Geodetic Synthetic Aperture Radar Interferometry

Sina Montazeri

Vollständiger Abdruck der von der Ingenieurfakultät Bau Geo Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender:
    Prof. Dr.-Ing. habil. Richard Bamler

Prüfende der Dissertation:
    1. Prof. Dr.-Ing. habil. Xiao Xiang Zhu
    2. Hon.-Prof. Dr. rer. nat. Michael Eineder
    3. Prof. Dr.-Ing. Uwe Sörgel,
       Universität Stuttgart

Die Dissertation wurde am 29.01.2019 bei der Technischen Universität München eingereicht und durch die Ingenieurfakultät Bau Geo Umwelt am 08.05.2019 angenommen.

Abstract

Spaceborne Interferometric Synthetic Aperture Radar (InSAR) is the technique of using two or more SAR images of the same scene to extract topography and deformation of the Earth's surface with high precision. The method has been considerably evolved in the last decades into an operational tool, which has contributed to advancements in the study of volcanoes, glaciers, forests and urban environments, to name a few examples. Although multi-temporal and multi-baseline methods such as Persistent Scatterer Interferometry (PSI) and SAR Tomography (TomoSAR) have been developed to tackle main shortcomings of conventional InSAR, they still suffer from the inherent limitations of SAR side-looking imaging, including the one-dimensional (1-D) deformation retrieval and the shadow phenomenon, as well as the relative nature of InSAR topographic and displacement measurements.

In the recent years, the launch of the German TerraSAR-X (TSX) satellite, that is characterized with high precision orbit determination and high spatial resolution, has led to the rise of new SAR geodetic applications. By applying a meticulous correction scheme to TSX timing measurements, SAR imaging geodesy allows the absolute two-dimensional (2-D) localization of bright point targets in SAR images with centimeter accuracy. The three-dimensional (3-D) absolute positioning has been also demonstrated through using stereo SAR. Nevertheless, the difficulty of identical target detection from multi-aspect data, which is the prerequisite for stereo SAR, and the fact that the precision of these methods are in general lower than phase-based methods due to the exploitation of amplitude information only, limit their ubiquitous applicability.

This thesis is dedicated to merging the capabilities of InSAR and SAR geodetic techniques, a framework referred to as *geodetic InSAR*. The ultimate goal is to overcome the pitfalls of both methods and produce detailed 3-D InSAR point clouds with absolute coordinates, with a focus on urban infrastructure monitoring.

Three approaches based on fusion of multi-track PSI point clouds, high resolution optical data and vector road network data are proposed for the *detection of identical point targets from multi-aspect SAR images*. Although the first technique is suitable for targets visible in a same-heading configuration, the latter two are able to detect targets visible in cross-heading configurations i.e., from orbits with opposite heading directions. This enhances the precision of the retrieved absolute 3-D coordinates using stereo SAR especially in the height component. With the aforementioned detection procedures thousands of Persistent Scatterer (PS) can be absolutely localized in urban areas, which has paved the way for *automatic SAR-based Ground Control Point (GCP) generation*. A practical demonstration of the GCP extraction framework is shown in this work using two pairs of cross-heading data over Oulu, Finland.

Abstract

The absolute height of the SAR GCPs and the radar timing corrections provided by the imaging geodesy method are exploited to introduce an approach for *geocoding error correction of InSAR point clouds*. The proposed method is flexible in the sense that it can be either used prior to the reference point selection in any InSAR stacking techniques or it can be utilized as a post-processing step after the generation and the geocoding of point clouds. Applications of this technique in both mentioned scenarios are shown for high resolution spotlight PSI and TomoSAR, where the latter is termed *geodetic SAR tomography*. Strategies have been also developed to evaluate the positioning accuracy of the produced *absolute InSAR point clouds* with respect to high accuracy Digital Terrain Model (DTM)s of urban areas. Horizontal and vertical accuracy of 20 cm and 10 cm has been achieved using PSI and TomoSAR processing over Berlin, respectively.

Multi-aspect absolute InSAR point clouds are used for the first demonstration of *geodetic point cloud fusion*, where no external positioning data is required for the coregistration and tying the point clouds to a geodetic reference frame. As the final contribution, an algorithm has been developed for the task of *1-D to 3-D motion decomposition* using fused TomoSAR point clouds. The technique is robust with respect to the high number of outliers available in TomoSAR estimates and allows for highly detailed deformation monitoring of urban areas using high resolution SAR data.

Zusammenfassung

Satellitengestütze Interferometriesches Synthetic Aperture Radar (InSAR) ist eine Technik die zwei oder mehr SAR Aufnahmen der selben Szene nutzt, um die Topographie und Deformation der Erdoberfläche mit hoher Genauigkeit abzuleiten. Die Methode wurde in den letzten Jahrzehnten zu einem operationellen Werkzeug weiterentwickelt und hat damit zu wissenschaftlichen Fortschritten in Feldern wie Vulkanologie, Glaziologie sowie der Beobachtung von Wäldern und urbanisierten Räumen beigetragen. Um Schwächen der konventionellen Radarinterferometrie zu beheben, wurden mit Persistent Scatterer Interferometry (PSI) und SAR Tomographie (TomoSAR) multi-temporale Methoden und Methoden mit mehreren Basislinien entwickelt. Diese haben jedoch immer noch inhärente Einschränkungen die aus der seitlichen Aufnahmegemoetrie des SAR Instruments stammen. Sie beinhalten die eindimensionale (1-D) Deformationsbestimmung, Phänomene des Radar Schattens sowie die relative Natur der InSAR Topographie und Deformationsmessungen.

Zuletzt hat die Inbetriebnahme des deutschen TerraSAR-X (TSX) Satelliten zu neuen geodätischen SAR Applikationen geführt, da dessen Orbit mit hoher Genauigkeit beschrieben werden kann und die Aufnahmen eine hohe räumliche Auflösung besitzen. Durch die Anwendung einer akkuraten Korrektur der TSX Zeitmessungen ist es möglich mit SAR gestützter Geodäsie, stark rückstreuende Punkte zwei-dimensional (2-D) mit Zentimetergenauigkeit zu vermessen. Die absolute drei-dimensionale (3-D) Lokalisierung kann durch die Benutzung der stereo SAR Technik erreicht werden. Voraussetzung dafür ist jedoch die schwierige Identifizierung von geeigneten Punkten aus mehreren Aufnahmen mit unterschiedlichen Winkeln. Das und die Erkenntnis, dass die Genauigkeit dieser amplitudengestützen Technik niedriger ist als die von phasenbasierten Methoden schränken die Benutzbarkeit allerdings ein.

Diese Arbeit beschäftigt sich mit der Verschmelzung von InSAR und geodätischen SAR Techniken, was wir folgend *geodätische Radarinterferometrie* nennen. Das Ziel ist es, die Schwächen beider Methoden auszugleichen, um 3-D InSAR Punktwolken mit absoluten Koordinaten zu generieren. Der Fokus liegt dabei auf der Beobachtung von urbaner Infrastruktur.

Es werden drei Herangehensweisen für die Erkennung von identischen Punktzielen aus mehreren Aufnahmen mit unterschiedlichen Winkeln vorgestellt. Diese basieren auf der Fusion von multi-track PSI Punktwolken, hochauflösenden optischen Daten und vektorisierten Daten des Straßennetzwerks. Dabei ist die erste Technik auf Ziele anwendbar, die auf Aufnahmen aus gleicher Satellitenflugrichtung zu erkennen sind. Mit den anderen zwei Methoden können Ziele in aufsteigender und absteigender Flugrichtung erkannt werden. So kann die Genauigkeit der mit stereo SAR abgeleiteten 3-D Koordinaten erhöht werden. Dabei wird speziell die Komponente der Höhe

verbessert. Mit den hier vorgestellten Erkennungstechniken können so tausende Persistent Scatterer (PS) in urbanen Räumen absolut positioniert werden. Damit ist die *Erstellung von SAR-basierten Punkten mit bekannten Koordinaten (sog. Ground Control Points (GCP))* möglich. In dieser Arbeit wird als praktisches Beispiel die Erstellung eines GCP Netzwerks in Oulu, Finnland gezeigt. Dazu wurden zwei Paare von SAR Aufnahmen aus unterschiedlicher Flugrichtung benutzt.

Die absolute Höhe der SAR GCPs und die Korrektur der Radar Zeitmessungen aus der SAR-gestützten Geodäsie werden benutzt um *Fehler in der Georeferenzierung von In-SAR Punktwolken zu korrigieren.* Die vorgestellte Methode kann dabei bereits vor der In-SAR Koregistrierung zur Bestimmung eines Referenzpunkts flexibel eingesetzt werden, oder der Nachprozessierung der georeferenzierten Punktwolken dienen. Hier zeigen wir die Anwendung dieser Technik in beiden Szenarien anhand hochauflösender spotlight PSI Prozessierung und TomoSAR Aufnahmen von Berlin.

Geodätische Fusion von Punktwolken wird anhand von InSAR Punktwolken demonstriert, die aus Szenen mit unterschiedlicher Aufnamegeometrie stammen. Dazu sind weder für die Koregistrierung noch zur Einhängung in eine geodätische Referenz externe Positionierungsdaten notwendig. Zuletzt, wird ein Algorithmus vorgestellt, der TomoSAR Punktwolken benutzt um *1-D in 3-D Bewegungen zu zerlegen.* Diese Technik ist stabil mit Bezug auf die hohe Anzahl von Ausreißern in den TomoSAR Schätzungen und macht es somit möglich mit hochauflösenden SAR Daten eine detaillierte Überwachung der Deformationen in urbanen Räumen durchzuführen.

Acknowledgments

In the past four years, many people have influenced my professional and personal life. The fact that I've received a PhD degree after all this effort is a wonderful thing but the memories of working and spending time with the colleagues and the friends I've made during this time is something I will cherish for a much longer time.

First and foremost, I would like to thank my first advisor Prof. Xiao Xiang Zhu. The journey under her supervision started with an internship in 2012, continued with a master thesis in 2014 and was finally concluded with a PhD dissertation in 2019. In the past six and a half years, she has guided me step by step through my research and has given me a high degree of independence, which has allowed me to collaborate with many people at DLR and at TUM. I am really grateful to her for all her support on scientific and administrative matters. I would also like to thank my second advisor Prof. Michael Eineder. From the starting phases of the PhD, he included me in all the geodetic SAR internal meetings and helped me to never feel as an outsider among the imaging geodesy experts. Michael provided me with many valuable inputs for my research and his attention to my work was very encouraging. I would also like to express my gratitude toward the remaining members of my examination committee Prof. Richard Bamler and Prof. Uwe Sörgel who guaranteed a smooth defense procedure.

I've had the honor to work alongside many excellent researchers at DLR in the last four years. I am indebted to Christoph Gisinger who has a great share in this dissertation. Our collaborations led to the publication of two journal papers included in this work. Christoph is without the doubt one of the most organized and thorough researchers I've ever worked with. His work in stereo SAR has played an integral part in the proposed method of geodetic InSAR. I would like to deeply thank Fernando Rodríguez González for helping me in many stages of my work. He spent an enormous amount of time supporting me in InSAR and PSI processing. Fernando is particularly gifted in explaining complex issues in an extremely easy way with his extraordinary domain knowledge. A large part of this PhD would have not been possible without him and his guidance. I would like to also thank Dr. Ulrich Balss for his patience and support on using the SGP software. I've learned a lot from his thorough explanations on basic and complex SAR concepts. I would like to thank my colleague Nan Ge for his help on various aspects of my PhD being from data ordering and processing to basic theories of SAR. I've always admired his way of research and his broad knowledge. The same level of gratitude goes to Dr. Yuanyuan Wang who was always ready to help a PhD student in trouble. I also thank Dr. Stefan Auer who acted as my mentor at DLR and was always interested in my research. I would like to further thank Bettina Bierkamp-Michalak from DLR, Dr. Annette Spengler and Dr. Michael Schmitt from TUM for their support during the course of my PhD. Michael in particular took care

Contents

Acronyms

1-D one-dimension

2-D two-dimensions

3-D three-dimensions

ADI Amplitude Dispersion Index

CR Corner Reflector

DEM Digital Elevation Model

DLR German Aerospace Center

DSM Digital Surface Model

GCP Ground Control Point

GIS Geographic Information System

GNSS Global Navigation Satellite System

IERS International Earth Rotation and Reference Systems Service

InSAR Interferometric SAR

LiDAR Light Detection and Ranging

LOS Line of Sight

MAD Median Absolute Deviation

MD Maximum Detection

NCC Normalized Cross-Correlation

PRF Pulse Repetition Frequency

PS Persistent Scatterer

PSI Persistent Scatterer Interferometry

PTA Point Target Analysis

RSF Range Sampling Frequency

SAR Synthetic Aperture Radar

SCR Signal-to-Clutter-Ratio

SGP SAR Geodesy Processor

SIFT Scale Invariant Feature Transform

SLC Single-Look Complex

SVD Singular Value Decomposition

TEC Total Electron Content

TMSP TerraSAR-X multimode SAR processor

TomoSAR Tomographic SAR

TSX TerraSAR-X

UTM Universal Transverse Mercator

VHR Very High Resolution

1 Introduction

1.1 Background and Motivation

Over the last three decades, Synthetic Aperture Radar (SAR) remote sensing has evolved into a major Earth observation technique. As a result of its coherent imaging nature, Interferometric SAR (InSAR) exploits the phase differences of two SAR images of the same scene to approximate the surface of the Earth in three-dimensions (3-D) [1]–[4]. In its differential form, InSAR extracts ground deformation from at least two SAR images that are temporally separated [5]. The high achievable precision in displacement retrieval, typically in the centimeter to millimeter regimes, as well as unique characteristics, such as high spatial resolution and large coverage, distinguish InSAR from its other remote sensing counterparts.

The launch of high resolution SAR sensors such as the German TerraSAR-X (TSX) and the Italian COSMO-Skymed in the past decade and multi-temporal and multi-baseline extensions of InSAR, such as the Persistent Scatterer Interferometry (PSI) [6], [7] and the Tomographic SAR (TomoSAR) [8]–[10] techniques, have further advanced the information retrieval from SAR data into an unprecedented level. Using these methods, it is nowadays possible to perform highly detailed 3-D mapping and to precisely track deformation of individual buildings in an entire urban area [11], [12]. However, all InSAR approaches are *relative* techniques, meaning that the height and the deformation parameters are evaluated with respect to an arbitrarily chosen reference point. For many cases no information about the 3-D position and the stability of the reference point is available. The systematic error caused due to the uncertainty of the coordinates of the reference point degrades the geo-localization accuracy of InSAR products. As a consequence the true source of deformation may not be correctly detected, which results in the misinterpretation of InSAR deformation maps. This is in particular critical for monitoring small-scale displacements in urban areas using high resolution SAR images for which the detailed structure of objects are acquired. Furthermore, InSAR products with low localization accuracy cannot be easily compared to or integrated into data from other sensors. Therefore, in order to obtain *absolute* coordinates and deformations, external data are required to link the InSAR estimates to a reference geodetic datum. Apart from the problem of relative estimates, single-aspect InSAR maps suffer from the main shortcomings caused by the SAR side-looking geometry, such as the retrieval of deformation in only one-dimension (1-D), the shadow effect and the layover phenomenon. This issue can be tackled by accurate multi-aspect coregistration of InSAR products from sufficiently separated viewing geometries.

It has been recently shown that aside from their imaging capabilities, modern SAR sensors, in particular TSX and TanDEM-X, have great potentials for achieving extraor-

dinary geometric accuracy [13], [14]. This can be partly attributed to the highly precise orbit determination of these satellites [15], [16] and, more importantly, to a comprehensive and accurate correction methodology aiming at mitigating SAR timing errors. For bright well-detectable point targets in a SAR image, the method, called *SAR imaging geodesy* [13], [17], provides absolute pixel localization accuracy in two-dimensions (2-D) within the centimeter and sub-centimeter regimes [18]. Additionally, the combination of 2-D absolute timing measurements of an identical target visible in SAR images from different viewing geometries leads to the target's absolute 3-D position through *stereo SAR* [19]. This has paved the way for SAR-based GCP generation. The tie-point selection step of stereo SAR is nevertheless a non-trivial task, especially when dealing with images from opposite geometries (cross-heading tracks). This limits the applicability of the method for localization of large number of scatterers.

In this thesis, strategies are proposed in order to overcome the aforementioned pitfalls of the *imaging* and the *localization* aspects of SAR. The research objectives of the thesis are introduced in Section 1.2 and finally the chapter is closed with outlining the structure of this thesis in Section 1.3.

1.2 Research objectives

The primary objective of this thesis is to explore the possibilities of merging the capabilities of both phase-based SAR techniques like conventional InSAR, PSI and TomoSAR and SAR geodetic techniques such as imaging geodesy and stereo SAR. We therefore aim at producing detailed multi-dimensional maps with absolute coordinates and identify applications, which can benefit from this improvement in the localization accuracy of InSAR products. To this end, the objectives of this work can be categorized as follows:

- **Automatic SAR GCP generation**: A processing chain for automatic detection and positioning of natural point scatterers to serve as SAR GCPs is desired. As the detection of identical scatterers solely based on SAR data is highly difficult, auxiliary optical or Geographic Information System (GIS) data are to be used to aid this process.

- **Absolute positioning of InSAR point clouds**: A framework should be proposed to exploit the absolute coordinates of the SAR-based GCPs to assist the relative InSAR-based methods. The methodology should be tested on the most common InSAR approaches for urban monitoring such as PSI and TomoSAR. Strategies should be introduced to evaluate the improvement in the localization accuracy of InSAR point clouds with respect to more accurate Digital Surface Model (DSM)s from other sensors like Light Detection and Ranging (LiDAR).

- **Geodetic multi-aspect InSAR point cloud fusion**: The availability of multiple absolute InSAR point clouds reconstructed from multi-aspect SAR data i.e., images from different orbit tracks, leads to substantial increase in the density of the

point clouds and will capture the entire extent of buildings if cross-heading tracks are used. Therefore, geodetic InSAR point cloud fusion approaches should be employed to make the coregistration of point clouds more accurate and objective than the currently available geometric methods.

- **3-D InSAR deformation monitoring**: A scheme should be designed to use fused absolute point clouds to reconstruct the full 3-D displacement vector of each scatterer from its observed 1-D deformation measurement. The method should robustly do so by inclusion of observations from all the available viewing geometries.

The objectives are defined in a way that the results of each step provide the input for the succeeding step. Each of the mentioned items will be treated in detail in Chapter 4.

1.3 Thesis Outline

This is a cumulative thesis. The main contributions of the author have been summarized in five peer-reviewed journal publications (three as the first-author and two as a co-author), which can be found in the appendices.

The introduction and the motivation of the dissertation have been already addressed in this chapter. Chapter 2 briefly reviews the basics of SAR and InSAR that are relevant for understanding this thesis and discusses the specific limitations of both methods, which are tackled in this work. Chapter 3 gives a comprehensive overview of the state-of-the-art of the methods used in this work and the summary of the contributions of the author is reported in Chapter 4. Finally, the conclusions as well as future research directions are discussed in Chapter 5.

2 Fundamentals

This chapter reviews the relevant basics of SAR and InSAR. Special emphasize is put on the inherent limitations of these methods, which are intended to be tackled by this thesis. The chapter also gives a quick recap on fundamentals of observation theory. For an in depth treatment of SAR data processing and image generation the reader is referred to [20]. The theoretical aspects of InSAR and its processing sequence are described in detail in [2]–[4].

2.1 SAR Imaging and Localization

Spaceborne SAR sensors use a side-looking geometry to map the Earth's surface by the *coherent* transmission of microwave pulses and the reception of the backscattered signals. The echoes are arranged in a 2-D matrix, where each echo is characterized with a *slant-range* coordinate (two-way pulse travel time) and an *azimuth* coordinate (the position along the sensor trajectory) [2]. The resolution in the slant-range direction ρ_r is a function of the pulse bandwidth β [20], [21]:

$$\rho_r = \frac{c}{2\beta},$$ (2.1)

where c denotes the speed of light in vacuum. The azimuth resolution, for a stripmap SAR, can be approximated as [21]:

$$\rho_a \approx \frac{d_a}{2},$$ (2.2)

which depends on the physical length of the radar antenna d_a. The aforementioned 2-D matrix is a raw SAR image, which requires extensive signal processing, known as SAR focusing, to achieve a more image-like product called a Single-Look Complex (SLC) image [2], [4], [20]. After the focusing, the azimuth time of each target is registered to a reference, commonly chosen as the time of the closest approach or the zero-Doppler time [20]. This compressed azimuth time t and the travel time of the transmitted and the received pulse τ denote the radar timing coordinates of each SLC pixel. In the SAR image space, the ith pixel is marked with a pixel number p_i in slant-range and with a line number l_i in azimuth, which are related to the aforementioned timing coordinates (t_i, τ_i) as:

$$l_i = PRF \cdot (t_i - t_0),$$ (2.3)

$$p_i = RSF \cdot (\tau_i - \tau_0),$$ (2.4)

where t_0 and τ_0 are the timings of the first imaged pixels in azimuth and range directions, and *PRF* and *RSF* denote the Pulse Repetition Frequency (PRF) and Range Sampling Frequency (RSF), which govern the pixel spacing in azimuth and range, respectively. The pixel located at (l_i, p_i) has a complex value, which can be written as [2]:

$$g(l_i, p_i) = A \cdot \exp(j\phi_i). \tag{2.5}$$

In Equation 2.5, the variable A denotes the amplitude of the signal, which is related to the intensity $I = A^2$ and therefore contains information about the brightness of the pixel. The term ϕ_i denotes the phase and includes contributions due to the distance of the sensor to the target, scattering mechanism, atmospheric delays and noise [4], [21].

The SAR products used in this thesis are exclusively from the Very High Resolution (VHR) spotlight mode of the TSX sensor [22]. As opposed to the conventional stripmap mode and modes designed for large swath coverage such as scanSAR [20] and TOPS [23], in the spotlight mode the antenna illumination time is increased to achieve high azimuth resolution with the cost of small and discontinuous swath coverage. The TSX spotlight images have a nominal spatial resolution of 1.1 m in azimuth and 0.6 m in range, when operating with $\beta = 300$ MHz. This level of resolution makes the spotlight images one of the most appropriate products for urban monitoring [24], which is the main focus of this work.

2.1.1 Limitations and Errors of SAR

Since SAR images provide a 2-D projection of the real 3-D world, acquired with a side-looking geometry, they introduce unpreventable *geometrical* distortions. These include the well-known forshortening, layover and shadow effects. Layover causes multiple scatterers, with equal slant-range coordinates, to be mapped into one resolution cell. This phenomenon is quite prevalent in high rise urban areas where backscattered contributions from the ground and the facade (sometimes also from the roof) of buildings cannot be distinguished in one pixel. The shadow in SAR images occurs when an area is not illuminated by the radar beam. Therefore no backscattered signal is received and the area in the image appears to be completely dark. These distortions as well as the fact that single SAR images cannot access the third dimension (the elevation angle) limit the interpretation and usability of SAR in urban areas.

Apart from the aforementioned geometrical peculiarities, the 2-D localization of targets using SAR can be biased. This bias may originate from actual errors in SAR timing measurements, such as a drift in SAR internal clock rate, or from indirect effects including propagation delays and position offsets caused by geodynamic phenomena. Hereafter both direct and indirect systematic effects on SAR measurements are referred to as *timing errors*. The origin of these errors and their physical properties will be later described in Section 3.1. For now it is important to mention that these timing errors are translated into 2-D shifts, affecting both t and τ, and thus change the position of targets within SAR images.

2.1.2 Tackling SAR Limitations: an Outlook

It has been mentioned that SAR images can only localize targets in 2-D and they include image distortions, such as layover and shadow, due to the side-looking nature of their acquisition geometry. Furthermore, the timing errors bias the 2-D position of targets in SAR images.

For 3-D reconstruction from SAR images, InSAR approaches can be employed (see Section 2.2). To mitigate the timing errors and obtain absolute 2-D SAR measurements, SAR imaging geodesy [13] can be used as will be described in Section 3.1. Layover separation can be tackled by SAR tomographic methods [8], [9], [12] as will be touched upon later in Section 3.4. In order to minimize shadowed areas, the combination of SAR images from opposite viewing geometries is required. This includes the coregistration of InSAR point clouds from different viewing geometries [25], [26], also known as point cloud fusion, as will be discussed in Section 3.5.

2.2 Interferometric Synthetic Aperture Radar (InSAR)

The phase of only one SLC cannot be directly used to extract information from SAR images. The interferometric phase $\Delta\phi$, as the phase difference between two complex-valued SAR measurements from different orbit positions and/or from different times, however, contains highly precise information about the topography and/or the deformation of the mapped area [3], [4]. Considering a *repeat-pass* scenario, assuming the scattering phase contributions between the two acquisitions cancel out, and ignoring the atmospheric, orbital and noise phase components, the wrapped interferometric phase ($-\pi < \Delta\phi_i \leq \pi$) at pixel ($l_i, p_i$) is caused solely by the range difference between the first and the second acquisitions and reads as [4]:

$$\Delta\phi_i = \frac{-4\pi}{\lambda}\left(-B_i^{\parallel} + \frac{B_i^{\perp}}{R_i \sin\theta i} \cdot h_i + d_i\right) - a_i \cdot 2\pi, \tag{2.6}$$

where λ is the radar wavelength. The three terms inside the parenthesis represent range difference components caused by flat earth, topography above the reference ellipsoid, and the displacement caused between the two acquisition times for the ith pixel, respectively. The variables $B_i^{\parallel}$ and $B_i^{\perp}$ are the parallel and the perpendicular components of the separation between the two satellite positions, known as *baseline*, R_i is the geometric distance from the center of the pixel to the satellite antenna, θ_i denotes the local incidence angle, h_i is the topographic height, d_i is the deformation in the radar Line of Sight (LOS) and $a_i \cdot 2\pi$ represents the phase ambiguity term. For the derivation of Equation 2.6 refer to [1].

If we consider a *single-pass* interferometric configuration for Equation 2.6 then $d_i = 0$. If we further remove the flat earth phase, the term $\frac{4\pi}{\lambda}B_i^{\parallel}$, by the knowledge of the satellite orbit and a reference ellipsoid [3] and assume the integer number of phase cycles a_i has been recovered through phase-unwrapping [2], then h_i is calculated through phase to height conversion [4]:

$$h_i = \frac{-\lambda}{4\pi} \cdot \frac{R_i \sin \theta_i}{B_i^\perp} \cdot \Delta\Phi_i, \tag{2.7}$$

where $\Delta\Phi_i = \Delta\phi_i + a_i \cdot 2\pi$ denotes the final absolute phase at the ith pixel. It is important to note that the phase-unwrapping procedure is relative in nature, which means $\Delta\Phi_i$ and consequently h_i are relative quantities defined with respect to a reference point selected by the user.

At this stage of InSAR processing, the ith pixel is defined within the 3-D radar datum in (l_i, p_i, h_i) or in (t_i, τ_i, h_i), which still needs a coordinate transformation to be defined in a common geodetic reference system. This is performed through *geocoding*, that is explained in the following.

2.2.1 InSAR Geocoding

For the pixel at (l_i, p_i) in the SAR image, geocoding is carried out by iterative solving of the Doppler-Range-Ellipsoid equations to retrieve the 3-D Cartesian coordinates of the target $\vec{T}_i = (X_i, Y_i, Z_i)$ from its radar timing and height coordinates (t_i, τ_i, h_i) [27]–[29]:

$$\left\| \dot{\vec{S}}(t_i) \cdot (\vec{S}(t_i) - \vec{T}_i) \right\| = 0 \tag{2.8}$$

$$\left\| \vec{S}(t_i) - \vec{T}_i \right\| - c \cdot \frac{\tau_i}{2} = 0 \tag{2.9}$$

$$\frac{X_i^2}{(a + H_i)^2} + \frac{Y_i^2}{(a + H_i)^2} + \frac{Z_i^2}{(b + H_i)^2} - 1 = 0, \tag{2.10}$$

where $\vec{S}$ and $\dot{\vec{S}}$ are the satellite state vector and its velocity vector, respectively. They are both functions of t_i. The first equation dictates the zero-Doppler condition. The second equation relates the geometric distance between $\vec{T}$ and $\vec{S}$ to the fast time of the target τ_i. The result from Equations 2.8 and 2.9 is intersected with a reference ellipsoid with semi-major axis a and semi-minor axis b, and with an estimated height above the ellipsoid H_i. Note that external height information, from at least one GCP or a Digital Elevation Model (DEM), is required for connecting the relative interferometric height h_i (see Equation 2.7) to H_i [3]. For convenience of interpretation (X_i, Y_i, Z_i) are usually converted to the Universal Transverse Mercator (UTM) map projection (E_i, N_i, H_i), where E_i and N_i denote the UTM Easting and Northing with H_i representing the height above the, global or local, reference ellipsoid.

2.2.2 Limitations and Errors of InSAR

2.2.2.1 Signal Decorrelation

One of the most important conditions for interferometry to work properly is that an imaged pixel stays coherent between the two acquisition times [4]. This condition however does not always hold due to the difference in the viewing geometries of the two

acquisitions, which causes geometrical decorrelation, or due to the temporal separation between the two images, which leads to temporal decorrelation [30]. Both of these issues cause a change in the scattering phase from the first to the second acquisitions leading to random errors in the interferometric phase. The phase noise complicates phase-unwrapping and consequently degrades the quality of the InSAR topographic and deformation maps.

2.2.2.2 Atmospheric Heterogeneity

The difference in atmospheric states of the first acquired SAR image and the second one is problematic for InSAR. This issue causes a systematic phase component, which affects the interferometric phase [4]. The spatial correlation of the effect does not allow a seamless separation of the atmospheric signal from the desired topography or deformation phase components using conventional InSAR [29] and thus leads to biased estimates.

2.2.2.3 Geocoding Errors

The systematic errors in radar timings, briefly mentioned in Subsection 2.1.1, and the fact that the InSAR height is relative in nature (see Section 2.2), manifest as position shifts in the geocoded InSAR products. In the following, it is characterized how these errors affect the final position of point targets in the UTM coordinate system. The radar timing coordinates of the assumed point targets are represented as (t_{az}, τ_{rg}) and their final geocoded height as H_T.

For t_{az} the errors cause shifts only in the along track direction. If we consider a straight satellite orbit trajectory and approximate the curved Earth geometry by a rectilinear one [20], as is visualized in Figure 2.1a, the ground shift δl_{az} due to an erroneous azimuth time t'_{az} is calculated in meters as:

$$\delta l_{az} = v_r \left(t'_{az} - t_{az} \right), \tag{2.11}$$

where t_{az} is the true azimuth time and v_r denotes the effective radar velocity of the satellite. Note that for the curved Earth geometry, the beam velocity v_g ($v_g < v_r$) should be used in Equation 2.11 [20]. The error δl_{az} affects only the horizontal geocoded coordinates and with the knowledge of the local heading angle of the satellite α, its effect can be projected onto the local east δl_{az}^E and north δl_{az}^N components (See Figure 2.1b):

$$\delta l_{az}^E = \delta l_{az} \sin \alpha, \tag{2.12}$$

$$\delta l_{az}^N = \delta l_{az} \cos \alpha. \tag{2.13}$$

Considering the near polar orbit of current SAR satellites, with heading angles close to 180° or 360°, Equations 2.12 and 2.13 show that error in t_{az} mostly affects the north component of the geocoded coordinates.
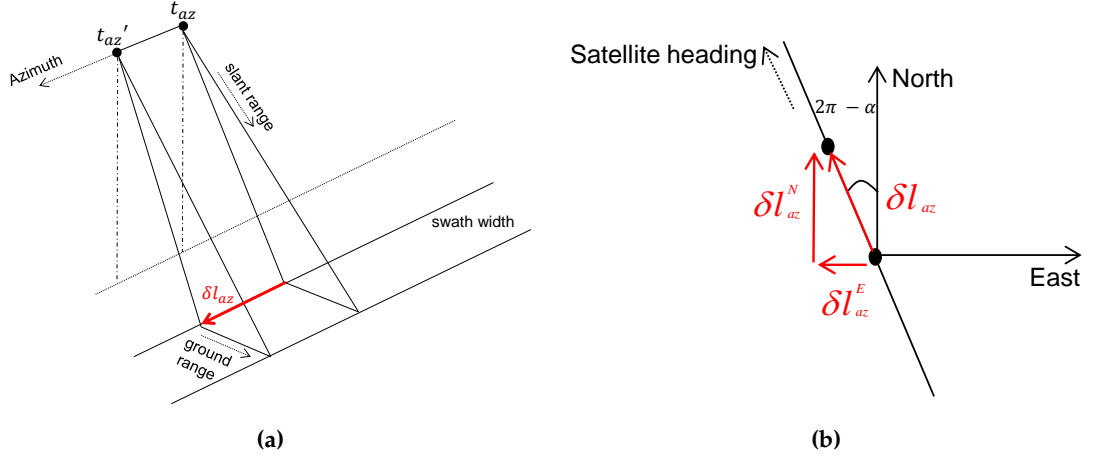
Figure 2.1: Timing bias in azimuth and its effect on the geocoded coordinates. (**a**) Azimuth timing error causing a displacement on the ground δl_{az}. (**b**) Projection of δl_{az} onto the east δl_{az}^E and north δl_{az}^N components using satellite heading angle α.

Errors in τ_{rg} cause a delay in the received radar pulse that affects the geometrical distance between the satellite and the target in the slant range direction δl_{sr}, which leads to a displacement on the ground range δl_{gr} depending on the the local incidence angle θ (see also Figure 2.2a):

$$\delta l_{sr} = c \; \frac{\left(\tau'_{rg} - \tau_{rg}\right)}{2} \tag{2.14}$$

$$\delta l_{gr} = \frac{\delta l_{sr}}{\sin \theta} \tag{2.15}$$

where τ_{rg} and τ'_{rg} are the true and the erroneous range time, respectively. Similar to δl_{az}, δl_{gr} is related to the east δl_{gr}^E and the north δl_{gr}^N components by a projection using α (see also Figure 2.2b):

$$\delta l_{gr}^E = \delta l_{gr} \cos \alpha, \tag{2.16}$$

$$\delta l_{gr}^N = -\delta l_{gr} \sin \alpha. \tag{2.17}$$

It is evident from Figure 2.2a and Equation 2.15 that δl_{gr} becomes larger with steeper incidence angles. It is also worth mentioning that according to Equations 2.16 and 2.17, the timing error in range manifests itself mostly in the east coordinate component.

Apart from the timing errors, the relativity of InSAR height estimates can be problematic for geocoding. The height value of each pixel in the interferogram is evaluated with respect to a reference point. It was mentioned in Subsection 2.2.1 that the height of the reference point is usually retrieved from an external source [3] or in modern In-
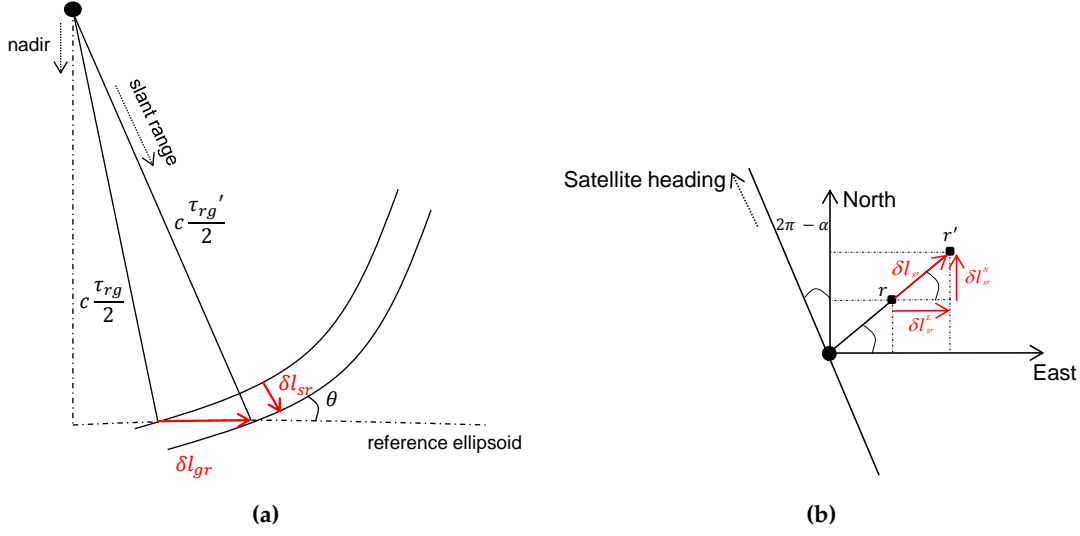
Figure 2.2: Timing bias in range and its effect on the geocoded coordinates. (**a**) Range timing error and its impact on geocoding in the ground range δl_{gr}. (**b**) Projection of δl_{gr} onto the east δl_{gr}^E and north δl_{gr}^N components using satellite heading angle α.
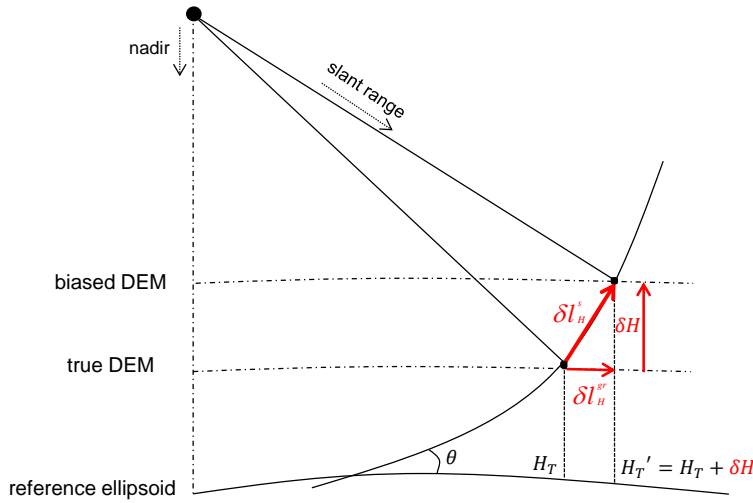
Figure 2.3: Depiction of height error δH due to unknown DEM error at the reference point and its effect on geocoded coordinates. It can be seen that this error causes a shift in the cross-range direction δl_H^s, which is decomposed into an offset in ground range δl_H^{gr} and a vertical component δH.

SAR processing chains, such as the one for TanDEM-X, through radargrammetry [31]. The final absolute height accuracy of InSAR products thus depends on the height error of this point, which is an overall unknown offset. This height error $\delta H = H_T' - H_T$ is almost constant for all pixels in the interferogram and has a significant effect on final
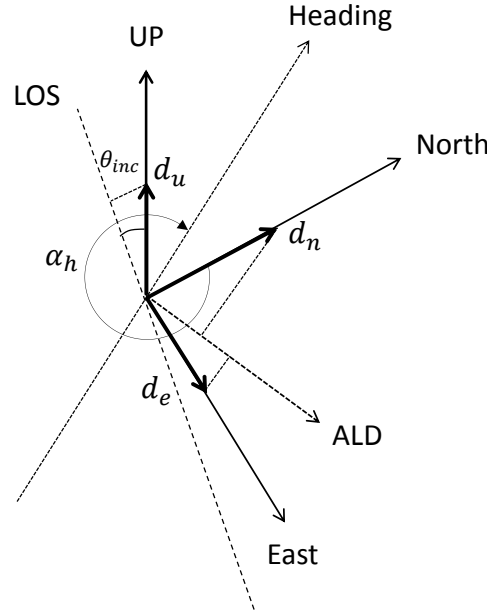
Figure 2.4: Projection of the original displacement vector **d** with components (d_e, d_n, d_u) onto the radar LOS.

geocoded coordinates both in the horizontal and in the vertical components, as can be seen in Figure 2.3. The variable δH causes a horizontal shift in the ground range δl_H^{gr} as:

$$\delta l_H^{gr} = \frac{\delta H}{\tan \theta}.$$ (2.18)

Similar to the error in range timing (see Equations 2.16 and 2.17), δl_H^{gr} can be projected onto the east and the north components by the knowledge of α.

2.2.2.4 InSAR 1-D Deformation

The deformation measurement of InSAR techniques d_{LOS} is the projection of the original 3-D displacement vector **d** with components d_e, d_n and d_u in east, north and up direction, respectively, onto the LOS. Following previous definitions of the local incidence angle θ and the local heading angle α, we can write [4]:

$$d_{LOS} = d_u \cos \theta - d_{ALD} \sin \theta$$ (2.19)

where d_{ALD} includes the projection of d_n and d_e on the azimuth look direction (ALD), which is perpendicular to the satellite flying direction and therefore is expressed as:

$$d_{ALD} = d_n \sin \alpha - d_e \cos \alpha.$$ (2.20)

Figure 2.4 depicts the aforementioned projection in 3-D. If Equation 2.20 is substituted in Equation 2.19, the explicit relation between d_{LOS} and the displacement components for a single pixel can be written as:

$$d_{LOS}^{D} = d_u \cos\theta + d_e \cos\alpha \sin\theta - d_n \sin\alpha \sin\theta, \tag{2.21}$$

where the superscript D denotes the descending viewing geometry. For an ascending observation, the second term in Equation 2.20 gets a positive sign, which means that the LOS deformation becomes [32]:

$$d_{LOS}^{A} = d_u \cos\theta - d_e \cos\alpha \sin\theta - d_n \sin\alpha \sin\theta. \tag{2.22}$$

Equations 2.21 and 2.22 indicate that the retrieval of the 3-D displacement components is not possible using only observations from a single viewing geometry. The 1-D deformation is ambiguous and depends on the satellite heading direction, which degrades the interpretation of InSAR displacement maps.

2.2.3 Tackling InSAR Limitations: an Outlook

The relevant basics of InSAR, geo-referencing its products, and some of the main limitations of the method have been discussed in this section. It was mentioned that Signal decorrelation leads to random noise on interferometric phase while atmospheric differences between the time of acquisitions, radar timing errors and height error of the InSAR reference point cause systematic errors, which eventually affect the final geocoded positions of InSAR products. The errors in range and azimuth measurements can be mitigated by using the SAR imaging geodesy method (see Section 3.1). This leads to correcting the horizontal geocoding errors of InSAR products. In order to compensate for the height error of the reference point, SAR-based GCPs can be used. These points are absolutely localized in 3-D by the *stereo SAR* method [19], that is explained in section 3.2. Exploiting multiple SAR images of the same scene, InSAR time-series approaches such as PSI [6], [7] can be used to overcome atmospheric disturbances and signal decorrelation. PSI restricts the topographic and deformation parameter retrieval to only point-like scatterers which remain coherent during the entire acquisition period. This method will be explained in Section 3.3. Finally, the 1-D deformation measurements of InSAR techniques can be, theoretically, decomposed into the original 3-D motion components, if InSAR displacement observations from at least three viewing geometries are available [4], [33], [34]. This requires the correct fusion of multi-aspect InSAR products, which is addressed in Section 3.5.

2.3 Basics of Observation Theory

2.3.1 Accuracy versus Precision

In geodesy, a distinction is usually made between *accuracy* and *precision*. In order to clarify these terms, let us assume a set of observations of the random variable $\underline{x}$ is avail-

able from measuring the *true* value μ_x of a quantity. The observations are considered accurate, or equivalently unbiased, if [35]:

$$E\{\underline{x}\} = \mu_x, \tag{2.23}$$

where $E\{\cdot\}$ is the mathematical expectation operator. Therefore, the accuracy of a set of observations is defined as their closeness to the true quantity. An inaccurate set of observations may be formulated as $E\{\underline{x}\} = \mu_x + \epsilon$, where ϵ denotes the magnitude of the bias with the condition $E\{\epsilon\} \neq 0$.

The precision is concerned with *variance*, which for the same assumptions as above is defined as [35]:

$$\sigma_x^2 = E\{(\underline{x} - \mu_x)^2\}. \tag{2.24}$$

According to Equation 2.24, the variance of $\underline{x}$, symbolized as σ_x^2, describes the degree of dispersion of the observations with respect to μ_x. Therefore σ_x^2 depends on how repeating the measurements generates consistent observation values. The lower values for σ_x^2 indicate smaller dispersion, smaller variance and consequently higher precision.

In this thesis, the term accuracy is only used when an external reference data or a ground truth is available for validation. Otherwise no conclusion about the accuracy of the observations or estimates can be made and therefore the statistical analysis will be restricted to precision only.

2.3.2 Parameter Estimation

Parameter estimation is at the core of many engineering problems. In the following, two common approaches for estimating unknown parameters from an overdetermined system of equations are briefly introduced. A distinction regarding their robustness with respect to outliers is made, which will be later utilized in Section 4.4.

Let us assume that an estimate of an unknown vector $\mathbf{x}_{n \times 1}$ is desired. With the knowledge of the design matrix $\mathbf{A}_{m \times n}$ and the observation vector $\mathbf{y}_{m \times 1}$, the functional model of the problem is written as [35]:

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{v}, \tag{2.25}$$

where $\mathbf{v}_{m \times 1} = \mathbf{y} - \mathbf{A}\mathbf{x}$ is the vector of residuals. Assuming that $\mathbf{A}$ has full rank and $m \geq n$ (overdetermined system), the estimation of unknowns is commonly carried out by minimizing the (weighted) sum of the square of residuals known as the (weighted) least squares or l_2-norm minimization [36]:

$$\hat{\mathbf{x}}_{l_2} = \arg\min_{\mathbf{x}} \left\{ ||\mathbf{y} - \mathbf{A}\mathbf{x}||_{\mathbf{W}}^{l_2} \right\} \tag{2.26}$$

with $\mathbf{W}$ being the weight matrix of observations. Equation 2.26 gives the unique solution of the problem, which guarantees unbiasedness and minimum variance when $\mathbf{W}$ is chosen equal to the inverse of the covariance matrix of the observations [35]. How-

ever, the aforementioned properties of the least squares method are only valid when observations are only influenced by random errors following a normal distribution. If multiple observations are affected by gross errors (outliers) to cause the errors follow a non-Gaussian and heavily tailed distribution, parameter estimation based on l_2-norm minimization is far from optimal [37]. The reason lies on the objective function described in Equation 2.26, which inherently gives more weight to outliers with large residual values rather than normal observations in the minimization process. Therefore, l_2-norm minimization is not robust in such cases due to the high sensitivity of the estimator toward outliers.

In order to ensure robustness from the system of equations outlined in Equation 2.25 an estimate of $\mathbf{x}$ can be obtained by minimizing the (weighted) sum of the absolute residuals [37]:

$$\hat{\mathbf{x}}_{l_1} = \arg\min_{\mathbf{x}} \left\{ \, |\mathbf{y} - \mathbf{A}\mathbf{x}|_{\mathbf{w}}^{l_1} \right\} \tag{2.27}$$

where $\mathbf{w}_{m \times 1}$ contains the diagonal elements of $\mathbf{W}$. Unlike least squares similar residual weights are used while solving Equation 2.27. This means l_1-norm adjustment does not magnify the effect of outliers. Therefore, it facilitates less biased estimation of the unknown parameters compared to least squares when there are gross errors in the observations, although the solution does not necessarily have the minimum variance [38].

3 State-of-the-art

3.1 Absolute 2-D Positioning Using SAR

The imaging capability of spaceborne SAR sensors and their ability to provide high precision relative measurements, using SAR interferometry, have been at the center of interest of the SAR community for decades. The pixel localization accuracy of SAR however, was not fully studied and exploited until a few years after the launch of high resolution satellites: TSX and COSMO-Skymed. In 2011, the SAR imaging geodesy method was introduced, which has enabled high precision 2-D absolute positioning of SAR images by exploiting the highly precise orbit of TSX [15] and by using correction techniques employed commonly in the field of geodesy [13].

In the following, the theoretical aspects as well as requirements and correction methodologies of imaging geodesy are treated along with summarizing the recent developments of the method.

3.1.1 SAR Imaging Geodesy

Following the notations in Chapter 2, a point target located within the pixel (l_i, p_i), in a SLC, is characterized with the biased azimuth and range times denoted by t_i' and τ_i', respectively. Considering the most prominent factors causing systematic errors, the timings can be written as:

$$t_i' = t_i + \Delta t_i^O + \Delta t_i^G + \Delta t^{cal} \tag{3.1}$$

$$\tau_i' = \frac{2R_i}{c} + \Delta \tau_i^O + \Delta \tau_i^I + \Delta \tau_i^T + \Delta \tau_i^G + \Delta \tau^{cal}, \tag{3.2}$$

where t_i is the true zero-Doppler time of the scatterer in seconds and R_i is the absolute geometric distance from the satellite to the center of the pixel in meters. The error terms Δt_i^O and $\Delta \tau_i^O$ denote the orbit inaccuracies and Δt_i^G and $\Delta \tau_i^G$ include the geodynamic effects on range and azimuth timings, respectively, while $\Delta \tau_i^I$ and $\Delta \tau_i^T$ are the ionospheric and the tropospheric delays on range timings. The geometrical calibration constants in azimuth and range are denoted by Δt^{cal} and $\Delta \tau^{cal}$, respectively. The magnitude of individual effects in Equations 3.1 and 3.2 can be scaled to units of length by multiplying the range error terms with $c/2$ and the azimuth error terms with the beam velocity of the satellite (for TSX $\approx 7050 \ m/s$).

Assuming that the goal is to estimate and remove the error terms of the mentioned point target, the imaging geodesy approach relies on the following requirements and the procedures:

- The accurate knowledge of the the SAR satellite's orbit through high precision orbit determination.

- High precision calibration of SAR instrument timings and accurate processing of SAR images into zero-Doppler geometry.

- High precision sub-pixel extraction of the point target response from corresponding SLC images through Point Target Analysis (PTA) and converting them to radar timings.

- Correction of external geodetic and atmospheric errors using available auxiliary data at the location and the time of the corresponding point target.

These items are linked with the limiting factors of Equations 3.1 and 3.2. In the following the error sources and the correction methods are described in more detail. The PTA procedure is not treated here and the interested reader is referred to [20] for the theory and to [39] for the method used in this thesis.

3.1.1.1 Orbit Accuracy

The accurate knowledge of orbit is an important requirement for imaging geodesy. The science orbit products of TSX have a 3-D accuracy of 4.2 cm as reported in [15]. This level of accuracy has been further improved to 1-2 cm by inclusion of models taking into account the solar radiation pressure and improved atmospheric density, and the use of sophisticated gravity field models [40]. These orbit solutions have been also enhanced by considering Global Navigation Satellite System (GNSS) ambiguity fixing and modeling of non-gravitational forces, and were validated using satellite laser ranging in [41] by analyzing a time-series of orbit solutions in a period of six years. If this level of accuracy on orbit determination is achieved, then the orbit effects do not play a significant role in degrading the accuracy of radar timings. In this thesis, since the enhanced orbit solutions were used, the orbit effects on 2-D and 3-D localization accuracy are deemed negligible.

3.1.1.2 Atmospheric Delays

The variable τ_i in Equation 3.2 is delayed due to the propagation of the SAR signal through different layers of atmosphere. The delay includes an ionospheric and a tropospheric component. For the TSX sensor, these effects and their impact on geolocation of Corner Reflector (CR)s were initially studied in [13], [14], [42], [43].

The frequency-dependent ionospheric delay at zenith is in the order of couple of centimeters for X-band SAR systems, if typical ionospheric conditions are assumed, and is inversely proportional to the $\cos \theta_i$, with θ_i denoting the local incidence angle of the radar beam [13]. This effect is corrected on base of the global Total Electron Content (TEC) maps [13] that use the single layer model as mapping function [44]. Using the orbit information of the SAR satellite, the point at which the line of sight of the satellite passes through the ionospheric layer to the point of interest is analytically localized.

The zenith TEC value is then determined at the location of the aforementioned point through spatial and temporal interpolation of the TEC maps [39], [45], [46] and is projected onto the radar LOS. At the final stage, an empirical weighting factor of 75% is used to account for the fact that TSX signals travel only partly through the ionosphere [18].

The tropospheric delay is considered the most prominent error source for TSX range measurements. Its magnitude can reach to over 4 m depending on the weather conditions, the target height and the steepness of θ_i [4], [13]. The effect of troposphere is computed through the 4-D integration of numerical weather data from European Center for Medium-Range Weather Forecasts (ECMWF). The method extracts the dataset from a local database, converts it to a conventional geographic coordinate system, performs a 3-D interpolation for defined integration points and eventually integrates the refractivity index along the integration path in the slant-range direction from the point of interest to the satellite [17], [47].

The ionospheric and tropospheric correction values are converted into timings $\Delta \tau_i^I$ and $\Delta \tau_i^T$ and are subtracted from the initial measurement τ_i' in Equation 3.2.

3.1.1.3 Geodynamic Effects and Plate Tectonics

The Geodynamic effects concern phenomena, which deform the surface of the earth and as a result change the position of targets on the ground [48]. The main contributors to geodynamic effects, namely the solid earth tides (SET), ocean tidal loading and pole tides were first considered for TSX range timings in [13]. The magnitude of SET is around 40 cm for the radial component and several centimeter for the horizontal components while the sum of other factors can affect the range measurements by several centimeters [13]. In order to improve the localization precision into sub-centimeter regimes, geodynamic effects with smaller magnitudes, such as atmospheric pressure loading, ocean pole tides and atmospheric tidal loading were further modeled in [49]. All of the effects, except for atmospheric pressure loading, are calculated according to the state-of-the-art models of the International Earth Rotation and Reference Systems Service (IERS) 2010 convention [48].

Apart from these periodical effects, the impact of cumulative linear shifts caused by plate tectonics, in the order of centimeters per year, are considered when aiming at centimeter regime localization accuracy [14], [17], [39], [47]. The plate tectonics should be considered because the reference frames within which the SAR images and the satellite orbits have been defined are different [14].

The corrections mentioned above are all projected into the slant range and azimuth coordinates and are subtracted from measured t_i' and τ_i' in Equations 3.1 and 3.2. For more in depth treatment of the geodynamic effects the interested reader is referred to [17], [48] and their references.

3.1.1.4 Satellite's Internal Effects and Calibration Constants

Important requirements for achieving 2-D localization in the centimeter and millimeter regimes are the accurate processing of SAR images to zero-Doppler geometry as well as precise measurement and annotation of internal timings. These errors affect both t_i' and τ_i' with magnitudes in the decimeter regimes [13]. The additional traveling time of the radar pulse through internal cables of the instrument, both in transmission and reception, influences τ_i' [19]. For t_i', the coarse clock rate which is used for annotating the raw acquisition time is one of the biggest limiting factors. In [50], the Instrument Fine Time (IFT) correction approach has been described in order to improve the annotation of the raw acquisition times. Another error source is the stop-go approximation, which assumes a static sensor during radar pulse transmission and reception. This introduces mainly azimuth timing errors with small secondary effects in range. The correction of the approximation is performed during azimuth focusing as is explained in [51]. All the mentioned factors are taken into account for TSX using the latest version of TerraSAR-X multimode SAR processor (TMSP) [51].

The variables Δt^{cal} and $\Delta \tau^{cal}$ compensate for the remaining effects that are not categorized in the previous error classes. They are evaluated after mitigating all the external errors and satisfying all the mentioned requirements regarding orbit accuracy and internal satellite dynamics. For this purpose, the difference between the radar timings measured from the SAR image and the expected values, surveyed with GNSS or from terrestrial surveys, yields the calibration constants. In this thesis, the used calibration constants are the results from long-term CR campaigns reported in [18], [19], [46].

3.1.2 Recent Advances in Imaging Geodesy

Using TSX, the imaging geodesy technique is capable of retrieving absolute 2-D coordinates of well-detectable point scatterers, like CRs, [13], [39] as well as providing accurate relative measurements by employing incoherent correlation of multiple SAR images [13]. The CR campaigns are useful for providing new geometrical calibration constants as has been demonstrated in [18], [39], [46]. Absolute range measurement ability of TSX and TanDEM-X has been exploited for maritime purposes in [52], [53]. The world-wide reproducibility of imaging geodesy was reported in [54] and the global applicability of the method has been enhanced by introducing tropospheric correction methods based on numerical weather data [17], [47], [55]. An operational processor, called SAR Geodesy Processor (SGP), was introduced in [56], which delivers the desired correction values on a coarse grid imposed on SAR images.

Geodynamic and atmospheric correction methods similar to the ones described in this Section, were also tested on Sentinel-1 stripmap products during its commissioning phase in [57] where the geolocation accuracy of SLCs was reported to be 6 cm in range and 17 cm in azimuth, respectively. These studies are still ongoing on interferometric wide swath (IW) products with preliminary results reported in [58]. The most recent results of Sentinel-1 geolocation accuracy using an array of 40 CRs in Australia was

reported in [59], which demonstrated measurements with precision of 6 cm in range and of approximately 28 cm in azimuth for both Sentinel-1 A and B IW products.

Apart from the above-mentioned developments, which mainly depend on synthetic bright scatterers, an approach similar to imaging geodesy was combined with single stack PSI results in [60] and [61] for improving geocoding accuracy of natural Persistent Scatterer (PS) and target association.

3.2 Absolute 3-D Positioning Using SAR

The SAR imaging geodesy technique, explained in Section 3.1, is a novel approach for absolute localization of bright scatterers only in 2-D. In order to retrieve the 3-D position of a point target, the visibility of the scatterer should be guaranteed in at least a second SAR image that is acquired from a different viewing geometry. This is the fundamental assumption that InSAR [2], radargrammetry [62] and stereo SAR [19] use for accessing the height of the target.

Stereo SAR, similar to radargrammetry, exploits only the amplitude information of the SAR images and therefore is not affected by the major difficulties in InSAR such as coherence loss and the phase unwrapping procedure. Unlike radargrammetry, which operates on homologous patches of the stereo image configuration for elevation reconstruction, stereo SAR directly estimates the 3-D position of single PS. It is important to note that unlike InSAR, stereo SAR and radargrammetry allow for the retrieval of absolute 3D coordinates without the use of a reference point. However, the precision of height estimation using these methods depends on the spatial resolution of the SAR images and is often much lower than the precision of reconstructed heights using InSAR.

In the following, the mathematical formulation and the estimation procedure of stereo SAR are described. The method used in this thesis is based on [19] and is called geodetic stereo SAR because the timing coordinates of targets involved in the stereo processing are *a priori* corrected using SAR imaging geodesy. For different approaches of tackling the stereo SAR problem the reader is referred to [63]–[65].

3.2.1 Geodetic Stereo SAR

We assume that for the pixel (l_i, p_i) in a SLC, the peak of the response of a dominating point scatterer is extracted by PTA and is converted to radar timings t_i' and τ_i'. Furthermore, the unbiased timings of the target t_i and τ_i are obtained by applying the corrections outlined in Subsection 3.1.1. Note that at this stage, the timings of a particular target are extracted from multiple SAR images separated in the temporal domain. The relation between these 2-D radar time coordinates of the target $\mathbf{x}_{T_i} = (t_i, \tau_i)$ and their corresponding 3-D Cartesian coordinates $\mathbf{X}_{T_i} = (X_i, Y_i, Z_i)$ is defined by the Range-Doppler equation system [20]:

$$\frac{\dot{\vec{S}}(t_i) \cdot (\vec{S}(t_i) - \vec{T}_i)}{\left\| \dot{\vec{S}}(t_i) \right\| \cdot \left\| \vec{S}(t_i) - \vec{T}_i \right\|} = 0 \tag{3.3}$$

$$\left\| \vec{S}(t_i) - \vec{T}_i \right\| - c \cdot \frac{\tau_i}{2} = 0 \tag{3.4}$$

with $\vec{S}(t_i)$ and $\dot{\vec{S}}(t_i)$ being the position and the velocity vector of the satellite relative to t_{az}, and τ_{rg} being the two-way traveled time from the satellite to the target. The variable t_{az} is implicitly included in Equation 3.3 relating the state-vector of the satellite to the time of the acquisition via a polynomial model [19]. Equation 3.4 defines a sphere centered on $\vec{S}(t_i)$ which reduces to a circle perpendicular to the satellite trajectory when coupled with the zero-Doppler plane described in Equation 3.3. Therefore $\mathbf{X}_{T_i}$ can be retrieved by including another set of timing observations from a different satellite position which evaluates the intersection point of the two circles, see Figure 3.1.

The estimation of the coordinates is carried out by means of least squares after linearization of the observation equations [36]. No initial weights are allocated to the timing observations [18], [19]. The estimation however is iterated and based on the residuals of each iteration, variance component estimation [19] is performed to allocate a common weight for all the range timings and a common weight for all the azimuth timings of the same target extracted from all available SAR images. The result is the 3-D coordinates of the target and *a posterior* covariance matrix of the estimates. The latter can be transformed to any frame orientations, for instance local East, North, Height, to be visualized as error ellipsoids [36]. The Precision of the estimated 3-D coordinates depends on the Signal-to-Clutter-Ratio (SCR) of the target, the precision of the external radar timing corrections, the separation in the viewing geometries and the number of acquisitions.

3.2.2 Recent Advances in Stereo SAR

The first results on absolute 3-D localization of CRs by means of geodetic stereo SAR was reported in [66], in which 3-D precision better than 4 cm and an absolute accuracy of 2-3 cm were achieved using TSX very high resolution spotlight products. Although 3-D positioning using multi-aspect TSX images had been previously demonstrated in [65], [67], [68], the results in [66] were unique in the sense that the stereo processing was carried out on thoroughly calibrated range and azimuth timings. The method was also used for 3-D positioning of multi-directional radar reflectors using both TSX high resolution and staring spotlight images in [69] and was extended to a differential form in [70]. The applicability of geodetic stereo SAR not only on CRs but on opportunistic non-ideal scatterers such as PS in an urban area was demonstrated in [19], with reported 3-D precision values better than 10 cm. This paved the way for new geodetic applications using SAR such as secular ground movement estimation using natural PS [71], [72], high precision mapping of road networks (DriveMark) [73], and highly precise auto-

(a) Cross-heading   **(b)** Same-heading
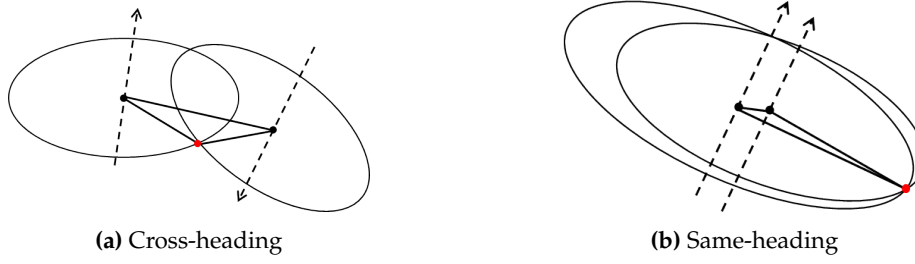
Figure 3.1: Localization of a point target (red dot) from (**a**) cross-heading and (**b**) same-heading satellite tracks using stereo SAR. The satellites are shown by black dots; their trajectories are presented by dashed lines and the baselines are depicted by solid lines between the satellite positions. The black circles are defined by the range-Doppler equations and their intersection leads to the 3-D position of the target.

matic SAR GCP generation using TSX [74]–[77] and using COSMO-SkyMed [78]. A combination of stereo SAR and InSAR for accurate DEM generation using TanDEM-X spotlight data has been used in [79].

3.3 Persistent Scatterer Interferometry (PSI)

The PSI technique was proposed in [6] in the late 1990's. It is a single-master multi-temporal extension of conventional InSAR that extracts phase-stable scatterers, the so-called PS, from a stack of SAR images and restricts the estimation of topography and deformation parameters to only these targets [6], [7]. The usage of multiple images instead of only two and considering only PS for the estimation allows PSI to overcome the main limitations of InSAR, namely atmospheric disturbances [4] and geometric and temporal decorrelation [30]. The PSI technique is highly effective in urban areas because of the availability of a high density of PS. In particular, coupling the technique with high resolution images, such as the ones from the TSX spotlight mode, produces on average between 40,000 to 100,000 PS per square kilometer [11], [80], which allows for detailed infrastructure monitoring.

In the following more details about PSI and its developments relevant to urban monitoring are given.

3.3.1 PSI Processing Sequence

Several implementations of PSI, suitable for different applications, are available. For a full review of these techniques the reader is referred to [81]. In the following, the most common processing steps of the method are listed and briefly explained.

1. Differential interferogram formation: From a stack of $N + 1$ coregistered SAR images, a master acquisition is selected. Subsequently N interferograms are computed, while their topographic phase components are removed using a reference digital elevation model (DEM).

2. Reference network construction: Scatterers presumed to be the most phase-stable ones are selected. The detection can be carried out using various methods, such as thresholding on the Amplitude Dispersion Index (ADI) [6] or on the signal-to-clutter ratio (SCR) [82]. These PS candidates are connected to form a reference network. Through the PS double-difference phase measurements, i.e., difference in time and space, differential topography and differential motion parameters are estimated on arcs.

3. Atmospheric phase estimation: The differential topography estimates are integrated with respect to an arbitrarily chosen reference point so that the topographic phase components are removed from the interferometric phases. The remaining phase contributions include deformation, atmosphere, and noise. Then a low-pass filtering in the spatial domain and a high-pass filtering in the temporal domain extracts the atmospheric component, which is interpolated over the entire scene and is subtracted from the differential interferograms.

4. PS densification: Additional PS are computed from the corrected differential interferograms. These PS are connected to the nearest point(s) in the reference network and their modeled parameters are estimated.

5. PS geocoding: The DEM height of each PS is added to its differential height estimate. The radar timing of each PS and its updated height are geocoded using satellite orbit and a reference ellipsoid to represent the PS coordinates in a common geodetic coordinate system.

3.3.2 Recent Advances in PSI

For each of the steps delineated above, numerous improvements have been suggested in the literature. In [83], a geometry-based SAR image coregistration method was developed to match PS from different images and hence avoid cross-correlation in the inteferogram generation part. In [84], [85], numerical weather data were used to simulate and mitigate tropospheric delay and provide aid in the master selection step. Apart from the reference PS technique which uses ADI, PS detection has been carried out by evaluation of target SCR [7], [82], by analyzing ADI plus spectral phase diversity [86] and by using criteria based on both amplitude and phase, which proved to be effective in non-urban areas [87]. In the reference network step, [7], [29] consider the geometry of the connections among arcs to construct a redundant reference network, while dense differential PS pairs were used in [88]. To improve the precision of the estimated parameters of PS in the reference network and define quality measures, the Least-Squares AMBiguity Decorrelation Adjustment (LAMBDA) method [89] was adapted to be used in PSI [90] which was further implemented in the STUN algorithm [7]. In terms of network inversion, to robustly retrieve the topography and deformation estimates of the PS in the reference network, a l_1 norm outlier rejection scheme was proposed after the LAMBDA estimation [91]. Apart from the default linear displacement model assumed

in the reference PS technique [6], the inclusion and estimation of seasonal ground motion were investigated in [92] for medium resolution data and in [80] for high resolution spotlight TSX images. For a detailed comparison of widely used PSI techniques, the interested reader is referred to [29].

There have been also major developments regarding applications exploiting the PSI method. Using medium resolution SAR data, accurate urban DEMs were generated in [93] and their sub-metric height accuracy was validated in [94]. The localization of PS and associating targets to real objects have been investigated in [61], [95], [96]. This aspect has been also studied extensively using high resolution spotlight TSX data in [11], [97], which demonstrated the concept of single-building deformation monitoring using PSI, and in [98] to evaluate PS localization accuracy at building facades.

3.4 Tomographic SAR Inversion (TomoSAR)

TomoSAR can be considered as one of the most advanced InSAR stacking techniques, especially for urban monitoring. Similar to PSI, it is a single-master method. However, the main difference with PSI is that, instead of only two complex SAR images for phase-difference exploitation, TomoSAR uses multiple SAR images to reconstruct the full reflectivity profile of each pixel along the cross-range direction [8]. Therefore, TomoSAR is capable of resolving the layover issue (see Subsection 2.1.1) and leads to the separation of different scattering mechanisms within a single pixel [9].

Since our focus in this thesis is on urban areas, in the following the point scatterer-based TomoSAR method is explained, for which the problem can be formulated with a deterministic model. For tomographic SAR inversion of volumetric and distributed scatterers with applications mostly in forest mapping, the interested reader is referred to [8], [99]–[101], to name a few examples.

In the following the imaging model of differential TomoSAR and the workflow for retrieving the elevation and deformation estimates of multiple scatterers using the method are described. The section is concluded with reviewing the most recent advances in TomoSAR for urban area mapping.

3.4.1 TomoSAR Imaging Model and Workflow

The TomoSAR imaging model for a case in which, three scattering contributions from the ground, the building facade and the building roof are mapped within one azimuth-range pixel, is visualized in Figure 3.2. We assume a stack of coregistered SLCs acquired in a repeat-pass configuration is available. Within the far-field approximation, the complex-valued measurement g_n at the pixel (l_i, p_i) is essentially the Fourier transform of the elevation s-dependent reflectivity profile γ of that pixel sampled at the corresponding elevation frequency ζ_n which, considering the temporal progression of the scatterers, for the nth acquisition at the time t_n ($n = 1, \cdots, N$), is written as [10], [102]:
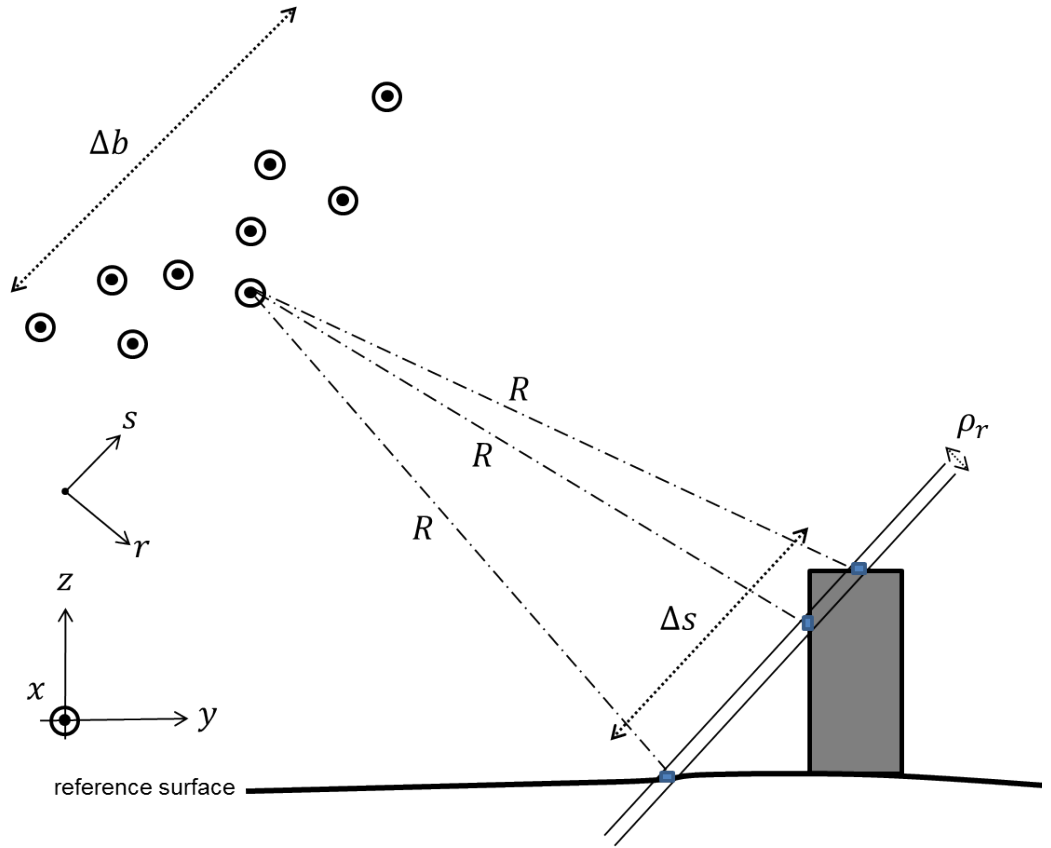
Figure 3.2: Imaging geometry of TomoSAR. The input images are acquired from different satellite positions with a total aperture length of Δb in elevation. Three backscattering contributions from the ground, the building facade and the building roof are mapped within one azimuth-range pixel and are visualized with small blue rectangles. TomoSAR estimates the reflectivity profile of the mapped pixel along s for a dicretized elevation extent Δs and subsequently separates the layovered scatterers.

$$g_n = \int_{\Delta s} \gamma(s) \exp\left(-j2\pi\left(\zeta_n s + \frac{2d(s, t_n)}{\lambda}\right)\right) ds, \tag{3.5}$$

where Δs is the elevation extent of the object, λ is the wavelength, and $\zeta_n = \frac{-2b_n}{\lambda R}$ where R is the slant range and b_n denotes the perpendicular baseline. The variable $d(s, t_n)$ is the relative LOS motion as a function of elevation and time, which can be modeled by a linear combination of M base functions $q_m(t_n)$ as [102], [103]:

$$d(s, t_n) = \sum_{m=1}^{M} c_m(s) q_m(t_n), \tag{3.6}$$

where $c_m(s)$ is the corresponding motion coefficient to be retrieved by differential To-moSAR. The relevant motion models to be considered in urban areas are linear and seasonal (thermal) deformation ($M = 2$) which, by defining the mth temporal frequency component at t_n as $\eta_{m,n} = \frac{2q_m(t_n)}{\lambda}$, are characterized as follows [102]:

- linear motion with $\eta_{1,n} = \frac{2t_n}{\lambda}$ and the coefficient $c_1(s)$, which stands for the LOS velocity v as a function of s.

- seasonal motion with $\eta_{2,n} = \frac{2\sin(2\pi(t_n-t_0))}{\lambda}$ and the coefficient $p_2(s)$, which stands for the amplitude of the periodic motion a while t_0 is the initial phase offset. Note that t_n and t_0 should be specified in year. Alternatively, a temperature-dependent motion can be considered.

The differential TomoSAR imaging model described in Equation 3.5 is further discretized along s and the motion parameters space (c_1, c_2). The discrete system model with the presence of noise ϵ is written as [103]:

$$\mathbf{g} = \mathbf{R}\gamma + \epsilon \tag{3.7}$$

where $\mathbf{g}$ is the measurement vector with N elements, γ is the reflectivity function to be reconstructed and $\mathbf{R}$ is an irregularly sampled mapping matrix. Since the dimension of γ is larger than $\mathbf{g}$, Equation 3.7 represents an underdetermined system of equations, which is solved through regularization. The TomoSAR results of this thesis have been obtained through l_2-norm regularization of γ, using Singular Value Decomposition (SVD) with Wiener filtering on $\mathbf{R}$ [12]:

$$\hat{\gamma} = \arg\min_{\gamma} \left\{ \|\mathbf{g} - \mathbf{R}\gamma\|_2^2 + \alpha_{l_2} \|\gamma\|_2^2 \right\}, \tag{3.8}$$

where α_{l_2} is the regularization parameter. For a list of more advanced spectral estimators commonly used to invert Equation 3.7, the reader is referred to [103].

With the theoretical background of differential TomoSAR already explained, the workflow of the method can be described in the following three steps:

1. Differential inteferograms are generated with respect to a single master similar to PSI. The atmospheric phase is mitigated from each SLC. Both the amplitude and the differential phase data are the input for TomoSAR processing.

2. The reflectivity profile γ of each pixel is reconstructed using Equation 3.8 or more advanced techniques if super resolution in elevation is desired.

3. The recovered $\hat{\gamma}$ of the previous step is used for the estimation of elevation and the deformation parameters of multiple scatterers within each pixel through the detection of peaks of the reflectivty profile.

In the following, the recent advances on spaceborne differential TomoSAR, mostly relevant to the utilization of different spectral estimation techniques (Step 2), is summarized.

3.4.2 Recent Advances in TomoSAR

The first tomographic experiments on real SAR data of an L-Band airborne system was reported in [8] in which an application of the method for forest height mapping was demonstrated. The problem of retrieving the radar reflectivity of multiple overlaid phase centers, taking into account the statistics of speckle, was addressed in [104]. The 3-D focusing of multipass SAR data was treated in [9] where the authors employed a truncated SVD approach for reconstructing γ.

The differential form of SAR tomography was introduced in [10] by taking into account the linear motion of multiple scatterers within a pixel. Practical demonstration of multiple scatterer detection using medium resolution ERS data and tracking their linear deformation were shown in [105]. The differential TomoSAR system model was extended in [102] to accomodate nonlinear motion models.

With the launch of high resolution SAR systems like TSX, the monitoring of urban areas became increasingly on focus. The usage of SVD with Wiener filtering was first shown in [12] using TSX spotlight products. This has been followed by introducing the compressive sensing based algorithm, SL1MMER, in [106] which retrieves γ by l_1-norm regularization. The robustness and the super resolution capabilities of SL1MMER have been studied in [107]–[109] and a fast algorithm for performing the inversion has been proposed in [110]. The SL1MMER algorithm has been also extended to exploit group sparsity in urban areas (M-SL1MMER), which produces similar results to SL1MMER with significant smaller number of data takes [111].

As these methods and in general TomoSAR algorithms are computationally expensive, the combination of PSI and TomoSAR has been also proposed in [88], [112].

3.5 Coregistration of Multi-Aspect InSAR Data

In the SAR community, the coregistration of multi-aspect InSAR data to produce a single detailed 3-D product is referred to as point cloud fusion. The fused result has a few advantages compared to single-aspect point clouds such as increase in the number of scatterers, recovering the radar shadow segments of each individual point cloud and enabling decomposing the inherent 1-D deformation of InSAR products into 3-D. For the latter two points, the availability of SAR images of cross-heading tracks is necessary.

In the following a short review on advances of point cloud fusion and motion decomposition based on multi-aspect data is provided. The review does not address the topic of multi-aspect InSAR DEM fusion [113]–[115] or multi-aspect SAR interferogram fusion [116] and treats the motion decomposition problem only for the case that SAR images from at least three nonplanar acquisition geometries are available. For a full review on different methods to generate 3-D InSAR deformation maps, using amplitude tracking techniques or a combination of InSAR and GNSS, the reader is referred to [34], [117].

3.5.1 InSAR Point Cloud Fusion

Relevant to PSI, one of the first fusion results on ERS data was presented in [118], which was based on the cross-correlation of binary images of multi-aspect geocoded PSI point clouds over Milan, Italy. Due to the usage of medium resolution data and the simple method to find a coarse 2-D shift between the point clouds, the precision of the method has been reported to be 5-10 m [118].

