extracted in both optical and SAR image data. In (Tupin et al. 1999), Dempster Shafer's theory is used in order to classify several detected structures, belonging to urban areas, roads, forests and rivers. In both cases, a large quantity of features are extracted for each class, allowing a learning of the parameters for the definition of the mass function. An alternative to the Dempster Shafer's theory is the Bayes' theory, which is a special case of the evidence theory, whereby all probabilities, causalities and mass functions are known *a priori*. In (Kim & Nevatia 2003), the authors use a Bayesian network in order to extract building parameters in multiple optical images. In (Hedman 2010), Bayesian networks are used to extract roads in multi-aspect SAR imagery. Finally, in (Brunner 2009), the author uses the Bayes' rule in order to identify damaged and undamaged buildings in SAR images. Bayes inference implies to learn the *a priori* parameters of the classes from a set of training data. Yet, for specific applications and knowing the classes to be extracted, it is possible for the operator to define empirically the *a priori* parameters.

In this work, the difficulty lies in the very few post-event data, making an adequate learning solely from data impossible. Nevertheless, Bayes' inference is applied, as the *a priori* probabilities and mass functions can be deduced empirically. Each facade has usually only one value for each parameter. Considering more configurations is possible, however it shouldn't be at the extent of the short time needed for a post-event analysis. Here, the analysis is limited to one configuration. In order to define the best empirical parameters, it is useful to analyze thoroughly the available parameters, the possible changes and their relation. Section 5.2.1 resumes theses issues, and Section 5.2.2 presents the adopted strategy, in theory and practice.

5.2.1 Possible changes

Figure 5.3 shows the different changes that may occur at building location, considering a rectangular building with flat roof, and its resulting signature in the disparity map. Only changes that occur in this work are represented, yet this model could be extended to other changes as to those presented in (Hommel 2010), considering complementary parameters.

Four possible building states are referenced, whereby three of them represent a damaged building, at different extent.

- Figure 5.3(1): the building is undamaged. In this case, $l_{InSAR} = l_{0_{radarg}} = l_{h_{radarg}}$, $h_{InSAR} = h_{radarg}$, and $\alpha_{InSAR} = \alpha_{radarg}$.

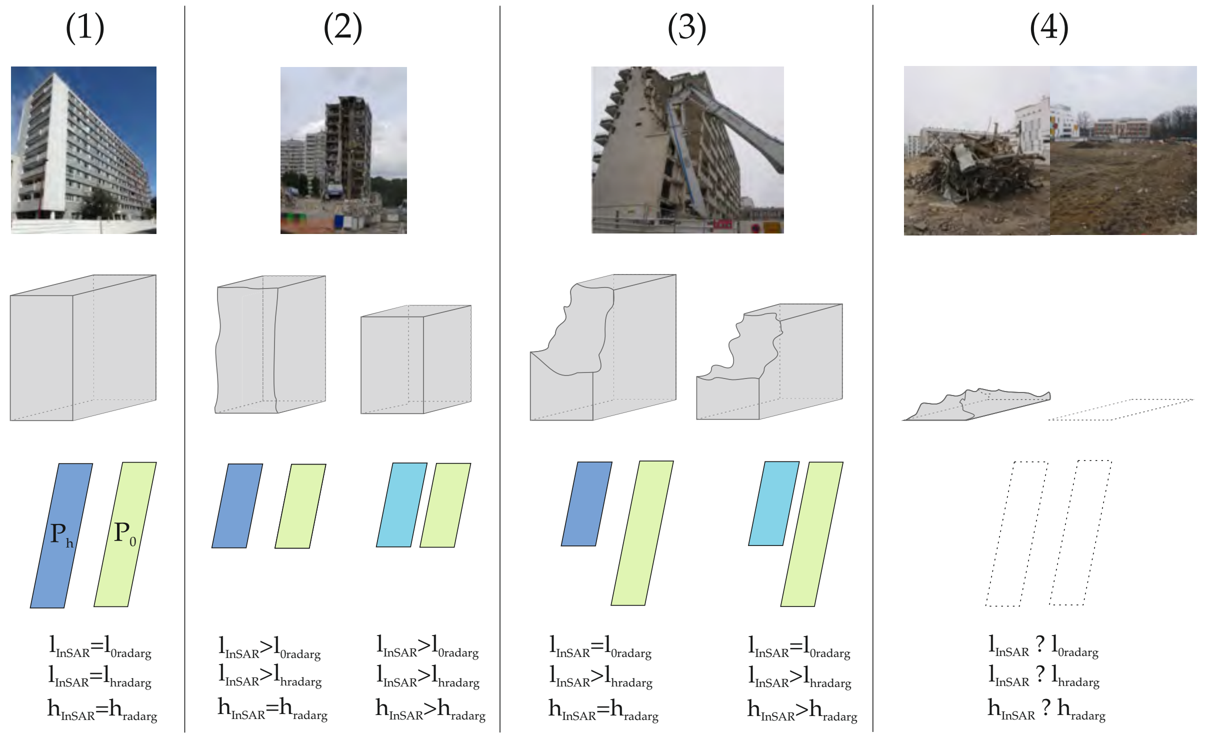


Figure 5.3: Building classes detected in this work; *first row*: example in an optical image; *second row*: schematic representation of the possible building states; *third row*: schematic representation of the corresponding building appearance in the radargrammetric disparity map; *fourth row*: correspondence between the parameters determined during the interferometric and radargrammetric processings.

- Figure 5.3(2): the building has been sliced vertically, leading to a change of the same amount in the lengths of P_0 and P_h. In this case, $l_{0_{radarg}} = l_{h_{radarg}}$, but $l_{InSAR} > l_{0_{radarg}}$. The height of the building may have stayed the same, as represented in the optical image Figure 5.3(2), but may also have not: $h_{InSAR} \geq h_{radarg}$.

- Figure 5.3(3): the building underwent a different change at the top than at the bottom, whereby the bottom of the building is unchanged. The parallelogram P_0 can be completely recovered, whereby P_h shows a shorter length: $l_{InSAR} = l_{0_{radarg}}$, $l_{0_{radarg}} \neq l_{h_{radarg}}$ and $l_{InSAR} > l_{h_{radarg}}$. As for the previous case, the height may have changed but not necessarily, so that $h_{InSAR} \geq h_{radarg}$.

- Figure 5.3(4): the building is gone, leaving only rubles or, in the case of a planned demolition, an empty space. However, the radargrammetric algorithm extracts parallelograms in the disparity map, which correspond in that case to no real building signature. Their length can be the same as the original building length from the InSAR data or vary, so can the building height and the orientation, therefore marked here with ‚?‘. However, the standard deviation of the disparity values inside the extracted parallelograms P_0 and P_h is much higher than the one in the extracted parallelograms when the building or a part of it is still standing. Indeed, when the building is still there, the standard deviation of the disparity values within the extracted parallelograms should be near 0 for both P_0 and P_h, as the disparities are homogeneous within one parallelogram (cf. Section 4.6.1).

The second and third possible building states could be subdivided into sub-classes, depending if the extracted building height h_{radarg} is the same as the original building height h_{InSAR} or not. Furthermore, a supplementary building state could be defined, considering the possibility of a new reconstruction at the same location as the old building. For the latter case, the assessment of the constant building orientation is not valid anymore, as for the building state (4) depicted in Figure 5.3(4). Additionally, the new length and height of the extracted building do not have to stay the same or be smaller than the original building parameters, but could become greater. Such cases are not implemented in this work, as they do not occur in the considered area.

From two original facade parameters (l_{InSAR}, h_{InSAR}), five parameters are derived from the radargrammetric processing ($l_{0_{radarg}}$, $l_{h_{radarg}}$, h_{radarg}, $\sigma_{0_{radarg}}$ and $\sigma_{h_{radarg}}$), which allow to distinguish the different building states, or *classes*, existing after the event. Having extracted a new facade, the five parameters, or *variables*, should allow its classification into one of the four defined classes. This is a difficult task if those variables are considered individually and simply compared to the original parameters, as some variables are the same in different classes. However, a combination of those variables into a global framework allows that a change in one variable influences the probability of this set of variables to belong to a specific class. Among the presented decision strategies, the Bayesian approach constitutes an appropriate framework, relying on the probability that a certain class occurs given a specific set of variables. In the following, the Bayes' theorem is recalled in some details, and its practical application in this work is described.

5.2.2 Bayes' Rule

In this subsection, the theory of the Bayes' rule is shortly recalled, with the appropriate notations. In a second part, the Bayes' rule is applied on the variables defined in this work, and their setting is explained in more details.

Theory

The Bayes' theorem describes the probability of an event to occur, knowing variables that *condition* the event. In other words, it describes the probability, given a certain combination of variables $X = x_1, x_2, ..., x_n$, that a certain event C occurs (Prince 2012). This probability is written as $P(C|X)$ and is called *posterior* probability. In order to calculate $P(C|X)$, prior information about the event and the variables are necessary:

- The probability $P(C)$ of the event itself, before considering any influence variables, has to be known. This probability is called *prior*.
- The conditional probability $P(X|C)$ of the variables given the event, called *likelihood* must be determined. It represents the tendency of the set of variables $X = x_1, x_2, ..., x_n$ to take specific values given a fix value of the event C.
- Finally, the *evidence* $P(X)$ has to be calculated. It officiates as norm factor within the Bayes' theorem, and represents the probability that the set of variables $X = x_1, x_2, ..., x_n$

occurs, without considering the event. If X is continuous, $P(X)$ is the distribution of X over C, expressed as a function of their joint distribution $P(X,C)$:

$$P(X) = \int P(X,C)dC = \int P(X|C)P(C)dC \tag{5.8}$$

In the case of discrete variables, the integral is replaced by a summation over all values of C:

$$P(X) = \sum_{k=1}^{m} P(X|C_k)P(C_k) \tag{5.9}$$

whereby the event C can take m discrete values.

The Bayes' rule intends to find a relationship between $P(X|C)$ and $P(C|X)$, and is expressed as:

$$P(C|X) = \frac{P(X|C)P(C)}{P(X)} \tag{5.10}$$

Knowing all the probabilities expressed in the right term, it is possible to calculate the probability of an event C given a certain occurrence of a combination of variables X.

As noticed from Equations (5.8) and (5.9), the probability distributions of X have to be known for determining $P(C|X)$. Depending on the domain of X, different models exist. X can be discrete, i.e. it can take only certain values of a predefined set, or continuous, i.e. it can take any real value. X can be finite, i.e. its range is limited, or infinite. X can be a single variable, in this case it is called *univariate*, or X can be a set of several variables (*multivariate*).
In this work, X is a set of six independent building parameters, which can take any values, some in a finite, and some in an infinite range. X is therefore multivariate, continuous and unbounded, which corresponds to a *multivariate normal distribution* (also called *multivariate Gaussian distribution*). An univariate normal distribution is defined by its mean μ and variance σ^2 as follows:

$$P(X) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-0.5\frac{(X-\mu)^2}{\sigma^2}\right] \tag{5.11}$$

Similarly, a multivariate normal distribution is defined by its mean μ and its covariance Σ, whereby μ is a nx1 vector describing the mean of each of the n variables, and Σ is a nxn matrix containing the single variances of each variable on its diagonal. The multivariate distribution is expressed as follows:

$$P(X) = \frac{1}{\sqrt{(2\pi)^{D/2}|\Sigma|^{1/2}}}\exp\left[-0.5(X-\mu)^T\Sigma^{-1}(X-\mu)\right] \tag{5.12}$$

Usually, the probability distribution of the variables is estimated using *training data*. For example, considering two events C_1 = *building undamaged* and C_2 = *building damaged*, one would express $P(X)$ using Equation (5.9). For this, the conditional probabilities $P(X|C_1)$ and $P(X|C_2)$ as well as the probabilities $P(C_1)$ and $P(C_2)$ would be needed. In the case of equiprobable events, $P(C_1) = P(C_2) = 0.5$. The estimation of $P(X|C_1)$ and $P(X|C_2)$ requires training data, i.e. building parameters for both undamaged and damaged buildings. Two normal distributions would be defined, corresponding to building parameters of undamaged, respectively damaged buildings. A lot of different combinations of variables are needed in order to train a

respective classifier and *learn* μ and Σ for each distribution. The more information about the data, i.e. the more buildings taken into account, the easier it is to fit a normal distribution to a set of variables. Afterwards, the estimation of the probability $P(C_k|X)$ of the event given the variables is facilitated. In this work, the total number of damaged buildings is too low to learn from them the building parameter distribution for damaged buildings. Furthermore, the parameters are dependent of each building so that each building has to be considered independently of the others, i.e. $P(C_k|X)$ has to be calculated for each building independently. Instead of fitting the probability distributions to training data, the probability distributions of the set of variables are defined in this work a priori, based on the theoretical behavior of the different buildings for the different damage classes. This step is detailed in the next subsection.
More information about the Bayes' rule, classifier and normal distribution can be found in (Prince 2012). Applications for classification using Bayes' rule and Bayesian networks can be found in (Brunner 2009) and (Hedman 2010), respectively.

Practical application

For a specific building, the building parameters before the event are l_{InSAR} and h_{InSAR}, and the building parameters deduced from the radargrammetric image pair are:

$$X = (l_{0_{radarg}}, l_{h_{radarg}}, h_{radarg}, \sigma_{0_{radarg}}, \sigma_{h_{radarg}}) \tag{5.13}$$

What is searched for is the probability, given the set of building parameters X, that the building belongs to one of the four building classes presented in Figure 5.3, i.e. $P(C_k|X)$.
Considering the Bayes' theorem, the *prior* $P(C_k)$ has to be determined first. Let's consider first the case where the probability that a damage occurs is the same as the probability of no damage, i.e:

$$P(C_1) = P(C_2) + P(C_3) + P(C_4) = 0.5, \text{ as } \sum_k P(C_k) = 1 \tag{5.14}$$

C_{2-4} are the three damage classes, and their occurrence is equiprobable, i.e.:

$$P(C_2) = P(C_3) = P(C_4) = 0.5 \cdot \frac{1}{3} \tag{5.15}$$

In Section 6.4.2, an example is given with other initial $P(C_k)$, which is closer to the reality in the case of planned demolitions.
Second, the *likelihood* $P(X|C_k)$ must be estimated. As mentioned previously, this corresponds to a multivariate normal distribution, defined in this work based on a theoretical analysis of the behavior of the building parameters considering the different classes. For every variable in the set of variables X, μ and σ corresponding to the behavior of the variable for the class C_k have to be estimated. Table 5.1 depicts the parameters μ and σ for each building parameter considering the corresponding building class. Here, the ratio $\frac{l_{0_{radarg}}}{l_{h_{radarg}}}$ is considered instead of $l_{h_{radarg}}$ only, in order to separate C_1 from C_3. The table has to be read as follows: for each building parameter x_j, the conditional probability $P(x_j|C_k)$ follows the normal distribution defined by μ and σ for the specified class C_k.

Class C_1 corresponds to an undamaged building. Therefore, as the probability is estimated for every building separately, it is evident that the building parameters follow the distribution of the

Table 5.1: Conditional likelihood of the building parameters given the classes

	$l_{0_{radarg}}$	$\frac{l_{0_{radarg}}}{l_{h_{radarg}}}$	h_{radarg}	$\sigma_{0_{radarg}}$	$\sigma_{h_{radarg}}$
C_1	$\mu = l_{InSAR}$ $\sigma = \sigma_{l_{0_{radarg}}}$	$\mu = 1$ $\sigma = SR$	$\mu = h_{InSAR}$ $\sigma = \sigma_{h_{radarg}}$	$\mu = M_{STD_0}$ $\sigma = S_{STD_0}$	$\mu = M_{STD_h}$ $\sigma = S_{STD_h}$
C_2	$\mu = l_{InSAR} - ML$ $\sigma = \sigma_{l_{0_{radarg}}}$	$\mu = 1$ $\sigma = SR$	$\mu = h_{InSAR} - MH$ $\sigma = \sigma_{h_{radarg}}$	$\mu = M_{STD_0}$ $\sigma = S_{STD_0}$	$\mu = M_{STD_h}$ $\sigma = S_{STD_h}$
C_3	$\mu = l_{InSAR}$ $\sigma = \sigma_{l_{0_{radarg}}}$	$\mu = MR$ $\sigma = SR$	$\mu = h_{InSAR} - MH$ $\sigma = \sigma_{h_{radarg}}$	$\mu = M_{STD_0}$ $\sigma = S_{STD_0}$	$\mu = M_{STD_h}$ $\sigma = S_{STD_h}$
C_4	$\mu = l_{InSAR}$ $\sigma = \sigma_{l_{0_{radarg}}} + \Delta_L$	$\mu = 1$ $\sigma = SR + \Delta_{SR}$	$\mu = 0$ $\sigma = \sigma_{h_{radarg}}$	$\mu = M_{STD}$ $\sigma = S_{STD_0}$	$\mu = M_{STD}$ $\sigma = S_{STD_h}$

original building, defined during the InSAR processing. For this class, the length of the double-bounce line should correspond to the length of the layover edge, therefore the ratio $\frac{l_{0_{radarg}}}{l_{h_{radarg}}}$ should be equal to 1. As the building is still standing, both parallelograms P_0 and P_h can be extracted over the whole building length, and their disparity values show a homogeneous behavior, leading to standard deviations close to 0. Experiences showed that these values yet differ from 0, and they are therefore set to M_{STD_0} and M_{STD_h}. Their respective values are given in Section 6.4.1. The building height should also correspond to the original building height derived by interferometry h_{InSAR}. The standard deviations of the parameters should however correspond to the accuracy with which those parameters can be estimated by radargrammetry. The standard deviation of the height parameter $\sigma_{h_{radarg}}$ also depends on the considered building model. In this work, only flat roof are considered, nevertheless in case of gabble roofs or other roof forms, this parameter could change.

Class C_2 is similar to class C_1, except that the extracted parallelograms are smaller than the original building length (of an amount of ML) and that the building height can vary (of an amount of MH).

For the class C_3, the building length determined with $l_{0_{radarg}}$ is the same as the original building length, but $l_{h_{radarg}}$ becomes smaller, leading to a ratio $\frac{l_{0_{radarg}}}{l_{h_{radarg}}}$ larger than 1, described here as MR. As for C_2, the building height can vary of an amount MH.

Finally, class C_4 shows high discrepancies to the other damage classes. As mentioned previously, the radargrammetric algorithm would extract two parallelograms. However, these correspond to nothing particular in the real word. Due to the higher number of pixels contained in the parallelogram, it is very likely that a parallelogram of length $l_{0_{radarg}} = l_{InSAR}$ is extracted. However, this is not mandatory, so that the standard deviation of this parameter is set higher as for the other classes. This explains the parameter setting for $l_{0_{radarg}}$ and $\frac{l_{0_{radarg}}}{l_{h_{radarg}}}$. As P_h contains disparity values corresponding to the ground, the expected extracted building height is 0. Finally, contrary to the other classes, the disparity values within both extracted parallelograms are heterogeneous, leading to higher standard deviations.

Every additional parameter presented in Table 5.1 ($SR, ML, MH, MR, MA, M_{STD}, M_{STD_0}, M_{STD_h}$), as well as the standard deviations corresponding to the radargrammetric processing are presented in Section 6.4.1. The standard deviations are calculated by evaluating the accuracy of the building parameters extracted by the radargrammetric processing, and the additional parameters are set by the operator, considering the obtained standard deviations and accuracy requirement for the decision making. The probability $P(X|C_k)$ is now estimated easily using the multivariate normal distribution described in Equation (5.12).

The last parameter needed for calculating $P(C_k|X)$ is the *evidence* $P(X)$, that is calculated straightforwardly using Equation (5.9).

Afterwards, the probability $P(C_k|X)$ can be calculated replacing every term in Equation (5.10) by its corresponding estimation. In this work, having extracted a specific set of variables X, $P(C_k|X)$ is calculated for $k = [1 : 4]$. The building is then classified to the class for which the maximum probability value $P(C_k|X)$ is achieved.

5.3 Conclusion

In this chapter, the fusion between interferometric and radargrammetric data has been thoroughly. First, the geocoding step, necessary for propagating the building parameters determined by InSAR into the radargrammetric geometry for further processing has been presented. After geocoding meshgrids containing the pixel coordinates of both InSAR and radargrammetric master images, correspondences of building corners can be established and building parameters transferred for the radargrammetric processing. After radargrammetric processing, the extracted parameters are compared to the parameters extracted by the InSAR processing in order to detect changes. For this purpose, a probabilistic approach using the Bayes' theorem has been presented, and its parameters have been set considering the different building damages taken into account in the scope of this work. In the next chapter, results of the presented change detection approach as well as the final parameter settings are shown.

6

Results and Evaluation

In this chapter, the results of both processing chains and of their fusion for detecting changes are presented. First, the chosen test area is introduced and the data acquired for pre- and post-event analysis are described (Section 6.1). Second, the results of building extraction by interferometry and radargrammetry are presented, in Sections 6.2 and 6.3 respectively. There, the parameter setting for specific processing steps and the achieved accuracy are discussed thoroughly. Consequently, the parameters for the change detection rule can be set and the new building states deduced (Section 6.4). A discussion of all the results is given in Section 6.5, where limitations and alternatives are presented.

6.1 Test Area

In the next paragraphs, the test site is described, considering the buildings' appearances and their surroundings. Furthermore, the acquired InSAR and radargrammetric data are specified, discussing the different incidence angles and baselines. An overview of the building states for the acquisition dates is given. Finally, the different reference data used for the evaluation of the developed algorithm are presented.

6.1.1 Test Site

The choice of the test site was related to the possibility of obtaining both pre- and post-event data. In case of a natural disaster, the pre-event data are the most difficult to obtain, as it is often not possible to determine where the event will take place. At the beginning of this thesis, the tandem mission whose goal is the creation of an accurate global DEM was not completed. Furthermore, the acquisition mode (StripMap) for the global coverage does not permit a detailed analysis of urban areas. Thus, it was necessary to find an area where changes were going to happen in order to command acquisitions with sufficient resolution for both pre- and post-event.
The test site chosen for this work is a part of two adjacent cities in the North East of Paris, Clichy-sous-Bois and Montfermeil. In those cities, a huge urbanization project (*Programme de Rénovation Urbaine - PRU*, urban renewal and rehabilitation program) brings several building

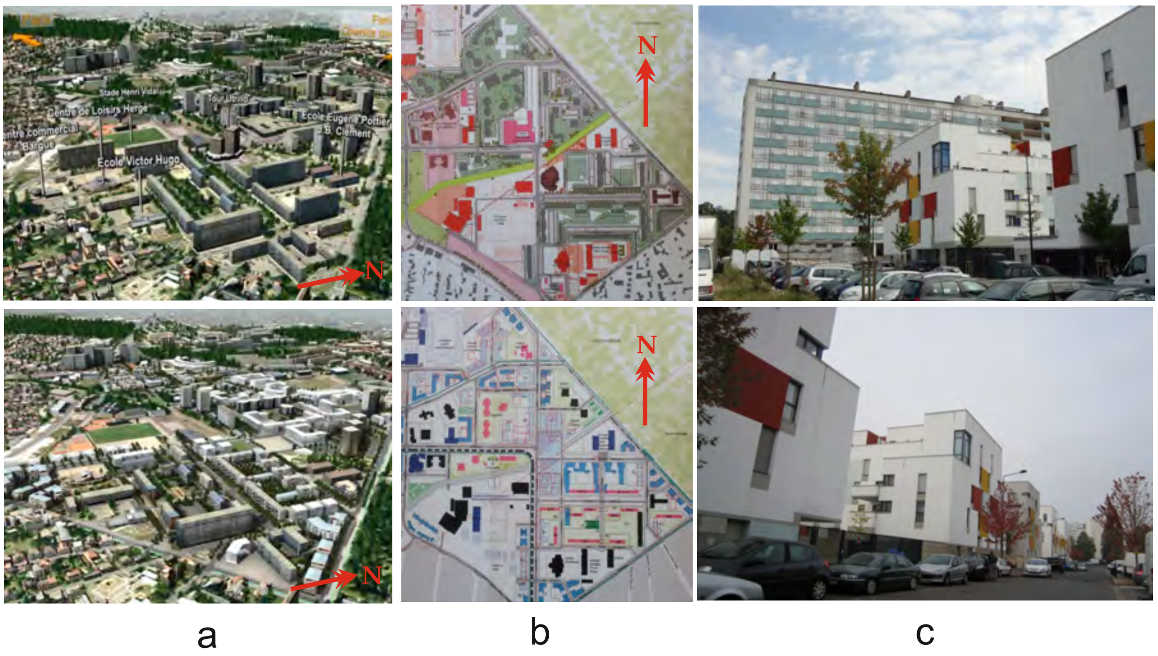


Figure 6.1: 3D and 2D overviews of the test site

demolitions and reconstructions about. The main goal of this project is to reinvigorate neglected suburbs and allow a reboost of the real estate market. In those districts, middle-rise building blocks offer lodging for numerous families. In such blocks, families own their own apartment, but the co-property charges are often too expensive to be paid for, and therefore left apart. This leads to a degradation of the common parts, resulting in precarious, and often insecure living conditions. This project should allow to demolish insalubrious lodging blocks for creating smaller housing, permitting the relocation of the families and possibility for new lodging. The co-property charges should be taken by the cities themselves, in order to keep order. Furthermore, architectural diversity and re-greening are two key aspects of the project that should break the monotony of the previous lodging blocks and allow to boost a real city dynamic and life.