Another notable approach was reported in [119], where the authors transformed multi-track PSI point clouds to a common datum defined by a reference track. They used the radar position of PS fields as well as their multi-image amplitude maps to estimate the datum parameters and avoided geocoding each PSI stack separately. Although the method is suitable for medium resolution data over flat areas, it is not usable for PSI results of high resolution products and also does not remove the vertical and horizontal bias of PS locations caused by the unknown height of the reference point of the reference track [97].

The fusion method described in [25] resolves the aforementioned problem of the reference point and has been used on multi-aspect PSI point clouds reconstructed from TSX high resolution spotlight images. It is based on a least-squares matching scheme that minimizes the distances between assumed identical points of two or more point clouds, from either adjacent or opposite viewing geometries [25], [97]. This method will be later explained in more detail in Subsection 4.2.1.

In [26], an alternative feature-based fusion algorithm is proposed, which is based on automatic detection and matching the so-called L-shapes of high rise buildings from InSAR point clouds. This method is computationally more efficient than the one introduced in [25] due to the reduced number of points in the matching step and is highly suitable for TomoSAR point clouds as the density of facade points are high.

In the airborne SAR domain, a radargrammetric approach using a quadruple CR as a single GCP for coregistration of two pair of cross-heading data sets were used in [120] and a method similar to the one described in [116] was proposed for multi-aspect multi-baseline InSAR over urban areas in [121].

3.5.2 1-D to 3-D Motion Decomposition

The availability of a fused InSAR product from multi-aspect data stacks enables tackling one of the major problems in InSAR namely decomposing the 1-D deformation into their original 3-D displacement vector. The theoretical aspects of the problem have been discussed in [122], [123].

In [124], two right-looking cross-heading pair of interferograms were used to recover the vertical and the East-West deformation, caused by an earthquake, while ignoring the North-South deformation.

For a special case, in regions with latitudes greater than $80°$, the full 3-D deformation of a glacier, considering also the North-South component, was reconstructed using two pairs of cross-heading data stacks of RADARSAT in [33].

The first result on small-scale deformation monitoring of single buildings was reported in [11], where the analysis was performed on a detailed fused TSX PSI point cloud over Berlin, Germany. In this study both the linear displacement rate and the seasonal deformation were decomposed to vertical and horizontal components while neglecting the deformation in the North-South direction.

4 Summary of the Contributions

The research objectives of this thesis, outlined in Section 1.2, have been addressed in five peer-reviewed journal papers that are found in the appendices. In this chapter, a summary of the articles is provided. The main addressed contributions are:

A) The proposal of a framework for automatic extraction of GCPs from multi-aspect high resolution SAR data.

B) The first demonstration of geocoding error correction of InSAR point clouds without the use of auxiliary positioning data.

C) The production of absolute 3-D point clouds and performing geodetic point cloud fusion by combining SAR geodetic techniques and InSAR-based approaches.

D) The first attempt for 1-D to 3-D motion decomposition exploiting multi-aspect TomoSAR point clouds.

This chapter continues with an introduction to the test sites as well as the SAR data used in this thesis in Section 4.1. Regarding the contributions, A) is summarized in Section 4.2 and is treated in detail in Appendix A.1. A summary of B) and C) is reported in Section 4.3 and in Appendix A.5 with demonstrations on PSI and TomoSAR point clouds in Appendices A.2 and A.3, respectively. Finally, the work related to D) is explained in Section 4.4, which is a summary of the article presented in Appendix A.4.

4.1 Data Set and Test Sites

The developments of this thesis have been tested on three data sets over Berlin, Germany and over Oulu located in Finland. The used SAR images are all from the high resolution spotlight mode of TSX with a range bandwidth of 300 MHz. The acquisition parameters of the SAR images are summarized in Table 4.1.

   The Oulu data set includes a total of 177 images in four stacks, two acquired from ascending orbit tracks and two acquired from descending ones. The coverage of each stack is marked with colored rectangles in Figure 4.1. Four PSI point clouds were generated from the Oulu data using the PSI-GENESIS [85] of the German Aerospace Center (DLR). The point clouds were used in Section 4.2 for automatic GCP generation and in Subsection 4.3.2 for the demonstration of geodetic point cloud fusion.

   The Berlin data set also includes four stacks of SAR images acquired from two pairs of cross-heading orbits similar to the ones from Oulu. The extent of Berlin data stacks
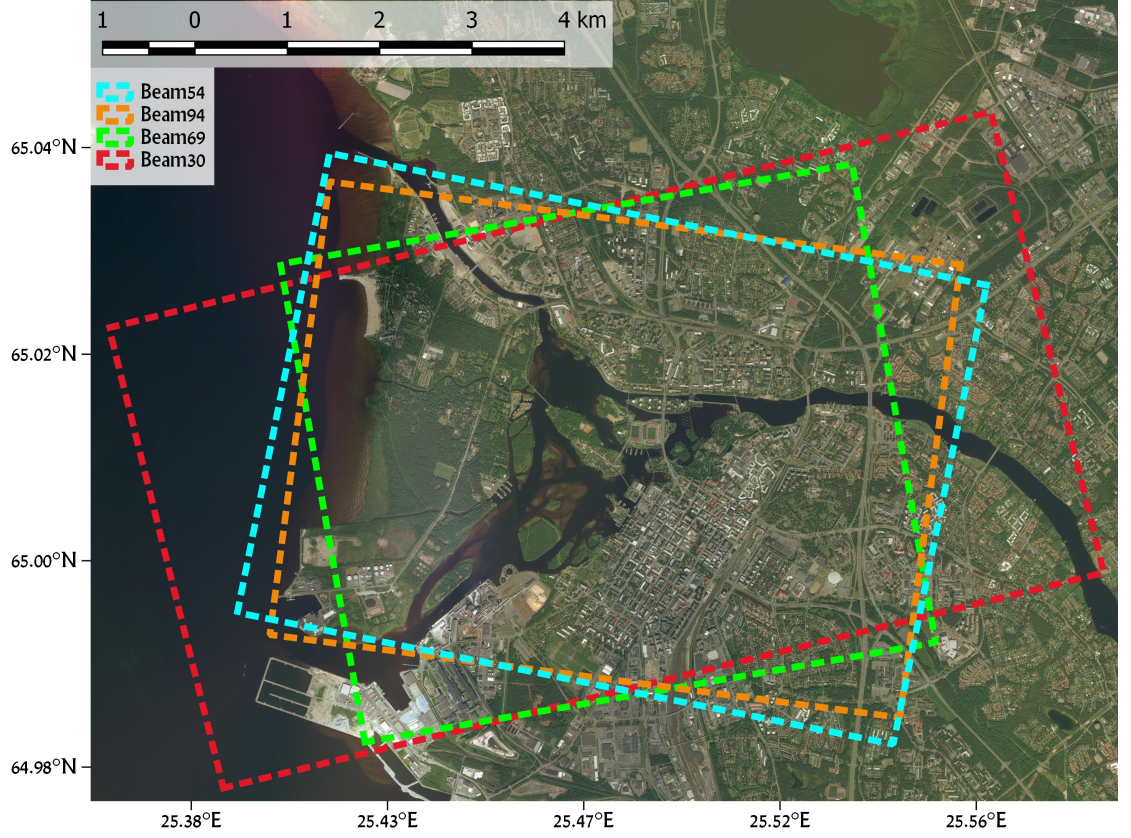
Figure 4.1: Optical image of Oulu taken from Bing images. The mean scene coverage of the TSX images are marked with colored rectangles.

Table 4.1: Acquisition parameters of the processed stacks of TSX high resolution spotlight images of Berlin and Oulu.

Processing Scenario	Beam Number	Period	Center θ	Average α	Orbit Direction	Number of Images
Berlin TomoSAR	42	2008−2013	36.1°	190.6°	Descending	109
	57	2008−2013	41.9°	350.3°	Ascending	102
	85	2008−2013	51.1°	352°	Ascending	111
	99	2008−2013	54.7°	187.2°	Descending	138
Berlin PSI	42	2010−2015	36.1°	190.6°	Descending	107
	57	2010−2015	41.9°	350.3°	Ascending	102
Oulu PSI	30	2014−2016	30.9°	346.1°	Ascending	44
	54	2014−2016	41.1°	191.4°	Descending	44
	69	2014−2016	46.2°	350°	Ascending	38
	94	2014−2016	53.4°	187.5°	Descending	51

and the underlying optical image of the city is shown in Figure 4.2. The data is categorized in two groups with two different time spans. The first group includes a total of 460 images from all stacks, which were used to produce four TomoSAR point clouds using the Tomo-GENESIS [125] of DLR. The TomoSAR point clouds are the basis of the work carried out in Sections 4.3 and 4.4. The second group consists of 214 images from
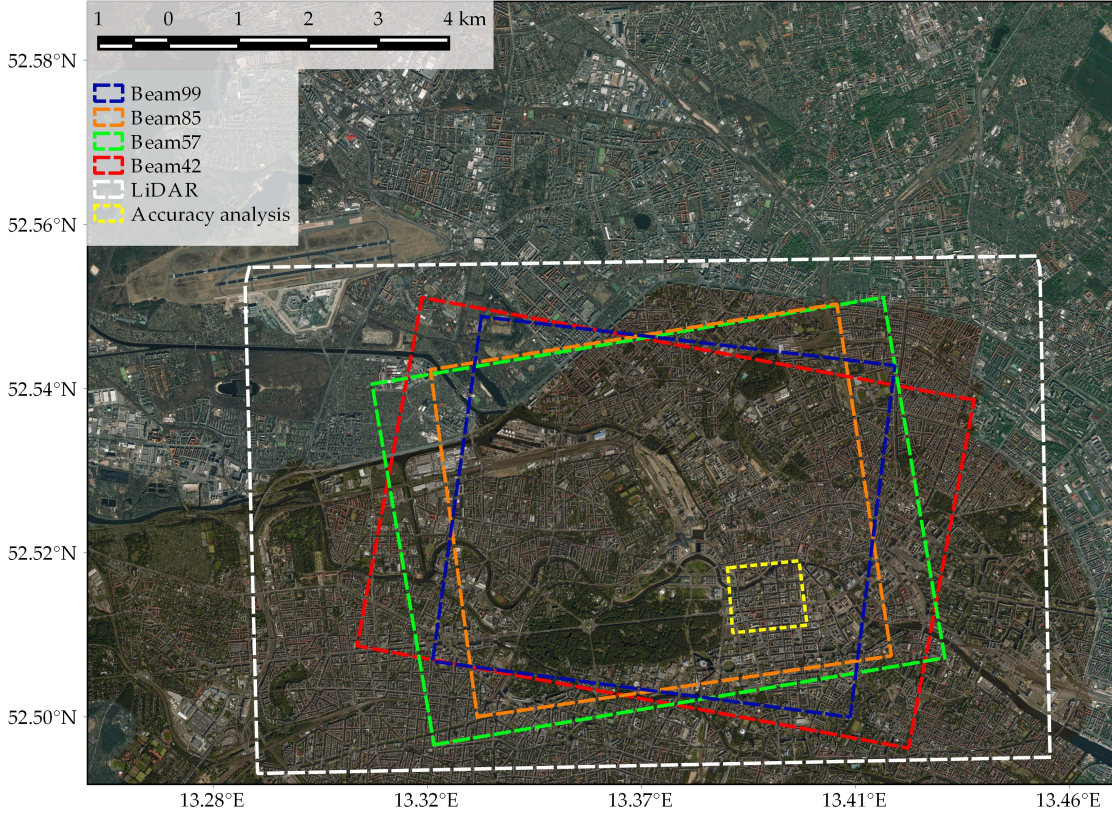
Figure 4.2: Optical image of Berlin taken from Bing images. The mean scene coverage of the TSX images as well as the extent of the reference LiDAR data is visualized. The test site used for localization accuracy analysis in Section 4.3 is also marked.

Beam42 and Beam57, acquired from 2010 to 2015, which were generated using the latest version of the TSX SAR processor and were processed to obtain two PSI point clouds. These products were used in Section 4.3.

In addition to SAR data, an aerial optical image with pixel ground spacing of seven centimeters is also available for Berlin as well as an accurate LiDAR point cloud. The latter serves as reference for 3-D localization accuracy analysis of InSAR point clouds. For Oulu, a GIS road network data of the city is available that is used for GCP candidate detection in Section 4.2. Furthermore, four frames of ortho-photo aerial images of Oulu are available for visualization purposes.

4.2 Automatic GCP Extraction from SAR Data

A framework for automatic identification and positioning of SAR-based GCPs from urban areas has been introduced in [77], which can be found in Appendix A.1. The processing chain requires multi-aspect SLC SAR images of the investigated scene and their corresponding product annotation files as input. It delivers absolute 3-D coordinates

of the GCPs as well as their posterior standard deviations as output. The algorithm consists of five major parts (for detailed flowchart refer to Appendix A.1):

1. Identification of identical scatterers visible in multi-aspect SAR images.

2. Precise extraction of scatterers' azimuth and range positions from SAR images at sub-pixel level.

3. Scatterer visibility check and initial removal of outliers from time series of phase noise.

4. Correction of radar timings for all the perturbing signals using the SAR imaging geodesy method.

5. Estimation of the 3-D absolute coordinates of the scatterers using stereo SAR.

The innovations of the framework are mostly related to the first and the third steps, which are briefly described in the following in Subsections 4.2.1 and 4.2.2. Furthermore, a practical demonstration of the positioning of large number of GCPs is given in Subsection 4.2.3.

4.2.1 Detection of Identical PS in Multi-Aspect SAR Images

There are two types of PS in urban areas, which can be considered as suitable GCP candidates. One includes the cluster of facade PS, which are assumed to originate as triple bounces from window corners of buildings, see Figure 4.3a. These PS are available in large numbers and form the basis of GCP generation using SAR images of same-heading orbit tracks. The other type of PS includes the isolated focused points such as lamp poles, traffic lights and targets alike, see Figure 4.3b. These targets are valuable in GCP generation, as they can be recognized in SAR images from cross-heading viewing geometries thanks to their simple reflection properties. GCPs localized from cross-heading orbits are always preferred because theoretically their 3-D position is retrieved with higher precision compared to the same-heading targets.

In any of the same- or cross-heading scenarios, the task of identical scatterer detection is challenging, and it is of course the essential prerequisite of SAR-based GCP generation. As reported in Appendix A.1, the candidate detection step is facilitated by selecting identical PS from same-heading orbits relying on fusion of multi-track PSI point clouds and from cross-heading orbits based on lamp pole detection using high resolution optical data or utilizing vector road network data. These techniques are more efficient in multi-aspect single PS matching than methods using Scale Invariant Feature Transform (SIFT) algorithm [126] or its modifications for SAR [127], [128]. The SIFT-based approaches are only applicable to SAR images taken from same-heading orbits with small difference of incidence angles.

(a)

(b)
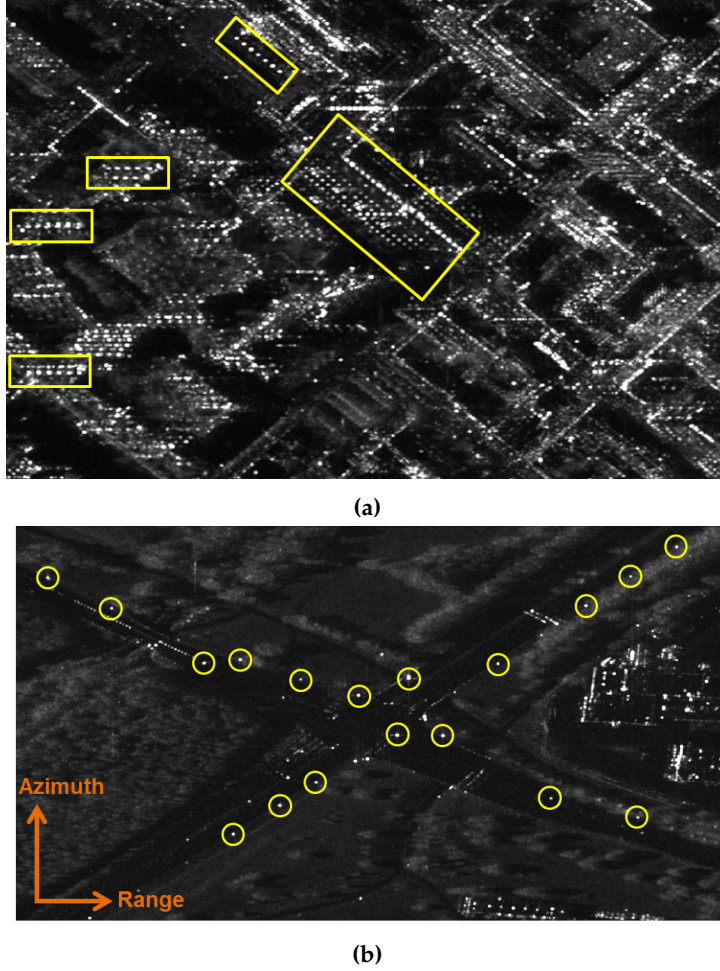
Figure 4.3: Suitable candidates for SAR-based GCP generation in urban areas. (**a**) shows examples of facade PS marked with yellow rectangles. (**b**) shows examples of lamp poles visible as bright focused isolated points within yellow circles.

4.2.1.1 Multi-Track PSI Point Cloud Fusion

In [25], a method for geometrical fusion of multi-track PSI point clouds has been proposed. The fusion algorithm, which is based on the geocoded PSI point cloud solutions of each geometry consists of three major parts, namely: 1) generation of initial point correspondences, 2) restricted least squares adjustment to minimize the distance between assumed identical points visible from different viewing geometries, and 3) adding a range-dependent shift to all PS using the result of the previous step for the final registration. For our task, which includes the detection of PS correspondences from same-heading tracks the following steps have been carried out:

1. Full PSI processing and geocoding of each SAR image stack separately, which has been performed by the PSI-GENESIS of DLR [84].

2. Coregistration of the geocoded PSI point clouds from different viewing geometries using the above-mentioned fusion algorithm. The result for two descending stacks of Oulu, colorized with different shades of gray, is shown in Figure 4.4a including 32000 PS correspondences visualized in yellow. The Euclidean distances between the matched pairs vary from 1.5 m to 5 m.

3. Reduction of point correspondences to the ones with high quality and closer distance by imposing a 10 m × 10 m grid on the point cloud. Within each grid cell only the PS pairs closer than 3 m and an ADI value lower than 0.25 [6] are chosen. In this way the homogeneity of the distribution of GCP candidates is also guaranteed. The results before and after reduction for the descending stacks of Oulu can be seen in a small area in Figure 4.4b and Figure 4.4c, respectively.

4. Radar-coding of the PS pairs obtained in the previous step to the corresponding non-coregistered SAR images.

The result of these steps is the coarse azimuth and range coordinates of detected PS pairs in SAR images from same-heading orbit tracks.

4.2.1.2 Template Matching on Optical Data

This method aims at the detection of identical PS from SAR images of cross-heading orbit tracks based on the available ortho-rectified high resolution optical image of the scene [74], [76]. It includes:

1. Preprocessing of the optical image such as smoothing, histogram equalization and simple sharpening [129].

2. Extraction of the template image based on the shadow of an arbitrary lamp pole visible in the optical image. The template is then correlated with the reference image to calculate the following similarity measure for each pixel (u,v) in the reference image [130]:

$$\rho(u,v) = \frac{\Sigma_{x,y}\left[\mathbf{I}(x,y) - \bar{\mathbf{I}}_{u,v}\right]\left[\mathbf{T}(x-u,y-v) - \bar{\mathbf{T}}\right]}{\sqrt{\Sigma_{x,y}\left[\mathbf{I}(x,y) - \bar{\mathbf{I}}_{u,v}\right]^2 \Sigma_{x,y}\left[\mathbf{T}(x-u,y-v) - \bar{\mathbf{T}}\right]^2}}, \qquad (4.1)$$

where $\mathbf{I}(x,y)$ and $\mathbf{T}(x,y)$ denote pixel values of the reference and the template image at (x,y), respectively, and $\Sigma_{x,y}$ stands for $\Sigma_{x=1}^{N_1}\Sigma_{y=1}^{N_2}$ with $N_1 \times N_2$ being the size of the template. The variables $\bar{\mathbf{I}}_{u,v}$ and $\bar{\mathbf{T}}$ denote the mean intensity values of the original image and the template, respectively. Equation 4.1 allows for calculation of the Normalized Cross-Correlation (NCC) value $\rho(u,v)$, which leads to the detection of the template location in the reference image after proper thresholding.

(a)

(b)   **(c)**
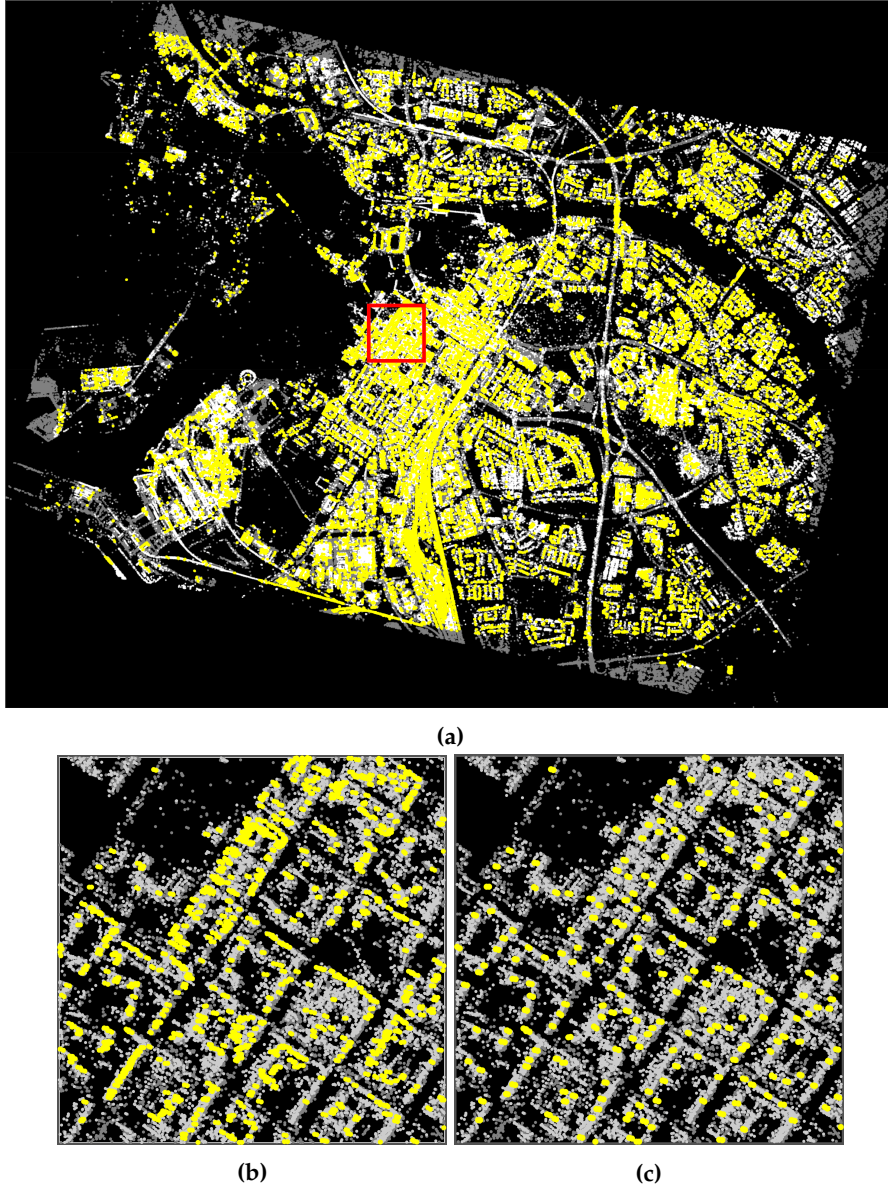
Figure 4.4: Depiction of PS correspondence detection from SAR images of same-heading orbit tracks of Oulu using the multi-track PSI fusion algorithm proposed in [25]. (**a**) shows the geocoded descending PSI point clouds of Oulu as light and dark gray points as well as the detected PS correspondences in yellow. (**b**) and (**c**) show a zoom-in area of (**a**) marked by the red rectangle before and after imposing a 10 m grid in which the pairs with closest distance are chosen.

3. The pixels with NCC value higher than 0.6 are selected and geo-referenced to represent the lamp poles. A nonparametric clustering using the mean shift concept is then carried out [131]:

$$\mathbf{M}(\mathbf{p}_i) = \frac{\Sigma_{j=1}^{n} \mathbf{p}_j \, g\left(\|\frac{\mathbf{p}_i - \mathbf{p}_j}{h}\|^2\right)}{\Sigma_{j=1}^{n} g\left(\|\frac{\mathbf{p}_i - \mathbf{p}_j}{h}\|^2\right)} - \mathbf{p}_i, \tag{4.2}$$

where $\mathbf{p}_i$ denotes a 3-D point for which the shift vector $\mathbf{M}(\mathbf{p}_i)$ is calculated. $\mathbf{p}_j$ represents the points in a neighborhood of $\mathbf{p}_i$, g is a kernel function with the bandwidth h and $\|\cdot\|$ is the Euclidean distance operator. The process is carried out iteratively until the length of $\mathbf{M}(\mathbf{p}_i)$ becomes equal or close to zero. For our application, since in any case there will be a mismatch between the detected points on optical data and the corresponding bright points in the SAR image, utilizing a flat kernel in Equation 4.2 suffices. This means the algorithm is simplified by calculating the sample mean in a specified radius of $\mathbf{p}_i$ and shifting the desired point towards the estimated center.

4. Radar-coding the cluster centroids, assumed to be the base of lamp poles, obtained in the previous step to the corresponding non-coregistered SAR images.

5. Matching the detected lamp poles to their corresponding bright points in the SAR images by employing the iterative closest point (ICP) algorithm [132].

An example of applying the aforementioned algorithm is shown in Figure 4.5 where 44 candidate GCPs were detected using an ascending and a descending stack of Berlin.

4.2.1.3 Vector Road Network Data

In urban areas, the cylindrical objects of our interest (lamp poles, road signs, traffic lights, etc.) are typically located along the roads. Therefore, with the availability of geospatial road data, either obtained from OpenStreetMap [133] or country-specific geoportals, and the projection of such maps on SAR images, one can search for PS in the neighborhood of the road data points. The method is applied to coregistered stacks of SAR images and consists of the following workflow:

1. Radar-coding the road data, which is usually delivered in the UTM coordinate system, based on the master orbit information of each stack. If no height information is available for the road data, a constant mean DEM height can be assumed for radar-coding.

2. Within a circular neighborhood around each road data point on the SAR image, the amplitude dispersion index (ADI) is evaluated:

$$D_a \approx \frac{\sigma_a}{\bar{a}}, \tag{4.3}$$

where σ_a and $\bar{a}$ are the temporal standard deviation and the temporal mean of calibrated amplitude values of the pixel, respectively, and D_a is a proxy for the phase dispersion.

(a)   **(b)**   **(c)**

(d)

(e)
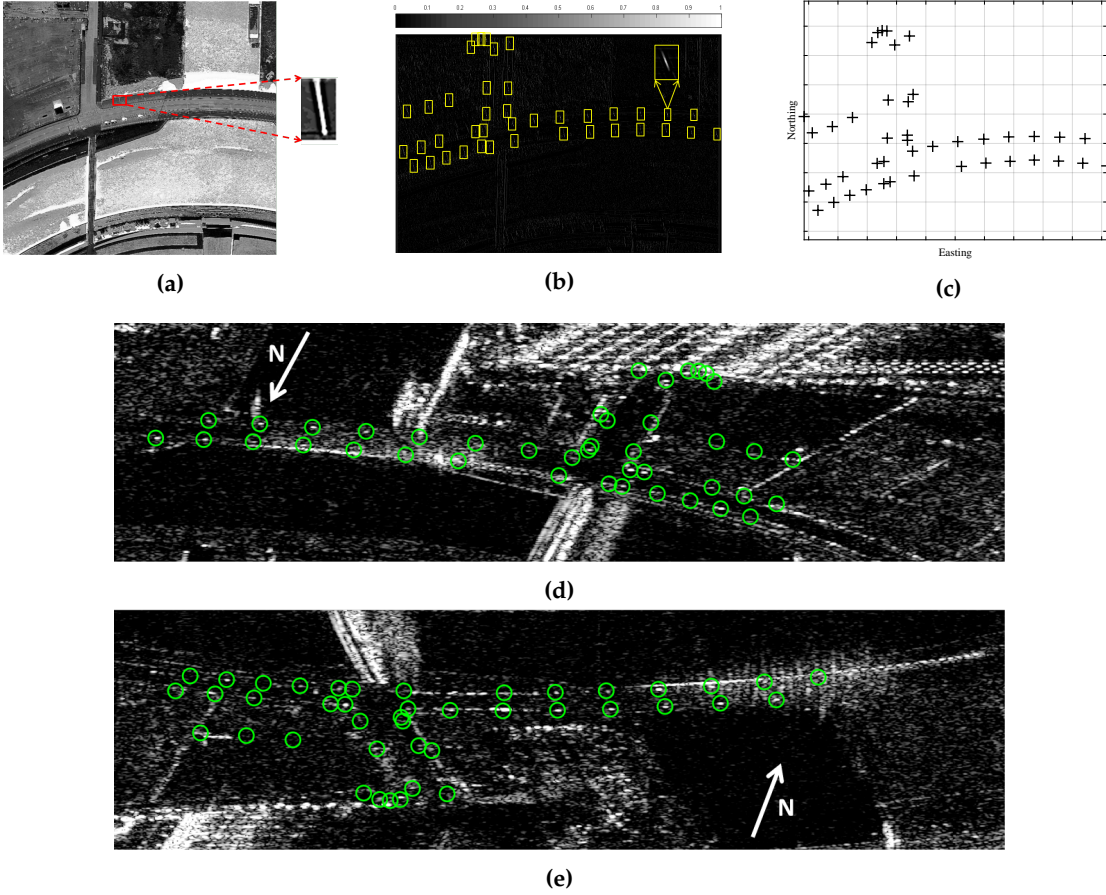
Figure 4.5: Demonstration of PS correspondence detection in Berlin based on high resolution optical data. (**a**) shows the preprocessed optical image and the extracted template. (**b**) is the calculated NCC map after correlating the extracted lamp pole template with the reference image in which the detected objects are marked by yellow rectangles. (**c**) shows the 44 detected objects in the UTM coordinate system after clustering. (**d**) and (**e**) show the results of matching after using the ICP algorithm on the descending and the ascending image, respectively. In the last two subfigures, it can be seen that the detected objects from the optical image (green circles) coincide with the visible bright points in the SAR images.

3. The pixel with the lowest value of D_a, i.e. the one with the highest phase stability is chosen as potential PS candidate. Since at this point, it is possible that many false pixels with relatively low D_a values in the neighborhoods are categorized as potential GCP candidates, a further thresholding on D_a is performed in SAR images from all available viewing geometries.

4. Geocoding of the presumable identical PS using the respective master orbit information and selecting the pairs, which have a distance less than a couple of meters as the final correspondences
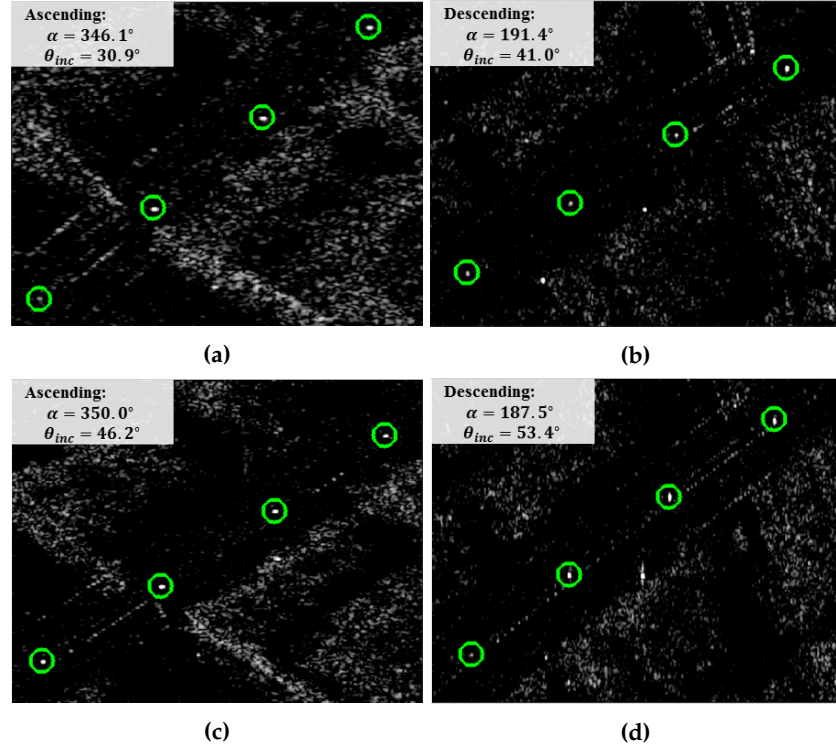
Figure 4.6: PS correspondence detection from four different viewing geometries. For each image, the respective averaged heading angle α and the averaged incidence angle θ_{inc} are stated.

This method is quite straightforward to implement and delivers an acceptable number of candidates visible in SAR images of cross-heading orbits. An example of the detected PS correspondences from four different viewing geometries (two ascending, two descending) is shown in Figure 4.6 for a highway in Oulu.

4.2.2 Robust Outlier Removal

The PS correspondence detection approaches, described in subsection 4.2.1, allow for the retrieval of approximate radar coordinates of the targets. These pixel positions are refined using PTA to obtain the sub-pixel location of the scatterers [20], [39]. After employing PTA for each scatterer in all the available non-coregistered SAR images, the following steps are carried out for removing outliers and performing a PS visibility check:

1. Based on PTA, the SCR of each PS is computed. For a PS in the ith SAR image, its SCR_i is related to its phase noise σ_{ϕ_i} in radians as [82]:

$$\sigma_{\phi_i} \approx \frac{1}{\sqrt{2\,SCR_i}}. \tag{4.4}$$

2. The time-series of phase noise σ_ϕ is analyzed for each GCP candidate. The values of σ_ϕ are non-negative and follow a right-skewed distribution. Therefore we use a method called adjusted boxplot for outlier removal [134]. The method modifies the original boxplot, described in [135], to include information about the skewness of the data. Thus, instead of classifying an observation as outlier if it lies outside of the interval defined by the boxplot method [135]:

$$\left[Q_1 - 1.5 \; IQR; \; Q_3 + 1.5 \; IQR\right], \tag{4.5}$$

adjusted boxplot declares an observation as outlier if its value exceeds the following interval [134]:

$$\left[Q_1 - 1.5 \; e^{(-4MC)} \; IQR; \; Q_3 + 1.5 \; e^{(3MC)} \; IQR\right]. \tag{4.6}$$

In Equation 4.5, Q_1 and Q_3 are the first and the third quartiles of the data, respectively and $IQR = Q_3 - Q_1$ denotes the interquartile range. In Equation 4.6, MC is the medcouple, a robust measure of the skewness of a univariate sample which for right-skewed distributions is always non-negative [136]. The exponential functions in Equation 4.6, which depend on the MC as well as the included coefficients, are chosen experimentally based on some well-known skewed distributions [134].

3. After the removal of outliers in the previous step, as a target visibility check, all acquisitions within which the σ_ϕ value of the PS are higher than 0.5 radians ($\approx$ 30°) are excluded.

The result of this step includes the outlier-free sub-pixel position of each candidate GCP, which is converted to radar timings. This provides the input for the last two remaining parts of the GCP extraction algorithm, i.e. usage of imaging geodesy for correction of timings and the eventual stereo SAR processing for estimating the absolute 3-D coordinates.

4.2.3 Positioning of GCPs

In the following, the applicability of the proposed automatic GCP extraction chain is demonstrated using the TSX data introduced in Section 4.1.

4.2.3.1 Berlin

After the detection of the initial GCP candidates using a high resolution optical image of Berlin, depicted in Figure 4.5, the absolute position of the 44 selected PS is calculated by performing steps two to five of the algorithm described in the beginning of Section 4.2. Only nine PS out of 44 were proved to be of high quality to be considered as the final GCP candidates. The reason is explained in detail in Appendix A.1. However,
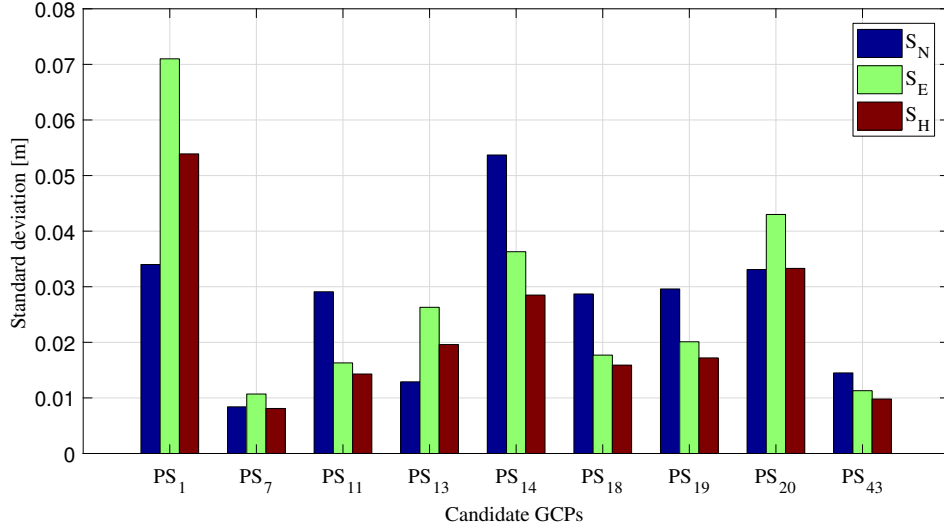
Figure 4.7: The standard deviations of the best nine GCP candidates estimated posterior to the geodetic stereo SAR processing. These values are defined in the local coordinate system of Berlin in North, East and Height (S_N, S_E, S_H) and are scaled to 95% confidence level.

considering the small spatial extent of the test site, nine high quality GCP will certainly suffice for any geodetic applications.

In Figure 4.7, the precision of the estimated coordinates is reported for the best nine GCPs. The standard deviations are defined in the local coordinate system of Berlin in the North, East, and Height (S_N , S_E, S_H) in the confidence level of 95%. The mean values of S_N, S_E and S_H are 2.7 cm, 2.8 cm and 2.2 cm, respectively. The cross-heading geometry has a significant impact on the precision of height components as it was expected.

Apart from the analysis of the posterior standard deviations of the GCPs, their height values were compared with the reference LiDAR data set of Berlin. We assume the phase centers of the detected GCPs are at the base of the lamp poles on the ground. Therefore, the cross-comparison includes finding the nearest neighbors of the GCP candidate in the LiDAR point cloud within the radius of 1 m, excluding the LiDAR points with large height values which originate from the top of the lamp pole, estimating the mode of the LiDAR heights to represent the reference height and evaluating the difference between the ellipsoidal height of stereo SAR results with respect to the reference height. The results of the cross-comparison are reported in Table 4.2. The estimated stereo SAR and approximated LiDAR reference heights are denoted by h_S and h_L, respectively, while their offset is represented by h_o. It is seen that for all except for one of the GCPs the height offset is below 20 cm. The results report a bias of 13 cm and a precision of 5 cm overall with respect to the LiDAR data. The bias roughly implies the absolute accuracy of the height estimation using the stereo SAR method.

Table 4.2: The result of cross-comparison between the estimated heights of stereo SAR h_S and their corresponding LiDAR heights h_L. The offset h_o is an indicator for the absolute accuracy of h_S.

GCP index	h_L [m]	h_S [m]	$h_o = h_S - h_L$ [m]
PS_1	74.64	74.80	0.16
PS_7	74.45	74.54	0.09
PS_{11}	74.83	74.99	0.16
PS_{13}	75.40	75.62	0.22
PS_{14}	73.87	73.96	0.09
PS_{18}	73.87	73.95	0.08
PS_{19}	75.59	75.76	0.17
PS_{20}	75.02	75.18	0.16
PS_{43}	79.78	79.85	0.07
Mean			**0.13**
Standard deviation			**0.05**

4.2.3.2 Oulu

The identical PS detection step for Oulu has been carried out using the PSI fusion algorithm for same-heading tracks and utilizing road network data for cross-heading tracks, as demonstrated in Figure 4.4 and Figure 4.6, respectively. By applying the remaining steps of the GCP generation algorithm, the absolute 3-D coordinates of the candidates are obtained.

The average coordinate statistics of the high quality GCPs in Oulu are reported in Table 4.3. The scatterers are categorized based on the geometry configuration used for their positioning. The results are all expressed in centimeter and are defined in the local east, north and vertical coordinates within 95% confidence interval. From Table 4.3, it is seen that the averaged standard deviations are smaller than two decimeter for all the cases. As it was expected, the localization quality boosts as the difference in the viewing geometries becomes larger, which is the case when changing from Ascending-Ascending (AA) or Descending-Descending (DD) to the Ascending-Descending (AD) and eventually the quad geometry (ADAD) configurations. It is also evident that for cross-heading geometries the retrieval of the height component is the most precise one while for the same-heading cases, the precision in the north component is the highest. The only remaining concern regarding localization using cross-heading tracks is the diameter of the lamp poles, which may worsen the accuracy in the east coordinate component. This bias can be estimated and removed if the scatterer is also visible from same-heading tracks or if the diameter of the lamp pole is known *a priori*.

The distribution of the total 2049 generated GCPs is visualized on the optical image of Oulu in Google Earth in Figure 4.8. The scatterers are color-coded based on the underlying geometry configuration used for their localization. It is seen that almost the entire area of Oulu is covered with the generated GCPs. The ones from the same-heading geometries cover the built areas while the ones from cross-heading configuration include the base of lamp poles, street lights and traffic lights.
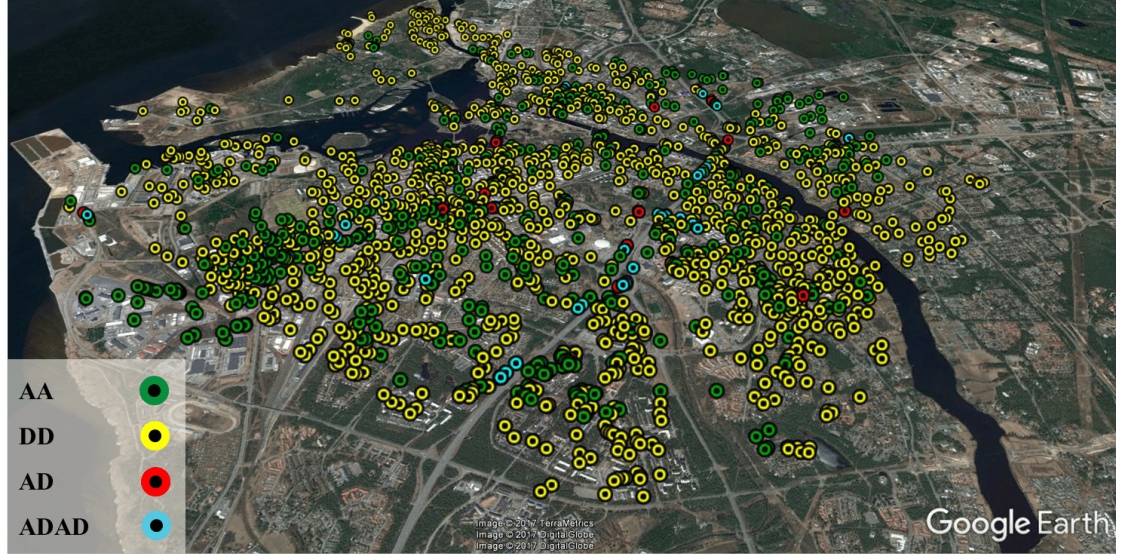
Figure 4.8: Total number of 2049 GCPs in Oulu color-coded based on the geometry configuration used for their positioning. The underlying optical image is taken from Google Earth.

Table 4.3: Averaged statistics based on the posterior 3-D coordinate standard deviations in Oulu. The letters A and D stand for ascending and descending geometries, respectively. The sample mean and standard deviation are denoted by μ and σ and $S_{[ENH]}$ represents the local coordinates standard deviations within a 95% confidence level.

Geometry	Number of Scatterers	μ_{s_E} [cm]	μ_{s_N} [cm]	μ_{s_H} [cm]	σ_{s_E} [cm]	σ_{s_N} [cm]	σ_{s_H} [cm]
AA	565	17.73	5.04	15.87	11.98	2.63	11.09
DD	1417	15.08	3.80	16.71	10.38	2.10	11.30
AD	24	2.26	2.50	1.75	0.99	1.11	0.83
ADAD	43	1.17	1.40	1.12	0.42	0.55	0.37

4.3 Geodetic InSAR

The concept of geodetic InSAR involves the utilization of SAR geodetic techniques in the framework of InSAR methods. This topic has been discussed in Appendix A.2 for correction of geocoding errors of PSI point clouds, in Appendix A.3 for absolute positioning of TomoSAR point clouds and for the demonstration of geodetic point cloud fusion, and also as a summary of the both aforementioned contributions in Appendix A.5.

4.3.1 Geocoding Error Correction for InSAR Point Clouds

The geocoding error correction methodology, proposed in Appendix A.2, mitigates the systematic effects caused by SAR timing errors, and estimates the unknown height of InSAR reference point (see Subsections 2.2.2 and 3.1.1). The flowchart of the approach
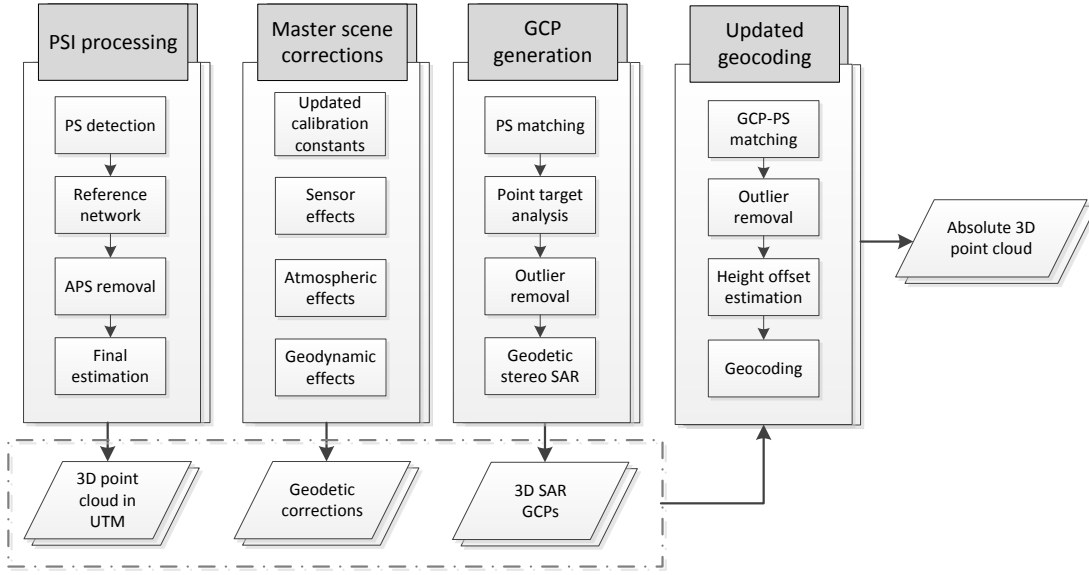
Figure 4.9: Workflow of the proposed geocoding error correction method. The processing steps, included in the big rectangles, are carried out independently and their results are shown by the parallelograms. The double shapes indicate that each processing is carried out for two or more SAR image stacks, which is the necessary condition for GCP generation.

is shown in Figure 4.9. Each block describes a separate processing, which produces an intermediary result shown as a parallelogram. The PSI processing in the first block [7], [83], results in the 3-D UTM coordinates of point clouds, for which the geodetic and atmospheric corrections, at the time of the master scene acquisitions, are calculated in the second block using the imaging geodesy approach [13], [56]. The third block includes the extraction of SAR-based GCPs for which either the explained method in Section 4.2 is used or the GCPs are generated using the SGP of DLR [75]. The results of the three previous steps, gathered in the dashed rectangle, are the input for the fourth step called updated geocoding, which produces the final 3-D absolute PSI point clouds. Note that the input for our approach includes a minimum of two SAR image stacks, indicated in Figure 4.9 as double shapes, acquired from different viewing geometries, i.e., from separate orbits. This is required for the extraction of highly precise 3-D GCPs. Since the first three blocks of the algorithm have been already explained, we focus here on briefly describing the processes of the fourth block. Starting from the extracted GCPs, the algorithm performs the following:

1. Filtering the GCPs based on their estimated posterior coordinate standard deviations. We impose a threshold of 10 cm for each coordinate component and further select the GCPs, which lie close to the road network data of the scene. This usually results in around 1000 high quality GCPs if the GCP extraction is performed using the method described in [75].

2. Radar-coding the GCPs onto the master scenes of each SAR image stack taking into account the full geodetic and atmospheric corrections. We then select the nearest neighbor of the GCP in the PSI point that has an ADI value lower than 0.4. The result is the correspondence of each GCP to its closest PS that most likely represents the base of lamp poles or traffic lights.

3. Removing wrong matched pairs from the previous step based on range and azimuth coordinate differences of the GCPs and their corresponding PS. Points with large coordinate differences are discarded based on the 2σ rule where the standard deviation is robustly estimated using Median Absolute Deviation (MAD) [137]:

$$\hat{\sigma}_{\mathrm{MAD}} = 1.4826 \, \mathrm{median}\Big(\big|\delta - \mathrm{median}(\delta)\big|\Big), \tag{4.7}$$

where δ denotes the vector of coordinate differences and 1.4826 is a correction factor making $\hat{\sigma}_{\mathrm{MAD}}$ consistent at Gaussian distributions.

4. Forming the ellipsoidal height difference histogram among GCPs and their corresponding PS, smoothing the histogram and detecting its peak as the approximate estimate of the DEM error. This process is carried out for each image stack separately.

5. Subtraction of the previously estimated DEM error from the ellipsoidal heights of the geocoded PSI point clouds to obtain absolute heights.

6. Finally, the absolute heights as well as the corrected range and azimuth timings of the scatterers and the orbit information of the master scenes are used for an updated geocoding, which produces the absolute 3-D PSI point clouds in the Cartesian coordinate system or in the UTM.

In the following, the applicability of the method is demonstrated using the Berlin TSX data introduced in Section 4.1. The generated absolute PSI point clouds are then compared to the reference LiDAR DSM of Berlin, which demonstrates the improvement in the geocoding accuracy of the PSI results.

4.3.1.1 Practical Demonstration in Berlin

The two stacks of Berlin, see Table 4.1, have undergone InSAR and PSI processing using the PSI-GENESIS of the DLR [82], [84], [85]. The PSI parameters of interest include DEM error, linear and periodic motion. The latter is modeled since we are analyzing an urban area with the acquisition time spanning over a few years. The geocoded topographic map generated from Beam42 is shown in Figure 4.10. It consists of approximately 1.4 million PS. A high density of PS is observed along the railways and also on building facades while white parts with no detected PS cover mostly vegetated areas.

At the next step, the geodetic and the atmospheric corrections are estimated and are subtracted from the SAR timings of all the PS using SGP [56]. The sum of all the
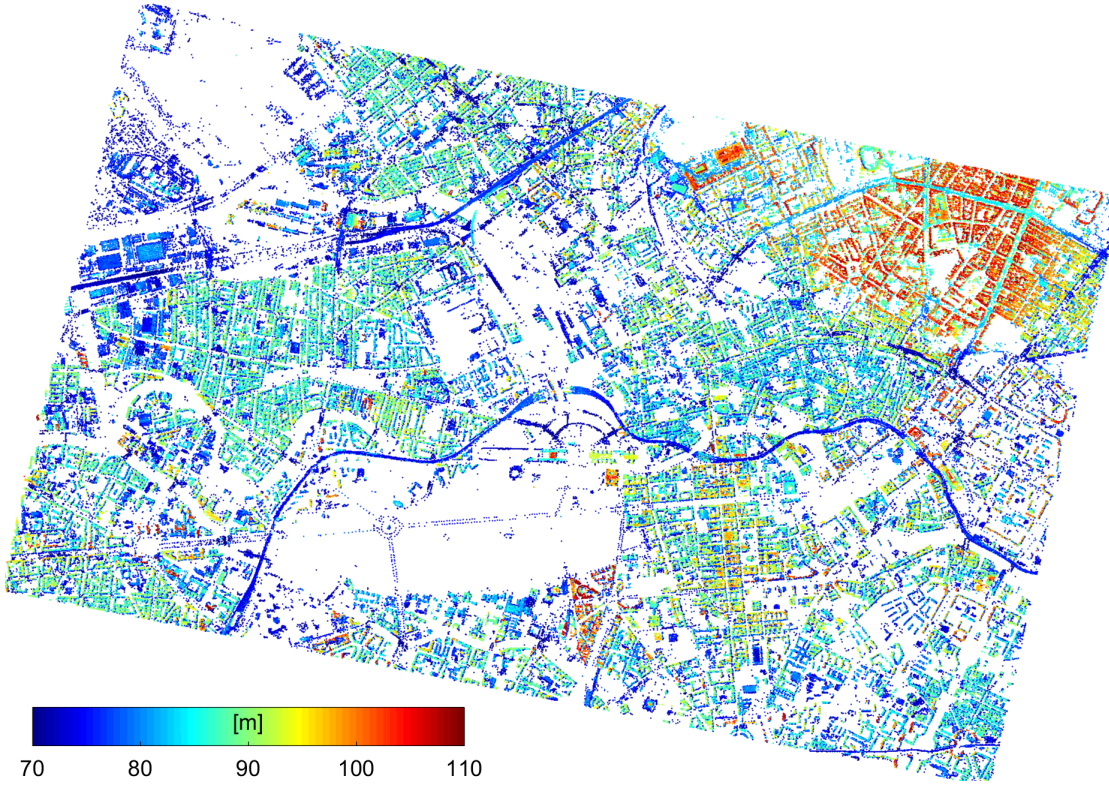
Figure 4.10: Geocoded descending PSI point cloud of Berlin in UTM coordinates. The ellipsoidal height is colorcoded. The *x*- and *y*-axis correspond to the UTM East and North, respectively.

corrections are reported in Figure 4.11 for the master acquisitions of each stack. The corrections are in the radar coordinate system where, the *y*-axis represents the azimuth coordinate and the *x*-axis depicts the range coordinate. The conversion to the unit of length is performed using Equation 2.14 for range errors and Equation 2.11 for azimuth errors, where $7050 m/s$ is assumed for the average beam velocity of the TSX satellite. It is seen that the magnitude of range errors is close to 3 m for Beam57 and approximately 2.75 m for Beam42, whereas the difference of maximum and minimum range errors over the scene is only 4 cm for both beams. The main contributions to the range errors come from the troposhperic error followed by geodynamic effects and finally the ionospheric delays. The azimuth errors are far less significant than range errors and manifest in sub-decimeter regimes, with millimeter variations across the scenes for both beams. The main source for azimuth errors is geodynamic effects followed by errors due to satellite dynamics. It is important to note that given the small squint angle of the TSX high resolution spotlight products and their operation in X-band, the tropospheric and ionospheric delays are negligible for azimuth timings and therefore are not considered in the SAR imaging geodesy method [18].

(a) Range error Beam57

(b) Azimuth error Beam57

(c) Range error Beam42

(d) Azimuth error Beam42

Figure 4.11: Range and azimuth error in the radar coordinate system (x-axis: range, y-axis: azimuth) of the master scenes of both investigated beams, which were acquired on 24.12.2011 (Beam57) and on 07.03.2012 (Beam42). Note that the scale of the colorbars for range error is in meter while the one for azimuth errors is in centimeter.

For GCP generation, 27 images from both beams are selected, which have been acquired within a period of 12 months, in 2014 and at the beginning 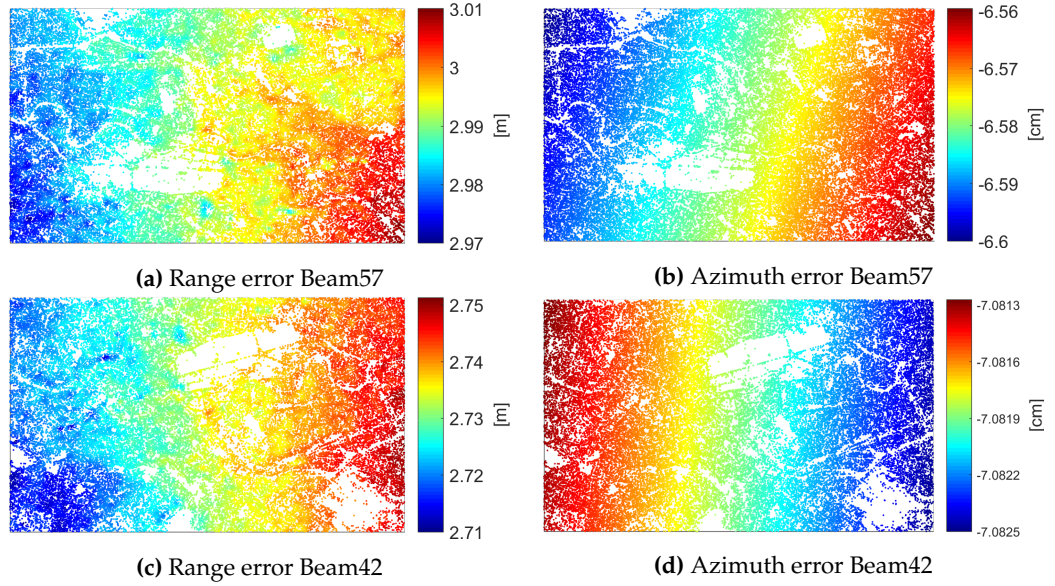of 2015. The reason for the time restriction is avoiding the impact of plate tectonics on the final GCP coordinates. The images are processed using SGP [75]. In total 17673 GCPs have been localized with precision of each GCP coordinate components being better than 1 m. The histogram of GCP precision values in local East σ_E, North σ_N and height σ_H is reported in Figure 4.12, where the vertical line at 10 cm denotes the applied threshold on the posterior coordinate standard deviations. The final GCPs are obtained after choosing only the ones close to the road network data of Berlin, which can be seen as purple dots overlaid on the geocoded PSI point clouds of both beams (in gray) in Figure 4.13. It is seen that the GCPs almost cover the full area within both PSI point clouds.

The matching of the GCPs to their corresponding PS in the PSI point clouds is carried out before evaluation of the DEM error of the reference point of each stack. This procedure is summarized in Figure 4.14. The left and the right subfigures correspond to results from Beam57 and Beam42, respectively. The 2-D scatterer plots show the distribution of range and azimuth coordinate differences among GCPs and their corresponding identified PS. In Figure 4.14a,b, the differences are plotted before outlier removal. It can be seen that very large deviations still occur after the correspondence detection step. Average deviations in the centimeter regime are observed in the range component while for the azimuth differences of Beam57, the mean deviation is close to 12 cm. The sample standard deviations, annotated in the subfigures, also show the large spread of the differences in both coordinate components in the meter regime. The
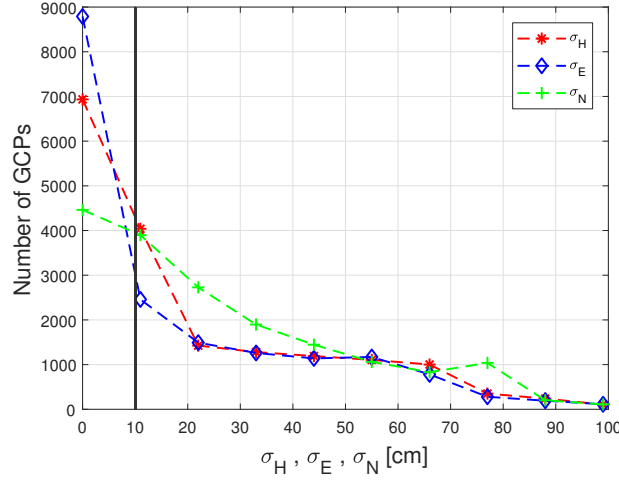
Figure 4.12: Histogram of GCP precision values in East σ_E, North σ_N and height σ_H.
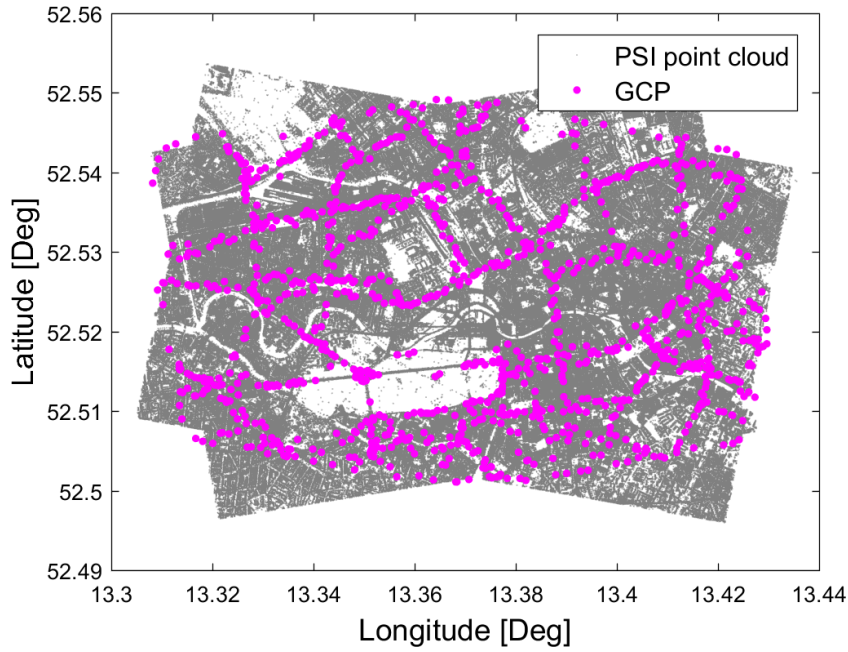
Figure 4.13: Distribution of the GCPs within the PSI point clouds of both beams.

second row of the subfigures, Figure 4.14c,d, depict the scatter plots after discarding outliers both in range and azimuth based on the 2σ rule. Note that the standard deviations for the outlier removal purpose are robustly estimated using MAD, as was explained in Subsection 4.3.1, while the annotated values in the subfigures are sample standard deviations, that are sensitive to outliers. From Figure 4.14c,d, it is clearly observed that the mean deviations decrease and now are close to 2 cm and 3 cm for range coordinate differences and approximately 2 cm and 1 cm for azimuth coordinate

differences of Beam57 and Beam42, respectively. It is important to note that the radar coordinate differences have been compensated for the continental drift which occurred between the master acquisition times of each beam, in 2011 and 2012, and the epoch with respect to which the GCPs have been localized, in 2015. This coordinate shift has been computed utilizing the average East, North, Up velocities from the nearest International GNSS Service (IGS) permanent station, in Potsdam approximately 45 km away from Berlin, and have been projected into the line-of-sight and the azimuth directions of each beam. Eventually, the third row, Figure 4.14e,f shows the histogram of height differences among GCP-PS pairs after discarding observations based on the 2σ rule. The peaks of the smoothed version of the height difference histograms provide the robust estimates of the DEM errors which are -4.06 m and -6.27 m for Beam57 and Beam42, respectively. It is expected that the DEM errors have different values because of non-identical PSI reference points. The estimated height shifts are subtracted from the final height estimates of PSI and along with the fully corrected range and azimuth timings, using SGP, an updated geocoding is carried out for each beam separately.

4.3.1.2 Cross-Comparison with LiDAR

The geocoded PSI point clouds before and after applying the aforementioned corrections are compared with the reference LiDAR data. It is important to note that, since an objective comparison is difficult to perform, due to the inherent differences between LiDAR and PSI point clouds, we avoid using the term validation and therefore reside with the word cross-comparison.

The cross-comparison for the 2-D horizontal coordinates is based on the extraction of facade PS from non-corrected and corrected PSI point clouds and evaluating their closeness to the corresponding building footprint in the LiDAR data. The facade extraction is performed by considering a neighborhood of 4 m around each PS, calculating the height variance in the neighborhood and considering a PS as facade point if the height variance is higher than 1.5 m [97]. For the 1-D vertical analysis, the facade points are excluded from the PSI point clouds and the LiDAR DSM, the histograms of ellipsoidal heights are formed containing two peaks which respectively represent ground and building roofs. The difference between the height values for which the ground peaks, in the PSI point clouds and in the LiDAR data occur, indicates the vertical accuracy.

The results of horizontal accuracy analysis for three selected test sites in Berlin is visualized in Figure 4.15. The green and red dots display the PS from ascending and descending point clouds, respectively, while the white dots show the extracted building footprints from LiDAR. All data points are presented in the UTM projection. The first and second rows correspond to the non-corrected and corrected results, respectively. Note that the subfigures have different scales. It is clearly seen that, before applying the geodetic corrections and DEM error compensation, red dots are located far away from the building footprints. This 2-D shift is largely compensated after applying the corrections in all the three cases. Moreover, the endpoint of each facade in the corrected point clouds matches with the endpoint of the facade from the point cloud from the
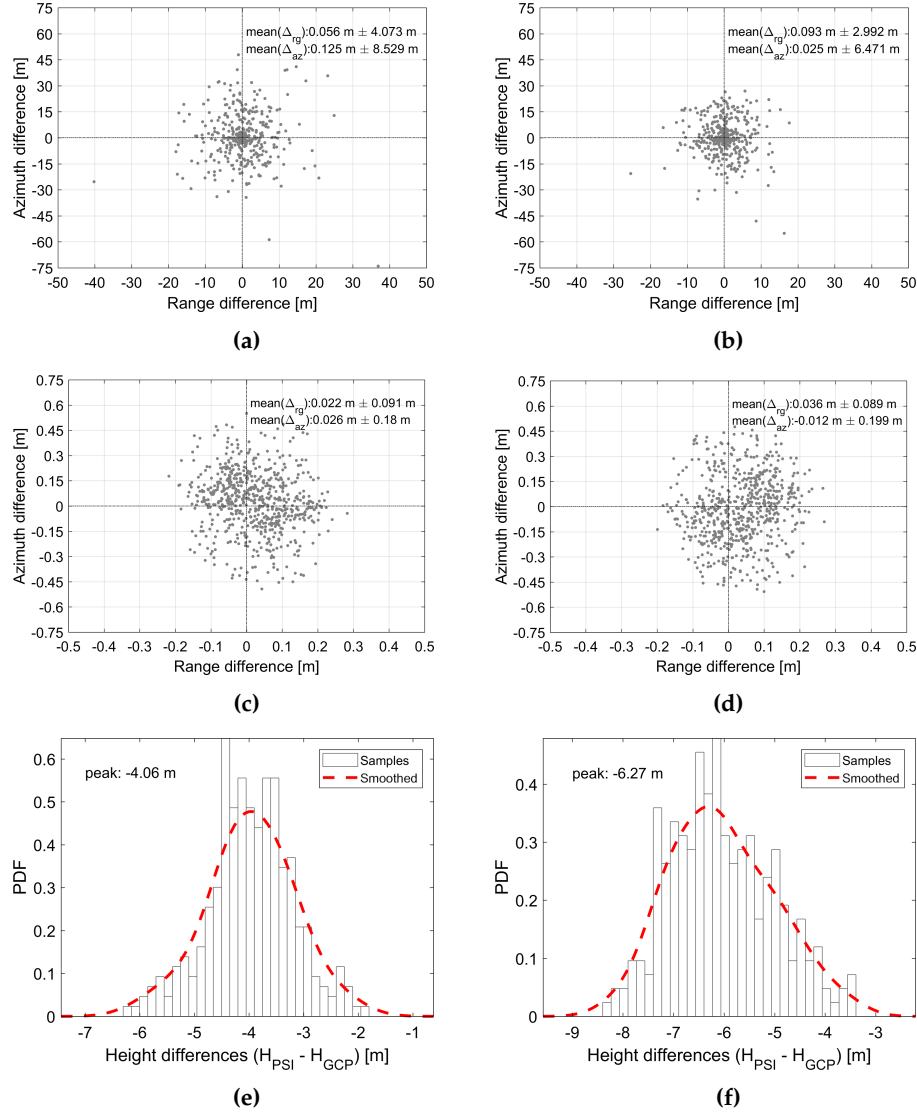
Figure 4.14: Restriction of the DEM error estimation to the true correspondences among GCPs and PS. (**a,b**) demonstrate the scatterplots of GCP-PS pairs before outlier removal, for Beam57 and Beam42, respectively; (**c,d**) show the scatterplots of the GCP-PS pairs after outlier removal. This causes a decrease in the bias and the standard deviation of the coordinate differences; (**e,f**) depict the robust height offset estimation after removal of the height differences in accordance with the 2σ rule. The peak of the smoothed histogram indicates the DEM error for each beam. Note that, for all the figures, the coordinate differences are defined as the GCP coordinates subtracted from their corresponding PS coordinates.

opposing geometry. This is not the case for non-corrected point clouds. Another observation is related to the ascending point clouds. The 2-D shift imposed on the green point cloud after correction is much smaller than the red point cloud. One of the rea-

(a)   **(b)**   **(c)**

(d)   **(e)**   **(f)**
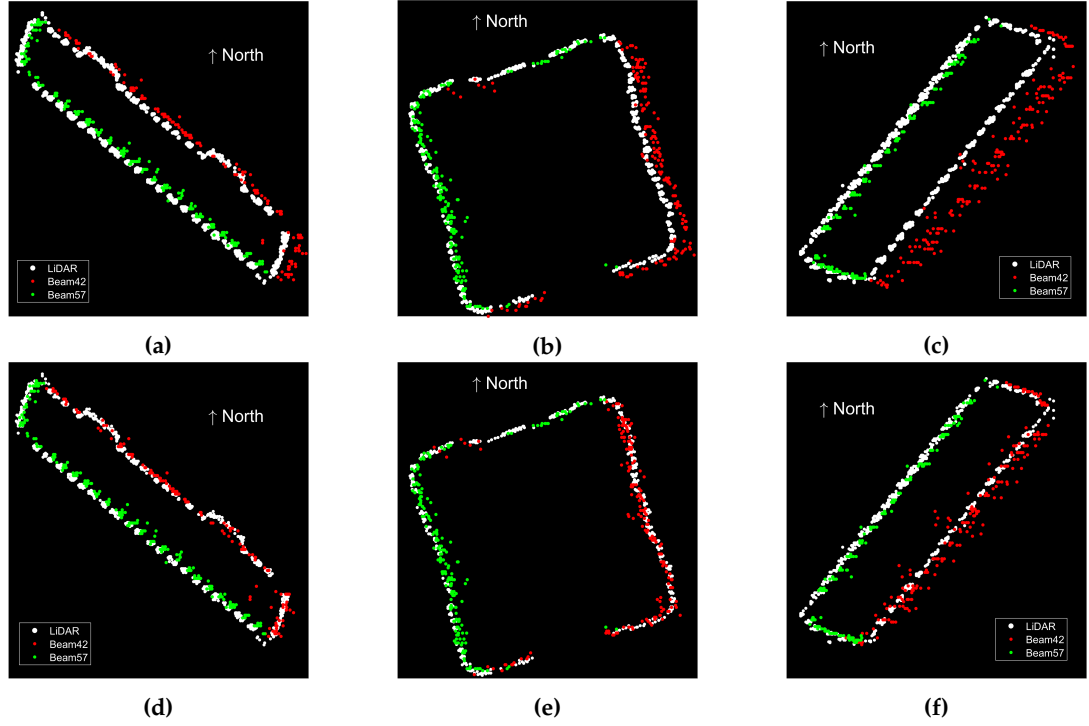
Figure 4.15: 2-D horizontal localization accuracy analysis of PSI point clouds of three selected test sites. (top: non-corrected, bottom: corrected). The x- and y-axes correspond to the Easting and Northing in UTM. Green and red dots show the ascending and descending point clouds while white dots show the extracted LiDAR footprints. The scale of figures is not identical meaning that the degree of zoom-in is higher in subfigures (**c**,**f**).

sons is that the DEM error of the ascending point cloud is approximately two meters smaller than the descending point cloud. Another explanation for this is that the center incidence angle is smaller for the descending point cloud compared to the ascending one (see Table 4.1). Therefore, according to Equations 2.15 and 2.18, the horizontal shifts are larger for the descending point cloud as has been expected.

In terms of a quantitative analysis, the mean value of the distances for the corrected point clouds and the non-corrected point clouds are calculated to be 0.4 m and 2.44 m, respectively. This proves that applying the corrections has certainly improved the overall geocoding accuracy of the PSI point clouds in the horizontal plane as the corrected point clouds are located closer to the LiDAR footprints compared to the non-corrected ones. It is important to note that a deviation equal to zero does not necessarily mean that the PSI point cloud is absolutely without any errors. The reason is that facade PS originate as triple-bounces from window corners of buildings which are not perfectly aligned with footprints extracted from LiDAR. Therefore, the difference in viewing geometries of both sensors does not allow a flawless comparison.

For the 1-D vertical accuracy analysis, we first visually inspect the vertical cross-section of corrected and non-corrected PSI point clouds overlaid onto their corresponding LiDAR point cloud for a test site. Figure 4.16 visualizes this cross-section. In the subfigures, the x- and y-axes correspond to the UTM Easting and ellipsoidal heights, respectively. The red and green colors represent the PSI point clouds of Beam42 and Beam57 while the reference LiDAR point clouds are in white. In the LiDAR data, the ground line and the building roofs are easily distinguishable as they contain the majority of the points and are seen as very bright lines in the figures. By looking at Figure 4.16a,c, a clear height offset between the ground lines in the non-corrected PSI point clouds of both beams and the ground line in the LiDAR data is observed. Moreover, in the non-corrected point clouds, the PS on buildings are not seamlessly aligned with building roofs of the LiDAR point clouds. The non-identical height shifts are compensated by employing the height offset estimation method, relying on SAR GCPs, as it can be seen in Figure 4.16b,d.

In order to quantify the degree of improvement in the absolute height of PSI point clouds with respect to the LiDAR heights, the façade PS are excluded and height histograms are formed. This procedure is carried out for a subset of the PSI point clouds and their corresponding LiDAR data marked with yellow rectangle in Figure 4.2. The ellipsoidal height histograms are reported in Figure 4.17 where the top and bottom subfigures correspond to Beam57 and Beam42, respectively. The red, green and blue colors describe the height histograms of the LiDAR, non-corrected PSI and corrected PSI point clouds. The first peak of the histograms relates to the ground points. It is observed that, after the DEM error of the reference point is estimated with the use of SAR-GCPs, the height histograms of the corrected PSI point clouds move toward the height histograms of the LiDAR data. Within each histogram, the height of the ground peaks are reported. The absolute differences in the height at the peaks of the non-corrected point clouds with respect to the peak of the LiDAR data are approximately 3.96 m and 6.43 m for Beam57 and Beam42, respectively. The differences reduce to 0.21 m and 0.11 m after the proposed height offset compensation method has been applied.

4.3.2 Geodetic Point Cloud Fusion

The idea behind point cloud fusion and its benefits have been discussed in Subsection 3.5.1. The Geodetic point cloud fusion approach is superior to its geometric counterparts [25], [26], [47] as it improves the geo-location accuracy of the scatterers while, similar to the geometric techniques, produces shadow-free point clouds. The fusion can be carried out either using a single stereo SAR-based GCP or with the availability of multiple GCPs as will be addressed in the following.

4.3.2.1 Fusion Using One GCP

In Appendix A.3 an approach, different from the one explained in Subsection 4.3.1, has been proposed for the production of absolute 3-D InSAR point clouds. The technique was the first attempt for merging the capabilities of geodetic SAR and InSAR-based

(a)

(b)

(c)

(d)
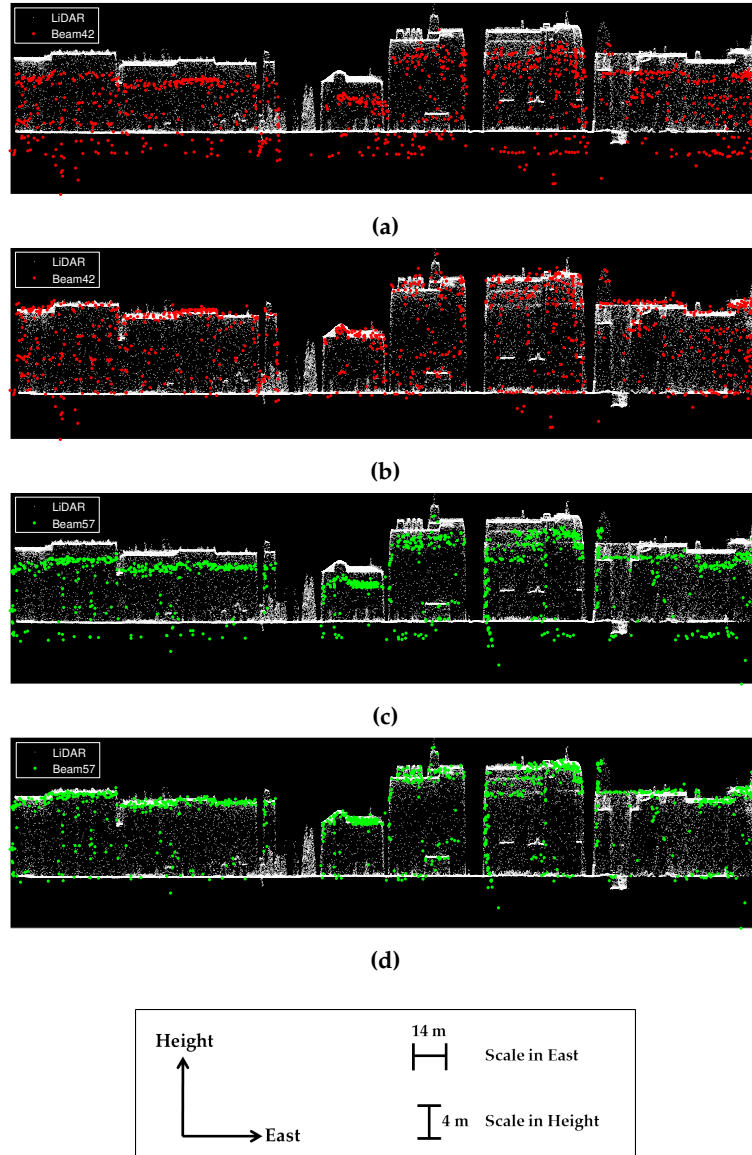
Figure 4.16: 1-D vertical localization accuracy analysis of PSI point clouds, before and after applying geodetic corrections and height offset compensation, in comparison with LiDAR. The x- and y-axes correspond to the Easting in UTM and ellipsoidal height, respectively. Green and red dots show the ascending (Beam57) and descending (Beam42) point clouds while white dots show the LiDAR data. (**a**,**c**) depict the non-corrected point clouds while (**b**,**d**) correspond to the corrected ones. The height shifts, caused by the unknown DEM error of the reference points are easily recognizable in non-corrected point clouds while the offsets are compensated for in the corrected results. Note that the Easting and the height are differently scaled in order to emphasize the vertical effect.

Figure 4.17: Ellipsoidal height histograms of non-façade points in LiDAR (red), non-corrected PSI (green) and corrected PSI point clouds (blue) corresponding to the yellow bounding box in Figure 4.2. (**a**,**b**) show the results for Beam57 and Beam42, respectively. For both beams, the height shifts are compensated for after correction, using SAR-GCPs, as the ground peaks of PSI point clouds and LiDAR data become aligned.

methods. It was tested on two pair of cross-heading TomoSAR point clouds reconstructed from the data introduced in Table 4.1. The individual steps of the framework, called geodetic SAR tomography, along with a practical demonstration in Berlin are reported in here.

1. Absolute 3-D localization of a PS with high SCR by stereo SAR (see Section 3.2), which is visible in all the available multi-aspect SAR image stacks. As was mentioned in Subsection 4.2.1, in urban areas lamp poles and traffic signs are some of the few examples that can satisfy the above-mentioned condition. The selected target for our Berlin experiment is marked in one ascending and one descending mean amplitude image in Figure 4.18.

2. TomoSAR processing of the available stacks following the workflow described in Subsection 3.4.1. The extracted GCP in the previous step is selected as the reference point during the TomoSAR processing of all the stacks. In our study, the tomographic inversion is carried out by the Maximum Detection (MD) algorithm of
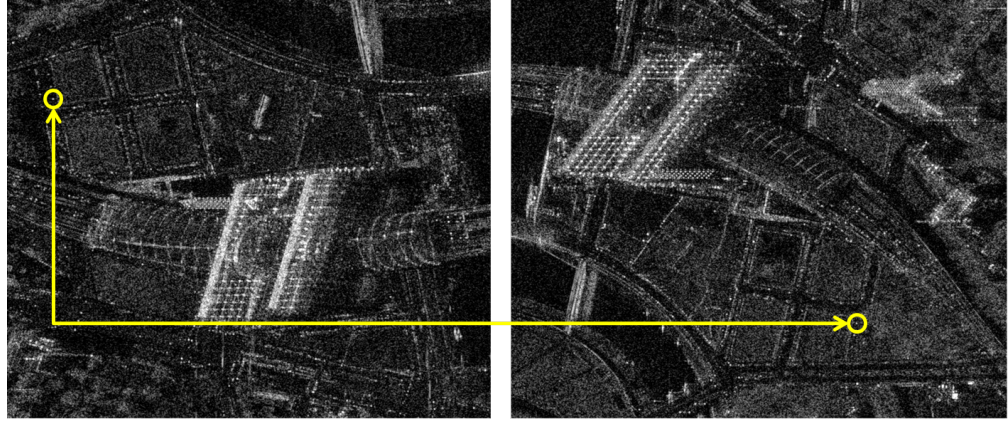
Figure 4.18: The selected identical PS, which is visible through a pair of cross-heading data stacks of Berlin. The scatterer is chosen as the reference point for the follow-on TomoSAR processing.

the Tomo-GENESIS [12], [125], which includes SVD-Wiener for the reconstruction of the reflectivity profile followed by peak detection, model order selection and final refinement of the amplitude and the phase of each scatterer to approximate the elevation [12]. The motion parameters include the linear deformation rate and the amplitude of seasonal deformation and are estimated using the Time-Warp method [102]. The geocoded topographic map of a descending stack of Berlin obtained from TomoSAR is visualized in Figure 4.19. By comparing the TomoSAR point cloud of Figure 4.19 to the PSI result of Figure 4.10, it is clearly seen that the TomoSAR approach is capable of 3-D mapping in much higher detail than PSI due to the layover separation capability.

3. Performing geodetic point cloud fusion for the generation of an absolute shadow-free dense 3-D point cloud. If the selected reference point has been localized using same-heading tracks, the geodetic point cloud fusion is carried out by shifting all of the geocoded point clouds toward the coordinates of the absolute reference point. If the selected reference point has been localized using cross-heading tracks, the bias, occurred due to the non-identical phase-centers of the reference point visible in multi-aspect SAR images, should be first resolved. Since the most probable candidate for reference points visible in cross-heading geometries are poles, the mentioned bias depends on the diameter of the pole, which can be considered to be in the order of 20 cm. The problem is sketched in Figure 4.20. The red dot P_{SS}, marking the location of the GCP obtained by stereo SAR, should be shifted to the green dots $P_{G_{Asc}}$ and $P_{G_{Dsc}}$, which denote the true geocoded location of the reference point in each stack. This requires the evaluation of the shift vectors ds_{Asc} and ds_{Dsc} for the ascending and the descending point clouds, respectively. Assuming the local incidence angles (θ_{Asc}, θ_{Dsc}) and heading angles
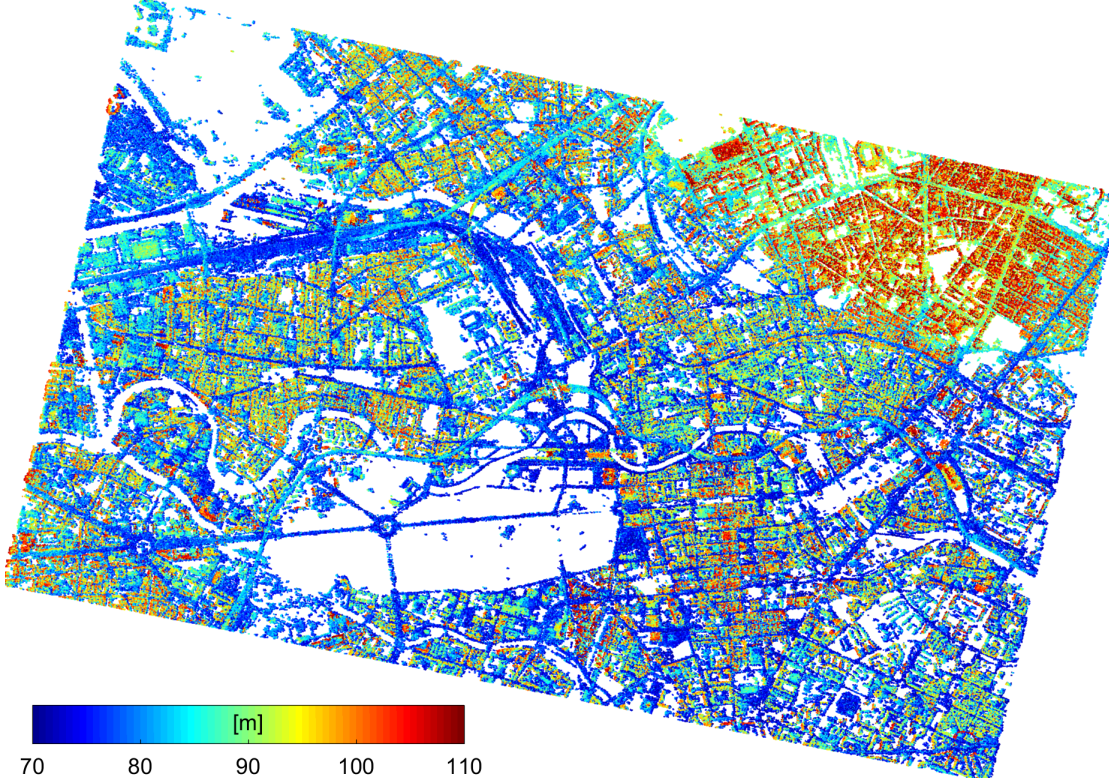
Figure 4.19: Geocoded descending TomoSAR point cloud of Berlin in UTM coordinates. The ellipsoidal height is colorcoded. The *x*- and *y*-axis correspond to the UTM East and North, respectively.

$(\alpha_{Asc}, \alpha_{Dsc})$ of each beam are available, and the diameter of the pole D is known, each component of the shift vectors is calculated as follows:

$$dz = \frac{D \tan \theta_{Asc} \tan \theta_{Dsc}}{\tan \theta_{Asc} + \tan \theta_{Dsc}} \tag{4.8}$$

$$dx_{Asc} = dxy_{Asc} \cos \alpha_{Asc} \tag{4.9}$$

$$dy_{Asc} = -dxy_{Asc} \sin \alpha_{Asc} \tag{4.10}$$

$$dx_{Dsc} = dxy_{Dsc} \cos \alpha_{Dsc} \tag{4.11}$$

$$dy_{Dsc} = -dxy_{Dsc} \sin \alpha_{Dsc}, \tag{4.12}$$

where $dxy_{Asc} = dz \cot \theta_{Asc}$ and $dxy_{Dsc} = dz \cot \theta_{Dsc}$. The variable dz is equal for all the stacks and is estimated using least squares adjustment based on different combinations of the ascending and descending tracks. Afterward dxy_{Asc} and dxy_{Dsc} are calculated and further projected onto the East and North components, which finally results in the three components of $\boldsymbol{ds}_{Asc}$ and $\boldsymbol{ds}_{Dsc}$:
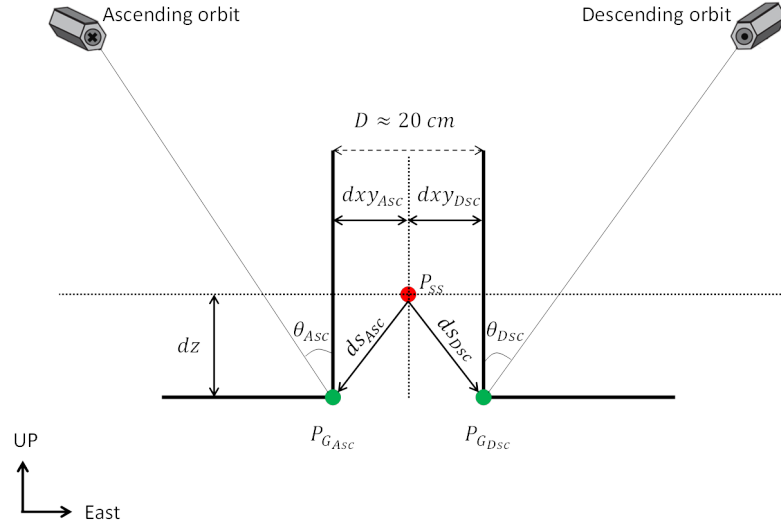
Figure 4.20: Illustration of the bias caused due to the non-identical phase centers of assumed common targets visible in cross-heading point clouds. The drawing is in the east-up plane where ascending and descending satellites fly perpendicular to the plane far away and toward the reader, respectively. For successful point cloud fusion, the coordinates resulted from stereo SAR P_{SS} should be shifted toward the expected true position of the reference point in each ascending and descending stack ($P_{G_{Asc}}$, $P_{G_{Dsc}}$). In order to do so, the corresponding shift vector in the elevation direction (ds_{Asc}, ds_{Dsc}) should be evaluated for each stack based on the known (θ_{Asc}, θ_{Dsc}) and (α_{Asc}, α_{Dsc}).

$$\boldsymbol{ds}_{Asc} = \left(dx_{Asc}, dy_{Asc}, dz_{Asc} \right)^T \tag{4.13}$$

$$\boldsymbol{ds}_{Dsc} = \left(dx_{Dsc}, dy_{Dsc}, dz_{Dsc} \right)^T. \tag{4.14}$$

Subsequently, P_{SS} is shifted to the position of $P_{G_{Asc}}$ and $P_{G_{Dsc}}$. The remaining errors of the geocoding are compensated by evaluating the difference between the geocoded reference point coordinates of each stack and the corresponding $P_{G_{Asc}}$ and $P_{G_{Dsc}}$. Finally, for each stack, the unique difference vector is added to the coordinates of all the scatterers to produce multiple absolutely localized TomoSAR point clouds.

Figure 4.21 shows the four absolutely localized TomoSAR point clouds, in different shades of gray, as well as the location of the reference point marked by the red asterisk. The total number of scatterers is approximately 63 million. A 3-D view of the point cloud obtained after the geodetic fusion is visualized in Figure 4.22, which shows the great level of details captured from the TomoSAR processing and the subsequent fusion

Figure 4.21: Geodetic fusion of four TomoSAR point clouds (in gray) with respect to the identical reference point marked by the red asterisk.

algorithm. The horizontal accuracy of the generated point cloud is approximately 20 cm, when compared to the reference LiDAR DSM (see Appendix A.3).

4.3.2.2 Fusion Using Multiple GCPs

Geodetic point cloud fusion can be also carried out as a post processing step when a network of GCPs, obtained from stereo SAR, is available. From the methodological point of view, the approach follows the same procedure of geocoding error correction explained in Subsection 4.3.1. In the following an example of such method fusing four PSI point clouds of Oulu, generated from the data of Table 4.1, is given.

The used GCPs include a subset of the same-heading candidates of Oulu (see Figure 4.8 and Table 4.3), which are filtered based on their posterior azimuth standard deviations. The total number of the utilized GCPs is 519 with their average statistics reported in Table 4.4. The GCPs correct for the height error of the reference point of each PSI stack while geodetic corrections are applied on the master scenes for mitigating the horizontal coordinate errors (see Subsection 4.3.1). A 2-D top view of the absolutely
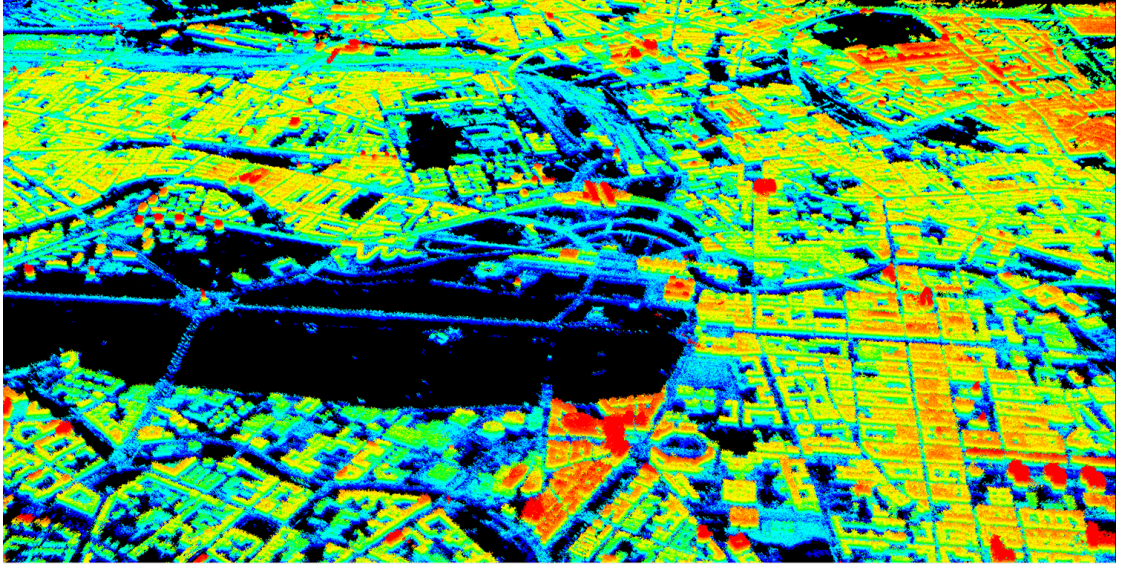
Figure 4.22: Absolute 3-D fused TomoSAR point cloud of the central area in Berlin. The ellipsoidal height values are colorcoded and range between 70 m to 110 m. The fusion of multitrack point clouds allow for a very detailed representation of the city where most structures can be easily recognized.

localized fused point cloud is shown in Figure 4.23. It includes approximately 7 million point scatterers covering an area of 50 km^2.

Table 4.4: Average statistics of Oulu GCPs with posterior azimuth standard deviation lower than 10 cm. These GCPs are used for geodetic fusion of four PSI stacks of Oulu.

	σ_N [cm]	σ_E [cm]	σ_H [cm]
Ascending GCPs	2.6 ± 0.83	8.03 ± 3.46	6.99 ± 3.17
Descending GCPs	1.99 ± 0.61	7.57 ± 2.24	8.41 ± 2.45

A small portion of the fused point cloud is overlaid onto the LiDAR data of Oulu in Figure 4.24 to show the approximate fit of the point clouds to the LiDAR data. Since no information about the absolute accuracy of the LiDAR data was available at the time of the study, no quantitative analysis for evaluating the height accuracy of the fused point cloud was performed. Therefore, the LiDAR DSM is used only for visualization purposes.

As the final showcase, for a test site, the 2-D alignment of one ascending and one descending point cloud of Oulu is checked in Figure 4.25 and in Figure 4.26. The ascending and descending point clouds are visualized as green and red dots, respectively. It is clear that in Figure 4.25a the end point of buildings captured from opposite sides do not match. This horizontal offset, which on average is approximately 4.7 m in East and almost 1.8 m in North, is largely mitigated after performing geodetic point cloud fusion in Figure 4.25b. This indicates the relative correctness of the fusion algorithm.
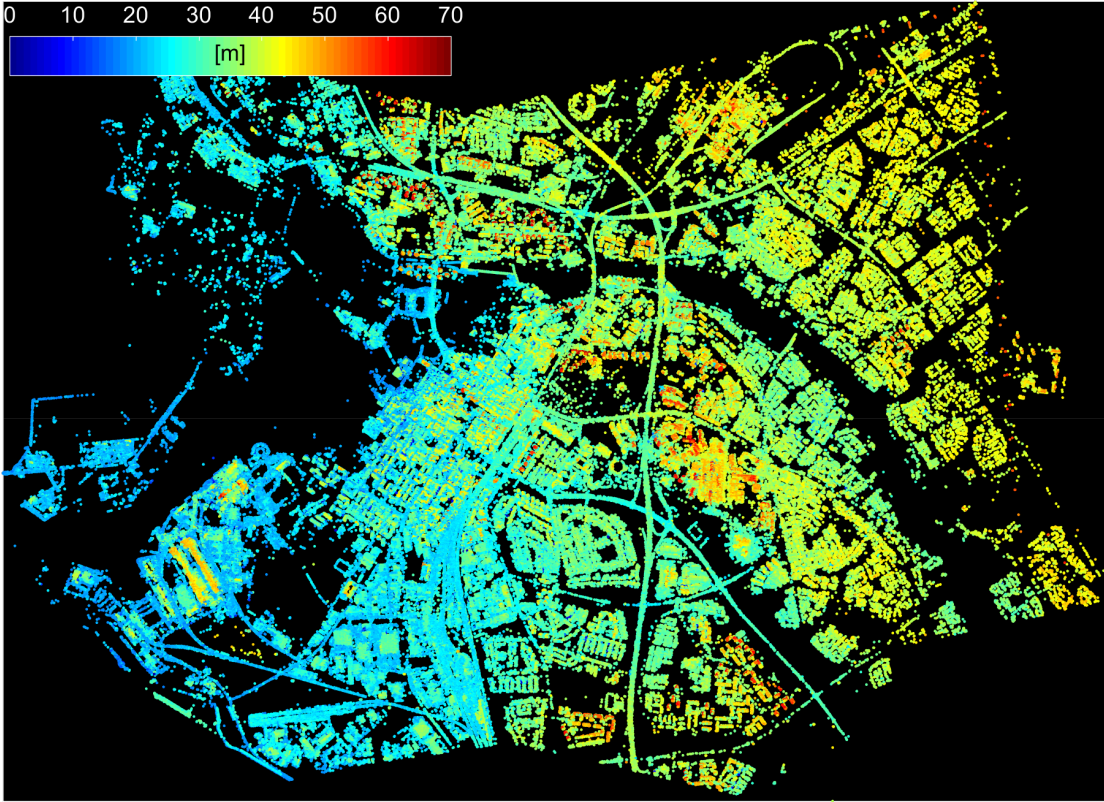
Figure 4.23: Geodetic fusion of four PSI point clouds of Oulu. The ellipsoidal height of the point clouds is colorcoded.

Figure 4.26a and Figure 4.26b show an overlay of the non-corrected and the corrected PSI point clouds of a test site in Oulu, respectively. It is observed that after the geodetic point cloud fusion, the point clouds are aligned with the corresponding side of the buildings seen in the aerial ortho-photos.

4.4 1-D to 3-D Motion Decomposition

The geodetic fusion of multi-aspect InSAR point clouds of urban areas allows to obtain a detailed shadow-free 3-D representation of the investigated scene, as was demonstrated in Subsection 4.3.2. Apart from this cartographic application, the fused point cloud facilitates the decomposition of 1-D InSAR deformation estimates into their original 3-D displacement vectors. Inspired from [11], [97], a 3-D motion decomposition algorithm has been proposed in [138] and applied to four TomoSAR point clouds of Berlin, which can be found in Appendix A.4. The method consists of the following steps:

61
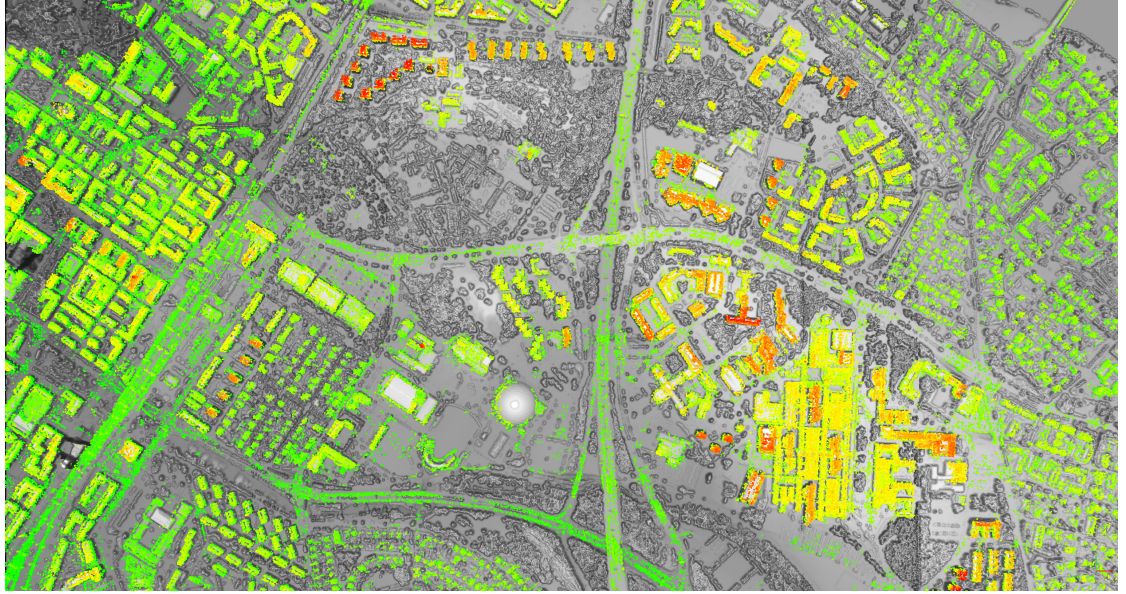
Figure 4.24: A 2-D view on the fusion of four PSI point cloud of Oulu. The LiDAR data is plotted in gray and serves as a reference for visualization purposes only. The height of the point clouds are colorcoded with green showing low heights while read shows high height values.

| (a) | (b) |
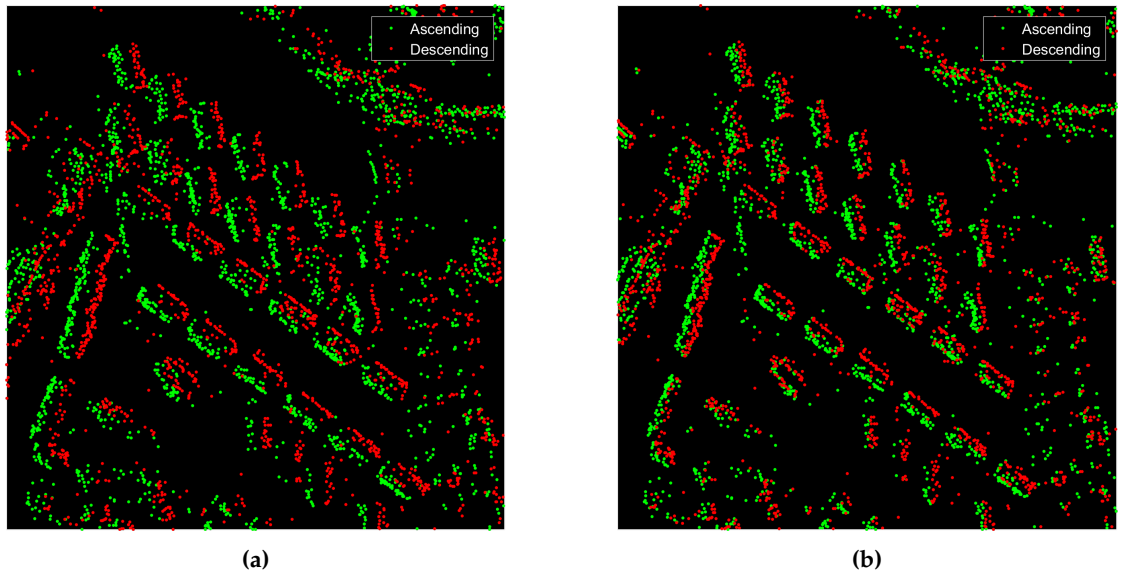
Figure 4.25: A 2-D view on the ascending and the descending PSI point clouds of Oulu (**a**) before and (**b**) after the geodetic point cloud fusion.

1. Geodetic fusion of multiaspect point clouds (see Subsections 4.3.2). The geodetic approach results in a more accurate and objective point cloud coregistration than

(a) **(b)**
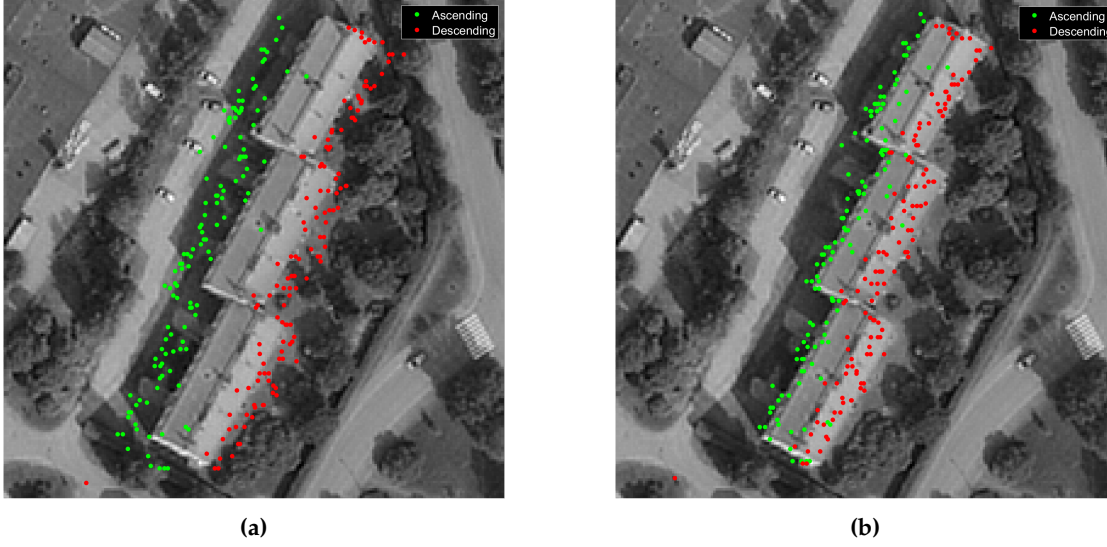
Figure 4.26: Ascending and descending point clouds of Oulu overlaid on the corresponding aerial ortho-photo of Oulu (**a**) before and (**b**) after the geodetic point cloud fusion. The results reported in (**b**) allow the assignment of the PS to building parts.

the geometric techniques reported in [25] and [26]. The fused absolute TomoSAR point cloud of Berlin can be observed in Figure 4.22, which is the basis for the motion decomposition algorithm.

2. Estimation of the 3-D motion components of each scatterer in local spatial cubes by minimizing the weighted sum of the absolute residuals of the functional model described in Equation 2.21 (see Subsection 2.3.2). The utilization of l_1-norm minimization instead of the more popular l_2-norm minimization is related to the existence of large number of outliers in the elevation and deformation estimates of TomoSAR point clouds [12], [107], [139].

Inside each cube, the unknowns include the three motion components of the central point, the observations consist of the linear or seasonal deformation parameters of the scatterers surrounding the central point and the design matrix is evaluated based on the local heading angle and the local incidence angle attributed to each scatterer. The weight vector is proportional to the inverse of the squared distance of each scatterer to the central point i.e., the points closer to the central point have higher weights in the estimation.

It is important to note that the size of the cube is chosen based on the spatial resolution of the SAR data and the inherent scale of the investigated objects. After an empirical study, for TSX high resolution spotlight data over European cities, we recommend to choose a cube dimension of 5 m in each coordinate direction.

In the following, the proposed motion decomposition algorithm is validated using simulated and real data. Furthermore, the effect of viewing geometries of the available data sets (see Table 4.1) on the retrieval of 3-D motion components is discussed.

4.4.1 Effect of Viewing Geometry

The quality of motion decomposition depends highly on the difference in the viewing geometry of the acquisitions [33]. If a point target is hypothetically visible in the four TomoSAR point clouds of Berlin summarized in Table 4.1, and we assume that LOS deformation measurements are uncorrelated and have equal standard deviations symbolized by σ, then the posterior covariance matrix of the target's displacement estimates $\mathbf{Q}_{\hat{X}}$ is equal to its dilution of precision (DOP) matrix [44], [124] if we set $\sigma = 1$:

$$\mathbf{Q}_{\hat{X}} = \begin{bmatrix} \sigma_u^2 & \sigma_{ue} & \sigma_{un} \\ \sigma_{eu} & \sigma_e^2 & \sigma_{en} \\ \sigma_{nu} & \sigma_{ne} & \sigma_n^2 \end{bmatrix} = (\mathbf{A}^T\mathbf{A})^{-1} = \begin{bmatrix} 43.3 & -0.8 & 277.8 \\ -0.8 & 0.51 & -5.4 \\ 277.8 & -5.4 & 1801.7 \end{bmatrix} \qquad (4.15)$$

where subscripts u, e and n stand for up, east and north, respectively and $\mathbf{A}$ denotes the design matrix that is evaluated based on the functional model described in Equation 2.21. A low diagonal value in the DOP matrix indicates the goodness of the geometry configuration for retrieving the corresponding component. The diagonal elements of Equation 4.15 show the strength of the current situation for retrieving the east-west motion and its weakness for retrieving the north-south displacement component. This has been expected due to the inherent near-polar orbit of TSX satellites. The off-diagonal elements of the DOP matrix show the covariance between the relative errors in motion components. The values show that although the sensitivity of the current geometry configuration is low with respect to the north-south component, it should not be omitted from the functional model of Equation 2.21. Not including the term related to d_n leads to the bias Δd_e in the motion decomposition:

$$\Delta d_e = d_n \tan\alpha, \qquad (4.16)$$

which can reach up to 18% of the deformation in the north-south direction with typical TSX azimuth values. Apart from this issue, the extremely high covariance value of 277.8 for north and up components shows that precise unambiguous retrieval of the north-south component is not possible with the available viewing geometries and also with currently operational SAR satellites [123].

4.4.2 Motion Decomposition on Simulated data

Motion decomposition is performed on simulated data solely for the purpose of comparing the results obtained from the l_1-norm and from the l_2-norm minimization. We assume a true deformation vector of $\mathbf{d}^t = \begin{bmatrix} d_u^t & d_e^t & d_n^t \end{bmatrix}^T = \begin{bmatrix} -10 & 1 & 2 \end{bmatrix}^T$ mm/year for the target. We further assume that the target is visible in data from the four viewing geometries of Berlin TomoSAR scenario. This results in four d_{LOS} calculated from Equa-
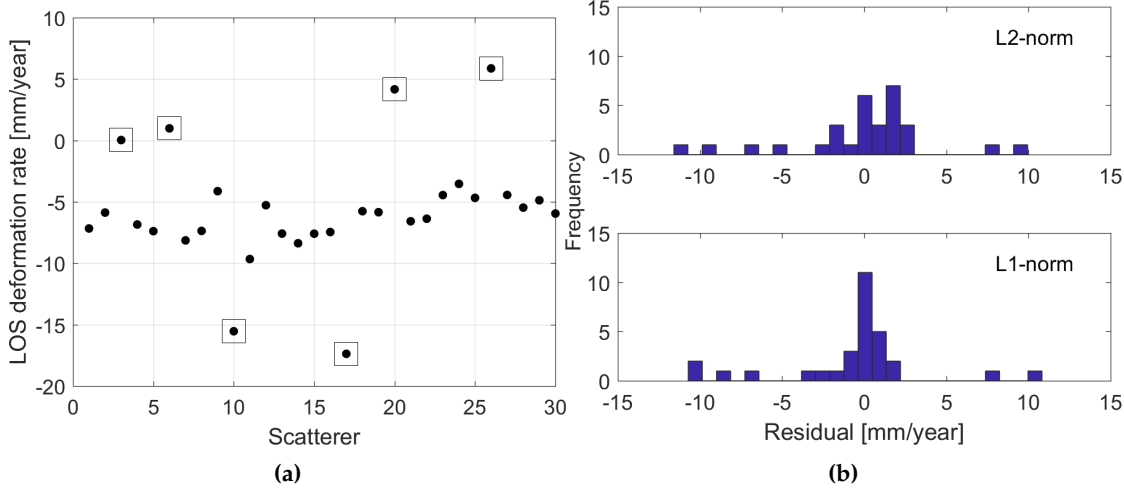
Figure 4.27: The simulated LOS deformation values for 30 points assumed to be inside an arbitrary spatial cube shown in **a**. In **b**, the histograms of residuals evaluated after the motion decomposition performed with l_2- and l_1-norm minimization are reported.

tion 2.21, which are used to realize LOS deformations of 30 point targets surrounding the central point in the cube in four groups differing in the associated incidence and heading angles. The 30 LOS observations are randomly realized following a Gaussian PDF with mean equal to the LOS deformation calculated using the mentioned forward model and the variance σ_v^2 apprxoimated by the Cramer-Rao bound of LOS linear deformation rate obtained from interferometric measurements [6]:

$$\sigma_v^2 = \left(\frac{\lambda}{4\pi}\right)^2 \frac{1}{2\,N\,SNR\,\sigma_t^2},\tag{4.17}$$

where N is the number of acquisition in a stack of SAR images, the variable SNR denotes the signal-to-noise ratio and σ_t^2 is the variance of the temporal baseline distribution. Considering a TSX data stack with average acquisition parameters taken from Table 4.1 with $\lambda = 31$ mm, and assuming that the $SNR = 10$ dB, $N = 115$ and $\sigma_t = 1.6$ years, then $\sigma \approx 0.03$ mm/year. With taking into account the atmospheric disturbances we assume a value of $\sigma_v = 1$ mm/year. At the final stage, we add 20% outliers, with magnitudes of 10 mm/year, to the 30 LOS observations. The scatter plot of the observations is visualized in Figure 4.27a. The 3-D motion components of the desired point has been retrieved by l_2-norm and l_1-norm minimization using Equation 2.26 and Equation 2.27, respectively. The histogram of residuals and the results of the estimations are reported in Figure 4.27b and Table 4.5. It is seen that the outliers have less effect on the results of the l_1-norm minimization rather than the least squares method. This fact is confirmed based on the more logical distribution of errors in the residual histogram of l_1-norm minimization as well as the higher degree of closeness of all the retrieved motion components with respect to the true ones compared to the results of l_2-norm minimization.

65

Table 4.5: Comparison of the results from the l_1- and the l_2-norm minimization applied to motion decompistion using the simulated case study. The vectors $\mathbf{d}^{l_2}$ and $\mathbf{d}^{l_1}$ denote the retrieved motion vectors from least squares and the l_1-norm minimization, respectively.

Components	$\mathbf{d}^t$ [mm/year]	$\mathbf{d}^{l_2}$ [mm/year]	$\mathbf{d}^{l_1}$ [mm/year]	$\mathbf{d}^{l_2} - \mathbf{d}^t$	$\mathbf{d}^{l_1} - \mathbf{d}^t$
d_u	-10	-5.57	-10.06	4.43	**-0.06**
d_e	1	2.21	0.88	1.21	**-0.12**
d_n	2	20.56	-4.08	18.56	**-6.08**

4.4.3 Motion Decomposition on Real data

The motion decomposition strategy has been applied to the data set of the Berlin TomoSAR scenario (see Table 4.1), which were used to generate an absolute fused point cloud visualized in Figure 4.22. The LOS linear and seasonal deformation maps of the entire city of Berlin have been previously produced and analyzed in [11], [97], [140]. Here the motion decomposition is only applied to the two test cases, marked in Figure 4.28, which previously showed remarkable deformation patterns.

Figure 4.29 shows the TomoSAR LOS seasonal deformation maps of the central of Berlin station available from each orbit (Figure 4.29a–d) as well as the decomposed motion maps in the vertical (Figure 4.29g) and in the east-west directions (Figure 4.29h). The vertical and east-west motion components evaluated by l_2-norm minimization are also shown in Figure 4.29e and Figure 4.29f. From the LOS maps, seasonal deformation with magnitudes up to 12 mm is visible. This is mainly due to the thermal-dilation-induced deformation of the railway station, which is made of steel. Figure 4.29g shows that the main parts of the station and also the bridge in the east of the station undergo vertical seasonal deformation with magnitudes up to 6 mm. The east hallway also shows magnitudes of 5 mm of vertical seasonal deformation (yellow part). The most interesting pattern is observed in Figure 4.29h. The right hallway undergoes heavy seasonal deformation in the east-west direction, with magnitudes equal to 12 mm (24 mm between summer and winter). By comparing the estimated components from l_1- and l_2-norm minimization another important observation is that in the results of l_2-norm minimization, most of the scatterers are filtered out because of low precision values. This is most apparent in Figure 4.29e, where the deforming hallway in the east direction of the station is not visible. On the contrary, l_1-norm minimization preserves the deformation pattern and hence demonstrates its robustness against outliers. Furthermore, to give a rough estimate on the noise level of the estimates obtained from the two mentioned estimators, the empirical standard deviation of the deformation values in the vertical direction was evaluated for both cases (see Figure 4.29e and Figure 4.29g). The calculated values are equal to 2.64 and 3.25 mm for l_1 and l_2, respectively.

Figure 4.30 shows the TomoSAR LOS linear deformation maps of the central station available from each orbit (Figure 4.30a–d) as well as the decomposed motion map in the vertical direction (Figure 4.30e). In general, the building and the rail tracks are not
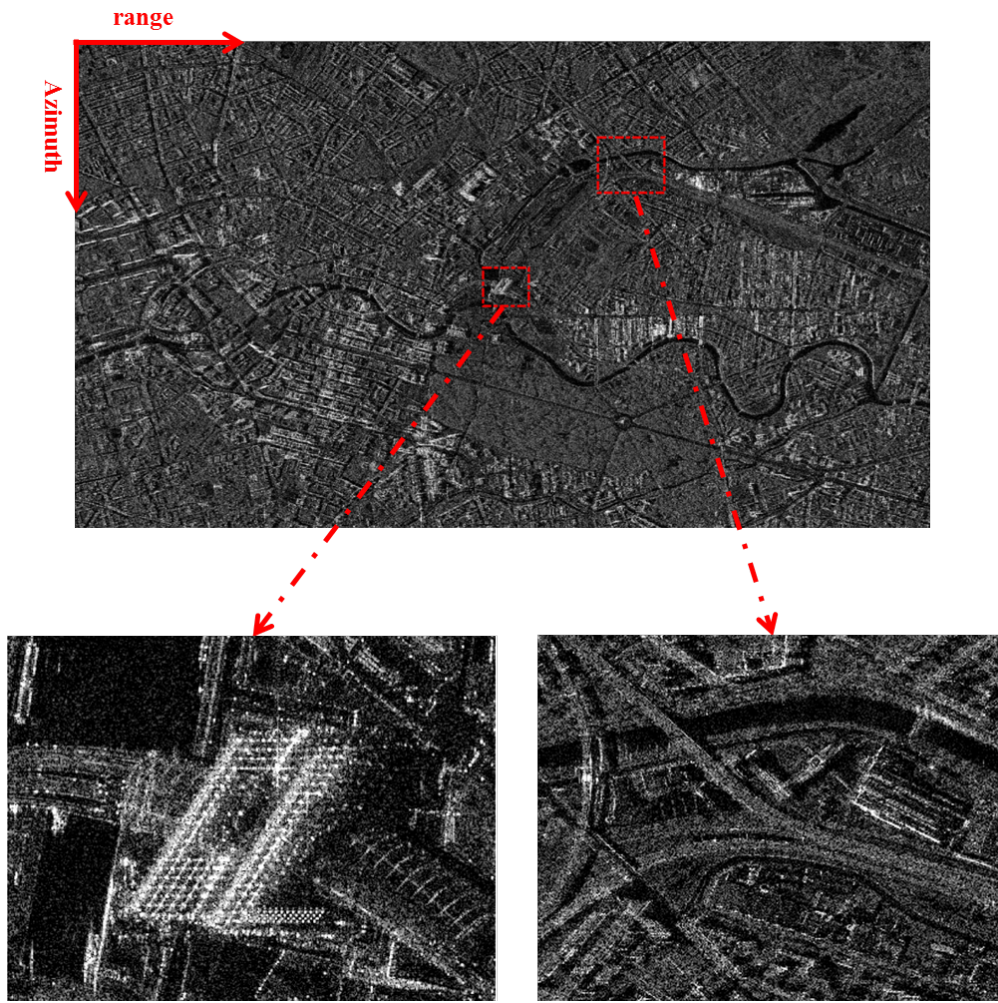
Figure 4.28: The mean calibrated amplitude image of Berlin from a descending orbit track and the test sites used in motion decomposition.

influenced by significant linear deformation as it is observed from the LOS maps. There are some parts at the left side of the station in red color which show a construction site and therefore cannot be interpreted as deforming areas. Furthermore, the main building parts of the station show LOS subsidence not higher than 4 mm/year. The decomposed linear vertical map (see Figure 4.30e) shows that the internal sections of the two main parts undergo vertical subsidence on the order of 2 to 4 mm/year, while the rail tracks and other parts of the station are stable.

For the second test site, the railway bridge, motion decomposition is only performed on the seasonal deformation map since there is no significant linear ground deformation visible in the area. The result is shown in Figure 4.31. From the seasonal LOS maps, it can be observed that the two sections on the railway bridge undergo seasonal deformation with magnitudes up to 7 mm. Another prominent pattern can be seen on the
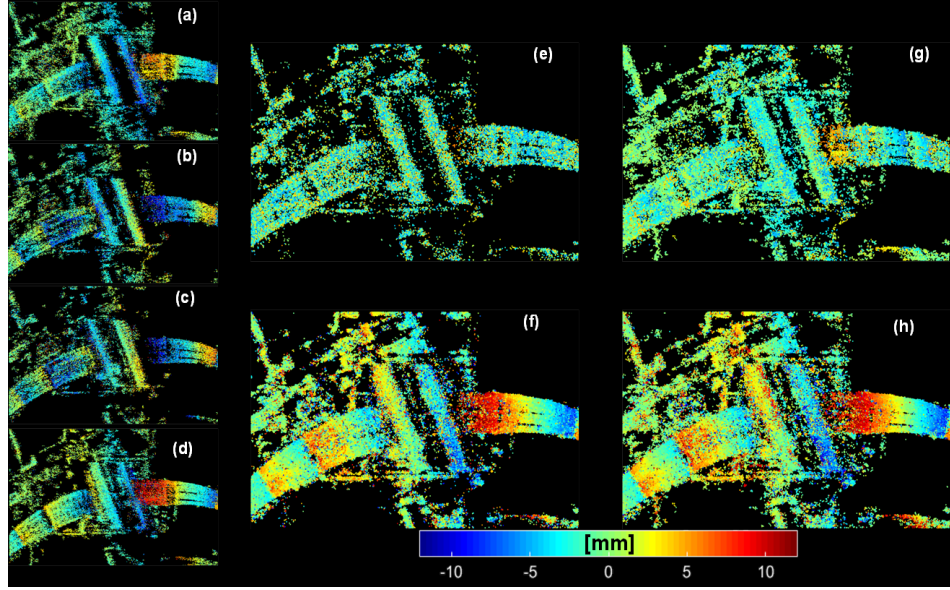
Figure 4.29: Motion decomposition of Berlin central railway station. LOS amplitudes of seasonal motion of (**a**) Beam42, (**b**) Beam57, (**c**) Beam85, and (**d**) Beam99. Decomposed seasonal deformation in (**e**) the vertical direction and in (**f**) the east-west direction by l_2-norm minimization. Decomposed seasonal deformation in (**g**) the vertical direction and in (**h**) the east-west direction by l_1-norm minimization. Deformation maps are in millimeters.
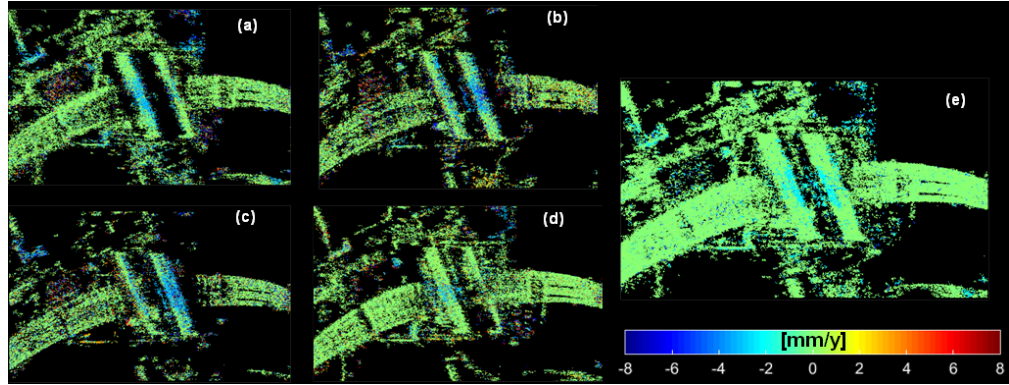
Figure 4.30: Motion decomposition of Berlin central railway station. LOS linear motion of (**a**) Beam42, (**b**) Beam57, (**c**) Beam85, and (**d**) Beam99. (**e**) shows the decomposed linear deformation rate in the vertical direction. Deformation maps are in mm/year.

building at the top left of each subfigure. From Figure 4.31e, it is seen that not much of the LOS seasonal deformation can be attributed to the motion in the vertical direction as, for most of the area in the scene, magnitudes not higher than 4 mm are visible. On the other hand, the rail tracks and the building located at the top left are highly influ-
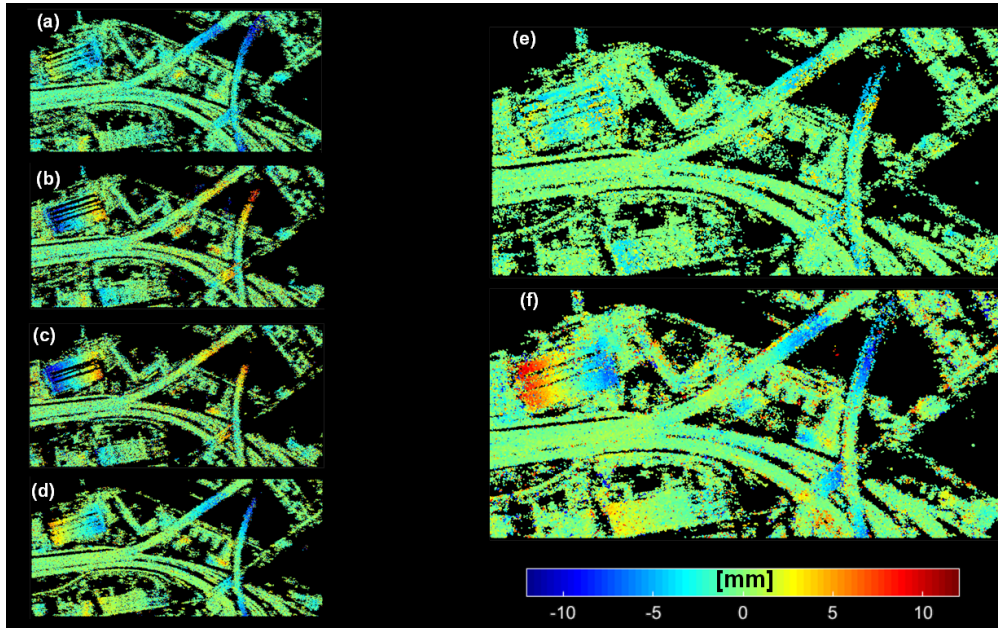
Figure 4.31: Motion decomposition of the railway bridge in Berlin. LOS amplitudes of seasonal motion of (**a**) Beam42, (**b**) Beam57, (**c**) Beam85, and (**d**) Beam99. Decomposed seasonal deformation in the vertical direction and in the east-west direction are shown in (**e**) and (**f**). Deformation maps are in millimeters.

enced by the seasonal deformation in the east-west direction with amplitudes as high as 12 mm.

5 Conclusions and Outlook

5.1 Summary and Conclusion

This dissertation discussed possibilities and methodologies for merging the capabilities of relative InSAR-based techniques and absolute SAR geodetic approaches. The research objectives delineated in Chapter 1 have been addressed with the focus on *the generation of highly detailed absolute InSAR point clouds* for the task of urban area monitoring. In this regard, the following conclusions can be drawn:

- Automatic extraction and positioning of GCPs is possible using multi-aspect high resolution SAR images of urban areas. The tedious task of detection of identical natural scatterers from SAR images of different viewing geometries can be aided by using auxiliary geospatial data or based on the geometric fusion of multiple PSI point clouds. It was concluded that the spatial resolution of the optical data used for detection of candidates should be high enough so that the proposed template matching scheme can reliably detect the desired targets. For the same task, it was seen that the localization accuracy of the input road network data does not necessarily need to be high.

- A quite flexible framework has been proposed for the generation of GCPs, where the user is able to limit or extend the number of GCP candidates either at the beginning of the procedure by selecting different distance thresholds for initial detection, or by trimming the data based on analyzing the phase noise time series, or at the final step of the processing based on the posterior quality indicators of stereo SAR.

- The precision of the coordinates of the stereo SAR GCPs depends highly on the angular difference of the used SAR images. In our experiments, candidate GCPs from same-heading tracks, which include mostly facade PS, had been localized with precision values better than 20 cm in East and Height and around 5 cm in North. Using data from cross-heading tracks, the candidates, which usually include lamp poles or traffic lights in cities, could be positioned with almost ten times lower standard deviations in East and Height and twice lower values in North. Therefore, it is recommended to use cross-heading tracks, upon availability, for GCP retrieval.

- The first demonstration of geocoding error correction and absolute InSAR point cloud production without the use of external positioning data has been provided. The coordinate shifts due to the unknown height of the InSAR reference points

can be largely compensated by the proposed calibration procedure using SAR-based GCPs. The correction method can be readily applied to previously generated point clouds as a post-processing step and does not require any modification in PSI or TomoSAR algorithmic steps. The InSAR products with enhanced localization accuracy allow the association of scatterers to their true positions, which is in particular important when dealing with high resolution SAR products. This leads to the correct detection of source of deformation in InSAR displacement maps.

- Geodetic multi-aspect fusion of InSAR point clouds paves the way for the generation of shadow-free products with high localization accuracy. The proposed fusion approach can be either carried out with a single SAR-based GCP or by using a GCP network. Apart from resolving the shadow phenomena using cross-heading data, the fusion of same-heading tracks leads to the substantial increase in the scatterer density of InSAR point clouds. Examples using TomoSAR and PSI point cloud fusion of entire urban areas were reported in this work with scatterer densities of 1,260,000 and 140,000 points per squared kilometers, respectively.

- The accuracy of the generated GCPs and the produced absolute PSI and TomoSAR point clouds were analyzed with respect to a reference LiDAR point cloud. For several GCPs an average height bias of 13 cm was reported. The horizontal accuracy of the point clouds were evaluated based on extraction of facade PS and comparing them with 2-D footprints from LiDAR, which resulted in accuracy of 20 cm and 40 cm for TomoSAR and PSI, respectively. For vertical accuracy analysis, the ground points of both InSAR and LiDAR point clouds were compared, which showed biases close to 10 cm. This level of accuracy is in line with the previously reported accuracy of the GCPs.

- The first demonstration of motion decomposition using multi-aspect TomoSAR point clouds has been given. The 3-D deformation components of scatterers were reconstructed based on the input of fused point clouds with absolute coordinates. By evaluating the effect of available viewing geometries, it was concluded that the reliable retrieval of the deformation in the North-South direction is not possible. This is the case in general with current SAR satellites operating in near-polar orbits. However, this component should not be omitted from the functional model in order to prevent biased deformation estimates in the remaining motion components.

- It has been shown that the utilization of l_1-norm minimization instead of least squares dramatically reduces the effect of outliers on the final displacement estimates using the proposed motion decomposition algorithm. Therefore, more accurate deformation mapping in higher level of details is possible by exploiting robust estimators. Motion decomposition allows for easier interpretation of deformation patterns rather using the 1-D LOS displacement observations of InSAR. However, it should be noted that l_1-norm regression requires certain number of

iterations, which makes the minimization procedure more computationally expensive than least squares.

5.2 Outlook

The findings of this thesis are initial steps toward reducing the gap between phase-based InSAR and geodetic SAR techniques. Several improvements can be made in different aspects of this work, which are briefly introduced in the following.

5.2.1 Alternative Remotely-Sensed GCP Generation

The detection of candidates based on high resolution optical data can be improved by employing more advanced object detection techniques instead of template matching. For instance, the usage of Histogram of Oriented Gradients (HoG) descriptors for this purpose have showed encouraging results [141]. For discarding unreliable candidates, in cases where a few point targets are located in close proximity of each other, analyzing PTA-based metrics such as Integrated Sidelobe Ratio (ISLR) can be helpful.

Apart from SAR-based generation of GCPs, methods based on SAR-Optical stereogrammetry can be used [142], [143]. However, it has to be mentioned that these approaches work only on data takes with small angular differences and therefore the accuracy of the extracted GCPs are lower than the proposed exclusively SAR-based method. The ongoing studies of matching optical and SAR imagery [144]–[146] can help in identifying more suitable GCP candidates and will deepen our understanding of the nature of these targets.

5.2.2 Employing GNSS Measurements

GNSS measurements can be used to validate the 3-D positioning accuracy of GCPs localized from stereo SAR. This can be performed by carrying out GNSS measurements on some selected lamp poles as has been demonstrated in [75]. This approach is more reliable than analyzing the absolute accuracy with respect to LiDAR DSMs. Furthermore, GNSS deformation observations can be considered as ground truth for the validation of InSAR displacement results obtained from motion decomposition. As GNSS measurements report the absolute deformation, they can be used to link the InSAR relative estimates into absolute reference frames. This has been recently done by augmenting a GNSS receiver on a radar transponder used as the reference point for PSI results in [147].

5.2.3 Alternative Validation Procedures

The cross-comparison between InSAR point clouds or GCPs and LiDAR DSMs should be improved by designing more sophisticated and objective comparison strategies. They should also be able to perform validation for an entire urban area and not just for selected test sites. Good examples of such approaches can follow the validation

techniques used in [94], [148], [149]. Also different ground truth data may be utilized for validation. For example, the 3-D position of GCPs or PS can be compared with the output of SAR simulation techniques [150] to improve target association and correct phase center detection.

5.2.4 Enhanced Parameter Estimation

In various parts of this research unknown parameters had to be estimated using redundant measurements. For the estimation of 3-D coordinates of GCPs using stereo SAR, no prior weights were assigned to the timing observations. A weighting based on the targets SCR or ADI can improve the results and decrease the effect of non-suitable observations. Similarly for compensating the DEM error of InSAR reference point, which is a critical step in improving the geocoding accuracy, a combination of weights using the posterior coherence estimates of PSI and the stereo SAR standard deviations, may improve the height offset estimation.

To exclude the effects of outliers in final estimates, parameter estimation using l_1-norm minimization was used for motion decomposition. Instead, a combined use of l_1 and l_2-norm minimization can be used in order to obtain posterior quality indicators for the deformation results. The usage of robust M-Estimators [151] is also recommended. These type of estimators may be also useful for stereo SAR processing in order to limit manual outlier rejection.

5.2.5 Extension to Medium Resolution SAR Data

As the final possible direction for future research, exploitation of the geodetic InSAR framework using SAR data with medium resolution, like Sentinel-1, is recommended. In this regard, the detection of identical targets from cross-heading tracks is a difficult task that has to be initially addressed. Alternatively, matching of PS detected in TSX and Sentinel-1 images of the same scene might be interesting. This can be used in stereo SAR to localize targets using timing measurements from TSX and Sentinel-1. Another possibility is to detect identical PS in both sensors to study the nature of the Sentinel-1 PS by analyzing its constituents in TSX. If the problem of matching is solved, then TSX images could be used for improving the 2-D geo-location of Sentinel-1 images.

A Appendices

A.1 Montazeri, S., Gisinger, C., Eineder, M., Zhu, X.X., 2018. Automatic Detection and Positioning of Ground Control Points Using TerraSAR-X Multiaspect Acquisitions. IEEE Transactions on Geoscience and Remote Sensing 56(5): 2613–2632

Automatic Detection and Positioning of Ground Control Points Using TerraSAR-X Multiaspect Acquisitions

Sina Montazeri[ID], Christoph Gisinger, Michael Eineder, *Fellow, IEEE*,
and Xiao Xiang Zhu[ID] *Senior Member, IEEE*

Abstract—Geodetic stereo synthetic aperture radar (SAR) is capable of absolute 3-D localization of natural persistent scatterers, which allows for ground control point (GCP) generation using only SAR data. The prerequisite for the method to achieve high-precision results is the correct detection of common scatterers in SAR images acquired from different viewing geometries. In this contribution, we describe three strategies for automatic detection of identical targets in SAR images of urban areas taken from different orbit tracks. Moreover, a complete workflow for automatic generation of large number of GCPs using SAR data is presented and its applicability is shown by exploiting TerraSAR-X high-resolution spotlight images over the city of Oulu, Finland, and a test site in Berlin, Germany.

Index Terms—Geodetic stereo synthetic aperture radar (SAR), ground control point, positioning, SAR, TerraSAR-X (TS-X).

I. INTRODUCTION

SYNTHETIC aperture radar (SAR) imaging geodesy and geodetic stereo SAR are relatively new techniques, which aim at high-precision absolute positioning of point targets in SAR images in two-dimensions (2-D) and three-dimensions (3-D), respectively [1]–[3]. The accuracy of both methods, when coupled with data from TerraSAR-X (TS-X) and TanDEM-X, is in the centimeter regime for targets with accurately known phase centers such as corner reflectors [3]. This level of accuracy is achievable due to the precise orbit determination [4] and instrument calibration of the aforementioned satellites followed by a thorough correction scheme which quantifies and removes the most prominent error sources affecting radar timing measurements. This paves the way for remotely sensed generation of ground control points (GCPs) using only SAR data.

The essential prerequisite for applying the geodetic stereo SAR method is the correct detection of identical scatterers in SAR images acquired from different geometries. In this regard, a target can be visible only from *same-heading* orbits, i.e., exclusively ascending or descending orbits, or also from *cross-heading* orbits, which include combinations of ascending and descending orbits. Conceptually, a target localized from the latter is favorable because of the more robust intersection geometry when compared with the former. This fact is demonstrated in Fig. 1(a) where the intersection angle occurs at almost 90° because of the large baseline between the satellites from cross-heading tracks. In Fig. 1(b), the target is localized with satellites from same-heading tracks, which force the baseline to be smaller and consequently the system of equations to solve for the 3-D coordinates to be less sensitive for the perpendicular height direction. However, the rare occurrence of identical scatterers visible from cross-heading configurations as well as the challenging task of automatically detecting such targets, either from same- or cross-heading tracks, currently limit the applicability of geodetic stereo SAR for localization of large number of Persistent Scatterer (PS)s.

To overcome the limitation to some extent, this paper describes an automatic algorithm for detection and absolute positioning of large number of natural PSs in SAR images of urban areas. The candidates are selected from both same-heading and cross-heading geometries based on the methods relying on fusion of multitrack PS interferometry (PSI) point clouds, correspondence detection with optical data, and utilizing vectorized road network data. The candidates are mainly chosen from the same-heading configuration because of the fact that for many PSs the phase centers are assumed to remain unchanged in SAR images. On the other hand, additional candidates are chosen from cross-heading geometries, although in a small number, because conceptually they can be localized more precisely compared to the candidates from same-heading geometries. Coupled with the subsequent geodetic stereo SAR, the proposed processing chain delivers sets of absolutely localized PSs in an investigated area.

The remainder of this paper is organized as follows. Section II reviews the theoretical background of the

techniques utilized in this study and gives an overview of the recent advances and the motivation for this paper. Section III describes the three methods for detecting identical Persistent Scatterer (PS)s visible in SAR images from same- and cross-heading tracks. In Section IV, the complete workflow for generating high-precision absolute GCPs is explained. In Section V, the applicability of the algorithm is demonstrated by exploiting TS-X high-resolution spotlight images over the city of Oulu, Finland, and a test site in Berlin, Germany, and finally, the conclusion is drawn in Section VI.

II. High-Precision Absolute 2-D and 3-D Positioning With TS-X

At the core of high-precision absolute positioning of candidate GCPs using SAR data are the imaging geodesy and the stereo SAR methods. These methods are described in this section followed by the recent advances and applications which rely on absolute localization capability of TS-X. It is important to note that the complete explanation of the theory of the methods and their practical implementations are not in the scope of this paper. For full treatment of these topics, the interested reader is referred to [1]–[7].

A. Background

The SAR imaging geodesy technique aims at achieving 2-D absolute pixel localization [1]. Based on the SAR measurement principle, a single pixel in a focused complex SAR image, processed to zero-Doppler coordinates, is characterized with two time tags: in the along-track direction, the time relative to the time of the closest approach defines the azimuth coordinate t_{az} and in the across-track direction, the difference in the time travel of the transmitted and the received chirp at t_{az} describes the range coordinate τ_{rg} [8]. If we measure the radar timing coordinates (t_{az}, τ_{rg}) for a point target located within the mentioned pixel, the following equations hold:

$$\tau_{rg} = \frac{2R}{c} + \delta\tau_{SD} + \delta\tau_O + \delta\tau_F + \delta\tau_I + \delta\tau_T + \delta\tau_G \quad (1)$$
$$t_{az} = t + \delta t_{SD} + \delta t_O + \delta t_F + \delta t_G \quad (2)$$

where R is the geometric distance from the satellite to the center of the pixel in meters and c is the speed of light in vacuum in m/s, while all the other terms are expressed in seconds. t is the raw acquisition time, $\delta\tau_{SD}$ and δt_{SD} are delays caused by satellite dynamics and electronics, $\delta\tau_O$ and δt_O are the orbit inaccuracies, $\delta\tau_F$ and δt_F are the feature localization errors, and $\delta\tau_G$ and δt_G include the geodynamic effects on all range and azimuth timings, respectively, while $\delta\tau_I$ and $\delta\tau_T$ are the ionospheric and the tropospheric delays considered only for range timings. The magnitude of the individual effects can be scaled to units of length by multiplying the range error terms with $(c/2)$ and the azimuth error terms with the platform's velocity. The outcomes vary from a couple of centimeters for the ionospheric effect, if the satellite operates in X-band, followed by decimeter regimes for satellite electronic delays and geodynamic effects for both components, to up to 4 m for the tropospheric effect depending on the average incidence angle of the acquired TS-X images.

Imaging geodesy corrects for all the error terms in (1) and (2) thus obtaining absolute range and azimuth timings. In this regard, the technique reduces the effects in satellite dynamics by avoiding the stop–go approximation in the TS-X multimode SAR processor and by taking into account the nonzero duration of the pulses and the internal delay caused by the instrument cables [9]. The propagation errors are estimated based on the path delays derived from the nearby global navigation satellite system (GNSS) stations or 3-D integration through weather models followed by appropriate mapping functions [2], [3], [10]. For the geodynamic effects such as solid earth tides, plate tectonics, ocean loading, and atmospheric loading, which change the position of a target on the ground, the corrections are applied based on models issued by the International Earth Rotation and Reference Systems Service (IERS) [11]. Taking into account all the mentioned factors, SAR imaging geodesy is currently capable of providing range and azimuth measurements with 1.16 and 1.85 cm standard deviations, respectively [12].

If a target is visible in SAR images acquired from two or more different viewing geometries, then stereo SAR retrieves the 3-D position of the target by combining the extracted timing information of the target from each SAR image. Furthermore, if the timing coordinates have been *a priori* corrected for the error sources expressed in (1) and (2), the method is called geodetic stereo SAR, which allows for absolute 3-D localization [3]. The relation between the 2-D radar time coordinates of a specific target in the SAR image $\mathbf{x}_T = (t_{az}, \tau_{rg})$ and its corresponding 3-D coordinates on the ground $\mathbf{X}_T = (X, Y, Z)$ is defined by the range-Doppler equation system [8]

$$|\mathbf{X}_S - \mathbf{X}_T| - c \cdot \tau_{rg} = 0 \quad (3)$$
$$\frac{\dot{\mathbf{X}}_S(\mathbf{X}_T - \mathbf{X}_S)}{|\dot{\mathbf{X}}_S||\mathbf{X}_T - \mathbf{X}_S|} = 0 \quad (4)$$

with $\mathbf{X}_S$ and $\dot{\mathbf{X}}_S$ being the position and velocity vector of the satellite relative to t_{az}, and τ_{rg} being the calibrated two-way traveled time from the satellite to the target. t_{az} is implicitly included in (4) relating the state vector of the satellite to the time of the acquisition via a polynomial model [3]. Equation (3) defines a sphere centered on $\mathbf{X}_S$, which reduces to a circle perpendicular to the satellite trajectory when coupled with the zero-Doppler plane described in (4). Therefore, $\mathbf{X}_T$ can be retrieved by including another set of timing observations from a different satellite position, which evaluates the intersection point of the two circles (see Fig. 1). The estimation of the coordinates is carried out by means of least squares plus the stochastic modeling of range and azimuth using the variance component estimation (VCE) [3]. Precision of the estimated 3-D coordinates depends on the signal-to-clutter-ratio (SCR) of the target, the precision of the external radar timing corrections, the separation in the viewing geometries, and the number of acquisitions. Geodetic stereo SAR has been proven to be able to localize corner reflectors with 3-D precision better than 4 cm and an absolute accuracy of 2–3 cm when compared to independently surveyed reference positions [3].
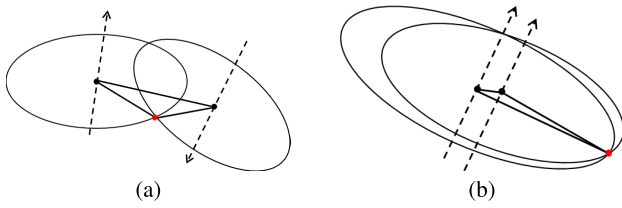
Fig. 1. Localization of a point target (red dot) from (a) cross-heading and (b) same-heading satellite tracks. The satellites are shown by black dots; their trajectories are presented by dashed lines and the baselines are depicted by solid lines between the satellite positions. The black circles are defined by the range-Doppler equations and their intersection leads to the 3-D position of the target.

B. Recent Advances and Motivation

In our previous research, geodetic stereo SAR has been also applied to a small number of natural PSs in urban areas where it could localize targets in 3-D with a precision better than 1 dm using TS-X high-resolution spotlight products [3]. The PSs were manually extracted from SAR images and originated from building facades for candidates visible in same-heading tracks or from the base of street lights for candidates visible from cross-heading tracks.

In [13], the first attempt for automatic timing extraction and matching of limited number of PSs originated from a building facade visible in TS-X images from two same-heading tracks was reported. In this paper, the geodetic stereo SAR method was extended to include the secular movement of the PSs as well as their 3-D absolute positions. The averaged 3-D precision was reported to be below 1 dm with encouraging results for estimating the plate tectonics using SAR data.

In [14], the concepts of imaging geodesy and stereo SAR were used to transform the relative estimates of SAR tomography into absolute 3-D point clouds by absolutely localizing the manually extracted reference point. The method, termed geodetic SAR tomography, allows for the generation of dense point clouds with an absolute localization accuracy in the order of 20 cm and is the basis for geodetic fusion of multiaspect interferometric SAR (InSAR) point clouds. The latter enables the decomposition of deformation estimates from SAR tomography into highly detailed 3-D displacement maps [15].

Automatic extraction of Ground Control Point (GCP)s from SAR products has been carried out for TS-X and COSMO-SkyMed in [16] and [17], respectively. Both methods focus on the detection of stereo candidates that presumably originate from street lights or traffic signs and are visible in SAR images as bright isolated points. Therefore, the majority of PSs in urban areas which stem from building corners or facades are not considered as potential candidates in these methods due to complex radar reflection properties in such scenarios. Furthermore, in [16], the 3-D positioning is done only with same-heading geometry configurations, and therefore, the error ellipsoid of the scatterers' coordinates is highly skewed in the cross-range direction [3], [18]. Nevertheless, the retrieved 3-D coordinates of several candidates were compared to their true positions observed with GNSS which confirmed an absolute accuracy better than 20 cm in each coordinate component [16].

Based on the above-mentioned studies, the motivation for carrying out this paper is fourfold.

1) The manual detection, extraction, and matching of PS candidates from SAR images acquired from different geometries are cumbersome and should be replaced by an automatic process.
2) The algorithm should be able to automatically detect and match identical PSs visible from cross-heading geometries in order to boost the precision of the retrieved 3-D coordinates.
3) The total number of high-quality PS candidates to be localized as GCPs should be as large as possible. This indicates that the majority of PSs in urban areas which stem from buildings should also be considered as candidates for 3-D absolute localization from the same-heading tracks.
4) The distribution of the GCPs should be as homogeneous as possible in the entire investigated area. This allows for the generation of an absolute reference network to be integrated into relative InSAR techniques.

III. DETECTION OF IDENTICAL PSs IN MULTIASPECT SAR IMAGES

Detection of identical PSs from SAR images acquired with different viewing geometries is a challenging task. This is because of the geometrical distortions of SAR images due to the oblique viewing geometry and less importantly the presence of speckle. Moreover, in urban areas captured by SAR sensors, which is the focus of this paper, the backscattering mechanism is highly complex because of the existence of several phase centers close to each other. Therefore, identical PS matching becomes even more difficult for multiaspect SAR images of urban areas.

In recent years, there have been several studies that explored the possibility to match features between SAR images. In [19], the capabilities of the conventional scale invariant feature transform (SIFT) algorithm [20], which is commonly used for feature extraction and matching between optical images, were extended to be suitable for SAR images. In [21], the SAR-SIFT algorithm has been proposed, which focuses on the efficient extraction of local descriptors from SAR images by modifying the SIFT algorithm to take into account the statistical properties of speckle. However, both of the aforementioned methods are applicable only to SAR images taken from same-heading orbits with small difference in the respective incidence angles. Specifically for the task of automatic 3-D positioning, Nitti *et al.* [17] have proposed to identify identical scatterers based on the detection of local features using the Harris corners. This is followed by constraining the search space by geocoding the local features, using an external digital elevation model and orbit information, and eventually using SIFT for the feature matching. Although the method is promising in terms of detection and positioning of targets even from cross-heading tracks, it only works on isolated PSs.

In the following, we describe in detail the three strategies we apply for the detection of identical PSs in SAR images acquired from the same- and cross-heading orbits. The methods do not tackle the detection problem directly within the

SAR images, but instead rely on external geospatial data and on limiting the search space on georeferenced positions of the PSs.

A. Multitrack PSI Point Cloud Fusion

In [22], a method for geometrical fusion of multitrack PSI point clouds has been proposed. The fusion algorithm, which is based on the geocoded PSI point cloud solutions of each geometry as well as information on heading and looking angle of the satellites, consists of three major parts, namely: 1) generation of initial point correspondences; 2) restricted least squares adjustment to minimize the distance between assumed identical points visible from different viewing geometries; and 3) adding a range-dependent shift to all PSs using the result of the previous step for the final registration. A summary of the method is described in the following. For a detailed description of the algorithm, the reader is referred to [22] and [23].

Since we are interested in the detection of large number of point correspondences, only the first part of the algorithm is relevant. This coarse registration is performed based on the cross correlation of a subset of geocoded PS point clouds from different geometries, after projection on a regular grid, in the xy-, xz-, and yz-planes. The subset is chosen based on the precision of height update estimates available for each PS after carrying out Persistent Scatterer Interferometry (PSI) [23], [24]. The resulting horizontal and vertical shifts from the mentioned cross correlation procedure are applied to the PSs of one point cloud to align them with the PSs of the other point cloud. The coarse shifts are further refined prior to the selection of corresponding PS pairs. The refinement is carried out inside a small neighborhood around each PS which includes the PSs from the other point cloud and tends to accomplish it by performing a statistical search to find the best fit between both 3-D point clouds [23]. The refined shift is applied to the PS point cloud of one acquisition geometry, and a one-by-one PS correspondence is detected in the other point cloud. At the final step, the 3-D coordinates of the geocoded PSs have to be projected on the SAR images of each orbit track, a process called radar-coding. Since the matching of the PSI results is performed on coordinates in the Universal Transverse Mercator (UTM) map projection, the coordinates are first converted to the Cartesian geocentric system as $(X, Y, Z)_i$ for the ith PS. Subsequently, the range-Doppler equations described in (3) and (4) are inverted to obtain the azimuth and range timing coordinates $(t_{az}, \tau_{rg})_i$, which can be easily expressed in pixels in the radar coordinate system $(L, P)_i$ by knowledge of the range sampling frequency, pulse repetition frequency, the first sampled azimuth time, and the first sampled range time for each acquisition. The latter information is stated in the product annotation files accompanied by the TS-X image products [25].

For the same-heading tracks, this method typically generates 200–2000 point correspondences per kilometer square depending on how densely constructed is a city which directly affects the total number of PSs in each point cloud.

B. Template Matching on Optical Data

Given the availability of suitable remotely sensed optical data, one can detect candidate objects from optical images which are probable to be observed in SAR images from different viewing geometries. In urban areas, scatterers, which are good candidates to be visible from both the same-heading and cross-heading tracks, usually originate from lamp poles or other cylindrical objects that are vertically oriented toward the sensor. Therefore, the basic idea when using optical data for the aid of GCP identification is to detect lamp poles and match the detected objects to the corresponding bright points in SAR images.

The method identifies lamp poles based on their distinctive shadows in optical images using a template matching scheme [26]. Prior to extracting the template, common preprocessing steps such as noise filtering and histogram equalization are carried out on the optical image. Additionally, in order to make the shadows of lamp poles more prominent, a simple sharpening procedure is carried out as follows:

$$\mathbf{I} = \mathbf{I}_o + a\mathbf{I}_m \tag{5}$$

where $\mathbf{I}$ is the sharpened image, $\mathbf{I}_o$ is the preprocessed original image, a is the scalar sharpening factor, and $\mathbf{I}_m$ is the unsharp mask. $\mathbf{I}_m$ is calculated as the difference between $\mathbf{I}_o$ and its blurred version. Higher values of factor a means the higher level of sharpening. The process expressed in (5) is called high boost filtering [27].

After the sharpening, the template is extracted based on the shadow of an arbitrary lamp pole visible in the optical image. The template is then correlated with the reference image to calculate the following similarity measure for each pixel (u, v) in the reference image [28]:

$$\rho(u, v) = \frac{\Sigma_{x,y}[\mathbf{I}(x, y) - \bar{\mathbf{I}}_{u,v}][\mathbf{T}(x - u, y - v) - \bar{\mathbf{T}}]}{\sqrt{\Sigma_{x,y}[\mathbf{I}(x, y) - \bar{\mathbf{I}}_{u,v}]^2 \ \Sigma_{x,y}[\mathbf{T}(x - u, y - v) - \bar{\mathbf{T}}]^2}} \tag{6}$$

where $\mathbf{I}(x, y)$ and $\mathbf{T}(x, y)$ denote the pixel values of the reference and the template image at (x, y), respectively, and $\Sigma_{x,y}$ stands for $\Sigma_{x=1}^{N_1} \Sigma_{y=1}^{N_2}$ with $N_1 \times N_2$ being the size of the template. Furthermore, $\bar{\mathbf{I}}_{u,v}$ and $\bar{\mathbf{T}}$ denote the mean intensity values of the original image and the template, respectively. Equation (6) allows for the calculation of the normalized cross correlation (NCC) value $\rho(u, v)$, which leads to the detection of the template location in the reference image after proper thresholding. It is important to note that due to the normalization carried out in the denominator of (6), $\rho(u, v)$ is independent of changes in brightness or contrast of the image and, therefore, improves the result of template matching.