In this work, only the part of building demolition is considered. The first draft of the project was drawn out in 2005 and it should finish in 2018. The main part of the demolitions should have taken place in 2010-2011, with respect to a report produced in 2009 (Consultants 2009). However, some delays due to rehousing of the displaced families postponed the main part of the demolitions between the end of 2012 and the beginning of 2014, parallel to the first constructions and reconstructions.

Figure 6.1 shows a reconstituted overview of the area in 2005 (first row) and a preview of 2018 (second row) in 3D, with corresponding 2D plans. In the latter, the old buildings are represented in grey for 2005, and delimited with red boundaries for 2018. The new constructions are marked in blue, and rehabilitations in pink. Overall, middle- and high-rise building blocks are replaced by more, smaller and colorful blocks (cf. Figure 6.1c), procuring the space for vegetated inner yards.

The planned demolitions are represented Figure 6.2c. A total of 21 demolitions has been planned during the whole project, comprising rectangular middle-rise blocks and octagonal high- and middle-rise lodging towers. Two more demolitions are likely to occur at the end of the project. They are not represented here, as they still did not take place at the time of this work. From

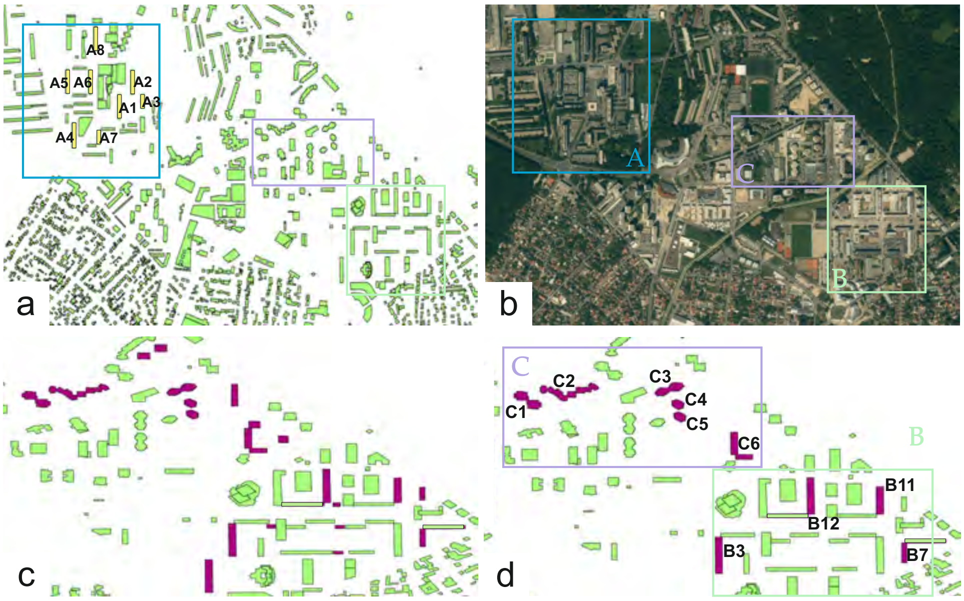


Figure 6.2: Test zones and planned demolition in zones B and C (violet) (source of the optical image: Bing Maps)

the 21 demolitions, already 11 have been fulfilled before the beginning of this work. For those buildings, no pre-event data have been acquired, thus no change detection can be conducted. The left 10 demolitions are represented in Figure 6.2d. Octagonal and rectangular buildings are present to the same amount. An overview of the appearance of the different buildings is given later in this paragraph. Considering the building appearances, the area is separated in this work into two zones B and C, corresponding also to the city delimitations: zone C is situated in Clichy-sous-Bois and zone B in Montfermeil. A further zone A, situated in Clichy-sous-Bois, is defined in this work (Figure 6.2a). In this zone, no demolition and no change occurs. However, the buildings are similar to the rectangular, flat roof buildings of zone B in their shape and dimensions. Furthermore, they are mostly unobstructed, and have the same orientation as the interesting buildings of zone B. They provide an optimal test area for understanding the outcomes of the algorithms of this work, before applying them for change detection, as no construction vehicle or specific disturbing object (crane, fence...) appear. The buildings' names used in the following are inscribed in Figure 6.2(a,d).

Figure 6.3 shows several buildings of the test areas. The first line represents buildings of zone B (respectively B3, B12 and B11) before demolition. Their facade is almost flat, alternating between tight french balconies and mosaic or metal plates, distributed arbitrarily on the facade. A part of the window front has already been removed, inducing a different appearance of the buildings in the SAR images for these areas, as double and single reflections of the signal may occur inside the building. B11 has still some windows, in comparison to B3 and B12, where all windows are removed and shut by wood hangings, or replaced by plastic hangings respectively. For B3, roller shutter are additionally distributed arbitrarily on the facade. Consequently, even if these buildings are similar in shape, dimension and material, their appearance in the SAR images differ a lot from one building to the other, due to the arbitrary state of the facades at the dates of the SAR acquisitions. Building A6 is characteristic for buildings of zone A. As for buildings of zone B, it has a rectangular shape with flat roof. Its facade presents also arbitrary structures,

Figure 6.3: Building appearances in the test area

alternating between windows, french balconies, mosaic plate and roller shutter. The buildings C1 and C2 are represented in Figure 6.3 as well. Both buildings have polygonal shape and present a lot of corners and cranny angles. They are consequently difficult to model and automatically extract from a SAR image. Both have small, regular balconies and their window front is covered by wooden hangings. Finally, for building C3, similar remarks concerning the shape can be made as for buildings C1 and C2: it consists of a polygon presenting a lot of corners and rounded angles. The windows are still in place, and there are small balconies aligned columnwise. The surroundings of the building are highly vegetated, making its recognition in the SAR images difficult. Corresponding SAR images of these areas will be shown in the following paragraphs. As a conclusion, the buildings of zones A and B are very similar, presenting a rectangular shape with flat roof, and very arbitrary facade structure. Those buildings are characterized by a height of about 10 to 15 floors (30 to 35 m), a width of 10 to 15 m and a length of 80 to 100 m. On the contrary, buildings of zone C present a more regular facade structure, with aligned balconies, but very arbitrary geometrical polygonal shapes that are difficult to recognize and automatically extract with the developed method. Moreover, the buildings of zone C are close to another and surrounded by vegetation, which makes their recognition and differentiation even more difficult as for the buildings of zones A and B. In particular, the buildings of zone A are unobstructed and separated by large parking lots, allowing a good recognition. The buildings of zone B are also well separated from another, yet construction machinery and surrounding construction objects may hinder their recognition. Despite their more noisy facade structures, buildings of zones A and B are preferable to analyze and recognize. Besides, the algorithms of this work were developed for rectangular shaped buildings, corresponding of zones A and B. The appearance of the buildings during and after demolition are represented on the last row of Figure 6.3 for different buildings and states. Corresponding to the different buildings states defined in Section 5.2.1,

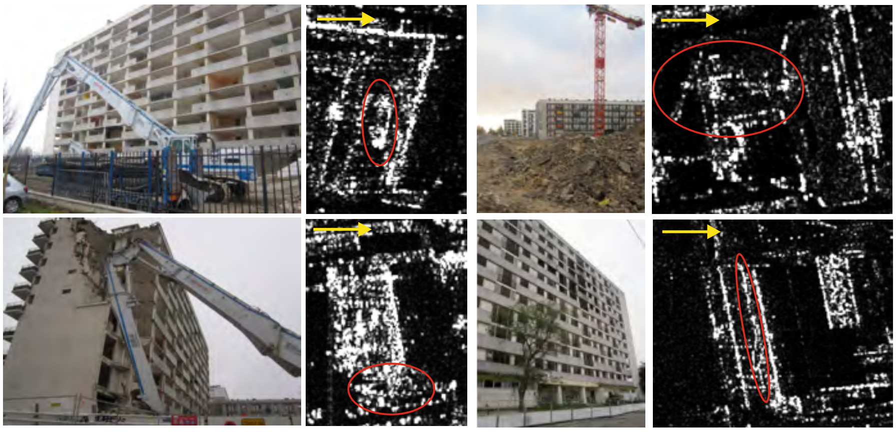


Figure 6.4: Additional challenges from a construction site

each building can have a different appearance in the SAR image, whereby either still a part of the facade is visible, everything is gone, or the remains are obstructed by construction machines and new construction.

The advantages of this test site are multiple. First, planed demolitions and reconstructions are considered, i.e. the monitoring and the subsequent command of SAR data for detecting changes is facilitated. Second, the demolitions happen in a ‚controlled‘ environment. They are restricted to a specific area and the rubbles are kept close to the buildings, so that the accessibility to the sites for ground truth assessment and field recognition is possible. Finally, the test area is situated close to Paris, in a well documented area, which ensure a robust reference (original building heights and dimensions).

In comparison to a site affected by a natural disaster, this test site presents additionally challenges. New objects appear near the buildings as bulldozers, cranes and fences that move between acquisitions and interfere with the building signature in the SAR image. Such objects and their signature are represented in Figure 6.4. Furthermore, due to the delay in the schedule of demolition, new construction of certain buildings happen at the same time as demolitions of other buildings, or directly after the demolition. This involves a very accurate planning of the acquisitions in order to detect the changes due to the demolitions and not to the reconstructions, as at least two acquisitions are required for performing radargrammetry at a specific building state. Finally, even if this is a huge scale urbanization project, only 10 buildings are analyzable during the time of this work, of which five have particular geometrical shapes. This is a very small number of buildings for a robust analysis, yet the work still shows that the developed algorithms can be applied with few data.

6.1.2 Interferometric Data

Table 6.1 gives an overview of the InSAR data and their respective acquisition parameters. Originally, only single-pass data from the tandem satellite mission should have been used, as they permit to get rid off noise due to time decorrelation. However, due to some delay for the acquisition of the science products, the first single-pass acquisition has been completed after three buildings (B7, B11, B12) already changed, as can be seen from Table 6.2. Green stays for ‚still standing‘, yellow for ‚under demolition‘, orange for ‚rubble/cleared‘ and violet for ‚new construc-

Table 6.1: InSAR acquisition parameters

Acquisition Date	Cycle	Orbit	Incidence Angle (°)	Effective Baseline (m)	Height of Ambiguity (m)
24.09/05.10.2012	Repeat-Pass	Descending	29	153	29.2
25.09/06.10.2012	Repeat-Pass	Descending	52	148	65.5
27.09/08.10.2012	Repeat-Pass	Ascending	36	532	10.8
30.09/11.10.2012	Repeat-Pass	Descending	42	31	230.9
02.10/13.10.2012	Repeat-Pass	Ascending	47	314	27.8
03.10/14.10.2012	Repeat-Pass	Ascending	22	381	8.2
04.11.2012	Single-Pass	Ascending	47	234	35.9
05.11.2012	Single-Pass	Ascending	22	186	17.1
20.11.2012	Single-Pass	Ascending	56	199	55.5
01.04.2013	Single-Pass	Ascending	56	172	64.7
10.05.2013	Single-Pass	Ascending	47	143	59.1
01.06.2013	Single-Pass	Ascending	47	151	55.9

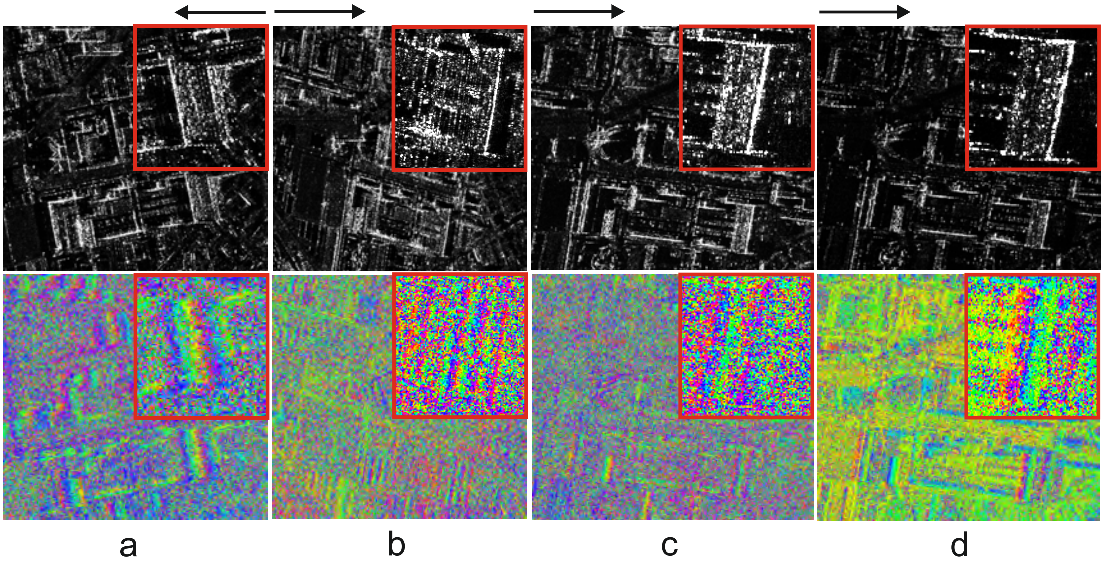


Figure 6.5: Examples of available InSAR data; first row: amplitude images; second row: corresponding phase images; (a) repeat-pass descending, 29°, $B_\perp = 153$ m, $h_a = 29.2$ m; (b) repeat-pass ascending, 22°, $B_\perp = 381$ m, $h_a = 8.2$ m ; (c) repeat-pass ascending, 47°, $B_\perp = 314$ m, $h_a = 27.8$ m ; (d) single-pass ascending, 47°, $B_\perp = 234$ m, $h_a = 35.9$ m

tion‘. Therefore, also repeat-pass data, acquired before the single-pass data and before building changes, have been used in this work. In order to analyze which interferometric configurations are the most suitable for building detection, different baselines and incidence angles are considered, leading to different height of ambiguity for the InSAR phase images.

For the repeat-pass data, three descending and three ascending configurations have been acquired, spanning an interval of about 25° of the incidence angles for each side. Besides, the effective baselines cover a very wide range, from a few tens of meters to a few hundreds of meters, leading to corresponding discrepancies of the height of ambiguities. A few examples are given in Figure 6.5, whereby it becomes obvious that the effective baseline influences the building appearance in the InSAR phase image. In figure 6.5b, the effective baseline is large, yielding a very small height of ambiguity. This results in multiple fringe patterns at building location. Furthermore, due to the very steep incidence angle, building layovers of adjacent building overlap, making the scene interpretation and the detection of single buildings difficult. A similar inci-

Table 6.2: InSAR building states - Green: ‚still standing‘; yellow: ‚under demolition‘; orange: ‚rubble/cleared‘; violet: ‚new construction‘

Acquisition Date	B3	B7	B11	B12	C1	C2	C3	C4	C5	C6
24.09/05.10.2012	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨
25.09/06.10.2012	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟨
27.09/08.10.2012	🟩	🟨	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟨
30.09/11.10.2012	🟩	🟨	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟨
02.10/13.10.2012	🟩	🟨	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟨
03.10/14.10.2012	🟩	🟨	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟨
04.11.2012	🟩	🟧	🟨	🟨	🟩	🟩	🟩	🟩	🟩	🟨
20.11.2012	🟩	🟧	🟨	🟧	🟩	🟩	🟨	🟨	🟨	🟧
01.04.2013	🟨	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
10.05.2013	🟧	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧

dence angle with a shorter baseline provides a larger fringe pattern in the InSAR phase image and can be interpreted more easily (Figure 6.5a).

For the single-pass data, only six ascending acquisitions are available, covering an interval of incidence angles slightly larger as for the repeat-pass data (over 30°). Yet, only three different incidence angles are available, of which only one differs from the repeat-pass data (56°). The provided baselines and subsequent height of ambiguities are all in a range enabling a good interpretation of the InSAR phase data at building location. Figure 6.5 shows the difference between repeat-pass (Figure 6.5c) and single-pass (Figure 6.5d) acquisitions of the same area under similar conditions (same incidence angle, 47°). A zoom onto the same building is represented for each dataset in the area delimitated in red. In both data, the characteristic layover fringes are recognizable, yet the repeat-pass data are more difficult to interpret because of the higher noise caused by time decorrelation. In the single-pass data, fixed urban objects smaller than the studied buildings are still well distinguishable from the ground.

Table 6.2 shows the states of the different buildings at the time of the different InSAR acquisitions. The building C6 is under demolition from the beginning of this work, thus the original parameters would correspond to the building extent under demolition at the time of acquisition. Only the first repeat-pass acquisition permits the determination of the parameters of all other buildings, as they are still standing. For the second repeat-pass acquisition, B7 has changed, and also here the calculated parameters by InSAR processing would correspond to the building extent under demolition. For the other repeat-pass acquisitions, B12 is under demolition, having similar consequences on the building parameter determination by interferometry. For the single-pass data acquired one month later, changes already occurred, except for the buildings B3, C1 and C2 that are still standing and whose original parameters can be determined. Finally, the InSAR data acquired in 2013 are not of high interest for determining the original building parameters, as most of them have already been demolished.

This is the reason why only the acquisitions marked in dark grey in Table 6.1 could be taken into account for determining pre-event building parameters. Depending on the considered building, more than one acquisition can be considered. The remaining acquisitions marked in light grey are only considered for assessing the suitability of the developed building detectors in the phase image for different configurations of acquisitions. It will be shown in Section 6.2.1 that acquisitions with an incidence angle of 22° are not suitable for building detection, therefore they will not be further considered for building extraction. Furthermore, the three single-pass data

Table 6.3: Same-side configurations

	Same-Side Configurations	
	Ascending	**Descending**
From single-pass InSAR	56°/47°	
	56°/22°	
From repeat-pass InSAR and radargrammetric acquisitions	47°/36°	52°/42°
	47°/22°	52°/29°
	36°/22°	42°/29°

acquired in 2013 are not suitable for determining the pre-event building parameters, as most of the buildings are already demolished at this time. Just two of them (01.04.2013 and 10.05.2013) are considered for the analysis of the building detectors using the zone A where no change occurs. They enable together with the previous InSAR data of 2012 the analysis of the parameter setting of the detectors under same incidence angle but different baseline configurations. After detector analysis (Section 6.2.1), only configurations permitting a reliable building detection are retained for further processing.
Useful for the recognition of the number of facades (Section 3.5), the range pixel spacing of those data is about 0.45 m in range and 0.86 m in azimuth direction. The exact spacing of each dataset is considered during processing.

In order to show the applicability of the developed algorithm for building detection and parameter extraction, another dataset representing a subset of the city center of Berlin is used. This is a single-pass interferogram acquired on 04.01.2012 in descending orbit, with 42° incidence angle, an effective baseline of 108.1 m and height of ambiguity of 65.4 m. It should assess the ability of the algorithm to detect and reconstruct single facades in the case where two facades of a building are visible. Indeed, in the Paris area, the orientation and extent of the buildings does not permit such an analysis.

6.1.3 Radargrammetric Data

Table 6.3 resumes the different radargrammetric configurations that can be built with the obtained radargrammetric data. Here, only same-side acquisitions are considered, as opposite-side configurations are not optimal in urban areas because they represent different building sides (Section 4.1). All the acquired data suitable for radargrammetric processing are listed in Table 6.4, showing the corresponding building states at the different acquisitions' dates. Here also, green stays for ‚still standing‘, yellow for ‚under demolition‘, orange for ‚rubble/cleared‘ and violet for ‚new construction‘. Some of the data used for the InSAR processing can be used for the radargrammetric processing, considering an earlier interferometric acquisition for pre-event analysis. There, only the magnitude image of one of the two interferometric acquisitions is used. Considering only the radargrammetric acquisitions capable of showing much changes, they were acquired over four periods of time: October 2012, January 2013, March 2013 and October 2014, permitting to analyze several changes of the buildings. For example, two states of change should be detectable for building B3: ‚under demolition‘ (March 2013) and ‚rubble/cleared‘ (October 2014).

Table 6.4: Radargrammetric building states - Green: ‚still standing‘; yellow: ‚under demolition‘; orange: ‚rubble/cleared‘; violet: ‚new construction‘

Acquisition Date	Orbit	Incidence Angle (°)	Heading Angle (°)	**B3**	**B7**	**B11**	**B12**	**C1**	**C2**	**C3**	**C4**	**C5**	**C6**
08.10.2012	Ascending	36	349.6	🟩	🟨	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟨
13.10.2012	Ascending	47	351.3	🟩	🟨	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟨
14.10.2012	Ascending	22	347.9	🟩	🟨	🟩	🟨	🟩	🟩	🟩	🟩	🟩	🟨
04.11.2012	Ascending	47	351.2	🟩	🟧	🟨	🟨	🟩	🟩	🟩	🟩	🟩	🟨
05.11.2012	Ascending	22	347.9	🟩	🟧	🟨	🟨	🟩	🟩	🟩	🟩	🟩	🟨
20.11.2012	Ascending	56	352.9	🟩	🟧	🟨	🟧	🟩	🟩	🟨	🟨	🟨	🟧
15.01.2013	Ascending	36	349.6	🟩	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
18.01.2013	Descending	42	189.5	🟩	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
20.01.2013	Ascending	47	351.3	🟩	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
23.01.2013	Descending	29	191.2	🟩	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
24.01.2013	Descending	52	187.9	🟩	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
01.02.2013	Ascending	22	347.9	🟩	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
19.03.2013	Descending	29	191.2	🟩	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
20.03.2013	Descending	52	187.9	🟩	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
22.03.2013	Ascending	36	349.6	🟩	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
25.03.2013	Descending	42	189.5	🟨	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
27.03.2013	Ascending	47	351.3	🟨	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
28.03.2013	Ascending	22	347.9	🟨	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
01.04.2013	Ascending	56	352.9	🟨	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
10.05.2013	Ascending	47	351.2	🟧	🟧	🟧	🟪	🟩	🟩	🟧	🟧	🟧	🟧
20.10.2014	Ascending	47	351.3	🟧	🟧	🟧	🟪	🟪	🟪	🟪	🟪	🟪	🟧
21.10.2014	Ascending	22	347.9	🟧	🟧	🟧	🟪	🟪	🟪	🟪	🟪	🟪	🟧
23.10.2014	Descending	29	191.2	🟧	🟧	🟧	🟪	🟪	🟪	🟪	🟪	🟪	🟧
24.10.2014	Descending	52	187.9	🟧	🟧	🟧	🟪	🟪	🟪	🟪	🟪	🟪	🟧
26.10.2014	Ascending	36	349.6	🟧	🟧	🟧	🟪	🟪	🟪	🟪	🟪	🟪	🟧
29.10.2014	Descending	42	189.5	🟧	🟧	🟧	🟪	🟪	🟪	🟪	🟪	🟪	🟧

For each period of acquisition, six acquisitions are available: three descending and three ascending, each presenting a different incidence angle (same as the available angles for the repeat-pass interferometry). These represent for each side three possible configurations, resumed in Table 6.3. As those configurations are available for each period, a thorough analysis of the best radargrammetric configuration for building extraction can be performed. Taking into account the amplitude images coming from the single-pass interferometric data, two more configurations are available, whereby 56°/22° is critical due to the very large intersection angle that provides low radiometric similarity.

Each of the presented same-side configurations can be obtained in a very short time span with TerraSAR-X and/or TanDEM-X: at least, one day is needed ($52°/29°, 47°/21°$), and at most five days ($52°/42°, 47°/36°$). These time spans are short enough for change detection analysis, especially because other sensor systems (e.g. COSMO-Skymed) allow an even shorter repetition cycle. In this work, changes occur on the buildings at the same time as the acquisitions. Thus, only specific configurations are available for specific building states. In particular, for the period of March 2013, one descending configuration can assess that building B3 is still standing ($52°/29°$), and one ascending configuration can already show some changes ($47°/22°$). Taking into account the amplitude image of the interferometric acquisition of 01.04.2013, an additional configuration could be possible ($56°/47°$) (cf. Table 6.4). However, as an incidence angle of 56° is not a standard product, it is not considered for later processing. Table 6.5 resumes for each acquisition period which building or zone undergoes what change. It here becomes obvious

Table 6.5: Evolution of the building states with the acquisition period

Period	**No Change from InSAR**	**Under Change**	**Rubbles/Cleared**
October 2012	B3, B7, B11, zone C	B12 (Ascending)	
November 2012	B3, zone C	B11 (Ascending), B12 (Ascending)	B7 (Ascending)
January 2013	B3, C1, C2		B7, B11, C3, C4, C5, C6
March/April 2013	B3 (Descending)	B3 (Ascending)	B7, B11, C3, C4, C5, C6
October 2014			B3, B7, B11, C6

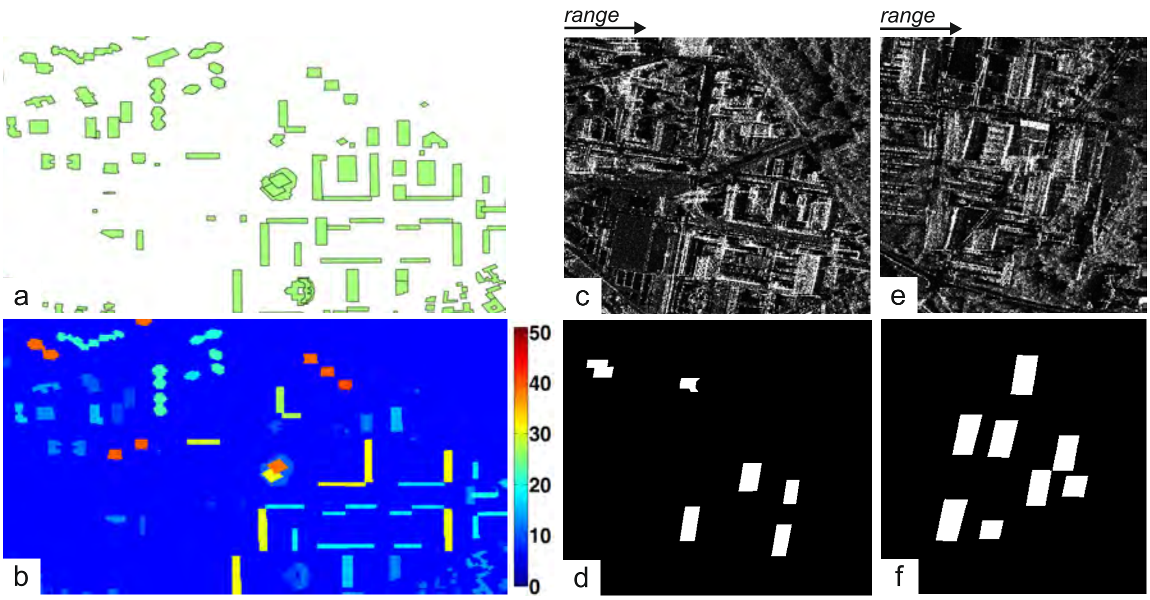


Figure 6.6: Ground truth data for pre-event analysis; (a) BDTopo©; (b) rasterized BDTopo© with respect to the building heights (m); (c) amplitude image of zone B with 36° ascending; (d) manually extracted corresponding layovers of the buildings of interest; (e) amplitude image of zone A with 36° ascending; (f) manually extracted corresponding layovers of the buildings of interest

that only the state ‚rubble/cleared' can be detected for the buildings of zone C, as only one acquisition has been achieved during demolition. Furthermore, for C1 and C2, no change can be detected, as new constructions are not considered in this work.