After detection of pixels that belong to the shadows of lamp poles, the result is georeferenced in the UTM coordinate system. Since more than one pixel exist for each lamp pole in the optical image, which represent the object, a subsequent clustering is performed. The clustering is carried out

nonparametrically using the mean shift concept [29]

$$\mathbf{M}(\mathbf{p}_i) = \frac{\Sigma_{j=1}^n \; \mathbf{p}_j \; g\left(\left\|\frac{\mathbf{p}_i - \mathbf{p}_j}{h}\right\|^2\right)}{\Sigma_{j=1}^n \; g\left(\left\|\frac{\mathbf{p}_i - \mathbf{p}_j}{h}\right\|^2\right)} - \mathbf{p}_i \qquad (7)$$

where $\mathbf{p}_i$ denotes a 3-D point for which the shift vector $\mathbf{M}(\mathbf{p}_i)$ is calculated. $\mathbf{p}_j$ represents the points in a neighborhood of $\mathbf{p}_i$, g is a kernel function with the bandwidth h, and $\|\cdot\|$ is the Euclidean distance operator. The main idea of the algorithm is to shift each point in a small neighborhood toward its weighted mean value and thus representing each cluster by its centroid [29]. The process is carried out iteratively until the length of $\mathbf{M}(\mathbf{p}_i)$ becomes equal or close to zero. For our application, since in any case there will be a mismatch between the detected points on optical data and the corresponding bright points in the SAR image, utilizing a flat kernel in (7) suffices. This means the algorithm is simplified by calculating the sample mean in a specified radius of $\mathbf{p}_i$ and shifting the desired point toward the estimated center.

In the next step, the clustered points with UTM coordinates should be radar-coded to all the available SAR images. As it was mentioned earlier, the positions of the detected lamp poles on the optical data and the bright PSs in the SAR image will most probably not coincide after radar-coding. This can be explained by height uncertainties of the georeferenced optical data and the fact that the data may not be perfectly orthorectified. Therefore, in the final step of the algorithm, the detected lamp poles are registered on the corresponding bright dots in the SAR image by employing the iterative closest point (ICP) algorithm [30]. To this end, binary masks are generated based on thresholding on the bright points on the SAR image and keeping only the detected lamp poles from the optical data. The ICP algorithm then finds for each individual point its closest point in the corresponding point set. It iteratively estimates the transformation parameters (translation and rotation) to minimize the mean squared error between the two point sets and finally registers one point cloud onto the other point cloud with the refined transformation parameters.

It is noteworthy that the positioning accuracy of the utilized optical image does not necessarily have to be high. A horizontal positioning accuracy in the order of couple of meters and an approximate knowledge of height based on freely available sources usually suffice for the procedure described in this section. If a mismatch occurs due to the low positioning accuracy, this will be compensated by the final step of the algorithm with applying the ICP. However, the spatial resolution of the optical data should be strictly high, 10 cm or better, in order to be able to accurately detect the shadows of the lamp poles.

C. Vector Road Network Data

In urban areas, the cylindrical objects of our interest (lamp poles, road signs, traffic lights, etc.) are typically located along the roads. Therefore, with the availability of geospatial road data, either obtained from OpenStreetMap or country-specific geoportals, and the projection of such maps on SAR images,

one can search for bright points in the neighborhood of the road data points.

The method is applied to co-registered stacks of SAR images. If SAR stacks from multiple viewing geometries are available, first the road data, which are usually delivered in the UTM coordinate system, are radar-coded based on the master orbit information of each stack. It is important to note that if the road data do not have any information about the ellipsoidal height, then for the radar-coding a constant height value based on prior knowledge can be chosen for all the road data points. Furthermore, the data with horizontal positioning accuracy in the order of couple of meters will suffice for the PS matching procedure as the PS correspondences are detected on a neighborhood-analysis basis, as is depicted in the following, which does not require the exact position of the road data point.

After radar-coding, a circular neighborhood is considered around each road data point. The radius of the circle depends on the typical width of streets and highways. Subsequently, for each pixel within the neighborhood the amplitude dispersion index (ADI) is evaluated [31]

$$D_a \approx \frac{\sigma_a}{\bar{a}} \qquad (8)$$

where σ_a and $\bar{a}$ are the temporal standard deviation and the temporal mean of calibrated amplitude values of the pixel, respectively, and D_a approximates the phase dispersion. The pixel with the lowest value of D_a, i.e., the one with the highest phase stability is chosen as potential PS candidate. This process is carried out for all of the available road data points. Since at this point, it is possible that many false pixels with relatively low D_a values in the neighborhoods are categorized as potential GCP candidates, a further thresholding on D_a is performed in SAR images from all available viewing geometries. This operation, in addition to constraining the approximate elevation of the PS candidates to be close to the ground, causes a dramatic decrease in the total number of detected candidates, but improves the accuracy of the detection.

Finally, the presumable identical PSs in all available geometries are geocoded using the respective master orbit information. In the geocoded results, the PSs which are close enough, in terms of coordinate differences, are selected as the final GCP candidates.

IV. AUTOMATIC GCP GENERATION: PROCESSING CHAIN

The processing chain for automatic detection and positioning of GCPs includes a set of procedures that starts from single-look slant range complex SAR images and their corresponding product annotation files to absolute 3-D coordinates of the chosen GCPs. The flowchart of the algorithm is shown in Fig. 2. It consists of the following major processes that have to be carried out in the stated order:
1) identification of identical scatterers visible in multiaspect SAR images.
2) precise extraction of scatterers' azimuth and range positions from SAR images at subpixel level.
3) scatterer visibility check and initial removal of outliers from time series of phase noise.

Fig. 2. Flowchart of the automatic GCP generation algorithm. The input data is the single-look slant range complex (SLC) SAR images and their accompanying L1B product annotation files. The identical PS detection part is drawn inside the gray dashed rectangle. In the case of PS detection from cross-heading viewing geometries, the algorithm receives as input the optical image or road network data of the scene upon availability. The processes are depicted with big blue rectangles, while the subprocesses are shown with small gray rectangles. The name of each process step is written in bold letters. The last blue rectangle includes the processes and subprocesses that all the detected candidates should go through independent of the detection methodology.

4) correction of radar timings for all the perturbing signals.
5) estimation of the 3-D absolute coordinates of the scatterers.

This section discusses each step. It is important to note that since some of the steps are well-established and well-documented techniques in the SAR community, all the details

will not be repeated here. Instead, for these processes, the relevant references are provided.

A. Identification of Identical PS

This section has been already covered in Section III and is shown as the processes inside the gray dashed rectangle in Fig. 2. For detection of PS candidates from same-heading orbits, the PSI processing is carried out following the guidelines from [24], [31]–[33] and the identical PS matching is done using the PSI multitrack fusion algorithm described in [22] and [23]. In the case of localization from cross-heading tracks, the PS candidate selection is carried out using the methods based on the optical data or on the road network data. Regardless of the detection method, the output of this step is the approximate radar coordinates of identical PSs in terms of lines and samples in all non-coregistered images of different orbit tracks.

B. Precise Extraction of PS Timings

The rough radar coordinates of the PSs from the previous step should be refined to subpixel level in order to extract the timings precisely. To this end, a process called point target analysis (PTA) is carried out [8]. In each image of the scene, a 32×32 window centered on the PS is extracted. In both range and azimuth directions, an oversampling by a factor of 32 is performed and the integer peak position of the PS response is measured. Subsequently in a 3×3 window centered on the peak position, a paraboloid interpolation is performed to refine the values around the maximum. This method allows the retrieval of the peak position of the PS with a sensitivity better than $(1/1000)$ of a pixel in each dimension [5]. These values are then converted to radar timings based on the product annotation files [25].

Based on the result of the PTA for each PS, the refined peak power P_{peak} and the clutter power P_{clutter} can be computed. These values, if expressed in decibel, are related to the SCR of the target as

$$\text{SCR} = 10^{\left(\frac{P_{\text{peak}} - P_{\text{clutter}}}{10} \right)} \tag{9}$$

which is expressed as a digital number.

C. PS Visibility Check and Initial Outlier Removal

The SCR of potential PS candidates should be high enough in the stack of SAR images so that the PS can be localized with high precision. Therefore, we analyze the time series of phase noise values σ_ϕ of the PSs to exclude potential outliers and check whether the scatterer is visible in one data take. The σ_ϕ of a PS in acquisition i is related to the Signal-to-Clutter-Ratio (SCR) of the target as [34]

$$\sigma_{\phi_i} \approx \frac{1}{\sqrt{2 \, \text{SCR}_i}} \tag{10}$$

which is expressed in radians. The values of σ_ϕ, for a specific PS in all data takes, are nonnegative and follow a right-skewed distribution. Removal of outliers based on statistical measures such as mean or median is not recommended

since many regular values can be categorized as outliers. Therefore, we use a method called *adjusted boxplot* which allows for robust elimination of outliers in univariate skewed distributions [35].

The main idea of the adjusted boxplot is to modify the original boxplot method, described in [36], to include information about the skewness of the data. Therefore, instead of classifying an observation as outlier if it lies outside of the interval defined by the boxplot method [36]

$$[Q_1 - 1.5 \, \text{IQR}; \quad Q_3 + 1.5 \, \text{IQR}] \tag{11}$$

the adjusted boxplot method declares an observation as outlier if its value exceeds the following interval [35]:

$$[Q_1 - 1.5 \, e^{(-4\text{MC})} \, \text{IQR}; \quad Q_3 + 1.5 \, e^{(3\text{MC})} \, \text{IQR}]. \tag{12}$$

In (11), Q_1 and Q_3 are the first and the third quartiles of the data, respectively, and $\text{IQR} = Q_3 - Q_1$ denotes the interquartile range. In (12), MC is the medcouple, a robust measure of the skewness of a univariate sample which for the right-skewed distributions is always nonnegative [37]. The exponential functions in (12), which depend on the MC as well as the included coefficients, are chosen experimentally based on some well-known skewed distributions. For more details on the theory and implementation of the adjusted boxplot method, the reader is referred to [35].

After the automatic identification and exclusion of σ_ϕ values which do not lie within the interval of (12), the remaining time series is analyzed to remove the data takes in which the specific PS is not visible. This is done by removing all σ_ϕ values which are above 0.5 rad ($\approx 30°$) as is stated in [24].

D. Correction of PS Timings in a Stack

The correction of the extracted PS timings is performed using the imaging geodesy technique [1], which was briefly introduced in Section II-A. It is worth mentioning that the tropospheric and ionospheric effects are corrected based on global numerical weather models and global ionospheric maps, respectively, if local GNSS receivers are not available in the vicinity of the investigated area. Along with these corrections comes the corresponding geometrical calibration of the SAR sensor in range and azimuth which ensures centimeter localization accuracy. The calibration is based on corner reflectors with known reference coordinates [1]. The output of this part is the absolute 2-D radar timing coordinates.

E. Absolute 3-D Localization of PS

At the final step in the processing chain, the corrected range and azimuth timings from the entire multiaspect set of SAR images are combined to retrieve the absolute 3-D position of the PS with the stereo SAR method described in Section II-A. Apart from the 3-D position of the target, stereo SAR reports on the standard deviation of each coordinate component (S_X, S_Y, and S_Z) as the by-product of the least squares adjustment. Furthermore, the observation quality of each PS, i.e., the azimuth and range standard deviations (S_{az} and S_{rg}), retrieved from applying VCE to residuals are delivered. It is important to note that the VCE is carried
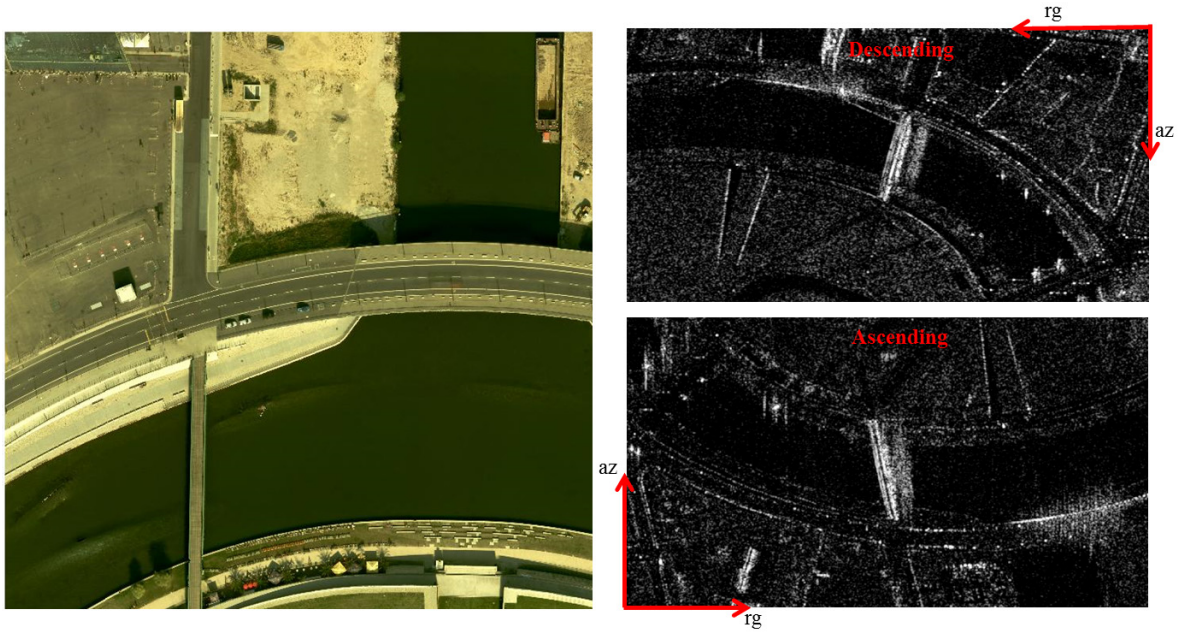
Fig. 3. (Left) Optical image and (right) SAR amplitude images of the test area in Berlin. The contrast of the optical image has been adjusted to illustrate the shadows of the lamp poles prominently.

out individually for each geometry which allows to judge the consistency of the observed geometries with respect to the underlying assumption that the intersection occurs at a common PS.

The residuals of the adjusted range and azimuth observations are the basis for the elimination of outliers after stereo SAR processing. Therefore, the processing is carried out repeatedly, where first the initial 3-D coordinates are estimated using the provided input timings. The range and azimuth residuals are analyzed to exclude observations with residual values larger than 0.6 m in range or larger than 1.1 m in azimuth. The upper bounds correspond to the nominal spatial resolution of TS-X high-resolution spotlight products used in this paper. Then the stereo SAR is performed again with the cleaned observations. This time, observations that show residuals larger than two times the S_{az} and the S_{rg} are removed. Additionally, to remove the PSs, which are not considered ideal for stereo SAR due to wrong correspondence matching caused by several scatterers being too close, the third step of data cleaning is performed. PSs having an S_{az} higher than 20 cm in any of their azimuth geometries are removed based on the assumption that the discrepancy should not exceed the typical size of the PS object, for instance, a lamp pole.

The estimated variance–covariance matrix of the 3-D position of the target is further used for error analysis. The matrix gives important information about the stability of the coordinates' results and is affected by the factors stated in Section II-A.

V. EXPERIMENTAL RESULTS

In this section, the workflow described in Section IV is applied on real data to produce remotely sensed SAR-based GCPs. In Section V-A, the results are reported for a small test site in Berlin where the detection of GCP

TABLE I
ACQUISITION PARAMETERS OF STACKS OF SAR IMAGES IN BERLIN

Beam Nr.	Incidence angle (degree)	Heading angle (degree)	Track type	Nr. of images
57	41.9	350.3	Asc	107
42	36.1	190.6	Dsc	107

candidates are carried out using cross-heading geometries. In Section V-B, the processing results are shown for the entire city of Oulu, Finland, where the detection and positioning are performed on PS candidates detected from both the same- and cross-heading orbit tracks using the methods described in Sections III-A and III-C, respectively.

A. Berlin

The first test site includes a small area close to the Berlin central railway station. The SAR data set, 214 images in total, consists of two stacks of TS-X very high-resolution spotlight images acquired with a range bandwidth of 300 MHz. The images cover a period of approximately six years from April 2010 to March 2015. In terms of viewing geometry, one stack was acquired from a descending orbit with images recorded at 05:20 coordinated universal time (UTC), and one stack was acquired from an ascending track with images recorded at 16:50 UTC. The acquisition parameters of each stack are summarized in Table I.

For the selected test site, an aerial optical image with ground spacing of 7 cm is also available. The optical image is orthorectified and was used in a stereo matching process to produce a digital surface model with decimeter accuracy [38]. The optical image of the test site and the corresponding SAR images are shown in Fig. 3.
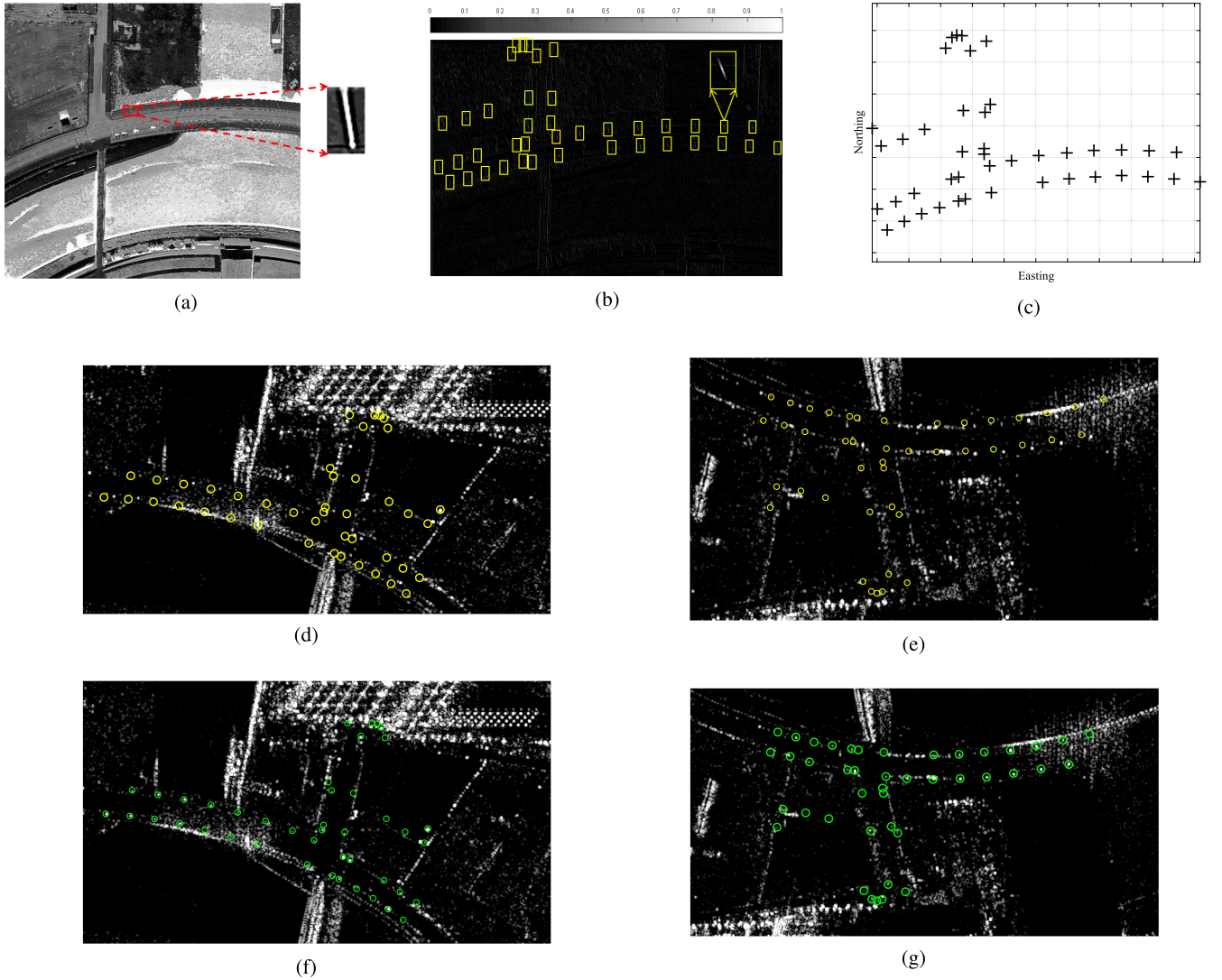
Fig. 4. Demonstration of PS correspondence detection in Berlin based on high-resolution optical data. (a) Preprocessed optical image after negative intensity transformation and the extracted template. (b) Calculated NCC map after correlating the extracted lamp pole template with the reference image in which the detected objects are marked by yellow rectangles. (c) 44 detected objects in the UTM coordinate system after clustering. In (d) and (e), the radar-coded results are depicted by yellow circles, which show offsets with respect to the bright points in the SAR images. (f) and (g) Results of matching after using the ICP algorithm on the descending and the ascending image, respectively. In (f) and (g), it can be seen that the detected objects from the optical image (green circles) coincide with the visible bright points in the SAR images.

The GCP candidate selection is carried out based on the optical data which includes the detection of lamp poles and their projection onto cross-heading SAR images (see Section III-B). The individual steps of this process, applied on the Berlin test site, are shown in Fig. 4. After the extraction of the template [see Fig. 4(a)], the NCC map is calculated and pixels with values higher than 0.6 are classified as parts of shadows of lamp poles as illustrated in Fig. 4(b). It is important to note that the exact tuning of the threshold is not necessary as long as the value is chosen low enough. If the threshold is strictly chosen as a high value, although we are selecting the most similar pixels to the template, we may ignore all the other candidates which show less similarity to the template but might have been potential candidates for stereo SAR processing and 3-D localization. Therefore,

in our processing chain, the default value is set to 0.6, which is slightly higher than the half of the Normalized Cross correlation (NCC) range [0,1], to guarantee a certain degree of similarity while selecting a large number of pixels. In the Berlin case study, this leads to the selection of 2030 pixels which are further clustered to represent the 44 detected objects in the UTM coordinate system [see Fig. 4(c)]. The objects include the lamp poles along the bridge and at the street perpendicular to the bridge as well as flag poles at the top left of the optical image close to the Berlin railway station. After radar-coding of the results onto the entire ascending and descending SAR images, the mismatch between the projected points and the actual bright points in the SAR images, as depicted in Fig. 4(d) and (e), is resolved using ICP. The detection outcome is marked with green circles
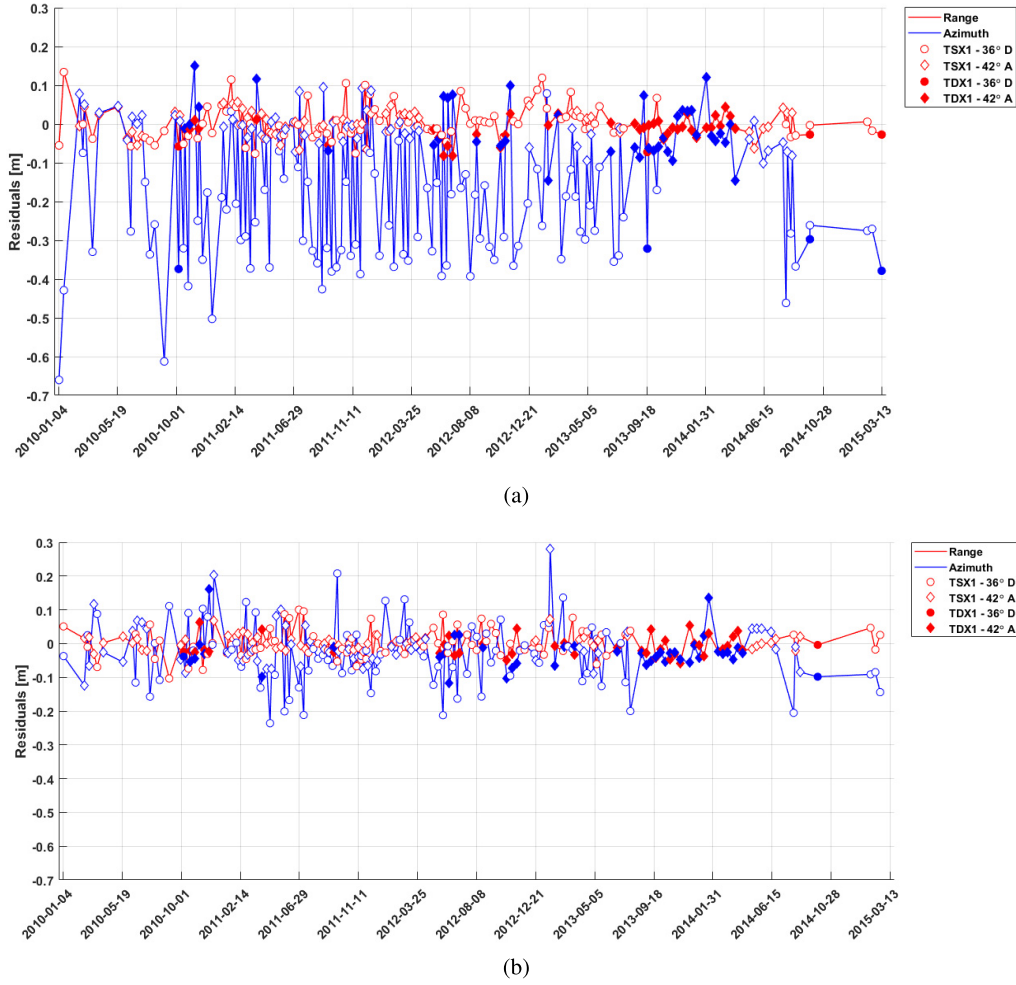
Fig. 5. Range and azimuth residuals for two PS examples of the Berlin test case after the stereo SAR processing. Step 1 (gross outlier detection) and step 2 (2σ test) have already been applied. (a) PS_{41}. (b) PS_{43}.

in Fig. 4(f) and (g). It is seen from the results that not all of the available street lights can be detected using the mentioned strategy because some can be occluded by cars or the object's shadow is not distinctive enough to match the extracted template. However, this is of low importance in our application since in such a small test site, with an area less than 2 km², two or three GCPs will certainly suffice. Moreover, if the method detects wrong candidates, which do not fall in the category of PSs, the subsequent Point Target Analysis (PTA) and phase noise analysis will discard these points.

The precise radar timings of the PSs are extracted using PTA (see Section IV-B). Subsequently for each PS candidate in each data take, SCR and σ_ϕ values are evaluated using (9) and (10), respectively. After excluding potential outliers with the adjusted boxplot method, the data takes in which the scatterer is not visible are discarded by thresholding on the σ_ϕ values (see Section IV-C). At this point, one might argue that analyzing the time series of the remaining σ_ϕ values can already give a hint if the scatterer is a suitable candidate for positioning or not. This statement is partially true since the mentioned analysis is only useful to separate the time-coherent scatterers from the noncoherent ones. It is

possible that several scatterers, located close to each other, are mapped as one bright point which results in low σ_ϕ value but of course not a suitable candidate for stereo SAR. These candidates are usually discarded in the stereo SAR processing due to large S_{az} values which indicate that not the same object has been detected from multiple viewing geometries. Therefore, in our processing chain, the detected candidates are not entirely removed based on their average SCR and they are all passed to the final stereo SAR processing.

As for the corrections, the ionospheric delay is estimated using global ionospheric maps. The tropospheric delay is estimated using the zenith path delay information of the closest permanent Global Navigation Satellite System (GNSS) station in Potsdam which is situated approximately 35 km away from the test site.

The positioning of the 44 PSs is carried out using stereo SAR (see Section II-A) and is followed by outlier elimination according to the criteria described in Section IV-E. Starting from the corrected input timings, the first solution is analyzed for gross outliers in the observations exceeding the resolution of the underlying TS-X high-resolution spotlight product. Applying these thresholds to the residuals of the
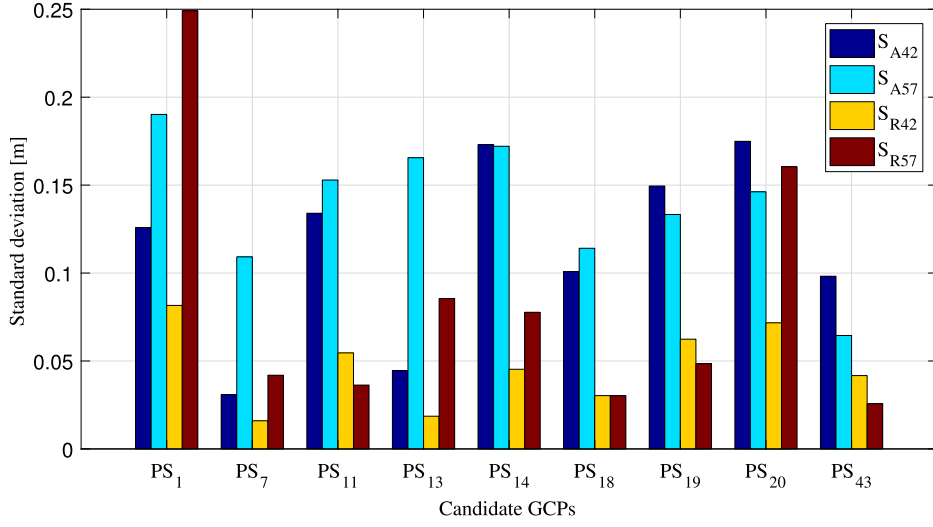
Fig. 6. Posterior azimuth and range standard deviations (S_{az}, S_{rg}) of the nine best GCP candidates estimated by VCE in the geodetic stereo SAR processing.
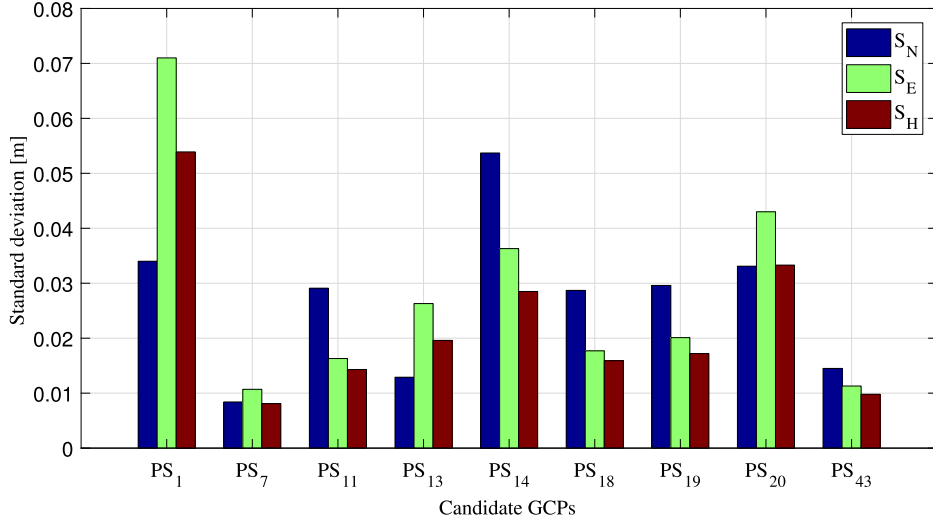

Fig. 7. Posterior standard deviations scaled to 95% confidence level of the nine best GCP candidates estimated posterior to the geodetic stereo SAR processing. The standard deviations are defined in the local coordinate system of Berlin in north, east, and vertical direction (S_N, S_E, and S_H).

adjusted observations reduces the number of solvable PSs from 44 to 42, because in the case of an obvious mismatch, all the observations of one geometry are removed. Moreover, the total number of observations is reduced by 25%, but this strongly varies across the individual PSs. After recomputation and application of the 2σ test using the estimated standard deviations from the VCE (S_{az}, S_{rg}), the number of solvable PSs remains 42 and the total amount of observations is reduced by another 8%. At this stage, the data is fairly cleaned at the observation level regarding the individual range and azimuth geometries, but their consistency has not been considered so far. Looking at the observation residuals of the two PSs displayed in Fig. 5 reveal that there are cases for which the azimuth of one geometry is clearly biased, because we try to combine data from different phase centers. In the ideal case, the algorithm yields a coordinate solution for which all sets of observations (two sets of azimuth and

two sets of range) can fulfill the range-Doppler positioning model of (3) and (4). For a mismatch or a spatial separation of the phase centers, the usually more precise range observations dominate the solution, and only one of the azimuth data sets may fit the estimated coordinates without a bias, but not the second set of azimuth data. Such a situation is illustrated by PS_{41} [see Fig. 5(a)], where the $36°$ azimuth displays a prominent bias of about -30 cm. To a certain degree this must be accepted since we cannot expect ideal multidirectional PSs, e.g., the lamp poles have a certain diameter. In this paper, as mentioned in Section IV-E, we define an empirical limit of 20 cm of what we consider acceptable, which removes candidates like the PS_{41} during the final processing step. Therefore, the residual results of scatterers such as PS_{43} [see Fig. 5(b)] may be seen as the best case scenario. The remaining difference in quality between range and azimuth is due to the nonsquare product reso-

TABLE II

RESULT OF CROSS COMPARISON BETWEEN THE ESTIMATED HEIGHTS
OF STEREO SAR h_S AND THEIR CORRESPONDING LiDAR
HEIGHTS h_L. THE OFFSET h_o IS AN INDICATOR
FOR THE ABSOLUTE ACCURACY OF h_S

GCP	h_L [m]	h_S [m]	$h_o = h_S - h_L$ [m]
PS_1	74.64	74.80	0.16
PS_7	74.45	74.54	0.09
PS_{11}	74.83	74.99	0.16
PS_{13}	75.40	75.62	0.22
PS_{14}	73.87	73.96	0.09
PS_{18}	73.87	73.95	0.08
PS_{19}	75.59	75.76	0.17
PS_{20}	75.02	75.18	0.16
PS_{43}	79.78	79.85	0.07
Mean			**0.13**
Standard deviation			**0.05**

TABLE III

ACQUISITION PARAMETERS OF STACKS OF SAR IMAGES IN OULU

Beam Nr.	Incidence angle (degree)	Heading angle (degree)	Track type	Nr. of images
30	30.9	346.1	Asc	44
54	41.1	191.4	Dsc	44
69	46.2	350.0	Asc	38
94	53.4	187.5	Dsc	51

lution, i.e., the TS-X spotlight SLC data have a resolution of 0.6 m × 1.1 m in range and azimuth, respectively [25].

For the aforementioned reasons, only nine PSs remain which we consider as good GCP candidates. The bar graphs of Figs. 6 and 7 summarize the quality of these PSs. Fig. 6 shows the posterior estimated standard deviations of observations. The S_{az} values vary from 3 to 19 cm with an average of 12 cm, while the S_{rg} values range from 1 to 24 cm with an average of 6 cm. This indicates that, for these natural PSs, the removal of error terms, as expressed in (1) and (2), and discarding the outliers allow subdecimeter and decimeter precision in range and azimuth, respectively. In Fig. 7, the positioning quality is assessed by reporting the precision of the estimated coordinates. The standard deviations are defined in the local coordinate system of Berlin in the north, east, and height (S_N, S_E, and S_H) in the confidence level of 95%. The mean values of S_N, S_E, and S_H are 2.7, 2.8, and 2.2 cm, respectively. The higher precision in the height direction is merely the effect of the cross-heading geometry used for the positioning.

Until now, the discussion was mostly focused on analyzing the relative accuracy of the coordinates based on the posterior precision estimates. Although the most reasonable procedure to validate the absolute accuracy is a point-wise comparison of stereo SAR coordinates with respect to the corresponding GNSS-surveyed ones, this was not applicable at the time of the study. Instead, the stereo SAR estimated ellipsoidal heights of the nine GCPs were compared with the LiDAR heights of the same area. We assume the phase centers of the detected GCPs are at the base of the lamp poles on the ground. Therefore, the cross comparison includes finding the nearest neighbors of the GCP candidate in the LiDAR point cloud within the radius of 1 m, excluding the LiDAR points with large height values which originate from the top of the lamp pole, estimating the mode of the LiDAR heights to represent the reference height, and evaluating the difference between the ellipsoidal height of stereo SAR results with respect to the reference height. The radius of the neighborhood is chosen in a way that the reference height calculation includes a reasonable number of samples and still be small enough to possibly

prevent the inclusion of different objects in the search window. It is also worth to note that the calculation of mode is carried out with the assumption that the majority of samples in the window stems from the ground. The results of the cross comparison are reported in Table II. The estimated stereo SAR and approximated LiDAR reference heights are denoted by h_S and h_L, while their offset is represented by h_o. It is seen that for all except for one of the GCPs, the height offset is below 20 cm. The results report a bias of 13 cm and a precision of 5 cm overall with respect to the LiDAR data which roughly implies the absolute accuracy of the height estimation using the stereo SAR method.

B. Oulu

The second test site covers the entire city of Oulu. The SAR data include four stacks of TS-X high-resolution spotlight products with 177 images in total. The images were acquired from May 2014 to October 2016, from two ascending orbits and two descending orbits. The acquisition parameters of the Oulu data set are reported in Table III, while the mean scene coverage and the acquisition time plot of the TS-X images are shown in Fig. 8. No images were ordered during the periods from November 2014 to March 2015 and November 2015 to March 2016 due to the accumulation of snow expected in Oulu during winter.

For Oulu, no optical images with sufficient spatial resolution were available for the detection of PS candidates from cross-heading geometries. Therefore, the road network data of Oulu was used instead. The data was freely accessed from the Finnish Transport Agency [39]. It was delivered in vector format in the UTM coordinate system and includes the main streets and highways of Oulu.

The detection of PS candidates from the same-heading tracks was carried out using the multitrack PSI fusion algorithm described in Section III-A. As the prerequisite of the algorithm, the PSI processing was performed by the PSI-GENESIS of the German Aerospace Center [33]. For each detected PS, the elevation and the deformation parameters (in this paper only a linear trend) were estimated. As an example, the radar-coded PS elevation map of the ascending stack of beam 30 is visualized in Fig. 9. After geocoding, the 3-D point clouds obtained from either ascending–ascending (AA) or descending–descending (DD) geometries form the initial input of the fusion algorithm as demonstrated in Fig. 10.

Fig. 10(a) shows the geocoded PS point clouds from the DD geometries, visualized in white and gray. The yellow points
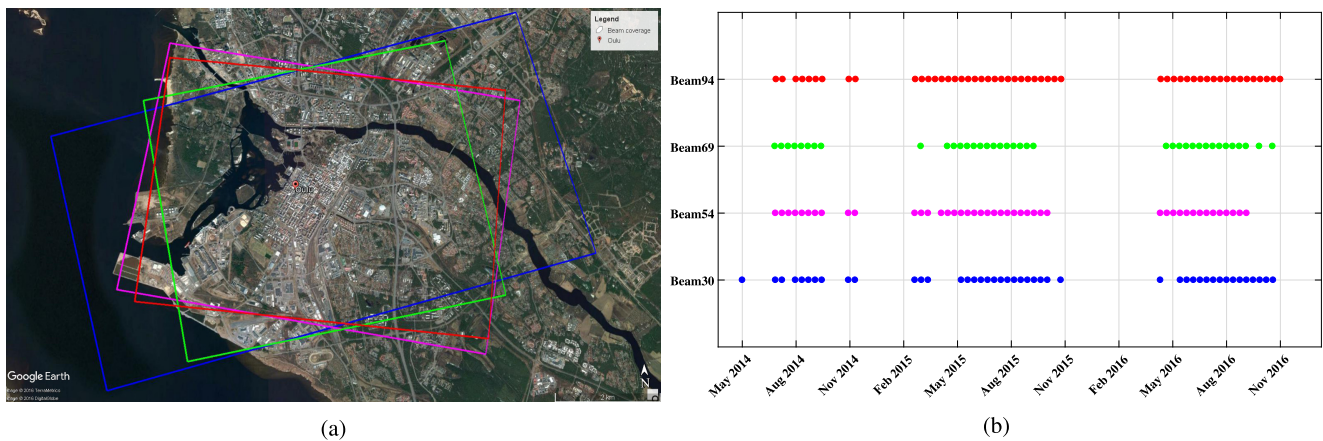
(a) (b)

Fig. 8. (a) Mean scene coverage of the TS-X images overlaid on the optical image of Oulu taken from Google Earth. (b) Acquisition time plot of the TS-X images of Oulu.
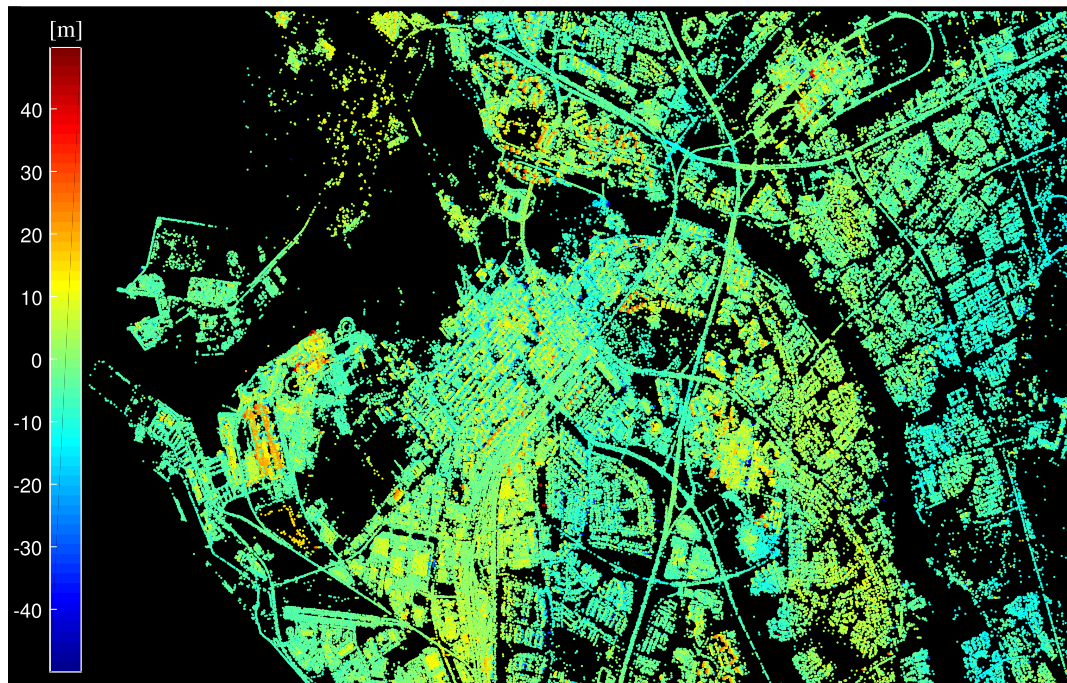

Fig. 9. PS elevation map obtained from PSI processing of an ascending stack of Oulu (beam 30). The total number of scatterers is approximately 540 000 after selecting only the PSs with posterior coherence values equal or higher than 0.7.

represent the identified PS pairs from the fusion algorithm. The total number of the correspondences is approximately 32 000 and the Euclidean distances between the matched PSs vary from 1.5 to 5 m. In order to reduce the number of PS correspondences to the ones with higher quality and closer distance, as well as to preserve the homogeneity of the distribution, a regular grid was imposed on the point clouds. Inside the 10 m × 10 m grid, the PS pairs that were closer together and had lower ADI values were selected to reduce the number of pairs from 32 000 to 10 000. The comparison between PS pairs before and after reduction can be seen in Fig. 10(b) and (c) (zoomed-in view), respectively. The same procedure depicted in Fig. 10 was also carried out for the

SAR images from the AA geometries and close to 9500 PS correspondences were detected. The results then were radar-coded for both geometry configurations.

The PS candidates to be localized from the cross-heading geometries, either ascending–descending (AD) configuration or quad geometry (ADAD) configuration, were selected based on the detection of bright points along the roads using the road network data as explained in Section III-C. The road network data were first radar-coded on the master scenes of all the four geometries as seen on the mean intensity images of each stack in Fig. 11. A circular neighborhood with a radius of 70 pixels was then considered around each road network node within which the ADI was
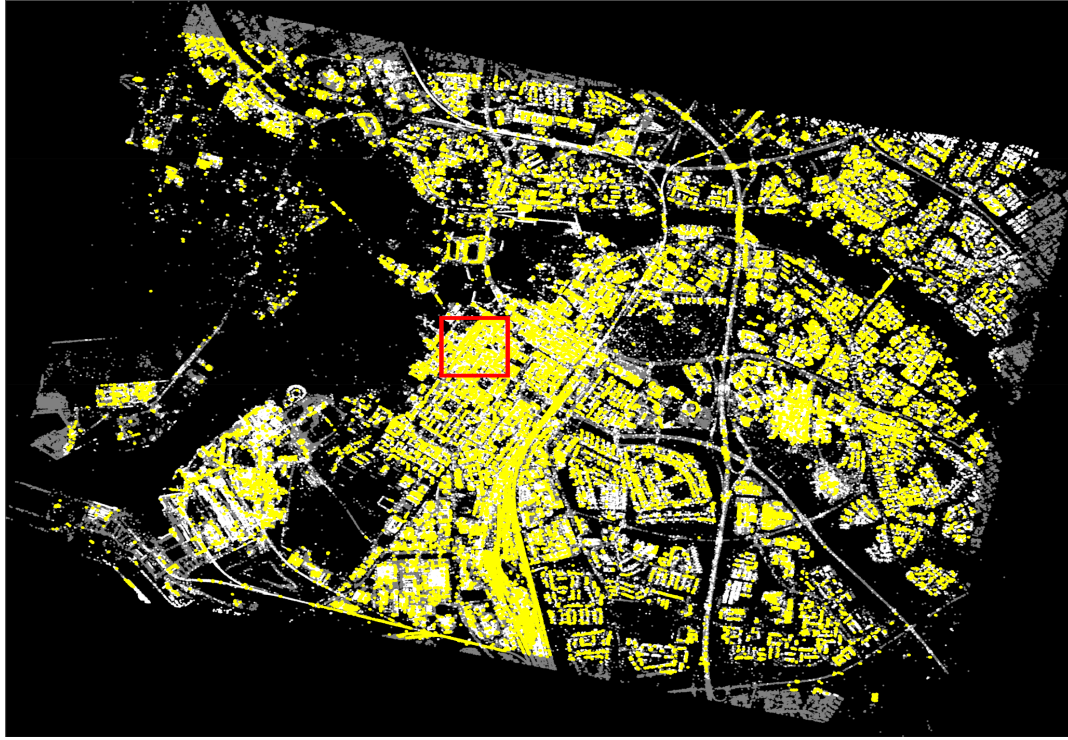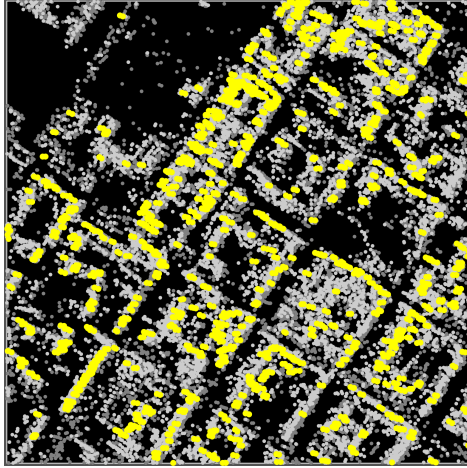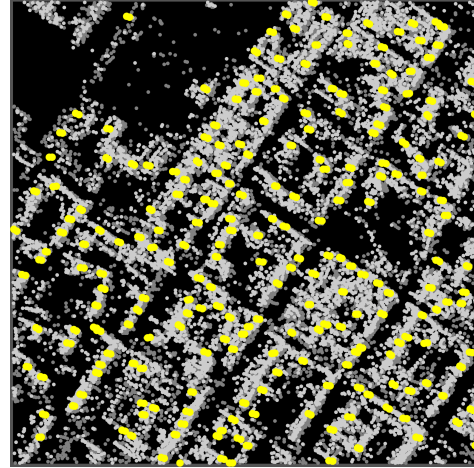
Fig. 10. Depiction of PS correspondence detection from SAR images of same-heading orbit tracks of Oulu using the first step of the multitrack PS fusion algorithm proposed in [22]. (a) Geocoded PS point clouds of beams 54 and 94 in a DD configuration as white and gray points as well as the detected PS correspondences in yellow. (b) and (c) Zoomed-in view of (a) marked by the red rectangle before and after imposing a 10-m grid in which the pairs with the closest distance are chosen.

evaluated for all pixels for all the four stacks. The neighborhood is chosen based on a rough knowledge on the maximum width of highways in Oulu ($\approx$35 m) and was adapted to the SAR data by taking into account the pixel spacing in range and azimuth direction and the oversampling factor used in the processing. After selecting the pixel with the lowest ADI in each neighborhood, a further threshold of 0.25 on Amplitude Dispersion Index (ADI) values, typically used in PSI processing [31], selects the stable bright points from each stack. After geocoding, the PSs from different stacks which are closer than

a threshold of 3 m are chosen as the final stereo candidates and are subsequently radar-coded on all the SAR images. The distance threshold depends on the geometry configuration from which the user is interested to localize the targets. If same-heading geometries are considered, the value should be lower than 3 m in order to ensure correct PS correspondence detection. An example of PS candidates visible from ADAD configuration in Oulu is given in Fig. 12. The candidates are all assumed to be bases of lamp poles and can be seen as bright points inside the green circles. The explained procedure
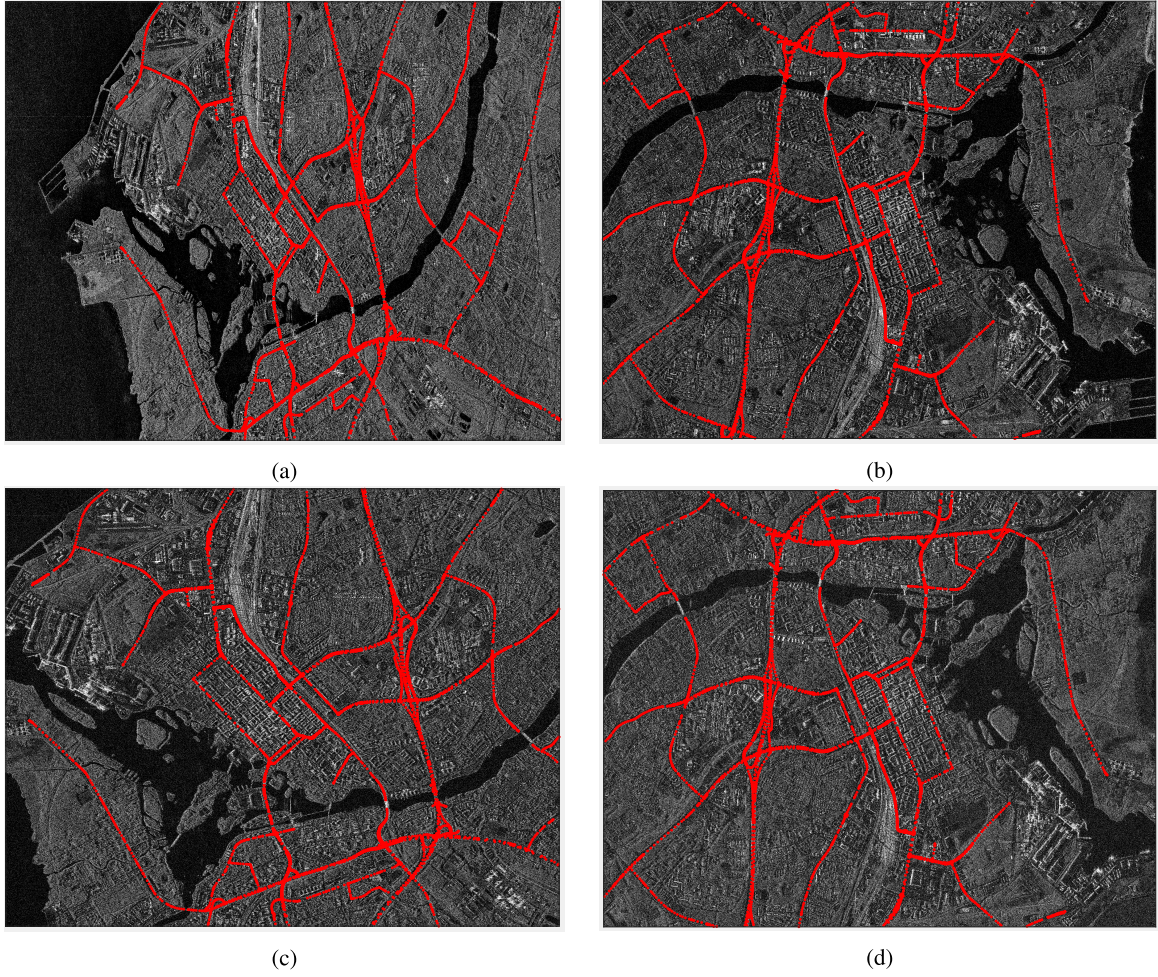
Fig. 11. Projection of the road network data of Oulu onto the master scene of (a) beam 30, (b) beam 54, (c) beam 69, and (d) beam 94. The road data is represented by red points and is the basis for detection of identical PS candidates from cross-heading orbit geometries.

produced 107 and 52 initial candidates from ADAD and AD geometry configurations, respectively. The quantity is lower in the latter because of the strict distance threshold of 1.5 m imposed on coordinate differences of PSs visible in different stacks. The threshold value is chosen empirically based on the histogram of minimum Euclidean distances evaluated between the PS pairs.

The PTA was performed on all the detected PS candidates in all the SAR images in which the candidate was visible. For each candidate a time series of σ_ϕ was evaluated using (10). Fig. 13 demonstrates the initial outlier removal on two PS candidates based on the adjusted boxplot method explained in Section IV-C. The σ_ϕ values of each scatterer are sorted in time in Fig. 13(a) and (b). The distribution of σ_ϕ is right-skewed in both cases as can be seen in Fig. 13(c) and (d). From the distribution plots, one can detect the samples which are not connected to the tail of the distribution and mark them as potential outliers. This process is done automatically and without the need for manual intervention using the interval defined in (12). In Fig. 13(e) and (f), the detected outliers are marked with red rectangles. For all the candidates, the mentioned procedure was carried out.

The atmospheric corrections were carried out using global ionospheric maps and the global tropospheric zenith path delays provided with the Vienna mapping function [40], since there was no access to the Oulu GNSS receiver at the time of the study. The geodynamic effects were fully considered according to the IERS conventions and all the effects were removed from the PS timings.

The final positioning was carried out by stereo SAR with all the mentioned outlier removal steps in Section IV-E. The last criterion, which includes the removal of PSs based on S_{az} values higher than 20 cm, reduces the amount of total PSs to only those points which can be considered ideal stereo candidates. The averaged quality of the estimated 3-D coordinates for all the remaining high-quality PSs are reported in Table IV. The scatterers are categorized based on the geometry configuration that is used for their positioning. The results are all expressed in centimeter and are defined in the local east, north, and vertical coordinates. The standard deviations are all defined within 95% confidence interval. From Table IV, it is seen that the averaged precisions are smaller than 2 dm for all the cases. As it was expected, the localization quality boosts as the difference in the viewing geometries becomes
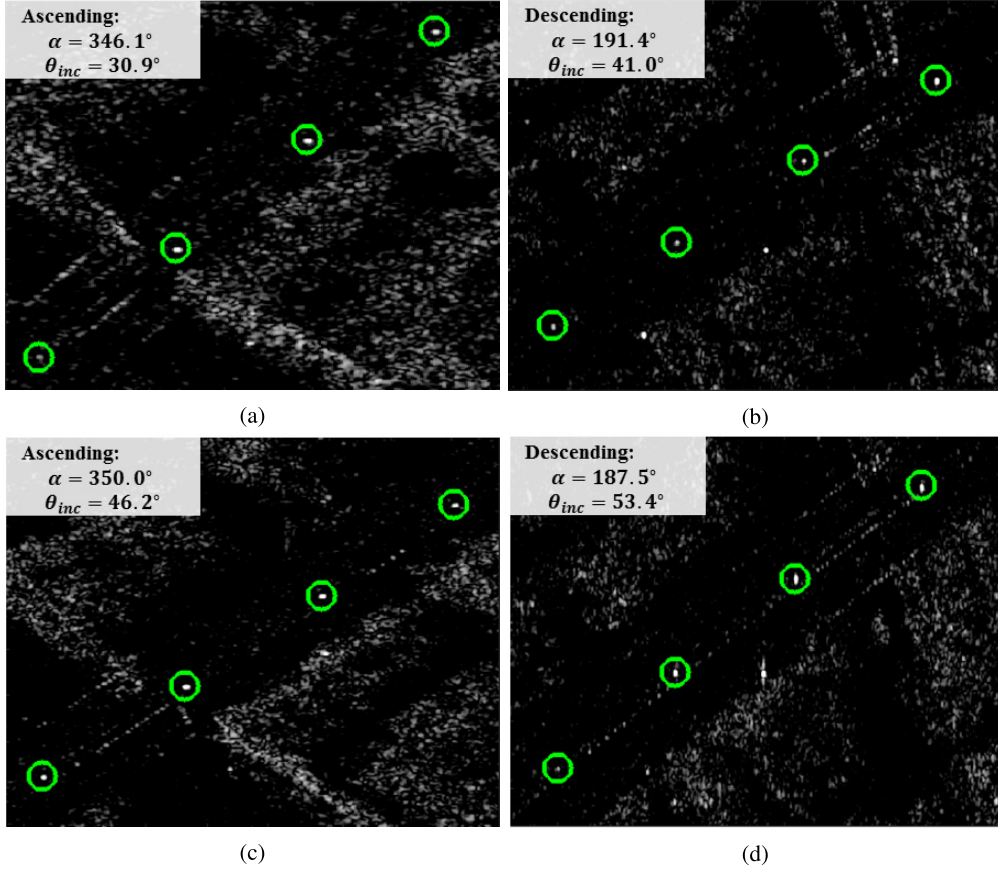
Fig. 12. PS correspondence detection from ADAD geometry configuration for (a) beam 30, (b) beam 54, (c) beam 69, and (d) beam 94. For each subfigure, the respective averaged heading angle (α) and the averaged incidence angle (θ_{inc}) are stated.

TABLE IV

AVERAGED STATISTICS BASED ON THE LEAST SQUARES ESTIMATED 3-D COORDINATE STANDARD DEVIATIONS IN OULU. THE LETTERS A AND D STAND FOR ASCENDING AND DESCENDING GEOMETRIES, RESPECTIVELY. THE SAMPLE MEAN AND STANDARD DEVIATION ARE DENOTED BY μ AND σ, AND $S_{\mathrm{[ENH]}}$ REPRESENT THE LOCAL COORDINATES STANDARD DEVIATIONS WITHIN 95% CONFIDENCE LEVEL

Geometry	Nr. Scatterers	μ_{s_E} [cm]	μ_{s_N} [cm]	μ_{s_H} [cm]	σ_{s_E} [cm]	σ_{s_N} [cm]	σ_{s_H} [cm]
AA	565	17.73	5.04	15.87	11.98	2.63	11.09
DD	1417	15.08	3.80	16.71	10.38	2.10	11.30
AD	24	2.26	2.50	1.75	0.99	1.11	0.83
ADAD	43	1.17	1.40	1.12	0.42	0.55	0.37

larger which is the case when changing from AA or DD to the AD and ADAD geometry configurations. It is also evident that for cross-heading geometries, the retrieval of the height component is the most precise one as for the same-heading cases, the precision in the north component is the highest. Therefore, in general, the localization of targets from cross-heading tracks is desirable. The only remaining concern regarding localization using cross-heading tracks is the diameter of the lamp poles which may worsen the accuracy in the east coordinate component. This bias can be estimated and removed if the scatterer is also visible from same-heading tracks, which is usually the case.

The distribution of the total 2049 generated GCPs is visualized on the optical image of Oulu in Google Earth in Fig. 14. The scatterers are color-coded based on the underlying geom-etry configuration used for their localization. It is seen that almost the entire area of Oulu is covered with the generated GCPs. The ones from the same-heading geometries cover the built areas, while the ones from cross-heading configuration include the base of lamp poles, street lights, and traffic lights.

VI. CONCLUSION

In this paper, we described a processing chain for automatic detection and positioning of opportunistic PSs, which are visible from multiaspect TS-X images. This paves the way for the generation of GCPs from SAR data only.

Three algorithms have been recommended for identical PS detection which are different in terms of number of generated GCPs and the subsequent positioning precision. The method based on the PSI fusion algorithm is able to provide point
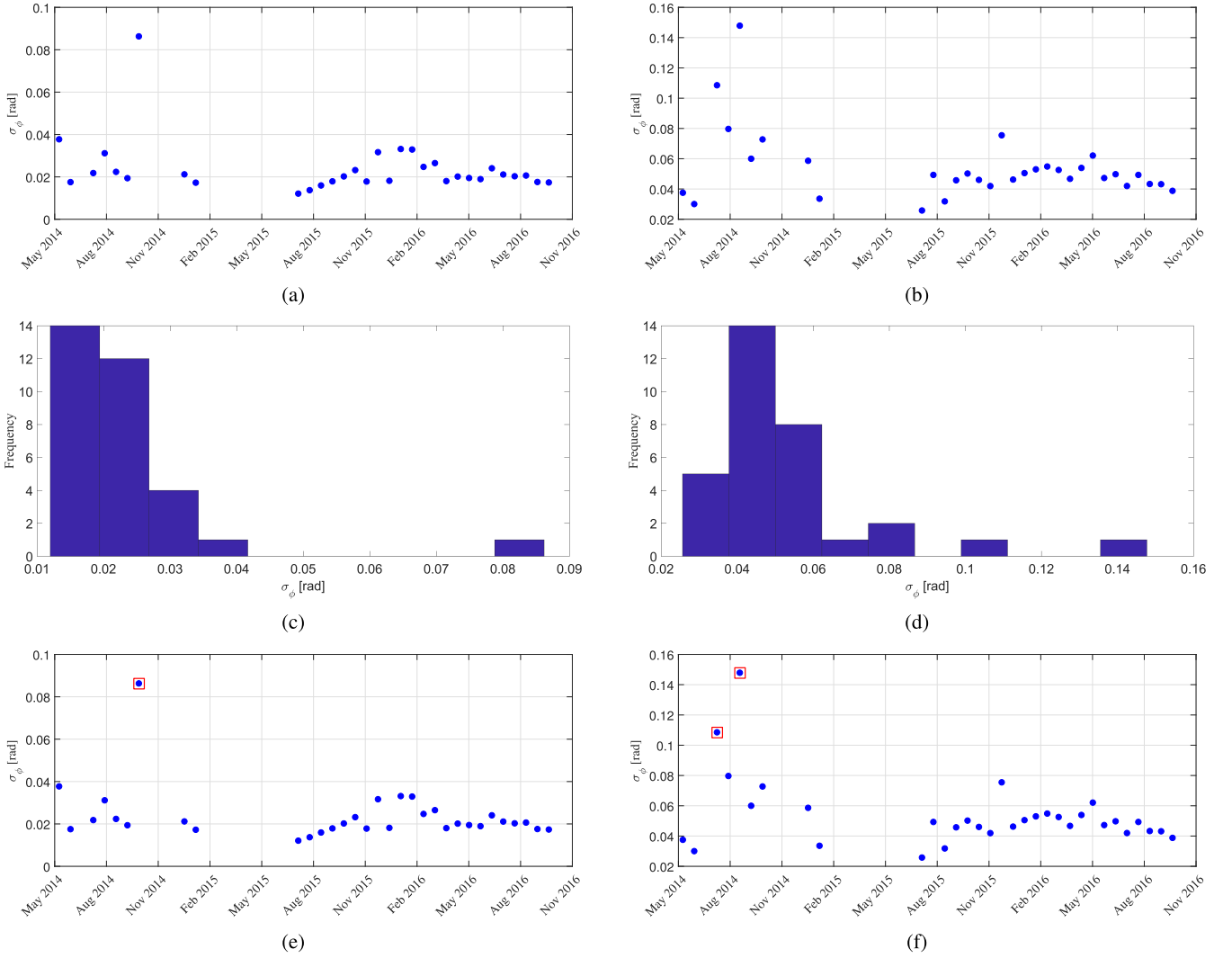
Fig. 13. Example of initial outlier removal based on phase noise time series of the detected candidates. In (a) and (b), the sorted σ_ϕ values of two random PSs are shown with their distributions depicted in (c) and (d). The outliers are identified and removed automatically based on the distributions and are shown in red rectangles in (e) and (f).

correspondences even on buildings and areas with complex scattering mechanisms in SAR images. Therefore, a large number of potential PS pairs can be obtained which normally cover the entire scene. The downside of the method is that SAR image stacks are required for which a complete PSI processing has to be performed separately before being able to find the PS correspondences. Furthermore, the method is usually applicable only for same-heading PS point clouds. Consequently in terms of localization precision with the subsequent stereo SAR, the relative error in the cross range is larger than the error for range and azimuth components. Another disadvantage is that many of the initial PS correspondences cannot be considered useful candidates for stereo SAR and have to be eliminated later in the processing, because the registration is only performed within the limits of the PSI 3-D localization quality for which PS pairs with distances of up to 5 m are detected. The detection algorithm based on the external optical data is quite straightforward to implement and provides

identical scatterers that are visible from cross-heading orbits leading to higher localization precision. The disadvantage of the method is that for reliable detection of lamp poles, the spatial resolution of the optical image should be in the subdecimeter regime. Moreover, the method is highly prone to detecting other linear structures as shadows of lamp poles and, therefore, more sophisticated object detection algorithms are recommendable. The method based on vector road network data, similar to the optical method, provides candidates which are suitable to be localized from cross-heading geometries. Also, the external data is freely accessible for most locations. The disadvantage of the method is that a co-registration on one master has to be carried out for each stack and the amplitude data must be calibrated.

It has been shown that the GCP generation processing chain is quite flexible as it allows the user to constraint the number and the quality of the candidate GCPs either from the start of the procedure, by selecting different distance thresholds for
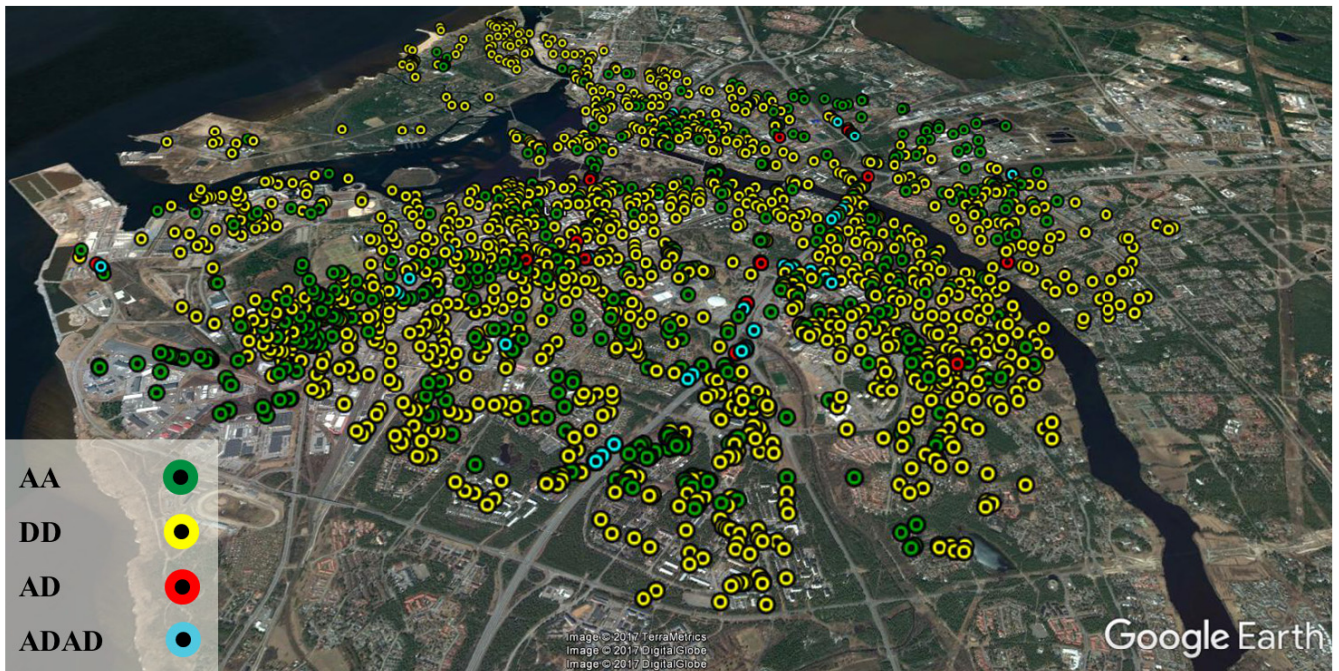
Fig. 14. Total number of 2049 GCPs in Oulu color-coded based on the geometry configuration used for their positioning. The underlying optical image is taken from Google Earth.

detection or trimming the data based on estimated phase noise time series, or at the final step of the procedure based on the posterior quality indicators obtained from stereo SAR.

By applying the algorithm to two test sites, it has been demonstrated that it is capable of positioning natural PSs with precision values ranging from 2 and 4–5 cm, for cross-heading AD and ADAD configurations, respectively, to approximately 20 cm for candidates extracted from same-heading geometries. As it was expected, the difference in the viewing geometries of the observed PS has the highest impact on the final localization precision followed by the number of acquisitions used in stereo SAR processing, the SCR of the target, and the quality of the external error corrections. Furthermore, as a preliminary cross comparison, the estimated ellipsoidal height of the retrieved candidates in Berlin was compared with the corresponding height of a LiDAR data which reported an average bias of approximately 13 cm.

The produced absolute GCPs have ample of applications in geodesy and absolute mapping. These points may substitute the conventional GCPs that are required for georeferencing of satellite imagery, which are usually surveyed at the field by GNSS. They can be further integrated as absolute reference points into multipass Interferometric SAR (InSAR) techniques. Furthermore, they can be used to detect long-term ground motions with small magnitudes and low spatial frequency, which are invisible to phase-based InSAR methods.

The future work will focus on smart preselection of the PS candidates by including the information obtained from PTA, using integrated sidelobe ratio, to robustly remove the PS candidates which are located too close to each other. Furthermore, the stereo SAR processing could benefit from weighting the initial timing observations of the PSs based on their respective

SCR or ADI values and also can be carried out with robust parameter estimation schemes such as M-estimator. Moreover, it is desirable to carry out GNSS measurements at selected test sites to be able to correctly validate the absolute accuracy of the generated GCPs. Finally, it is important to note that the proposed methodology is tailored to detection and absolute localization of GCPs in an urban area, where a large number of PSs are available. In applications where the investigated scene includes mainly a nonurban area, it is recommendable to employ artificial PSs such as corner reflectors or active transponders.

REFERENCES

[1] M. Eineder, C. Minet, P. Steigenberger, X. Y. Cong, and T. Fritz, "Imaging geodesy—Toward centimeter-level ranging accuracy with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 661–671, Feb. 2011.

[2] X. Cong, U. Balss, M. Eineder, and T. Fritz, "Imaging geodesy—Centimeter-level ranging accuracy with TerraSAR-X: An update," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 5, pp. 948–952, Sep. 2012.

[3] C. Gisinger *et al.*, "Precise three-dimensional stereo localization of corner reflectors and persistent scatterers with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1782–1802, Apr. 2015.

[4] S. Hackel, O. Montenbruck, P. Steigenberger, U. Balss, C. Gisinger, and M. Eineder, "Model improvements and validation of TerraSAR-X precise orbit determination," *J. Geodesy*, vol. 91, no. 5, pp. 547–562, Dec. 2016.

[5] U. Balss *et al.*, "High precision measurement on the absolute localization accuracy of TerraSAR-X," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2012, pp. 1625–1628.

[6] U. Balss, C. Gisinger, X. Y. Cong, R. Brcic, S. Hackel, and M. Eineder, "Precise measurements on the absolute localization accuracy of TerraSAR-X on the base of far-distributed test sites," in *Proc. 10th Eur. Conf. Synth. Aperture Radar (EUSAR)*, Jun. 2014, pp. 1–4.

[7] U. Balss, H. Breit, T. Fritz, U. Steinbrecher, C. Gisinger, and M. Eineder, "Analysis of internal timings and clock rates of TerraSAR-X," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2014, pp. 2671–2674.

[8] I. G. Cumming and F. H. Wong, *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation [With CDROM]* (Artech House Remote Sensing Library). Boston, MA, USA: Artech House, 2005.

[9] H. Breit, E. Börner, J. Mittermayer, J. Holzner, and M. Eineder, "The TerraSAR-X multi-mode SAR processor—Algorithms and design," in *Proc. 5th Eur. Conf. Synth. Aperture Radar (EUSAR)*, 2004, pp. 501–503.

[10] C. Gisinger, "Atmospheric corrections for TerraSAR-X derived from GNSS observations," M.S. thesis, Inst. Astron. Phys. Geodesy, Tech. Univ. München, München, Germany, 2012.

[11] G. Petit and B. Luzum, Eds., "Frankfurt am Main: Verlag des Bundesamts für Kartographie und Geodäsie," IERS Conventions, Frankfurt, Germany, Appl. Note 36, 2010, p. 179.

[12] M. Eineder *et al.*, "SAR imaging geodesy–recent results for TerraSAR-X and for Sentinel-1," in *Proc. 10th Int. Workshop Adv. Sci. Appl. SAR Interferometry Sentinel-1 InSAR FRINGE*, Jun. 2017.

[13] C. Gisinger *et al.*, "Absolute 4-D positioning of persistent scatterers with TerraSAR-X by applying geodetic stereo SAR," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2015, pp. 2991–2994.

[14] X. X. Zhu, S. Montazeri, C. Gisinger, R. F. Hanssen, and R. Bamler, "Geodetic SAR tomography," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 18–35, Jan. 2016.

[15] S. Montazeri, X. X. Zhu, M. Eineder, and R. Bamler, "Three-dimensional deformation monitoring of urban infrastructure by tomographic SAR using multitrack TerraSAR-X data stacks," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 6868–6878, Dec. 2016.

[16] U. Balss, H. Runge, S. Suchandt, and X. Y. Cong, "Automated extraction of 3-D ground control points from SAR images—An upcoming novel data product," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 5023–5026.

[17] D. O. Nitti *et al.*, "Automatic GCP extraction with high resolution COSMO-SkyMed products," in *SAR Image Analysis, Modeling, and Techniques XVI*, vol. 10003, C. Notarnicola, S. Paloscia, N. Pierdicca, and E. Mitchard, Eds. Bellingham, WA, USA: SPIE, Oct. 2016, p. 1000302.

[18] P. Dheenathayalan, D. Small, A. Schubert, and R. F. Hanssen, "High-precision positioning of radar scatterers," *J. Geodesy*, vol. 90, no. 5, pp. 403–422, May 2016.

[19] S. Suri, P. Schwind, J. Uhl, and P. Reinartz, "Modifications in the SIFT operator for effective SAR image matching," *Int. J. Image Data Fusion*, vol. 1, no. 3, pp. 243–256, Sep. 2010.

[20] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004.

[21] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SAR-SIFT: A sift-like algorithm for SAR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 1, pp. 453–466, Jan. 2015.

[22] S. M. Gernhardt, X. Y. Cong, M. Eineder, S. Hinz, and R. Bamler, "Geometrical fusion of multitrack PS point clouds," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 1, pp. 38–42, Jan. 2012.

[23] S. M. Gernhardt, "High precision 3D localization and motion analysis of persistent scatterers using meter-resolution radar satellite data," M.S. thesis, Chair Remote Sens. Technol., Tech. Univ. München, München, Germany, 2012.

[24] B. M. Kampes, *Radar Interferometry: Persistent Scatterer Technique*. Dordrecht, The Netherlands: Springer, 2006.

[25] T. Fritz, *et al.*, "TerraSAR-X ground segment basic product specification document," German Aerosp. Center (DLR), Oberpfaffenhofen, Germany, Tech. Rep. TX-GS-DD-3302, 2008, no. 1.5.

[26] S. Montazeri *et al.*, "SAR ground control point identification with the aid of high resolution optical data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2016, pp. 3205–3208.

[27] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 2nd ed. Upper Saddle River, NJ, USA: Prentice-Hall, 2002.

[28] K. Briechle and U. D. Hanebeck, "Template matching using fast normalized cross correlation," in *Optical Pattern Recognition XII*, vol. 4387, D. P. Casasent and T.-H. Chao, Eds. Bellingham, WA, USA: SPIE, Mar. 2001, pp. 95–102.

[29] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.

[30] P. J. Besl and D. N. McKay, "A method for registration of 3-D shapes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 239–256, Feb. 1992.

[31] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 8–20, Jan. 2001.

[32] N. Adam, B. M. Kampes, M. Eineder, J. Worawattanamateekul, and M. Kircher, "The development of a scientific permanent scatterer system," in *Proc. Joint Workshop-High Resolution Mapping Space (ISPRS/EARSeL)*, Hannover, Germany, 2003, pp. 1–6.

[33] N. Adam, F. R. Gonzalez, A. Parizzi, and R. Brcic, "Wide area persistent scatterer interferometry: Current developments, algorithms and examples," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2013, pp. 1857–1860.

[34] N. Adam, B. M. Kampes, and M. Eineder, "The development of a scientific persistent scatterer system: Modifications for mixed ERS/ENVISAT time series," in *Proc. Envisat ERS Symp.*, Salzburg, Austria, 2004, pp. 1–9.

[35] M. Hubert and E. Vandervieren, "An adjusted boxplot for skewed distributions," *Comput. Stat. Data Anal.*, vol. 52, no. 12, pp. 5186–5201, Aug. 2008.

[36] J. W. Tukey, "Exploratory data analysis," *Addison-Wesley Series in Behavioral Science*. Reading, MA, USA: Addison-Wesley, 1977.

[37] G. Brys, M. Hubert, and A. Struyf, "A robust measure of skewness," *J. Comput. Graph. Stat.*, vol. 13, no. 4, pp. 996–1017, Dec. 2004.

[38] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 328–341, Feb. 2008.

[39] (Jan. 2017). *Finnish Transport Agency*. [Online]. Available: http://www.liikennevirasto.fi/web/en

[40] J. Boehm, B. Werl, and H. Schuh, "Troposphere mapping functions for GPS and very long baseline interferometry from European Centre for Medium-Range Weather Forecasts operational analysis data," *J. Geophys. Res., Solid Earth*, vol. 111, no. B2, pp. 1–9, Feb. 2006.

Sina Montazeri received the B.Sc. degree in geodesy and surveying engineering from the University of Isfahan, Isfahan, Iran, and the M.Sc. degree in geomatics from the Delft University of Technology (TU Delft), Delft, The Netherlands, in 2011 and 2014, respectively. He is currently pursuing the Ph.D. degree with the German Aerospace Center (DLR), Remote Sensing Technology Institute (IMF), Oberpfaffenhofen, Germany.

In 2012, he spent two weeks with the Laboratoire des Sciences de l'Image, de l'Informatique et de la Télédétection, University of Strasbourg, Strasbourg, France, as a Research Team Member. From 2013 to 2014, he was a Research Assistant with DLR, IMF, where he was involved in absolute localization of point clouds obtained from SAR tomography. His research interests include advanced InSAR techniques for deformation monitoring of urban infrastructure, SAR imaging geodesy, SAR positioning, image and signal processing relevant to radar imagery, and incorporation of geodetic estimation framework into multipass InSAR.

Mr. Montazeri was one of the recipients of the DLR Science Award and the IEEE Geoscience and Remote Sensing Society Transactions Prize Paper Award, in 2016 and 2017, respectively.

Christoph Gisinger received the Diploma degree in geodesy from Technische Universität Graz, Graz, Austria, and the M.Sc. degree in earth oriented space science and technologies from Technische Universität München, Munich, Germany. He is currently pursuing the Ph.D. degree in joint processing of synthetic aperture radar (SAR) and global navigation satellite system to retrieve geophysical signals.

From 2012 to 2017, he was with the Chair of Astronomical and Physical Geodesy, Technische Universität München, where he was involved in SAR imaging geodesy and the corrections of absolute SAR observations. He is currently a Research Scientist with the Remote Sensing Technology Institute, German Aerospace Center (DLR), Oberpfaffenhofen, Germany, where he is involved in the geodetic SAR processing of the missions TerraSAR-X and Sentinel-1. His research interests include stereo SAR, correction modeling for absolute SAR observations, and radargrammetric image processing with a focus on fusing SAR with geodetic observations to make use of SAR in geodetic applications.

Michael Eineder (SM'10–F'16) received the Diploma degree in electrical engineering from Technische Universität München (TUM), Munich, Germany, in 1990, and the Dr.rer.nat. degree from the University of Innsbruck, Innsbruck, Austria, in 2004.

He joined the German Aerospace Center (DLR), Oberpfaffenhofen, Germany, in 1990, where he was involved in a variety of international missions including SIR-C/X-SAR and SRTM/X-SAR in cooperation with NASA, ERS-1 (ESA), TerraSAR-X, and TanDEM-X (Germany). Since 2013, he has been a TUM Honorary Professor. He is currently the Head of the SAR Signal Processing Department, Remote Sensing Technology Institute, DLR. His responsibilities encompass the development of SAR and InSAR processing systems for current and future radar missions. He is currently a Part-Time Lecturer of remote sensing with TUM. His research interests include the design of SAR missions for geophysical applications and imaging geodesy, an absolute positioning technique exploiting high-resolution SAR.

Xiao Xiang Zhu (S'10–M'12–SM'14) received the bachelor's degree in space engineering from the National University of Defense Technology, Changsha, China, in 2006, and the M.Sc., Dr.-Ing., and Habilitation degrees in signal processing from Technische Universität München (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

She was a Guest Scientist or Visiting Professor with the Italian National Research Council, Naples, Italy; Fudan University, Shanghai, China; the University of Tokyo, Tokyo, Japan; and the University of California, Los Angeles, CA, USA, in 2009, 2014, 2015, and 2016, respectively. Since 2011, she has been the Head of the Signal Analysis Team, Remote Sensing Technology Institute, DLR, Oberpfaffenhofen, Germany, and the Head of the Helmholtz Young Investigator Group "SiPEO," DLR and TUM, since 2013. Since 2015, she has been a Professor of signal processing in earth observation with TUM and DLR. Her research interests include advanced InSAR techniques such as high-dimensional tomographic SAR imaging and SqueeSAR; computer vision in remote sensing including object reconstruction and multidimensional data visualization; big data analysis in remote sensing; and modern signal processing, including innovative algorithms such as sparse reconstruction, nonlocal means filter, and robust estimation and deep learning, with applications in the field of remote sensing such as multi/hyperspectral image analysis.

Dr. Zhu is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Article

Geocoding Error Correction for InSAR Point Clouds

Sina Montazeri [1] iD **, Fernando Rodríguez González** [1] **and Xiao Xiang Zhu** [1,2,*] iD

[1] Remote Sensing Technology Institute, German Aerospace Center, 82234 Weßling, Germany; sina.montazeri@dlr.de (S.M.); fernando.rodriguezgonzalez@dlr.de (F.R.G.)
[2] Signal Processing in Earth Observation, Technical University of Munich, 80333 Munich, Germany
[*] Correspondence: xiaoxiang.zhu@dlr.de; Tel.: +49-8153-28-3531

check for updates

Abstract: Persistent Scatterer Interferometry (PSI) is an advanced multitemporal InSAR technique that is capable of retrieving the 3D coordinates and the underlying deformation of time-coherent scatterers. Various factors degrade the localization accuracy of PSI point clouds in the geocoding process, which causes problems for interpretation of deformation results and also making it difficult for the point clouds to be compared with or integrated into data from other sensors. In this study, we employ the *SAR imaging geodesy* method to perform geodetic corrections on SAR timing observations and thus improve the positioning accuracy in the horizontal components. We further utilize *geodetic stereo SAR* to extract large number of highly precise ground control points (GCP) from SAR images, in order to compensate for the unknown height offset of the PSI point cloud. We demonstrate the applicability of the approach using TerraSAR-X high resolution spotlight images over the city of Berlin, Germany. The corrected results are compared with a reference LiDAR point cloud of Berlin, which confirms the improvement in the geocoding accuracy.

1. Introduction

In the past two decades, Interferometric Synthetic Aperture Radar (InSAR) and its multitemporal extensions have proved their ability for continuous mapping and deformation monitoring of the surface of the Earth. Among the existing InSAR approaches, Persistent Scatterer Interferometry (PSI) is nowadays considered one of the most operational techniques for wide area processing. PSI is a single-master InSAR technique that extracts phase-stable scatterers, the so-called Persistent Scatterers (PS), from a stack of SAR images and retrieves their heights along with their deformation history through the exploitation of their interferometric phase [1,2]. Its usage of multiple images and the restriction of the estimation of parameters to only PS overcomes the main limitations of InSAR, namely atmospheric disturbances and geometric and temporal decorrelation [3]. The PSI technique is highly effective in urban areas because of the availability of a high density of PS. In particular, coupling the technique with high resolution images, such as the ones from the TerraSAR-X spotlight mode, produces on average between 40,000 to 100,000 PS per square kilometer [4,5], which allows for detailed infrastructure monitoring.

Similar to other InSAR approaches, PSI is a *relative* technique. This means that the height and the deformation estimates of all PS are relative with respect to a reference point, which is chosen during the PSI processing [2]. For most of the cases, no information about the exact height and the stability of the reference point is available. Therefore, PSI assumes a height value extracted from the available external Digital Elevation Model (DEM) for this point and considers a height update equal to zero. Evidently, if the reference point is not stable in time and/or its true absolute height value

is not equal to the corresponding DEM height, due to an erroneous DEM, the deformation and/or the height estimates of all PS will be biased. The height of PS points is an influential factor in the *geocoding* process. Geocoding assigns each PS to its 3D position in a geodetic reference system using the orbit parameters of the master acquisition and the estimated interferometric height [6–8]. Therefore, if the height estimates are biased, then the 3D coordinates of the PS will have offsets with respect to their true positions. Furthermore, the timing information of the master orbit, which is also directly involved in the geocoding process, could be erroneous due to atmospheric and geodynamic effects [9], which again hampers the localization accuracy of the PS. Assigning PS to wrong locations can lead to misrepresentation of the origin of the deformation and thus degrades the interpretation of the signal. This is in particular important for monitoring small-scale deformations in urban areas using high resolution SAR products for which the detailed structure of buildings and other objects are captured. Additionally, wrongly geocoded PSI point clouds can be neither integrated into nor compared with data from other sensors.

Several studies have been carried out in order to improve the geocoding accuracy of InSAR point clouds. In [10], the authors used an external Digital Surface Model (DSM) to evaluate a height shift between the PSI point cloud and the DSM in the vicinity of the point cloud. In [11], in addition to the usage of the DSM similar to the approach in [10], the authors installed a corner reflector in one scene and measured its position by Global Navigation Satellite System (GNSS). They radar-coded the 3D coordinates of the corner reflector and calculated a shift in range and azimuth with respect to the position of the object in the SAR image. Eventually, they corrected the entire point cloud using the three shifts: in range, azimuth and cross-range. In an alternative approach, a manual shift based on an overlay of a PSI point cloud on an ortho-rectified aerial image was used to compensate for the geocoding errors in [12]. Apart from the mentioned methods which all rely on external 2D or 3D data to correct for the errors of geocoding, two geometrical point cloud fusion approaches have been proposed in [13], which effectively compensates for the DEM error of the reference points in each of the multiaspect point clouds. The first method, which is further elaborated in [14], is based on a least squares PS matching scheme while the second approach, described in detail in [15], employs the iterative closest point (ICP) algorithm [16] to perform multiaspect point cloud matching. The first geodetic approach for reducing geocoding errors has been described in [17], in which the concepts of SAR imaging geodesy [9] and stereo SAR [18] were combined with urban SAR tomography [19] in order to absolutely localize a large number of scatterers by absolutely geolocating the manually extracted reference point.

Motivated by our work in [17] and the recently developed frameworks which automatically extract and localize Ground Control Points (GCP) from SAR images from different orbit tracks [20,21], we propose a SAR-based method for improving the geocoding accuracy of PSI point clouds. We will show that by employing SAR GCPs and applying geodetic corrections in the entire scene, one can correct for the geocoding errors as a post-processing step after the PSI processing. The applicability of the method is demonstrated in an urban area using high resolution TerraSAR-X spotlight images of Berlin, Germany. Furthermore, the results are compared with an external aerial Light Detection and Ranging (LiDAR) data, demonstrating the improvement in the geocoding accuracy of the PSI point clouds.

The remainder of the paper is organized as follows. Section 2 reviews the geocoding procedure of InSAR products and their error sources. Section 3 describes the methods used for improving the geocoding accuracy of PSI point clouds. The test site and the dataset are introduced in Section 4. The results and discussion are reported in Section 5 and the conclusions are drawn in Section 6.

2. InSAR Geocoding: Principle and Error Sources

Geocoding is performed as the final step of any InSAR processing in order to report the results in a common geodetic reference system. It is indeed a coordinate transformation from radar datum to an earth-fixed geodetic datum [7]. For each pixel in the SAR image, geocoding is carried out by iterative

solving of the Doppler-Range-Ellipsoid equations to retrieve the 3D Cartesian coordinates of the target $\vec{T} = (X_T, Y_T, Z_T)$ [6,8,22]:

$$\left\| \dot{\vec{S}}(t_{az}) \cdot (\vec{S}(t_{az}) - \vec{T}) \right\| = 0, \tag{1}$$

$$\left\| \vec{S}(t_{az}) - \vec{T} \right\| - c \cdot \frac{\tau_{rg}}{2} = 0, \tag{2}$$

$$\frac{X_T{}^2}{(a + H_T)^2} + \frac{Y_T{}^2}{(a + H_T)^2} + \frac{Z_T{}^2}{(b + H_T)^2} - 1 = 0, \tag{3}$$

where $\vec{S}$ and $\dot{\vec{S}}$ are the satellite state vector and its velocity vector, respectively. They are both functions of azimuth time t_{az}, which is the time of the closest approach, when the two-way travel time of the radar chirp τ_{rg} is recorded [23]. The first equation dictates the zero-Doppler condition. The second equation means that the geometric distance between $\vec{T}$ and $\vec{S}$ should be equal to τ_{rg} multiplied by the speed of light in vacuum c divided by two. The result from Equations (1) and (2) is intersected with a reference ellipsoid with semi-major axis a and semi-minor axis b and with an estimated height above the ellipsoid H_T that is obtained from InSAR. The final coordinates are calculated in the Cartesian coordinate system. For convenience of interpretation, these coordinates are usually converted to the geographic coordinates (Φ,Λ,H) or to the Universal Transverse Mercator (UTM) map projection (E,N,H), where Φ and Λ are the geographic latitude and longitude while E and N denote the UTM Easting and Northing and H representing the height above the, global or local, reference ellipsoid.

Looking closely at Equations (1)–(3), the factors which influence the final position of the target after geocoding are the orbit accuracy, the radar timings (t_{az}, τ_{rg}) and the estimated height of the target H_T. For modern SAR sensors such as TerraSAR-X, the orbit accuracy is in the centimeter regime [24,25] and therefore its effect on localization inaccuracy of targets is negligible. In the following, the other factors, namely t_{az}, τ_{rg} and H_T, and their error sources are briefly discussed. It is also characterized how these errors affect the final position of the target in the UTM coordinate system. It is important to note that only the errors common to all PS in a PSI point cloud, causing biased observations, are addressed here. For stochastic effects influencing individual PS, such as feature localization error depending on the Signal-to-Clutter-Ratio (SCR) of the target, the reader is referred to [22].

2.1. Errors in Azimuth and Range Times

The radar timings of a target can be affected by various error sources. These errors include:

- satellite dynamic effects such as incorrect annotation of t_{az} in the time of radar pulse reception or incorrect annotation of τ_{rg} due to instrument cable delays in the satellite [9].
- propagation delays caused by the ionosphere and the troposphere, from which the latter is considered the most prominent error source on τ_{rg} for SAR satellites operating in X-band [7,9].
- geodynamic effects that change the position of a target on the ground, including solid earth tides, plate tectonics, ocean loading and atmospheric loading [9,18].

These errors cause biases in the timings, which are directly propagated into the final coordinates of the geocoded target through Equations (1) and (2). For t_{az}, the errors cause shifts only in the along track direction. If we consider a straight satellite orbit trajectory and approximate the curved Earth geometry by a rectilinear one [23], as is visualized in Figure 1a, the ground shift δl_{az} due to an erroneous azimuth time t'_{az} is calculated in meters as:

$$\delta l_{az} = v_g (t'_{az} - t_{az}), \tag{4}$$

where t_{az} is the true azimuth time and v_g denotes the ground-track velocity of the satellite [23]. The error affects only the horizontal geocoded coordinates and with the knowledge of the local

heading angle of the satellite α, its effect can be projected onto the local East δl_{az}^E and North δl_{az}^N components (see Figure 1b):

$$\delta l_{az}^E = \delta l_{az} \sin \alpha, \tag{5}$$

$$\delta l_{az}^N = \delta l_{az} \cos \alpha. \tag{6}$$

Considering the near polar orbit of current SAR satellites, with heading angles close to 180° or 360°, Equations (5) and (6) show that error in t_{az} mostly affects the North component of the geocoded coordinates.
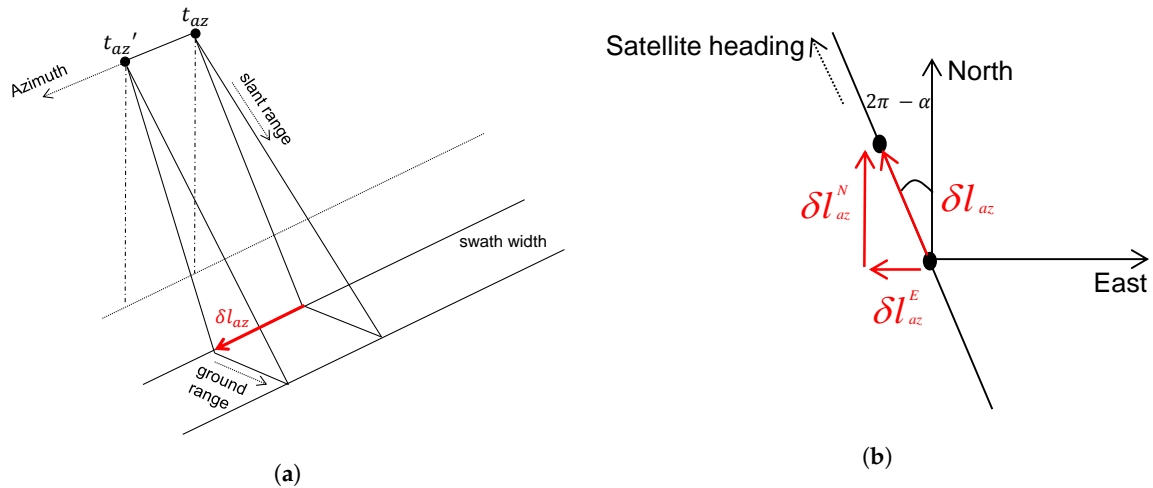

(**a**)　　　　　　　　　　　　　　　　　　　　　　　　　　(**b**)

Figure 1. Timing bias in azimuth and its effect on the geocoded coordinates. (**a**) azimuth timing error causing a displacement on the ground δl_{az}; (**b**) projection of δl_{az} onto the East δl_{az}^E and North δl_{az}^N components using satellite heading angle α.

Errors in τ_{rg} cause a delay in the received radar pulse that affects the geometrical distance between the satellite and the target in the slant range direction δl_{sr}, which leads to a displacement on the ground range δl_{gr} depending on the the local incidence angle θ (see also Figure 2a):

$$\delta l_{sr} = c\, \frac{(\tau'_{rg} - \tau_{rg})}{2}, \tag{7}$$

$$\delta l_{gr} = \frac{\delta l_{sr}}{\sin \theta}, \tag{8}$$

where τ_{rg} and τ'_{rg} are the true and the erroneous range time, respectively. Similar to δl_{az}, δl_{gr} is related to the East δl_{gr}^E and the North δl_{gr}^N components by a projection using α (see also Figure 2b):

$$\delta l_{gr}^E = \delta l_{gr} \cos \alpha, \tag{9}$$

$$\delta l_{gr}^N = -\delta l_{gr} \sin \alpha. \tag{10}$$

It is evident from Figure 2a and Equation (8) that δl_{gr} becomes larger with steeper incidence angles. It is also worth mentioning that, according to Equations (9) and (10), the timing error in range manifests itself mostly in the East coordinate component.
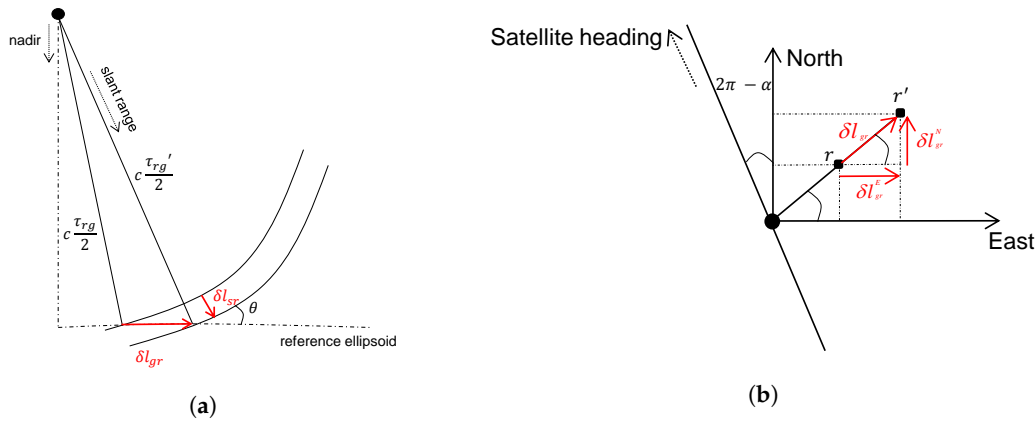
Figure 2. Timing bias in range and its effect on the geocoded coordinates. (**a**) range timing error and its impact on geocoding in the ground range δl_{gr}; (**b**) projection of δl_{gr} onto the East δl_{gr}^{E} and North δl_{gr}^{N} components using satellite heading angle α.

2.2. Error in the Height of PS

In Section 1, it was mentioned that the height estimates in InSAR approaches are defined with respect to a reference point. This implies that no DEM error is assumed for this point and its final height is equal to its corresponding DEM height. In the specific case of PSI, exploiting the differential interferograms, the DEM error for each PS is estimated relative to the reference point and at a later step before geocoding, the DEM height of each PS is added to its DEM error to obtain the final PS height denoted by H_T'. It is obvious that the final absolute height accuracy of all PS depends on the DEM error of the reference point, which is an overall unknown offset. This height error $\delta H = H_T' - H_T$ is constant for all PS and has a significant effect on final geocoded coordinates both in the horizontal and in the vertical components, as can be seen in Figure 3. The variable δH causes a horizontal shift in the ground range δl_H^{gr} as:

$$\delta l_H^{gr} = \frac{\delta H}{\tan \theta}. \tag{11}$$

Similar to the error in range timing (see Equations (9) and (10)), δl_H^{gr} can be projected onto the East and the North components by knowledge of α.
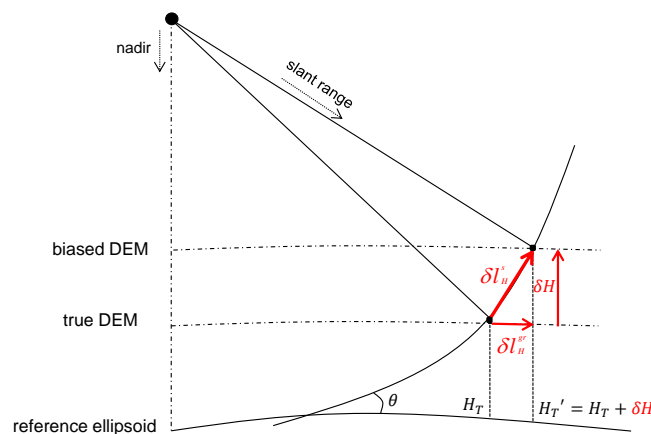

Figure 3. Depiction of height error δH due to unknown DEM error at the reference point and its effect on geocoded coordinates. It can be seen that this error causes a shift in the cross-range direction δl_H^s, which is decomposed into an offset in ground range δl_H^{gr} and a vertical component δH.

3. Methodology

The error sources and their impact on the localization of InSAR point clouds have been addressed in Section 2. In this section, we present our proposed framework for the estimation and removal of the aforementioned biases, which leads to an improvement of the overall geocoding accuracy of PSI point clouds. The method is divided into four parts, which are briefly explained in the following and are depicted in Figure 4. Each processing step is carried out independently and produces an intermediary result, which is shown as a parallelogram in Figure 4. The double shapes indicate that two or more SAR image stacks are involved in each processing. The intermediary results, gathered in the dashed gray rectangle, are the input for the fourth step, updated geocoding, which produces the final absolute 3D PSI point clouds. Please note that the input for our approach includes a minimum of two SAR image stacks acquired from different viewing geometries i.e., from separate orbits. This is required for the extraction of highly precise GCPs. Apart from the methodology description, the procedure to compare the geocoded PSI point clouds with external LiDAR data is also explained in this section.
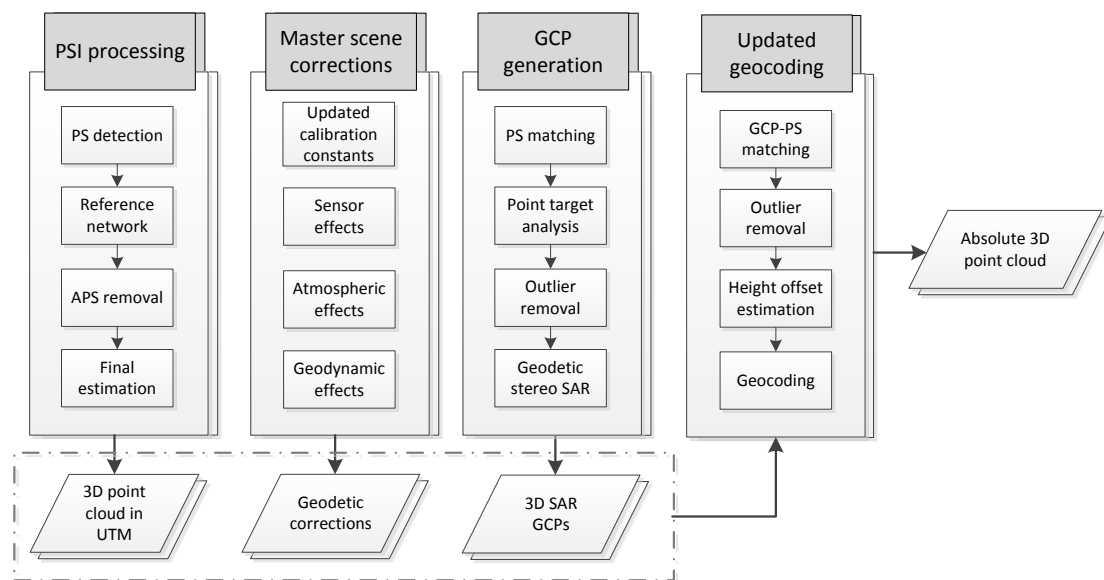

Figure 4. Workflow of the utilized geocoding error correction method. The processing steps, included in the big rectangles, are carried out independently and their results are shown by the parallelograms. The double shapes indicate that each processing is carried out for two or more SAR image stacks, which is the necessary condition for GCP generation.

3.1. PSI Processing

A concise description of the PSI technique has been given in Section 1. The major steps of the processing chain are outlined in Figure 4. The output of the PSI processing are the geocoded 3D UTM coordinates of PS and their displacement parameters. For details about the theory of the method and its processing sequence, the interested reader is referred to [1,2].

3.2. Master Scene Corrections

Geocoding is commonly carried out based on the orbit and the timing information of the master SAR image. In order to provide unbiased t_{az} and τ_{rg}, the errors stated in Section 2.1 are mitigated by the SAR imaging geodesy approach [9,26–29]. The ionospheric path delay is corrected on base of the global Total Electron Content (TEC) maps that use the single layer model [30] for ionosphere. Using the orbit information of the SAR satellite, the point at which the line of sight of the satellite passes through the ionospheric layer to the point of interest is analytically localized. The zenith TEC

value is then determined at the location of the aforementioned point through spatial and temporal interpolation of the TEC maps [27,31,32]. Subsequently, the vertical TEC value at the point is projected onto the line of sight of the SAR satellite using a proper mapping function [31]. The path delay caused by troposphere is computed through the 4D integration of numerical weather data from European Centre for Medium-Range Weather Forecasts (ECMWF). The method extracts the dataset from a local database, converts it to a conventional geographic coordinate system, performs a 3D interpolation for defined integration points and eventually integrates the refractivity index along the integration path in the slant-range direction from the point of interest to the satellite [15,26]. The geodynamic effects are removed based on the state-of-the-art models according to the International Earth Rotation and Reference Systems Service (IERS) guidelines [33]. The effects are reported in the horizontal and the radial coordinate components, which are transformed to the radar timing coordinate system [9]. The coordinate transformation residuals between object and sensor coordinate systems are compensated for by taking into account the plate tectonics effect and referring the observation to a reference epoch [26]. Finally, the geometric calibration constants in range and azimuth are updated based on the most recent studies concerning long-term corner reflector experiments [34]. For this purpose, all the aforementioned effects are initially mitigated using the described methods and then the median of offsets between the expected position of the corner reflector, surveyed with GNSS, and its corresponding position in a time series of SAR images is considered as new re-calibration constants [34]. All of the mentioned corrections are computed for a coarse grid and are further interpolated for targets inside the grid [35]. The corrections are subtracted from the annotated SAR measurements to achieve absolute t_{az} and τ_{rg}. For more information regarding SAR imaging geodesy and its operational implementation, the reader is referred to [9,26,34,35].

3.3. GCP Generation

The first step regarding the GCP generation is the correct detection of identical PS from multiaspect SAR images. After the calculation of corrections (see Section 3.2), for each individual Single Look Complex (SLC) image, bright point-like targets are extracted through the spatial analysis of their SCR. It is important to mention that for the extraction of GCP candidates, PS from the previously generated PSI point clouds can be used as well, as has been shown in [21]. However, since our aim is to show the flexibility of the geocoding error correction approach, independent from which GCP candidate detection method is used, we use a completely separate processing chain from PSI processing in order to extract GCPs [20]. After the GCP candidate detection, target radar coordinates are geocoded using an external DEM and a point-matching scheme is carried out to find the correspondence of each target in other SLCs with different viewing geometries [20]. The method then extracts the precise timing information of each point target in all SLCs through Point Target Analysis (PTA) [23]. Finally, identical PS timings from multiaspect data are combined using stereo SAR to estimate the absolute 3D coordinates of each PS [18]. The individual steps are also depicted in the workflow in Figure 4, where the last step, called geodetic stereo SAR, refers to the combination of the SAR imaging geodesy and the stereo SAR methods. The output of this part is several thousands of GCPs with absolute 3D coordinates.

A further filtering of the GCPs is carried out based on their posterior coordinate standard deviations, which are also the output of the GCP extraction framework. The threshold for this filtering is chosen based on our previous studies dealing with the stereo localization of natural PS [18,21]. As a final refinement, we filter the GCPs based on their closeness to the road network data of the scene, which can be freely accessed either from OpenStreetMap [36] or from country-specific geo-portals as demonstrated in [21]. The reason for the latter filtering is that the most reliable GCP targets are the ones that can be easily recognized from cross-heading geometries. In urban areas, these targets include lamp poles and traffic signs, which are commonly found close to streets and roads. For different approaches of SAR GCP generation using high resolution SAR data, the interested reader is recommended to consult [20,21].

3.4. Height Offset Estimation and Updated Geocoding

The objective of this processing step is to calibrate the height coordinates of the PS in PSI point clouds by using the generated SAR-based GCPs. In order to estimate the DEM error of the reference point of each PSI point cloud, the first step is to find for each GCP its corresponding point in the PSI point cloud since different processing chains were used for PSI processing and GCP generation (see Section 3.3). To this end, we radar-code the GCP coordinates onto the master scenes of each SAR image stack taking into account the full geodetic corrections. We then select the nearest neighbor of the GCP in the PS point cloud that has an Amplitude Dispersion Index (ADI) lower than a threshold [1]. In this way, we can correspond each GCP to its closest PS that most likely represents the base of lamp poles or traffic signs. However, this approach is highly prone to detecting wrong correspondences since the ADI of the lamp pole might be higher than the predefined threshold. To overcome this limitation, we further refine the GCP-PS correspondence detection as follows:

1. Coordinate differences are calculated in range and azimuth of the GCPs and their assumed corresponding bright point in the PSI point cloud.
2. Points with coordinate differences larger than two times the standard deviation of differences (2σ rule) are discarded first in range and then in azimuth.

Note that, due to the probable existence of outliers, the standard deviation of coordinate differences is robustly estimated using the Median Absolute Deviation (MAD) measure [37]. The reason for the strict measure in the second step is that the geodetic and the propagation errors on range and azimuth of a TerraSAR-X spotlight scene usually vary within centimeter and millimeter regimes, respectively. Therefore, large coordinate difference variations show that the wrong correspondence between the GCP and the bright point has been detected and lead to discarding the candidate.

After the correspondence detection has been refined, we need to robustly estimate a single height offset among GCP heights and their corresponding PS heights. Initially, the GCP-PS ellipsoidal height differences are calculated. The dispersion of height differences is estimated using MAD. Furthermore, large height difference observations are discarded based on the 2σ rule, while the distance of each observation is evaluated with respect to the median of the differences to ensure robustness [37]. Subsequently, the outlier-free height difference histogram is formed. The peak of the smoothed histogram is detected as the mode of height differences, which is the approximate estimate of the DEM error of the reference point.

The estimated height offset is subtracted from the ellipsoidal heights of the geocoded PSI point cloud to obtain correct absolute heights. At the final step, the corrected heights as well as the corrected range and azimuth timings (see Section 3.2) are used for an updated geocoding using Equations (1)–(3). The result is the corrected geocoded 3D PSI point cloud in either Cartesian or UTM coordinate system.

3.5. Cross-Comparison with LiDAR

In order to check the overall localization accuracy of PSI point clouds after correcting for the geocoding errors, we carry out qualitative and quantitative horizontal and vertical analyses with respect to a reference LiDAR data. Since data from LiDAR are acquired with a nadir-looking geometry in contrast to the slant-viewing geometry of SAR sensors, we should make the data from both sensors as compatible as possible. To this end, we first extract the PS originated from façade of buildings using an approximate approach proposed in [14]. For each PS, a neighborhood is selected in the horizontal plane within which the variance of PS heights is evaluated. If the variance is higher than a threshold, the point is considered to be a façade PS [14]. A similar approach is carried out to extract the 2D building footprints from the LiDAR point cloud. After the separation of façade and non-façade points in the UTM coordinate system, the localization accuracy analysis is carried out as follows:

- 2D horizontal accuracy: For a number of test sites, a line is robustly fitted to each side of the building footprint in the LiDAR point cloud in the East–North plane. The mean distance of the

façade PS in the PSI point clouds are evaluated with respect to the corresponding footprint line in the LiDAR data. The average of all the deviations for corrected and non-corrected PSI point clouds are separately calculated, which demonstrates the closeness of each point cloud to the reference LiDAR data.

- 1D vertical accuracy: For an identical area in the LiDAR data and both corrected and non-corrected PSI point clouds, the façade points are excluded. Then for each of the three point clouds, ellipsoidal height histograms are formed containing two peaks which respectively represent ground and building roofs from the selected test site. The difference between the height values for which the ground peaks, in the PSI point cloud and in the LiDAR point cloud occur, indicates the vertical accuracy.

4. Area of Interest and Data Set

The investigated test site includes the central urban area of Berlin located in northeastern Germany. Our data set includes two stacks of TerraSAR-X high-resolution spotlight products with nominal spatial resolution of 1.1 m in azimuth and 0.6 m in range. The total number of images is 214, which cover a period of approximately five years from April 2010 to March 2015. In terms of viewing geometry, one stack is acquired from a descending orbit track and the other one from an ascending orbit with approximate incidence angles of 36° and 42°, respectively. In addition to the SAR data, an aerial LiDAR point cloud with a 3D accuracy in the decimeter regime is available, for which the acquisition period is from January to March 2009. The LiDAR data is used as reference to check the absolute 3D localization accuracy of PSI point clouds. The acquisition parameters of each SAR image stack are reported in Table 1. The mean scene coverage of the SAR images, the extent of the LiDAR data and the test sites used for vertical and horizontal accuracy analysis are marked in Figure 5.

Table 1. Acquisition parameters of stacks of SAR images in Berlin.

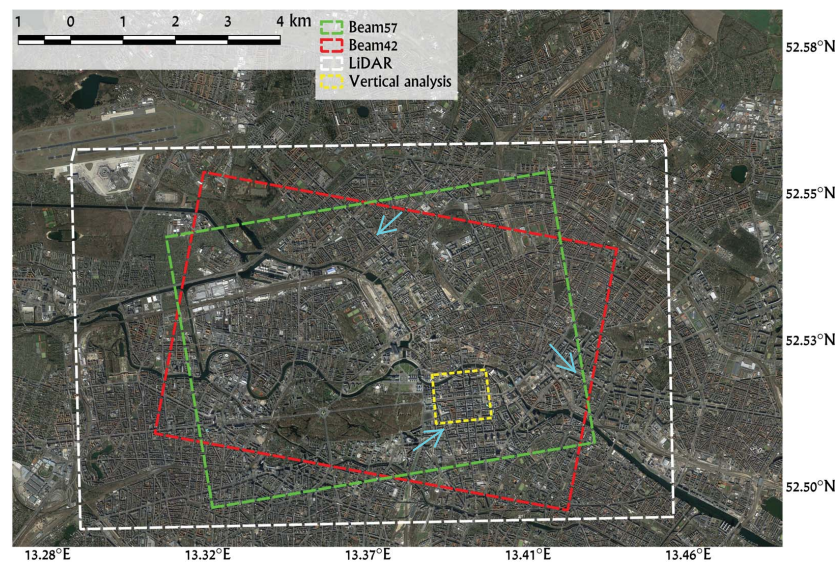| Beam Nr.	Center θ (Degree)	Average α (Degree)	Orbit Direction	Area (km^2)	Polarization	Nr. of Images
42	36.1	350.3	descending	40.94	VV	107
57	41.9	190.6	ascending	42.52	VV	107

Figure 5. Optical image of Berlin taken from Google Earth. The mean scene coverage of the TerraSAR-X images as well as the extent of the LiDAR data are marked with colored rectangles. The small yellow rectangle shows the test site within which the vertical localization accuracy of PSI point clouds has been analyzed. The cyan arrows indicate the three test sites used for horizontal accuracy analysis.

5. Results and Discussion

5.1. Berlin InSAR and PSI Processing

The two stacks of Berlin have undergone InSAR and PSI processing using the PSI-GENESIS of the German Aerospace Center (DLR) [38–41]. The master scenes for both stacks are chosen in winter, on 24 December 2011 for beam 57 and on 7 March 2012 for beam 42, and central with respect to the temporal and perpendicular baselines. The baseline distribution of the SAR images of both beams can be seen in Figure 6. Prior to the InSAR stacking, complex SAR images are oversampled by a factor of two in both azimuth and range in order to avoid signal aliasing of the interferograms and amplitudes. The coregistration of slave images onto the master scene is carried out geometrically using a Shuttle Radar Topography Mission (SRTM) DEM and precise orbits and further applying an offset correction using PS [42]. Differential interferograms are then formed by removing the topographic phase contribution using the aforementioned DEM.
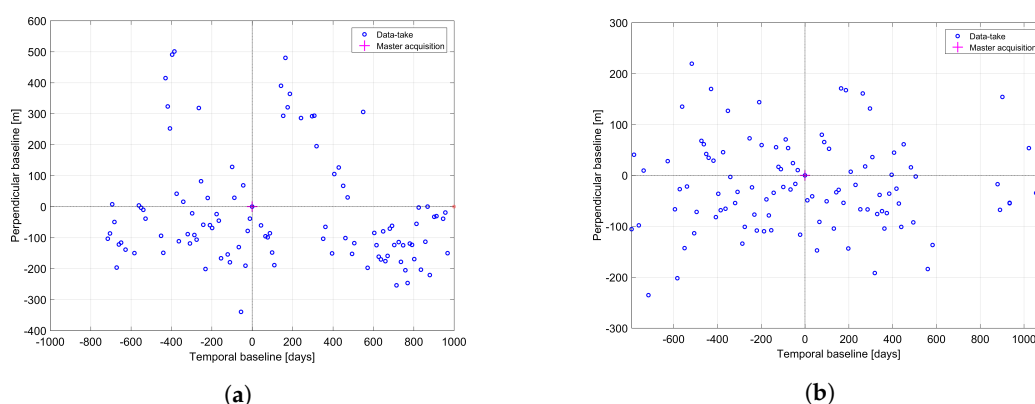

Figure 6. Temporal-perpendicular baseline distribution of the SAR image stacks in Berlin. (**a**,**b**) correspond to the data stacks of Beam57 and Beam42, respectively.

The PSI processing is initiated by extracting the PS candidates by thresholding on the SCR of the targets estimated in the coregistered SAR image stack [2,38]. After geocoding of the PS candidates using their corresponding DEM heights, a grid with cell size of 250 m is superimposed on all PS and the PS with the highest phase stability is selected within each cell using the locally estimated temporal coherence. These PS form the reference network, which is the basis for the estimation of height and deformation parameters as well as the estimation and removal of the atmospheric phase screen [1,2]. The parameters of interest in this study include DEM error, linear and periodic motion since we are analyzing an urban area with the acquisition time spanning over a few years. After the estimation of differential parameters on the arcs of the network, a robust ℓ_1-norm network inversion for outlier rejection is carried out and a high quality PS is selected as the reference point [40]. Subsequently, the DEM error and the deformation at each point of the reference network are estimated using an ℓ_2-norm network inversion relative to the previously selected reference point. After compensating the estimated APS, the PS that are not part of the reference network are connected to the nearest reference network points and their parameters are estimated. Later, the DEM height of each PS is added to its corresponding differential residual height estimate. The radar timing of each PS and its updated height are used for geocoding. The final results of the PSI processing are the topographic and deformation maps as well as displacement history of each PS. Since the focus of this paper is on geocoding accuracy of the point clouds only the topographic maps are reported. The results are shown in Figures 7 and 8. The ascending and descending point clouds consist of approximately 1.3 and 1.4 million PS, respectively. The higher density of PS is seen along the railways and also on building façades. The white parts with no detected PS include mostly the vegetated areas such as the well known Tierpark situated at the lower left of the center of the scene. Deformation maps of Berlin

obtained from TerraSAR-X PSI and SAR tomography have been previously analyzed [4,14,43,44] and are not further discussed here.
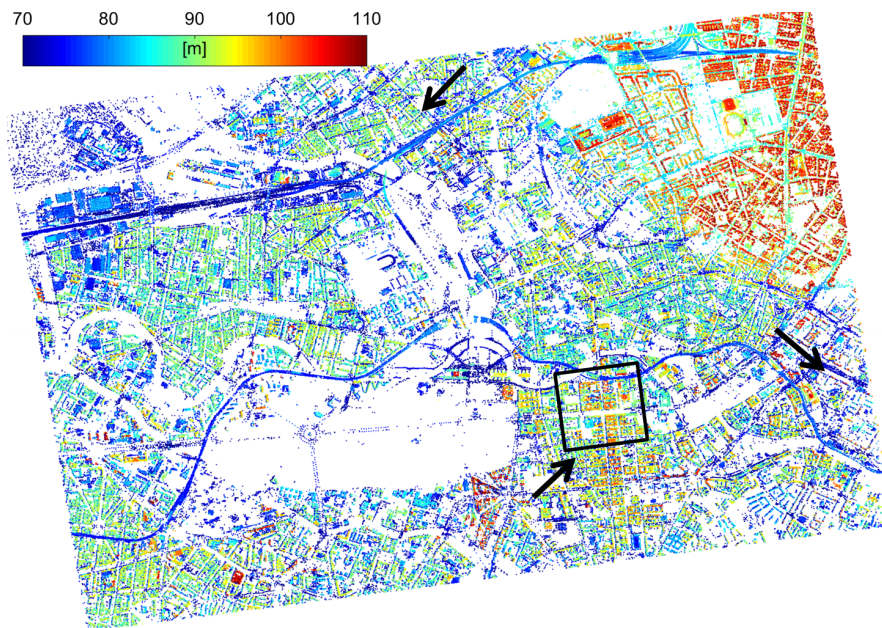

Figure 7. Geocoded PSI point cloud of Berlin, reconstructed from Beam57, in UTM coordinates. The ellipsoidal height is colorcoded. The *x*- and *y*-axis correspond to the UTM East and North, respectively. The test sites used for vertical and horizontal accuracy analysis are marked with black rectangle and black arrows, respectively.
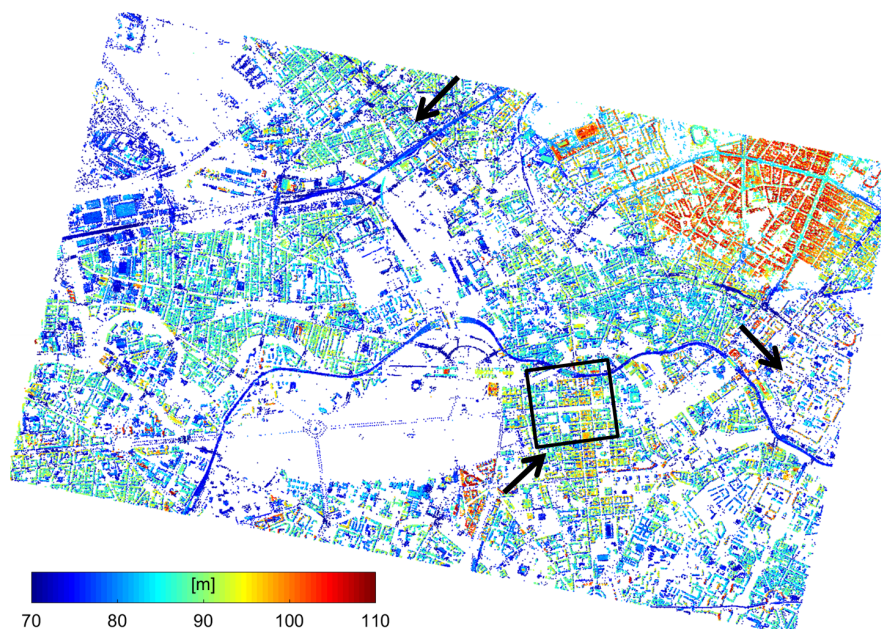

Figure 8. Geocoded PSI point cloud of Berlin, reconstructed from Beam42, in UTM coordinates. The ellipsoidal height is colorcoded. The *x*- and *y*-axis correspond to the UTM East and North, respectively. The test sites used for vertical and horizontal accuracy analysis are marked with black rectangle and black arrows, respectively.

5.2. Geodetic Corrections, GCP Generation and Height Offset Estimation

Geodetic error corrections for individual SAR images and GCP generation have been carried out with the SAR Geodesy Processor (SGP) of the DLR [20,35]. All the timing corrections stated in Section 3.2 are evaluated for a coarse grid of 200 m in the slant range geometry. The sum of all corrections in the range and azimuth components of the master scene acquisitions of both beams are visualized in Figure 9. The corrections are in the radar coordinate system where, the *y*-axis represents the azimuth coordinate and the *x*-axis depicts the range coordinate. The errors are defined in the unit of length by multiplying range errors with $c/2$ for one-way measurements and azimuth errors with an average TerraSAR-X ground track velocity of 7050 m/s. It is seen that the magnitude of range errors is close to 3 m for Beam57 and approximately 2.75 m for Beam42, whereas the difference of maximum and minimum range errors over the scene is only 4 cm for both beams. The main contributions to the range errors come from the troposhperic error followed by geodynamic effects and finally the ionospheric delays. The azimuth errors are far less significant than range errors and manifest in sub-decimeter regimes, with millimeter variations across the scenes for both beams. The main source for azimuth errors is geodynamic effects followed by errors due to satellite dynamics. It is important to note that given the small squint angle of the TerraSAR-X high resolution spotlight products and their operation in X-band, the tropospheric and ionospheric delays are negligible for azimuth timings and therefore are not considered in the SAR imaging geodesy method [34].
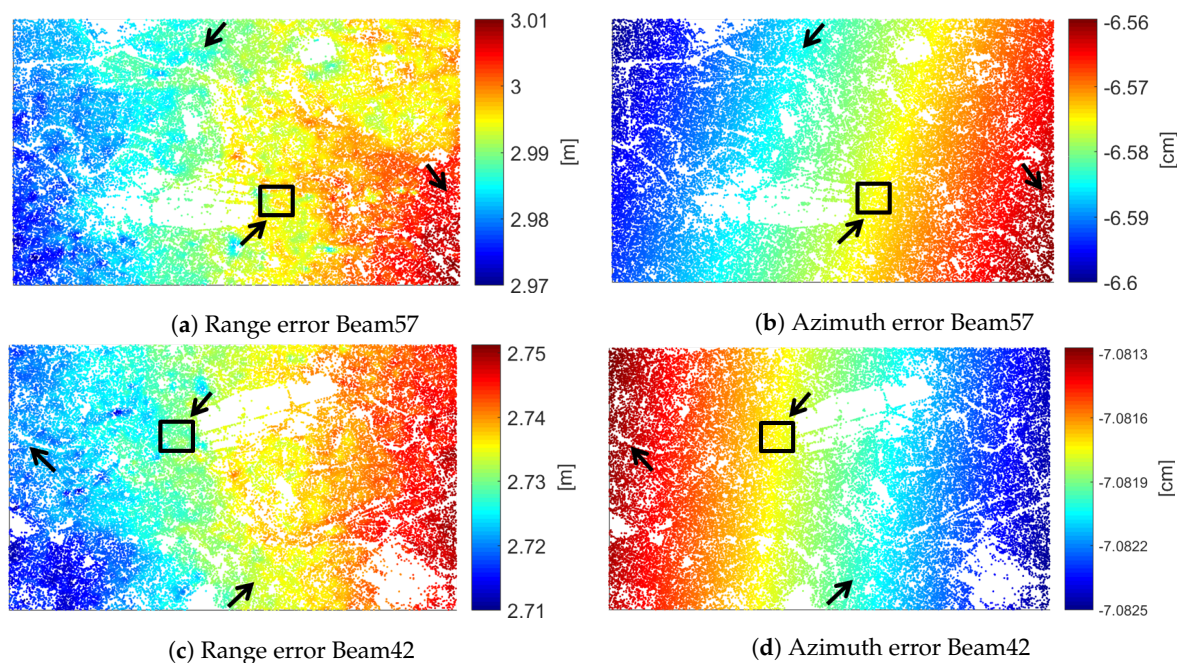

(**a**) Range error Beam57

(**b**) Azimuth error Beam57

(**c**) Range error Beam42

(**d**) Azimuth error Beam42

Figure 9. Range and azimuth error in the radar coordinate system (*x*-axis: range, *y*-axis: azimuth) of the master scenes of both investigated beams, which were acquired on 24 December 2011 (Beam57) and on 7 March 2012 (Beam42). Note that the scale of the colorbars for range error is in meters while the one for azimuth errors is in centimeters. The test sites used for vertical and horizontal accuracy analysis are marked with black rectangle and black arrows, respectively.

For GCP generation, 27 images from both beams are selected, which have been acquired within a period of 12 months, in 2014 and at the beginning of 2015. The reason for the time restriction is avoiding the impact of plate tectonics on the final GCP coordinates. Bright point targets are precisely extracted from the SLCs using PTA with a maximum oversampling factor of 32 in both azimuth and range followed by a 2D kernel interpolation in a 3×3 window [27]. The signal path delays and geodetic errors are corrected for the location of each point target using their extracted timing information. After coarse geocoding of these points using a SRTM DEM, for each target, its corresponding points

from other SLCs are detected if they are located within a distance of couple of meters relative to each other. The timing information of assumed identical targets are then combined through stereo SAR to estimate the absolute 3D coordinates of the GCP. The total number of GCPs localized from the combination of the aforementioned 27 SLCs is 17,673. The SGP removes all the GCPs with localization precision worse than 1 m based on the 1σ rule. We further filter the results based on the posterior precision of GCP coordinates in the local East, North and ellipsoidal height. For precision values, we impose a threshold of 10 cm for all coordinate components. It is observed from Figure 10 that, by imposing the 10 cm threshold, there still remains an acceptable number of GCPs. This number is equal to 3800. As the final step, as was mentioned in Section 3.3, the GCPs are further filtered based on their closeness to the road network data of the scene, which reduces the number of GCPs to approximately 1000. The final obtained GCPs are visualized (in purple) on top of the geocoded PSI point clouds (in gray) in Figure 11. It is seen that the GCPs almost cover the full area within both PSI point clouds.
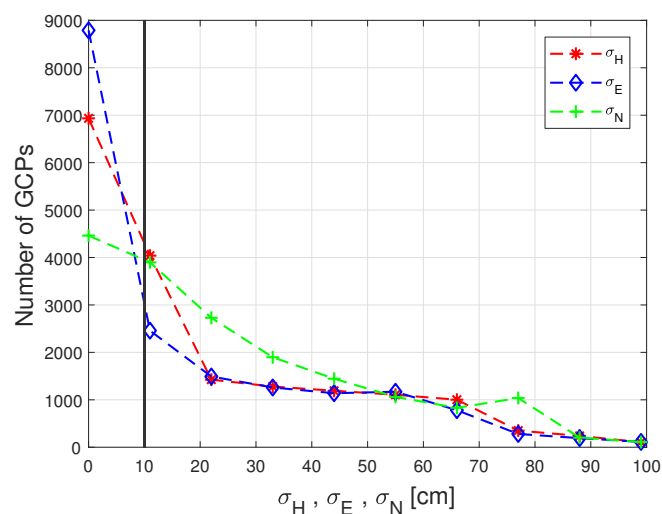

Figure 10. Histogram of GCP precision values in East σ_E, North σ_N and height σ_H.
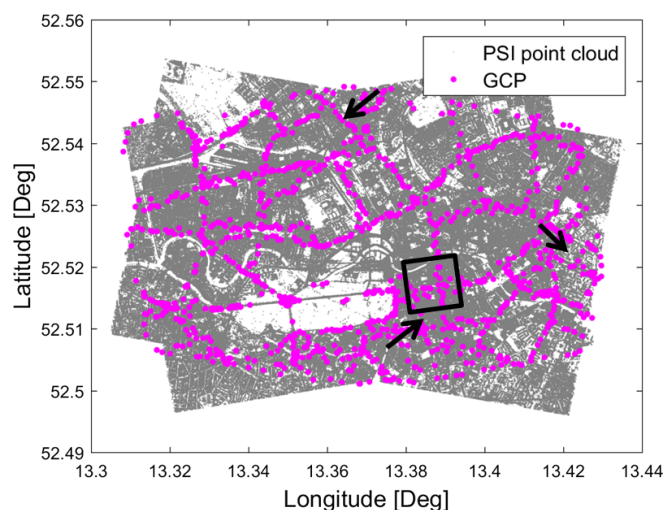

Figure 11. Distribution of the GCPs within the PSI point clouds of both beams. The test sites used for vertical and horizontal accuracy analysis are marked with black rectangle and black arrows, respectively.

The generated SAR-GCPs are used to estimate the DEM error of the reference point of each PSI point cloud. For each radar-coded GCP, its closest PS with an ADI value lower than 0.4 is selected. The higher threshold value of 0.4 is preferred instead of the recommended ADI threshold of 0.25 [1],

since lamp poles and traffic signs are assumed to be less stable than the more common PS which usually occur on building façades and corners. The correspondence detection is further refined by removing points which show large deviations with respect to the radar-coded GCPs (see Section 3.4). After the correspondence detection, the DEM error is estimated by analyzing the histogram of ellipsoidal height differences. Figure 12 summarizes the robust shift estimation procedure for both beams of Berlin. The left and right subfigures correspond to results from Beam57 and Beam42, respectively. The 2D scatterer plots show the distribution of range and azimuth coordinate differences among GCPs and their corresponding identified PS. In Figure 12a,b, the differences are plotted before outlier removal. It can be seen that very large deviations still occur after the correspondence detection step. Average deviations in the centimeter regime is observed in the range component while for the azimuth differences of Beam57, the mean deviation is close to 12 cm. The sample standard deviations, annotated in the subfigures, also show the large spread of the differences in both coordinate components in the meter regime. The second row of the subfigures, Figure 12c,d, depict the scatter plots after discarding outliers both in range and azimuth based on the 2σ rule. Note that the standard deviations for the outlier removal purpose are robustly estimated using MAD, as was explained in Section 3.4, while the annotated values in the subfigures are sample standard deviations, that are sensitive to outliers. From Figure 12c,d, it is clearly observed that the mean deviations decrease and now are close to 2 cm and 3 cm for range coordinate differences and approximately 2 cm and 1 cm for azimuth coordinate differences of Beam57 and Beam42, respectively. It is important to note that the radar coordinate differences have been compensated for the continental drift which occurred between the master acquisition times of each beam, in 2011 and 2012, and the epoch with respect to which the GCPs have been localized, in 2015. This coordinate shift has been computed utilizing the average East, North, Up velocities from the nearest International GNSS Service (IGS) permanent station, in Potsdam approximately 45 km away from Berlin, and have been projected into the line-of-sight and the azimuth directions of each beam. Eventually, the third row, Figure 12e,f shows the histogram of height differences among GCP-PS pairs after discarding observations based on the 2σ rule. The peaks of the smoothed version of the height difference histograms provide the robust estimates of the DEM errors which are -4.06 m and -6.27 m for Beam57 and Beam42, respectively. It is expected that the DEM errors have different values because of non-identical PSI reference points. The estimated height shifts are subtracted from the final height estimates of PSI and along with the fully corrected range and azimuth timings, using SGP, an updated geocoding is carried out for each beam separately. The localization accuracy of the corrected point clouds are evaluated in the next subsection.

5.3. Cross-Comparison with LiDAR

The geocoded PSI point clouds before and after applying the aforementioned corrections are compared with the reference LiDAR data. It is important to note that, since an objective comparison is difficult to perform, due to the inherent differences between LiDAR and PSI point clouds, we avoid using the term validation and therefore reside with the word cross-comparison.

In 2D, the corrected and non-corrected PSI point clouds of ascending and descending tracks are overlaid on the LiDAR data. For a few individual buildings, their façade PS are extracted by the method explained in Section 3.5. For this procedure, the neighborhood size is chosen to be 4 m while the height variance threshold is 1.5 m as recommended by [14]. The results for three buildings, indicated with cyan arrows in Figure 5, are visualized in Figure 13. The green and red dots display the PS from ascending and descending point clouds, respectively, while the white dots show the extracted building footprints from LiDAR. All data points are presented in the UTM projection. The first and second rows correspond to the non-corrected and corrected results, respectively. Note that the subfigures have different scales. It is clearly seen that, before applying the geodetic corrections and DEM error compensation, red dots are located far away from the building footprints. This 2D shift is largely compensated after applying the corrections in all the three cases. Moreover, the endpoint of each façade in the corrected point clouds match with the endpoint of the façade from the point cloud from

the opposing geometry. This is not the case for non-corrected point clouds. Another observation is related to the ascending point clouds. The 2D shift imposed on the green point cloud after correction is much smaller than the red point cloud. One of the reasons is that the DEM error of the ascending point cloud is approximately two meters smaller than the descending point cloud. Another explanation for this is that the center incidence angle is smaller for the descending point cloud compared to the ascending one (see Table 1). Therefore, according to Equations (8) and (11), the horizontal shifts are larger for the descending point cloud as has been expected.
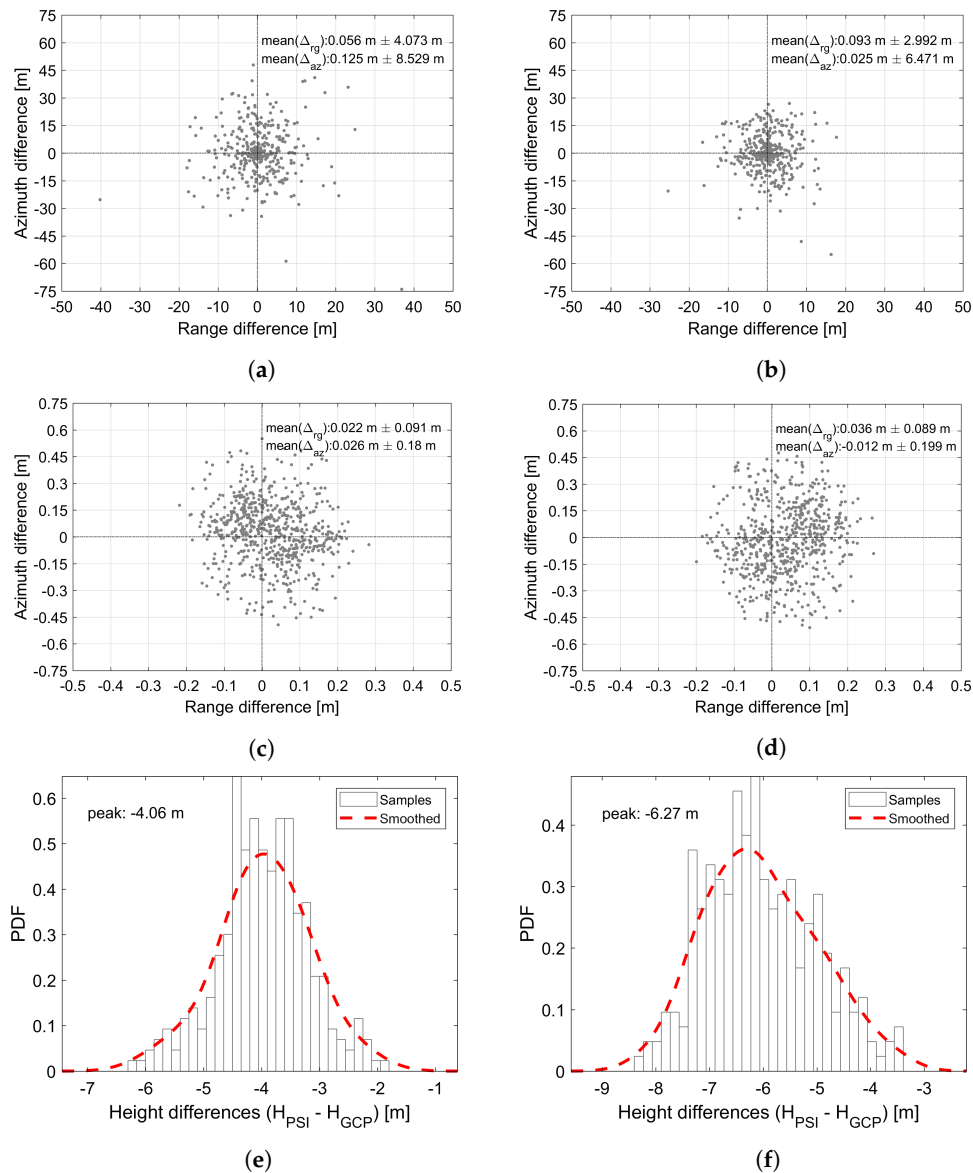

Figure 12. Restriction of the DEM error estimation to the true correspondences among GCPs and PS. (**a**,**b**) demonstrate the scatterplots of GCP-PS pairs before outlier removal, for Beam57 and Beam42, respectively; (**c**,**d**) show the scatterplots of the GCP-PS pairs after outlier removal. This causes a decrease in the bias and the standard deviation of the coordinate differences; (**e**,**f**) depict the robust height offset estimation after removal of the height differences in accordance with the 2σ rule. The peak of the smoothed histogram indicates the DEM error for each beam. Note that, for all the figures, the coordinate differences are defined as the GCP coordinates subtracted from their corresponding PS coordinates.

In order to quantify the approximate deviation of each point cloud from the reference LiDAR data, the distance of each façade PS was calculated with respect to a 2D line fitted to the corresponding LiDAR façade of all the three cases reported in Figure 13. The mean value of the distances for the corrected point clouds and the non-corrected point clouds are 0.4 m and 2.44 m, respectively. This proves that applying the corrections have certainly improved the overall geocoding accuracy of the PSI point clouds in the horizontal plane as the corrected point clouds are located closer to the LiDAR footprints compared to the non-corrected ones. It is important to note that a deviation equal to zero does not necessarily mean that the PSI point cloud is absolutely without any errors. The reason is that façade PS originate as triple-bounces from window corners of buildings which are not perfectly aligned with footprints extracted from LiDAR. Therefore, the difference in viewing geometries of both sensors does not allow a flawless comparison.
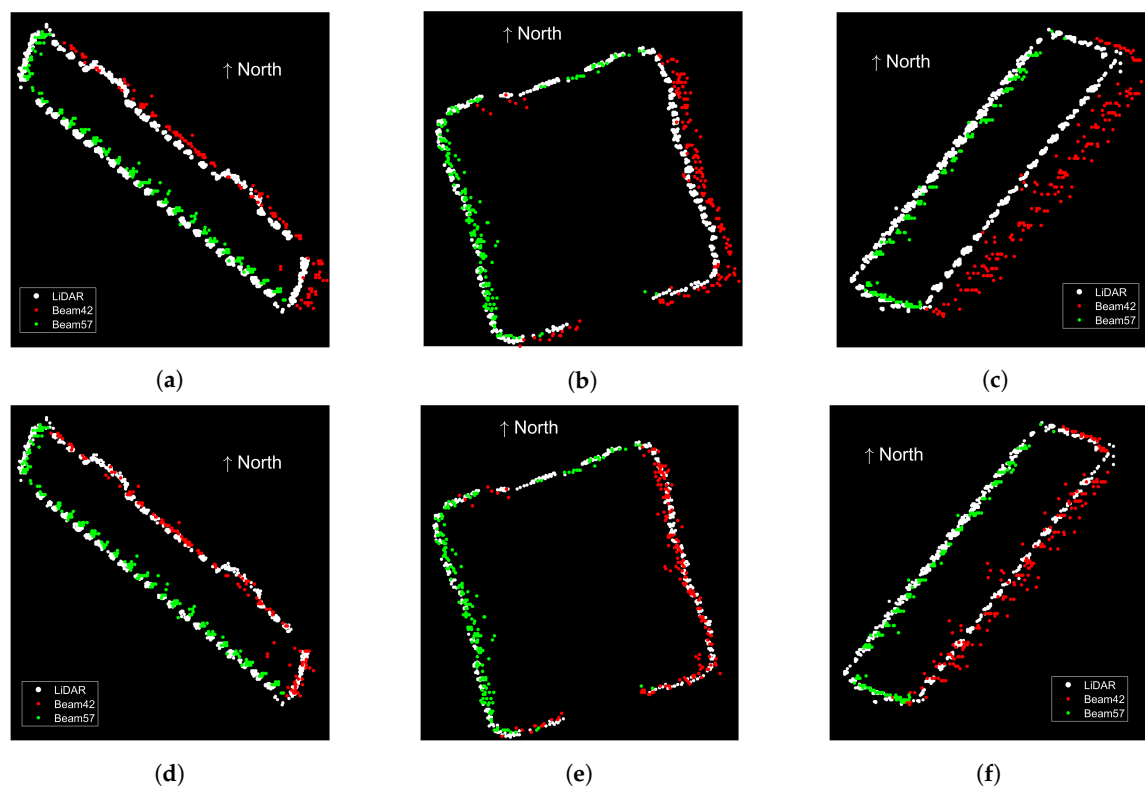

Figure 13. 2D horizontal localization accuracy analysis of PSI point clouds of the test sites marked with cyan arrows in Figure 5 (top: non-corrected, bottom: corrected). The *x*- and *y*-axes correspond to the Easting and Northing in UTM. Green and red dots show the ascending and descending point clouds while white dots show the extracted LiDAR footprints. The scale of figures is not identical meaning that the degree of zoom-in is higher in subfigures (**c**,**f**).

For the 1D vertical accuracy analysis, we first visually inspect the vertical cross-section of corrected and non-corrected PSI point clouds overlaid onto their corresponding LiDAR point cloud for a test site. Figure 14 visualizes this cross-section. In the subfigures, the *x*- and *y*-axes correspond to the UTM Easting and ellipsoidal heights, respectively. The red and green colors represent the PSI point clouds of Beam42 and Beam57 while the reference LiDAR point clouds are in white. In the LiDAR data, the ground line and the building roofs are easily distinguishable as they contain the majority of the points and are seen as very bright lines in the figures. By looking at Figure 14a,c, a clear height offset between the ground lines in the non-corrected PSI point clouds of both beams and the ground line in the LiDAR data is observed. Moreover, in the non-corrected point clouds, the PS on buildings are not seamlessly aligned with building roofs of the LiDAR point clouds. The non-identical height

shifts are compensated by employing the height offset estimation method, relying on SAR GCPs, as it can be seen in Figure 14b,d.
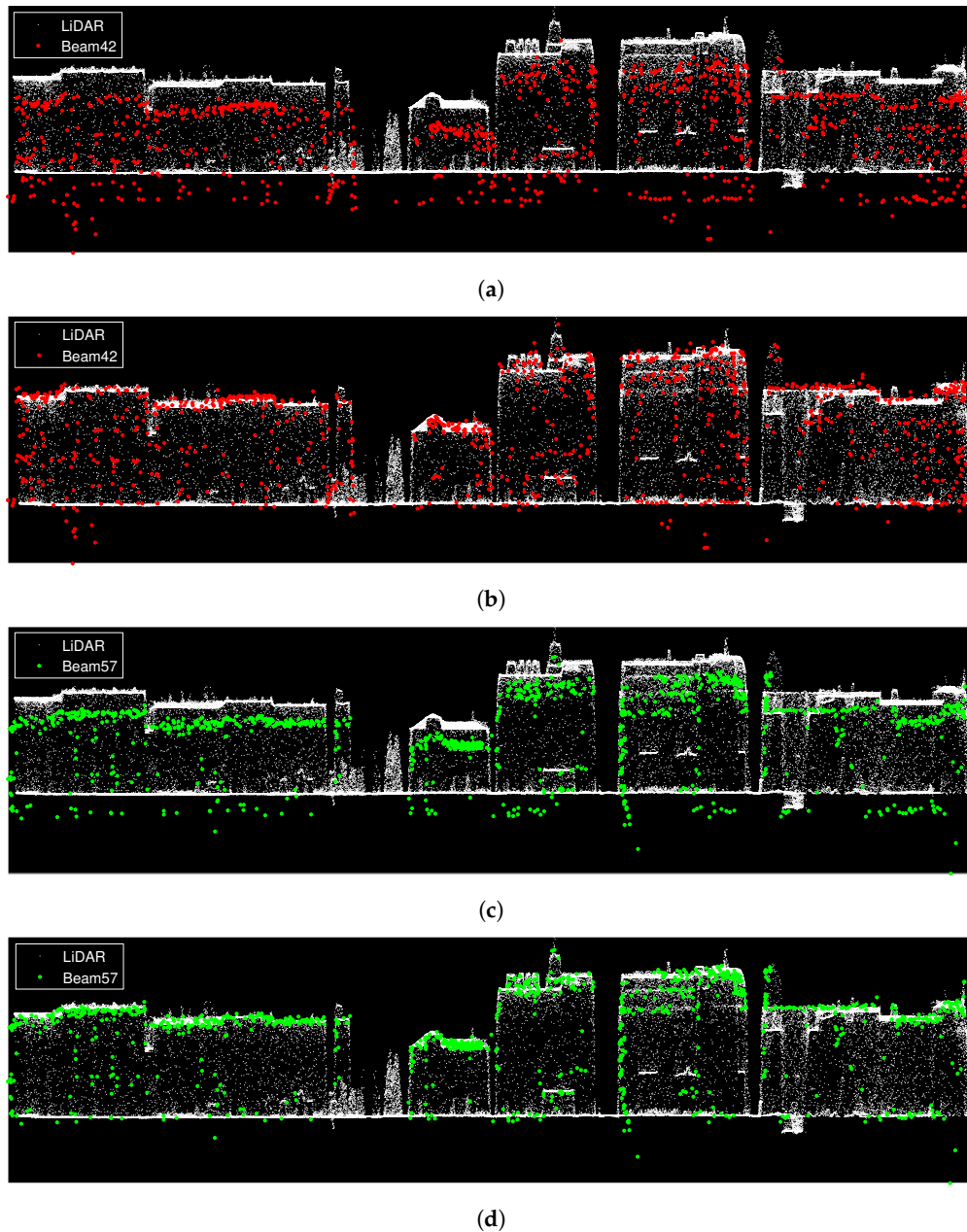

(**a**)

(**b**)

(**c**)

(**d**)

Figure 14. 1D vertical localization accuracy analysis of PSI point clouds, before and after applying geodetic corrections and height offset compensation, in comparison with LiDAR. The *x*- and *y*-axes correspond to the Easting in UTM and ellipsoidal height, respectively. Green and red dots show the ascending (Beam57) and descending (Beam42) point clouds while white dots show the LiDAR data. (**a**,**c**) depict the non-corrected point clouds while (**b**,**d**) correspond to the corrected ones. The height shifts, caused by the unknown DEM error of the reference points are easily recognizable in non-corrected point clouds while the offsets are compensated for in the corrected results. Note that the Easting and the height are differently scaled in order to emphasize the vertical effect.

In order to quantify the degree of improvement in the absolute height of PSI point clouds with respect to the LiDAR heights, the façade PS are excluded and height histograms are formed following the approach described in Section 3.5. This procedure is carried out for a subset of the PSI point clouds

and their corresponding LiDAR data marked with yellow rectangle in Figure 5. The ellipsoidal height histograms are reported in Figure 15 where the top and bottom subfigures correspond to Beam57 and Beam42, respectively.

The red, green and blue colors describe the height histograms of the LiDAR, non-corrected PSI and corrected PSI point clouds. The first peak of the histograms relates to the ground points. It is observed that, after the DEM error of the reference point is estimated with the use of SAR-GCPs, the height histograms of the corrected PSI point clouds move toward the height histograms of the LiDAR data. Within each histogram, the height of the ground peaks are reported. The absolute differences in the height at the peaks of the non-corrected point clouds with respect to the peak of the LiDAR data are approximately 3.96 m and 6.43 m for Beam57 and Beam42, respectively. The differences reduce to 0.21 m and 0.11 m after the proposed height offset compensation method has been applied.
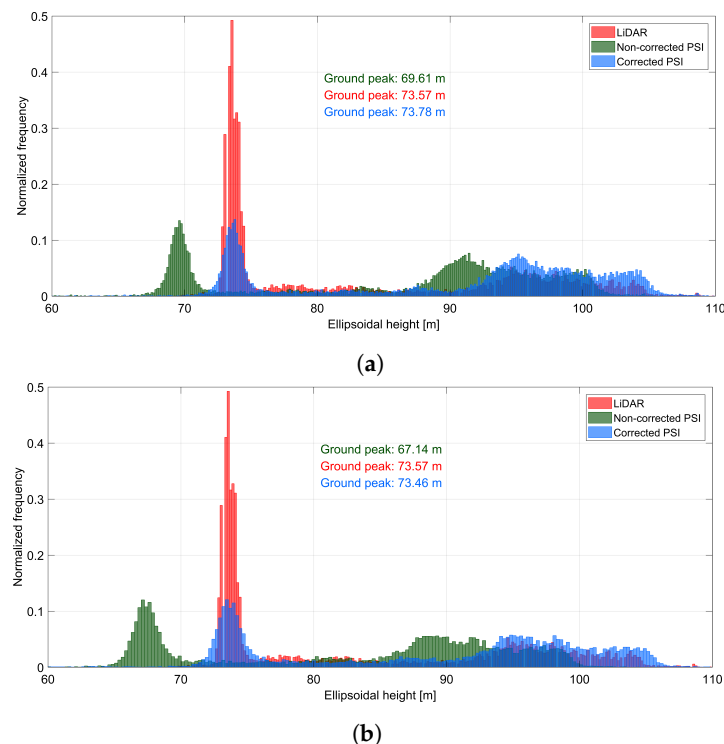

(**a**)

(**b**)

Figure 15. Ellipsoidal height histograms of non-façade points in LiDAR (red), non-corrected PSI (green) and corrected PSI point clouds (blue) corresponding to the yellow bounding box in Figure 5. (**a**,**b**) show the results for Beam57 and Beam42, respectively. For both beams, the height shifts are compensated for after correction, using SAR-GCPs, as the ground peaks of PSI point clouds and LiDAR data become aligned.

6. Conclusions

In this paper, a SAR-based method has been introduced in order to improve the geocoding accuracy of PSI point clouds. It has been shown that the geocoding errors related to height can be compensated for by the use of GCPs which are extracted and localized using the SAR data itself. The corrections in range and azimuth timings of the PS also influence the horizontal accuracy and have been fully considered using the SAR imaging geodesy technique. The method does not rely on any external positioning data and can be carried out as a post-processing step after PSI point cloud generation. A few test cases have been visually inspected, which showed the significance of the applied corrections, especially for the estimation of the DEM error of PSI reference points. By cross-comparison of corrected PSI point clouds with respect to a reference LiDAR data, it has been shown that a 2D horizontal accuracy around 40 cm is achievable while the height accuracy was reported to be close to 20 cm for one beam and 10 cm for the other beam. The improvement in the geocoding accuracy is

highly beneficial for very localized small-scale deformation monitoring when, for instance, a certain part of an individual building is of interest rather than a large area surrounding it. PSI products with high localization accuracy can also be more easily compared with or incorporated into data from other sensors in comparison with the out-of-the-box PSI results.

Although the described methodology has been applied to high resolution X-band SAR data, the concept of geodetic corrections using SAR imaging geodesy is applicable to medium resolution products, such as the ones from the Sentinel-1 C-band SAR images, as well. The main limitation of the proposed method for lower resolution SAR products is related to the GCP identification part. In this case, the SCR of good GCP candidates should be high enough so that they can be visible in medium and low resolution SAR images. We are currently investigating this issue on Sentinel-1 images of urban areas. It is also noteworthy that the explained method is mostly suitable for urban areas where a large number of PS and GCPs are expected. However, in rural areas close to road networks, where lamp poles and traffic lights are also available, the exact same workflow explained in this paper can be used for geocoding error correction.

Future studies will focus on the integration of SAR-GCPs inside the PSI processing chain. Furthermore, more sophisticated and objective comparison strategies should be introduced for validation of PSI results with respect to accurate DSMs.