6.1.4 Reference Data

The advantage of choosing a test area in a very urbanized zone near a capital is that many reliable ground truth data already exist, especially for the pre-event analysis. Different reference data are used for the assessment of pre-event building reconstruction and the post-event change detection.

For the pre-event analysis, GIS data provided by the IGN (*Institut National de l'information Geographique et forestière*- national institute of geographical and forest information) are used. The database BDTopo© is used, permitting a geocoded vector representation of the buildings. The building lengths, widths and orientations can be directly measured on the data, and their height is given as attribute. The BDTopo© database is created based on the fusion of cadastral information, optical imagery and interpolated DEM, leading to different precision levels depending on the considered buildings, specified in the attribute table of each building. Planimetric and altimetric accuracies are given, whereby the altimetric accuracy corresponds to the absolute

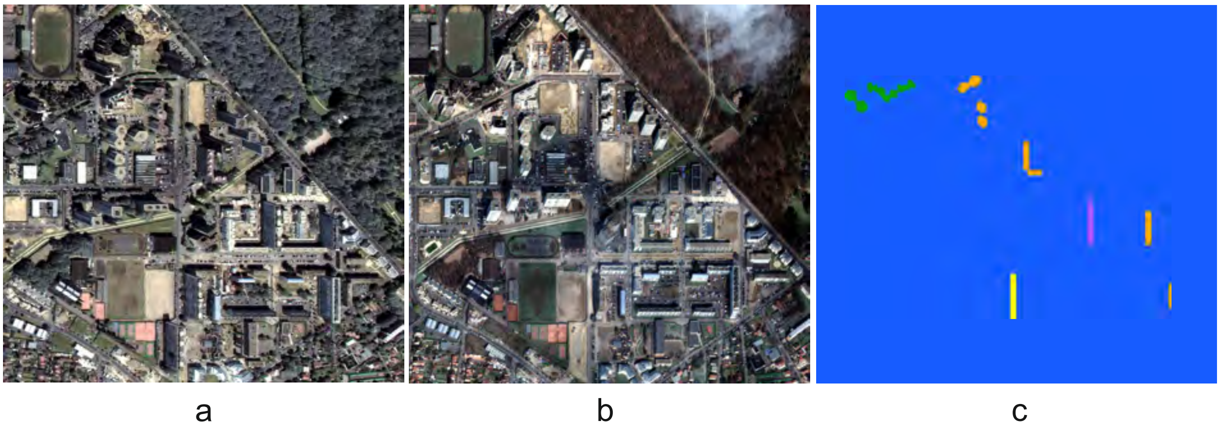


Figure 6.7: Ground truth data for post-event analysis: WorldView 2 data; (a) archive image - 29.06.2010; (b) new image - 01.04.2013; (c) manual classification of the building states for the buildings of interest - different color corresponds to different building states - Green: ‚still standing‘; yellow: ‚under demolition‘; orange: ‚rubble/cleared‘; violet: ‚new construction‘

height accuracy of the building footprint. All buildings situated in zone B have a planimetric precision of 1.5 to 2.5 m, and an altimetric precision of 1 to 2.5 m, except B11, whose precision specifications are even better (0.5 to 1.5 m in planimetry and less than 1 m in altimetry). The buildings of zone C have a similar planimetric accuracy, but the altimetric accuracy degrades to 20 m. However this specification applies only on absolute height of the building footprint. A french architecture database PSS© provides for those buildings altimetric information, permitting to assess the relative buildings' heights given by the BDTopo©. This data is represented in Figure 6.6(a,b), whereby Figure 6.6b represents the rasterised data, where the attribute ‚building height‘ is considered for the colored representation.

As the buildings' layovers are detected in order to retrieve the building parameters, layover ground truth has been produced manually in this work in order to assess the completeness and correctness of the extracted facades before parameter determination. Parallelograms corresponding to the layover areas have been drawn out for the zones of interest, using the amplitude and phase images of the InSAR data for different incidence angles. An overview for one incidence angle is given in Figure 6.6(d,f) with the corresponding amplitude images in Figure 6.6(c,e) for zones B and A, respectively. Ground truth for zone C was more difficult to obtain, as the buildings are barely recognizable for an operator in the InSAR data.

For the change detection analysis, two different kinds of data are available, permitting to estimate if a change occurred. An accurate determination of the building extent is not possible anymore, but the determination of which change occurred (‚still standing‘, ‚under demolition‘, ‚rubbles/cleared‘) and its categorization into one of the building classes defined in Section 5.2.1 is possible.

First, two multispectral satellite images from WorldView 2 (WV2) are available. The first image has beeen acquired on 29.06.2010, when the buildings were still standing. The second image has been acquired during demolition, on 01.04.2013. The exact building states at this date can be read from Table 6.4, as one single-pass InSAR acquisition has been performed at the same date. Some changes occurred after the acquisition date, but most of the changes can already be observed. Both multispectral images have a resolution of 2 m in the multispectral channels

Date	Type
01-02.08.2012	Field Inspection
16.10.2012	Pictures from Construction Supervisor
04.11.2012	Field Inspection
26.12.2012	Field Inspection
31.01.2013	Field Inspection
21-22.03.2013	Field Inspection
26.20.2014	Field Inspection

Figure 6.8: Ground truth data for post-event analysis: field inspections

and 0.5 m in the pancromatic. A difference map has been built manually in order to depict the changes. They were mapped using the ground truth provided by the BDTopo©. These reference data are shown in Figure 6.7.

Second, regular field inspection and close contact with the construction supervisors enabled the acquisition of several pictures showing the evolution of the demolitions at the different periods of SAR acquisitions. In particular, these pictures and field inspection permitted the creation of Tables 6.2, 6.4 and 6.5 showing the different building states at the SAR acquisition dates. Some of the pictures are shown in Figure 6.8, and the accompanying table depicts the dates of the field inspections.

6.2 Results of Interferometric Processing

In this section, important results of the interferometric processing are presented. First, the influence of different parameters on the detector results is shown. Second, results and accuracy of the layover extraction are determined, using the most suitable detector parameters. Finally, results of building parameter determination are given and their accuracy is discussed.

6.2.1 Parameter Setting of the Detectors

In section 3.4, the importance of the chosen synthetic phase ramp for both detectors has been explained. First test were performed and discussed in (Dubois et al. 2015) for the phase ramp detector, showing that the preferable ramp length L_{SR} highly depends on the considered building height. In this work, the analysis of the optimal ramp lengths L_{SR} is extended for both detectors on several acquisition configurations. As for the first analysis, parallelograms corresponding to the layover extents were created manually for each different configuration. The optimal ramp length corresponds to the length L_{SR} for which the maximum difference between the mean value of the detector results within the parallelogram and its direct surrounding is achieved, i.e. better contrast (cf. Section 3.4.3). Considering several buildings of the same height or length for a specific configuration, it is possible to determine a mean optimal ramp length L_{SR} for a specific height or length for this configuration. Table 6.6 depicts the obtained mean optimal ramp length

Table 6.6: Influence of the acquisition parameters on the parameter setting - Phase Ramp Detector

Building Height (m)	Incidence Angle (°)	Cycle / Orbit	Height of Ambiguity (m)	Effective Baseline (m)	Mean Maximal Difference	Mean Ramp Length L_{SR} (pix)	Standard Deviation Ramp Length (pix)
21	47	RP asc.	27.8	314	0.005	16	-
		SP asc.	35.9	234	0.044	19	-
33	22	RP asc.	8.2	381	0.024	3	1.9
	29	RP desc.	29.2	153	0.119	23	2.6
	36	RP asc.	10.8	532	0.086	9	0.7
	42	RP desc.	230.9	31	0.221	17	4.7
	47	RP asc.	27.8	314	0.080	17	1.1
		SP asc.	35.9	234	0.170	20	1.6
		SP asc.	59.1	143	0.249	33	2.9
	52	RP desc.	65.5	148	0.249	17	4.7
	56	SP asc.	55.5	199	0.227	24	2.2
		SP asc.	64.7	172	0.224	27	2.7
44	47	RP asc.	27.8	314	0.173	20	-
		SP asc.	35.9	234	0.165	25	-
		SP asc.	59.1	143	0.271	35	-

L_{SR} for the phase ramp detector, for different building heights and configurations. Three main building heights were analyzed: 21 m, 33 m and 44 m. An exhaustive analysis is given for the building heights of 33 m, as it corresponds to the average height of most buildings of zone A and zone B. A reliable average calculation is possible over all the buildings. The standard deviation of the determined optimal ramp length L_{SR} is calculated over all buildings of height close to 33 m. Depending on the demolished buildings at the different acquisition times, around 10 buildings are available for each acquisition. The standard deviation of the reference heights of those buildings is around $\sigma_h = 1$ m, changing only slightly depending on the considered acquisition and available reference building. For the other building heights, no standard deviation can be estimated, as 21 m and 44 m corresponds to single isolated buildings.

In general, considering for each height the same configuration, it is observable that higher buildings are better detected with longer ramps L_{SR}. This trend was already described thoroughly in (Dubois et al. 2015) and is not further developed in this work. Considering a specific building height, it is interesting to analyze the influence of the different configurations. The results of Table 6.6 are represented graphically in Figure 6.9 for the building height of 33 m. The discrete results are connected with simple lines in order to detect some trends. The optimal phase ramp L_{SR} was tested in 2 pixel steps for each building and each configuration. A first observation indicates that the standard deviation of the optimal ramp length for all considered buildings is low, always less than 5 pixel. Recognizing that the reference height itself has a standard deviation $\sigma_h = 1$ m, the optimal ramp length is stable for each configuration.

Considering the baseline parameter, it is observable that the maximum difference of the detector results between building and surroundings, i.e. the best contrast, is obtained for baselines between 150 m and 250 m, yielding phase ramps between 17 pixel and 33 pixel. Really long baselines provide very small optimal phase ramps lengths of less than 10 pixel, and the contrast between layover and surrounding is very low (blue line in Figure 6.9). Figure 6.10(b,c) shows two obtained difference images from the phase ramp detector -second row- for two different baselines. The optimal phase ramp L_{SR} of Table 6.6 has been considered in each case. The corresponding

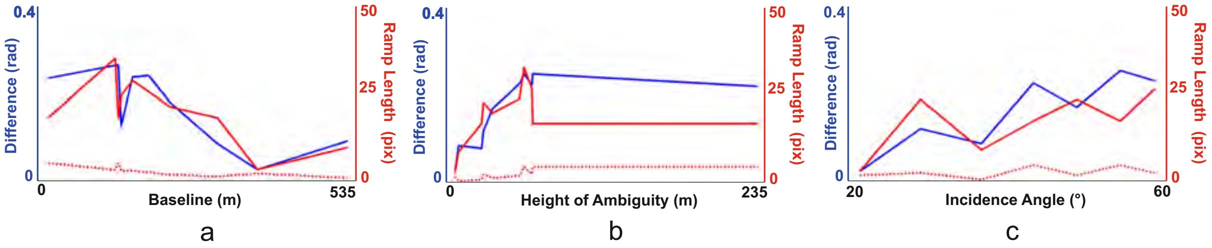


Figure 6.9: Graphical representation of the influence of different parameters on the optimal ramp length L_{SR} for the phase ramp detector and building height of 33 m; (a) baseline; (b) height of ambiguity; (c) incidence angle

phase images are depicted in the first row. It is obvious that the layovers are better recognizable with the smallest baseline.

Opposite observations can be made for the height of ambiguity. The available heights of ambiguity in this work are principally between 8 m and 65 m, except for one that reaches 230 m. Higher differences between building and surroundings are obtained for heights of ambiguity higher than 30 m, yielding an optimal ramp length L_{SR} higher than 17 pixel. Even if phase unwrapping is considered within the phase ramp detector, small heights of ambiguity seem to lead to very small optimal ramp lengths, with very low contrast between building and surroundings, as shown in Figure 6.10a. An interesting result is for the very long height of ambiguity, represented in Figure 6.10d. Even if the phase ramp, i.e. the fringe pattern, is hardly discernible in the phase image, the detector manages to recognize it.

Finally, considering the influence of the incidence angle on the optimal ramp length L_{SR} for a specific building height, no particular trend can be observed. For the incidence angles where more than one configuration is available (47° and 56°), an average of the results has been built. Only very steep incidence angles (22°) produce very bad contrast between building and surroundings (Figure 6.10a). Considering only the punctual results of Figure 6.9c, it could be possible to fit a line with a slightly ascending trend towards higher incidence angles, for both difference and optimal ramp length. Using larger incidence angles seems to provide a better contrast and yield a longer optimal ramp length. However, this observation has to be considered carefully, as with higher incidence angles, the layover area becomes smaller, and smaller ramp lengths would be expected.

Similar results and a comparison are shown for the constant phase detector in Table 6.7 and Figure 6.11. Instead of different building heights, which have no influence on the results, different building lengths of the test area have been considered. Similarly to the phase ramp detector, only the building length 90 m is analyzed here thoroughly, as it corresponds to the length of many buildings of the area and allows a robust average calculation of the optimal ramp length L_{SR} over all buildings. The building with a length of 70 m is only visible in one configuration, as it has been demolished and consequently does not appear in the other configurations.

Generally, the mean differences between building layover and building surroundings are for most configurations much smaller than with the phase ramp detector. Two main reasons exist: first, as the phase noise is considered uniform, the average difference between a ground pixel of the building surroundings and its neighbors will be near zero. Second, constant phase trends can also be observed in building surroundings, e.g. roads.

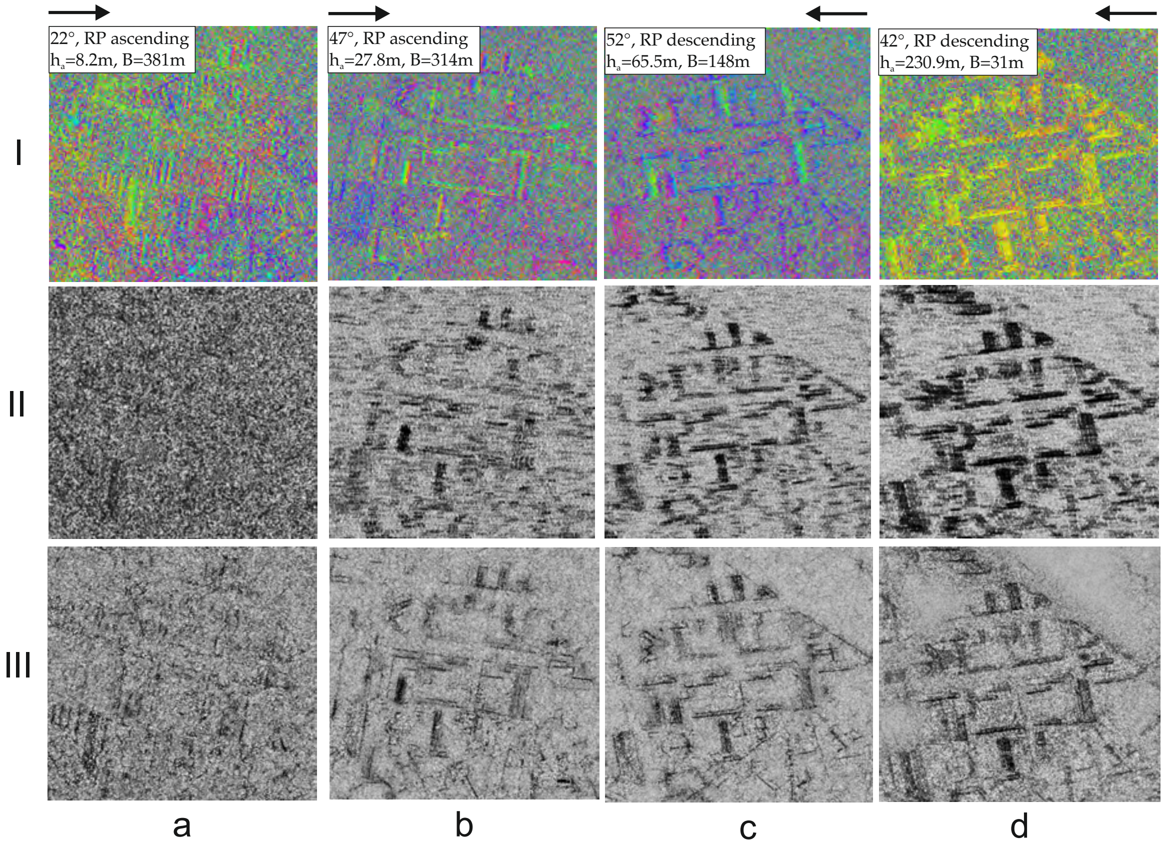


Figure 6.10: Detector results for different configurations; (I) phase images; (II) phase ramp detector; (III) constant phase detector

Furthermore, contrary to the phase ramp detector for which the optimal phase ramp is dependent on the building height, the optimal segment length L_{SR} of the constant phase detector does not seem to depend on the building length. Except from some configurations for which a particular optimal length L_{SR} is achieved, most optimal lengths range between 15 pixel and 25 pixel, for the three presented building lengths.

Figure 6.11 shows a graphical representation of the results for a building length equal to 90 m, corresponding to most buildings of the test area. Here also, the influence of baseline and incidence angle is analyzed. The constant phase detector is independent of the building height, i.e. also of the height of ambiguity. Instead, it should depend on the building orientation. In the test area, most buildings of 90 m have the same North-South orientation, so that an analysis of the influence of the building orientation on the parameter setting cannot be performed.

As for the phase ramp detector, baselines between 150 m and 250 m seem to provide the better contrast between layover and surroundings, and achieve the smallest standard deviation of the optimal segment length L_{SR}. An optimal segment length of $20-25$ pixel is achieved for those baselines. A descending trend can be observed in the difference values by increasing the baseline, showing that longer baselines are not suitable for discerning layover from the surroundings.

Having a look at the incidence angle, an optimal segment length of 74 pixel is obtained for 42°. Yet, the standard deviation is also very high (26.2 pixel), so that this result should be considered carefully. Apart from this configuration, a general slightly ascending trend can be observed for both difference and segment length, as for the phase ramp detector.

Table 6.7: Influence of the acquisition parameters on the parameter setting - Constant Phase Detector

Building Length (m)	Incidence Angle (°)	Effective Baseline (m)	Mean Maximal Difference	Mean Ramp Length (pix)	Standard Deviation Ramp Length (pix)
50	47	314	0.072	20	7.1
		234	0.079	20	7.1
		143	0.072	20	0
70	47	314	0.018	15	-
90	22	381	0.027	9	1.9
	29	153	0.049	26	7.5
	36	532	0.038	13	6.1
	42	31	0.095	74	26.2
	47	314	0.040	18	3.7
		234	0.098	22	8.1
		143	0.073	27	7.6
	52	148	0.095	24	1.9
	56	199	0.117	23	4.1
		172	0.113	23	2.7

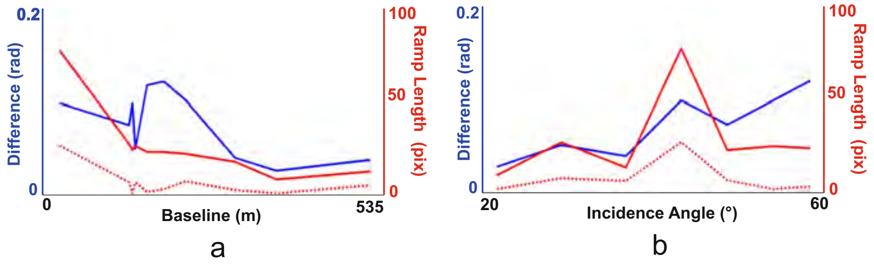


Figure 6.11: Graphical representation of the influence of different parameters on the optimal ramp length L_{SR} for the constant ramp detector for a building of 90 m length

Considering now both detector results, the best suited configurations and ramp/segment lengths can be used for further processing. It is observable that both single-pass and repeat-pass data provide satisfying results for specific configurations. Due to less decorrelation of the single-pass data, higher differences between buildings and surroundings or a more robust estimation of L_{SR} would have been expected for those data. Instead, both data types show similar results and can be used. This is of great interest for this work, as the single-pass data has been acquired when some buildings were already demolished, whereas the repeat-pass configurations present the original building states. Besides, it shows the robustness of the developed detectors against time induced decorrelation.

The configurations marked in grey in Tables 6.6 and 6.7 are considered for the following processing. These four configurations permit the analysis of single-pass and repeat-pass data for further processing, and also a comparison of the building extraction from ascending and descending configurations. For the reasons explained previously, configurations with a too short height of ambiguity are left aside ($h_a = 8.2$ m and $h_a = 10.8$ m) as they do not provide a robust distinction of the buildings from the surroundings. The configuration of 42° incidence angle is taken into consideration, even if the results of the azimuth detector show a high standard deviation. A visual observation of the results in Figure 6.10d reveals that both detectors allow to recognize building layovers. From the configuration presenting 47° incidence angle, only two datasets are retained, one single-pass and one repeat-pass. The single-pass acquisition with a

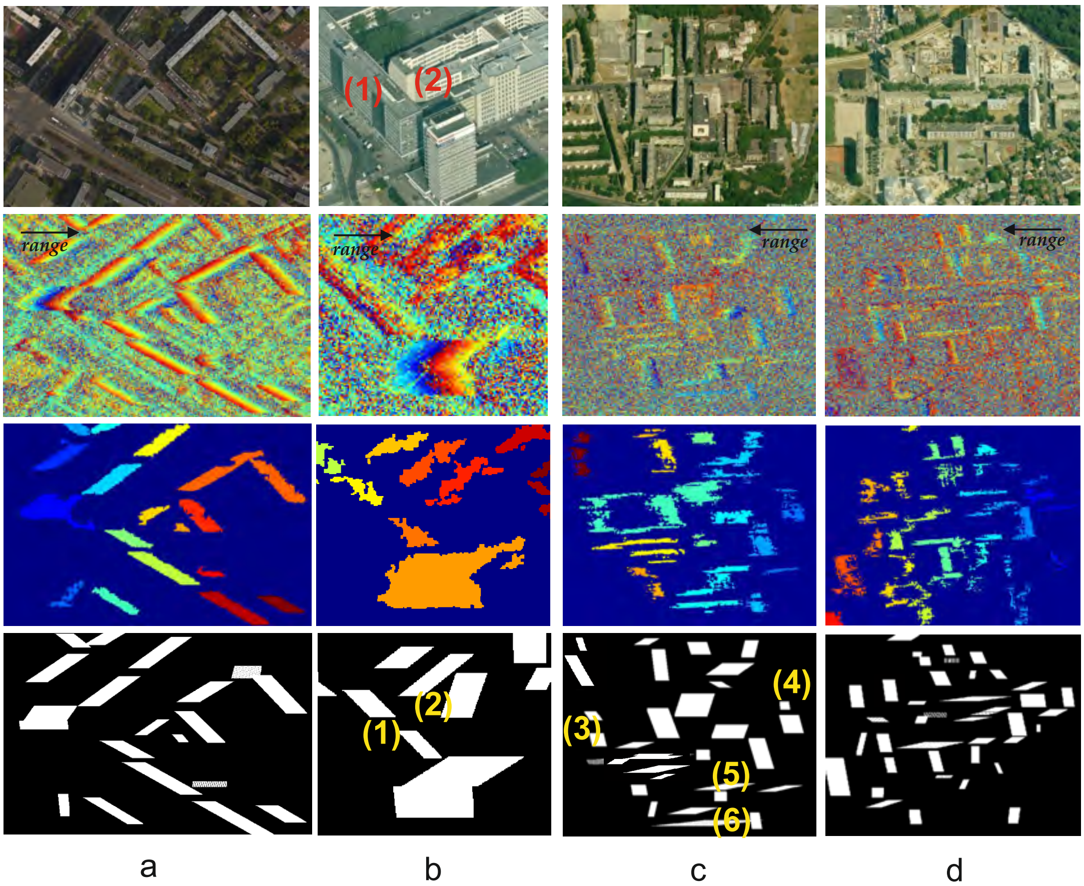


Figure 6.12: Results of Building Hypothesis Segmentation and subsequent Parallelogram Fitting for (a) and (b): two subsets of the Berlin dataset; (c) Paris zone A for 52° (d) Paris zone B for 52° - (source of the optical image: Bing Maps)

height of ambiguity of 35.9 m is preferred, as it approximately corresponds to the mean building height of the buildings of interest. The configuration with 52° incidence angle is retained as it shows satisfying results for both detectors and allows to have another descending dataset. The configurations acquired in single-pass with 56° incidence angle are left apart, as they have been acquired as most buildings of interest were already demolished. For the four considered configurations, the optimal estimated ramp/segment length is set for each detector.