Abbreviations

The following abbreviations are used in this manuscript:

ADI	Amplitude Dispersion Index
DEM	Digital Elevation Model
DLR	German Aerospace Center
DSM	Digital Surface Model
ECMWF	European Centre for Medium-Range Weather Forecasts
GCP	Ground Control Point
GNSS	Global Navigation Satellite Systems
IERS	International Earth Rotation and Reference Systems Service
IGS	International GNSS Service
InSAR	Interferometric SAR
LiDAR	Light Detection and Ranging
MAD	Median Absolute Deviation
PSI	Persistent Scatterer Interferometry
PS	Persistent Scatterer
PTA	Point Target Analysis
SAR	Synthetic Aperture Radar
SCR	Signal-to-Clutter-Ratio
SGP	SAR Geodesy Processor
SLC	Single Look Complex
SRTM	Shuttle Radar Topography Mission
TEC	Total Electron Content
UTM	Universal Transverse Mercator

References

1. Ferretti, A.; Prati, C.; Rocca, F. Permanent scatterers in SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 8–20. [CrossRef]
2. Kampes, B.M. Radar Interferometry: Persistent Scatterer Technique. In *Number v. 12 in Remote Sensing and Digital Image Processing*; Springer: Dordrecht, The Netherlands, 2006.
3. Zebker, H.; Villasenor, J. Decorrelation in interferometric radar echoes. *IEEE Trans. Geosci. Remote Sens.* **1992**, *30*, 950–959. [CrossRef]
4. Gernhardt, S.; Bamler, R. Deformation monitoring of single buildings using meter-resolution SAR data in PSI. *ISPRS J. Photogramm. Remote Sens.* **2012**, *73*, 68–79. [CrossRef]
5. Gernhardt, S.; Adam, N.; Eineder, M.; Bamler, R. Potential of very high resolution SAR for persistent scatterer interferometry in urban areas. *Ann. GIS* **2010**, *16*, 103–111. [CrossRef]
6. Schwabisch, M. A fast and efficient technique for SAR interferogram geocoding. In Proceedings of the 1998 IEEE International Geoscience and Remote Sensing (IGARSS), Seattle, WA, USA, USA, 6–10 July 1998; Volume 2, pp. 1100–1102. [CrossRef]
7. Hanssen, R.F. Radar Interferometry: Data Interpretation and Error Analysis. In *Number v. 2 in Remote Sensing and Digital Image Processing*; Kluwer Academic: Dordrecht, The Netherlands; Boston, MA, USA, 2001.
8. Eineder, M. Efficient simulation of SAR interferograms of large areas and of rugged terrain. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 1415–1427. [CrossRef]
9. Eineder, M.; Minet, C.; Steigenberger, P.; Cong, X.Y.; Fritz, T. Imaging Geodesy—Toward Centimeter-Level Ranging Accuracy With TerraSAR-X. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 661–671. [CrossRef]
10. Chang, L.; Hanssen, R.F. Detection of cavity migration and sinkhole risk using radar interferometric time series. *Remote Sens. Environ.* **2014**, *147*, 56–64. [CrossRef]
11. Yang, M.; Dheenathayalan, P.; Chang, L.; Wang, J.; Lindenbergh, R.; Liao, M.; Hanssen, R. High-precision 3D geolocation of persistent scatterers with one single-Epoch GCP and LIDAR DSM data. In *Proceedings of the Living Planet Symposium*; European Space Agency: Paris, France, 2016; Volume 740, p. 398.
12. Bateson, L.; Novali, F.; Cooksley, G. *Terrafirma User Guide: A Guide to the Use and Understanding of Persistent Scatterer Interferometry in the Detection and Monitoring of Terrain-Motion*; Technical Report 19366/05/I-EC; ESA: Paris, France, 2010.
13. Gernhardt, S.; Cong, X.Y.; Eineder, M.; Hinz, S.; Bamler, R. Geometrical Fusion of Multitrack PS Point Clouds. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 38–42. [CrossRef]
14. Gernhardt, S. High Precision 3D Localization and Motion Analysis of Persistent Scatterers Using Meter-Resolution Radar Satellite Data. Ph.D. Thesis, Technische Universität München, München, Germany, 2012.
15. Cong, X. SAR Interferometry for Volcano Monitoring: 3D-PSI Analysis and Mitigation of Atmospheric Refractivity. Ph.D. Thesis, Technische Universität München, München, Germany, 2014.
16. Besl, P.; McKay, N.D. A method for registration of 3-D shapes. *IEEE Trans. Pattern Anal. Mach. Intell.* **1992**, *14*, 239–256. [CrossRef]
17. Zhu, X.X.; Montazeri, S.; Gisinger, C.; Hanssen, R.F.; Bamler, R. Geodetic SAR Tomography. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 18–35. [CrossRef]
18. Gisinger, C.; Balss, U.; Pail, R.; Zhu, X.X.; Montazeri, S.; Gernhardt, S.; Eineder, M. Precise Three-Dimensional Stereo Localization of Corner Reflectors and Persistent Scatterers With TerraSAR-X. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1782–1802. [CrossRef]
19. Zhu, X.X.; Bamler, R. Very high resolution spaceborne SAR tomography in urban environment. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 4296–4308. [CrossRef]
20. Balss, U.; Runge, H.; Suchandt, S.; Cong, X.Y. Automated extraction of 3-D Ground Control Points from SAR images—An upcoming novel data product. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 5023–5026. [CrossRef]
21. Montazeri, S.; Gisinger, C.; Eineder, M.; Zhu, X.X. Automatic Detection and Positioning of Ground Control Points Using TerraSAR-X Multiaspect Acquisitions. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 2613–2632. [CrossRef]
22. Van Leijen, F. Persistent Scatterer Interferometry Based on Geodetic Estimation Theory. Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 2014.

23. Cumming, I.G.; Wong, F.H.C. *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*; Artech House Remote Sensing Library, Artech House: Boston, MA, USA, 2005.

24. Yoon, Y.; Eineder, M.; Yague-Martinez, N.; Montenbruck, O. TerraSAR-X Precise Trajectory Estimation and Quality Assessment. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1859–1868. [CrossRef]

25. Hackel, S.; Montenbruck, O.; Steigenberger, P.; Balss, U.; Gisinger, C.; Eineder, M. Model improvements and validation of TerraSAR-X precise orbit determination. *J. Geod.* **2017**, *91*, 547–562. [CrossRef]

26. Cong, X.Y.; Balss, U.; Eineder, M.; Fritz, T. Imaging Geodesy—Centimeter-Level Ranging Accuracy with TerraSAR-X: An Update. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 948–952. [CrossRef]

27. Balss, U.; Cong, X.Y.; Brcic, R.; Rexer, M.; Minet, C.; Breit, H.; Eineder, M.; Fritz, T. High precision measurement on the absolute localization accuracy of TerraSAR-X. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 1625–1628. [CrossRef]

28. Balss, U.; Gisinger, C.; Cong, X.Y.; Brcic, R.; Hackel, S.; Eineder, M. Precise measurements on the absolute localization accuracy of TerraSAR-X on the base of far-distributed test sites. In Proceeding of the 10th European Conference on Synthetic Aperture Radar (EUSAR), Berlin, Germany, 3–5 June 2014.

29. Balss, U.; Breit, H.; Fritz, T.; Steinbrecher, U.; Gisinger, C.; Eineder, M. Analysis of internal timings and clock rates of TerraSAR-X. In Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium, Quebec City, QC, Canada, 13–18 July 2014; pp. 2671–2674. [CrossRef]

30. Misra, P.; Enge, P. *Global Positioning System: Signals, Measurements, and Performance*; Misra, P., Enge, P., Eds.; Jamuna Press: Warsaw, Poland, 2006.

31. Gisinger, C. Atmospheric Corrections for TerraSAR-X Derived from GNSS Observations. Master's Thesis, Technische Universität München, München, Germany, 2012.

32. Balss, U.; Gisinger, C.; Cong, X.Y.; Brcic, R.; Steigenberger, P.; Eineder, M.; Pail, R.; Hugentobler, U. High resolution geodetic earth observation with TerraSAR-X: Correction schemes and validation. In Proceedings of the 2013 IEEE International Geoscience and Remote Sensing Symposium, Melbourne, VIC, Australia, 21–26 July 2013; pp. 4499–4502. [CrossRef]

33. Petit, G.; Luzum, B. *IERS Conventions*; IERS Technical Note No. 36; Verlag des Bundesamts für Kartographie und Geodäsie: Frankfurt am Main, Germany, 2010.

34. Balss, U.; Gisinger, C.; Eineder, M. Measurements on the Absolute 2-D and 3-D Localization Accuracy of TerraSAR-X. *Remote Sens.* **2018**, *10*, 656. [CrossRef]

35. Eineder, M.; Balss, U.; Suchandt, S.; Gisinger, C.; Cong, X.; Runge, H. A definition of next-generation SAR products for geodetic applications. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium, Milan, Italy, 26–31 July 2015; pp. 1638–1641. [CrossRef]

36. OpenStreetMap Contributors. Planet Dump. 2017. Available online: https://www.openstreetmap.org (accessed on 18 August 2018).

37. Rousseeuw, P.J.; Hubert, M. Anomaly detection by robust statistics. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2018**, *8*, 1–14. [CrossRef]

38. Adam, N.; Kampes, B.M.; Eineder, M. The development of a scientific persistent scatterer system: Modifications for mixed ERS/ENVISAT time series. In Proceedings of the Envisat and ERS Symposium, Salzburg, Austria, 6–10 September 2004; pp. 1–9.

39. Adam, N.; Gonzalez, F.R.; Parizzi, A.; Brcic, R. Wide area Persistent Scatterer Interferometry: Current developments, algorithms and examples. In Proceedings of the 2013 IEEE International Geoscience and Remote Sensing Symposium, Melbourne, VIC, Australia, 21–26 July 2013; pp. 1857–1860. [CrossRef]

40. Rodriguez Gonzalez, F.; Bhutani, A.; Adam, N. L1 network inversion for robust outlier rejection in persistent Scatterer Interferometry. In Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 24–29 July 2011; pp. 75–78. [CrossRef]

41. Rodriguez Gonzalez, F.; Adam, N.; Parizzi, A.; Brcic, R. The Integrated Wide Area Processor (IWAP): A Processor for Wide Area Persistent Scatterer Interferometry. In *Proceedings of the ESA Living Planet Symposium 2013*; ESA: Edinburgh, UK, 2013; pp. 1–4.

42. Adam, N.; Kampes, B.M.; Eineder, M.; Worawattanamateekul, J.; Kircher, M. The Development of a Scientific Permanent Scatterer System. In Proceedings of the Joint ISPRS/EARSeL Workshop on High Resolution Mapping from Space 2003, Edinburgh, UK, 9–13 September 2003.

43. Montazeri, S. The Fusion of SAR Tomography and Stereo-SAR for 3D Absolute Scatterer Positioning. Master's Thesis, Delft University of Technology, Delft, The Netherlands, 2014.

44. Montazeri, S.; Zhu, X.X.; Eineder, M.; Bamler, R. Three-Dimensional Deformation Monitoring of Urban Infrastructure by Tomographic SAR Using Multitrack TerraSAR-X Data Stacks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6868–6878. [CrossRef]

A.3 Zhu, X.X., Montazeri, S., Gisinger, C., Hanssen, R.F., Bamler, R., 2016. Geodetic SAR Tomography. IEEE Transactions on Geoscience and Remote Sensing 54(1): 18-35

Geodetic SAR Tomography

Xiao Xiang Zhu, *Senior Member, IEEE*, Sina Montazeri, Christoph Gisinger,
Ramon F. Hanssen, *Senior Member, IEEE*, and Richard Bamler, *Fellow, IEEE*

Abstract—In this paper, we propose a framework referred to as "geodetic synthetic aperture radar (SAR) tomography" that fuses the SAR imaging geodesy and tomographic SAR inversion (TomoSAR) approaches to obtain *absolute* 3-D positions of a large amount of natural scatterers. The methodology is applied on four very high resolution TerraSAR-X spotlight image stacks acquired over the city of Berlin. Since all the TomoSAR estimates are relative to the same reference point object whose absolute 3-D positions are retrieved by means of stereo SAR, the point clouds reconstructed using data acquired from different viewing angles can be geodetically fused. To assess the accuracy of the position estimates, the resulting absolute shadow-free 3-D TomoSAR point clouds are compared with a digital surface model obtained by airborne LiDAR. It is demonstrated that an *absolute positioning* accuracy of around 20 cm and a meter-order *relative positioning* accuracy can be achieved by the proposed framework using TerraSAR-X data.

Index Terms—Absolute positioning, geodetic SAR tomography, geodetical fusion, SAR geodesy, SAR tomography, stereo SAR, synthetic aperture radar (SAR), TerraSAR-X.

I. INTRODUCTION

SPACEBORNE tomographic synthetic aperture radar (SAR) inversion (TomoSAR) [1]–[9] uses stacks of SAR images acquired at slightly different positions over a certain time period in a repeat–pass manner, like all other advanced InSAR techniques, such as persistent scatterer interferometry (PSI) [10]–[14], small baseline subset (SBAS) [15]–[17], SqueeSAR [18]–[20], and CEASAR [21]. They all aim at retrieving the 3-D position and the parameters of the undergoing motion of point, surface, and/or volumetric scatterers. Among them, TomoSAR is the only technique that can reconstruct nontrivial reflectivity profiles along the third native coordinate of SAR—elevation s—for each azimuth–range $(x - r)$ pixel. In particular, using stacked very high resolution (VHR) SAR images delivered by modern spaceborne SAR sensors, TomoSAR allows us to retrieve not only the most detailed 3-D geometrical shape but also the undergoing temporal motion of individual buildings and urban infrastructures in the centimeter or even millimeter scale [4], [5], [22], [23]. The resulting 4-D point clouds have a point (scatterer) density that is comparable to LiDAR. Experiments using TerraSAR-X high-resolution spotlight data stacks show that a scatterer density on the order of one million points per square kilometers can be achieved by TomoSAR [24]. However, similar to conventional InSAR and PSI, the elevation and deformation rates are estimated with respect to a previously chosen reference point that makes them *relative 3-D* estimates [25]–[27].

Another attractive feature of modern SAR sensors, in particular of TerraSAR-X and TanDEM-X, is the precise orbit determination and high geometrical localization accuracy. After compensating for the most prominent geodynamic and atmospheric error sources, the *absolute* 2-D (range and azimuth) positions of targets such as corner reflectors and persistent scatterers can be estimated to centimeter-level accuracy—a method called "SAR imaging geodesy" [28], [29]. Moreover, using two or more SAR observations acquired from different satellite orbits, their *absolute* 3-D positions can be retrieved by means of stereo SAR [30]. However, common scatterers that appear in SAR images acquired from different geometries, in particular from cross heading orbits, are very rare. This limits the application in 3-D absolute scattering positioning.

In this paper, we propose a framework referred to as "geodetic TomoSAR" that fuses the SAR image geodesy and TomoSAR approaches to obtain *absolute 3-D positions* of a large amount of natural scatterers. We work on four stacks of TerraSAR-X high-resolution spotlight images over the city of Berlin, among them two are acquired from ascending orbits and two from descending ones. First, tens of opportunistic (or natural) point scatterers that appear in all image stacks are manually identified. Their *absolute* 2-D SAR range and azimuth positions are calculated using imaging geodesy by compensating all the error sources, and their absolute 3-D positions are then calculated using stereo SAR. the most precisely localized point target is then chosen as reference point for the follow-on TomoSAR processing. Since the TomoSAR estimates are relative to the identical reference point whose absolute 3-D positions are known, the resulting point clouds are geodetically fused. Finally, to assess the position estimates, the resulting absolute 3-D TomoSAR point clouds are compared with a

digital surface model (DSM) obtained by airborne LiDAR. Experimental results demonstrate that the absolute positioning accuracy using TerraSAR-X is around 20 cm. The elevation estimation accuracy of TomoSAR depends on the number of the used images, SNR, baseline distribution, orbit height, and wavelength. In our experiments using TerraSAR-X, it is around 1 m [31].

II. STEREO SAR

A. SAR Imaging Geodesy for Absolute Ranging

The general imaging principle of SAR is based on the transmission of pulses and the reception of their echoes reflected back from the surface. Therefore, the location of a pixel in a radar image corresponds to the two-way round-trip time t_R (= range) as well as the mean time of transmission and reception t_A (= azimuth). Since the position of the satellite with respect to time is known from precise orbit determination, the azimuth is referred to an absolute location in 3-D space. The geometric distance R from this satellite position to the surface is obtained by scaling the two-way round-trip time with the velocity of light c.

In the case of a point scatterer, the two radar observations, i.e., the azimuth time and the geometric range, can be extracted from the focused SAR image through point target analysis (PTA), which yields the center coordinates of the scatterer's signature at subpixel level. If the errors present in this type of observations (atmospheric signal delays including ionospheric and tropospheric delays, geodynamic displacements such as solid Earth tides, continental drift, atmosphere pressure loading, ocean tidal loading, pole tides, ocean pole tides, and atmosphere tidal loading) are corrected by external models and the remaining unknown effects, e.g., time delays induced by cables and electronics, are calibrated for, the outcome are absolute 2-D radar observations.

For TerraSAR-X and TanDEM-X, this whole process has been mastered down to the 1–2 cm level [32]. It involves an accurate SAR processor to generate the focused SAR images, the PTA to extract the radar coordinates, the computation of the external corrections, and the geometrical calibration [32]. Since the process combines correction principles used in geodesy with SAR, we refer to it as SAR imaging geodesy [28]. The following provides a short summary on important elements of the process, and the reader interested in the details can find them in [28], [32].

In order to generate SAR images that are accurately focused in zero-Doppler geometry, the TerraSAR-X multimode SAR processor (TMSP) [33] avoids approximations often used to reduce the computational effort, e.g., the stop–go approximation, which assumes a static satellite during the transmission and reception of one pulse. Furthermore, effects such as the nonzero duration of the pulses are taken into account. For the external corrections, we distinguish between the signal propagation delays due to the atmosphere and the geodynamical effects (solid Earth tides, plate tectonics, ocean loading, atmospheric loading, etc.), causing a displacement of a target on ground. The geodynamic effects are considered by models following the conventions issued by the International Earth Rotation and Reference Systems Service (IERS) [34], whereas the atmosphere is corrected through path delays derived from complementary Global Navigation Satellite System (GNSS) observations. Thus, the concept of separating the atmospheric delay into a nondispersive part (usually called tropospheric delay—even if it involves also contributions from other atmosphere layers) and a dispersive part (ionospheric delay) which is well established in the field of GNSS [35] can be adapted to SAR.

It is worth to mention for medium-resolution sensors, such as Sentinel-1, the positioning capability is less precise than for high-resolution sensors. In fact, a Sentinel-1 artificial corner reflector needs to be 4.5 m large, which is unrealistic, in order to achieve 1-cm range accuracy. Shoebox-sized compact active transponders would be a convenient and highly demanded alternative [36].

Based on these thoroughly corrected 2-D observations, our stereo SAR approach allows the straightforward retrieval of absolute 3-D coordinates by combining sets of range and azimuth observations of a target in a joint parameter estimation.

B. Stereo SAR for Absolute 3-D Positioning of Corner Reflectors or Persistent Scatterers

The details as well as the validation of the stereo SAR method outlined in this section are given in [37]. Its analysis was carried out for TerraSAR-X and TanDEM-X data of our corner reflectors located at the geodetic observatories Wettzell (Germany) and Metsähovi (Finland). The reference coordinates of both reflectors are known from onsite geodetic surveying with accuracy better than 5 mm, and the comparison with the solution computed by stereo SAR showed differences at the 2–3-cm level [37].

The geometry of a SAR observation of a single point target at zero-Doppler location is given by the well-known range-Doppler equation system [38], i.e.,

$$|\mathbf{X}_S - \mathbf{X}_T| - R = 0 \tag{1}$$

$$\frac{\dot{\mathbf{X}}_S(\mathbf{X}_T - \mathbf{X}_S)}{|\dot{\mathbf{X}}_S||\mathbf{X}_T - \mathbf{X}_S|} = 0 \tag{2}$$

where $\mathbf{X}_s$ and $\dot{\mathbf{X}}_s$ are the position and velocity vectors of the sensor with respect to the azimuth time t_A, $\mathbf{X}_T$ denotes the unknown position vector of the target, and R is the observed range derived from the two-way round-trip time after compensation of all system and atmospheric delays. If the relationship between the sensor trajectory during an acquisition (position and velocity) and the azimuth time is expressed by an analytical model, e.g., polynomials, and introduced into the range-Doppler equations, the unknown target position $\mathbf{X}_T$ can be resolved in absolute 3-D by combining at least two acquisitions. In terms of geometry, this corresponds to the intersection of two or more circles that are perpendicularly oriented with respect to the sensor trajectories since the target is considered to be at zero-Doppler location (see Fig. 1).

The mathematical problem becomes overdetermined for two or more radar acquisitions because every acquisition i provides two equations (1) and (2) that relate the observations $t_{A,i}$ and $t_{R,i}$ with the three target coordinates $\mathbf{X}_T = [x\ y\ z]$. However, an optimal solution to this problem can be found by applying general least squares parameter estimation [37]. This allows a straightforward computation of absolute 3-D coordinates in the
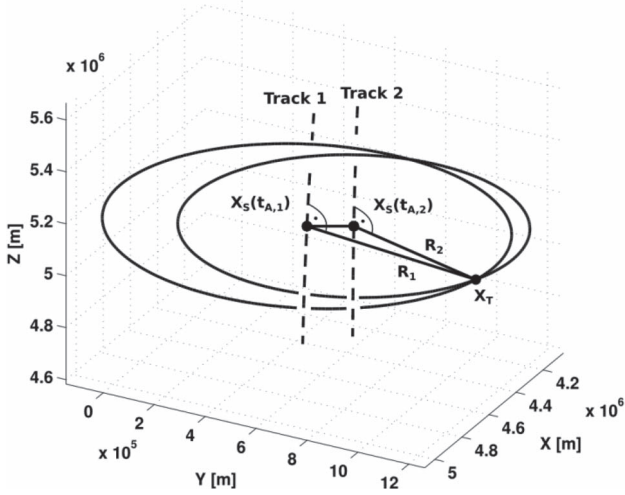
Fig. 1. Graphical representation of a stereo SAR acquisition for a target at zero-Doppler location. Note that due to the availability of repeat–pass data stacks, we use multiple images from the same geometry in stereo SAR configuration for a more accurate and robust absolute 3-D location estimation.

global reference frame of the satellite orbit, and all observations taken from orbit tracks with visibility on the target can be combined into a single estimate. In addition to the actual coordinate solution, the parameter estimation provides the corresponding standard deviations that allow the quality analysis of the solution. The standard deviations are influenced by the noise of the observations, the intersection geometry given by the separation of the orbit tracks (= stereo baseline), and the total amount of acquisition used to compute the solution.

In order to apply the stereo SAR method with TerraSAR-X in practice, the following steps have to be performed.

1) Focusing of the TerraSAR-X image by the TMSP. The step is fulfilled by ordering the TerraSAR-X L1B data, which provides not only the accurately focused single-look slant range complex (SLC) images but also extensive metainformation [39]. This metainformation includes the trajectory of the SAR antenna phase center in the International Terrestrial Reference Frame (ITRF) 2008, which is described by discrete position and velocity vectors (10-s sampling interval). Polynomials are used to interpolate the trajectory from the state vectors of the science orbit.

2) Locating the maxima of the point scatterer responses visible in the images taken from different orbit tracks by using PTA. The PTA algorithm must avoid aliasing and employ best practice interpolation.

3) Computation of the external radar corrections introduced in Section II-A according to the time and the observation geometry of every acquisition.

4) Stereo SAR processing using the range and azimuth observations of Step 2 corrected by the results of Step 3, as well as the polynomial trajectory models obtained in Step 1.

Note that Steps 3 and 4 have to be performed in iteration: An initial 3-D position is computed without correcting the observation data. This solution is introduced into the correction computation, which yields a first set of corrections. After

repeating this process one to two times, both the corrections and the position solution become stable, and the final result is achieved.

III. TomoSAR

TomoSAR, including SAR tomography and differential SAR tomography, uses stacks of SAR images taken from slightly different positions over a long period in a repeat–pass manner and uses the stacks to reconstruct the 3-D positions of coherent objects and their undergoing motion by means of spectral estimation. According to the scattering mechanism, the coherent targets, i.e., the signal, to be resolved can be categorized as discrete scatterers and volumetric scatterers. The reflectivity along elevation of discrete scatterers can be characterized by several δ-functions, i.e., the signal can be described by a deterministic model with a few parameters. Volumetric scatterers have a continuous backscatter profile associated with completely random scattering phases, i.e., the signal can only be described by stochastic models. Our target application is urban infrastructure monitoring, i.e., the resolution of discrete scatterers with motion. For tomographic SAR reconstruction of distributed scatterers, the readers are recommended to consult [40], [41].

Among various TomoSAR system models, differential SAR tomography was originally proposed in [2] for estimating linear motion of multiple scatterers inside a pixel. Motion, however, is often nonlinear (periodic, accelerating, stepwise, etc.). Therefore, conventional differential SAR tomography has been extended to estimate multicomponent nonlinear motion in [5] by means of the generalized "time warp" method. It rewrites the D-TomoSAR system model to an $M + 1$-dimensional standard spectral estimation problem, where M indicates the user-defined motion model order and hence enables the motion estimation for all possible complex motion models. In this section, this generalized model will be briefly described.

The focused complex-valued measurement g_n at an azimuth–range pixel for the nth acquisition at time t_n ($n = 1, \ldots, N$) is [3]

$$g_n = \int_{\Delta s} \gamma(s) \exp\left(-j2\pi\left(\xi_n s + 2d(s, t_n)/\lambda\right)\right) ds \quad (3)$$

where $\gamma(s)$ represents the reflectivity function along elevation s with an extent of Δs, and $\xi_n = -2b_n/(\lambda r)$ is the spatial (elevation) frequency proportional to the respective aperture position (baseline) b_n, with λ being the wavelength and r being the range. $d(s, t_n)$ is the line-of-sight (LOS) motion as a function of elevation and time. The motion relative to the master acquisition may be modeled using a linear combination of M base functions $\tau_m(t_n)$ [5], i.e.,

$$d(s, t_n) = \sum_{m=1}^{M} p_m(s)\tau_m(t_n) \quad (4)$$

where $p_m(s)$ is the corresponding motion coefficient to be estimated. Later, we will show that $\tau_m(t_n)$ can be interpreted as a warped time variable if we choose the units of the coefficients

appropriately. The choice of the base functions depends on the underlying physical motion processes, e.g., linear, seasonal, step function, temperature history, or even measured GPS deformation series of ground control points.

Let us define the mth temporal frequency component at t_n as $\eta_{m,n} = 2\tau_m(t_n)/\lambda$. Then, (3) can be rewritten as an $M + 1$-dimensional Fourier transform of $\gamma(s)\delta(p_1 - p_1(s), \ldots, p_M - p_M(s))$, which is a delta line in the $M + 1$ elevation-motion parameter space, i.e., [5]

$$g_n = \int\limits_{\Delta p_M} \cdots \int\limits_{\Delta p_1} \int\limits_{\Delta s} \gamma(s)\delta\left(p_1 - p_1(s), \ldots, p_M - p_M(s)\right)$$
$$\cdot \exp\left(-j2\pi(\xi_n s + \eta_{1,n}p_1 + \cdots + \eta_{M,n}p_M)\right)$$
$$\times ds\, dp_1, \ldots, dp_M, \; n = 1, \ldots, N. \tag{5}$$

After discretizing (5) along s and motion parameter space, in the presence of noise ε, the discrete-TomoSAR system model can be written as

$$\mathbf{g} = \mathbf{R}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \tag{6}$$

where $\mathbf{g}$ is the measurement vector with N elements, $\boldsymbol{\gamma}$ is the reflectivity function along elevati on uniformly sampled along elevation s_l $(l = 1, \ldots, L)$ and motion parameter space p_{m,l_m} $(l_m = 1, \ldots, L_m)$. $\mathbf{R}$ is an irregularly sampled M-dimensional discrete Fourier transform mapping matrix sampled at ξ_n and $\eta_{m,n}$. In practice, $\mathbf{R}$ and $\boldsymbol{\gamma}$ are reshaped to 2-D matrices, with a dimension of $N \times (L \times L_1 \times \cdots L_M)$. This renders tomographic SAR inversion a higher dimensional spectral estimation problem that can be again solved by the well-established spectral estimation methods. For more details, the readers are recommended to consult [5].

In our test sites, the following two component motion base functions, i.e., M = 2, are assumed.

- *Linear motion*: $\eta_{1,n} = 2t_n/\lambda$, and the coefficient $p_1(s)$ stands for the LOS velocity (v) as a function of s.
- *Seasonal motion*: $\eta_{2,n} = 2\tau_2(t_n)/\lambda$ where $\tau_2(t_n) = \sin(2\pi(t_n - t_0))$ can be interpreted as a warped time variable modeling the seasonal movement evolving over time, and $p_2(s)$ stands for the amplitude (a) of the periodic motion; t_0 is the initial phase offset.

IV. Geodetic SAR Tomography

Here, we will introduce the proposed framework *geodetic SAR tomography*, which consists of four main steps, namely, identification of reference point candidates, absolute positioning of reference point candidates, TomoSAR processing, and fusion of geodetic point clouds. To make the procedure more accessible for the readers, we explain the framework together with practical examples.

A. Data Sets

In this paper, the investigated test site includes the central area of the city of Berlin, Germany. The available data set consists of four stacks of TerraSAR-X VHR spotlight images acquired with a range bandwidth of 300 MHz. The images have
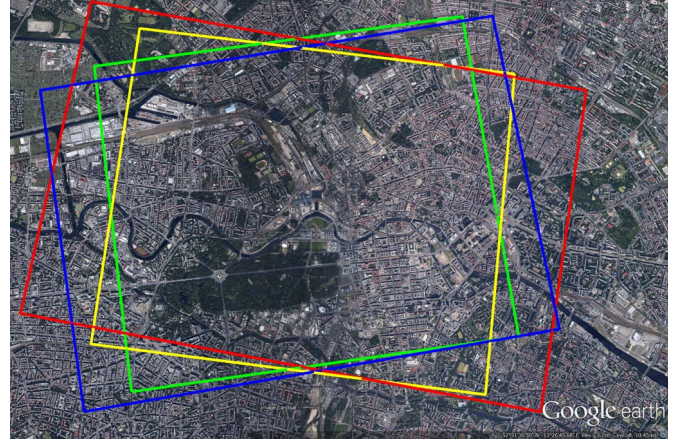

Fig. 2. Optical image of the city of Berlin (Google Earth). Rectangles mark the coverage of the four TerraSAR-X data stacks.

TABLE I
ACQUISITION PARAMETERS OF EACH STACK INCLUDING THE AVERAGE INCIDENCE ANGLE, THE FLYING DIRECTION OR AZIMUTH, THE TRACK TYPE AND THE NUMBER OF AVAILABLE IMAGES

Beam	Incidence angle	Heading angle	Track type	Nr. of Images
57	41.9°	350.3°	Ascending	102
85	51.1°	352°	Ascending	111
42	36.1°	190.6°	Descending	109
99	54.7°	187.2°	Descending	138

an azimuth resolution of 1.1 m and a slant-range resolution of 0.6 m covering an area of 10 km × 5 km. Two stacks are acquired from descending orbits with images recorded at 05:20 Coordinated Universal Time (UTC), and two stacks are acquired from ascending tracks with images recorded at 16:50 UTC. Fig. 2 shows the mean scene coverage of individual stacks overlaid on the optical image of Berlin. Furthermore, the details about the system parameters and properties of each stack are summarized in Table I.

Since Berlin is regularly monitored by TerraSAR-X, a large number of images are available for each stack ranging from 102 to 138 with a time span of approximately five years from February 2008 to March 2013 with the acquisition repeat cycle of 11 days. The four stacks consist of noncoregistered complex images.

In addition to the TerraSAR-X data sets, a point cloud of the test area obtained from aerial laser scanning is available (provided by "Land Berlin" and "Business Location Service," supported by "Europäischer Fonds für Regionale Entwicklung"). This data set is used to construct a DSM, which serves as a reference for the localization accuracy analysis of the TomoSAR point clouds. The LiDAR point cloud corresponding to the Reichstagsgebäude, Berlin, Germany, is visualized in Fig. 3.

B. Absolute Positioning of the Reference Point

As a characteristic of all interferometric SAR techniques, the height and deformation updates are estimated relative to a reference point. Although special care is taken to choose the reference point in an area close to the reference digital elevation model (DEM) and most plausibly not affected by
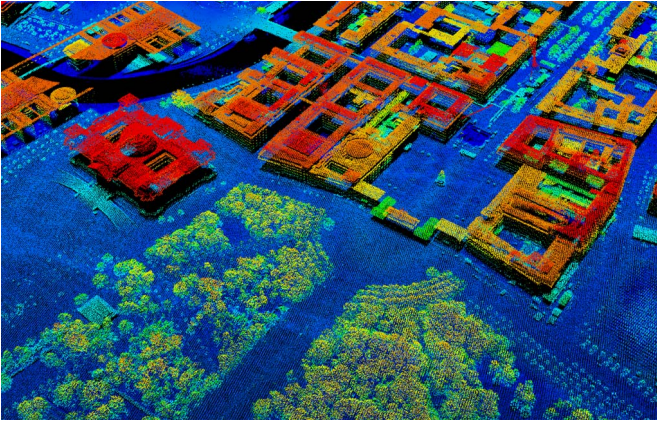
Fig. 3. LiDAR point cloud of a small area in Berlin in UTM coordinate system based on WGS 84 data. High absolute localization accuracy, on the order of 10 cm, makes the point cloud an excellent source to serve as a reliable reference surface model.

deformation, this however cannot be fully guaranteed and leads to complication in interpretation of the final results. Moreover, the exact 3-D position of the reference point is not known. Therefore, it is more likely that the final geocoded results will show offsets or even small scaling effect with respect to their true positions. The latter can also be problematic when it is desired to fuse the results obtained from different (same-heading or cross-heading) tracks in order to produce shadow-free point clouds. In this case, a lack of knowledge about the exact height of the reference point leads to inconsistencies between the point clouds [42]–[44].

In the following, the approach to select and absolutely localize natural point scatterers in SAR images is reported. The exact 3-D positions of the points are achieved with the stereo SAR method explained in Section II. Finally, the point with the highest quality is used as the reference point in the TomoSAR processing of all the four stacks which is extensively treated in Section IV-C.

1) Identification of Reference Point Candidates: The point targets, on which the 3-D stereo SAR reconstruction is performed, should have certain characteristics to be considered suitable reference point candidates. The criteria are the following.

- The target should be located in an isolated area.
- The target should be a single scatterer of high SNR, which is visible through the entire stack of SAR images.
- The target should be visible, at least, in two stacks of images acquired from different geometries.

The first condition should be satisfied in order to minimize the impact of interference caused by neighboring targets' responses on the reference candidate point. This is met by visual inspection of the mean amplitude image of each stack and the corresponding optical image of the scene to identify isolated targets.

The second condition ensures that the later tomographic reconstruction reaches in a higher 3-D localization as these points will also be served as reference points while forming differential observations. This is dealt with by calculation of the normalized amplitude dispersion index [10].

The third condition is vital from the radargrammetric point of view. Although, in optical imagery, selection of identical targets is commonly carried out with well-established algorithms such as scale-invariant feature transform (SIFT) [45] and Kanade–Lucas–Tomasi feature tracker [46], in SAR images, this cannot be done due to the existence of speckle [47]. For this reason and also considering the low number of candidates, in this paper, identical targets were selected manually by visual investigation of mean amplitude images of different stacks.

2) Absolute Positioning of the Reference Point: The outcome of the aforementioned procedure is eight point scatterers chosen from the central area of the city of Berlin (see Figs. 4 and 5). All of the scatterers are assumed the base of lamp posts located in the area, which typically have cylinder shapes that can reflect back radar signals from all illumination angles as shown in the right plot of Fig. 4. The scatterers are from three different types categorized based on the combination of geometry used for 3-D positioning, namely, ascending–ascending (AA), descending–descending (DD), and ascending–descending (AD). Fig. 6 shows the two different stereo orbit configurations that are used for point scatterers in Berlin. For each target, the time coordinates are retrieved by PTA from a number of SAR images in the stack. These time measurements are first corrected and then used in the zero-Doppler equations, outlined in Section II, to retrieve the 3-D coordinates. Table II gives an overview of data-take configurations, the time period within which the time coordinates were measured and the number of images used in PTA.

Among the candidates reported in Table II and visualized in Figs. 4 and 5, the one with the highest quality, i.e., the lowest 3-D standard deviation is selected as the reference point for TomoSAR processing of all the SAR image stacks. The stability of the results depends on the geometry of the observations, the number of observations, and the SNR of the targets. The geometrical configuration is the most important factor as for AD geometry the intersection occurs at almost $90°$ angle providing a well-conditioned system of equations. This effect can be clearly seen in Fig. 6(b), whereas a large baseline between the ascending and the descending acquisitions is achievable. On the other hand, AA or DD configurations [see Fig. 6(a)] result in a more ill-posed system due to a rather small baseline. In order to support the aforementioned discussion, the coordinate standard deviations of the scatterers are plotted in Fig. 7 (left subfigure). The horizontal axis consists of the names of the scatterers with subscripts denoting the geometry used for the 3-D positioning. The vertical axis describes the standard deviation values ranging from 1 to 9 cm, which are at least one order of magnitude better than the relative estimates achieved by repeat–pass InSAR. In addition, the graph demonstrates that the standard deviation values are lower in the x-direction. Moreover, as it was expected, the results from P_{AD1} and P_{AD2}, which are calculated from the cross-heading orbits, are more precise than the other points. It is worth mentioning that restricting the quality control of estimates solely based on the standard deviations may not be a reliable criterion. This is mainly due to the presence of covariance between the coordinate stochastics. Therefore, it is meaningful to analyze the error ellipsoid that is obtained

Fig. 4. (Left) Selected reference point candidates visualized as red dots in the optical image of Berlin (Google Earth). All of the candidates are assumed base of lamp posts. (Right) Photograph of one of such lamp posts in Berlin.
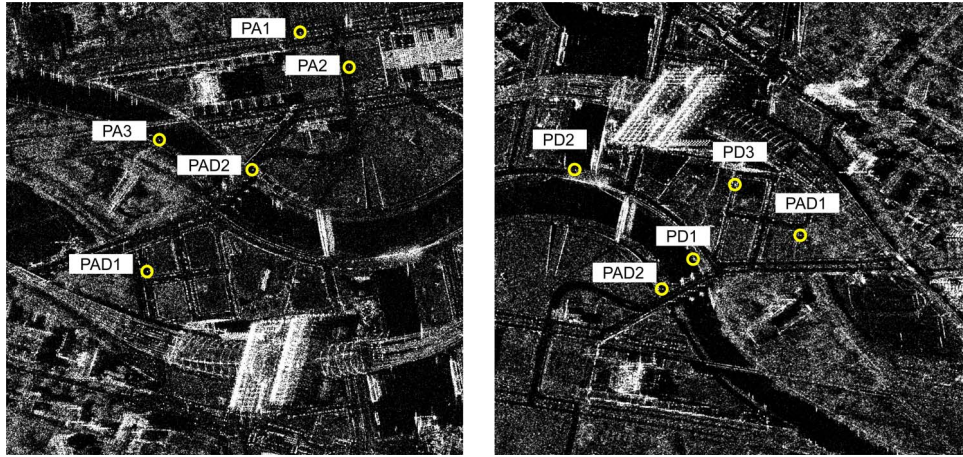

Fig. 5. Selected reference point candidates distinguished with yellow circles in the SAR images taken from the ascending and descending geometries with an incidence angle between 36° and 55°, as detailed in Table I.
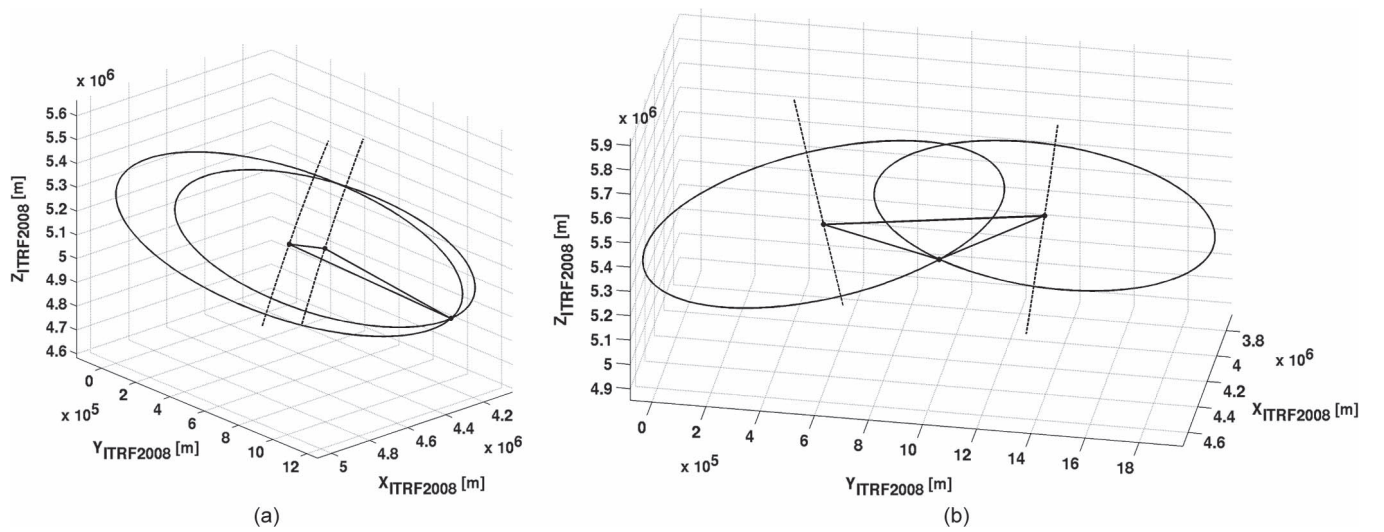

Fig. 6. Different orbit stereo configurations taken into account for 3-D scatterer reconstruction in Berlin. (a) Same heading orbits. (b) Cross-heading orbits.

by transformation of the posterior variance–covariance matrix of the estimates to the uncorrelated diagonal matrix of the eigenvalues based on eigenvector decomposition. In this case, the diagonal elements of the decomposed matrix represent the stochastics in the inherent SAR coordinate system. The mentioned approach was carried out on the variance–covariance

Scatterer	Geometry	Period	Nr. of Data-takes	Beams
P_{AD1}	AD	2008-2011	33	42,57
P_{AD2}	AD	2008-2011	30	42,57
P_{A1}	AA	2008-2011	22	57,85
P_{A2}	AA	2010-2012	9	57,85
P_{A3}	AA	2010-2012	6	57,85
P_{D1}	DD	2010-2011	11	42,99
P_{D2}	DD	2010-2011	8	42,99
P_{D3}	DD	2010-2011	11	42,99

matrix of each point, and the result is plotted in Fig. 7 (right subfigure) where σ_r, σ_a, and σ_s denote the ellipsoid axes that orient themselves with respect to the range, the azimuth, and the elevation of the SAR geometry, respectively. It is observed that, for the scatterers, which are localized from the same-heading geometries, the range component is minimal followed by azimuth, whereas the highest uncertainties are allocated to elevation components varying from 4 to 14 cm. However, it is seen that due to the almost optimum geometry configuration of cross-heading tracks, the elevation components of P_{AD1} and P_{AD2} have the smallest value better than 2 cm.

Based on the given discussion, the matter of configuration leads to discarding the point targets identified from the same-heading tracks narrowing the selection between P_{AD1} and P_{AD2}. Among them, P_{AD1} was selected as the reference point since it has slightly better precision, and it was also visible in all the four stacks. Fig. 8 shows the selected target in the mean amplitude images of one ascending and one descending TerraSAR-X spotlight image. The target is a lamp pole in a pedestrian area near the Berlin central station. Its absolute positions in the ITRF 2008 are as follows:

$$[X \ Y Z] = [3783630.014 \pm 0.010 \text{ m } 899035.0040$$
$$\pm 0.010 \text{ m } 5038487.589 \pm 0.011 \text{ m}].$$

It has to be emphasized that the listed 1-cm level uncertainties refer to the variance–covariance information provided by the position computation with stereo SAR. Thus, these values are a measure for the consistency of the stereo-based scatterer retrieval. It must be considered however that the satellite illuminates different sides of the lamp post from ascending and descending orbits. Regarding the absolute accuracy of the reference point, we expect a bias on the order of 20 cm, which depends on the diameter of the lamp post. We will analyze this possible bias in Section IV-D.

C. Tomographic Processing

After choosing the aforementioned absolutely geo-positioned scatterer as the reference point, the InSAR stacking and TomoSAR processing were done by the PSI-GENESIS [11] and Tomo-GENESIS system [48], [49] of the Remote Sensing Technology Institute of DLR, respectively.

The Tomo-GENESIS processing chain consists of three main steps, namely, preprocessing, tomographic processing, and fusion of point clouds. In this paper, mainly the first two steps are concerned, which are briefly outlined as follows. Furthermore, instead of geometrical fusion of point clouds as previously used in Tomo-GENESIS, a geodetical fusion method will be introduced.

1) Preprocessing: The processing starts from the stack of coregistered complex SAR images. The task of preprocessing is to estimate and remove the atmospheric phase screen (APS) of each image in the stack. The core feature of the preprocessing is *spatial difference*. It makes use of the assumption that APS is spatially slowly varying but highly uncorrelated from one image to another. Therefore, the estimates using spatially differential measurements should be "APS free." Such a method was already described in [50]. We customized it to adopt our problem [51].

The preprocessing procedures are described as follows: Images in the stack are downsampled if they are VHR. In the downsampled images stack, pixel pairs with spacing (arc) shorter than the atmospheric correlation length (typically recommended to be 2 km [52] in this paper, due to the high density of bright points, an even shorter distance of 250 m is chosen), are then selected and connected. Spatial differential measurements are calculated between the pixels in a pair. Then, we estimate the topography and motion parameters based on the single point scatterer phase model. The differential topography estimates is further integrated globally, and the topographic phase contribution is removed from each image. The remaining phase should be fairly flat. It consists of only the deformation phase, APS, and stochastic scattering phase. The unwrapped residual phase is first low-pass filtered in space to remove random noise, and high-pass filtered in time to eliminate deformation signals. The result in this step is the APS of each image on the network made up of selected pixels. After interpolation and upsampling, the APS-induced phase is removed for each image of the stack.

2) Tomographic Processing: After APS removal, tomographic processing is applied to each pixel of the stacked images aiming at the retrieval of the elevation and motion parameters of multiple scatterers inside one azimuth–range pixel. Here, the generalized time warp model with $M = 2$ is used; we estimate linear and periodic seasonal motion.

Depending on the applications, different algorithms can be chosen for tomographic reconstruction.

- *PSI*: PSI is a special case of TomoSAR that attempts to separate the following phase contributions: elevation of the point, deformation parameters (e.g., deformation rate and amplitude of seasonal motion), orbit errors, and tropospheric water vapor delay. This is done by assuming the presence of only a single scatterer in the pixel. This restriction brings the big advantage of computational efficiency. It is recommended for large-scale urban monitoring.
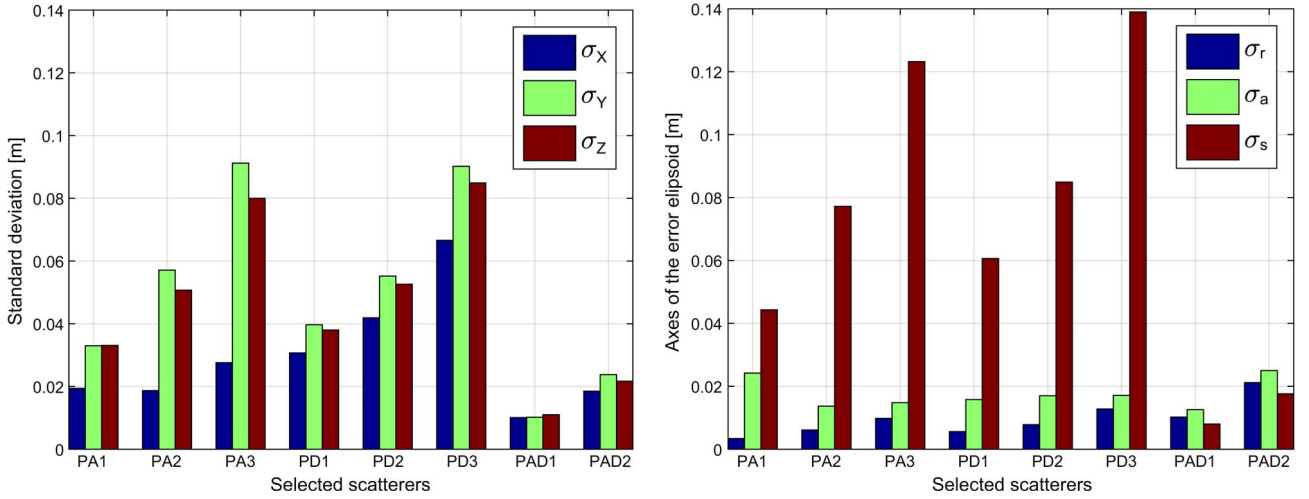
Fig. 7. Estimated coordinate stochastics for three components obtained from stereo SAR. The importance of geometry configuration on 3-D positioning is clearly seen as the scatterers localized from cross-heading tracks are more precise.
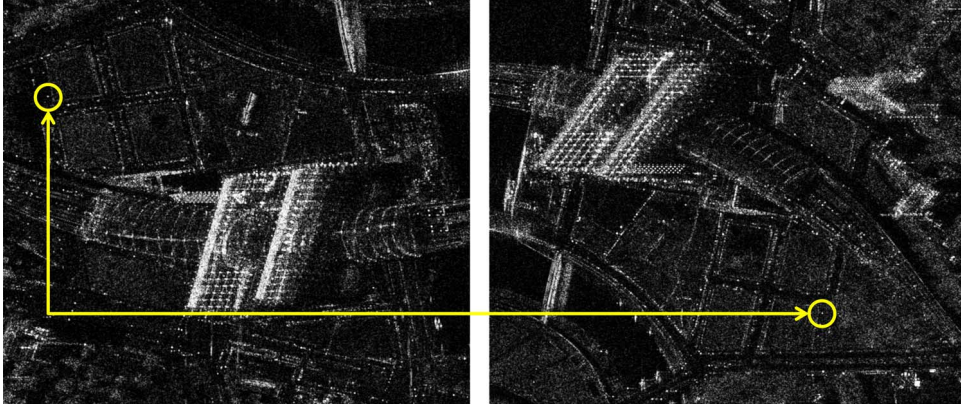

Fig. 8. Common target visible in all ascending and descending data stacks, which is selected as the reference point for the follow-on TomoSAR processing. The point can be observed as a bright dot inside the yellow circles.

- *Maximum detection (MD)*: MD stands for SVD-Wiener (linear MAP) reconstruction followed by peak detection and model order selection and final refinement of the amplitude and phase estimates [4]. This algorithm is computationally efficient and is not sensitive to irregular sampling. As a linear method, MD has almost no super-resolution capability. Taking account to its fast computation, MD is recommended if the native elevation resolution is sufficient for the application.

- *Scale-down by L_1 norm minimization, model selection, and estimation reconstruction (SL1MMER)*: The SL1MMER algorithm is proposed in [53]. It consists of three main steps: 1) a dimensionality scale down by L_1 norm minimization; 2) model selection; and 3) parameter estimation. In [24] and [54], this algorithm is demonstrated to give robust estimation with very high elevation resolution. In the relevant parameter range for TomoSAR, super-resolution factors of 1.5–25 (compared with the Rayleigh resolution unit) can be expected. SL1MMER can offer so far ultimate 4-D SAR imaging; however, it is computationally very expensive. Therefore, it is recommended for the monitoring of individual high rise buildings.

- *Integrated Approach*: Considering the high computational cost of TomoSAR, tomographic SAR inversion is integrated with PSI for operational use in [51]. With the integration of PSI, the processing is 30–50 times faster than SL1MMER alone, and still comparable results can be achieved. Thus, it gives a good compromise of the aforementioned three methods.

Since our aim is to demonstrate the framework of *geodetical SAR tomography*, the simple MD method, i.e., SVD-Wiener followed by model order selection and parameter estimation, is chosen for an efficient TomoSAR processing. Starting from SLCs, for an input data stack, the Tomo-GENESIS system retrieves the following information: number of scatterers inside each azimuth–range pixel, amplitude and phase, topography and motion parameters (e.g., linear deformation velocity and amplitude of thermal dilation induced seasonal motion) of each detected scatterer with respect to a reference point.

The final elevation estimates of two of the four data stacks using the same reference point are exemplified in Fig. 9. The elevation is color-coded. Fig. 9(a) and (b) refers to the result for beams 42 and 57, respectively. For each of the retrieved
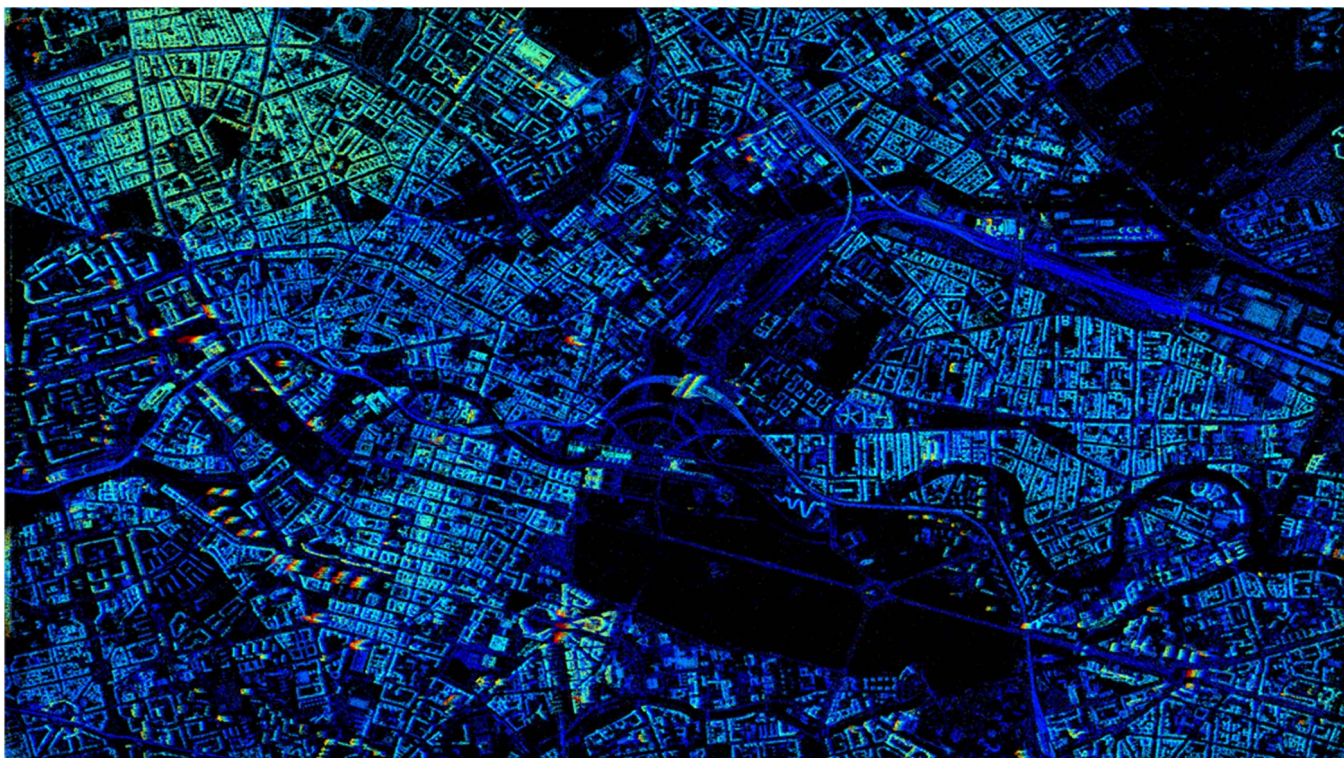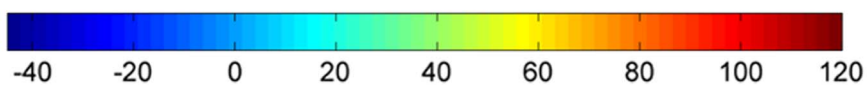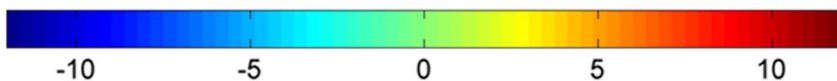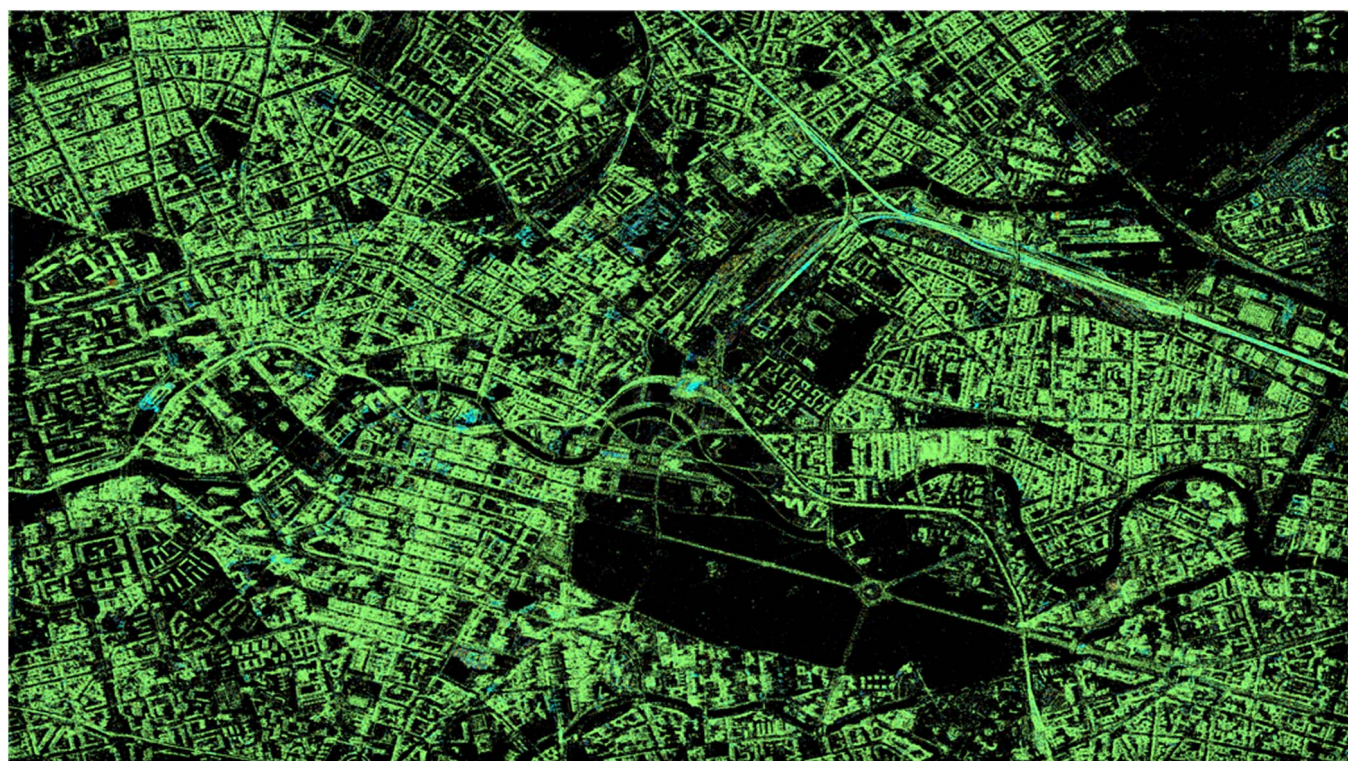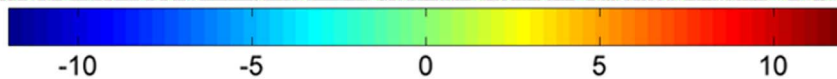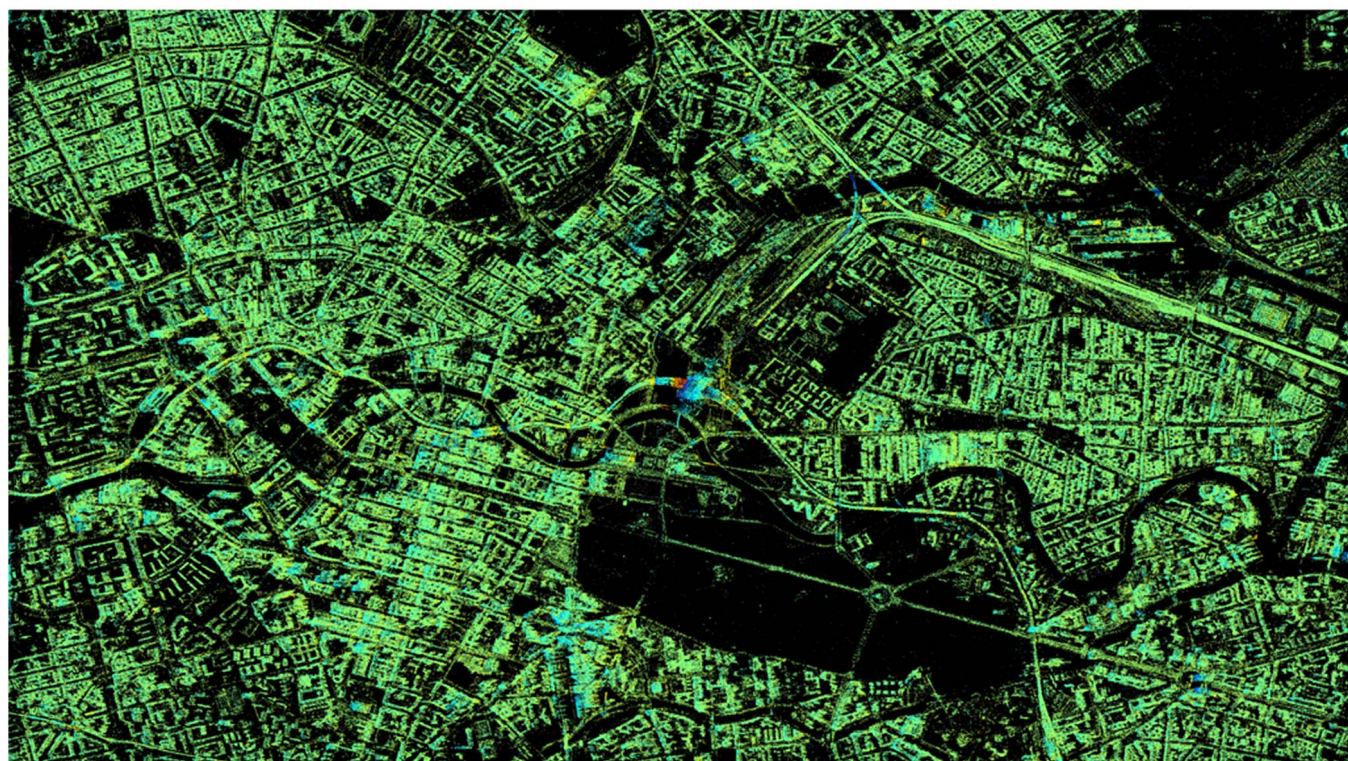
(a)
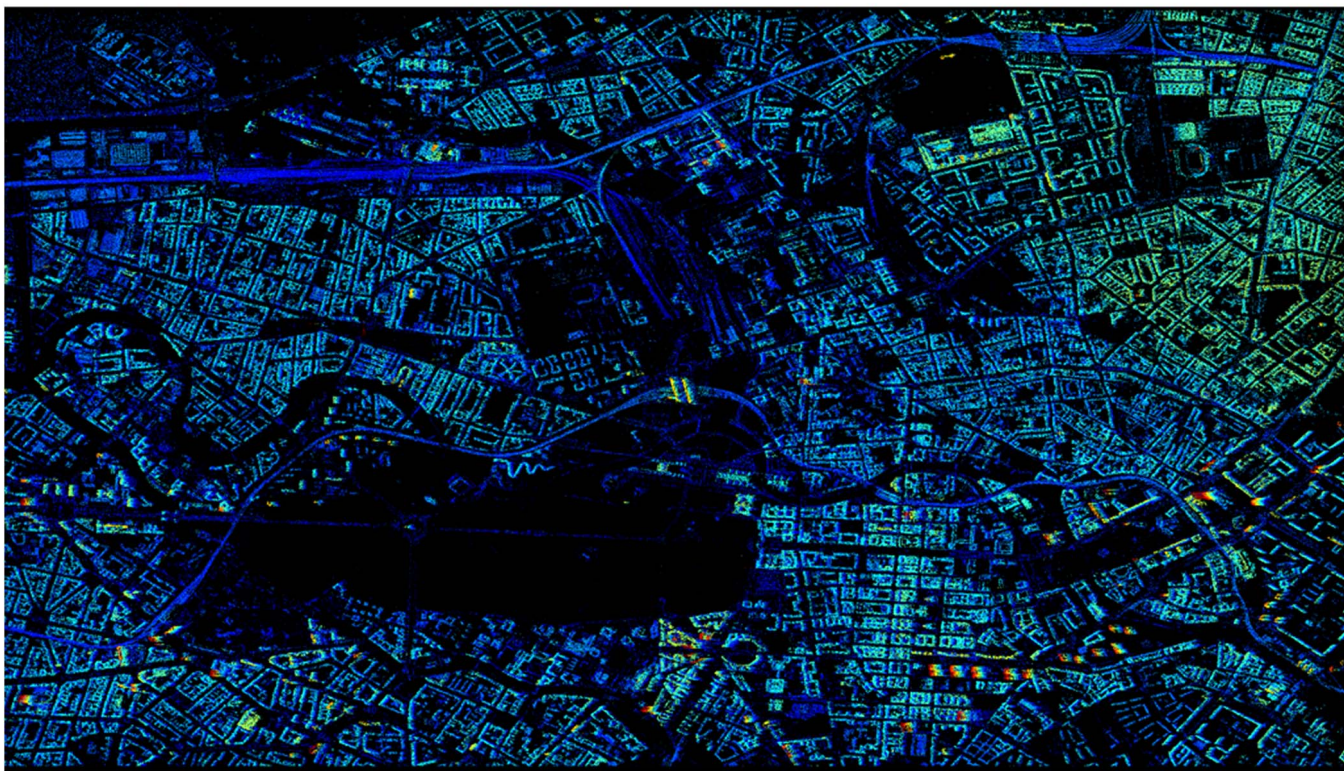
(b)

Fig. 9. TomoSAR results. Elevation estimates of two of the four stacks using the same reference point; elevation is color-coded [unit: m]. (a) Beam 42, descending. (b) Beam 57, ascending.

(a)

(b)

Fig. 10. TomoSAR results. Estimated LOS linear deformation rate (a) and amplitude of seasonal motion (b) of beam 42. Motion parameter is color-coded. (a) Linear deformation rate [mm/y]. (b) Amplitude of seasonal motion [mm].

scatterers/points, its undergoing amplitude of seasonal motion and linear deformation rate are also estimated.

Since, in this paper, mainly the absolute positions of these scatterers are concerned, the motion results will not be discussed further. To give the readers an impression of the estimated motion parameters, Fig. 10 shows the estimated LOS linear deformation (a) rate and (b) amplitude of seasonal motion of beam 42 as an example. It can be observed that Berlin is rather stable, i.e., there is no significant ground deformation pattern. Most of the buildings and other man-made urban infrastructures mainly undergo temperature changes and induced seasonal deformation with amplitude of up to 15 mm. Some railway sections, the buildings along them, and several buildings in construction undergo a linear subsidence with a rate of up to 8 mm/y. For a more meaningful analysis, LOS deformation estimates of different viewing angles need to be fused [55].

D. Geodetic Fusion of TomoSAR Point Clouds

The side-looking geometry of SAR sensors only allows for mapping the illuminated sides of buildings. In order to produce a shadow-free point cloud, coregistration of results from, at least, one set of cross-heading orbits is required. For PSI and TomoSAR whose estimates are relative, geocoded point clouds obtained from different acquisition geometries cannot be directly coregistered. This is mainly due to the offsets in the elevation direction that are caused by selection of reference points with unknown heights, during the processing. The coregistration task of two unstructured 3-D InSAR point clouds is referred to as point cloud fusion in the SAR community. In [42], a method for fusion of multitrack PSI results is proposed based on a least-squares matching scheme that minimizes the distances between assumed identical points of two point clouds. The method aims to estimate the offset between the identical points in the elevation direction. In [44], an alternative feature-based fusion algorithm is proposed, which is based on automatic detection and matching the so-called *L-shapes* of high rise buildings from InSAR point clouds. This method is computationally more efficient than the one introduced in [42] due to the reduced number of points in the matching step. Relevant work in the airborne research domain can be found in [56] and [57]. It is important to note that all mentioned existing methods perform the point cloud fusion *geometrically*. It can be shown that by merging the capabilities of the stereo SAR (see Section II) and TomoSAR (see Section III), it is possible to perform *geodetic* point cloud fusion.

In the framework of geodetical SAR tomography, TomoSAR processing is based on the selection of an identical reference point whose 3-D positions are retrieved by means of stereo SAR. After geocoding, it is therefore expected that the point clouds will be automatically fused without further manipulations. However, this is not true for the following reasons.

- Each point cloud is geocoded separately based on the corresponding noncorrected range and azimuth timing information. As a consequence, the geocoded coordinates
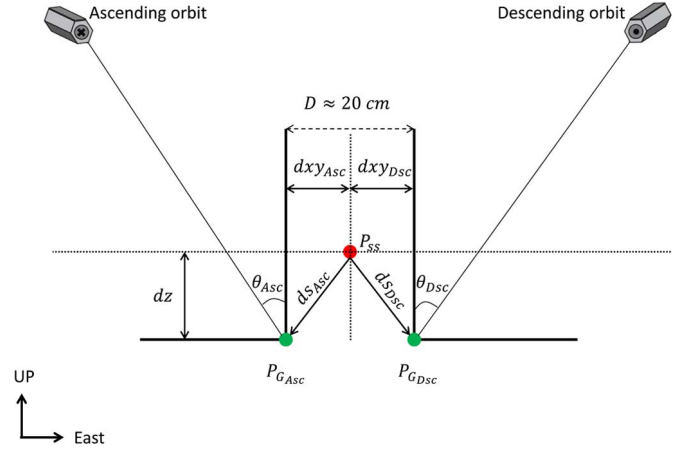

Fig. 11. Shift between the geocoded reference points in TomoSAR point clouds obtained from cross-heading geometries relevant to our specific case. The drawing is in the east-up plane while ascending and descending satellites fly perpendicular to the plane far away and toward the reader, respectively. For successful point cloud fusion, the coordinates resulted from stereo SAR (P_{SS}) should be shifted toward the expected true position of the reference point in each ascending and descending stack ($P_{G_{Asc}}, P_{G_{Dsc}}$). In order to do so, the corresponding shift vector in the elevation direction (ds_{Asc}, ds_{Dsc}) should be evaluated for each stack based on the known local incidence angles ($\theta_{Asc}, \theta_{Dsc}$) and the heading angles of the satellites ($\alpha_{Asc}, \alpha_{Dsc}$).

of the reference point, in each stack, show offsets with respect to the reference point coordinates obtained from stereo SAR, as well as to the geocoded reference point coordinates of other stacks.

- Scatterers visible in SAR images acquired in urban environment from both ascending and descending orbits, which are assumed identical in stereo SAR processing, are often lamp posts [37]. The SAR illuminates different sides of the lamp post from ascending and descending orbits. This means the identical scatterer assumption is not fully valid. Under the assumption that these two points are identical in 3-D, the reference point coordinates obtained from stereo SAR is situated on the body between the two sides of the lamp post. The coordinate offsets of the true reference points of individual stacks and the results obtained from stereo SAR depend on the diameter of the lamp post and incidence angles of each stack.

In order to compensate for the mentioned offsets, corrections are further required. For successful fusion of TomoSAR point clouds obtained from different viewing geometries, the coordinate shifts between the geocoded coordinates of the reference point in each stack and the reference point coordinates obtained from stereo SAR should be resolved.

Fig. 11 describes the problem stated earlier related to the nonmodeled diameter of the lamp post. The drawing is depicted in the (approximately) east-up plane where it is assumed that the ascending (left) and descending (right) satellites fly away and toward the reader, respectively. The diameter of the lamp post (D) is approximately 20 cm. The red dot denotes the approximate position of the reference point whose coordinates are retrieved with the stereo SAR method P_{SS}. The two green dots indicate the true positions on which the geocoded reference points in the point clouds obtained from ascending ($P_{G_{Asc}}$) and the descending ($P_{G_{Dsc}}$) tracks should be located. In order

Fig. 12. Optical image of Federal Intelligence Building (Google Earth). The part in the red ellipse is investigated later to compare the fusion results before and after applying the coordinate corrections.

to compensate the offset between the two cross-heading point clouds, caused by the diameter of the lamp post, for each of the ascending and descending stacks, $\mathbf{P}_{SS}$ should be shifted to the corresponding illuminated base of the lamp post by evaluation of the shift vectors $\mathbf{ds}_{Asc}, \mathbf{ds}_{Dsc}$ in the elevation direction. With the knowledge of the local incidence angle of each beam at the location of the lamp post $(\theta_{Asc}, \theta_{Dsc})$ and the satellites heading angles $(\alpha_{Asc}, \alpha_{Dsc})$, each component of the shift vectors can be calculated as follows:

$$dz = \frac{D \cdot \tan(\theta_{Asc}) \cdot \tan(\theta_{Dsc})}{\tan(\theta_{Asc}) + \tan(\theta_{Dsc})} \tag{7}$$

$$dx_{Asc} = dxy_{Asc} \cdot \cos(\alpha_{Asc}) \tag{8}$$

$$dy_{Asc} = -dxy_{Asc} \cdot \sin(\alpha_{Asc}) \tag{9}$$

$$dx_{Dsc} = dxy_{Dsc} \cdot \cos(\alpha_{Dsc}) \tag{10}$$

$$dy_{Dsc} = -dxy_{Dsc} \cdot \sin(\alpha_{Dsc}). \tag{11}$$

$dxy_{Asc} = dz \cdot \cot(\theta_{Asc})$ and $dxy_{Dsc} = dz \cdot \cot(\theta_{Dsc})$. Shift in the upward direction (dz) is equal for all the stacks and therefore can be evaluated using a least squares adjustment based on different combinations of the ascending and descending tracks. Afterward, the individual horizontal shifts (dxy_{Asc}, dxy_{Dsc}) are calculated for each stack, and with the known azimuth angles, they are projected into the east and north directions $(dx_{Asc}, dy_{Asc}, dx_{Dsc}, dy_{Dsc})$. The shift vectors thus are formed as

$$\mathbf{ds}_{Asc} = [dx_{Asc}, dy_{Asc}, dz_{Asc}]^T \tag{12}$$

$$\mathbf{ds}_{Dsc} = [dx_{Dsc}, dy_{Dsc}, dz_{Dsc}]^T. \tag{13}$$

Subsequently, $\mathbf{P}_{SS}$ is shifted to the position of $\mathbf{P}_{G_{Asc}}$ or $\mathbf{P}_{G_{Dsc}}$

that is dependent on the acquisition geometry of the stack

$$\mathbf{P}_{G_{Asc}} = \mathbf{P}_{SS} - \mathbf{ds}_{Asc} \tag{14}$$

$$\mathbf{P}_{G_{Dsc}} = \mathbf{P}_{SS} - \mathbf{ds}_{Dsc}. \tag{15}$$

The remaining errors of the geocoding is compensated by evaluating the difference between the geocoded reference point coordinates of each stack and the corresponding $\mathbf{P}_{G_{Asc}}$ or $\mathbf{P}_{G_{Dsc}}$. Finally, for each stack, the unique difference vector is added to the coordinates of all the scatterers to produce four absolutely localized corrected point clouds.

Compensation for the coordinate shifts between the geocoded reference points and true position of the stereo SAR results allows for seamless geodetic fusion of TomoSAR point clouds. These corrections have been applied to each point cloud separately. To confirm that the corrections are necessary, a small test site including the Federal Intelligence Service (BND) building in Berlin is chosen to compare the fusion results before and after applying the coordinate corrections. Fig. 12 shows the optical image of the building. The red ellipse marks the building section that is investigated in Fig. 13.

In Fig. 13, the results from the ascending stacks are visualized in blue, and the descending point clouds are shown in red. In the noncorrected fusion (left), the black arrow represents the shift available between the same heading tracks. Moreover, the black ellipse highlights that the result from descending stacks (red) does not match with the building fraction captured from ascending stacks (blue). The right figure illustrates the fused point clouds by applying all aforementioned corrections. The good match of all four point clouds confirms the effectiveness of the proposed fusion strategy.

Fig. 14 illustrates the fusion of two ascending and two descending absolute TomoSAR point clouds in 2-D over the city of Berlin. The coordinates are expressed in Universal Transverse Mercator (UTM) coordinate system. It is seen that the point clouds are reasonably overlaid on each other after applying the corrections. This part is finalized by 2-D and 3-D visualizations of the fused point cloud of the central urban area of Berlin illustrated in Fig. 15. The absolute point clouds are plotted in the UTM coordinate system, and the height of each scatterer is color-coded with respect to the WGS84 ellipsoid.

The coverage of the test site is approximately 10 km × 5 km and the number of captured absolute positioned scatterers is 63 million. It is clearly seen that the fusion of TomoSAR point clouds obtained from different geometries allows for highly detailed 3-D mapping of the city. It is worth to mention that the aforementioned point density is obtained using the computationally efficient MD estimator, processing the same data sets using more expensive algorithms, e.g., SL1MMER, will even lead to a significantly higher point density [54]. This aspect is however outside the scope of this paper.

V. LOCALIZATION ACCURACY ANALYSIS OF THE FUSED TOMOSAR POINT CLOUD

The localization accuracy of the fused TomoSAR point cloud is assessed by comparing the results with an accurate DSM calculated from a point cloud obtained from aerial laser scanning
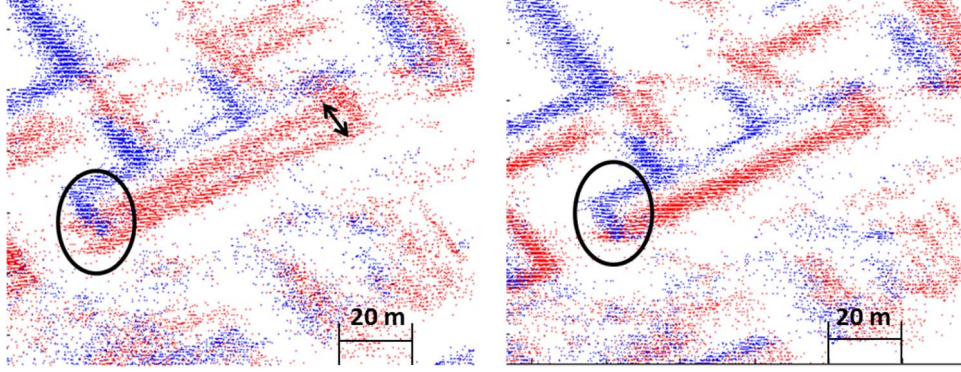
Fig. 13. Comparison between the fusion results before (left) and after (right) applying the reference point coordinate correction. The result from the ascending stacks is visualized in blue, and the descending point clouds are shown in red. The noncorrected fusion (left) includes certain offsets between the results from same-heading tracks (the black arrow) and wrong intersection of different building fractions captured from cross-heading tracks (the back ellipse).
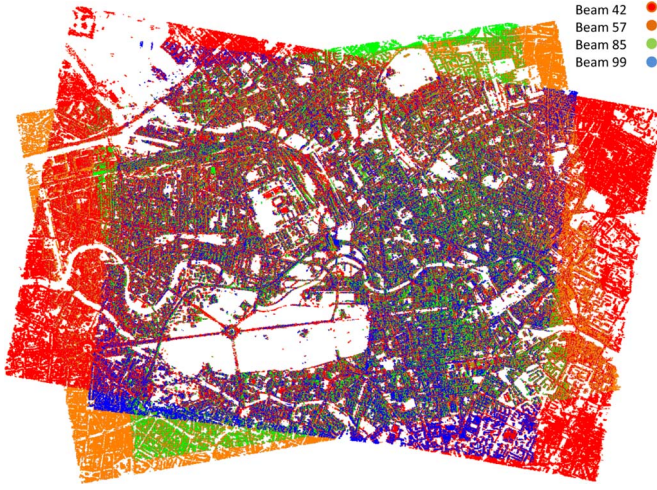

Fig. 14. Fusion result of two ascending and two descending tracks over the city of Berlin. The point clouds are absolutely localized after correcting the geocoded coordinates of the reference point and are geodetically fused.

characterized with a large number of data points and high absolute geolocalization accuracy on the order of 10 cm. This allows for a quantitative analysis on the positioning accuracy of the TomoSAR point cloud.