6.2.2 Results of Layover Extraction

Figure 6.12 shows the results of the segmentation of the building hypothesis (third row) and subsequent parallelogram fitting (fourth row) for two different datasets. Figure 6.12(a,b) represent two different zones of the city center of Berlin, where two-facades buildings are observable. Figure 6.12(c,d) represent respectively zone A and zone B of Paris, with the repeat-pass acquisition of incidence angle 52°. Having a look at the corresponding phase images (second row), it is obvious that the building layovers of Berlin are better recognizable than those of the Paris dataset, which are more noisy. However, having a look at the third row of Figure 6.12, the building hypotheses have been correctly segmented for both datasets. Furthermore, two-facades buildings were well recognized and corresponding single facades reconstructed. In Figure 6.12a,

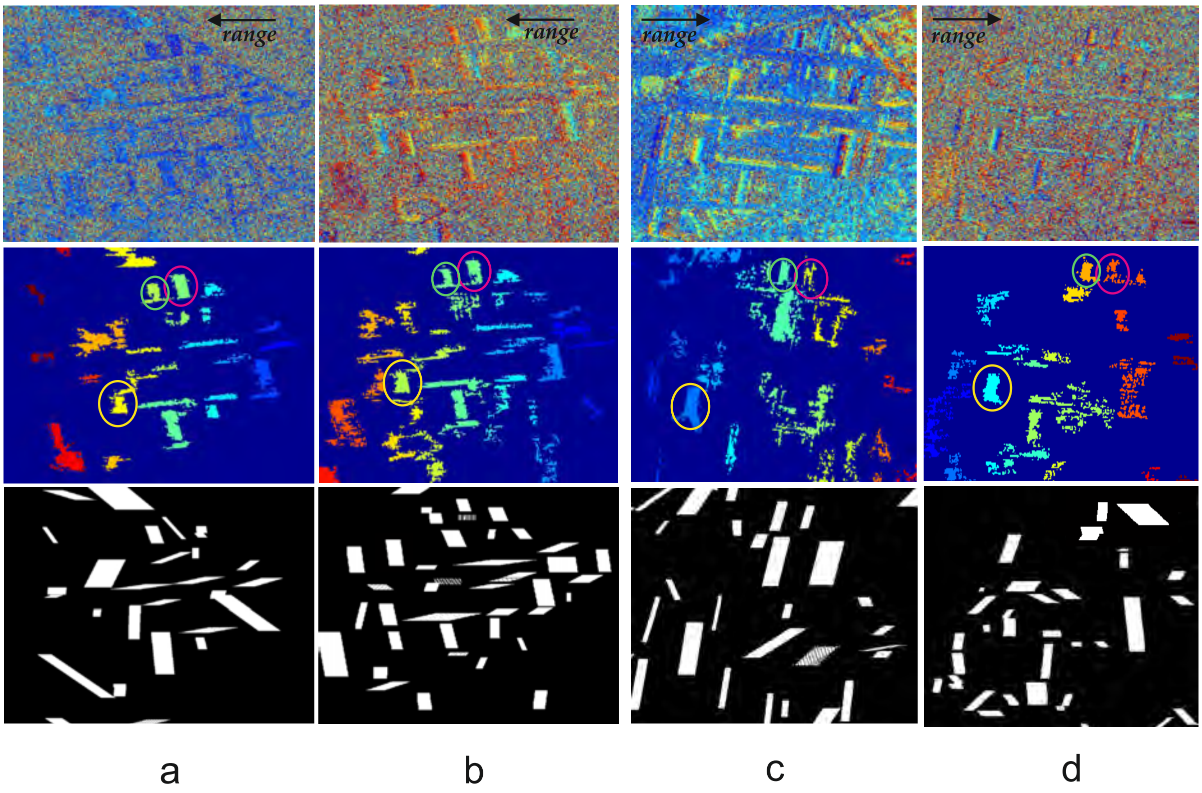
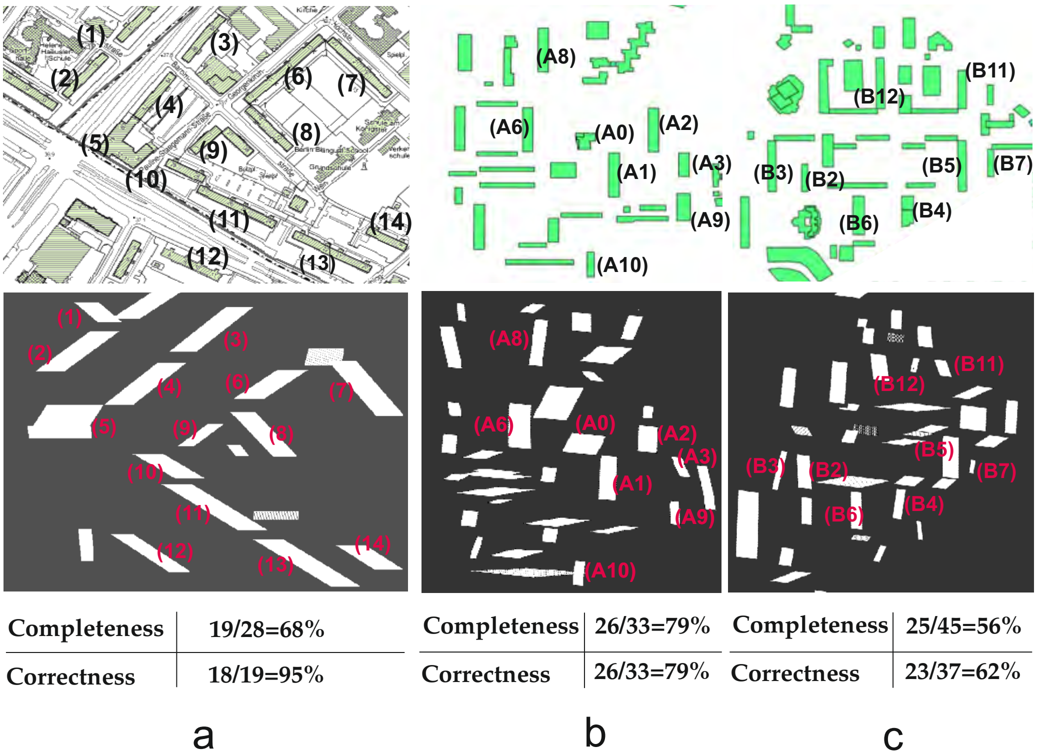


Figure 6.13: Results of building hypothesis segmentation and subsequent parallelogram fitting for zone B with the four different datasets: (a) 42° descending; (b) 52° descending; (c) 47° ascending single-pass; (d) 47° ascending repeat-pass

the segments are very dense and have sharp borders, leading to almost perfectly fitted parallelograms, as visible in the fourth row of Figure 6.12a. In Figure 6.12b, two buildings (1) and (2) have been separated into several segments, leading to several fitted parallelograms. This can be explained by the noisier phases at these building locations, leading to poorer detector results. A closer look at the corresponding optical image shows that metallic structures are situated on and in front of building (1), destructing the layover appearance. Furthermore, building (2) corresponds to a complex building with several inner yards. For the Paris dataset, the segments generally appear more disconnected than for the Berlin dataset. As explained previously, both detectors perform similarly on repeat-pass and single-pass data. This sparseness is mainly due to the appearance of the buildings in Paris, presenting irregular facade structures compared to the buildings of Berlin, which may hinder a consistent reflection of the signal. Due to this slightly worse segmentation, the subsequent parallelogram fitting tends to over-separate building segments, as for buildings (3) and (4) in Figure 6.12c. Yet, it is observable that also buildings oriented almost perpendicular to flight direction and very difficult to observe in the phase image have been well extracted (e.g. buildings (5) and (6)).

Figure 6.13 shows for zone B of Paris the different extraction results obtained with the four considered datasets, two descending and two ascending ones. Considering the segmented building hypotheses, some similitude is observable in all datasets. Examples of corresponding building hypotheses are encircled in the same color in Figure 6.13 for better interpretation. For this area, it seems however that the descending datasets allow a better building recognition as the ascend-



Completeness	19/28=68%
Correctness	18/19=95%

Completeness	26/33=79%
Correctness	26/33=79%

Completeness	25/45=56%
Correctness	23/37=62%

Figure 6.14: Comparison of the extracted parallelograms with ground truth; (a) Berlin; (b) Paris zone A - 52°; (c) Paris zone B - 52°

ing datasets. There is more similarity between the fitted parallelograms of the two descending datasets, even if the data are acquired from different incidence angles (42° and 52°). On the contrary, both ascending datasets are acquired from the same incidence angle of 47°, but the extracted parallelograms present high discrepancies. Furthermore, much less building hypotheses are detected within the ascending datasets and the segmented hypotheses are quite disconnected. Even if the phase image of Figure 6.13c is less noisy than the one of Figure 6.13d due to the single-pass acquisition, most of the building hypotheses are too sparse and do not allow a correct parallelogram fitting. Only dense building hypotheses are correctly fitted with parallelograms. Consequently, for the further building analysis and parameter extraction, only the descending datasets are considered.

Figure 6.14 allows a comparison between the extracted building layovers, i.e. the fitted parallelograms, and a reference. As the layovers represent the building facades, they are situated on the side of the building footprints that looks towards the sensor, and an overlapping representation of reconstructed layovers and footprint of the reference would be confusing. Instead, the buildings have been enumerated in both images in order to allow better comparison. The extracted parallelograms for the Paris dataset have additionally been georeferenced. A coarse estimation of the completeness and correctness of the extracted parallelograms has been performed. The completeness is estimated by counting the number of fitted parallelograms that correspond to the reference, compared to the number of buildings present in the reference. The correctness is the number of fitted parallelograms that actually correspond to the reference, compared to the total number of extracted parallelograms. For one-facade buildings separated into two or more parallelograms, only one parallelogram is counted as correct. The highest completeness

Table 6.8: Accuracy of parallelogram fit

	m_0 (pixel)	Σ_α (°)	Σ_{s_x}	Σ_{t_x} (pixel)
Berlin	0.34	0.006	0.002	0.047
Paris zone A	0.44	0.003	0.003	0.056
Paris zone B	0.52	0.014	0.004	0.089

is reached for zone A of Paris, with 79%. This corresponds also to the correctness of the extracted parallelograms, meaning that 79% of the extracted parallelograms actually correspond to buildings. The best correctness is achieved for Berlin, were almost all fitted parallelograms do correspond to a real building.

Complementary to the coarse evaluation of the parallelograms' fits, the standard error of unit weight m_0 was evaluated for each parallelogram during the least squares fit. This error represents how well the extracted parallelograms fit the segmented building hypothesis. A lower value corresponds to a good fit between parallelogram and building hypothesis. The mean over all extracted parallelograms is evaluated for each scene and written in Table 6.8, first column. The scene of Berlin obtains the best result, whereby the fits of zone B shows the worst result, due to the sparseness of the building hypotheses. In addition, the internal accuracy of the least squares fit is evaluated by determining the covariance of each parameter (shearing α, scale s_x and translation t_x, cf. Section 3.6.1). These covariances are also listed in Table 6.8. Same conclusions can be deduced, whereby the scene of Berlin provides a better accuracy and the zone B of Paris shows higher variances.

In general, the developed method for the extraction of the building layovers performs well, allowing a good segmentation of the building hypotheses. Depending on the appearance of the buildings in the different datasets, but also to their orientation or direct surroundings, the hypotheses are more or less dense and yield a correspondingly well defined final parallelogram fit, from which the building parameters can be deduced. The extracted building parameters from the parallelograms and their accuracy are given in the next section.

6.2.3 Results of Building Parameter Determination

Table 6.9 depicts the extracted building parameters for the buildings enumerated in Figure 6.14, for the scene of Berlin and the two zones of Paris. The reference parameters obtained from GIS data and the BDTopo© are mentioned, and the differences between extracted parameters and reference are listed for each building. As mentioned in Section 3.6, three parameters are determined for each facade: length l, orientation α and height h. The height h is determined in two ways, either considering the parallelogram extent in range direction (h_p), or the underlying filtered phase values (h_ϕ).

As expected from the previous analysis of the quality of the building segmentation and parallelogram fit, the determined parameters for the scene of Berlin are very accurate. Only a few wrong estimates arise. The longest facade of the two-facades building (5) was not well retrieved. The fitted parallelogram presents a slightly different orientation as in reality (16°). This induces a wrong estimation of the facade length and height, as their calculation depends on the parallelogram extents in both range and azimuth direction. However, the pixel spacing in azimuth direction is larger than in range direction. Therefore, for the small facade of building (5), the

Table 6.9: Estimated building parameters by InSAR; l: length; h: height (h_p: height determined by the parallelogram extent; h_φ: height determined by the phase values); α: orientation; Δ: differences between determined parameter and reference; buildings marked in grey are clear false extraction that were left apart for subsequent precision analysis

Building	Reference			ExtractedParallelograms				Differences			
	l (m)	h (m)	α (°)	l (m)	h (m)		α (°)	Δ_l (m)	Δ_h (m)		Δ_α (°)
					h_p	h_φ			Δ_{h_p}	Δ_{h_φ}	
Berlin											
(1)	39	25	143	40	29	21	140	1	4	-4	-3
(2)	80	29	38	88	32	26	35	8	3	-3	-3
(3)	80	29	33	94	30	29	37	14	1	0	4
(4)	95	29	37	90	32	34	37	-5	3	5	0
(5)	19	75	140	18	70	71	96	-1	-5	-4	-44
	53	75	41	35	62	61	57	-18	-13	-14	16
(6)	134	30	31	69	34	21	32	-65	4	-9	1
(7)	108	30	133	98	31	24	132	-10	1	-6	-1
(8)	82	35	131	82	37	20	134	0	2	-15	3
(9)	49	30	33	49	18	13	35	0	-12	-17	2
(10)	55	30	149	62	33	27	151	7	3	-3	2
(11)	111	30	150	113	38	31	150	2	8	1	0
(12)	84	30	150	87	25	36	148	3	-5	6	-2
(13)	110	30	150	115	32	31	150	5	2	1	0
(14)	56	30	150	57	32	27	148	1	2	-3	-2
Paris - zone A - 52°											
(A0)	34	46	84	64	55	15	78	30	9	-31	-6
(A1)	88	34	103	84	28	27	101	-4	-6	-7	-2
(A2)	89	36	101	50	33	33	100	-39	-3	-3	-1
(A3)	49	35	104	39	18	17	125	-10	-17	-18	21
(A6)	89	34	101	83	41	43	104	-6	7	9	3
(A8)	90	30	101	86	23	22	95	-4	-7	-8	-6
(A9)	56	15	102	43	16	15	103	-13	1	0	1
(A10)	51	17	102	45	17	16	94	-6	0	-1	-8
Paris - zone B - 52°											
(B2)	52	16	101	60	28	24	99	8	12	8	-2
(B3)	94	32	103	72	13	32	80	-22	-19	0	-23
(B4)	55	8	96	55	19	18	90	0	11	10	-6
(B5)	93	33	102	75	34	28	99	-18	1	-5	-3
(B6)	70	13	102	68	24	24	101	-2	11	11	-1
(B7)	51	16	100	24	9	12	90	-27	-7	-4	-10
(B11)	71	34	101	31	25	20	123	-10	-3	-14	22
(B12)	94	34	102	47	35	22	110	-47	13	-12	8

wrong orientation has a lower influence on the results as its extent in azimuth direction is small. Building (6) shows a good orientation and height estimate. However, its length is underestimated. Having a look at Figure 6.12a, it is observable that the building has been separated into two parallelograms, from which only one is considered for parameter estimation. The underestimation of the building height for building (9) is caused by the incomplete detection of the layover, probably due to disturbing trees situated in front of the building.

Both scenes of Paris present higher discrepancies in the determined building parameters. For example, building A2 has been separated into two parallelograms (cf. Figure 6.14), leading to an significant underestimation of the building length. Similarly, a few buildings of zone B are separated into several parallelograms due to the sparseness of the detection, or are just partly detected, leading to an underestimation of their length (e.g. buildings B11 and B12). In Fig-

Table 6.10: Average and standard deviation of the difference between estimated and reference building parameters

		Δ_l (m)	Δ_{h_p} (m)	Δ_{h_ϕ} (m)	Δ_α (°)
Berlin	μ	3.3	2.0	3.8	12.8
	σ	3.9	3.5	4.6	5.5
Paris- zone A	μ	10.4	6.7	10.6	6.7
	σ	8.5	5.2	10.0	6.3
Paris- zone B	μ	7	8.8	8.5	3
	σ	7	4.5	2.3	1.9

ure 6.12d, some surroundings of building B3 are contained in the building hypothesis, leading to the wrong orientation of the final parallelogram.

In general, for the other well recognized buildings, the estimation of the building parameters is accurate. Only a slight underestimation of the height is generally observed using the phase values. The underestimated phase values arise if the determined building orientation is slightly different from the real orientation. Indeed, the phase values are summed up along the building orientation in order to add constant phase values together. If the determined orientation varies, different phase values are summed up, decreasing the final building height. Figure 6.15 shows this phenomenon on two extracted buildings, the first one providing the correct building height (building (13)), and the second providing an underestimated height (building (5)).

Table 6.10 depicts for each scene the mean and standard deviation of the differences between extracted and reference parameters, leaving clear outliers apart. The outliers are marked in grey in Table 6.9 for each scene.

For the Berlin dataset, all parameters are estimated very accurately (about 3 m for length and height). The worst orientation accuracy is only caused by the small part of building (5), otherwise it would be of less than 3°. Such results are very encouraging and show that the solely use of interferometric phase images can provide sufficient accuracy for the estimation of pre-event data for change detection.

For the Paris dataset, the building orientation is accurate, but the accuracy of the other parameters is between 7 m and 10 m. As already mentioned, this is mainly due to the higher amount of noise in the dataset. The building facades present many arbitrary patterns that disturb a consistent path of the signal for the whole facades. Besides, in the Berlin dataset, the buildings are situated almost isolated, near main and large traffic axes, which allows an undisturbed layover analysis. In the Paris dataset, the buildings are closer together, near parking lots and trees. Finally for zone B, demolition machines and fences present an additional disturbance.

In general, the developed method shows high potential for the reconstruction of well isolated buildings presenting an uniform facade structure. As the extraction accuracy of the building parameters for Paris, especially for zone B, is not sufficient and the original building parameters of B3, B11 and B12 cannot be retrieved correctly, the further radargrammetric analysis relies on GIS data in order to have an accurate estimate of the pre-event building parameters.

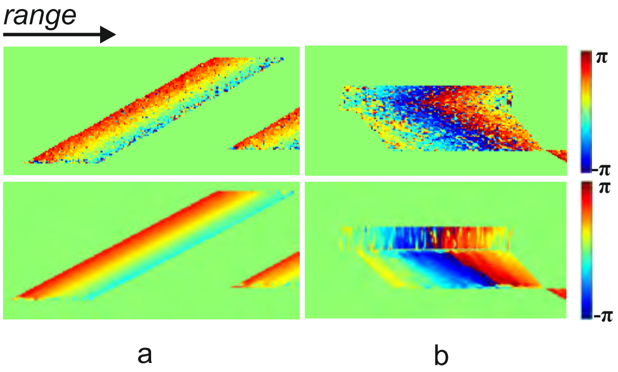


Figure 6.15: Filtered phase values depending on building orientation; (a) building (13); (b) building (5)

6.3 Results of Radargrammetric Processing

In this subsection, the results of the radargrammetric processing chain are given and discussed. First, the accuracy of the new coregistration method is evaluated and compared to standard methods. Second, the parameter setting for the calculation of the disparity map is analyzed. In a third part, results of building parameter estimation are given, for damaged and undamaged buildings. Hereby, the quality of the results obtained using different configurations is discussed.

6.3.1 Accuracy of the Coregistration with SAR-SIFT

In this section, both settings of the SAR-SIFT parameters and its comparison with standard methods are presented.

Setting of the SAR-SIFT Parameters

For the SAR-SIFT algorithm presented in Section 4.4.2, two parameters have a strong influence on the coregistration result: the threshold set on the new gradient by ratio (GR) for detecting the keypoints and the size of the neighborhood for orientation and descriptor estimation. The size of the neighborhood corresponds to the number of 4x4 pixel windows taken around each keypoint for orientation and descriptor estimation. For each configuration, different thresholds on the GR ($t_{GR} = [0.005, 0.02, 0.05]$) and different sizes of the neighborhood ($NS = [2, 8, 12]$) have been tested and the achieved accuracies in range and azimuth direction after coregistration have been determined. For this purpose, the offsets of the positions of at least eight homologous ground points between both images has been manually estimated and averaged in both range and azimuth direction. Table 6.11 gives for each configuration the best achieved accuracies in range and azimuth direction, and the corresponding set of parameters. For configurations presenting an intersection angle greater than 23°, the images cannot be coregistered using the presented SAR-SIFT algorithm. Their radiometric differences are too large to find any similarity. However, the SAR-SIFT algorithm performs well up to an intersection angle of 14°. The achieved accuracies are less than 0.5 m, except for the configuration presenting an intersection angle of 14°, for which the range accuracy is of 1 m. A large size of the neighborhood NS ($NS \in [8, 12]$) seems to be optimal, as well as a medium threshold t_{GR} ($t_{GR} \in [0.02, 0.05]$).

Table 6.11: Parameters and corresponding best achieved accuracy for each configuration

Configuration	Size of Neighborhood	Threshold on GR	Range Accuracy (m)	Azimuth Accuracy (m)
47°/36°	8	0.02	0.1	0.2
47°/22°	-	-	-	-
36°/22°	8	0.05	1.0	0.2
52°/42°	8	0.05	0.1	0.3
52°/29°	-	-	-	-
42°/29°	12	0.02	0.1	0.3

Table 6.12: Comparison of SAR-SIFT with standard methods in terms of accuracy; Config.: configuration; Correl.: coregistration using cross correlation; SIFT (Lowe): coregistration using the standard SIFT; SIFT (M.-H.): coregistration using multiscale Harris on SIFT; SAR-SIFT: coregistration using the proposed method based on SAR-SIFT; 2D Acc: two-dimensional accuracy; Nb KP: number of extracted keypoints; Nb Match: number of matches before RANSAC; Nb IN: number of matches (inliers) after RANSAC

Config.	Correl.	SIFT (Lowe)				SIFT (M.-H.)				SAR-SIFT			
	2D Acc (m)	Nb KP (10^3)	Nb Match	Nb IN	2D Acc (m)	Nb KP (10^3)	Nb Match	Nb IN	2D Acc (m)	Nb KP (10^3)	Nb Match	Nb IN	2D Acc (m)
47°/36°	-	2.2/2.8	310	9	-	1.3/1.7	356	34	5.4	4.5/5.2	300	99	0.2
36°/22°	-	2.8/1.7	199	21	-	2.5/1.3	194	18	-	4.3/2.0	199	13	1
52°/42°	7.2	5.3/2.7	594	8	-	2.9/1.5	583	76	2.4	8.0/4.4	601	255	0.3
42°/29°	-	2.7/1.9	111	7	-	0.8/0.5	129	17	-	5.8/3.5	111	46	0.3

Comparison with standard methods

In order to assess the contribution of the SAR-SIFT algorithm for the coregistration purpose, it is interesting to compare its results with some standard methods. SAR-SIFT relies only on the radiometric information content of the image, i.e. no external information such as DEM is used but image registration is performed by matching radiometric similarities in both images. Three comparable methods have been tested here. First, registration has been performed using standard cross correlation (Correl.) between both images. This method is usually employed for coregistering interferometric image pairs, and is implemented on standard SAR softwares like the open source software NEST used in this work for this task. Second, the performance of the standard SIFT algorithm (Lowe 1999) has been evaluated. This algorithm detects keypoints by determining the extrema of a difference of Gaussian scale space. Finally, the third tested algorithm is a modification of the standard SIFT algorithm using a multiscale Harris (M.-H.) image space instead of a difference of Gaussian scale space. Basically, second derivatives of the intensity values are used by the Harris detector. In order to allow a reliable comparison of the algorithms' performances, the parameters are set for each approach so that approximately the same number of matches before RANSAC (Nb Match) is obtained for all methods. Results of those methods in terms of 2D accuracy (2D Acc), number of detected keypoints (Nb KP) and number of final inliers after RANSAC (Nb IN) for determining the parameters of the projective transformation are presented in Table 6.12. The 2D accuracy of the SAR-SIFT algorithm is calculated using the best range and azimuth accuracies achieved for each configuration (Config.) presented in Table 6.11.

Obviously, the correlation-based method and the original SIFT method are both outperformed for the considered configurations. The correlation-based method manages to register the images

Table 6.13: Constraint values for the available configurations

Radargrammetric Configuration	22°/36°	22°/47°	36°/47°	29°/42°	29°/52°	42°/52°
ζ (°)	1.7	3.4	1.7	1.7	3.3	1.6
da_{max} (pix)	2.1	5.2	1.3	1.6	3.8	1.1
dr_{max} (pix)	64.1	108.5	28.8	41.2	71.4	23.0

only for the configuration presenting the smallest intersection angle (10° for 52°/42°), but the 2D registration accuracy is higher than $7m$, which makes this method useless for further processing. Even if the number of detected keypoints with the original SIFT algorithm allows to obtain enough matches, the estimated transformation is wrong due to the high outlier ratio, i.e. RANSAC cannot find a proper solution. The 2D accuracy cannot be estimated. Results using the multiscale Harris detector are slightly better, and the images can be coregistered with the two configurations presenting the smallest intersection angles (10° for 52°/42° and 11° for 47°/36°). However, the number of final inliers for the projective transformation is low compared to the SAR-SIFT, for a similar number of matches (Nb Match) before RANSAC. Only the configuration 36°/22° shows a slightly larger number of inliers than for the SAR-SIFT, but the estimated transformation does not permit the coregistration. With the SAR-SIFT, more keypoints are detected for the same number of matches, but in general much more inliers subsist after RANSAC. Furthermore, coregistration is possible and very accurate for each presented configuration of Table 6.12, with a lowest accuracy of 1 m, which is still very satisfactory.

6.3.2 Parameter Setting for the Disparity Map Calculation

In Section 4.5.3, two constraints da_{max} and dr_{max} were introduced in order to narrow down the search area for matching of layover areas. They correspond to the azimuth, respectively range component of the total disparity d. As the original building height h is known from the InSAR processing - respectively here GIS data - and the difference of heading angle ζ is known from the product metadata, both constraints can be evaluated. The average building height in the test area is about 33 m. In order to leave a buffer for matching, a maximum expected building height of 40 m is considered for the calculation of da_{max} and dr_{max}. Ideally, these parameter could be evaluated for each building separately. Table 6.13 shows the results for different configurations. According to Equations (4.21) and (4.19), subsequent template and search window sizes can be defined that permit to retrieve the correct disparity d, depending on the acquisition configuration. Table 6.14 shows for the configuration 47°/36° possible dimensions of template and search windows. For all presented combinations, the disparity map of building B5 (cf. Figure 6.14) is created and its height calculated within the manually defined P_h using Equation (4.24). The reference height of B5 is 33 m. Table 6.15 shows the standard deviation of the disparity values within P_h for all corresponding combinations of window dimensions. They are normalized by the total number of pixel within P_h that changes depending on the chosen template size l_m (cf. Section 4.5.3). All results of Tables 6.14 and 6.15 are given for two matching criteria: the standard normalized cross-correlation ρ (cf. Equation (4.3)) and the criterion based on the coefficient of variation υ (cf. Equation (4.4)), in order to compare their performance.

Table 6.14: Influence of the window size on the mean building height (m)

			Normalized Cross Correlation					Coefficient of Variation					σ_ρ (m)	σ_υ (m)
		w_m (pix)	1	3	3	3	11	1	3	3	3	11		
		w_s (pix)	5	7	9	15	25	5	7	9	15	25		
l_m (pix)	l_s (pix)													
9	95		58	38	38	40	42	45	42	43	43	45	8.4	1.3
11	69		43	33	34	36	37	36	35	33	34	34	3.9	1.1
15	119		58	43	43	46	48	48	46	46	46	51	6.2	2.2
21	79		40	31	33	35	37	33	32	33	34	33	3.5	0.7
21	119		52	42	42	44	46	45	43	43	44	52	4.1	3.8
35	95		42	30	32	34	33	33	33	33	35	34	4.6	0.9
STD (m)			8.2	5.6	4.7	5.0	5.8	6.8	5.9	6.1	4.6	8.9		

Table 6.15: Influence of the window size on the standard deviation of the disparities within P_h (pixel)

			Normalized Cross Correlation					Coefficient of Variation				
		w_m (pix)	1	3	3	3	11	1	3	3	3	11
		w_s (pix)	5	7	9	15	25	5	7	9	15	25
l_m (pix)	l_s (pix)											
9	95		0.009	0.006	0.005	0.005	0.006	0.006	0.006	0.006	0.006	0.008
11	69		0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.006
15	119		0.005	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.005
21	79		0.003	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.002	0.003
21	119		0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.004
35	95		0.002	0.001	0.002	0.001	0.002	0.001	0.001	0.001	0.001	0.002

It is observable from Table 6.14 that for both evaluated criteria, the dimension of template and search window along the range direction (l_m and l_s) plays an important role for the retrieval of the correct building height. Particularly when replacing l_m and l_s in Equation (4.21) by the values of Table 6.14, a better height estimate is achieved if the resulting dr corresponds to the values of Table 6.13. Larger differences between l_m and l_s induce larger height estimates, as either the search area is much larger, or the template shorter, and consequently more vulnerable to similar radiometric zones or inherent noise. Depending on the building surroundings, the highest criterion value can then be obtained for a disparity that does not correspond to the layover border but shows more similar radiometry.

The result of the height estimate does not seem to depend on the chosen window widths (w_m and w_s) in azimuth direction, as it is quite similar for every combination of window widths for a specified window length (l_m and l_s). Yet, higher discrepancies are observable for very thin window widths ($w_m = 1$ and $w_s = 5$) for criterion ρ and for large window widths for criterion υ ($w_m = 11$ and $w_s = 25$). This is confirmed by having a look in Table 6.14 at the standard deviation STD of the height values over the different length combinations using these widths. They show standard deviations of 8.2 m and 8.9 m respectively, although the standard deviation of the height values for the other widths are principally between 4 m and 6 m, for both criteria. Except from these two dimensions, both criteria seem to perform similarly for fixed window widths w_m and w_s. Considering now the standard deviations σ_ρ and σ_υ of the height values within fixed window lengths, the criterion υ based on the coefficient of variation outperforms the normalized cross correlation ρ. For most length combinations, σ_υ is of about 1 m whereas σ_ρ comes to about 4 m. Here also, the standard deviation of the height estimate is very low for chosen l_m and l_s that define a dr similar to the one of Table 6.13. Similar observations are

Table 6.16: Building parameters from radargrammetry - zone A

		A1	A2	A5	A6	A8
Reference	l (m)	90	90	90	90	90
	h (m)	34	36	32	34	30
36°/22°	$l_{0_{radarg}}$	88	88	88	88	88
	$l_{h_{radarg}}$	88	88	88	88	88
	h_{radarg}	26	33	31	32	35
	$\sigma_{0_{radarg}}$	0.0098	0.0080	0.0107	0.0102	0.0119
	$\sigma_{h_{radarg}}$	0.0155	0.0153	0.0187	0.0148	0.0141
47°/36°	$l_{0_{radarg}}$	88	88	88	88	88
	$l_{h_{radarg}}$	88	88	88	88	88
	h_{radarg}	26	30	29	28	28
	$\sigma_{0_{radarg}}$	0.0042	0.0044	0.0049	0.0042	0.0044
	$\sigma_{h_{radarg}}$	0.0040	0.0045	0.0062	0.0050	0.0040
52°/42°	$l_{0_{radarg}}$	88	88	88	88	88
	$l_{h_{radarg}}$	88	88	88	88	88
	h_{radarg}	25	29	28	29	29
	$\sigma_{0_{radarg}}$	0.0032	0.0026	0.0024	0.0028	0.0021
	$\sigma_{h_{radarg}}$	0.0043	0.0043	0.0054	0.0042	0.0045
42°/29°	$l_{0_{radarg}}$	88	88	88	88	88
	$l_{h_{radarg}}$	88	88	88	88	88
	h_{radarg}	28	22	27	22	26
	$\sigma_{0_{radarg}}$	0.0061	0.0066	0.0046	0.0052	0.0036
	$\sigma_{h_{radarg}}$	0.0083	0.0078	0.0083	0.0068	0.0077

made for the standard deviation of the disparity values within the extracted parallelogram P_h (Table 6.15).

Considering those remarks, the criterion based on the coefficient of variation is chosen for the processing of all buildings, as it permits to obtain a more robust height estimate for chosen window lengths. Furthermore, in order to have sufficiently large windows sizes for considering slight misregistration after SAR-SIFT whilst reducing the risk of wrong matching, the template size is set to $w_m = 3$ and $l_m = 21$ for this work. The search window size is adapted subsequently for each radargrammetric configuration, considering the constraints of Table 6.13.

6.3.3 Results of Building Parameter Determination

Table 6.16 shows for five buildings of zone A the extracted building parameters $l_{0_{radarg}}$, $l_{h_{radarg}}$, h_{radarg} as well as the standard deviations of the disparity values within P_0 and P_h, $\sigma_{0_{radarg}}$ and $\sigma_{h_{radarg}}$ respectively. The disparities have been weighted by their corresponding matching criterion value and normalized by the number of pixel in the corresponding parallelogram. Figure 6.16 shows the extracted parallelograms from the disparity map for some of those buildings. In this zone, no change occurred and the buildings still have their original dimensions. As stated in Section 6.2.3, the original building positions and heights are taken from available GIS data and projected into the master geometry for each configuration.

In general, P_0 is more easily recognizable in the disparity map than P_h. For the same building, P_h is better recognizable for configurations with a small intersection angle, e.g. 47°/36° and 52°/42°, than for configurations presenting a larger intersection angle (36°/22° and 42°/29°). This is due to the more similar radiometric appearance of two images taken under narrower intersection angle, which allow a more unambiguous matching. For acquisitions with a large

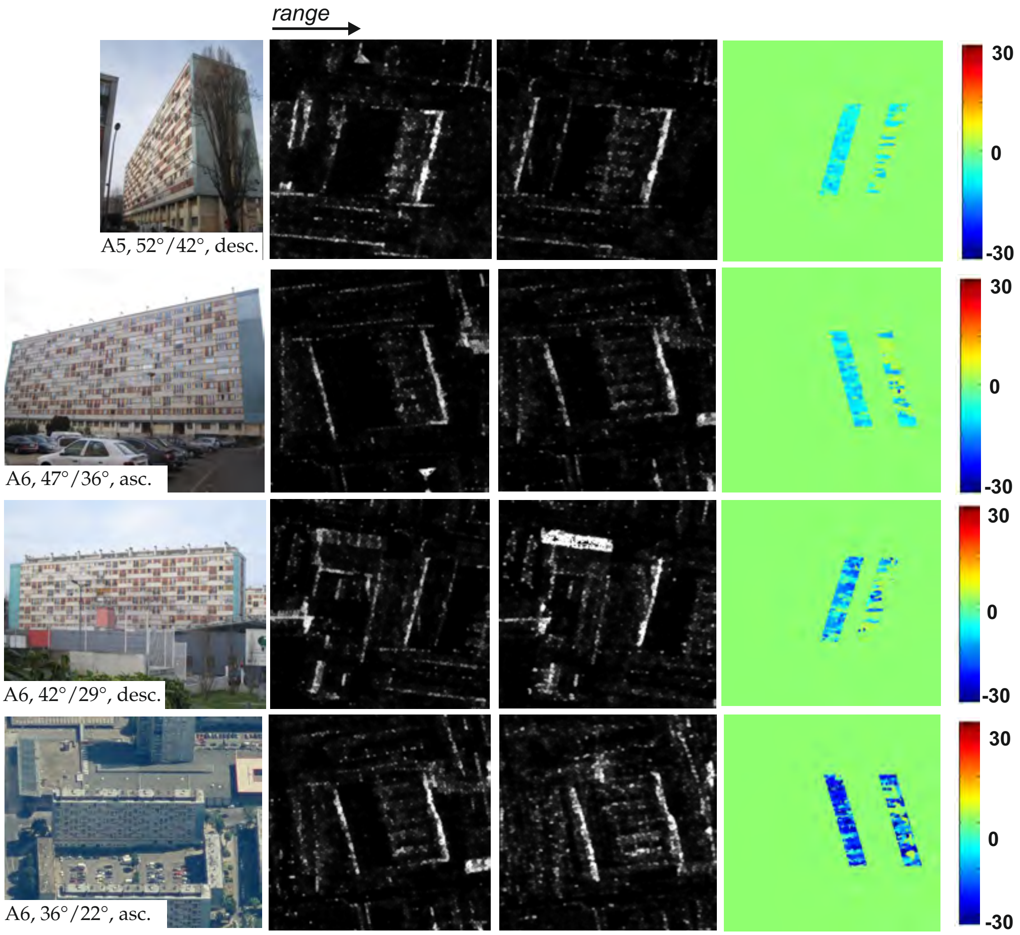


Figure 6.16: Results of parallelogram extraction for buildings of zone A (source of the optical image of the fourth row: Bing Maps)

intersection angle, overlap effects with neighbor buildings (as for building A6 for configuration 42°/29°) make the matching more ambiguous. This effect is confirmed by having a look at the weighted standard deviation of the disparity values within the extracted parallelograms. They are in general smaller within P_0 ($\sigma_{0_{radarg}}$) than within P_h ($\sigma_{h_{radarg}}$) and generally smaller for configurations with a smaller intersection angle.

Considering the extracted lengths, they are the same for both parallelograms as expected, e.g. $l_{0_{radarg}} = l_{h_{radarg}}$ for each building, and correspond to the original building lengths. In the presented cases, the estimated lengths show all a slight underestimation of 2 m for each building, which may be due to the projection step of the GIS coordinates into the geometry of the master SAR image.

The extracted building heights mostly correspond to the original building height. In general, except for building A8 under the configuration 36°/22°, a small underestimation compared to the reference is observable. The absolute differences between estimation and reference are reported in Table 6.17 for all configurations, and an average is calculated for each building (μ_{blg}) and each configuration (μ_{config}). The buildings A5 and A8 clearly show a better estimation as the others, e.g. show lower difference between estimation and reference for over all configurations, i.e. lower μ_{blg}. A mean height difference μ_{blg} of about 3 m is achieved over all configurations for those buildings. Considering their appearance in the master and slave SAR images, is is clear that

Table 6.17: Differences of height estimates to reference

	A1	A2	A5	A6	A8	μ_{config}
Reference h (m)	34	36	32	34	30	
36°/22°	8	3	1	2	5	3.8
47°/36°	8	6	3	6	2	5
52°/42°	9	7	4	5	1	5.2
42°/29°	6	14	5	12	4	8.2
μ_{blg}	7.8	7.5	3.3	6.3	3	

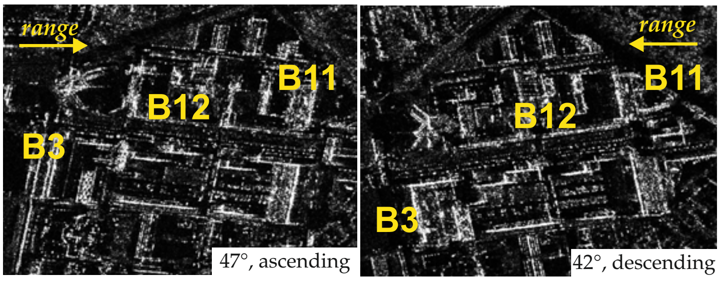


Figure 6.17: Appearance of zone B in ascending and descending intensity images

both buildings are well isolated from the surroundings and the layover borders are well defined and distinguishable. It is shown in Figure 6.16 for configuration 52°/42° for building A5. The matching of the layover borders does not suffer of any disturbing object. Considering now the average of the height differences for a same configuration μ_{config}, it is obvious that configuration 42°/29° is not suited for this test area. Especially the heights of buildings A2 and A6 are highly underestimated. Having a look at their surroundings, their facade layover overlap with neighbor buildings under descending configuration with steep incidence angles, leading to wrong matching and subsequent wrong height estimation (Figure 6.16 for building A6). P_h shows heterogeneous disparity values, yielding an underestimation of the building height after filtering. Both configurations with larger incidence angles and shorter intersection angles (47°/36° and 52°/42°) show an average height difference of less than 6 m, i.e. less than two floors, which is sufficient for rapid analysis after a disaster. For these configurations, the extent of the facade layover is smaller than with steep incidence angle, so that there is less risk of overlapping with neighboring objects. Best results are achieved with the configuration 36°/22°, presenting the largest intersection angle and the steepest incidence angle of 22°. In this case, no disturbing object appears in the facade layover and the matching performs well, even if the radiometric difference is higher than for the configurations with a smaller intersection angle (Figure 6.16 for building A6). The buildings of zone B for which changes occur, present similar properties as those of zone A. Considering ascending and descending acquisitions (cf. Figure 6.17), it seems that the facade layovers of the buildings of interest present less disturbance due to surrounding objects in the ascending configuration. Therefore, both ascending configurations 36°/22° and 47°/36° are considered for the following analysis of zone B.

Table 6.18 shows the results of parameter extraction by radargrammetry for the buildings of interest of zone B, at different dates that correspond to different building states (cf. Table 6.4 for the corresponding building states). In general, similar observations can be made as for zone A concerning the accuracy of length and height estimates as well as the behavior of the standard deviations of the disparity values. In particular, it is observable that for a building that underwent changes between the different acquisition dates, specific parameters change. Especially, the automatic adaptation of the parallelogram extent can be observed. For example, for building B12, the detected parallelogram lengths and heights differ from date to date, letting expect changes. Indeed, depending on the acquisition date and the configuration, either the extracted parallelogram lengths or the heights or both differ from the original building parameters. The case of this building is however more complex, as reconstruction occurs in a short time span after demolition (cf. Table 6.4). For building B7, already demolished in October 2012, the standard deviation of the disparity values within the extracted parallelograms is always higher as for the other not demolished buildings. Finally, building B11 is still standing in October 2012 but demolished in January 2013. Even if the extracted parallelograms have the same original length, the height parameter undergoes a higher change, and the standard deviations become higher from January 2013. A probabilistic analysis of the changes for zone B is given in Section 6.4.

6.4 Results of Change Detection

This section aims at analyzing the capacity of the developed methodology for categorizing changes at building level. First, the parameters of the conditional likelihoods presented in Table 5.1 are defined with respect to empirical knowledge of the data. Second, Bayes' rule is applied for each building and configuration in order to determine possible changes. The buildings of zone B presented in Table 6.18 are here considered for changes, using the acquisition configurations 36°/22° and 47°/36°.

6.4.1 Determination of the Bayes' Parameters

As explained in Section 6.2.3, the original building parameters are taken from GIS data for the radargrammetric analysis of zone B, as the buildings were not sufficiently recognized during interferometric processing. Therefore, the parameters μ and σ of the normal distributions of the conditional likelihoods presented in Table 5.1 are defined here with respect to the GIS data. A total of 15 parameters has to be set before applying the Bayes' rule for change detection. Usually, those parameters are learned using numerous training data. In this work, the results of zone A using GIS data are used for the parameter setting, as the buildings present similar shapes, orientation and appearance in the SAR image. As the results depend on the radargrammetric configuration, the parameter setting varies with respect to the chosen configuration. The setting of all parameters is described in the following:

- l_{InSAR}: adapted for each building, it is the original length of the building to be analyzed for changes, taken here from GIS data: $l_{InSAR} = l_{GIS}$.
- h_{InSAR}: similar to l_{InSAR}, this parameter is adapted for each building. It is the original height of the building to be analyzed for changes, taken here from GIS data: $h_{InSAR} = h_{GIS}$.
- ML: this parameter is the amount of change in the building length that should be categorized. In this work, a change of 20% of the building length is chosen, i.e. if $l0_{radarg}$ is about

Table 6.18: Building parameters of zone B at different dates

		B3	B5	B7	B11	B12
Reference	l (m)	90	90	50	70	90
	h (m)	32	33	16	34	34
October 2012						
36°/22°	$l_{0_{radarg}}$	89	89	49	68	91
	$l_{h_{radarg}}$	89	89	49	68	82
	h_{radarg}	28	29	34	28	27
	$\sigma_{0_{radarg}}$	0.0065	0.0114	0.0257	0.0132	0.0131
	$\sigma_{h_{radarg}}$	0.0152	0.0164	0.0275	0.0174	0.0136
47°/36°	$l_{0_{radarg}}$	91	91	49	68	68
	$l_{h_{radarg}}$	91	91	49	68	68
	h_{radarg}	31	31	25	26	32
	$\sigma_{0_{radarg}}$	0.0025	0.0025	0.0092	0.0039	0.0049
	$\sigma_{h_{radarg}}$	0.0050	0.0051	0.0099	0.0066	0.0079
January 2013						
36°/22°	$l_{0_{radarg}}$	89	91	49	68	91
	$l_{h_{radarg}}$	89	91	23	68	65
	h_{radarg}	31	31	29	21	27
	$\sigma_{0_{radarg}}$	0.0082	0.0084	0.0243	0.0209	0.0133
	$\sigma_{h_{radarg}}$	0.0201	0.0159	0.0419	0.0208	0.0204
47°/36°	$l_{0_{radarg}}$	89	91	49	70	93
	$l_{h_{radarg}}$	89	91	23	70	93
	h_{radarg}	32	31	24	14	24
	$\sigma_{0_{radarg}}$	0.0022	0.0038	0.0087	0.0054	0.0036
	$\sigma_{h_{radarg}}$	0.0062	0.0056	0.0179	0.0065	0.0049
October 2014						
36°/22°	$l_{0_{radarg}}$	91	91	49	70	26
	$l_{h_{radarg}}$	91	91	49	70	26
	h_{radarg}	43	26	32	19	22
	$\sigma_{0_{radarg}}$	0.0079	0.0093	0.0304	0.0263	0.0402
	$\sigma_{h_{radarg}}$	0.0196	0.0166	0.0220	0.0150	0.0510
47°/36°	$l_{0_{radarg}}$	91	91	49	70	91
	$l_{h_{radarg}}$	91	91	49	70	91
	h_{radarg}	32	34	18	10	28
	$\sigma_{0_{radarg}}$	0.0009	0.0042	0.0067	0.0054	0.0039
	$\sigma_{h_{radarg}}$	0.0066	0.0061	0.0062	0.0044	0.0045

20% shorter than the original building length l_{GIS}, the probability for this building to be categorized into class (2) increases.

- MH: this is the equivalent of ML for the height, chosen to be here also 20% of the original building height.

- MR: this parameter allows the differentiation of class (3) from classes (1) and (2). Here also, a difference of 20% between $l_{0_{radarg}}$ and $l_{h_{radarg}}$ would be considered as a building change corresponding to class (3). MR is greater than 1, specifying that $l_{h_{radarg}}$ should be shorter than $l_{0_{radarg}}$.

- M_{STD_0}: this parameter is the mean standard deviation of the disparity values within P_0. In the ideal case, all disparity values within P_0 are equal to 0 and their standard deviation is 0. However, depending on the configuration and also on the building surroundings that influence the disparity calculation, the standard deviation of the disparity values is not 0 in reality. This parameter is determined by averaging the standard deviation within P_0 for all buildings of zone A, for each configuration. The parameter set $M_{STD_0} =$

$[0.0102, 0.0044, 0.0026, 0.0053]$ pixel is obtained for the configurations 36°/22°, 47°/36°, 52°/42° and 42°/29°, respectively. For larger intersection angles, the standard deviation is higher, showing that the matching of homologous points is more ambiguous due to the radiometric difference.

- M_{STD_h}: this is equivalent to M_{STD_0} for parallelogram P_h. The set of parameters $M_{STD_h} = [0.0158, 0.0048, 0.0046, 0.0078]$ pixel is obtained for the configurations 36°/22°, 47°/36°, 52°/42° and 42°/29°, respectively. It is observable that M_{STD_h} is higher than M_{STD_0}, the high intensity of the double-bounce line yielding P_0 permitting a more reliable matching.

- M_{STD}: this is equivalent of the two previous parameters for class (4). The building is not present anymore, thus there is no double-bounce line. The disparity values around the original double-bounce line are not homogeneous anymore. Besides, the extracted P_0 does not need to correspond to the original building position. The standard deviation of the disparity values within P_0 is therefore higher. In this work, M_{STD} is set to 0.0286 pixel and 0.0098 pixel, for the configurations 36°/22° and 47°/36°, respectively. They correspond to the mean values of $\sigma_{0_{radarg}}$ of building B7 above the three acquisition dates for both configurations (cf. Table 6.18). Indeed, B7 is the only building that is demolished from the beginning of the analysis. This parameter could be enhanced if more training data were available.

Those parameters correspond to the expected means of the normal distributions for each variable of the conditional probability. The parameters l_{InSAR} and h_{InSAR} vary for each building depending on its original dimensions, but the other parameters are fixed for the defined building classes. The seven remaining parameters correspond to the standard deviations of the normal distributions for each variable. They are also defined in this work based on the results obtained for the training zone A.

- $\sigma_{l0_{radarg}}$: this parameter represents the deviation of the estimated building length compared to the original length. It corresponds to the precision with which the building length $l0_{radarg}$ is retrieved by radargrammetry. In this work, it is set to be the standard deviation between the original GIS length, considered as desired value, and the obtained building lengths, for a specific configuration. It was calculated for each configuration over all building lengths, leading each time to $\sigma_{l0_{radarg}} = 2$ m.

- $\sigma_{h_{radarg}}$: this parameter is equivalent to $\sigma_{l0_{radarg}}$ for the height estimate and is determined in the same way, using all buildings of zone A. This leads to $\sigma_{h_{radarg}} = [4.5, 5.5, 5.9, 9.1]$ m for the configurations 36°/22°, 47°/36°, 52°/42° and 42°/29°, respectively.