To check the overall accuracy of the point cloud, Fig. 16(a) shows the fused TomoSAR point cloud of a small area of Berlin overplotted onto the corresponding area in LiDAR DSM. The heights of the scatterers are color-coded with blue to red values indicating lower to higher heights. The DSM is plotted in gray for better visualization purposes. The good fit of the TomoSAR point cloud on the DSM is visually observable. For a closer speculation, a cross section through the buildings, marked with a white rectangle in Fig. 16(a), is visualized in Fig. 16(b). The figure shows a slice in the xz plane with the height values color coded. In order to validate the results illustrated in Fig. 16, the accuracy analysis is carried out for horizontal and vertical directions separately as reported in the following.

A. Horizontal Accuracy Analysis

The optimum way to assess the accuracy of the fused To-moSAR point cloud is a point-wise comparison with respect to the LiDAR point cloud. However, this is not feasible as LiDAR sensors map the surface with a nadir looking angle, whereas SAR sensors capture the scene from a side-looking geometry. The difference in acquisition geometry leads to different mapping of the same object and therefore complicates the comparison.

In this paper, the horizontal assessment is carried out by evaluating the mean façade points from the TomoSAR point cloud with respect to the extracted corresponding façade line from the LiDAR DSM.

- Building façades from the LiDAR DSM are estimated as follows: Based on the top view extent of the building, an area is cropped, which contains the desired building fraction. Centered on each point inside the cropped scene, a vertical cylinder with the radius of 2 m is considered. Inside the cylinder, the height variance is calculated. Points with height variances higher than a threshold are labeled as façade points. The building façade surface is assumed vertical. Therefore, at the final step, the footprint of the façade in 2-D, i.e., on the ground plane, is estimated by fitting a line to the façade points using reweighted least squares with a bisquare weighting function.

- For the TomoSAR point cloud, points belonging to a specific façade are extracted using the algorithms proposed in [58] and [59]: First, the scatterer density in the horizontal plane is estimated in adaptive windows varying dependent on the orientation of the façades. Based on the estimated point density and local normal directions, points belonging to individual building façades are extracted.

The perpendicular distance between the façade points extracted from TomoSAR point clouds and the corresponding façade lines estimated from LiDAR DSM is calculated. The mean value of the deviations is regarded as the horizontal bias between the TomoSAR point cloud and the LiDAR DSM.

B. Vertical Accuracy Analysis

The vertical accuracy is analyzed by comparing identical flat grounds mapped in LiDAR DSM and in the fused TomoSAR
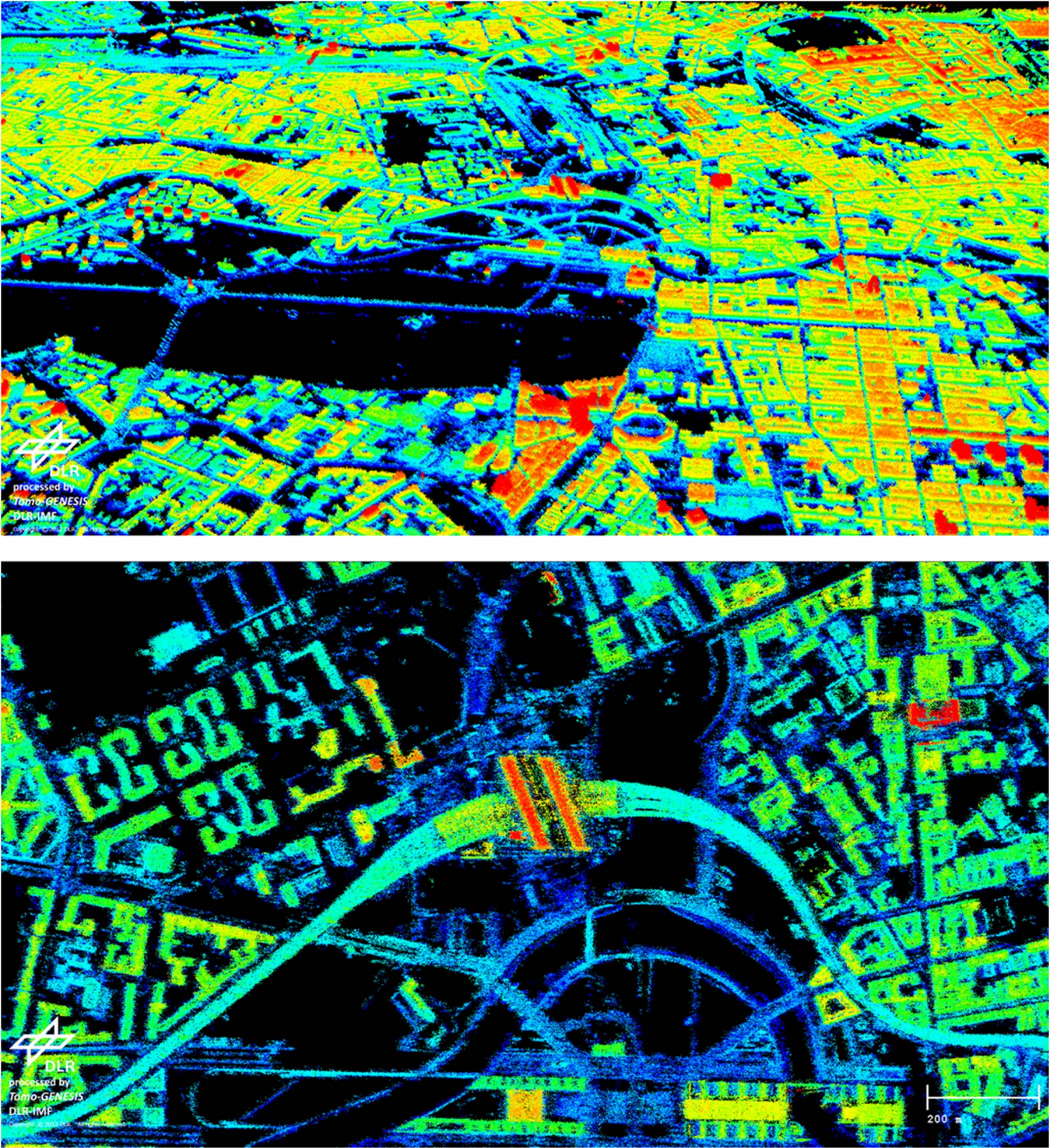
Fig. 15. Three-dimensional absolutely positioned TomoSAR point clouds in 3-D (top) and 2-D (bottom). The absolute height values are color-coded and range between 70 m to 110 m. Clearly, the fusion of multitrack point clouds allow for a very detailed representation of the city where most of the structures can be easily recognized.

point cloud. An identical patch is selected from both the TomoSAR point cloud and the reference surface. Inside the patch, the height deviation of scatterers from the TomoSAR point cloud with respect to the flat scene in the LiDAR DSM is evaluated by calculating the root mean square error.

C. Test Sites and Discussion

Different test sites from Berlin are chosen to validate the positioning accuracy of the fused TomoSAR point cloud. For horizontal analysis, two different buildings are selected, and for each of them, the extracted façade in the LiDAR DSM is
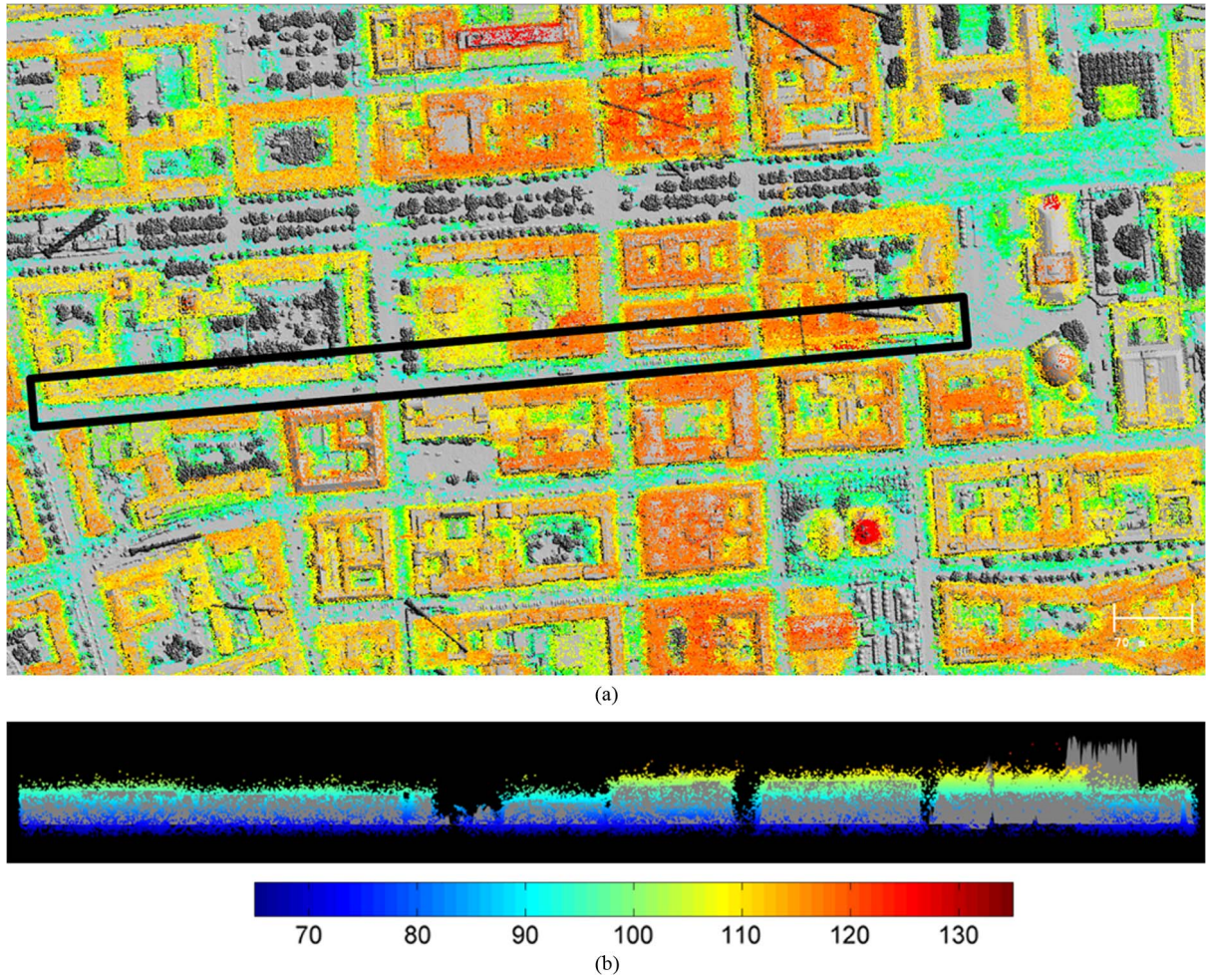
Fig. 16. (a) Fused TomoSAR point cloud overplotted onto the reference DSM. The DSM is plotted in gray for better visualization purposes. The height of the TomoSAR point cloud is color coded. It is seen that the scatterers are fitted to the corresponding building parts on the DSM as no large deviations are visible from the plot. The cross section in the xz plane visualized in (b) confirms the mentioned statement. [unit: meters].

compared with the detected façade points in the TomoSAR point cloud. Fig. 17 shows the findings of the first test site including a high-rise building in Postdamer Platz located at a distance of approximately 2 km from the reference point. Fig. 17(a) shows the optical image of the area where the investigated façade is marked with a white rectangle. Fig. 17(b) and (c) shows the mapped area in the LiDAR DSM and the TomoSAR point cloud in the UTM coordinate system with the facade distinguished with a white rectangle and a black rectangle, respectively. Façade points from the LiDAR DSM are extracted based on a threshold of 20 m on the height standard deviation values estimated in a vertical cylinder with the radius of 2 m around each point. The façade footprint, i.e., a line on the ground plane, is then the estimated. Fig. 17(d) shows the LiDAR façade points, color-coded based on the height variance values, and the estimated façade line in red. Façade points in the TomoSAR point cloud are approximated by applying a threshold of 13 points/m² on the scatterer density estimates. Moreover, the TomoSAR point cloud is clipped within height values of 80 to 160 m. The latter is carried out based on the information from the LiDAR DSM, which shows that the façade points are most likely located within the mentioned height interval. The

result is shown in Fig. 17(e), where the detected TomoSAR façade points are color-coded based on the scatterer density estimates and are plotted along with the estimated façade line from the LiDAR DSM. Eventually, Fig. 17(f) shows the histogram of the perpendicular distances between the TomoSAR façade points and the fitted LiDAR façade line. The mean value of the distances implies the bias between the TomoSAR point cloud and the LiDAR DSM for this specific test site, which is equal to -0.184 m. The standard deviation hints the estimation accuracy of TomoSAR, which is equal to 1.17 m. It matches well with the meter-order elevation estimatio accuracy calculated from the derived Cramér–Rao lower bound [31].

For vertical accuracy analysis, several test sites are selected. The important criterion in the selection of the test sites is that they should be absolutely located on flat ground. Experiences show that the deviations within a couple of meters exist between the TomoSAR point cloud and the LiDAR DSM. This corresponds to: 1) the meter-level elevation estimation accuracy of repeat–pass InSAR, such as TomoSAR; and 2) the fact the roof points or ground points selected for evaluation are very unlikely being precisely located on flat surfaces as desired.
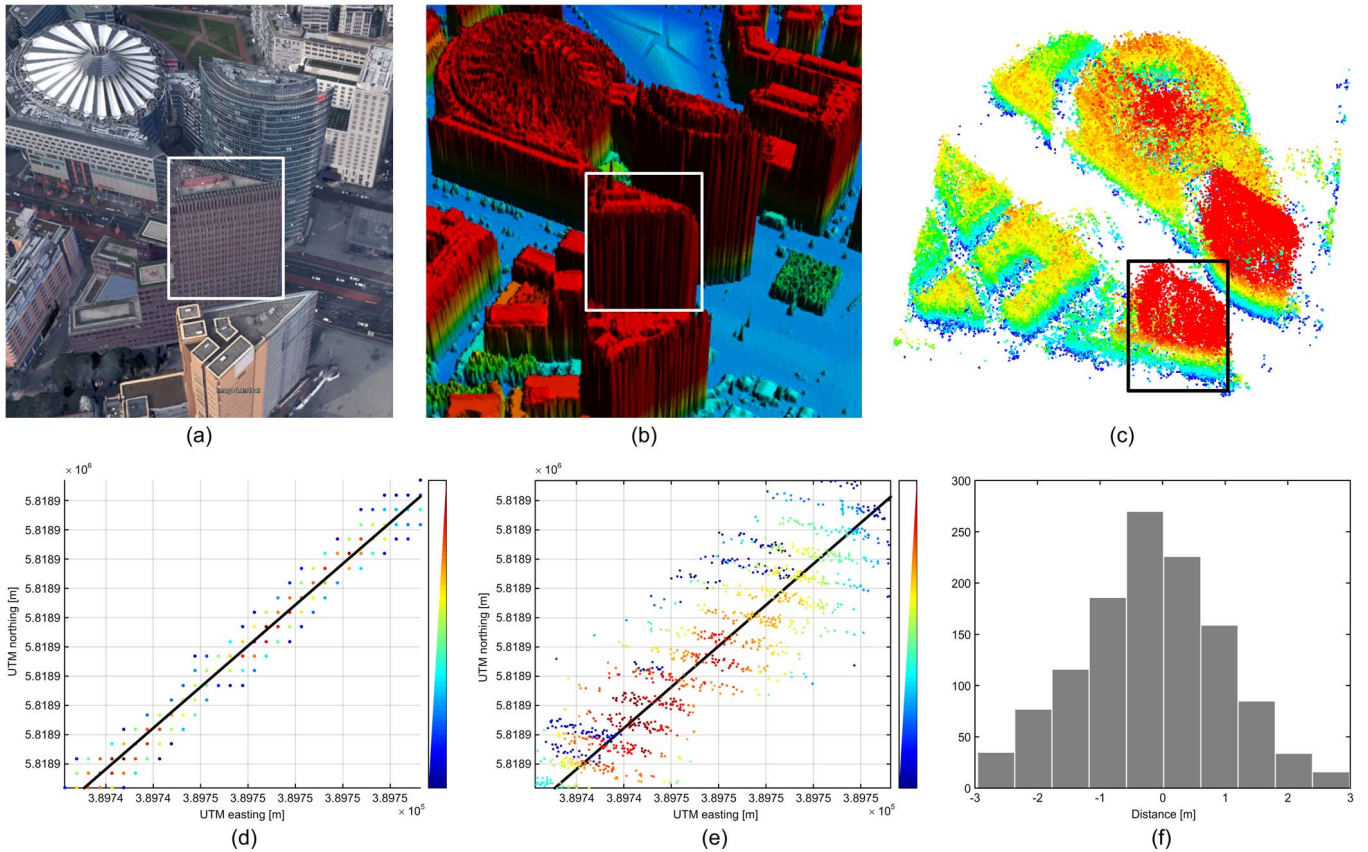
Fig. 17. First test site for the horizontal accuracy analysis of the TomoSAR point cloud. The investigated façade is marked with white rectangles. (a) Optical image of Postdamer Platz (Google Earth). (b) Corresponding LiDAR DSM. (c) Corresponding TomoSAR point cloud. d) Detected façade points from the LiDAR DSM and the estimated façade line in UTM projection. The façade points are color-coded based on the height variance values and the façade line is expressed in black. (e) Detected façade points from the TomoSAR point cloud, color-coded based on the scatterer density values, and the façade line fitted to LIDAR DSM drawn in black. (f) Histogram of differences between TomoSAR façade points and the façade line in LiDAR DSM. The histogram is centered on −0.184 m, which hints the achieved absolute positioning accuracy.

VI. Conclusion

In this paper, we have proposed the "geodetic TomoSAR" framework that fuses SAR image geodesy and SAR tomography to obtain *absolute* 3-D positions of a large amount of natural scatterers. An absolute 3-D TomoSAR point cloud with 63 million points covering an area of 10 km × 5 km over the city of Berlin is presented. Compared with a high-precision LiDAR DEM, the absolute positioning accuracy of the proposed approach reaches 20 cm. It demonstrates the applicability of the proposed approach. Future work concentrates on: 1) automatic identification of common scattering objects appearing in SAR images obtained from different geometries, e.g., by exploring the regular pattern attributed to building façades (for same heading orbits) or street lamps (for cross heading orbits); and 2) investigation on absolute deformation estimates for large area, e.g., by cooperating GPS measurements.

References

[1] G. Fornaro, F. Serafino, and F. Soldovieri, "Three-dimensional focusing with multipass SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 507–517, Mar. 2003.

[2] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 37–44, Jan. 2005.

[3] G. Fornaro, D. Reale, and F. Serafino, "Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 1, pp. 224–237, Jan. 2009.

[4] X. X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4296–4308, Dec. 2010.

[5] X. X. Zhu and R. Bamler, "Let's do the time warp: Multicomponent nonlinear motion estimation in differential SAR tomography," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 735–739, Jul. 2011.

[6] Y. Wang, X. X. Zhu, Y. Shi, and R. Bamler, "Operational TomoSAR processing using terrasar-X high resolution spotlight stacks from multiple view angles," in *Proc. IEEE IGARSS*, 2012, pp. 7047–7050.

[7] P. Ma, H. Lin, H. Lan, and F. Chen, "On the performance of reweighted minimization for tomographic SAR imaging," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 4, pp. 895–899, Apr. 2015.

[8] L. Wei, T. Balz, L. Zhang, and M. Liao, "A novel fast approach for SAR tomography: Two-step iterative shrinkage/thresholding," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 6, pp. 1377–1381, Jun. 2015.

[9] O. Frey, I. Hajnsek, U. Wegmuller, and C. L. Werner, "SAR tomography based 3-D point cloud extraction of point-like scatterers in urban areas," in *Proc. IEEE IGARSS*, 2014, pp. 1313–1316.

[10] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 8–20, Jan. 2001.

[11] N. Adam, B. Kampes, M. Eineder, J. Worawattanamateekul, and M. Kircher, "The development of a scientific permanent scatterer system," in *Proc. ISPRS Workshop High Resolution Mapping Space, Hannover, Germany*, 2003, vol. 2003, p. 6.

[12] B. M. Kampes, *Radar Interferometry—Persistent Scatterer Technique*, vol. 12. Dordrecht, The Netherlands: Springer-Verlag, 2006.

[13] S. Gernhardt, N. Adam, M. Eineder, and R. Bamler, "Potential of very high resolution SAR for persistent scatterer interferometry in urban areas," *Ann. GIS*, vol. 2, no. 16, pp. 103–111, Sep. 2010.

[14] Z. Wang, D. Perissin, and H. Lin, "Subway tunnels identification through Cosmo-SkyMed PSInSAR analysis in Shanghai," in *Proc. IEEE IGARSS*, 2011, pp. 1267–1270.

[15] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti, "A new algorithm for surface deformation monitoring based on small baseline differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 40, no. 11, pp. 2375–2383, Nov. 2002.

[16] R. Lanari *et al.*, "A small-baseline approach for investigating deformations on full-resolution differential SAR interferograms," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 7, pp. 1377–1386, Jul. 2004.

[17] K. Goel and N. Adam, "An advanced algorithm for deformation estimation in non-urban areas," *ISPRS J. Photogramm. Remote Sens.*, vol. 73, pp. 100–110, Sep. 2012.

[18] A. Ferretti *et al.*, "A new algorithm for processing interferometric datastacks: SqueeSAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 9, pp. 3460–3470, Sep. 2011.

[19] Y. Wang, X. X. Zhu, and R. Bamler, "Retrieval of phase history parameters from distributed scatterers in urban areas using very high resolution SAR data," *ISPRS J. Photogramm. Remote Sens.*, vol. 73, pp. 89–99, Sep. 2012.

[20] Y. Wang, X. X. Zhu, and R. Bamler, "Robust covariance matrix estimation with application to volcano monitoring using SAR image stacks," in *Proc. EUSAR*, 2014, pp. 1–4.

[21] G. Fornaro, A. Pauciullo, D. Reale, and S. Verde, "Improving SAR tomography urban area imaging and monitoring with CAESAR," in *Proc. EUSAR*, 2014, pp. 1–4.

[22] X. X. Zhu, N. Adam, and R. Bamler, "First demonstration of spaceborne high resolution SAR tomography in urban environment using TerraSAR-X data," in *Proc. CEOS SAR Workshop*, Oberpfaffenhofen, Germany, 2008, pp. 1–8.

[23] D. Reale, G. Fornaro, A. Pauciullo, X. Zhu, and R. Bamler, "Tomographic imaging and monitoring of buildings with very high resolution SAR data," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 661–665, Jul. 2011.

[24] X. Zhu and R. Bamler, "Demonstration of super-resolution for tomographic SAR imaging in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 8, pp. 3150–3157, Aug. 2012.

[25] R. Bamler, M. Eineder, N. Adam, X. Zhu, and S. Gernhardt, "Interferometric potential of high resolution spaceborne SAR," *Photogramm.—Fernerkund.—Geoinf.*, vol. 2009, no. 5, pp. 407–419, Nov. 2009.

[26] G. Fornaro, F. Lombardini, A. Pauciullo, D. Reale, and F. Viviani, "Tomographic processing of interferometric SAR data: Developments, applications, and future research perspectives," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 41–50, Jul. 2014.

[27] X. X. Zhu and R. Bamler, "Superresolving SAR tomography for multidimensional imaging of urban areas: Compressive sensing-based TomoSAR inversion," *IEEE Signal Process. Mag.*, vol. 31, no. 4, pp. 51–58, Jul. 2014.

[28] M. Eineder, C. Minet, P. Steigenberger, X. Cong, and T. Fritz, "Imaging geodesy—Toward centimeter-level ranging accuracy with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 661–671, Feb. 2011.

[29] X. Cong, U. Balss, M. Eineder, and T. Fritz, "Imaging geodesy—Centimeter-level ranging accuracy with TerraSAR-X: An update," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 5, pp. 948–952, Sep. 2012.

[30] F. W. Leberl, *Radargrammetric Image Processing*. Rijeka, Croatia: InTech, 1990.

[31] X. X. Zhu and R. Bamler, "Super-resolution power and robustness of compressive sensing for spectral estimation with application to spaceborne tomographic SAR," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 1, pp. 247–258, Jan. 2012.

[32] U. Balss *et al.*, "Precise measurements on the absolute localization accuracy of TerraSAR-X on the base of far-distributed test sites," in *Proc. EUSAR*, 2014, pp. 1–4.

[33] H. Breit, E. Börner, J. Mittermayer, J. Holzner, and M. Eineder, "The TerraSAR-X multi-mode SAR processor—Algorithms and design," in *Proc. EUSAR*, 2004, pp. 501–503.

[34] G. Petit and B. Luzum, *IERS Conventions (2010)*. Frankfurt, Germany: Verlag des Bundesamts für Kartographie und Geodäsie, 2010.

[35] B. Hofmann-Wellenhof, H. Lichtenegger, and E. Wasle, *GNSS—Global Navigation Satellite Systems: GPS, GLONASS, Galileo, and More*. Wein, Austria: Springer-Verlag, 2007.

[36] P. S. Mahapatra, S. Samiei-Esfahany, H. van der Marel, and R. F. Hanssen, "On the use of transponders as coherent radar targets for SAR Interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1869–1878, Mar. 2014.

[37] C. Gisinger *et al.*, "Precise three-dimensional stereo localization of corner reflectors and persistent scatterers with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1782–1802, Apr. 2015.

[38] I. G. Cumming and F. H. Wong, *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*, Auflage: Har/Cdr. Boston, MA, USA: Artech House, 2005.

[39] T. Fritz *et al.*, "TerraSAR-X Ground Segment Level 1b Product Format Specification." TX-GS-DD-3307, v1.3, 2007.

[40] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.

[41] E. Aguilera, M. Nannini, and A. Reigber, "Multi-signal compressed sensing for polarimetric SAR tomography," in *Proc. IEEE IGARSS*, 2011, pp. 1369–1372.

[42] S. Gernhardt, X. Cong, M. Eineder, S. Hinz, and R. Bamler, "Geometrical fusion of multitrack PS point clouds," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 1, pp. 38–42, Jan. 2012.

[43] M. Schmitt, O. Maksymiuk, C. Magnard, and U. Stilla, "Radargrammetric registration of airborne multi-aspect SAR data of urban areas," *ISPRS J. Photogramm. Remote Sens.*, vol. 86, pp. 11–20, Dec. 2013.

[44] Y. Wang and X. X. Zhu, "Feature-Based Fusion of TomoSAR Point Clouds from Multiview TerraSAR-X Data Stacks," presented at the Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International, 2013, pp. 85–88.

[45] D. G. Lowe, "Object Recognition From Local Scale-Invariant Features," in *Proc. ICCV*, 1999, vol. 2, pp. 1150–1157.

[46] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. 7th Int. Joint Conf. Artif. Intell.*, San Francisco, CA, USA, 1981, vol. 2, pp. 674–679.

[47] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SAR-SIFT: A SIFT-like algorithm for applications on SAR images," in *Proc. IEEE IGARSS*, 2012, pp. 3478–3481.

[48] X. Zhu, "Very high resolution tomographic SAR inversion for urban infrastructure monitoring: A sparse and nonlinear tour," Ph.D. dissertation, Deutsche Geodätische Kommission, Munich, Germany, 2011.

[49] X. Zhu, Y. Wang, S. Gernhardt, and R. Bamler, "Tomo-GENESIS: DLR's tomographic SAR processing system," in *Proc. JURSE*, 2013, pp. 159–162.

[50] G. Fornaro, A. Pauciullo, and F. Serafino, "Deformation monitoring over large areas with multipass differential SAR interferometry: A new approach based on the use of spatial differences," *Int. J. Remote Sens.*, vol. 30, no. 6, pp. 1455–1478, Mar. 2009.

[51] Y. Wang, X. Zhu, and R. Bamler, "An efficient tomographic inversion approach for urban mapping using meter resolution SAR image stacks," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 7, pp. 1–5, Jul. 2014.

[52] R. F. Hanssen, *Radar Interferometry: Data Interpretation and Error Analysis*. Dordrecht, The Netherlands: Springer-Verlag, 2001.

[53] X. Zhu and R. Bamler, "Tomographic SAR inversion by L1-norm regularization—The compressive sensing approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3839–3846, Oct. 2010.

[54] X. X. Zhu and R. Bamler, "Super-Resolution of Sparse Reconstruction for Tomographic SAR Imaging—Demonstration with Real Data," in *Proc. 9th EUSAR*, 2012, pp. 215–218.

[55] S. Gernhardt and R. Bamler, "Deformation monitoring of single buildings using meter-resolution SAR data in PSI," *ISPRS J. Photogramm. Remote Sens.*, vol. 73, pp. 68–79, Sep. 2012.

[56] M. Schmitt, "Reconstruction of urban surface models from multi-aspect and multi-baseline interferometric SAR," Ph.D. dissertation, Technische Univ. München, München, Germany, 2014.

[57] M. Schmitt, "Three-dimensional reconstruction of urban areas by multi-aspect TomoSAR data fusion," in *Proc. Joint Urban Remote Sens. Event*, Lausanne, Switzerland, 2015, (On CD).

[58] X. Zhu and M. Shahzad, "Façade reconstruction using multiview space-borne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3541–3552, Jun. 2014.

[59] M. Shahzad and X. X. Zhu, "Robust reconstruction of building facades for large areas using spaceborne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 752–769, Feb. 2015.

Christoph Gisinger received the Diploma in geodesy from Technische Universität Graz, Graz, Austria, and the M.Sc. degree in Earth-oriented space science and technologies from Technische Universität München, Munich, Germany. He is currently toward the Ph.D. degree in joint processing of SAR and global navigation satellite systems to retrieve geophysical signals at the Technische Universität München.

He is currently the Chair of the Institute for Astronomical and Physical Geodesy, Technische Universität München. His research interests include stereo SAR, correction modeling for absolute SAR observations, and radargrammetric image processing with focus on fusing SAR with geodetic observations to measure deformations.

Xiao Xiang Zhu (S'10–M'12–SM'14) received the Bachelor's degree in space engineering from the National University of Defense Technology, Changsha, China, in 2006 and the M.Sc. and Dr.-Ing. degrees and the Habilitation in signal processing from Technische Universität München (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

In 2009, she was a Guest Scientist or a Visiting Professor with the Italian National Research Council (CNR-IREA), Naples, Italy; in 2014, with Fudan University, Shanghai, China; and in 2015, with the University of Tokyo, Tokyo, Japan. Since 2011, she has been a Scientist with the Remote Sensing Technology Institute, German Aerospace Center, Wessling, Germany, where she is currently the Head of the Team Signal Analysis. Since 2013, she has been a Helmholtz Young Investigator Group Leader and appointed as TUM Junior Fellow. In 2015, she has been appointed as the Professor for Signal Processing for Earth Observation at TUM. Her main research interests include advanced InSAR techniques such as high-dimensional tomographic SAR imaging and SqueeSAR, computer vision in remote sensing including object reconstruction and multidimensional data visualization, and modern signal processing, including innovative algorithms such as compressive sensing and sparse reconstruction, with applications in the field of remote sensing such as multi/hyperspectral image analysis.

Ramon F. Hanssen (ME'04–SM'15) received the M.Sc. degree in geodetic engineering and the Ph.D. degree (*cum laude*) from the Delft University of Technology, Delft, The Netherlands, in 1993 and 2001, respectively.

He was with the International Institute for Aerospace Survey and Earth Science (ITC), Enschede, The Netherlands; the University of Stuttgart, Stuttgart, Germany; and the German Aerospace Center (DLR), Wessling, Germany. He was a Fulbright Fellow at Stanford University, Stanford, CA, USA. He was also with Scripps Institution of Oceanography, La Jolla, CA, working in the field of microwave remote sensing, radar interferometry, signal processing, and geophysical application development. Since 2008, he has been an Antoni van Leeuwenhoek Professor in Earth observation with the Delft University of Technology and has been leading the Department of Geoscience and Remote Sensing since 2011. He is the author of a textbook on radar interferometry.

Richard Bamler (M'95–SM'00–F'05) received the Diploma in Electrical engineering, the Doctor of Engineering degree, and the Habilitation in signal and systems theory from the Technische Universität München (TUM), Munich, Germany, in 1980, 1986, and 1988, respectively.

From 1981 to 1989, He worked at TUM on optical signal processing, holography, wave propagation, and tomography. Since 1989, he has been with the German Aerospace Center (DLR), Wessling, Germany, where he is currently the Director of the Remote Sensing Technology Institute. In early 1994, he was a Visiting Scientist with the Jet Propulsion Laboratory, in preparation of the SIC-C/X-SAR missions. In 1996, he was a Guest Professor with the University of Innsbruck, Innsbruck, Austria. Since 2003, he has been a Full Professor in remote sensing technology at the TUM as a double appointment with his DLR position. His teaching activities include university lectures and courses on signal processing, estimation theory, and SAR. He, as well as his team and institute, has been working on SAR and optical remote sensing, image analysis and understanding, stereo reconstruction, computer vision, ocean color, passive and active atmospheric sounding, and laboratory spectrometry. They were and are responsible for the development of the operational processors for SIR-C/X-SAR, SRTM, TerraSAR-X, TanDEM-X, Tandem-L, ERS-2/GOME, ENVISAT/SCIAMACHY, MetOp/GOME-2, Sentinel 5p, EnMAP, etc. His research interests include algorithms for optimum information extraction from remote sensing data with emphasis on SAR. This involves new estimation algorithms, such as sparse reconstruction and compressive sensing. He has devised several high-precision algorithms for monostatic and bistatic SAR processing, SAR calibration and product validation, ground moving target identification for traffic monitoring, SAR interferometry, phase unwrapping, persistent scatterer interferometry, differential SAR tomography, and data fusion.

Sina Montazeri received the Bachelor's degree in geodesy and surveying engineering from the University of Isfahan, Isfahan, Iran and the Master's degree in geomatics from the Delft University of Technology, Delft, The Netherlands in 2011 and 2014, respectively. He is currently working toward the Ph.D. degree with the Remote Sensing Technology Institute, German Aerospace Center (DLR), Wessling, Germany.

From April 2013 to May 2014, he was a Part-Time Research Assistant with the Remote Sensing Technology Institute, DLR, working on absolute localization of point clouds obtained from SAR tomography. His research interests include, InSAR processing, image and signal processing relevant to radar imagery, parameter estimation, and geodetic InSAR.

A.4 Montazeri, S., Zhu, X.X., Eineder, M., Bamler, R., 2016. Three-Dimensional Deformation Monitoring of Urban Infrastructure by Tomographic SAR Using Multitrack TerraSAR-X Data Stacks. IEEE Transactions on Geoscience and Remote Sensing 54(12): 6868-6878

Three-Dimensional Deformation Monitoring of Urban Infrastructure by Tomographic SAR Using Multitrack TerraSAR-X Data Stacks

Sina Montazeri, Xiao Xiang Zhu, *Senior Member, IEEE*, Michael Eineder, *Fellow, IEEE*, and Richard Bamler, *Fellow, IEEE*

Abstract—**Differential synthetic aperture radar tomography (D-TomoSAR), similar to its conventional counterparts such as differential interferometric SAR and persistent scatterer interferometry, is only capable of capturing 1-D deformation along the satellite's line of sight. In this paper, we propose a method based on L1-norm minimization within local spatial cubes to reconstruct 3-D displacement vectors from TomoSAR point clouds available from at least three different viewing geometries. The methodology is applied on two pairs of cross-heading—combination of ascending and descending—TerraSAR-X (TS-X) spotlight image stacks over the city of Berlin. The linear deformation rate and the amplitude of seasonal deformation are decomposed, and the results from two test sites with remarkable deformation pattern are discussed in detail. The results, to our knowledge, demonstrate the first attempt for motion decomposition using TomoSAR data from multiple viewing geometries.**

Index Terms—**Differential SAR tomography (D-TomoSAR), geodetic point cloud fusion, L1-norm minimization, motion decomposition, synthetic aperture radar (SAR).**

I. Introduction

TOMOGRAPHIC synthetic aperture radar (SAR) inversion (TomoSAR) is a multibaseline extension of conventional interferometric SAR (InSAR) which allows for multiple scatterer discrimination within each resolution cell [1]–[4]. Therefore, it is favored over similar InSAR approaches, such as persistent scatterer interferometry (PSI) [5], [6], for studying urban areas in which the prevalent occurrence of layover violates the single scatterer assumption. By multitemporal analysis of SAR image stacks, TomoSAR is capable of estimating the underlying deformation of individual scatterers in the scene, a framework called differential TomoSAR (D-TomoSAR) [7]. Using TerraSAR-X (TS-X) high-resolution spotlight data stacks, experiments show that D-TomoSAR is capable of providing 4-D maps with a scatterer density on the order of one million points per square kilometers [8].

One of the limitations of InSAR techniques, including TomoSAR, is that they only measure deformation along the radar line of sight (LOS). Therefore, apart from the fact that no information can be retrieved from the shadowed parts, due to the dependence of the LOS motion on the satellite's position, interpreting the deformation pattern can be complicated. In order to enhance the understanding of the undergoing deformation, a decomposition of the observed LOS displacement into the 3-D deformation vector is beneficial. One method to tackle the aforementioned problem is using a combination of deformation observations from an ascending and a descending track (cross-heading) where two out of three components can be retrieved [9]. For particular cases, to retrieve the third component, a prior knowledge about the characteristics of the displacement is necessary [9], [10]. If LOS observations from more than two geometry configurations are available, then it is possible to reconstruct the full 3-D displacement vector in the local coordinate system without the use of auxiliary data [11]. However, since the LOS deformation estimates obtained from different viewing geometries do not necessarily originate from the same object, a strategy should be introduced to overcome this problem.

In this paper, we propose a method to reconstruct 3-D deformation vectors of urban areas from TomoSAR point clouds available from at least three different viewing geometries. The method has the following strengths.

1) It benefits from multitrack data which, in the cross-heading case, allow for shadow-free deformation monitoring and, in the same-heading case, provide substantial increase in the scatterer density of the displacement map.
2) For coregistration of the multitrack point clouds, the method uses a geodetic approach unlike geometric approaches such as the ones reported in [12] and [13]. The coregistration therefore is more accurate and objective in terms of geolocalization.
3) It estimates the full 3-D displacement vector of each point by minimizing the L1-norm of the residuals instead of

the more popular L2-norm. This approach is more robust with respect to outliers and retains the high density of TomoSAR point clouds.

To our knowledge, this is the first attempt for motion decomposition using TomoSAR data from multiple viewing geometries.

The rest of this paper is organized as follows. In Sections II and III, the theoretical aspects related to TomoSAR and InSAR LOS deformation are described, respectively. Section IV explains the proposed algorithm. Section V reports on the results of applying the proposed method on simulated data as well as real data, and the conclusion is drawn in Section VI.

II. D-TomoSAR

Conventional D-TomoSAR, proposed in [7], assumes solely the linear motion of multiple scatterers in each azimuth-range resolution cell. However, dealing with X-band data of urban areas, nonlinear motion parameters such as the periodic seasonal displacement should also be taken into account, which does not fit into the framework of spectral estimation [1], [14]. In [14], a generalized time warp method is proposed in order to allow for multicomponent nonlinear motion retrieval of multiple scatterers which rewrites the D-TomoSAR system model to an $M + 1$-dimensional spectral estimation problem, where M denotes the user-defined motion model order. The specific type of the time warp method relevant to our application is described in this section.

It is well known that, within the Fraunhofer far-field approximation, the focused complex-valued SAR measurement (g_n) at an azimuth-range pixel is essentially the Fourier transform of the elevation (s)-dependent reflectivity profile (γ) of that pixel sampled at the corresponding elevation frequency (ξ_n) which, considering the temporal progression of the scatterers, for the nth acquisition at time t_n $(n = 1, \ldots, N)$, is written as [14]

$$g_n = \int_{\Delta s} \gamma(s) \exp\left(-j2\pi\left(\xi_n s + \frac{2d(s, t_n)}{\lambda}\right)\right) ds \quad (1)$$

where Δs is the elevation extent of the object, λ is the wavelength, r is the slant range, and $\xi_n = (-2b_n)/\lambda r$, with b_n denoting the perpendicular baseline. $d(s, t_n)$ is the relative LOS motion as a function of elevation and time which can be modeled by a linear combination of M base functions $\tau_m(t_n)$ as [14]

$$d(s, t_n) = \sum_{m=1}^{M} p_m(s)\tau_m(t_n) \quad (2)$$

where $p_m(s)$ is the corresponding motion coefficient to be retrieved. As discussed earlier, the relevant motion models to be considered in urban areas are linear and seasonal (thermal) deformation $(M = 2)$ which, by defining the mth temporal frequency component at t_n as $\eta_{m,n} = 2\tau_m(t_n)/\lambda$, are characterized as follows [14]:

1) *linear motion* with $\eta_{1,n} = 2t_n/\lambda$ and the coefficient $p_1(s)$ which stands for the LOS velocity (v) as a function of s;
2) *seasonal motion* with $\eta_{2,n} = 2\sin(2\pi(t_n - t_0))/\lambda$ and the coefficient $p_2(s)$ which stands for the amplitude

(a) of the periodic motion while t_0 is the initial phase offset. Note that t_n and t_0 should be specified in years. Alternatively, a temperature-dependent motion can be considered.

As the base function of the seasonal motion is chosen as a sinusoidal signal, (1) implies a nonlinear problem. The time warp method deals with this issue by rearranging the image acquisition times in order to imitate a linear motion trend [14]. Therefore, the D-TomoSAR system model, for our specific case $(M = 2)$, is extended to

$$g_n = \int_{\Delta a} \int_{\Delta v} \int_{\Delta s} \gamma(s)\delta\left(v - v(s), a - a(s)\right)$$
$$\cdot \exp\left(-j2\pi(\xi_n s + \eta_{1,n} v + \eta_{2,n} a)\right) \times dsdvda$$
$$n = 1, \ldots N. \quad (3)$$

Equation (3) is further discretized along s and the motion parameters space (p_1, p_2). The discrete D-TomoSAR system model with the presence of noise ε can be written as [4]

$$\mathbf{g} = \mathbf{R}\boldsymbol{\gamma} + \boldsymbol{\varepsilon} \quad (4)$$

where $\mathbf{g}$ is the measurement vector with N elements, $\boldsymbol{\gamma}$ is the reflectivity function, and $\mathbf{R}$ is the mapping matrix. For more information on how to solve (4) for retrieving the relevant unknown parameters, the interested reader is referred to [1] and [14].

III. InSAR LOS Deformation

The deformation measurement of SAR techniques d_{LoS} is the projection of the original 3-D displacement vector $\mathbf{d}$ with components d_e, d_n, and d_u in east, north, and up direction, respectively, onto the LOS. Assuming a local incidence angle of θ_{inc} and a satellite orbit with heading angle α_h, we can write [9]

$$d_{\text{LoS}} = d_u \cos(\theta_{\text{inc}}) - d_{\text{ALD}} \sin(\theta_{\text{inc}}) \quad (5)$$

where d_{ALD} includes the projection of d_n and d_e on the azimuth look direction (ALD), which is perpendicular to the satellite flying direction and therefore is expressed as

$$d_{\text{ALD}} = d_e \cos(\alpha_h) - d_n \sin(\alpha_h). \quad (6)$$

Fig. 1 depicts the aforementioned projection in 3-D. If (6) is substituted in (5), the explicit relation between deformation measurement d_{LoS} and the displacement components for a single pixel can be written as

$$d_{\text{LoS}} = d_u \cos(\theta_{\text{inc}}) - d_e \cos(\alpha_h)\sin(\theta_{\text{inc}}) + d_n \sin(\alpha_h)\sin(\theta_{\text{inc}}). \quad (7)$$

From (7), it is inferred that, in order to be able to solve for the three deformation components, at least three LOS observations from different nonplanar acquisition geometries are required. Another issue regarding motion decomposition is the sensitivity of InSAR observations with respect to each component. Considering the near polar orbit of TS-X satellites, for instance, with the heading angle of 190.6° and an incidence angle of 36.1°, the sensitivity decomposition of LOS deformation is
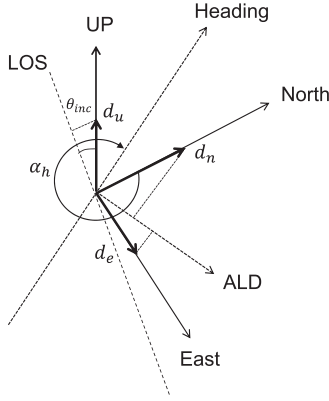
Fig. 1. Projection of the original displacement vector $\vec{d}$ with components (d_e, d_n, d_u) onto the radar LOS.

$[0.8, 0.58, -0.1] \cdot [d_u, d_e, d_n]^T$. Therefore, it is seen that observations are most sensitive to the deformation in the vertical direction and least sensitive to the deformation in the north direction. This fact should not be falsely interpreted as ignoring the deformation component d_n in the functional model of the decomposition problem stated in (7), like in [15]. According to (6), converting d_{ALD} to d_e while ignoring d_n results in the bias Δd_e in the east-west motion component which is expressed as

$$\Delta d_e = d_n \cdot \tan(\alpha_h). \tag{8}$$

With typical satellite azimuth values of TS-X, (8) demonstrates that the systematic error in east-west component can reach up to 18% of the deformation in the north-south direction if d_n is omitted.

IV. Motion Decomposition Using Multitrack InSAR Observations

The workflow of the proposed algorithm is shown in Fig. 2. It uses as input the TomoSAR point clouds from different viewing geometries and delivers 3-D deformation maps. In this section, as the main focus of this paper, the procedures inside the dashed rectangle are explained in detail. In each section, at first, a small motivation is given, followed by describing the proposed methodology.

A. Geodetic Point Cloud Fusion

InSAR point cloud fusion is referred to as the coregistration of at least two point clouds obtained from different orbit tracks, either same or cross-heading. Such coregistration is necessary because the reference point chosen during the PSI and TomoSAR processing is different for each stack of SAR data with unknown height which causes the point clouds to show certain offsets with respect to their true positions as well as to each other [12]. For solving this problem, relevant to meter-resolution SAR data, two *geometrical* approaches have been introduced in [12] and [13]. Although with such methods the coregistration is carried out successfully, there will be no guarantee that the yielded fused point cloud is accurately geolocalized.
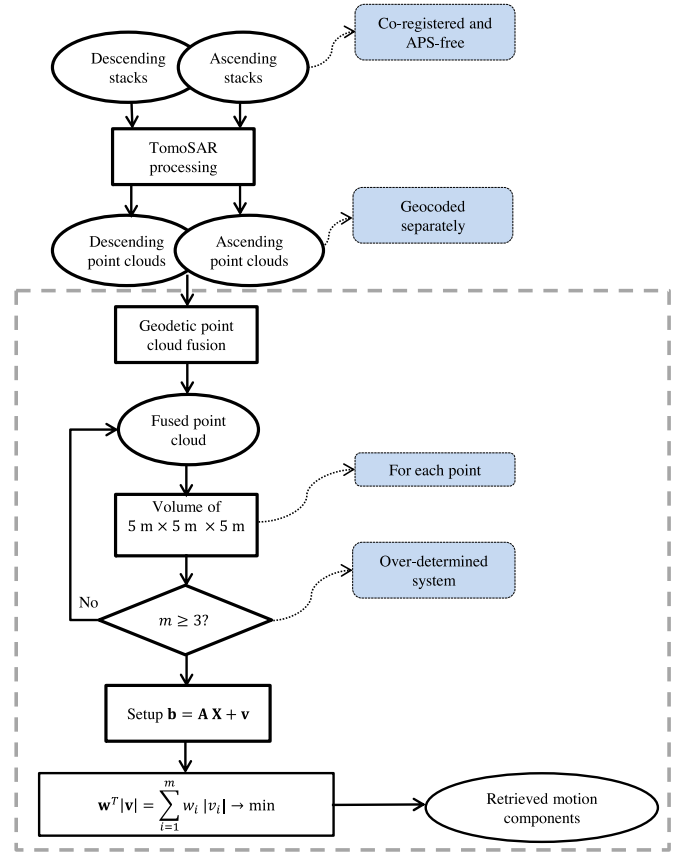

Fig. 2. Flowchart of the proposed motion decomposition algorithm. The processes inside the dashed rectangle are explained in Section IV.

In this paper, point cloud fusion is carried out *geodetically* using the recently developed *geodetic SAR tomography* framework [16]. The basic idea is to select an identical reference point during TomoSAR processing of each stack whose absolute position is retrieved from the SAR data itself. The position of this reference point is calculated by StereoSAR [17], i.e., combining more than two SAR measurements, available from different orbits, whose range and azimuth times are corrected for the most prominent error sources, a method called *imaging geodesy* [18], [19]. Since the elevation and deformation of scatterers of each stack are evaluated with respect to this point, the *absolute* point clouds are coregistered after geocoding with some refinements. The required steps for performing geodetic point cloud fusion are summarized in the following. For more details, the interested reader is recommended to consult [16].

1) Identification of time-coherent reference point candidates to be visible in SAR images from two or more viewing geometries. The stability of the reference points is measured by evaluating the amplitude dispersion index [5].
2) Extraction of 2-D radar coordinates of the candidates with subpixel accuracy from all available SAR images using point target analysis [20].
3) Correction of the radar time coordinates for the atmospheric, geodynamic, and satellite effects depending on the corresponding acquisition time and geometry of the SAR observation using *imaging geodesy* [18], [19].

4) Solving the range-Doppler equations for identical targets visible from different viewing geometries using the stereo SAR method to retrieve the absolute 3-D coordinates [17] and selecting the point with the highest quality visible in all stacks as the reference point.

5) Performing TomoSAR processing of each stack based on the identical reference point chosen in step 4. The results are then independently geocoded.

6) If the selected reference point was localized from the same-heading tracks, the geodetic point cloud fusion is performed by shifting all of the geocoded point clouds toward the coordinates of the absolute reference point. If the selected reference point was localized from the cross-heading tracks, since most probably these targets are vertically oriented cylinders such as lamp poles, the bias in the coordinates of the reference point should be resolved, depending on the diameter of the object, before shifting the point clouds toward the coordinates of the reference point.

The aforementioned framework leads to a shadow-free dense TomoSAR point cloud with *absolute* coordinates which is far more accurate than the geometrically fused point clouds in terms of geolocalization. In [16], it was reported that the horizontal localization accuracy of such geodetically fused point clouds is approximately 20 cm when compared to highly accurate digital surface models obtained from LiDAR. Considering the spatial extent of buildings in urban areas, it is important to note that the mentioned accuracy is sufficient for the task of motion decomposition.

B. Problem Formulation Within Spatial Cubes

The absolutely localized fused point cloud, which is the result of the previous step, is the basis for motion decomposition. Since it is unlikely that the same scatterer can be visible from different viewing geometries, especially from cross-heading tracks, the problem of decomposition is defined in a spatial volume centered on the target point for which the 3-D motion retrieval is desired. The cube is sliding on the points until the entire area of interest is covered. Inside the cube, the necessary information for the follow-on estimation is recorded. Assuming m as the number of points inside the cube, except for the central point, which is constrained to be equal or larger than three in order to guarantee an overdetermined system, and n as the number of unknowns, which is equal to three, the observation vector $\mathbf{b}$ is of size $m \times 1$ including either LOS linear rates or LOS amplitudes of seasonal motion attributed to the corresponding scatterers; $\mathbf{X}$ is an $n \times 1$ vector consisting of the three unknown motion components of the central point, $\mathbf{A}$ is the design matrix of size $m \times n$, which is evaluated based on the heading angle of the satellite and the unique local incidence angle associated to each scatterer [see (7)], and $\mathbf{W}$ is the weight matrix of deformation observations of size $m \times m$ that is evaluated based on the inverse squared distance of each scatterer relative to the central point, which means that points closer to the central point have higher weights.

An issue regarding this approach is that scatterers that originated from different objects may be included in one cube.

Therefore, it is important to choose a proper dimension for the spatial cube. It should be small enough to only include, as much as possible, scatterers originating from an identical object and also big enough to include appropriate number of scatterers to prevent the decomposition problem from becoming underdetermined or unreliable. This choice depends on the inherent scale of the objects, as well as the spatial resolution of the SAR data. After an empirical study, for TS-X high-resolution spotlight data over European cities, we recommend to choose a cube size of 5 m × 5 m × 5 m, which is used in this paper.

Eventually, after building the required matrices and vectors, we move to the next step which is the estimation of the 3-D motion components of the central point inside the cube.

C. Three-Dimensional Motion Retrieval

With the necessary vector notations already introduced in Section IV-B, the functional model of our problem inside each cube is written as

$$\mathbf{b} = \mathbf{A}\mathbf{X} + \mathbf{v} \qquad (9)$$

with $\mathbf{v}$ being the vector of residuals with the same size as the observation vector, i.e., $m \times 1$. Equation (9) depicts an overdetermined $(m \geq n)$ system for which the estimation of unknowns is commonly carried out by minimizing the (weighted) sum of the square of residuals known as the (weighted) least squares (L2-norm minimization) [21], [22]

$$\mathbf{v}^T \mathbf{W} \mathbf{v} = \sum_{i=1}^{m} v_i \mathbf{W}_{i,i} v_i^T \to \min. \qquad (10)$$

Equation (10) gives the unique solution of the problem which guarantees unbiasedness and minimum variance when matrix $\mathbf{W}$ is chosen equal to the inverse of the covariance matrix of the observations [22]. However, the aforementioned properties of the least squares method are valid when observations are only influenced by random errors following a normal distribution. If several observations are affected by gross errors (outliers) to cause the errors follow a non-Gaussian and heavily tailed distribution, parameter estimation based on L2-norm minimization is far from optimal [23]. The reason lies on the objective function described in (10) which inherently gives more weight to outliers with large residual values rather than normal observations in the minimization process. Therefore, L2-norm minimization is not robust in such cases due to the high sensitivity of the estimator toward outliers.

For our specific case study which relies on point clouds obtained from TomoSAR, there always exists a large amount of outliers in the elevation and deformation estimates for the reasons adequately given in [1], [4], [8], and [24]. Thus, an estimator which is robust with respect to outliers should be employed for motion decomposition.

For the aforementioned reasons, inside each cube centered on the target point, we carry out the estimation of motion components by L1-norm adjustment.

For the system of equations outlined in (9), the unknown vector can be retrieved by minimizing the (weighted) sum of
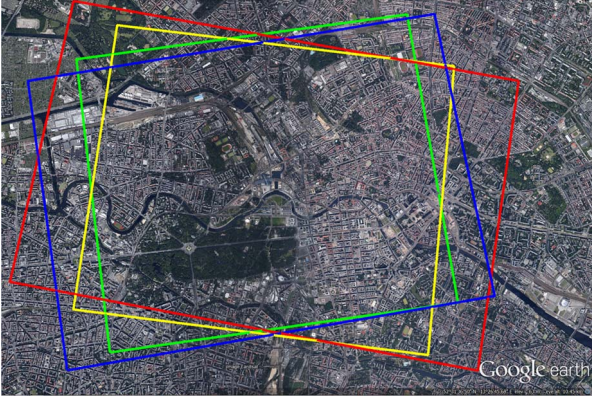
Fig. 3. Optical image of the city of Berlin (Google Earth). Rectangles mark the coverage of the four TS-X data stacks.

the absolute residuals [25]

$$\mathbf{w}^T |\mathbf{v}| = \sum_{i=1}^{m} w_i |v_i| \rightarrow \min \qquad (11)$$

where $\mathbf{w}$ is an $m \times 1$ vector which contains the diagonal elements of the weight matrix $\mathbf{W}$. The optimization problem outlined in (11) is dealt with linear programming. Unlike least squares, while solving (11), similar weights are associated with the residuals in the minimization process, i.e., L1-norm adjustment does not magnify the effect of outliers. Therefore, it facilitates less biased estimation of the unknown parameters compared to least squares, although the solution does not necessarily have a minimum variance [25].

V. APPLICATION ON SIMULATED DATA AND REAL DATA

In this section, the proposed method described in Section IV is validated using simulated and real data. In Section V-A and B, the data set is introduced, and the effect of the viewing geometry of the SAR satellites on the retrieval of the 3-D deformation vector is inspected, respectively. In Section V-C, a simple simulation is carried out to justify the utilization of L1-norm adjustment instead of least squares for the task of motion decomposition. Section V-D and E are dedicated to the TomoSAR processing and the motion decomposition carried out on real data, respectively. Finally, some remarks on the computational complexity of the entire procedure are given in Section V-F.

A. Test Area and Data Set

The investigated test site includes the central area of the city of Berlin, Germany. The available data set, 460 images in total, consists of four stacks of TS-X very high resolution spotlight images acquired with a range bandwidth of 300 MHz. The images have an azimuth resolution of 1.1 m and a slant range resolution of 0.6 m covering an area of 10 km × 5 km in a period of five years from March 2008 to March 2013. In terms of viewing geometry, two stacks were acquired from descending orbits with images recorded at 05:20 Coordinated Universal Time (UTC), and two stacks were acquired from ascending tracks with images recorded at 16:50 UTC. Fig. 3 shows the mean

TABLE I
ACQUISITION PARAMETERS OF EACH STACK INCLUDING THE AVERAGE
INCIDENCE ANGLE, THE AZIMUTH, THE TRACK TYPE,
AND THE NUMBER OF AVAILABLE IMAGES

Beam	Incidence angle [degree]	Heading angle [degree]	Track type	Nr. of Images
57	41.9	350.3	Ascending	102
85	51.1	352	Ascending	111
42	36.1	190.6	Descending	109
99	54.7	187.2	Descending	138

scene coverage of individual stacks overlaid on the optical image of Berlin. Furthermore, the details about the system parameters and properties of each stack are summarized in Table I.

B. Effect of Viewing Geometry

The quality of motion decomposition depends highly on the difference in the viewing geometry of the acquisitions [26]. This effect can be analyzed by a concept equivalent to the dilution of precision (DOP) commonly used with GNSS measurements [11], [27]. If we assume that a single point is observed from multiple viewing geometries where LOS measurements are uncorrelated and have equal standard deviations σ, one can evaluate the posterior covariance matrix of the displacement estimates $(\mathbf{Q}_{\hat{X}})$ as

$$\mathbf{Q}_{\hat{X}} = \begin{bmatrix} \sigma_u^2 & \sigma_{ue} & \sigma_{un} \\ \sigma_{eu} & \sigma_e^2 & \sigma_{en} \\ \sigma_{nu} & \sigma_{ne} & \sigma_n^2 \end{bmatrix} = \sigma^2 \cdot (\mathbf{A}^T \mathbf{A})^{-1}. \qquad (12)$$

The square root of the diagonal elements of $\mathbf{Q}_{\hat{X}}$ denotes the precision of the displacement components, where subscripts u, e, and n stand for up, east, and north, respectively. If we set $\sigma = 1$, then the diagonal terms represent the effect of geometry on these estimates. A low diagonal value indicates the goodness of the geometry configuration for retrieving the corresponding component. Using the incidence angle and azimuth values of Table I in (7), the DOP matrix is calculated for a point target hypothetically visible in all of the available four viewing geometries. The matrix is equal to

$$\begin{bmatrix} 43.3 & -0.8 & 277.8 \\ -0.8 & 0.51 & -5.4 \\ 277.8 & -5.4 & 1801.7 \end{bmatrix}.$$

The following conclusions can be drawn from the calculated DOP matrix.

1) The square root of the diagonal elements is approximately equal to [6.6 0.7 42.5], which shows the strength of the current geometry configuration to retrieve the east-west motion component and its weakness for retrieving the north-south motion component. Therefore, it is expected that the motion in the north-south direction cannot be reliably resolved due to the inherent near-polar orbit of TS-X satellites.

2) The off-diagonal elements show the covariance between relative errors in motion components. Thus, the north-south component should not be omitted from the

functional model of the decomposition problem to prevent the propagation of error on other components.

3) According to [28], the correlation between the motion components poses another problem to the motion decomposition. This value for the relative error in the north-south and the vertical components is calculated to be the following: $\rho_{nu} = \sigma_{nu}/(\sigma_n \cdot \sigma_u) = 277.8/(\sqrt{1801.7} \times \sqrt{43.3}) \approx 0.994$. This means that not only the north-south component should not be omitted from the functional model, but also with the current geometry configuration, the separation of the north-south and the vertical motion component is ambiguous; an issue that is elaborated in [28].

Based on these findings, the analysis of motion will be restricted to the east-west and vertical directions while retaining the motion in the north-south component in the functional model of the problem. It is essential to add that precise unambiguous retrieval of the north-south component, in general, is not possible using the geometry configuration of current SAR satellites [28].

Additionally, it is important to note that the analysis given in this section only considers the best possible combination of viewing geometries, i.e., existence of observations from the four available satellite orbits inside the spatial volume. Obviously, it is possible that inside each cube observation from less than four geometries will be available. For instance, in a test case, that involves the motion decomposition of a high-rise building for which observations from one side of the façade cannot include LOS measurements from the cross-heading tracks. Nevertheless, the concept described in this section can be used to analyze the impact of viewing geometry on each motion component in the aforementioned cases as well.

C. Motion Decomposition on Simulated Data

Motion decomposition is performed on simulated data solely for the purpose of comparing the results obtained from minimizing the L1-norm and the L2-norm of residuals. To this end, motion decomposition is carried out for a point target assumed to undergo a subsidence in the vertical direction with a magnitude of 10 mm/year and lateral deformations in the east and north directions with 2 and 1 mm/year, respectively, i.e., $\boldsymbol{d_T} = \begin{bmatrix} d_{u,T} = -10 \\ d_{e,T} = 1 \\ d_{n,T} = 2 \end{bmatrix}$ mm/year. It is important to note that, since the goal is to compare the two mentioned estimation methods, the magnitude and the distribution of the deformation are chosen arbitrarily with no actual physical meaning.

For the reconstruction of $\boldsymbol{d_T}$, assuming that the point is visible from the four viewing geometries listed in Table I, four LOS displacements are calculated using (7). These four values are then used to realize 30 LOS observations in four groups differing in the associated incidence and heading angles to each point. The value 30 is chosen based on the average number of points visible in a $5 \text{ m} \times 5 \text{ m} \times 5 \text{ m}$ cube in the fused TomoSAR point cloud of the city of Berlin. The 30 LOS deformation observations are randomly realized following a Gaussian probability density function with the mean value equal to the LOS deformation calculated using the mentioned
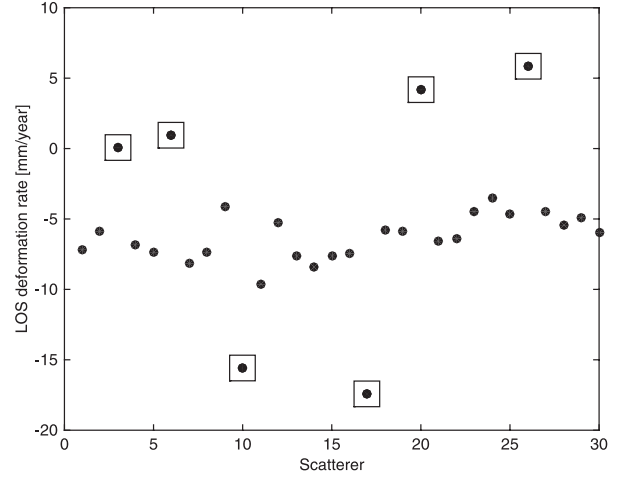

Fig. 4. Simulated LOS deformation values for 30 points assumed to be inside a cube. The simulation uses the system parameters of the four available data stacks outlined in Table I. The outliers are plotted inside squares, while the other dots represent the normal observations.

forward model and the variance (σ_v^2) approximated by the Cramér–Rao bound of LOS linear deformation rate obtained from interferometric measurements [5], [28], [29]

$$\sigma_v^2 = \left(\frac{\lambda}{4\pi}\right)^2 \cdot \frac{1}{2\text{NOA} \cdot \text{SNR}\sigma_t^2} \quad (13)$$

where NOA is the number of acquisitions in a stack of SAR data, SNR denotes the signal-to-noise ratio, and σ_t^2 is the variance of the temporal baseline distribution. If we consider a TS-X data stack with $\lambda = 31$ mm, assuming that SNR = 10 dB, NOA = 115, and $\sigma_t \approx 1.6$ years, evaluated based on the acquisitions listed in Table I, then $\sigma_v \approx 0.03$ mm/year. Of course, σ_v is higher in practice than the reported value mainly due to the atmospheric disturbances. Nevertheless, the exact inclusion of the atmospheric error was ignored since for us vital is the comparison between the two aforementioned parameter estimation methods. In order to have a more realistic scenario, we set the standard deviation of the estimated deformation rate of all scatterers inside the cube to be $\sigma_v = 1$ mm/year.

The next step includes the addition of deliberate outliers in the data. Based on the variability in the LOS deformation measurements observed in a number of cubes in the real data of Berlin, for the simulation, we assume that approximately 20% of the observations inside the cube (six in our case) are corrupted by gross errors. Without loss of generality, we presume that the linear deformation rate of these points is off with an error with a magnitude of 10 mm/year. Fig. 4 shows the scatter plot of the simulated LOS deformation values. Black dots inside the squares represent the outliers, while the other dots are normal observations. We proceed now with estimation of the 3-D deformation components of the desired point scatterer which is carried out by minimizing the L2- and L1-norms of the residuals following (10) and (11), respectively. Fig. 5 shows the histogram of residuals [see (9)] obtained from each estimator. It is seen that more residuals are close to zero in the L1 results rather than the L2. Also, it is observed that the L2-norm minimization tends to distribute the errors caused by
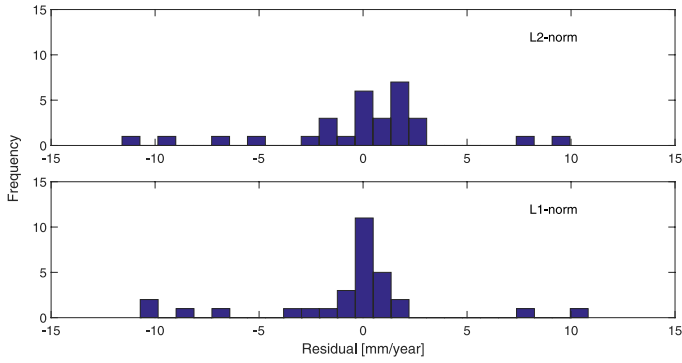
Fig. 5. Histograms of residuals evaluated after the motion decomposition performed with L2- and L1-norm minimization.

TABLE II
TRUE MOTION COMPONENTS AND THE RETRIEVED ONES FROM L2- AND L1-NORM MINIMIZATION FOR THE SIMULATED CASE STUDY

Components	True motion components [mm/year]	Estimated with L2-norm adjustment [mm/year]	Estimated with L1-norm adjustment [mm/year]	Difference w.r.t true motion components [mm/year]	
				L2-norm	L1-norm
d_u	−10	−5.57	−10.06	4.43	**−0.06**
d_e	1	2.21	0.88	1.21	**−0.12**
d_n	2	20.56	−4.08	18.56	**−6.08**

outliers onto other normal observations, while this effect is less significant for L1-norm minimization. Additionally, Table II reports the true motion components as well as the estimated ones with L2- and L1-norm minimization. The difference between the true components and the estimated ones is also reported in the fourth column. It can be observed that L1-norm minimization outperforms L2-norm minimization for all of the retrieved components.

D. Real-Data TomoSAR Processing and Geodetic Point Cloud Fusion

The InSAR stacking and TomoSAR processing of the Berlin data were carried out by the GENESIS [30] and the Tomo-GENESIS system [1], [31] of the Remote Sensing Technology Institute, German Aerospace Center (DLR). Starting from SLCs, for an input data stack, the Tomo-GENESIS system retrieves the following information: the number of overlaid scatterers inside each azimuth-range pixel, amplitude and phase, topography, and motion parameters of each detected scatterer. In this paper, the tomographic inversion was performed by the Maximum Detection (MD) algorithm of the Tomo-GENESIS. The MD algorithm uses SVD-Wiener for reconstruction of the reflectivity profile, followed by peak detection, model order selection, and final refinement of the amplitude and phase estimates [4].

According to previous studies on Berlin reported in [15] and [16], the city is rather stable, i.e., there is no significant linear ground deformation. Most of the buildings and other man-made urban infrastructure mainly undergo temperature-change-induced seasonal deformation [15], [32]. The LOS linear and
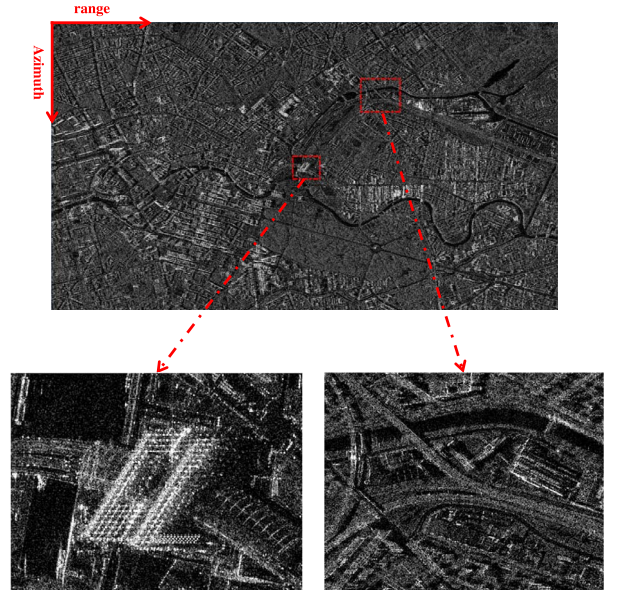

Fig. 6. (Top) Mean calibrated amplitude image of Berlin from a descending orbit track. (Bottom left) Berlin central railway station. (Bottom right) Railway bridge.
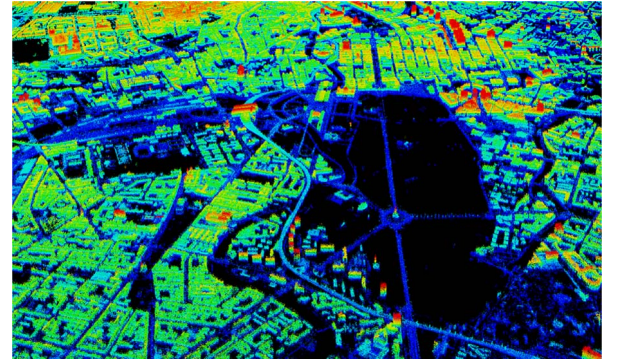

Fig. 7. Geodetically fused TomoSAR point cloud of Berlin in 3-D. Height is color-coded.

seasonal deformation maps of the entire Berlin can be found in [16], [32], and [33]. This study only focuses on two small test sites including the Berlin central railway station and the railway bridge which, in previous studies, showed a remarkable pattern of seasonal deformation [15], [34]. These test sites are marked with red rectangles on a descending SAR intensity map in Fig. 6.

For a meaningful analysis on the mentioned areas, LOS deformation estimates are to be decomposed (see Section V-E). For this purpose, TomoSAR point clouds obtained from multiple viewing angles are fused using the *geodetic SAR tomography* framework [16]. Fig. 7 illustrates the fused TomoSAR point cloud in the UTM coordinate system. The point cloud covers an area of 10 km × 5 km and contains approximately 63 million point scatterers. It is observed that fusion of point clouds obtained from cross-heading orbit tracks provides a highly detailed shadow-free point cloud which is the basis for the motion decomposition. Such shadow-free TomoSAR point clouds can be further utilized to reconstruct dynamic city models [35]–[38].
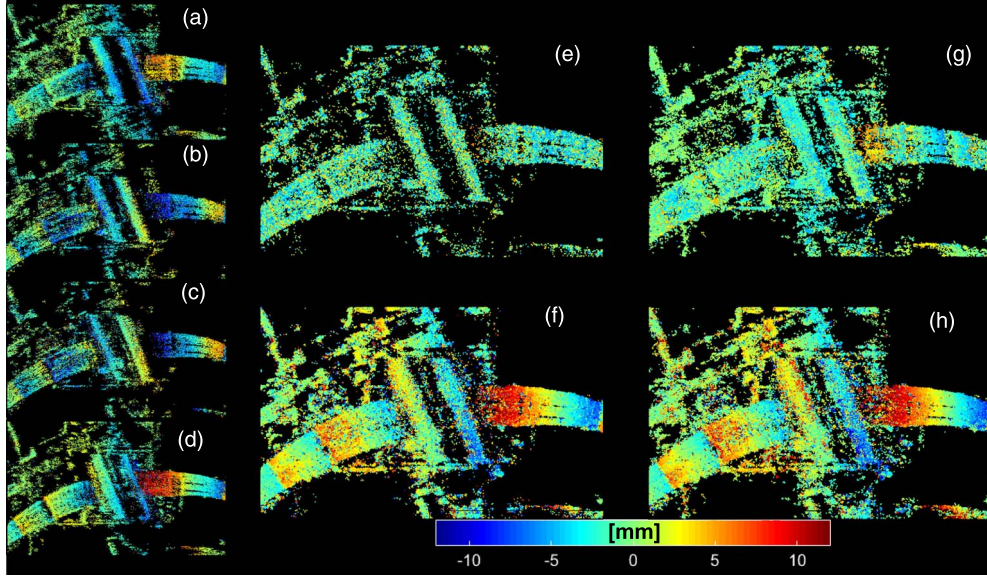
Fig. 8. Berlin central railway station. LOS amplitude of seasonal motion of (a) beam 42, (b) beam 57, (c) beam 85, and (d) beam 99. Decomposed seasonal deformation in (e) the vertical direction and (f) the east-west direction by L2-norm minimization. Decomposed seasonal deformation in (g) the vertical direction and (h) the east-west direction by L1-norm minimization. Deformation maps are in millimeters.

E. Motion Decomposition and Discussion

Using the method introduced in Section IV, motion decomposition is done on the two mentioned test cases. Fig. 8 shows the TomoSAR LOS seasonal deformation maps of the central station available from each orbit [Fig. 8(a)–(d)] as well as the decomposed motion maps in the vertical [Fig. 8(g)] and east-west directions [Fig. 8(h)]. In order to justify why L1-norm minimization was preferred over L2-norm, the vertical and east-west motion components evaluated by L2-norm minimization are also shown in Fig. 8(e) and (f). From the LOS maps, seasonal deformation with magnitudes up to 12 mm is visible. This is mainly due to the thermal-dilation-induced deformation of the construction material—steel—of the railway station. Fig. 8(g) shows that the main parts of the station and also the bridge in the east of the station undergo vertical seasonal deformation with magnitudes up to 6 mm. This describes 12 mm of motion between summer and winter due to the expansion and contraction of steel material. The east hallway also shows magnitudes of 5 mm of vertical seasonal deformation (yellow part). The most interesting pattern is observed in Fig. 8(h). The right hallway undergoes heavy seasonal deformation in the east-west direction, with magnitudes equal to 12 mm (24 mm between summer and winter). The railway bridge next to the hallway has smaller magnitudes of motion and moves in the reverse direction with respect to the hallway. The rail tracks located on the west side of the station show smaller values of horizontal seasonal motion on the order of 5 mm.

By comparing the estimated components from L1- and L2-norm minimization, another important observation is that, in the results of L2-norm minimization, most of the scatterers are filtered out because of low precisions. This is most apparent in Fig. 8(e), where the deforming hallway in the east direction of the station is not visible. On the contrary, L1-norm minimization preserves the deformation pattern and hence demonstrates its robustness against outliers. Furthermore, to give a rough estimate on the noise level of the estimates obtained from the two mentioned estimators, the empirical standard deviation of the deformation values in the vertical direction was evaluated for both cases [see Fig. 8(e) and (g)]. The calculated values are equal to 2.64 and 3.25 mm for L1 and L2, respectively.

Fig. 9 shows the TomoSAR LOS linear deformation maps of the central station available from each orbit [Fig. 9(a)–(d)] as well as the decomposed motion map in the vertical direction [Fig. 9(e)]. In general, the building and the rail tracks are not influenced by significant linear deformation as it is observed from the LOS maps. There are some parts at the left side of the station in red color which show a construction site and therefore cannot be interpreted as deforming areas. Furthermore, the main building parts of the station show LOS subsidence not higher than 4 mm/year. The decomposed linear vertical map [see Fig. 9(e)] shows that the internal sections of the two main parts undergo vertical subsidence on the order of 2–4 mm/year, while the rail tracks and other parts of the station are stable.