- SR: considering only $l0_{radarg}$ and $l_{h_{radarg}}$ from zone A, SR should be equal to 0, as it is the standard deviation of the ratio of both parameters. Yet in this work, it is set to 0.25, in order to give the possibility that the building does not show a change even if $l0_{radarg}$ and $l_{h_{radarg}}$ differ of a small amount, for example due to disturbing objects in one layover corner.

- S_{STD_0}: this parameter represents the standard deviation of the standard deviation values within parallelogram P_0. It is calculated over all buildings of zone A, for each configuration. The obtained values are $S_{STD_0} = [0.0014, 0.0003, 0.0004, 0.0012]$ pixel for the configurations 36°/22°, 47°/36°, 52°/42° and 42°/29°, respectively. It is observable, as for M_{STD_0}, that the configuration presenting the largest intersection angle shows the highest standard deviation values.

- S_{STD_h}: it is the equivalent of S_{STD_0} for parallelogram P_h, yielding $S_{STD_h} = [0.0018, 0.0009, 0.0005, 0.0006]$ pixel for the configurations 36°/22°, 47°/36°, 52°/42° and 42°/29°, respectively.

- Δ_L and Δ_{SR}: these two parameters are used for class (4), in order to give an imprecision on the parameters $l_{0_{radarg}}$ and $l_{h_{radarg}}$. Indeed, for class (4), no parallelogram is recognizable in the disparity map but the developed algorithm extracts in any case two parallelograms from the data. Their dimension and length ratio cannot be predicted with high precision, and can have a high discrepancy to the original values. These parameters are set empirically to $\Delta_L = 2$ m and $\Delta_{SR} = 0.5$.

In this work, the building orientation α is considered as exact, as coming from GIS data, and has not been evaluated. Furthermore, the accuracy of the GIS data has not been taken into account for two reasons. First, the given planimetric accuracy of the GIS data (2.5 m) corresponds to the absolute accuracy of the building corner position in the RGF93 reference frame with the associated Lambert projection. In this work, the length is considered, which represents a relative measurement between two corners. Second, the height specification is also given with an absolute precision of 2.5 m. However, for change detection, it is important to know with which accuracy the developed algorithm can estimate the relative building height. This is why the results of the training zone A are used for setting the uncertainties of the parameters of the conditional likelihoods.

6.4.2 Results of Change Detection

Table 6.19 shows the results of change detection for buildings B3, B5, B7, B11 and B12 of zone B, at three different dates, relying on the building extraction results presented in Table 6.18 and the parameter setting for the Bayes' inference defined in the previous section. Results of both configurations 36°/22° and 47°/36° are listed. Figure 6.18 displays for configuration 36°/22° a corresponding map of the buildings states at the three different dates (second row), and the corresponding reference maps (first row).

The detected changes mostly agree with the reference. Among the five presented buildings, only B5 undergoes no demolition, which is well recognized by all configurations and at all dates with the defined set of parameters. Building B7 is already demolished in October 2012 and the building site is planned to stay bare. This is also well recognized by the algorithm, except for configuration 47°/36° in October 2014, where the probability that B7 is still standing, i.e. belongs to class (1), is the highest ($P1 = 0.83$). Having a look at Table 6.18, both parallelogram lengths $l_{0_{radarg}}$ and $l_{h_{radarg}}$ are the same and correspond to the original building length, the determined

Table 6.19: Class probabilities of the buildings of zone B for different configurations and dates - the priors are set as in Equation (5.14)

		B3	B5	B7	B11	B12
October 2012						
36°/22°	P1	0.72	0.73	0	0.58	0.47
	P2	0	0	0	0	0
	P3	0.28	0.27	0	0.42	0.53
	P4	0	0	1	0	0
47°/36°	P1	0.85	0.82	0	0.54	0
	P2	0	0	0	0	1
	P3	0.15	0.18	0	0.46	0
	P4	0	0	1	0	0
January 2013						
36°/22°	P1	0.87	0.84	0	0	0.40
	P2	0	0	0	0	0
	P3	0.13	0.16	0	0	0.60
	P4	0	0	1	1	0
47°/36°	P1	0.87	0.82	0	0.06	0.42
	P2	0	0	0	0	0
	P3	0.13	0.18	0	0.94	0.58
	P4	0	0	1	0	0
October 2014						
36°/22°	P1	0.99	0.59	0	0	0
	P2	0	0	0	0	0
	P3	0.01	0.41	0	0	0
	P4	0	0	1	1	1
47°/36°	P1	0.87	0.90	0.83	0.03	0.65
	P2	0	0	0	0	0
	P3	0.13	0.10	0.17	0.97	0.35
	P4	0	0	0	0	0

building height h_{radarg} is very similar to the original height, and the standard deviation within the extracted parallelograms is very low, letting expect that the building is still standing. Having a look at the intensity images used for matching (Figure 6.19), the shadow of B5 produces very low intensity values for matching, leading in this case to more homogeneous values within P_h for B7. The original position of the double-bounce line of B7 is indicated in red. Building B11 is still standing in October 2012 (class (1)), the terrain is cleared in January 2013 and October 2014 (class (4)). The changes have been well recognized at all dates for the configuration 36°/22°. In January 2013 and October 2014, the building is categorized as ‚under demolition‘ (class (3)) with the configuration 47°/36°, instead of class (4). Here also, the standard deviation within the extracted parallelograms is too low for being considered as demolished. The case of building B12 is more ambiguous, as changes occurred between two acquisition dates. A ground truth picture from October, 16th 2012 (Figure 6.19) shows that B12 is partly demolished and that the demolition corresponds to class (2). For configuration 47°/36°, this change has been well recognized and classified. For configuration 36°/22°, a change has been recognized but classified into class (3). From Table 6.4, five days passed between the acquisitions of configuration 47°/36°, and six days between the acquisitions of configuration 36°/22°. However, acquisitions 47° and 22° were acquired during the week-end, so that there should be no change between both configurations. The ground truth picture was acquired only two days later, so that there is an uncertainty about the real appearance of B12 for both configurations. However, it is sure that a change occurred with respect to the original building dimensions and that this change can be categorized either into

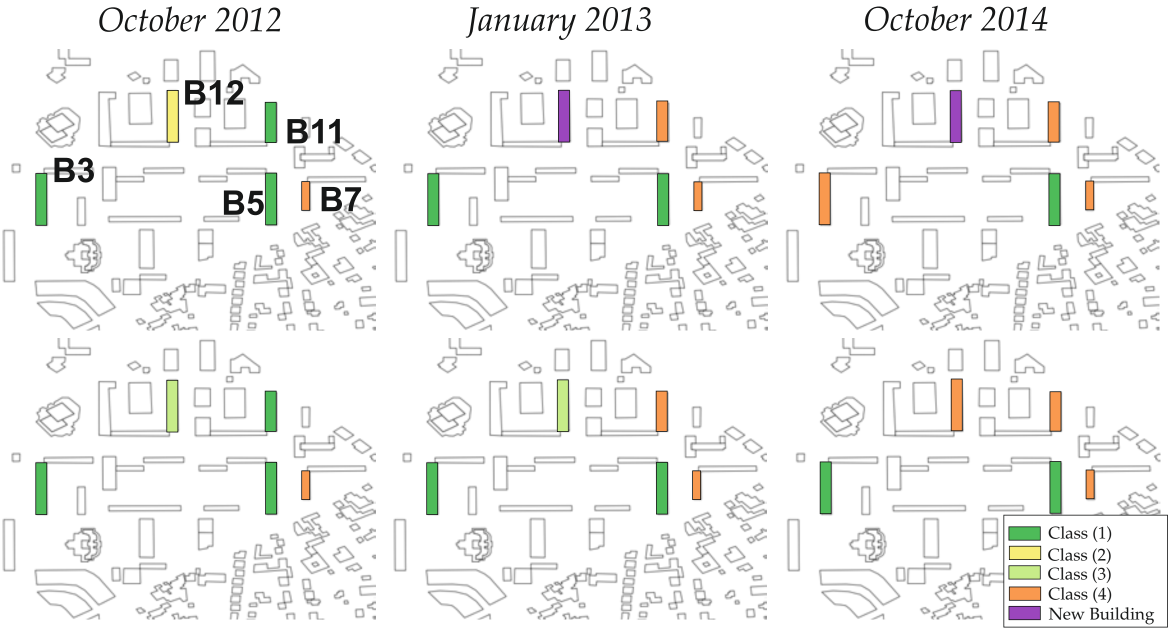


Figure 6.18: Change detection map for configuration 36°/22° at different date

class (2) or into class (3), which has been detected by both configurations. In January 2013, building B12 was demolished but another building was being built at the same site, explaining why the building is classified into class (3). In October 2014, the new construction is finished, but presents no regular rectangular geometry (cf. Figure 6.19), which explains why the results differ for both configurations. Finally, building B3 is still standing in October 2012 and January 2013 (class (1)), and the terrain is cleared in October 2014, with no new construction (class (4)). The building is well classified with both configurations for October 2012 and January 2013. Yet, it is further classified as ‚still standing' (class (1)) in October 2014. A look at the disparity map in Figure 6.19 and at the results in Table 6.18 reveals that a parallelogram P_0 of zero disparity values, with very low standard deviation, has been extracted in front of the original position of the double-bounce line (red), leading to a categorization into class (1). This is due to a fence situated in front of the building and yielding very high intensity values, similar to the ones of a double-bounce line (Figure 6.19). In this case, the use of acquisitions taken from the opposite side (i.e. descending) could improve the detection result.

In conclusion, almost all changes are detected for both configurations. For few particular cases, the category of change mismatches the reference, but the occurrence of a change was always detected, except for building B3. Due to the sparseness of the demolitions, those results represent a very small subset of data for performing a robust evaluation of the developed method. The use of more buildings both for training the Bayes' parameters and for change analysis would on the one hand permit a more robust parameter definition, and on the other hand allow an exhaustive evaluation of the method. However, it is already observable that the use of further Bayes' parameters, e.g. the distance between expected and real position of the parallelogram P_0, could improve the results. Furthermore, the results of both configurations were analyzed here separately. Combining results of several configurations, considering also acquisitions taken from the opposite building side could be helpful. In this case, new rules would be necessary to define

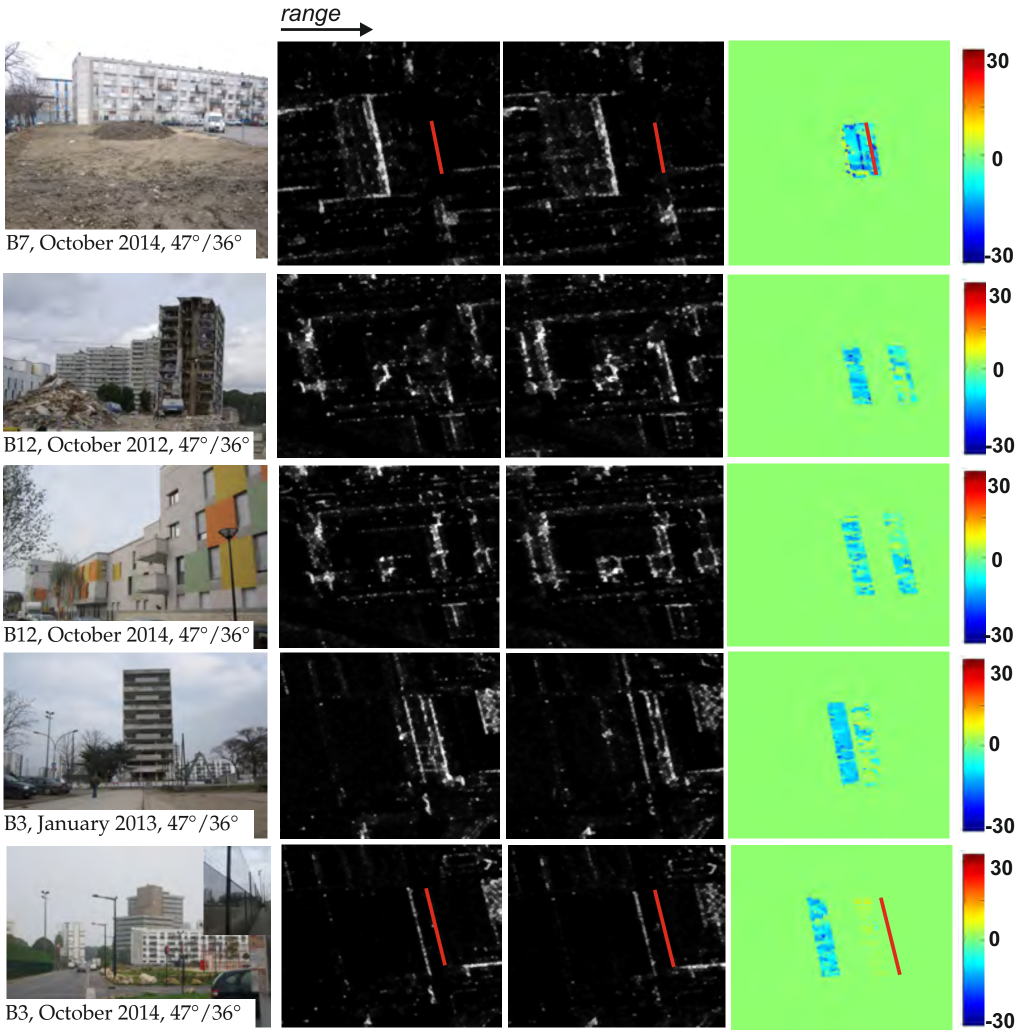


Figure 6.19: Results of parallelogram extraction for buildings of zone B

if the building is well visible on both sides or if one side is totally obstructed with overlapping effects due to neighboring objects.

Finally in this analysis, the occurrence of a change has been set to be equally likely to no change: $P(C_1) = P(C_2)+P(C_3)+P(C_4) = 0.5$ (Equation 5.14). In case of a disaster, an estimation of the total amount of destruction can be done within a few minutes from the local authorities, which permits to adjust these parameters. In this work, no such information is available. However, it is obvious that less than half the buildings are demolished. An analysis of the change detection results under a lower damage probability is therefore meaningful and will allow the analysis of the influence of these parameters. Table 6.20 shows the change detection results for the same previous buildings, considering $P(C_1) = 0.85$ and $P(C_2) + P(C_3) + P(C_4) = 0.15$. The Bayes' parameters defined in Section 6.4.1 are not modified. Even if with these new parameters, the probability that a building is still standing increases, similar results are observable for all buildings as with the previous class probabilities. Especially building B7 is still classified as ‚cleared‘ (class (4)) and the changes categorized at building B11 are correct, equal to those determined in Table 6.19. Only for building B12 in January 2013, the determined classes are not correct.

In conclusion, it seems that the chosen parameters for the Bayes' inference are correct and allow to categorize the different changes occurring at building locations. The chosen class probability

Table 6.20: Class probabilities of the buildings of zone B for different configurations and dates - the priors are set as follows: $P(C_1) = 0.85$ and $P(C_2) + P(C_3) + P(C_4) = 0.15$

		B3	B5	B7	B11	B12
October 2012						
36°/22°	P1	0.94	0.94	0	0.89	0.84
	P2	0	0	0	0	0
	P3	0.06	0.06	0	0.11	0.16
	P4	0	0	1	0	0
47°/36°	P1	0.97	0.96	0	0.87	0
	P2	0	0	0	0	1
	P3	0.03	0.04	0	0.13	0
	P4	0	0	1	0	0
January 2013						
36°/22°	P1	0.98	0.97	0	0.42	0.79
	P2	0	0	0	0	0
	P3	0.02	0.03	0	0.58	0.21
	P4	0	0	1	0	0
47°/36°	P1	0.98	0.96	0	0.28	0.80
	P2	0	0	0	0	0
	P3	0.02	0.04	0	0.72	0.20
	P4	0	0	1	0	0
October 2014						
36°/22°	P1	1	0.85	0	0	0
	P2	0	0	0	0	0
	P3	0	0.15	0	0	0
	P4	0	0	1	1	1
47°/36°	P1	0.98	0.98	0.97	0.13	0.91
	P2	0	0	0	0	0
	P3	0.02	0.02	0.03	0.87	0.09
	P4	0	0	0	0	0

has an influence on the final detection results in cases where the building category is not well defined (e.g. building B12, new reconstruction). Using more training data, and increasing the number of parameters for the definition of the different building states, would make the results more robust.

6.5 Discussion and Conclusion

In this chapter, the results of building parameter estimation by interferometry and radargrammetry have been shown and analyzed. Focusing on rectangular, isolated, flat roof buildings with predominant layover areas, several acquisition configurations have been evaluated.

During interferometric analysis, results and parameterisation of both layover detectors were analyzed. A ramp length of about a third of the layover extent in range direction has been shown as optimal for obtaining good contrast between building and surroundings with the phase ramp detector. With the constant ramp detector, the results are more spread, as they also depend on the building orientation. However, it seems that the same ratio can be considered as for the previous detector, of about a third of the building length, for optimal distinction from the surroundings. Both detectors were tested on repeat-pass and single-pass data. An important conclusion is that they both work similarly for both data types. They are therefore robust to

decorrelation effects. Furthermore, they perform better with baselines between 150 m and 350 m and incidence angles greater than 40°.
The subsequent layover extraction and parallelogram fitting shows promising results. Two-facades buildings are recognized and analyzed as two separated facades. In this work, only an angle of 90° between both facades has been permitted but the approach could easily be enhanced to other angles in order to analyze buildings with other shapes. Even in noisy data, layover are well extracted and parallelogram fitted. Also buildings whose longest side is oriented almost orthogonal to the flight direction are well extracted, yielding a high completeness of the result. For the scene of Paris, there exist a high discrepancy between fitted parallelograms of a same area considering different incidence angles and orbit orientations. This is due to the higher amount of noise in the data, not only due to time decorrelation, but particularly due to the buildings themselves and their surroundings. Indeed, the analyzed buildings present arbitrary facade structures, alternating between concrete, mosaic, french balconies and roller shutters which are not consistent between the acquisitions. Furthermore, their layover overlaps with parking lots whose appearance changes significantly between repeat-pass acquisitions, influencing the layover appearance. Finally, some of the buildings' signatures are influenced by additional disturbing objects as cranes, construction machines and fences. In order to fit more robust parallelograms, a fusion of the extracted layover segments, i.e. building hypotheses, of different acquisition configurations taken from the same direction is worth considering. Furthermore, an additional fusion with different acquisition directions would allow the reconstruction of the whole buildings, as information from both sides would be available. For the buildings that were well recognized, the extracted building parameters are very promising. The height determination using the parallelogram extent instead of the inherent phase values yields a slightly better result. Indeed, the height determination using the phase values is influenced by potentially wrong orientation of the extracted parallelogram. Moreover, the filtering of the phase values tends to yield underestimated heights. An analysis of the underlying amount of noise in the layover could be used in order to chose one method or the other for height determination for each building.
The developed interferometric approach relies only on the phase data for detecting the building and extracting subsequent information. It shows very promising results, comparable to approaches using more information, such as intensity values, simulated or GIS data. The use of such additional information could improve the robustness of the results of the developed algorithm.

Also during the radargrammetric processing, the most favorable configurations have been analyzed. It has been shown that the modified SAR-SIFT algorithm manages to match images up to an intersection angle of 15°, which is an improvement compared to standard feature matching algorithms. An analysis of the parameter setting for disparity calculation has been performed, which allows to reduce the risk of wrong matching whilst reducing the computation time. A consequence of this analysis is the choice of a matching criterion based on the coefficient of variation instead of the standard normalized cross correlation. Considering the subsequently chosen parameters, an analysis of the extracted building parameters considering different configurations has been performed. Due to the correct propagation of the original building parameters into the geometry of the disparity map, the building length can be determined with high accuracy.

The second parallelogram P_h, leading to the height determination, is more easily distinguishable in configurations with smaller intersection angles. Such configurations show more similar radiometry, facilitating the pixel-based matching. Also the standard deviation of the disparity values within P_h is smaller by configurations presenting a small intersection angle. However, the final height estimate seems independent on the chosen configuration, showing similarly good estimates. A general underestimation of the height parameter is observable, probably due to the filtering step, which is yet necessary for obtaining homogeneous disparity values within P_h. In this work, ascending configurations show better results than descending configurations. However, this result is highly dependent on the building neighborhood and should be considered carefully for each scene. For very well isolated buildings, there should be no preferential acquisition side. Two important remarks can be done regarding the radargrammetric processing. First, the dimensions of template and search window for matching have been set considering the accuracy of height estimate and the lowest standard deviation of the disparity values within P_h. This leads to very specific window sizes, whose dimension in range direction differs at least by a factor of ten with the azimuth dimension. This difference is necessary in order to obtain the sufficient disparity on the layover border in near range. However, the images are coregistered on ground level and the double-bounce lines overlap. Such large windows at the double-bounce line induce wrong matches and slow down the matching. A dynamic adaptation of the dimensions of search and template window depending on their respective position within the expected layover could improve the result and reduce the standard deviation of the disparity values within the extracted parallelograms, especially within P_0. Second, in order to improve the matching process, the algorithm favors matches lying in a specific quarter of the search window, which corresponds to the expected position of the layover border in the slave image compared to the master image. This creates a bias of the disparity values that should be taken into account, especially if building changes are expected, in which case this bias tends to overestimate the building height. A combined approach using feature detection could be useful in order to determine if any layover pattern is distinguishable in the front of the building. In the contrary case, no matching direction should be favored, so that no ‚wrong‘ building height is extracted.

The change detection approach presented in this work not only proposes the detection of the changes, but also their categorization. Beside the building class for ‚no change‘, three other change classes have been defined. For the analyzed buildings, the change detection results at three different dates show very promising results. The radargrammetric approach showed its capacity to extract smaller parallelograms, indicating a possible change. The four chosen parameters for the Bayes' inference show their potential for the categorization of the changes. The parameters were trained using similar buildings of another zone. Yet, due to the small amount of available buildings both for training and for change detection, a full evaluation of the algorithm is not possible. Anyway, using different probabilities of the occurrence of the classes, the algorithm seems already robust. Class (1) (‚building is still standing‘) and class (3) (‚building damaged only on part of the roof‘) seem to show a higher correlation. Additional parameters for the Bayes' inference could permit to separate them more easily. For example, feature extraction and texture analysis could be used in the single SAR images in order to distinguish and analyze the layover border corresponding to the roof. Moreover, the distance between extracted parallelograms and original building position could be taken into account. Furthermore in this work, the orientation α of the building has not been taken into account for change detection, as the used pre-event GIS data were considered to have an accurate orientation. Leaving the possibility to detect parallelograms in other orientations could improve the categorization of class (4), where no building is distinguishable. In all cases, the results of change detection would highly benefit of using more data, both for learning the parameters and then for analyzing the changes.

7

Conclusion and Outlook

This final chapter summarizes the presented work, showing the benefits of the developed approach for answering the different questions addressed in Chapter 1. The different contributions are resumed and their respective results are recapitulated, showing the possibilities of improvement for future works as well as desirable requirements for future satellite missions.

The first question ‚**Which SAR system(s) and technique(s) enable(s) a rapid and global application?**‘ was answered by the use of two different SAR techniques for pre-event and post-event analysis, both available on the same SAR system. The previously existent problem of the poor availability of pre-event data is solved in this work by using interferometric data of the tandem mission of the German satellites TerraSAR-X and TanDEM-X, whose goal is to provide a global Earth coverage. The use of globally available data ensures a worldwide applicability of the presented approach. The need of rapid response for post-event analysis has been addressed in this work by using radargrammetric stereo data for post-event analysis. Due to the different incidence angles, such data are acquired faster than repeat-pass interferometric data, and allow a more robust 3D scene analysis as single SAR images.

In order to detect and analyze changes in urban areas, the second question ‚**How do buildings look like, using the chosen technique(s)?**‘ had to be answered first. Consequently, methods had to be developed that permit to detect and extract the recognized pattern. In this work, the focus was put on isolated middle-rise buildings, presenting rectangular shape and flat roof. Their appearance varies between interferometric and radargrammetric data, but also depending on the acquisition configuration.
Starting with the interferometric data, the focus was put in this work on the phase images, in which the buildings have similar appearance independently of their facade structure and material. In particular, the layover areas were analyzed, characterized by a descending phase ramp in range direction and constant phase values along the building orientation. Complementary to an existing detector whose goal is to find phase ramps in range direction, a detector was developed, looking for constant phase values in several orientations. Their combination for segmenting building candidates has proven successful. Regarding different interferometric pairs, configura-

tions with effective baselines between 150 m and 300 m provide the better segmentation results. Based on the extracted segments, an algorithm was developed that permits to determine if the segmented building candidates contain one or two facades, and in the latter case, to separate them. A subsequent building extraction based on the parallelogram shape of the facade layover was implemented, using least squares fit. Here also, results were very encouraging, showing the capacity of the algorithm to recognize single facades and extract their shape. The subsequently deducted building parameters are good, showing up to 3 m accuracy in length and 2 m in height. The appearance of buildings in radargrammetric data varies depending on the incidence angle of the single acquisitions. Yet, matching two images in a stereoscopic way permits to derive a general building appearance for all configurations and subsequently extract the building height. In this work, a new methodology for matching was developed, combining the advantages of several existing approaches. The automatic coregistration of both images was addressed by modifying a SAR-SIFT algorithm in order to favor linear structures. The performance of this feature-based coregistration was compared to standard methods and showed high improvements, especially for large intersection angles. In order to keep the whole image information, a pixel-based matching method was developed. Several improvements to standard matching methods were implemented, focusing on the preservation of edges and linear structures. Refined Lee speckle filtering combined with a Laplacian image pyramid showed less edge smoothing than a standard Gaussian pyramid. Furthermore, a matching criterion based on the coefficient of variation lead to a more homogeneous estimation of the disparities around linear structures than normalized cross-correlation. Independently of the acquisition configuration, the appearance of buildings in the resulting disparity map is characterized by two parallelograms of homogeneous disparity values, situated around the layover borders of the buildings, at near and far range. The parallelogram extraction relies on statistical analysis of the disparity image. The disparities of the near range parallelogram can be interpreted to height values, corresponding to the relative building height to the ground. In this work, the difference of heading angles of both acquisitions was taken into account for a more exact height estimation. Furthermore, the parameter setting for matching was analyzed in order to avoid wrong matches and keep the processing time low. Due to the transmission of pre-event building parameters for the building detection in the disparity map, the accuracy of the determined building length is very good, up to 2 m. Also the height estimation is satisfying, up to 4 m for the best acquisition configuration. Acquisition configurations with small intersection angle show more homogeneous height estimates than configurations with large intersection angles, where the radiometry of both matched images presents high dissimilarities. In all cases, the use of ascending or descending orbit is highly depending on the building surroundings on each side.

Finally, the third question ‚**Is it possible to detect changes at building level with respect to the applied technique(s)?**‘ has been addressed in this work by using Bayes’ inference of pre- and post-event building parameters. The theory of transmitting the building parameters from one geometry into the other was explained, and the combination of interferometric and radargrammetric data in a Bayesian framework was highlighted. Besides, the parallelogram extraction during radargrammetric processing takes into account a possible reduction of the building length and height. In order to categorize the changes, four different building states have been defined, based on the expectancy on the building signature in the radargrammetric disparity map. Due to the too small amount of buildings undergoing changes in the test area, and to their bad recognition in the interferometric phase images due to surrounding construction machines and their arbitrary facade patterns, it was remedied in this work to pre-event GIS data for change analysis. The parameter of the Bayes' inference were deduced using empirical values deduced from a small training zone. Results of change detection for a few buildings at different dates are promising, but this approach would benefit of more data for the learning of the parameters of the different building classes.

Outlook

The presented approach for building detection and extraction in interferometric and radargrammetric SAR data, as well as the change detection framework, show already very good and promising results. Still, the robustness and global applicability of the approach could benefit from a few improvements. Furthermore, preferences for future satellite missions in terms of acquisition geometry, can already be deducted from the presented results.

First of all, very high resolution Spotlight data were used for the interferometric processing. Even if by now most of the big cities and risk areas have been acquired under this resolution, the global Earth coverage of the tandem mission of TerraSAR-X and TanDEM-X is ensured by StripMap data, having lower resolution. It could therefore be interesting to analyze the applicability of the developed detectors for lower resolution data, which present a shorter layover area in terms of number of pixels. In such a case, the use of additional information, as intensity data, or even GIS or simulated data, could be useful, as the fringe pattern characterizing the layover areas is reduced to a few pixels.

On the contrary, the radargrammetric processing could benefit of using even higher resolution, available since 2014 on TerraSAR-X with the Staring Spotlight mode. Such data provide more details about facade features and surroundings. The presented pixel-based matching could be fused to a feature extraction step, permitting a hybrid matching taking into account both radiometric and geometric differences between images taken under different incidence angles. Not only the homogeneity of the disparity in both extracted parallelograms could be improved, but new patterns could be extracted and matched. For example, linear structures corresponding to the different window lines could be extracted and found in the different acquisitions using a frequency-based approach that considers the incidence angle. An example of such an extraction is given in Figure 7.1 for Very High Resolution Spotlight data in a single amplitude image. There, the lines of higher intensities have been extracted by a weighted Hough transform. Defining the line frequency and using more detailed intensity images could improve the extraction, especially for images acquired with steeper incidence angle. Such a hybrid approach using both pixel-based and feature-based matching could also improve the processing time and the robustness of the algorithm, as less constraints should be defined for the search area of the matching. For exam-

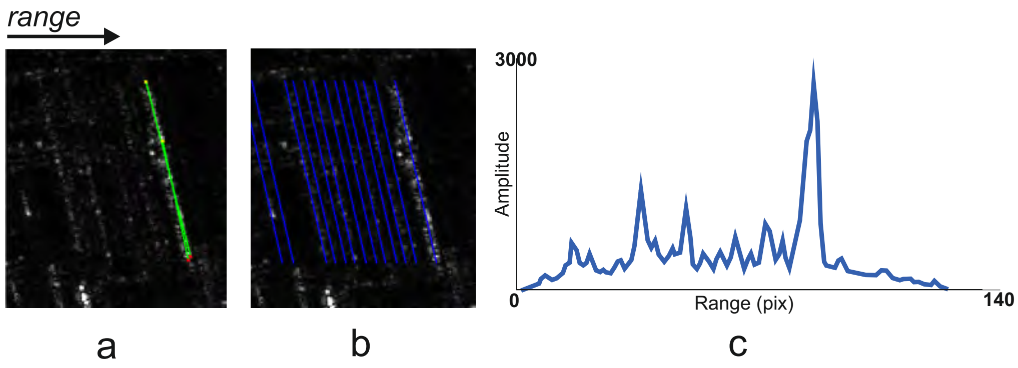


Figure 7.1: Window lines extraction using weighted Hough transform; (a) double-bounce line; (b) window lines along the facade layover; (c) profile view of the summed intensity values along the line direction

ple, a coarse matching could be performed using only the extracted features, and a pixel-based refinement could be applied subsequently.
Furthermore, the presented approach is until now limited to isolated middle-rise buildings with rectangular shape and flat roof. An analysis of the influence of neighbor objects and other building shapes on the building signature in the disparity map could permit its application in denser city areas where buildings are closer and present multiple shapes. Such an analysis would complete the analysis presented in (Thiele 2014) for interferometric phase images. Depending on the building shape, it could also be necessary to consider complementary information, e.g. from the building roof, if the layover area is not always well defined.

Moreover, the use of several acquisition configurations for the determination of the building parameters in both interferometric and radargrammetric cases could enhance the result. Indeed, using opposite-side configurations by interferometry could allow the determination of the full building footprint, and the use of opposite-sides, but also multiple same-side configurations by radargrammetry could allow a more robust parameter estimation, especially in cases where steep incidence angles would imply an overlap of the layover with neighbor objects. For fast building parameter estimation by radargrammetry, the total time span between all acquisitions should yet be considered.

Finally, considering all the previous remarks, combining information of several acquisition configurations and using additional features would provide more parameters for the change detection. New building state categories could be defined, but particularly the current categories could be better defined and separated from each other. The utilization of existing CAD models of different building damages, as presented in (Hommel 2010) for LIDAR data, could permit to broaden the applicability of the presented method to standard categories of damages, for example corresponding to the European Macroseismic Scale, whilst permitting a more robust estimation of the buildings or change parameters.

The knowledge gained by this work, as well as the possible presented improvements are valuable for future requirements on satellite missions and rapid mapping tasks.

Bibliography

Adam, N., Eineder, M., Yague-Martinez, N. & Bamler, R. (2008), High resolution interferometric stacking with TerraSAR-X, *in* 'Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International', Vol. 2, IEEE, pp. II–117.

Arciniegas, G., Bijker, W., Kerle, N., Tolpekin, V. et al. (2007), 'Coherence-and amplitude-based analysis of seismogenic damage in Bam, Iran, using Envisat ASAR data', *Geoscience and Remote Sensing, IEEE Transactions on* **45**(6), 1571–1581.

Askne, J. & Hagberg, J. (1993), Potential of interferometric SAR for classification of land surfaces, *in* 'Geoscience and Remote Sensing Symposium, 1993. IGARSS'93. Better Understanding of Earth Environment., International', IEEE, pp. 985–987.

Askne, J. I., Dammert, P. B., Ulander, L. M. & Smith, G. (1997), 'C-band repeat-pass interferometric SAR observations of the forest', *Geoscience and Remote Sensing, IEEE Transactions on* **35**(1), 25–35.

Atwood, D., Meyer, F. & Arendt, A. (2010), 'Using L-band SAR coherence to delineate glacier extent', *Canadian Journal of Remote Sensing* **36**(sup1), S186–S195.

Auer, S. & Gernhardt, S. (2014), 'Linear Signatures in Urban SAR Images-Partly Misinterpreted?', *Geoscience and Remote Sensing Letters, IEEE* **11**(10), 1762–1766.

Auer, S., Gernhardt, S. & Bamler, R. (2011a), Investigations on the nature of Persistent Scatterers based on simulation methods, *in* 'Urban Remote Sensing Event (JURSE), 2011 Joint', IEEE, pp. 61–64.

Auer, S. J. et al. (2011b), 3D synthetic aperture radar simulation for interpreting complex urban reflection scenarios, PhD thesis, Technische Universität München.

Baillard, C. (1997), Analyse d'images aériennes stéréoscopiques pour la restitution 3-D des milieux urbains. Détection et caractéristique du sursol, PhD thesis, Ecole Nationale Supérieure des Télécommunications de Paris.

Balz, T. & Liao, M. (2010), 'Building-damage detection using post-seismic high-resolution SAR satellite data', *International Journal of Remote Sensing* **31**(13), 3369–3391.

Bamler, R. & Hartl, P. (1998), 'Synthetic aperture radar interferometry', *Inverse problems* **14**(4), R1.

Baran, I., Stewart, M. P., Kampes, B. M., Perski, Z. & Lilly, P. (2003), 'A modification to the Goldstein radar interferogram filter', *IEEE Transactions on Geoscience and Remote Sensing* **41**(9), 2114–2118.

Bhattacharya, A. & Touzi, R. (2012), 'Polarimetric SAR urban classification using the Touzi target scattering decomposition', *Canadian Journal of Remote Sensing* **37**(4), 323–332.

Bignami, C., Chini, M., Pierdicca, N. & Stramondo, S. (2004), Comparing and combining the capability of detecting earthquake damages in urban areas using SAR and optical data, *in* 'Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International', Vol. 1, IEEE.

Bo, G., Dellepiane, S. & Beneventano, G. (1999), A locally adaptive approach for interferometric phase noise reduction, *in* 'Geoscience and Remote Sensing Symposium, 1999. IGARSS'99 Proceedings. IEEE 1999 International', Vol. 1, IEEE, pp. 264–266.

Bolles, R. C. & Fischler, M. A. (1981), A RANSAC-Based Approach to Model Fitting and Its Application to Finding Cylinders in Range Data., *in* 'IJCAI', Vol. 1981, pp. 637–643.

Bolter, R. & Leberl, F. (2000*a*), Detection and reconstruction of human scale features from high resolution interferometric SAR data, *in* 'Pattern Recognition, 2000. Proceedings. 15th International Conference on', Vol. 4, IEEE, pp. 291–294.

Bolter, R. & Leberl, F. (2000*b*), *Shape-from-shadow building reconstruction from multiple view SAR images*, na.

Brett, P. T. & Guida, R. (2013), 'Earthquake damage detection in urban areas using curvilinear features', *Geoscience and Remote Sensing, IEEE Transactions on* **51**(9), 4877–4884.

Brunner, D. (2009), Advanced methods for building information extraction from very high resolution SAR data to support emergency response, PhD thesis, University of Trento.

Brunner, D., Lemoine, G. & Bruzzone, L. (2010), 'Earthquake damage assessment of buildings using VHR optical and SAR imagery', *Geoscience and Remote Sensing, IEEE Transactions on* **48**(5), 2403–2420.

Budillon, A., Pascazio, V. & Schirinzi, G. (2013), 'Joint estimation of moving target reflectivity and velocity via AT-InSAR systems based on complex interferometric data', *ISPRS Journal of Photogrammetry and Remote Sensing* **75**, 1–10.

Bürgmann, R., Rosen, P. A. & Fielding, E. J. (2000), 'Synthetic aperture radar interferometry to measure Earth's surface topography and its deformation', *Annual Review of Earth and Planetary Sciences* **28**(1), 169–209.

Burt, P. J. & Adelson, E. H. (1983), 'The Laplacian pyramid as a compact image code', *Communications, IEEE Transactions on* **31**(4), 532–540.

Butenuth, M., Frey, D., Nielsen, A. A. & Skriver, H. (2011), 'Infrastructure assessment for disaster management using multi-sensor and multi-temporal remote sensing imagery', *International journal of remote sensing* **32**(23), 8575–8594.

Capaldo, P., Crespi, M., Fratarcangeli, F., Nascetti, A., Pieralic, F., Agugiaro, G., Poli, D. & Remondino, F. (2012), DSM generation from optical and SAR high resolution satellite imagery: Methodology, problems and potentialities, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International', IEEE, pp. 6936–6939.

Capaldo, P., Crespi, M., Fratarcangeli, F., Nascetti, A. & Pieralice, F. (2011), 'High-resolution SAR radargrammetry: A first application with COSMO-SkyMed spotlight imagery', *Geoscience and Remote Sensing Letters, IEEE* **8**(6), 1100–1104.

Carrasco, D., Diaz, J., Broquetas, A., Arbiol, R., Castillo, M. & Pala, V. (1997), Ascending-descending orbit combination SAR interferometry assessment (DEM), *in* 'Third ERS Symposium on Space at the service of our Environment', Vol. 414, p. 1789.

Chan, Y. K. & Koo, V. C. (2008), 'An introduction to synthetic aperture radar (SAR)', *Progress In Electromagnetics Research B* **2**, 27–60.

Chaussard, E., Wdowinski, S., Cabral-Cano, E. & Amelung, F. (2014), 'Land subsidence in central Mexico detected by ALOS InSAR time-series', *Remote Sensing of Environment* **140**, 94–106.

Chhikara, R. (1988), *The Inverse Gaussian Distribution: Theory: Methodology, and Applications*, Vol. 95, CRC Press.

Chini, M., Pierdicca, N. & Emery, W. J. (2009), 'Exploiting SAR and VHR optical images to quantify damage caused by the 2003 Bam earthquake', *Geoscience and Remote Sensing, IEEE Transactions on* **47**(1), 145–152.

Chini, M., Piscini, A., Cinti, F. R., Amici, S., Nappi, R. & DeMartini, P. M. (2013), 'The 2011 Tohoku (Japan) tsunami inundation and liquefaction investigated through optical, thermal, and SAR data', *Geoscience and Remote Sensing Letters, IEEE* **10**(2), 347–351.

Consultants, A. (2009), DIVERSIFICATION DE L'HABITAT ET MIXITE SOCIALE DANS LES QUARTIERS EN RENOVATION URBAINE - Le project de Clichy-sous-Bois/Montfermeil, Technical report.

Cossu, R., Dell'Acqua, F., Polli, D. A. & Rogolino, G. (2012), 'SAR-based seismic damage assessment in urban areas: scaling down resolution, scaling up computational performance', *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* **5**(4), 1110–1117.

Crosetto, M. & Pérez Aragues, F. (2000), Radargrammetry and SAR Interferometry for DEM Generation: Validation and Data Fusion., *in* 'SAR workshop: CEOS Committee on Earth Observation Satellites', Vol. 450, p. 367.

Dai, E., Jin, Y.-Q., Hamasaki, T. & Sato, M. (2008), 'Three-dimensional stereo reconstruction of buildings using polarimetric SAR images acquired in opposite directions', *Geoscience and Remote Sensing Letters, IEEE* **5**(2), 236–240.

Daubechies, I. et al. (1992), *Ten lectures on wavelets*, Vol. 61, SIAM.

Deledalle, C.-A., Denis, L., Tupin, F., Reigber, A., Jäger, M. et al. (2014), 'NL-SAR: a unified Non-Local framework for resolution-preserving (Pol)(In) SAR denoising'.

Deledalle, C.-A., Tupin, F. & Denis, L. (2010), A non-local approach for SAR and interferometric SAR denoising, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International', IEEE, pp. 714–717.

Dell'Acqua, F., Bignami, C., Chini, M., Lisini, G., Polli, D. A. & Stramondo, S. (2011), 'Earthquake damages rapid mapping by satellite remote sensing data: L'Aquila april 6th, 2009 event', *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* **4**(4), 935–943.

Dell'Acqua, F. & Gamba, P. (2001), 'Detection of urban structures in SAR images by robust fuzzy clustering algorithms: the example of street tracking', *Geoscience and Remote Sensing, IEEE Transactions on* **39**(10), 2287–2297.

Dell'Acqua, F., Gamba, P., Soergel, U. & Thoennessen, U. (2001), Fully 3D and joint 2D/3D segmentation of InSAR data over urban environments, *in* 'Remote Sensing and Data Fusion over Urban Areas, IEEE/ISPRS Joint Workshop 2001', IEEE, pp. 328–331.

Dellinger, F., Delon, J., Gousseau, Y., Michel, J. & Tupin, F. (2015), 'SAR-SIFT: A SIFT-Like Algorithm for SAR Images'.

Desai, M. D. (1997), 'Spotlight mode SAR stereo technique for height computation', *Image Processing, IEEE Transactions on* **6**(10), 1400–1411.

Do, M. N. & Vetterli, M. (2001), Frame reconstruction of the Laplacian pyramid, *in* 'Acoustics, Speech, and Signal Processing, 2001. Proceedings.(ICASSP'01). 2001 IEEE International Conference on', Vol. 6, IEEE, pp. 3641–3644.

Dong, L. & Shan, J. (2013), 'A comprehensive review of earthquake-induced building damage detection with remote sensing techniques', *ISPRS Journal of Photogrammetry and Remote Sensing* **84**, 85–99.

Dong, Y., Forster, B. & Ticehurst, C. (1997), 'Radar backscatter analysis for urban environments', *International Journal of Remote Sensing* **18**(6), 1351–1364.

Dong, Y., Li, Q., Dou, A. & Wang, X. (2011), 'Extracting damages caused by the 2008 Ms 8.0 Wenchuan earthquake from SAR remote sensing data', *Journal of Asian Earth Sciences* **40**(4), 907–914.

Dubois, C., Thiele, A. & Hinz, S. (2012), Adaptive filtering of interferometric phases at building location, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International', IEEE, pp. 6625–6628.

Dubois, C., Thiele, A. & Hinz, S. (2013), 'Pixel-based approach for building heights determination by sar radargrammetry', *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* **1**(3), 19–24.

Dubois, C., Thiele, A. & Hinz, S. (2014), Building Detection in TanDEM-X Data, *in* 'EUSAR 2014; 10th European Conference on Synthetic Aperture Radar; Proceedings of', VDE, pp. 1–4.

Dubois, C., Thiele, A., Welte, A. & Hinz, S. (2015), Parameterisation of layover detector for urban analysis in InSAR data, *in* 'Urban Remote Sensing Event (JURSE), 2015 Joint', IEEE, pp. 1–4.

Fayard, F., Méric, S., Pottier, E. et al. (2007), 'Matching stereoscopic SAR images for radargrammetric applications', *Proceeding IGARSS 2007* .

Ferraioli, G. (2010), 'Multichannel InSAR building edge detection', *Geoscience and Remote Sensing, IEEE Transactions on* **48**(3), 1224–1231.

Ferretti, A., Monti Guarnieri, A., Prati, C. & Rocca, F. (1997), Multi-baseline interferometric techniques and applications, *in* 'ERS SAR Interferometry', Vol. 406, p. 243.

Ferretti, A., Monti-Guarnieri, A., Prati, C., Rocca, F. & Massonet, D. (2007), *InSAR Principles-Guidelines for SAR Interferometry Processing and Interpretation*, Vol. 19.

Ferretti, A., Prati, C. & Rocca, F. (2000), 'Nonlinear subsidence rate estimation using permanent scatterers in differential SAR interferometry', *Geoscience and Remote Sensing, IEEE transactions on* **38**(5), 2202–2212.

Fjortoft, R., Lopes, A., Marthon, P. & Cubero-Castan, E. (1998), 'An optimal multiedge detector for SAR image segmentation', *Geoscience and Remote Sensing, IEEE Transactions on* **36**(3), 793–802.

Franceschetti, G., Iodice, A. & Riccio, D. (2002), 'A canonical problem in electromagnetic backscattering from buildings', *Geoscience and Remote Sensing, IEEE Transactions on* **40**(8), 1787–1801.

Frey, O., Siddique, M. A., Hajnsek, I., Wegmueller, U. & Werner, C. L. (2014), Combining SAR tomography and a PSI approach for highresolution 3-D imaging of an urban area, *in* 'EUSAR 2014; 10th European Conference on Synthetic Aperture Radar; Proceedings of', VDE, pp. 1–4.

Fritz, T., Eineder, M., Mittermayer, J., Schättler, B., Balzer, W., Buckreuß, S. & Werninghaus, R. (2008), 'TerraSAR-X ground segment basic product specification document', *Cluster Applied Remote Sensing. TX-GS-DD-3302* .

Fritz, T. & Werninghaus, R. (2007), 'TerraSAR-X ground segment level 1b product format specification', *Clustert Applied Remote Sensing (CAF), German Aerospace Center (DLR), Tech. Rep. TX-GS-DD-3307* .

Frost, V. S., Stiles, J. A., Shanmugan, K. S. & Holtzman, J. (1982), 'A model for radar images and its application to adaptive digital filtering of multiplicative noise', *Pattern Analysis and Machine Intelligence, IEEE Transactions on* (2), 157–166.

Fua, P. (1991), Combining stereo and monocular information to compute dense depth maps that preserve depth discontinuities, *in* 'International Joint Conference on Artificial Intelligence (IJCAI)', number CVLAB-CONF-1991-001.

Fullerton, J., Leberl, F. & Marque, R. E. (1986), 'Opposite-side SAR image processing for stereo viewing', *Photogrammetric Engineering and Remote Sensing* **52**(9), 1487–1498.

Gamba, P., Houshmand, B. & Saccani, M. (2000), 'Detection and extraction of buildings from interferometric SAR data', *Geoscience and Remote Sensing, IEEE Transactions on* **38**(1), 611–617.

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. & 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.

Ghiglia, D. C. & Pritt, M. D. (1998), *Two-dimensional phase unwrapping: theory, algorithms, and software*, Wiley New York:.

Goel, K. & Adam, N. (2012), 'Three-dimensional positioning of point scatterers based on radargrammetry', *Geoscience and Remote Sensing, IEEE Transactions on* **50**(6), 2355–2363.

Goldstein, R. M. & Werner, C. L. (1998), 'Radar interferogram filtering for geophysical applications', *Geophysical Research Letters* **25**(21), 4035–4038.

Goldstein, R. M., Zebker, H. A. & Werner, C. L. (1988), 'Satellite radar interferometry: Two-dimensional phase unwrapping', *Radio Science* **23**(4), 713–720.

Gong, W., Meyer, F., Webley, P. W. & Lu, Z. (2011), Methods of INSAR atmosphere correction for volcano activity monitoring, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International', IEEE, pp. 1654–1657.

Gonzalez, J. H., Bachmann, M., Böer, J., Fiedler, H., Krieger, G. & Zink, M. (2009), 'TanDEM-X Mission and DEM Accuracy'.

Goodman, J. W. (1976), 'Some fundamental properties of speckle', *JOSA* **66**(11), 1145–1150.

Gosselin, G., Touzi, R. & Cavayas, F. (2014), 'Polarimetric Radarsat-2 wetland classification using the Touzi decomposition: case of the Lac Saint-Pierre Ramsar wetland', *Canadian Journal of Remote Sensing* **39**(6), 491–506.

Guida, R., Iodice, A. & Riccio, D. (2010), Monitoring of collapsed built-up areas with high resolution SAR images, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International', IEEE, pp. 2422–2425.

Guo, R. & Zhu, X. X. (2014), High-Rise Building Feature Extraction Using High Resolution Spotlight TanDEM-X Data, *in* 'EUSAR 2014; 10th European Conference on Synthetic Aperture Radar; Proceedings of', VDE, pp. 1–4.

Harris, F. J. (1978), 'On the use of windows for harmonic analysis with the discrete Fourier transform', *Proceedings of the IEEE* **66**(1), 51–83.

Hartigan, J. A. & Wong, M. A. (1979), 'Algorithm AS 136: A k-means clustering algorithm', *Applied statistics* pp. 100–108.

Hazewinkel, M. (2001), *Encyclopaedia of mathematics, supplement III*, Vol. 13, Springer Science & Business Media.

Hedman, K. (2010), 'Statistical fusion of multi-aspect Synthetic Aperture Radar data for automatic road extraction', *Deutsche Geodätische Kommission, Reihe C* (654).

Heliere, F., Lin, C.-C., Fois, F., Davidson, M., Thompson, A. & Bensi, P. (2009), BIOMASS: a P-band SAR earth explorer core mission candidate, *in* ‘Radar Conference, 2009 IEEE’, IEEE, pp. 1–6.

Hinz, S., Weihing, D., Suchandt, S. & Bamler, R. (2008), Detection and velocity estimation of moving vehicles in high-resolution spaceborne synthetic aperture radar data, *in* ‘Computer Vision and Pattern Recognition Workshops, 2008. CVPRW’08. IEEE Computer Society Conference on’, IEEE, pp. 1–6.

Hirschmüller, H. (2005), Accurate and efficient stereo processing by semi-global matching and mutual information, *in* ‘Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on’, Vol. 2, IEEE, pp. 807–814.

Hoffmann, J. (2007), ‘Mapping damage during the Bam (Iran) earthquake using interferometric coherence’, *International Journal of Remote Sensing* **28**(6), 1199–1216.

Hommel, M. (2010), Detektion und Klassifizierung eingestürzter Gebäude nach Katastrophenereignissen mittels Bildanalyse, PhD thesis, Karlsruher Inst. für Technologie, Diss., 2010.

Hu, J., Li, Z., Ding, X., Zhu, J., Zhang, L. & Sun, Q. (2012), ‘3D coseismic displacement of 2010 Darfield, New Zealand earthquake estimated from multi-aperture InSAR and D-InSAR measurements’, *Journal of Geodesy* **86**(11), 1029–1041.

Iqbal, M., Chen, J., Yang, W., Wang, P. & Sun, B. (2013), ‘SAR image despeckling by selective 3D filtering of multiple compressive reconstructed images’, *Progress In Electromagnetics Research* **134**, 209–226.

Ito, Y., Hosokawa, M., Lee, H. & Liu, J. G. (2000), ‘Extraction of damaged regions using SAR data and neural networks’, *International Archives of Photogrammetry and Remote Sensing* **33**(B1; PART 1), 156–163.

Janoth, J., Gantert, S., Schrage, T. & Kaptein, A. (2014), From TerraSAR-X towards TerraSAR Next Generation, *in* ‘EUSAR 2014; 10th European Conference on Synthetic Aperture Radar; Proceedings of’, VDE, pp. 1–4.

Jin, H.-y., Yang, X.-h., Jiao, L.-C. & Liu, F. (2005), A Method of Reducing Speckle Noise in SAR Image Based on Laplacian Pyramid Directional Filter Banks, *in* ‘Neural Networks and Brain, 2005. ICNN&B’05. International Conference on’, Vol. 2, IEEE, pp. 862–866.

Kaufman, L. & Rousseeuw, P. J. (2009), *Finding groups in data: an introduction to cluster analysis*, Vol. 344, John Wiley & Sons.

Khazaei, H. & Georgiadis, G. (n.d.), ‘Compressing the Laplacian Pyramid’.

Kim, Z. & Nevatia, R. (2003), ‘Expandable Bayesian networks for 3D object description from multiple views and multiple mode inputs’, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **25**(6), 769–774.

Klausing, H. & Holpp, W. (2000), ‘Radar mit realer und synthetischer Apertur’, *Verlag Oldenburg: Oldenburg* .

Krieger, G., Moreira, A., Fiedler, H., Hajnsek, I., Werner, M., Younis, M. & Zink, M. (2007), ‘TanDEM-X: A satellite formation for high-resolution SAR interferometry’, *Geoscience and Remote Sensing, IEEE Transactions on* **45**(11), 3317–3341.

Kuan, D. T., Sawchuk, A. A., Strand, T. C. & Chavel, P. (1985), ‘Adaptive noise smoothing filter for images with signal-dependent noise’, *Pattern Analysis and Machine Intelligence, IEEE Transactions on* (2), 165–177.

Kuny, S. & Schulz, K. (2014), Debris detection in SAR imagery using statistics of simulated texture, *in* ‘in Proc. of Pattern Recognition in Remote Sensing, IAPR Workshop, Stockholm, Sweden’.

Kuruoğlu, E. E. & Zerubia, J. (2004), 'Modeling SAR images with a generalization of the Rayleigh distribution', *Image Processing, IEEE Transactions on* **13**(4), 527–533.

Lachaise, M., Fritz, T. & Breit, H. (2014), InSAR processing and dual-baseline phase unwrapping for global TanDEM-X DEM generation, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International', IEEE, pp. 2229–2232.

Leberl, F. (1976), 'Mapping of lunar surface from side-looking orbital radar images', *The moon* **15**(3-4), 329–342.

Leberl, F. & Maurice, K. (1993), 'Stereo-Mapping of Planet Venus from Magellan SAR Images: A Status Report', *INTERNATIONAL ARCHIVES OF PHOTOGRAMMETRY AND REMOTE SENSING* **29**, 795–795.

Leberl, F. W. (1990), 'Radargrammetric image processing'.

Leberl, F. W., Maurice, K., Thomas, J. K. & Millot, M. (1994), 'Automated radar image matching experiment', *ISPRS journal of photogrammetry and remote sensing* **49**(3), 19–33.

Lee, J.-S. (1980), 'Digital image enhancement and noise filtering by use of local statistics', *Pattern Analysis and Machine Intelligence, IEEE Transactions on* (2), 165–168.

Lee, J.-S. (1981), 'Refined filtering of image noise using local statistics', *Computer graphics and image processing* **15**(4), 380–389.

Lee, J.-S., Jurkevich, L., Dewaele, P., Wambacq, P. & Oosterlinck, A. (1994), 'Speckle filtering of synthetic aperture radar images: A review', *Remote Sensing Reviews* **8**(4), 313–340.

Lee, J.-S., Papathanassiou, K. P., Ainsworth, T. L., Grunes, M. R. & Reigber, A. (1998), 'A new technique for noise filtering of SAR interferometric phase images', *Geoscience and Remote Sensing, IEEE Transactions on* **36**(5), 1456–1465.

Li, G., Lin, H., Li, Y., Zhang, H. & Jiang, L. (2014), Monitoring glacier flow rates dynamic of Geladandong Ice Field by SAR images Interferometry and offset tracking, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International', IEEE, pp. 4022–4025.

Li, Z., Ding, X.-L. & Liu, G. (2004), 'Modeling atmospheric effects on InSAR with meteorological and continuous GPS observations: algorithms and some test results', *Journal of Atmospheric and Solar-Terrestrial Physics* **66**(11), 907–917.

Liu, W., Yamazaki, F., Gokon, H. & Koshimura, S. (2012), Extraction of damaged buildings due to the 2011 Tohoku, Japan earthquake tsunami, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International', IEEE, pp. 4038–4041.

Lopes, A., Laur, H. & Nezry, E. (1990), Statistical distribution and texture in multilook and complex SAR images, *in* 'Geoscience and Remote Sensing Symposium, 1990. IGARSS'90.'Remote Sensing Science for the Nineties'., 10th Annual International', IEEE, pp. 2427–2430.

Lopes, A., Nezry, E., Touzi, R. & Laur, H. (1993), 'Structure detection and statistical adaptive speckle filtering in SAR images', *International Journal of Remote Sensing* **14**(9), 1735–1758.

Lowe, D. G. (1999), Object recognition from local scale-invariant features, *in* 'Computer vision, 1999. The proceedings of the seventh IEEE international conference on', Vol. 2, Ieee, pp. 1150–1157.

Mansouri, B., Shinozuka, M. & Nourjou, R. (2007), SAR remote sensing for urban damage assessment for Tehran, *in* 'Proceedings of the 5th International Workshop Remote Sensing Application Natural Hazards, Washington, DC, USA', pp. 10–11.

Marconcini, M., Marmanis, D., Esch, T. & Felbier, A. (2014), A novel method for building height estmation using TanDEM-X data, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International', IEEE, pp. 4804–4807.

Martinez, J., Beaudoin, A., Wegmuller, U. & Strozzi, T. (1998), Classification of land-cover and forest types using multidate ERS tandem data acquired over hilly terrain, *in* 'Geoscience and Remote Sensing Symposium Proceedings, 1998. IGARSS'98. 1998 IEEE International', Vol. 4, IEEE, pp. 1809–1811.

Massonnet, D., Briole, P. & Arnaud, A. (1995), 'Deflation of Mount Etna monitored by spaceborne radar interferometry', *Nature* **375**(6532), 567–570.

Matsuoka, M. & Yamazaki, F. (2000), 'Use of interferometric satellite SAR for earthquake damage detection', *Sat* **2**, z1.

Matsuoka, M. & Yamazaki, F. (2004), 'Use of satellite SAR intensity imagery for detecting building areas damaged due to earthquakes', *Earthquake Spectra* **20**(3), 975–994.

Méric, S., Fayard, F., Pottier, É. et al. (2009), 'Radargrammetric SAR image processing', *Geoscience and remote sensing* pp. 421–454.

Michaelsen, E., Thiele, A., Cadario, E. & Soergel, U. (2008), 'Building extraction based on stereo analysis of high-resolution SAR images taken from orthogonal aspect directions', *Pattern Recognition and Image Analysis* **18**(2), 231–235.

Mittermayer, J., Wollstadt, S., Prats-Iraola, P. & Scheiber, R. (2014), 'The TerraSAR-X staring spotlight mode concept'.

Mokadem, A., Thirion-Lefevre, L., Colin-Koeniguer, E., Guinvarc'h, R. & Nouvel, J.-F. (2013), Analysing urban areas in the frame of non line of sight target detection. Electromagnetic modelling, validation and application to real data in Ka-band, *in* 'Electromagnetics in Advanced Applications (ICEAA), 2013 International Conference on', IEEE, pp. 543–546.

Moreira, A. (2000), *Radar Mit Synthetischer Apertur: Grundlagen und Signalverarbeitung*, Habilitationsschrift, Universität Karlsruhe.

Moreira, A., Hajnsek, I., Krieger, G., Papathanassiou, K., Eineder, M., De Zan, F., Younis, M. & Werner, M. (2009), Tandem-L: monitoring the Earth's dynamics with InSAR and Pol-InSAR, *in* 'Proceedings International Workshop on Applications of Polarimetry and Polarimetric Interferometry (Pol-InSAR)'.

Moreira, A., Krieger, G., Hajnsek, I., Hounam, D., Werner, M., Riegger, S. & Settelmeyer, E. (2004), TanDEM-X: a TerraSAR-X add-on satellite for single-pass SAR interferometry, *in* 'Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International', Vol. 2, IEEE, pp. 1000–1003.

Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I. & Papathanassiou, K. P. (2013), 'A tutorial on synthetic aperture radar', *Geoscience and Remote Sensing Magazine, IEEE* **1**(1), 6–43.

Nascetti, A. (2013), High resolution radargrammetry: development and implementation of an innovative image matching strategy, PhD thesis, University of Rome "La Sapienza".

Nocera, L. (1996), Reconstruction du relief en imagerie radar à ouverture de synthèse: stéréoscopie et ombrages, PhD thesis.

Oriot, H. & Cantalloube, H. (2008), Circular SAR imagery for urban remote sensing, *in* 'Synthetic Aperture Radar (EUSAR), 2008 7th European Conference on', VDE, pp. 1–4.

Paar, G. & Pölzleitner, W. (1992), Robust disparity estimation in terrain modeling for spacecraft navigation, *in* 'Pattern Recognition, 1992. Vol. I. Conference A: Computer Vision and Applications, Proceedings., 11th IAPR International Conference on', IEEE, pp. 738–741.

Paris, S., Hasinoff, S. W. & Kautz, J. (2011), 'Local Laplacian filters: edge-aware image processing with a Laplacian pyramid.', *ACM Trans. Graph.* **30**(4), 68.

Perko, R., Raggam, H., Deutscher, J., Gutjahr, K. & Schardt, M. (2011), 'Forest assessment using high resolution SAR data in X-band', *Remote Sensing* **3**(4), 792–815.

Petit, D., Adragna, F. & Durou, J.-D. (2000), Filtering of layover areas in high-resolution IFSAR for building extraction, *in* 'Europto Remote Sensing', International Society for Optics and Photonics, pp. 230–240.

Polli, D., Dell'Acqua, F., Gamba, P. & Lisini, G. (2010), Earthquake damage assessment from post-event only radar satellite data, *in* 'Proceedings of the eighth international workshop on remote sensing for disaster response', Vol. 30.

Poulain, V., Inglada, J., Spigai, M., Tourneret, J.-Y. & Marthon, P. (2011), 'High-resolution optical and SAR image fusion for building database updating', *Geoscience and Remote Sensing, IEEE Transactions on* **49**(8), 2900–2910.

Prince, S. J. (2012), *Computer vision: models, learning, and inference*, Cambridge University Press.

Rabin, J., Delon, J., Gousseau, Y. & Moisan, L. (2010), MAC-RANSAC: a robust algorithm for the recognition of multiple objects, *in* 'Fifth International Symposium on 3D Data Processing, Visualization and Transmission (3DPTV 2010)', p. 051.

Rabus, B., Eineder, M., Roth, A. & Bamler, R. (2003), 'The shuttle radar topography mission-a new class of digital elevation models acquired by spaceborne radar', *ISPRS Journal of Photogrammetry and Remote Sensing* **57**(4), 241–262.

Raney, R. K. & Wessels, G. J. (1988), 'Spatial considerations in SAR speckle consideration', *Geoscience and Remote Sensing, IEEE Transactions on* **26**(5), 666–672.

Rejichi, S., Chaabane, F., Tupin, F. & Bloch, I. (2010), Morphological filtering of SAR interferometric images, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International', IEEE, pp. 1581–1584.

Romeiser, R., Johannessen, J., Chapron, B., Collard, F., Kudryavtsev, V., Runge, H. & Suchandt, S. (2010), Direct surface current field imaging from space by along-track InSAR and conventional SAR, *in* 'Oceanography from Space', Springer, pp. 73–91.

Rosen, P. A., Hensley, S., Joughin, I. R., Li, F. K., Madsen, S. N., Rodriguez, E. & Goldstein, R. M. (2000), 'Synthetic aperture radar interferometry', *Proceedings of the IEEE* **88**(3), 333–382.

Rossi, C., Eineder, M., Duque, S., Fritz, T. & Parizzi, A. (2014a), Principal slope estimation at SAR building layovers, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International', IEEE, pp. 5–8.

Rossi, C., Eineder, M. & Fritz, T. (2014b), Detecting Building Layovers in a SAR Interferometric Processor Without External References, *in* 'EUSAR 2014; 10th European Conference on Synthetic Aperture Radar; Proceedings of', VDE, pp. 1–4.

Rossi, C. & Gernhardt, S. (2013), 'Urban DEM generation, analysis and enhancements using TanDEM-X', *ISPRS Journal of Photogrammetry and Remote Sensing* **85**, 120–131.

San A, D. K. et al. (2007), 'Automatic building extraction from high resolution stereo satellite images'.

Schreier, G. (1993), *SAR geocoding: data and systems*, Wichmann.

Schubert, A. (2004), *Stereo-assisted interferometric SAR*, Remote Sensing Laboratories, Dept. of Geography, University of Zurich.

Schubert, A., Faes, A., Kääb, A. & Meier, E. (2013), ‘Glacier surface velocity estimation using repeat TerraSAR-X images: Wavelet-vs. correlation-based image matching’, *ISPRS Journal of Photogrammetry and Remote Sensing* **82**, 49–62.

Schubert, A., Small, D., Holecz, F., Meier, E. & Nüesch, D. (2002), Extraction of surface topography from SAR stereo pairs using an airborne X-band sensor: preliminary results, *in* ‘Retrieval of Bio-and Geo-Physical Parameters from SAR Data for Land Applications’, Vol. 475, pp. 151–156.

Seguin, G., Srivastava, S. & Auger, D. (2014), ‘Evolution of the RADARSAT Program’, *Geoscience and Remote Sensing Magazine, IEEE* **2**(2), 56–58.

Sheng, Y. & Xia, Z.-G. (1996), A comprehensive evaluation of filters for radar speckle suppression, *in* ‘Geoscience and Remote Sensing Symposium, 1996. IGARSS’96.’Remote Sensing for a Sustainable Future.’, International’, Vol. 3, IEEE, pp. 1559–1561.

Shinozuka, M., Ghanem, R., Houshmand, B. & Mansouri, B. (2000), ‘Damage detection in urban areas by SAR imagery’, *Journal of Engineering Mechanics* **126**(7), 769–777.

Simonetto, E. (2002), Extraction 3D de structures industrielles sur des images RAMSES haute résolution par radargrammetrie, PhD thesis, Université de Rennes.

Simonetto, E., Oriot, H. & Garello, R. (2005), ‘Rectangular building extraction from stereoscopic airborne radar images’, *Geoscience and Remote Sensing, IEEE Transactions on* **43**(10), 2386–2395.

Singh, G., Yamaguchi, Y., Boerner, W.-M. & Park, S.-E. (2013), ‘Monitoring of the March 11, 2011, off-Tohoku 9.0 Earthquake with super-Tsunami disaster by implementing fully polarimetric high-resolution POLSAR techniques’, *Proceedings of the IEEE* **101**(3), 831–846.

Small, D. & Schubert, A. (2008), ‘Guide to ASAR geocoding’, *Issue* **1**(19.03), 2008.

Sörgel, U. (2003), Iterative Verfahren zur Detektion und Rekonstruktion von Gebäuden in SAR-und InSAR-Daten, PhD thesis, Fakultät für Elektrotechnik und Informationstechnik, Leibniz Universität Hannover.

Sörgel, U., Michaelsen, E., Thiele, A., Cadario, E. & Thoennessen, U. (2009), ‘Stereo analysis of high-resolution SAR images for building height estimation in cases of orthogonal aspect directions’, *ISPRS Journal of Photogrammetry and Remote Sensing* **64**(5), 490–500.

Stramondo, S., Bignami, C., Chini, M., Pierdicca, N. & Tertulliani, A. (2006), ‘Satellite radar and optical remote sensing for earthquake damage detection: results from different case studies’, *International Journal of Remote Sensing* **27**(20), 4433–4447.

Strehl, A. (2002), ‘Relationship-based Clustering and Cluster Ensembles for High-dimensional Data Mining’.

Strozzi, T., Dammert, P. B., Wegmuller, U., Martinez, J.-M., Askne, J. I., Beaudoin, A. & Hallikainen, M. T. (2000), ‘Landuse mapping with ERS SAR interferometry’, *Geoscience and Remote Sensing, IEEE Transactions on* **38**(2), 766–775.

Tao, J., Auer, S., Reinartz, P. & Bamler, R. (2013), Object-based change detection for individual buildings in SAR images captured with different incidence angles, *in* ‘Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International’, IEEE, pp. 1238–1241.

Tao, J., Palubinskas, G., Reinartz, P. & Auer, S. (2011), Interpretation of SAR images in urban areas using simulated optical and radar images, *in* 'Urban Remote Sensing Event (JURSE), 2011 Joint', IEEE, pp. 41–44.

Thiele, A. (2014), *3d Building Reconstruction from High-resolution Multi-aspect Lnterferometric Synthetic Aperture Radar Data*, Bayerische Akademie der Wissenschaften.

Thiele, A., Cadario, E., Schulz, K., Thoennessen, U. & Soergel, U. (2007a), 'InSAR phase profiles at building locations', *Proceeding of ISPRS photogrammetric image analysis* **36**(part 3), W49A.

Thiele, A., Cadario, E., Schulz, K., Thonnessen, U. & Soergel, U. (2007b), 'Building recognition from multi-aspect high-resolution InSAR data in urban areas', *Geoscience and Remote Sensing, IEEE Transactions on* **45**(11), 3583–3593.

Thiele, A., Wurth, M., Even, M. & Hinz, S. (2013), 'Extraction of building shape from TanDEM-X data', *Proc. Int. Archives Photogramm., Remote Sens. Spatial Inf. Sci* **1**, W1.

Tison, C., Tupin, F. & Maitre, H. (2004), Retrieval of building shapes from shadows in high resolution SAR interferometric images, *in* 'Geoscience and Remote Sensing Symposium, 2004. IGARSS'04. Proceedings. 2004 IEEE International', Vol. 3, IEEE, pp. 1788–1791.

Toutin, T., Blondel, E., Clavet, D. & Schmitt, C. (2013), 'Stereo Radargrammetry With Radarsat-2 in the Canadian Arctic', *Geoscience and Remote Sensing, IEEE Transactions on* **51**(5), 2601–2609.

Toutin, T. & Gray, L. (2000), 'State-of-the-art of elevation extraction from satellite SAR data', *ISPRS Journal of Photogrammetry and Remote Sensing* **55**(1), 13–33.

Trianni, G. & Gamba, P. (2008), 'Damage detection from SAR imagery: Application to the 2003 Algeria and 2007 Peru earthquakes', *International Journal of Navigation and Observation* **2008**.

Tupin, F., Bloch, I. & Maître, H. (1999), 'A first step toward automatic interpretation of SAR images using evidential fusion of several structure detectors', *Geoscience and Remote Sensing, IEEE Transactions on* **37**(3), 1327–1343.

Tupin, F. & Nicolas, J. (2002), 'Matching criteria for radargrammetry', *Geoscience and Remote Sensing, IGARSS'02, Toronto, Canada* **5**, 2608–2610.

Upret, P. & Yamazak, F. (2012), Use of high-resolution SAR intensity images for damage detection from the 2010 Haiti earthquake, *in* 'Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International', IEEE, pp. 6829–6832.

Vasile, G., Trouvé, E., Ciuc, M. & Buzuloiu, V. (2004), 'General adaptive-neighborhood technique for improving synthetic aperture radar interferometric coherence estimation', *JOSA A* **21**(8), 1455–1464.

Vasile, G., Trouvé, E., Lee, J.-S. & Buzuloiu, V. (2006), 'Intensity-driven adaptive-neighborhood technique for polarimetric and interferometric SAR parameters estimation', *Geoscience and Remote Sensing, IEEE Transactions on* **44**(6), 1609–1621.

Wang, C., Zhang, H., Wu, F., Zhang, B., Tang, Y., Wu, H., Wen, X. & Yan, D. (2009), 'Disaster phenomena of Wenchuan earthquake in high resolution airborne synthetic aperture radar images', *Journal of Applied Remote Sensing* **3**(1), 031690–031690.

Wang, T.-L. & Jin, Y.-Q. (2012), 'Postearthquake building damage assessment using multi-mutual information from pre-event optical image and postevent SAR image', *Geoscience and Remote Sensing Letters, IEEE* **9**(3), 452–456.

Wegner, J. D., Ziehn, J. R. & Soergel, U. (2014), 'Combining High-Resolution Optical and InSAR Features for Height Estimation of Buildings With Flat Roofs', *Geoscience and Remote Sensing, IEEE Transactions on* **52**(9), 5840–5854.

Yonezawa, C. & Takeuchi, S. (2001), ‘Decorrelation of SAR data by urban damages caused by the 1995 Hyogoken-nanbu earthquake’, *International Journal of Remote Sensing* **22**(8), 1585–1600.

Yonezawa, C., Tomiyama, N. & Takeuchi, S. (2002), Urban damage detection using decorrelation of SAR interferometric data, *in* ‘Geoscience and Remote Sensing Symposium, 2002. IGARSS’02. 2002 IEEE International’, Vol. 4, IEEE, pp. 2051–2053.

Zhu, X. X. & Bamler, R. (2010), ‘Very high resolution spaceborne SAR tomography in urban environment’, *Geoscience and Remote Sensing, IEEE Transactions on* **48**(12), 4296–4308.

Zhu, X. X. & Shahzad, M. (2014), ‘Façade reconstruction using multiview spaceborne TomoSAR point clouds’, *Geoscience and Remote Sensing, IEEE Transactions on* **52**(6), 3541–3552.

Zink, M., Krieger, G., Fiedler, H., Hajnsek, I. & Moreira, A. (2008), The TanDEM-X mission concept, *in* ‘Synthetic Aperture Radar (EUSAR), 2008 7th European Conference on’, VDE, pp. 1–4.

Acknowledgments

Foremost, I would like to express my sincere gratitude to my advisor Prof. Dr.-Ing. Stefan Hinz, whose expertise, enthusiasm and immense knowledge were a considerable support for my Ph.D study and research.

A particular thank goes also to my second advisor, Prof. Dr.-Ing. Uwe Sörgel, for accepting the supervision of my work in a short term, and for his insightful comments that helped and encourage me for the final stretch.

Besides my advisors, I would like to thank the rest of my thesis committee: Prof. Dr.-Ing. Dr. h.c. Bernard Heck and PD Dr.-Ing. Boris Jutzi for their critical suggestions and encouragement.

My sincere thanks also go to Dr.-Ing. Stefan Auer, whose questions and understanding yielded new ideas and development after each meeting.

I would also like to express my very grateful thanks to my friend, mentor and colleague Dr.-Ing. Antje Thiele, whose continuous support and guidance are the reasons why I first envisaged a PhD study and the rocks that contributed to my perseverance during difficult moments.

Furthermore, I would like to thank the GRAduate school of Climate and Environment (GRACE) of the KIT for encouraging and supporting my further education and facilitating my networking through the founding of different conference visits and a stay abroad. All the same, I am grateful to the Karlsruhe House of Young Scientist (KHYS) of the KIT for founding a further stay abroad, which was very helpful for the completion of my thesis. The founded stays abroad were for me not only the opportunity to have new ideas and create new projects with the partner institute, but also were a great human experience.

To this sense, I would also like to thank Prof. Mattia Crepi and Dr. Andrea Nascetti for offering me the possibility to stay twice at the Institute of Geodesy and Geomatics of the University of Rome ‚La Sapienza‘. Both stays were very fruitful on the technical point of view, and I wish that our collaboration would continue further after my dissertation. A particular thank goes to Elisa Benedetti, Mara Branzati, Martina Di Rita and Carla Garagnani for their help, support and friendship that made my stays unforgettable.

A big thank goes also to my fellow colleagues of the Institute of Photogrammetry and Remote Sensing (IPF) of the KIT for stimulating discussions, recommendations, and for all the fun we have had in the last four years. I would like to address a particular thank to Dr.-Ing. Jens Leitloff, Dr.-Ing. Martin Weinmann and Dipl.-Ing. Steffen Urban for their proofreading of my thesis, and to M.A. Elis Engelmann for her help with all the administrative burden scholarship student are confronted for.

Last but not least, I would like to thank my friends and family for their continuous support, that started long before my Ph.D. and will continue even after. I am very grateful to have such a comprehensive family that forget all possible mood swings and always find encouraging words: to this sense: Merci Maman, Nanou and Danke Steffen!