For the second test site, the railway bridge, motion decomposition is only performed on the seasonal deformation map since there is no significant linear ground deformation visible in the area. The result is shown in Fig. 10. From the seasonal LOS maps [see Fig. 10(a)–(d)], it can be observed that the two sections on the railway bridge undergo seasonal deformation with magnitudes up to 7 mm. Another prominent pattern can be seen on the building at the top left of each subfigure. From Fig. 10(e), it is seen that not much of the LOS seasonal deformation can be attributed to the motion in the vertical direction as, for most of the area in the scene, magnitudes not higher than 4 mm are visible. On the other hand, the rail tracks and the building located at the top left are highly influenced by the seasonal deformation in the east-west direction with amplitudes as high as 12 mm.
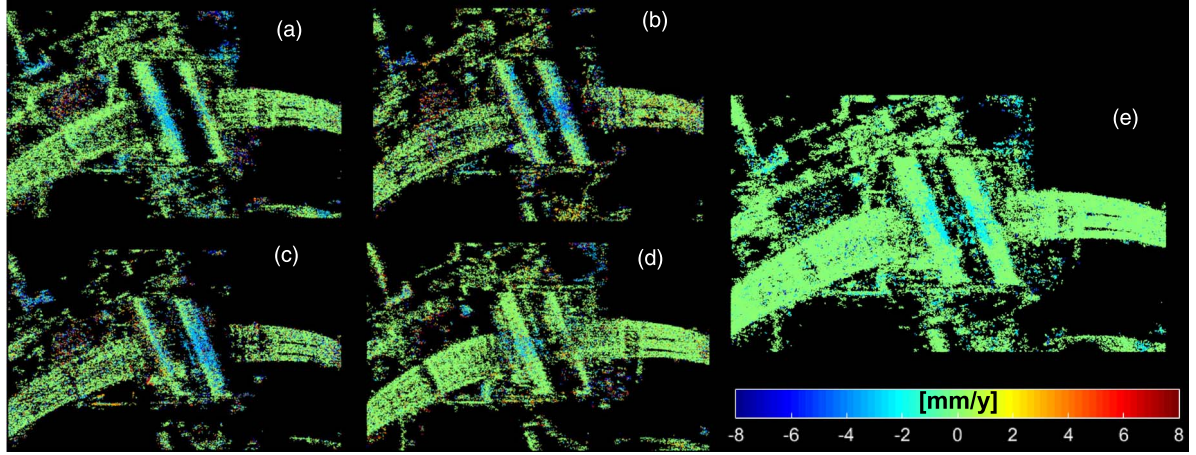
Fig. 9. Berlin central railway station. LOS linear motion of (a) beam 42, (b) beam 57, (c) beam 85, and (d) beam 99. (e) Decomposed linear deformation rate in the vertical direction. Deformation maps are in millimeters per year.
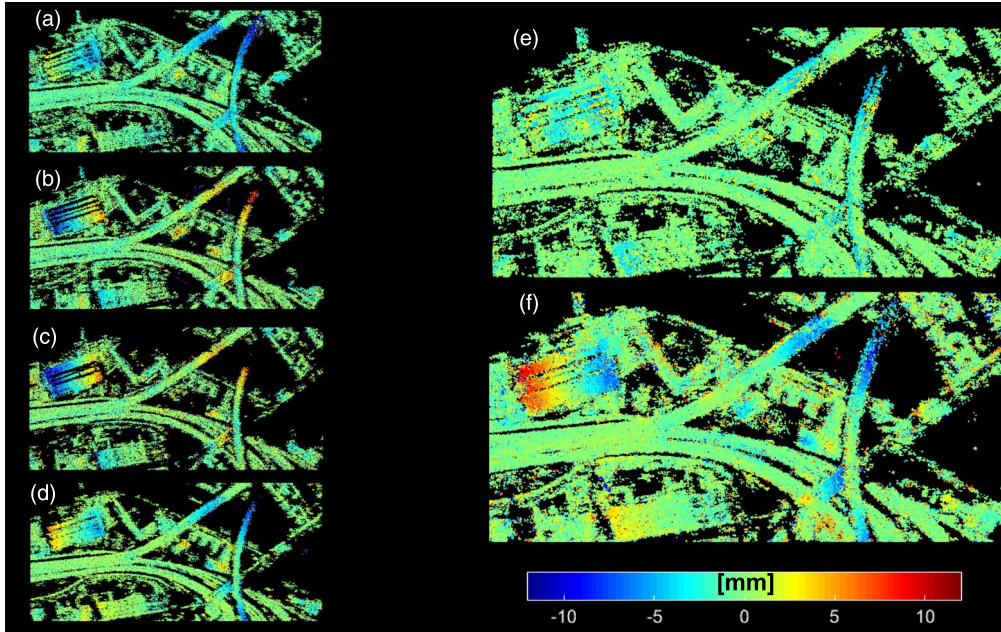

Fig. 10. The railway bridge. LOS amplitude of seasonal motion of (a) beam 42, (b) beam 57, (c) beam 85, and (d) beam 99. Decomposed seasonal deformation in (e) the vertical direction and (f) the east-west direction. Deformation maps are in millimeters.

F. Note on the Computational Complexity

In this section, the computational effort required for the TomoSAR processing and the follow-on motion decomposition is discussed. If we define O (1) to be the computation time for one addition and multiplication, since the tomographic inversion is performed using the MD algorithm, according to [39], it requires at least $O\left(N^2 + NL\right)$, where N is the number of images in the stack (in our study ranging from 102 to 138) and L is the level of discretization in the elevation direction (usually 50–200 for TS-X).

The motion decomposition was carried out by L1-norm minimization, which generally requires more computational effort than the standard least squares method. Following the aforementioned notation, the computational complexity of ordinary L2-norm minimization in a linear functional model is $O\left(C^2 N\right)$, where N denotes the number of observations (inside a cube varying from 3 to 140) and C is the number of unknowns, which, in our study, is always equal to three. For parameter estimation using L1-norm minimization, the analysis is not straightforward since the complexity depends on the type of solver that is used for minimization, the number of iterations, and also the required accuracy to achieve convergence. In this experiment, the L1-norm minimization was carried out by CVX, a package for specifying and solving convex programs [40], [41]. It uses as default the SDTP3 solver [42] to perform the optimization iteratively with roughly ten iterations relevant to our case and the required accuracy of $\epsilon^{1/2}$, where ϵ is the machine precision approximately equal to 2.22×10^{-16}. With these considerations, for the first test site (Berlin central station) which has a total number of scatterers approximately equal to 250 000 from the four stacks, the motion decomposition using L1-norm minimization takes about 700 min, while the adjustment based on L2-norm takes about 40 min on a personal computer utilizing eight-core 3.60-GHz CPUs with a physical

memory of 32 GB. Thus, the former is 17.5 times more computationally expensive than the latter.

VI. CONCLUSION AND OUTLOOK

In this paper, a methodology has been proposed for reconstructing the 3-D displacement components from LOS observations available from tomographic processing of SAR image stacks acquired from multiple viewing geometries. The coregistration of multitrack TomoSAR point clouds was performed using the *geodetic SAR tomography* framework, which provides absolutely localized TomoSAR point clouds. Based on the fused TomoSAR point cloud, the 3-D deformation vector of each scatterer was constructed inside a spatial cube by means of L1-norm minimization, which is more robust with respect to outliers than the more popular L2-norm minimization. The effect of the geometry configuration of the four available TS-X satellites on the motion decomposition was analyzed. As expected, it was seen that, with the current SAR satellites operating in near-polar orbits, the reliable retrieval of deformation in the north-south direction is not possible. However, the component should be considered in the functional model in order to prevent biased estimation of the two other motion components. The decomposition methodology was used on TS-X very high resolution spotlight data available from multiple viewing geometries over the city of Berlin. The results showed that man-made objects constructed with steel material undergo seasonal deformation with magnitudes up to 24 mm between summer and winter in the east-west direction.

Future work will concentrate on the following: 1) combining both L1-norm and L2-norm minimization in the estimation process to provide robust displacement estimates accompanied by quality indicators and 2) performing GPS measurements in order to produce absolute deformation maps.

REFERENCES

[1] X. X. Zhu, "Very High Resolution Tomographic SAR Inversion for Urban Infrastructure Monitoring—A Sparse and Nonlinear Tour," Ph.D. dissertation, Dept. Civil, Geo Environmental Eng., Tech. Univ. München, München, Germany, 2011.

[2] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.

[3] G. Fornaro, F. Serafino, and F. Soldovieri, "Three-dimensional focusing with multipass SAR data," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 3, pp. 507–517, Mar. 2003.

[4] X. X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 12, pp. 4296–4308, Dec. 2010.

[5] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 1, pp. 8–20, Jan. 2001.

[6] B. M. Kampes, *Radar Interferometry: Persistent Scatterer Technique.* Dordrecht, The Netherlands: Springer-Verlag, 2006.

[7] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 37–44, Jan. 2005.

[8] X. X. Zhu and R. Bamler, "Demonstration of super-resolution for tomographic SAR imaging in urban environment," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 8, pp. 3150–3157, Aug. 2012.

[9] R. F. Hanssen, *Radar Interferometry: Data Interpretation and Error Analysis.* Dordrecht, The Netherland: Kluwer, 2001.

[10] J. Hu, Z. W. Li, X. L. Ding, J. J. Zhu, L. Zhang, and Q. Sun, "Resolving three-dimensional surface displacements from InSAR measurements: A review," *Earth-Sci. Rev.*, vol. 133, pp. 1–17, Jun. 2014.

[11] T. J. Wright, B. E. Parsons, and Z. Lu, "Toward mapping surface deformation in three dimensions using InSAR," *Geophys. Res. Lett.*, vol. 31, no. 1, pp. 1–5, 2004.

[12] S. Gernhardt, X. Y. Cong, M. Eineder, S. Hinz, and R. Bamler, "Geometrical fusion of multitrack PS point clouds," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 1, pp. 38–42, Jan. 2012.

[13] Y. Wang and X. X. Zhu, "Automatic feature-based geometric fusion of multiview TomoSAR point clouds in urban area," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 3, pp. 953–965, Mar. 2015.

[14] X. X. Zhu and R. Bamler, "Let's do the time warp: Multicomponent nonlinear motion estimation in differential SAR tomography," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 4, pp. 735–739, Jul. 2011.

[15] S. Gernhardt and R. Bamler, "Deformation monitoring of single buildings using meter-resolution SAR data in PSI," *ISPRS J. Photogramm. Remote Sens.*, vol. 73, pp. 68–79, Sep. 2012.

[16] X. X. Zhu, S. Montazeri, C. Gisinger, R. F. Hanssen, and R. Bamler, "Geodetic SAR tomography," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 1, pp. 18–35, Jan. 2016.

[17] C. Gisinger *et al.*, "Precise three-dimensional stereo localization of corner reflectors and persistent scatterers with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1782–1802, Apr. 2015.

[18] M. Eineder, C. Minet, P. Steigenberger, X. Y. Cong, and T. Fritz, "Imaging geodesy—Toward centimeter-level ranging accuracy with TerraSAR-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 661–671, Feb. 2011.

[19] X. Y. Cong, U. Balss, M. Eineder, and T. Fritz, "Imaging geodesy—Centimeter-level ranging accuracy with TerraSAR-X: An update," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 5, pp. 948–952, Sep. 2012.

[20] I. G. Cumming and F. H. Wong, *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation.* Boston, MA, USA: Artech House, 2005.

[21] E. M. Mikhail and F. E. Ackermann, *Observations and Least Squares.* Washington, DC, USA: Univ. Press Amer., 1982.

[22] P. J. G. Teunissen, D. G. Simons, and C. C. J. M. Tiberius, *Probability and Observation Theory.* Delft, The Netherlands: Delft Univ. Technol., 2009.

[23] S. C. Narula and J. F. Wellington, "The minimum sum of absolute errors regression: A state of the art survey," *Int. Statist. Rev./Revue Int. de Statistique*, vol. 50, no. 3, p. 317, Dec. 1982.

[24] S. Auer, S. Gernhardt, and R. Bamler, "Ghost persistent scatterers related to multiple signal reflections," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 5, pp. 919–923, Sep. 2011.

[25] A. Amiri-Simkooei, "Formulation of L1 norm minimization in Gauss–Markov models," *J. Surveying Eng.*, vol. 129, no. 1, pp. 37–43, Feb. 2003.

[26] L. Gray, "Using multiple RADARSAT InSAR pairs to estimate a full three-dimensional solution for glacial ice movement," *Geophys. Res. Lett.*, vol. 38, no. 5, pp. 1–6, Mar. 2011.

[27] P. Misra and P. Enge, *Global Positioning System: Signals, Measurements, and Performance*, 2nd ed. Lincoln, MA, USA: Ganga-Jamuna Pr, 2006.

[28] H. Ansari, F. De Zan, A. Parizzi, M. Eineder, K. Goel, and N. Adam, "Measuring 3-D surface motion with future SAR systems based on reflector antennae," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 2, pp. 272–276, Feb. 2016.

[29] R. Bamler, M. Eineder, N. Adam, X. Zhu, and S. Gernhardt, "Interferometric potential of high resolution spaceborne SAR," *Photogrammetrie—Fernerkundung—Geoinf.*, vol. 2009, no. 5, pp. 407–419, Nov. 2009.

[30] N. Adam, F. R. Gonzalez, A. Parizzi, and R. Brcic, "Wide area persistent scatterer interferometry: Current developments, algorithms and examples," in *Proc. IEEE IGARSS*, 2013, pp. 1857–1860.

[31] X. X. Zhu, Y. Wang, S. Gernhardt, and R. Bamler, "Tomo-GENESIS: DLR's tomographic SAR processing system," in *Proc. IEEE JURSE*, 2013, pp. 159–162.

[32] S. Gernhardt, "High Precision 3D Localization and Motion Analysis of Persistent Scatterers Using Meter-Resolution Radar Satellite Data," Ph.D. dissertation, Tech. Univ. München, München, Germany, 2012.

[33] S. Montazeri, "The Fusion of SAR Tomography and Stereo-SAR for 3D Absolute Scatterer Positioning," M.S. thesis, Dept. Geosci. Remote Sens. (GRS), Delft Univ. Technol., Delft, The Netherlands, 2014.

[34] S. Gernhardt, N. Adam, M. Eineder, and R. Bamler, "Potential of very high resolution SAR for persistent scatterer interferometry in urban areas," *Ann. GIS*, vol. 16, no. 2, pp. 103–111, Aug. 2010.

[35] X. X. Zhu and M. Shahzad, "Facade reconstruction using multiview spaceborne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3541–3552, Jun. 2014.

[36] M. Shahzad and X. X. Zhu, "Robust reconstruction of building facades for large areas using spaceborne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 752–769, Feb. 2015.

[37] M. Shahzad and X. X. Zhu, "Automatic detection and reconstruction of 2-D/3-D building shapes from spaceborne TomoSAR point clouds," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1292–1310, Mar. 2016.

[38] Y. Wang, X. X. Zhu, B. Zeisl, and M. Pollefeys, "Fusing meter-resolution 4-D InSAR point clouds and optical images for semantic urban infrastructure monitoring," *IEEE Trans. Geosci. Remote Sens.*, in press.

[39] Y. Wang, X. X. Zhu, and R. Bamler, "An efficient tomographic inversion approach for urban mapping using meter resolution SAR image stacks," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 7, pp. 1250–1254, Jul. 2014.

[40] M. Grant and S. Boyd, *CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1*, 2014.

[41] M. Grant and S. Boyd, "Graph implementations for nonsmooth convex programs," in *Recent Adv. Learn. Control*, vol. 371, V. D. Blondel, S. P. Boyd, and H. Kimura, Eds. London, U.K.: Springer-Verlag, 2008, pp. 95–110.

[42] R. H. Tütüncü, K. C. Toh, and M. J. Todd, "Solving semidefinite-quadratic-linear programs using SDPT3," *Math. Programm.*, vol. 95, no. 2, pp. 189–217, Feb. 2003.

Sina Montazeri received the Bachelor's degree in geodesy and surveying engineering from the University of Isfahan, Isfahan, Iran, in 2011 and the Master's degree in geomatics from the Delft University of Technology, Delft, The Netherlands, in 2014. He is currently working toward the Ph.D. degree in the Remote Sensing Technology Institute, German Aerospace Center (DLR), Wessling, Germany.

From April 2013 to May 2014, he was a Part-Time Research Assistant with the Remote Sensing Technology Institute, DLR, working on absolute localization of point clouds obtained from SAR tomography. His research interests include advanced InSAR techniques for deformation monitoring of urban infrastructure, image and signal processing relevant to radar imagery, and incorporation of geodetic estimation framework in multipass InSAR.

Xiao Xiang Zhu (S'10–M'12–SM'14) received the Bachelor's degree in space engineering from the National University of Defense Technology, Changsha, China, in 2006 and the M.Sc. degree, the Dr.-Ing. degree, and the "Habilitation" in the field of signal processing from the Technical University of Munich (TUM), Munich, Germany, in 2008, 2011, and 2013, respectively.

Since 2011, she has been a Scientist with the Remote Sensing Technology Institute, German Aerospace Center (DLR), Oberpfaffenhofen, Germany, where she is the Head of the Team Signal Analysis. Since 2013, she has also been a Helmholtz Young Investigator Group Leader and appointed as TUM Junior Fellow. In 2015, she was appointed as a Professor with Signal Processing in Earth Observation, TUM. She was a Guest Scientist or Visiting Professor with the Italian National Research Council (CNR-IREA), Naples, Italy, Fudan University, Shanghai, China, University of Tokyo, Tokyo, Japan, and University of California, Los Angeles, CA, USA, in 2009, 2014, 2015, and 2016, respectively. Her main research interests are the following: advanced InSAR techniques such as high-dimensional tomographic SAR imaging and SqueeSAR, computer vision in remote sensing including object reconstruction and multidimensional data visualization, big data analysis in remote sensing, and modern signal processing, including innovative algorithms such as sparse reconstruction, nonlocal means filter, robust estimation, and deep learning, with applications in the field of remote sensing such as multi-/hyperspectral image analysis.

Dr. Zhu is an Associate Editor of the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING.

Michael Eineder (M'02–SM'10–F'16) received the Diploma degree in electrical engineering from the Technische Universität München (TUM), Munich, Germany, in 1990 and the Dr. rer. nat. degree from the University of Innsbruck, Innsbruck, Austria, in 2004.

He has been a TUM Honorary Professor since 2013. He is currently the Head of the SAR Signal Processing Department, Remote Sensing Technology Institute, German Aerospace Center (DLR), Wessling, Germany. His responsibilities encompass the development of SAR and InSAR processing systems for current and future radar missions. Furthermore, he is a Part-Time Lecturer of remote sensing with the TUM. He joined DLR in 1990, where he has worked on a variety of international missions including SIR-C/X-SAR, SRTM/X-SAR in cooperation with NASA, ERS-1 (ESA), TerraSAR-X, and TanDEM-X (Germany). His current research interest is focused on the design of SAR missions for geophysical applications and on imaging geodesy, an absolute positioning technique exploiting high-resolution SAR.

Richard Bamler (M'95–SM'00–F'05) received the Diploma degree in electrical engineering, the Doctorate in Engineering, and the "Habilitation" in the field of signal and systems theory from the Technische Universität München (TUM), Munich, Germany, in 1980, 1986, and 1988, respectively.

He was with the TUM from 1981 to 1989, working on optical signal processing, holography, wave propagation, and tomography. He joined the German Aerospace Center (DLR), Oberpfaffenhofen, Germany, in 1989, where he is currently the Director of the Remote Sensing Technology Institute. In early 1994, he was a Visiting Scientist with the Jet Propulsion Laboratory (JPL) in preparation of the SIC-C/X-SAR missions, and in 1996, he was a Guest Professor with the University of Innsbruck, Innsbruck, Austria. Since 2003, he has held a full professorship in remote sensing technology with the TUM as a double appointment with his DLR position. His teaching activities include university lectures and courses on signal processing, estimation theory, and SAR. He, his team, and his institute have been working on SAR and optical remote sensing, image analysis and understanding, stereo reconstruction, computer vision, ocean color, passive and active atmospheric sounding, and laboratory spectrometry. They were and are responsible for the development of the operational processors for SIR-C/X-SAR, SRTM, TerraSAR-X, TanDEM-X, Tandem-L, ERS-2/GOME, Envisat/SCIAMACHY, MetOp/GOME-2, Sentinel 5p, EnMAP, etc. His research interests are the algorithms for optimum information extraction from remote sensing data with emphasis on SAR. This involves new estimation algorithms, like sparse reconstruction and compressive sensing. He has devised several high-precision algorithms for monostatic and bistatic SAR processing, SAR calibration and product validation, GMTI for traffic monitoring, SAR interferometry, phase unwrapping, persistent scatterer interferometry, differential SAR tomography, and data fusion.

A.5 Zhu, X.X., Wang, Y., Montazeri, S., Ge, N., 2018. A Review of Ten-Year Advances of Multi-Baseline SAR Interferometry Using TerraSAR-X Data. Remote Sensing 10(9), 1374: 1-32

remote sensing

MDPI

Review

A Review of Ten-Year Advances of Multi-Baseline SAR Interferometry Using TerraSAR-X Data

Xiao Xiang Zhu [1,2,*] [ID], **Yuanyuan Wang** [2] [ID], **Sina Montazeri** [1] [ID] and **Nan Ge** [1]

[1]   Remote Sensing Technology Institute, German Aerospace Center, 82234 Weßling, Germany;
      Sina.Montazeri@dlr.de (S.M.); Nan.Ge@dlr.de (N.G.)
[2]   Signal Processing in Earth Observation, Technical University of Munich, 80333 Munich, Germany;
      wang@bv.tum.de
[*]   Correspondence: xiaoxiang.zhu@dlr.de; Tel.: +49-8153-28-3531

Abstract: Since its launch in 2007, TerraSAR-X has continuously provided spaceborne synthetic aperture radar (SAR) images of our planet with unprecedented spatial resolution, geodetic, and geometric accuracy. This has brought life to the once inscrutable SAR images, which deterred many researchers. Thanks to merits like higher spatial resolution and more precise orbit control, we are now able to indicate individual buildings, even individual floors, to pinpoint targets within centimeter accuracy. As a result, multi-baseline SAR interferometric (InSAR) techniques are flourishing, from point target-based algorithms, to coherent stacking techniques, to absolute positioning of the former techniques. This article reviews the recent advances of multi-baseline InSAR techniques using TerraSAR-X images. Particular focus was put on our own development of persistent scatterer interferometry, SAR tomography, robust estimation in distributed scatterer interferometry and absolute positioning using geodetic InSAR. Furthermore, by introducing the applications associated with these techniques, such as 3D reconstruction and deformation monitoring, this article is also intended to give guidance to wider audiences who would like to resort to SAR data and related techniques for their applications.

Keywords: multi-baseline; multi-pass; PS; DS; geodetic; TomoSAR; D-TomoSAR; PSI; robust estimation; covariance matrix; InSAR; SAR; review

1. Introduction

1.1. Overview of Multi-Baseline InSAR

Since its launch in 2007, TerraSAR-X has continuously revealed synthetic aperture radar (SAR) images of unprecedented high resolution from space. This has brought life to the once obscure and sometimes inscrutable SAR images that deterred many researchers. Figure 1 shows a comparison of the medium resolution ERS image and a high resolution TerraSAR-X spotlight image of the same area in Las Vegas. Individual buildings are for the first time interpretable by the naked eye from spaceborne SAR images, because the 1-m resolution in spotlight mode is well beyond the inherent scale of the 3-m floor height typical of urban infrastructure. This marks the start of an era of urban infrastructure monitoring using spaceborne SAR images. Currently, the staring spotlight mode provides images with a resolution up to 25 cm, from which the mapping of individual window edges is even possible. This breakthrough in spatial resolution, together with the precise orbit determination with sub-centimeter accuracy [1,2], positions TerraSAR-X images as a perfect dataset for long-term repeated monitoring of large areas with precision and high resolution.
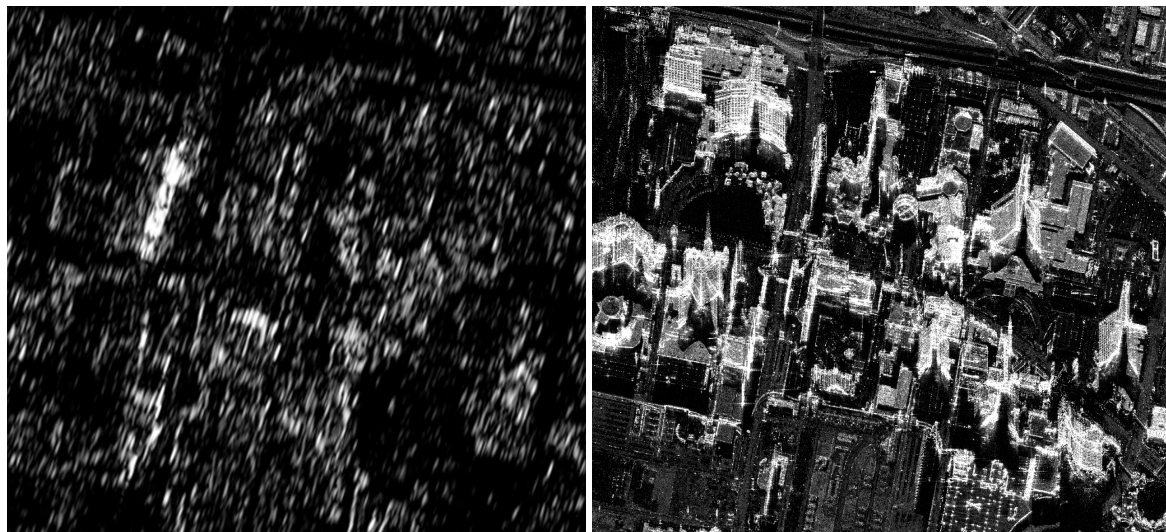
Figure 1. Comparison of medium (ERS) and high (TerraSAR-X) resolution SAR images of downtown Las Vegas [3].

Among the many promising InSAR techniques that prospered in the past decade, multi-baseline, especially multi-pass, InSAR techniques are undoubtedly one of the jewels in the crown. They build up invaluable data cubes of long-term image time series. For example, the TerraSAR-X revisit time of 11 days allows monthly deformation signals of the Earth's surface, such as ground subsidence, to be retrieved using techniques like persistent scatterer interferometry (PSI). For monitoring dense urban areas, SAR tomography (TomoSAR) and its differential form, D-TomoSAR inversion, are the most competent techniques because of their capability of layover separation. They generate point clouds with density comparable to that of a LiDAR. Both PSI and TomoSAR produce highly accurate parameter estimates, because they work on highly coherent point targets. Therefore, they are often the workhorses for deformation monitoring and 3D reconstruction, especially in urban areas. To complement these techniques, distributed scatterer (DS)-based techniques, such as SqueeSAR [4], robust InSAR optimization (RIO) [5] and coherence tomography enable dense monitoring of deformation in areas of low interferometric coherence, such as volcanic areas. Among them, some algorithms, such as RIO, address the statistical robustness of estimators to ensure the reliability of the accuracy of the estimates in operational processing over large areas. Despite the many advantages of multi-baseline InSAR, they are still relative measures, as the estimates are often relative to a local reference point whose 3D position is unknown. Such differential operation is often performed in multi-baseline InSAR in order to mitigate some common phase errors, such as atmospheric delay. It was only until recently, that geodetic InSAR [6] bridged the gap between multi-baseline InSAR techniques and absolute positioning using SAR imaging geodesy [7] to produce absolute 3D (and higher dimensional) InSAR point clouds. It is an important piece of the components of the ecosystem of Earth observation using SAR data. Multi-baseline InSAR techniques that were once only a relative measure can now be employed as geodetic techniques to provide centimeter-level absolute positioning and millimeter-level relative deformation monitoring.

1.2. Principle of Multi-Baseline InSAR

InSAR is the technique of using SAR as an interferometer. Multi-baseline InSAR techniques exploit the interferometric phase (i.e., the phase difference) of multiple complex-valued SAR images. These images are acquired at different satellite positions, time, or frequency, and hence, they create spatial, temporal baselines, or Δk-radar when forming interferograms. For TerraSAR-X images, such multi-baseline configuration is usually acquired in a repeat-pass manner (hence "multi-pass"), except if the twin satellite TanDEM-X was employed. Figure 2 shows the multi-baseline InSAR

configuration in an urban scenario at a fixed azimuth position. The TerraSAR-X satellite flies perpendicular into the screen/paper. The term *r* indicates the line of sight (LOS), i.e., the slant range direction, of the sensor; *s* is the elevation direction that is perpendicular to the range and azimuth. The blue outline on the surface indicates the area illuminated by radar pulses. The elongated ellipse is the range-elevation tube within which all the objects are imaged into a single pixel in the focused SAR image. The cross-section of the tube naturally depends on the range and azimuth resolution of the sensor. The extent of the tube Δ*s* is much larger than the dimension of the cross-section because of the large distance between the sensor and the object, as well as the small angular diversity among different acquisitions. Therefore, it is common that several objects, such as a building roof, tree and ground, are layovered in a single pixel in a TerraSAR-X image.
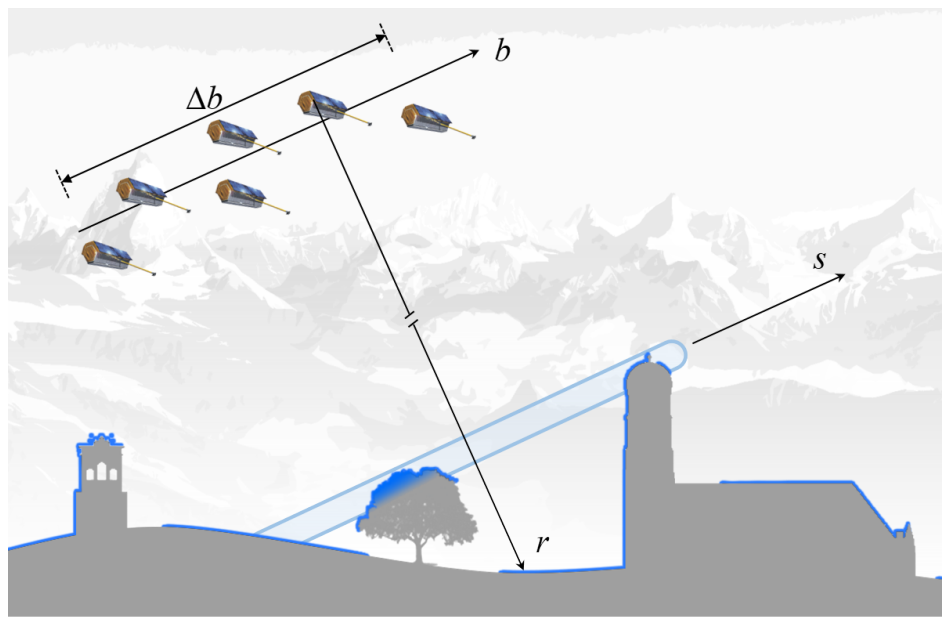

Figure 2. Schematic drawing of the principle of multi-baseline InSAR at a fixed azimuth position (modified after [8]). The TerraSAR-X satellite flies away from the reader into the screen/paper. The line-of-sight, i.e., the range direction, of the sensor is indicated by *r*. The range timing is always delayed after propagation through the atmosphere. The term *s* is the elevation direction that is perpendicular to the range. The blue outline on the surface indicates the area illuminated by radar pulses. The elongated ellipse is the range-elevation resolution cell with in which all the objects are imaged into a single pixel in the final SAR image. It is very common that several objects, such as a building roof, tree and ground, are layovered in a single pixel in a TerraSAR-X image.

If one considers a single phase center in the range-azimuth-elevation tube without layover (i.e., single scatterer model), the absolute interferometric phase of the *n*-th measurement in a multi-baseline InSAR stack is [9]:

$$\phi_n = -\frac{4\pi}{\lambda}\frac{b_n s}{R} + \phi_{defo} + \phi_{atmo} + \phi_{error}, \tag{1}$$

where λ is the wavelength of the SAR electromagnetic wave, b_n is the baseline of the *n*-th image, *s* is the elevation of the single scatterer and *R* is the nominal range which is the distance of the SAR sensor to a zero-elevation point. The deformation phase ϕ_{defo} is often modeled as a function $d(t_n)$ (e.g., linear or periodic) of the acquisition time t_n. The interferometric phase is always delayed due to atmospheric propagation. In multi-baseline InSAR, such atmospheric phase delay ϕ_{atmo} is mitigated by subtracting a nearby reference point. This renders multi-baseline InSAR a relative measure, unless the absolute position of the reference point is known a priori.

Based on Equation (1), the forward system model of multi-baseline InSAR measurement can be expressed as Equation (2), where g_n is the pixel value at the *n*-th image, and $\gamma(s)$ is the reflectivity profile along the elevation direction. Since a far-field antenna acts like a Fourier transform to the signal in the resolution cell, each measurement is actually the Fourier transform at a specific frequency that is linearly proportional to the perpendicular baseline b_n to the master satellite position. This is also known as the system model for TomoSAR [10–13].

$$g_n = \int_{\Delta s} \gamma(s) \exp\left(-j\frac{4\pi b_n}{\lambda R}s\right) ds. \tag{2}$$

In the case of differential TomoSAR (D-TomoSAR), Equation (2) is extended into higher dimensions [14–16].

Equation (2) can be written in a more compact matrix form as:

$$\mathbf{g} = \mathbf{R}\gamma, \tag{3}$$

where $\mathbf{R}$ and γ are the discretized Fourier matrix and the reflectivity profile along the direction *s*, respectively. Estimating γ is essentially a spectral estimation problem. In the case of PSI- or DS-based interferometry that assumes a single phase center in the resolution cell, it is basically a spectral estimation of a single frequency. Equation (3) will degenerate to either Equation (4) for the deterministic PS mode or Equation (5) for the stochastic DS model.

$$\mathbf{g} = \mathbf{r}(s_1)\gamma_1, \tag{4}$$

$$E\left(\mathbf{g}\mathbf{g}^H\right) = \mathbf{r}(s_1)\mathbf{C_{gg}}\mathbf{r}(s_1)^H, \tag{5}$$

where s_1 is the elevation of the single phase center, $\mathbf{r}(s_1)$ is the column in $\mathbf{R}$ associated with s_1, γ_1 is the complex-valued brightness of the PS and $\mathbf{C_{gg}}$ is the covariance matrix of the DS.

1.3. The Structure of This Paper

The rest of this article introduces the recent development of the aforementioned techniques, each in its respective section. In Section 2, we introduce the development of PS-based methods following their improvements in estimates accuracy that in turn refers to the reconstructed point density, as well as the reduction of the required number of images for a reliable estimation. Section 3 focuses on the development of robust InSAR techniques based on DS. Section 4 focuses on the evolution of TerraSAR-X absolute positioning from a single target to many targets and eventually to the fusion with multi-baseline InSAR.

2. Advances in Point Scatterer-Based Methods

This section focuses on the advances of PS-based methods, i.e., PSI and TomoSAR/D-TomoSAR in urban areas. Their development mainly focuses on the improvement of estimation accuracy, which in turn increases the density of the retrieved point cloud or reduces the number of interferograms required for a reliable estimation.

Both PSI and TomoSAR utilize a single-master configuration to extract time-coherent scatterers from SAR images. The major difference between the two methods is the number of scatterers that are assumed within a resolution cell, which requires different spectral estimators to be employed in the parameter retrieval. However, over the past two decades, PSI has made substantial development, so that it usually refers to a full processing chain including interferogram formation and reference network construction. Therefore, PSI is often employed as a preprocessing step for TomoSAR. Several variations of PSI that differ in algorithmic details have been introduced in recent years. For a full review

of these techniques the reader is referred to [17]. Although the specifics of existing PSI algorithms are different, the following workflow is widely acknowledged:

Step 1   Differential interferogram formation: From a stack of $N + 1$ co-registered SAR images, a master acquisition is selected. Subsequently N interferograms are computed, while their topographic phase components are removed using a reference digital elevation model (DEM).

Step 2   Reference network construction: Scatterers presumed to be the most phase-stable ones are selected. The detection can be carried out using various methods, such as thresholding on the amplitude dispersion index (ADI) [18] or on the signal-to-clutter ratio (SCR) [19]. These PS candidates are connected to form a reference network. Through the PS double-difference phase measurements, i.e., difference in time and space, differential topography and differential motion parameters are estimated on arcs.

Step 3   Atmospheric phase estimation: The differential topography estimates are integrated with respect to an arbitrarily chosen reference point so that the topographic phase components are removed from the interferometric phases. The remaining phase contributions include deformation, atmosphere, and noise. Then a low-pass filtering in the spatial domain and a high-pass filtering in the temporal domain extracts the atmospheric component, which is interpolated over the entire scene and subtracted from the differential interferograms.

Step 4   PS densification: Additional PS are computed from the corrected differential interferograms. These PS are connected to the nearest point(s) in the reference network and their modeled parameters are estimated.

Step 5   PS geocoding: The DEM height of each PS is added to its differential height estimate. The radar timing of each PS and its updated height are geocoded using satellite orbit and a reference ellipsoid to represent the PS coordinates in a common geodetic coordinate system.

The processing steps for TomoSAR are similar to those of PSI, except the fourth step is replaced with higher-order spectral estimators that can be enumerated as followings.

- The full reflectivity profile is reconstructed using higher-order spectral estimation techniques.
- The scatterers' positions and motion parameters are determined by detecting maxima on the reflectivity profile.

2.1. Overview of Advances

For each of the steps delineated above, numerous improvements have been suggested in the literature. For example, in the reference network step, [20,21] consider the geometry of the connections among arcs to construct a redundant reference network, while dense differential PS pairs were used in [22] to form the network. In terms of network inversion, to robustly retrieve the topography and deformation estimates of the PS in the reference network, a ℓ_1 norm outlier rejection scheme was proposed after the LAMBDA estimation [23]. In [24,25], numerical weather data were used to simulate and mitigate tropospheric delay. For a detailed comparison of widely used PSI techniques, the interested reader is referred to [21].

The development of TomoSAR has been mainly focused on the improvement of the spectral estimator and the scatterer detector. Studies have been conducted to improve the maximum likelihood estimator (MLE) by restricting the support of the signal (i.e., nonlinear least square) [3,13], by ℓ_2 norm regularization (i.e., the Tikhonov method) [15], and by ℓ_1 norm regularization (i.e., compressive sensing-based method) [26,27]. The SL1MMER algorithm proposed in [27] was also recently extended to the M-SL1MMER [28], which exploits group sparsity in the urban environment. M-SL1MMER achieves a comparable result with far fewer images than SL1MMER and other algorithms. Several studies have also addressed the efficiency and robustness of the detection of scatterers. For example, [29] describes the optimal detection of multiple scatterers, and [27,30–32] address scatterer detection in the super-resolution regime where the distance among scatterers is less than the elevation resolution.

In general, TomoSAR is so far the most competent multi-baseline InSAR method for urban area monitoring. However, the relatively high computational cost limits it for extensive uses like PSI, especially for the CS-based TomoSAR algorithms. Therefore, combining PSI and TomoSAR has also been proposed to improve the computational efficiency of TomoSAR processing [22,33,34]. Only recently, an efficient sparse recovery algorithm was proposed, which made city-scale 3D/4D reconstruction directly using SL1MMER operational [35].

2.2. Very High Resolution PSI

PSI is undoubtedly the workhorse for deformation monitoring of large areas, owing to its computational efficiency and reliability in the accuracy of the deformation estimates. As mentioned earlier, estimating the unknown elevation and deformation parameters in PSI is a spectral estimation of a single frequency. The spectral estimator is essentially a periodogram that can be expressed as follows.

$$\hat{\theta} = \arg\max_{\theta} \left\{ \left| \frac{1}{N} \sum_{n=1}^{N} g_n \exp\left(-j\phi_n\left(\theta\right)\right) \right| \right\} \approx \arg\max_{\theta} \left\{ \left| \frac{1}{N} \sum_{n=1}^{N} \frac{g_n}{|g_n|} \exp\left(-j\phi_n\left(\theta\right)\right) \right| \right\}, \quad (6)$$

where θ denotes the parameters, including the elevation s and the deformation parameters, and $\phi_n\left(\theta\right)$ is the modeled phase of the PS in the n-th image (i.e., Equation (1)). Often, the amplitude of **g** is dropped in the estimation [18], since it barely changes the estimates for PS of high signal-to-noise ratio (SNR).

Employing very high resolution (VHR) PSI, it is now possible to detect very localized deformation patterns even on different parts of a single building [36]. Apart from its deformation monitoring capability, VHR PSI leads to detailed 3D reconstruction of urban areas owing to the high density PSI point clouds. It can typically produce 40,000 to 100,000 PS per square kilometer using TerraSAR-X high resolution spotlight images [37,38]. The 3D reconstruction capability has even been strengthened by the geometrical fusion of PSI point clouds obtained from different viewing geometries, i.e., along-heading and cross-heading orbits [39]. Especially in the case of cross-heading orbits, that is, the combination of point clouds from ascending and descending orbits, point cloud fusion provides a shadow-free point cloud of the observed area. It also allows a decomposition of the raw LOS PSI deformation measurements into 3D displacement vectors in geodetic coordinate system [36,40,41].

2.3. Differential TomoSAR

Unlike PSI, D-TomoSAR retrieves the full reflectivity profile γ, and detects prominent peaks from it. Therefore, D-TomoSAR is inherently a more competent method for urban area monitoring than PSI. The MLE (under complex Gaussian noise) of γ can be expressed as follows.

$$\hat{\gamma}_{MLE} = \arg\min_{\gamma} \frac{1}{2} \|\mathbf{g} - \mathbf{R}\gamma\|_2^2. \quad (7)$$

During the last decade, we have developed a suite of algorithms named Tomo-GENESIS [42] to address both the methodological and practical aspects of D-TomoSAR. For example, the Tomo-GENESIS suite includes both conventional linear estimators [15] and the compressive sensing (CS)-based estimator that works in the superresolving regime [27,30,31], as well as a computationally efficient processing pipeline [22], the fusion of TomoSAR point clouds from multiple aspects [43] and 3D object reconstruction from TomoSAR point clouds [44–46].

2.3.1. Conventional (Non-Superresolving) D-TomoSAR

For spaceborne data, the number of acquisitions is usually far less than the discretization of γ. Therefore, Equation (3) is often under-determined. A popular method before the invention of CS-based

TomoSAR techniques to regularize the equation system was to employ the ℓ_2 norm regularization that is also known as Tikhonov regularization. The regularized estimator is shown as follows.

$$\hat{\boldsymbol{\gamma}}_{\ell_2} = \arg\min_{\boldsymbol{\gamma}} \frac{1}{2} \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2^2 + \lambda_{\ell_2} \|\boldsymbol{\gamma}\|_2^2, \tag{8}$$

where λ_{ℓ_2} is the regularization parameter. We have implemented the estimator using singular value decomposition with Wiener filtering on the system matrix $\mathbf{R}$. Therefore this algorithm is also known as SVD-Wiener in the community [15].

This type of estimator is also a maximum a posteriori (MAP) estimator. It is the optimal Bayesian estimator that minimizes posterior expected loss. Experiments showed promising performance on TerraSAR-X image stacks [13]. However, in the classical Nyquist–Shannon sampling regime, the resolution of the reconstructed reflectivity profile is limited by the so-called Rayleigh resolution (see Equation (9)) [15] that is governed by the spread of the baseline Δb.

$$\rho_s = \frac{\lambda R}{2\Delta b}. \tag{9}$$

2.3.2. Super-Resolving D-TomoSAR

For dense urban areas, closely spaced objects often coexist in a range-azimuth-elevation resolution cell. These objects cannot be resolved by conventional tomographic inversion algorithms. This is where CS-based super-resolving tomographic inversion comes to play, as it can achieve super-resolution in the estimate of $\boldsymbol{\gamma}$, if it is sparse. The CS-based TomoSAR estimator can be generally expressed in a similar form as Equation (8), except that the ℓ_2 regularization term is replaced by the signal sparsity term, i.e., the ℓ_0 norm. Because of the nonconvexity of the ℓ_0 norm, it is often relaxed by the ℓ_1 norm in optimization, such as the SL1MMER "scale-down by ℓ_1 norm minimization, model selection, and estimation reconstruction" algorithm proposed in [27]. The $\ell_2 + \ell_1$ norm estimator can be expressed as follows.

$$\hat{\boldsymbol{\gamma}}_{\ell_0} = \arg\min_{\boldsymbol{\gamma}} \frac{1}{2} \|\mathbf{g} - \mathbf{R}\boldsymbol{\gamma}\|_2^2 + \lambda_K \|\boldsymbol{\gamma}\|_1, \tag{10}$$

where λ_K is a regularization parameter (K being the sparsity, i.e., the number of discrete scatterers). In practice, the minimization of the ℓ_0 norm is often relaxed by the ℓ_1 norm for better convexity in the optimization.

Because of their super-resolving ability and the robustness of the ℓ_1 norm minimization, CS-based D-TomoSAR algorithms are the state of the art in term of the accuracy of the parameter estimate and the performance of scatterer detection. This in turn increases the density of the reliable points. Figure 3 is a comparison of the point cloud retrieved by PSI and SL1MMER of the same building (Bellagio Hotel, Las Vegas). SL1MMER retrieves many more points than PSI. Yet, CS-based algorithms are less computationally efficient than the conventional TomoSAR. To cope with large area processing, we enriched Tomo-GENESIS with an approach [22,47] that integrates PSI, conventional TomoSAR, and super-resolving TomoSAR. Recently, we have developed a fast and accurate ℓ_1-regularized least square solver with application to D-TomoSAR [35]. This new solver offers a speedup of one or two orders of magnitude than typical second order cone programming. With above-mentioned advances, we are able to reconstruct a high-quality TomoSAR point cloud of an entire city with density comparable to that of LiDAR. For a better overview of the capability of the aforementioned methods, Table 1 summarizes the typical density of the point cloud reconstructed by PSI and D-TomoSAR using a TerraSAR-X high resolution sliding spotlight image stack.

Table 1. Comparison of the typical density of the point cloud reconstructed by PSI and D-TomoSAR using TerraSAR-X high resolution spotlight image stack.

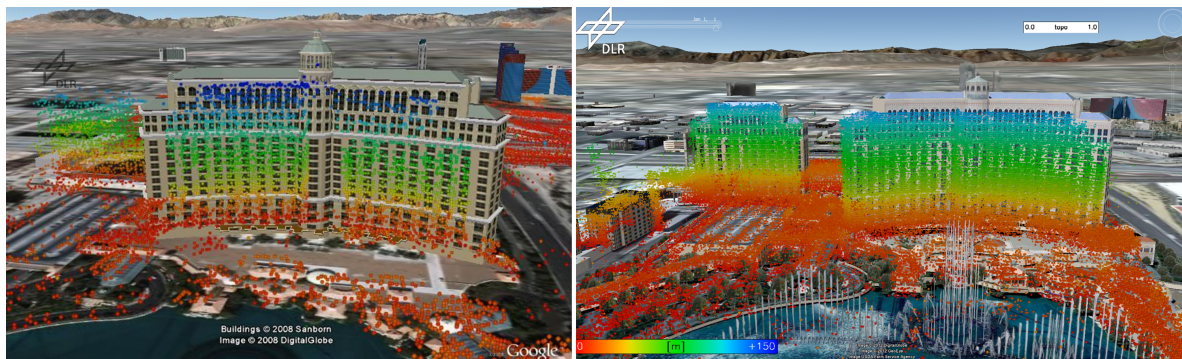|	Density (thousand/km^2)
PSI [38]	40–100
D-TomoSAR (non-superresolving) [15]	150–250
D-TomoSAR (SL1MMER) [48]	500–1500

Figure 3. Comparison of the density of the 3D point cloud retrieved by PSI (**left**) and TomoSAR (**right**) of Bellagio Hotel, Las Vegas [8].

2.3.3. Staring Spotlight TomoSAR

In spotlight modes, the radar beam is steered back and forth toward a common reference target in order to increase its illumination time t_{AP} (see Figure 4). The beam sweep rate controls the balance between the scene spatial extent and the azimuth resolution. In the TerraSAR-X sliding spotlight mode, the radar beam is swept at a moderate rate with a squint angle range up to ±0.75 degrees [49]. While in its staring spotlight mode, the beam sweep rate is set to equal the frequency modulation (FM) rate of the reference target. In other words, the radar beam is configured to exactly follow the target over time and the squint angle range can be up to ca. ±2.2 degrees. As a result, the azimuth resolution is maximized. Nevertheless, the improved azimuth resolution comes at the cost of reduced scene extent: the time span of a focused image Δt_{image} is considerably shorter. Needless to say, the slant range resolution stays unchanged for both modes, as long as the same range bandwidth is employed during imaging.
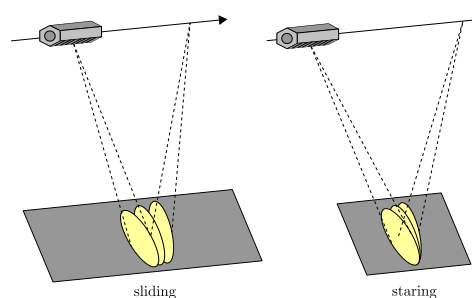
Figure 4. TerraSAR-X sliding (**left**) and staring (**right**) spotlight imaging geometry [48].

The transition from sliding to staring spotlight requires several adaptations in SAR focusing and InSAR processing. In the staring spotlight mode, the satellite can no longer be assumed to be standing still during chirp transmission and reception, or to follow a linear trajectory. In addition, variations of tropospheric and ionospheric delay within the large squint angle span also need to be corrected. Another major challenge is to estimate Doppler centroid frequency as a function of focused image time.

The TerraSAR-X multimode SAR processor [50,51] and the integrated wide area processor [24,25] were revised accordingly.

For PSI and TomoSAR in urban areas, the improved azimuth resolution has at least two advantages. PSs in the same resolution cell in the sliding spotlight mode may be resolved in different resolution cells in the staring spotlight mode. This leads to an increase in the density of the resulted 4D point cloud. Furthermore, the clutter in each resolution cell may be significantly suppressed, thanks to the increased azimuth resolution. Consequently, the SCR of PSs increases, which in turn leads to a better lower bound on the variance of elevation estimates [52].

In order to demonstrate these improvements, we processed two interferometric stacks of the City of Las Vegas in the sliding and staring spotlight modes using the SL1MMER algorithm. Each stack consists of 12 scenes acquired from October 2014 to February 2015. In each mode, 11 coregistered complex interferograms were used for the TomoSAR reconstruction of two regions of interests (ROIs).

One of the ROIs is a relatively flat region that was selected mainly for the assessment of relative vertical accuracy. The mean intensity maps in both modes are shown in Figure 5a. In the staring spotlight mode, point-like targets, such as the six bright points aligned at each side of the central field, are more focused. Even for DS, clutter appears to be more suppressed and thus the boundaries between areas of different smoothness are easier to recognize. This indicates an increased SCR. As a result, the reconstructed TomoSAR point cloud from staring spotlight images has a significant increase in the number of points compared to that of the sliding spotlight images. Indeed, the total number of scatterers in the staring spotlight is approximately 5.5 times as high, and the scatterer density is up to circa 13.5 million points per km^2 in this small region. In addition, the better SCR also improves the relative accuracy of height estimates. In fact, the relative accuracy of height estimates using staring spotlight images is approximately 1.7 times as high [48].
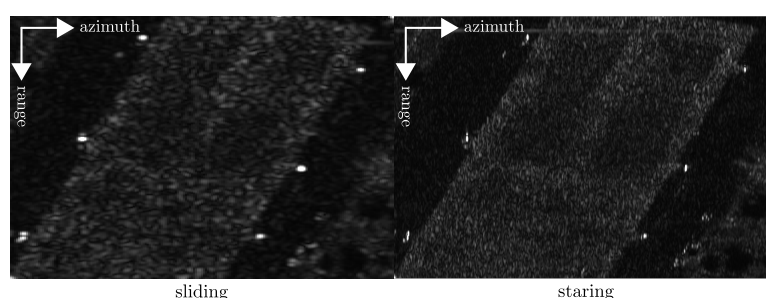
(**a**) ROI #1 that contains a flat area.
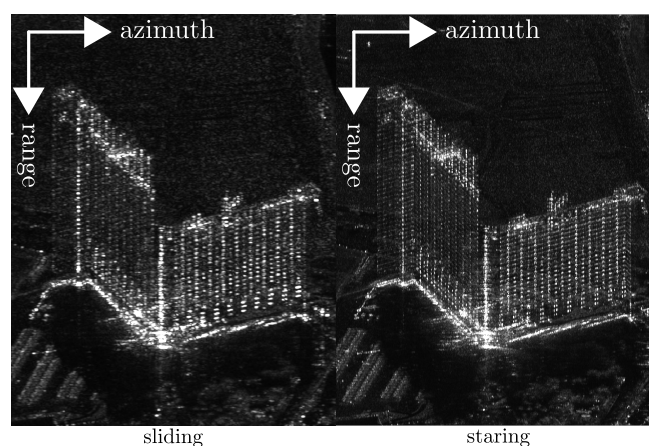
(**b**) ROI #2 that contains two high-rise buildings.

Figure 5. Mean intensity map of two ROIs in the sliding (**a**) and staring (**b**) spotlight modes [48].

Another ROI contains two high-rise buildings (Hilton Grand Vacations on the Las Vegas Strip), which were chosen as a demonstration of layover separation. The mean intensity maps are shown in Figure 5b. Similarly, point-like targets stand out more prominently from clutter in the staring spotlight mode and the regularities on the building facades are more clearly visible. The TomoSAR point clouds of single and double scatterers are shown in Figure 6. For this ROI, a substantial increase in the number of (single and double) scatterers was also observed. The scatterer density in the staring spotlight mode is approximately 5.1 times as high, see Table 2. The number of double scatterers in the staring spotlight mode almost rivals the number of single scatterers in the sliding spotlight mode.

single scatterers (sliding)　　　　　　　　　　single scatterers (staring)

double scatterers (sliding)　　　　　　　　　　double scatterers (staring)

600　620　640　660　680　700
Topography [m]
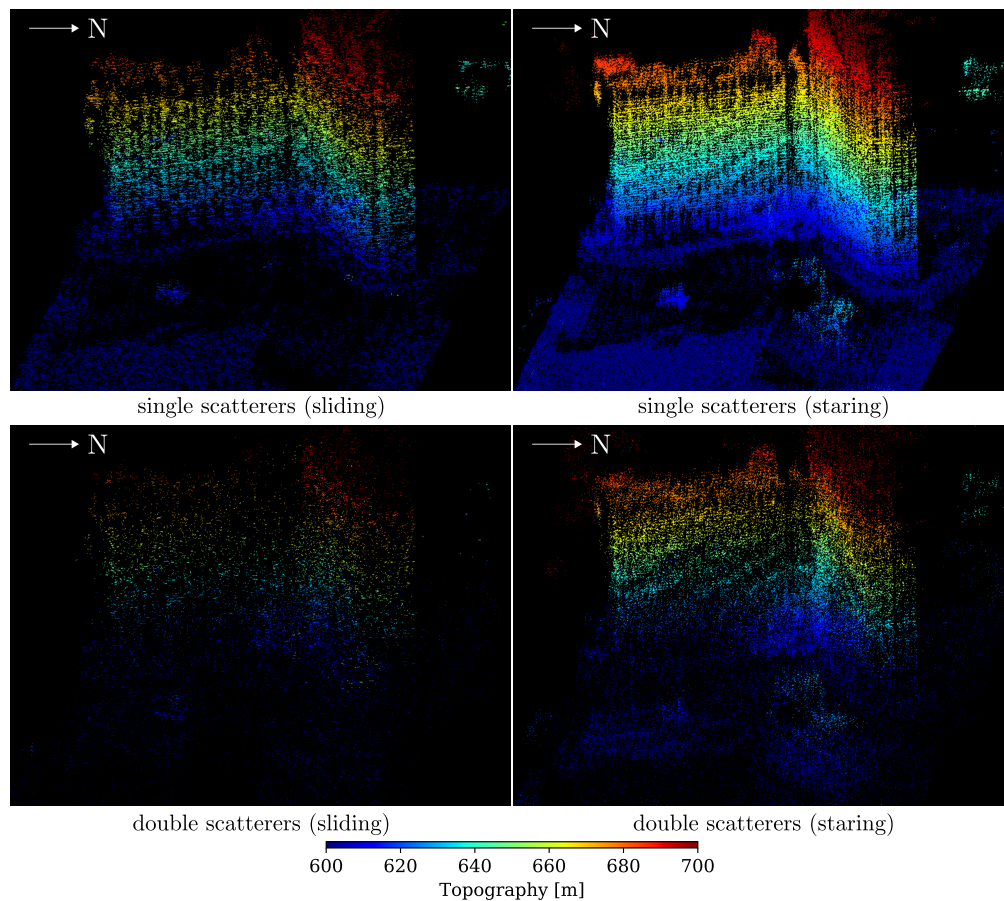
Figure 6. Updated topography (m) of the region in Figure 5b with 12 TerraSAR-X images in the sliding (left column) and staring (right column) spotlight modes, respectively. The upper and lower rows show single and double scatterers, respectively [48].

Table 2. Statistics of the point clouds in Figure 6.

	Sliding	Staring	Ratio [1]
No. of single scatterers	148, 646	740, 656	4.98
No. of double scatterers	21, 576	124, 546	5.77
Total no. of scatterers	170, 222	865, 202	5.08
Single-to-double-scatterer ratio	6.89	5.95	1.16
Scatterer density (million/km^2)	1.56	7.91	5.08

[1] The ratio was calculated by dividing the larger by the smaller value.

2.3.4. Point Cloud Fusion

Both PSI and D-TomoSAR deliver 4D point clouds relative to their reference points. They need to be co-registered when considering the results from multiple SAR image stacks. Although general

point cloud fusion is a classic topic in the computer vision field, there is very little literature addressing InSAR point cloud fusion, especially for point clouds from image stacks of cross-heading orbits. This is because the fusion of two point clouds requires the identification of common points in the two point clouds. There is theoretically no common point from such two point clouds due to the cross-heading geometry.

The first attempt to fuse cross-heading TerraSAR-X point clouds in an urban area was presented in [36]. This algorithm employs RANSAC to robustly match the ground points of two cross-heading TerraSAR-X PSI point clouds. The point correspondences are found by searching closely spaced point pairs on the ground surface. Therefore, this algorithm does not address the exact point correspondence. To find the exact point correspondence, Wang and Zhu detected the end positions of L-shaped facades in the two TomoSAR point clouds where the two point clouds converge [43]. In [6,53], the authors located dozens of pairs of street lampposts in two point clouds as point correspondences, additionally taking into account the diameter of the lampposts.

The fusion of along-heading (either both ascending or both descending) InSAR point clouds is less challenging. Classical point cloud co-registration methods such as iterative closest point (ICP) can be directly applied. Gernhardt et al. have demonstrated the direct application of ICP on multiple InSAR point clouds of a volcano [39].

2.3.5. 3D Motion Decomposition

A natural step after the fusion of multiple D-TomoSAR point clouds from different aspects is the decomposition of the 1D LOS displacement vector into its original 3D motion components. A 3D deformation vector in a geographic coordinate system is highly beneficial to improving the interpretation of the deformation pattern. A 3D motion decomposition algorithm was proposed and validated on four TomoSAR point clouds in [54]. The method relies on either geometrical [39,43] or geodetic fusion [6,55] of multi-aspect TomoSAR point clouds as input. It estimates the 3D motion components of the queried point target by inclusion of observations from all different viewing geometries and robust inversion with ℓ_1 norm minimization in a local neighborhood. The method allows for highly detailed and shadow-free 3D deformation monitoring, as has been demonstrated in [54]. An example of seasonal motion decomposition on a small test site in Berlin is demonstrated in Figure 7, where it shows the vertical (up), and horizontal (east-west) linear deformation of a railway bridge.
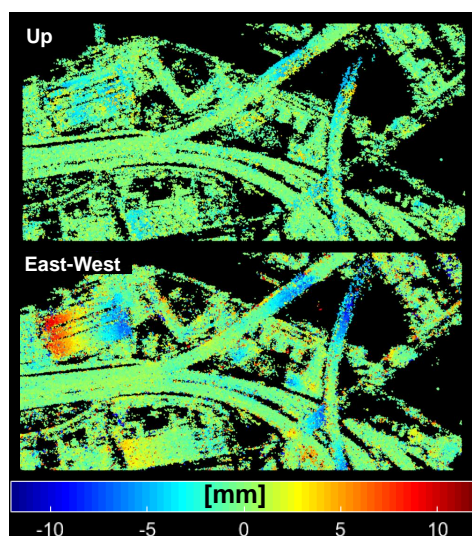

Figure 7. Decomposed seasonal deformation of a railway bridge located in the northeast of Berlin, Germany [54].

2.3.6. Object Reconstruction

Due the development described above, the quality of TomoSAR point cloud, including point density and relative accuracy, has become sufficient for the reconstruction of 3D models of individual objects. We have developed a suite of algorithms that have proved effective for tasks ranging from reconstructing vertical facade [44,45] (see Figure 8), to the detection and reconstruct of a LOD1 model of individual buildings [46,56].
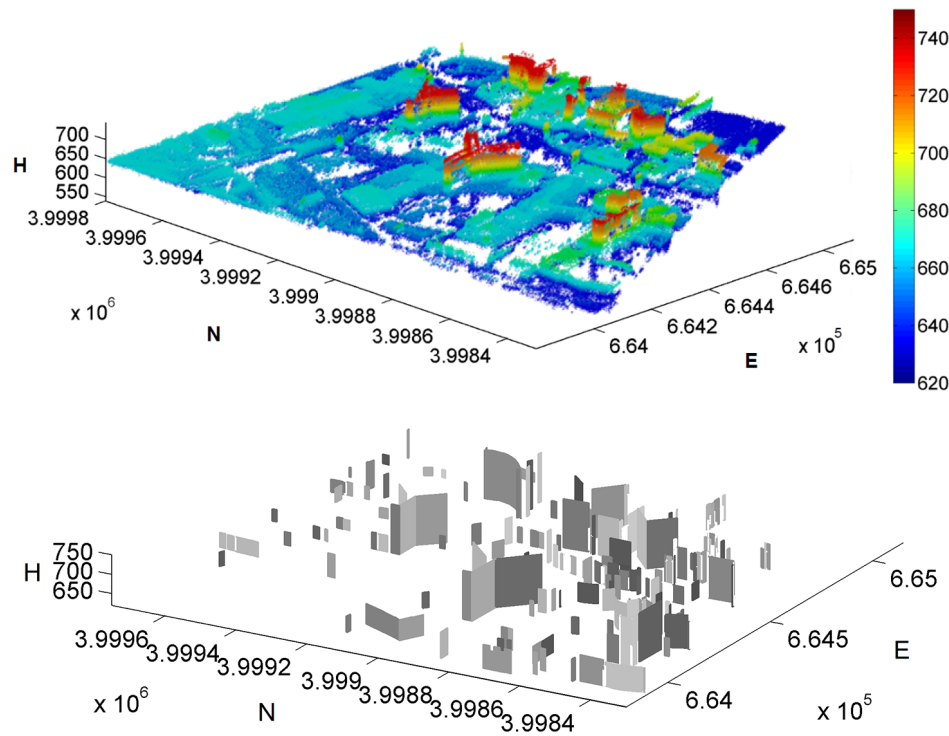

Figure 8. A TomoSAR point cloud of Las Vegas (upper), and the reconstructed facades (lower) [45]. The color of the point cloud represent its height above ellipsoid.

2.4. Object-Based InSAR Algorithms

The reconstruction of such high quality dense point clouds, as in aforementioned examples, are only possible with a stack of fairly high number of images. In practice, we are often faced with a limited number of images. In such situations, a proper algorithm should exploit information from neighboring pixels in order to reduce the number of images needed for a reliable reconstruction, such as adaptive filtering and nonlocal filtering that have been extensively described in previous literature, such as [4,57,58] and [59–61], respectively. However, this section goes beyond these pixel cluster-based methods. It focuses on the recent development of object-based algorithms that explicitly exploit geometric and semantic information to support parameter retrieval in multi-baseline InSAR. To this end, this section introduces the M-SL1MMER algorithm [28], which exploits the freely available building footprint from OpenStreetMap (OSM), and RoMIO (Robust Multi-pass InSAR via Object-based low rank decomposition) [62,63], which exploits the smoothness prior and low rank property of the InSAR data stack of individual objects.

2.4.1. M-SL1MMER

Multiple-snapshot SL1MMER (M-SL1MMER) is an extension of the original SL1MMER algorithm for joint tomographic reconstruction of resolution cells containing scatterers that share, up to quantization errors, the same height (hereafter referred to as "iso-height resolution cells") [28].

Similar approaches based on multiple snapshots or polarimetric channels can be found, for example, in [64–66].

In M-SL1MMER, the iso-height resolution cells are detected by projecting the freely available OSM building footprint [67] to the SAR image, and shifting it toward the near range direction until it reaches the top of the building. Each shifted position of the footprint represents a cluster of iso-height resolution cells. Let a specific iso-height cluster contain M resolution cells; the InSAR measurements $\mathbf{g}_m \in \mathbb{C}^N$ (N being the number of interferograms) of the m-th resolution cell can be approximated with the linear model (see Equation (3)) $\mathbf{g}_m \approx \mathbf{R}_m \boldsymbol{\gamma}_m$ for all $m = 1, \ldots, M$. In addition, we assume without loss of generality that $\mathbf{R}_1 \approx \cdots \approx \mathbf{R}_M$ and rewrite the M linear models in the more compact form $\mathbf{G} \approx \mathbf{R}\boldsymbol{\Gamma}$, where the m-th columns of $\mathbf{G}$ and $\boldsymbol{\Gamma}$ equal $\mathbf{g}_m$ and $\boldsymbol{\gamma}_m$, respectively. A key element of M-SL1MMER involves solving the following $\ell_{2,1}$ regularization problem:

$$\hat{\boldsymbol{\Gamma}} = \arg\min_{\boldsymbol{\Gamma}} \frac{1}{2}\|\mathbf{R}\boldsymbol{\Gamma} - \mathbf{G}\|_F^2 + \lambda_{2,1}\|\boldsymbol{\Gamma}\|_{2,1}, \tag{11}$$

where $\|\cdot\|_F$ and $\|\boldsymbol{\Gamma}\|_{2,1} = \sum_{i=1}^{L}\|\boldsymbol{\gamma}^i\|_2$ denote the Frobenius and $\ell_{2,1}$ norms, respectively, and $\boldsymbol{\gamma}^i$ is the i-th row of $\boldsymbol{\Gamma}$. The $\ell_{2,1}$ norm is known to promote the entries of $\boldsymbol{\Gamma}$ to be jointly sparse among columns. In other words, nonzero rows can be expected in $\hat{\boldsymbol{\Gamma}}$ or its submatrices. Solving the minimization problem in (11) is followed by model selection and amplitude debiasing independently for each resolution cell, as in the SL1MMER algorithm (see Section 2.3).

As a practical demonstration, we reconstructed the elevation of two high-rise buildings using 6 TanDEM-X bistatic sliding spotlight interferograms. The elevation estimates of the upper and lower layers are depicted in Figure 9. In the case of layover, the higher and lower scatterers can be found in the upper and lower layers, respectively. The smooth color transition on the reconstructed building facades already indicates its high quality. Roof-facade and facade-ground interactions are clearly visible in the near and far range, respectively. This can also be observed in the elevation difference map under layover (see Figure 10). The color change from deep blue (near range) to cyan (far range) corresponds to increasing elevation distance between building roof and facade.
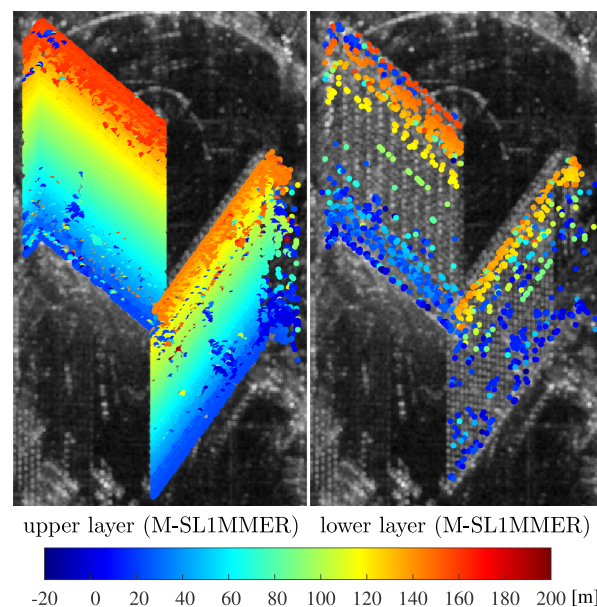

upper layer (M-SL1MMER)　　lower layer (M-SL1MMER)

-20　0　20　40　60　80　100　120　140　160　180　200 [m]

Figure 9. Elevation estimates of two test buildings with M-SL1MMER using 6 TanDEM-X bistatic sliding spotlight interferograms. In the case of layover, the higher and lower scatterers appear in the upper and lower layers, respectively [28].
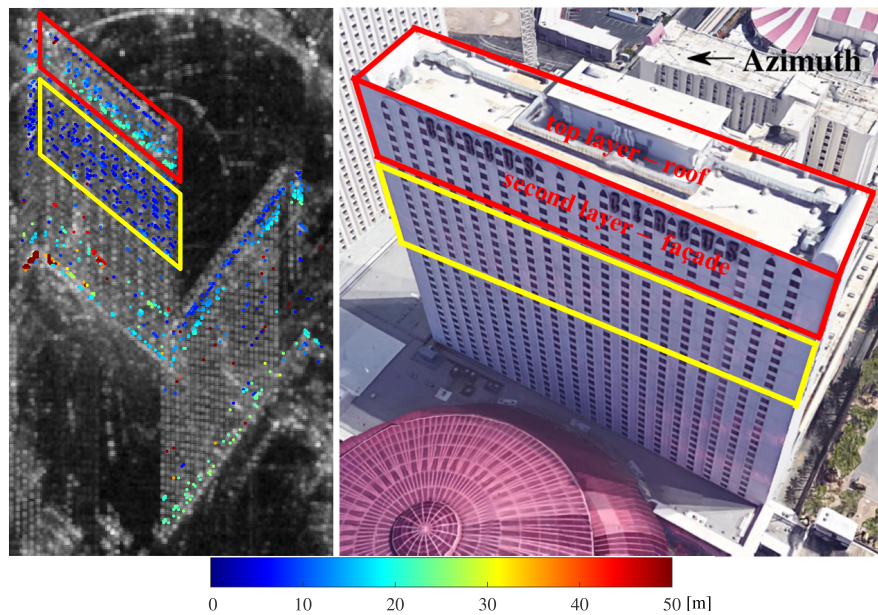
Figure 10. Difference of elevation estimates of higher and lower scatterers in Figure 9 subject to layover. The red and yellow rectangles mark areas where roof-facade and facade-facade interactions are expected, respectively [28].

2.4.2. RoMIO

As a complement to M-SL1MMER, RoMIO does not necessarily require explicit information of the footprints of the objects in the image. It is a more general framework that exploits the low rank property of InSAR phase tensor, because the low rankness of a tensor describes its information entropy, which requires looser signal support than the explicit iso-height line required in M-SL1MMER. RoMIO filters the InSAR data tensor by robustly minimizing its rank. Therefore, it can be regarded as a filtering step in prior to multi-baseline InSAR algorithms. The core RoMIO estimator can be expressed as follows.

$$\{\hat{\mathcal{X}}, \hat{\mathcal{E}}\} = \arg\min_{\mathcal{X}, \mathcal{E}} \operatorname{rank}(\mathcal{X}) + \lambda_{rank} \|\mathcal{E}\|_0, \quad s.t. \ \mathcal{X} + \mathcal{E} = \mathcal{G}, \tag{12}$$

where $\mathcal{G}$ is the observed InSAR phase tensor, $\mathcal{X}$ and $\mathcal{E}$ model the tensor of the true signal, and the sparse outliers, respectively, $\hat{\mathcal{X}}, \hat{\mathcal{E}}$ are the recovered outlier-free phase tensor and the estimated outlier tensor, respectively, $\operatorname{rank}(\mathcal{X})$ refers to the multilinear rank of $\mathcal{X}$, and λ_{rank} is the regularization parameter. In practice, the multilinear rank and the ℓ_0 norm are relaxed by the tensor nuclear norm $\|\mathcal{X}\|_*$ and ℓ_1 norm, respectively.

RoMIO reaches filtering performance comparable to state-of-the-art filtering algorithms, i.e., nonlocal means filtering [59,61]. However, it outperforms nonlocal means filtering by a factor of two in terms of the interferometric phase variance when the interferogram is corrupted by 50% outliers [63]. The merit of this extreme robustness in turn improves parameter estimation in multi-baseline InSAR algorithms. In typical settings of the TerraSAR-X high-resolution spotlight image stack, i.e., 10–20 images, SNR of 0–5 dB, a combination RoMIO and PSI outperforms the original PSI by a factor of 10 to 30 in the accuracy of the linear deformation estimates [63].

While optimizing the deformation parameters using multi-baseline InSAR algorithms, e.g., PSI, RoMIO can also make use of the explicit support of objects, such as a given segmentation mask of the SAR image. RoMIO includes a spatial regularization term, e.g., smoothness, of the 2D matrices of

the parameters in the estimator [62]. A general form of such regularized estimators can be expressed as follows.

$$\{\hat{\mathbf{S}}, \hat{\mathbf{P}}\} = \underset{\mathbf{S},\mathbf{P}}{\arg\min} \frac{1}{2} \|\mathcal{W} \odot (\mathcal{G} - \overline{\mathcal{G}}(\mathbf{S},\mathbf{P}))\|_F^2 + \lambda_{TV} f(\mathbf{S},\mathbf{P}), \tag{13}$$

where $\mathbf{S}$ and $\mathbf{P}$ are the matrices of the elevation and deformation parameters. Similar to other MAP estimators, e.g., Equation (8), the first term on the right-hand side of the estimator is a data fidelity term that calculates a weighted log likelihood between the observed InSAR phase tensor $\mathcal{G}$ and the modeled tensor $\overline{\mathcal{G}}$, where $\mathcal{W}$ denotes an optional weighting tensor, and $\odot$ denotes the element-wise product between two tensors. An example of the weighting tensor can be a tensor comprised of coherence matrices of each interferogram. Pixels of higher coherence are given higher weights. The function $f(\mathbf{S},\mathbf{P})$ denotes the regularization term that represents the spatial prior of $\mathbf{S}$ and $\mathbf{P}$. The regularization parameter $\lambda_T V$ controls the balance between these two terms. In [62], we made use of the popular total variation as a smoothness prior.

3. Advances in Robust Estimation

3.1. Overview of Advances

Robust estimation in multi-baseline InSAR was sporadically mentioned in previous literature. Some examples include using an adaptive window to improve the covariance matrix estimation [4,57,68], improving the PSI reference network by ℓ_1 norm minimization [23,24], and robust detection of multiple scatterers in TomoSAR [31,33]. However, it was not systematically addressed until [5]. Wang and Zhu pointed out that, due to the existence of non-Gaussian samples and unmodeled phase, e.g., the atmospheric phase, robust estimation in multi-baseline InSAR lies on the following two fundamental problems:

- covariance matrix estimation for DS, due to the existence of non-Gaussian and nonstationary samples
- phase history parameters estimation for both DS and PS, due to observations with large unmodeled phase

The impact of non-robust covariance estimation and the existence of nonstationary phase on parameter estimation in multi-baseline InSAR has been confirmed in several recent works, such as [69–72], and [58,73], respectively. The following sections will elaborate on these two points. The development of robust estimation is greatly associated with DS-based InSAR. Please refer to [74] for a recent review of DS-based InSAR techniques.

3.2. Robust Covariance Matrix Estimation

The estimation of the covariance matrix of a pixel is usually carried out by the sample covariance matrix. Its estimator is shown in Equation (14), where $\mathbf{g}$ is the multivariate observation, and $\mathbf{G}$ is the matrix consisting of M spatial samples, that is $\mathbf{G} = [\mathbf{g}_1, \mathbf{g}_2, ..., \mathbf{g}_M]$. Equation (14) is also the MLE if the samples are complex circular Gaussian (CCG) distributed. Unfortunately, this equation does not always hold in real data. This is why a robust estimator is necessary. A robust covariance estimator should consider the following two scenarios (and the mixture of both):

- the selected samples are non-Gaussian (possibly heavily tailed distribution)
- the expected interferometric phase of the samples is nonstationary, e.g., very strong underlying topographic phase

$$\hat{\mathbf{C}}_{\mathbf{gg}} = \frac{1}{M} \sum_{m=1}^{M} \mathbf{g}_m \mathbf{g}_m^H = \frac{1}{M} \mathbf{G}\mathbf{G}^H \tag{14}$$

The following content will summarize the robust covariance estimators, focusing on the points above.

3.2.1. Non-Gaussian Samples

For the first scenario, [5] proposed that the sample covariance matrix can be made more robust by an M-estimator, which is essentially an iterative reweighted sample covariance matrix [75,76]:

$$\hat{\mathbf{C}}_{k+1} = \frac{1}{M} \sum_{m=1}^{M} w\left(\mathbf{g}_m^H \hat{\mathbf{C}}_k^{-1} \mathbf{g}_m\right) \mathbf{g}_m \mathbf{g}_m^H, \tag{15}$$

where m and k are the sample index and the iteration index, respectively, and $w(x)$ is a weighting function of the negative log-likelihood of the sample $\mathbf{g}_m$ to the CCG probability density function (PDF). The weighting function down-weights highly deviated samples whose log-likelihood is small. Equation (15) is solved iteratively. The authors of [5] also proposed an approximation to drop the iterative process, which is the sign covariance matrix (SCM) [77,78]. Extending it to complex number, it is:

$$\hat{\mathbf{C}}_{SCM} = \frac{1}{M} \sum_{m=1}^{M} \|\mathbf{g}_m\|^{-2} \mathbf{g}_m \mathbf{g}_m^H \tag{16}$$

SCM is an engineering solution for fast processing under the general M-estimator's framework. The weighting function is replaced by the inverse of the ℓ_2 norm of the sample. Therefore, only the direction (or sign) of each multivariate sample is considered.

3.2.2. Non-Gaussian Samples with Nonstationary Interferometric Phase

It is often the case that the interferometric phase of the selected samples are not stationary, due to varying topography and motion or other factors. Usually, this type of deterministic phase is estimated and mitigated in prior to covariance estimation. For example, [58] proposed a multi-resolution defringe algorithm to mitigate such nonstationary phase.

Nevertheless, poor estimates significantly affect the covariance matrix estimation. Therefore, [5] proposed a new covariance estimator rank M-estimator (RME) for complex multivariate. The RME is derived by replacing the multivariate $\mathbf{g}$ with its rank $\mathbf{r}$ in Equation (15):

$$\hat{\mathbf{C}}_{RME,k+1} = \frac{1}{M} \sum_{m=1}^{M} w\left(x_m\left(\hat{\mathbf{C}}_{RME,k}\right)\right) \hat{\mathbf{r}}_m \cdot \hat{\mathbf{r}}_m^H \tag{17}$$

The complex rank vector $\mathbf{r}$, analogous to its real number version [78], is defined as follows:

$$\hat{\mathbf{r}}_m = \frac{1}{J} \sum_{j=1}^{J} \frac{\mathbf{g}_m \odot \mathbf{g}_j^*}{\left\| \mathbf{g}_m \odot \mathbf{g}_j^* \right\|}, \tag{18}$$

where $\mathbf{g}_j$ is a direct neighborhood sample of $\mathbf{g}_m$, and $\odot$ denotes the Hadamard product. The multiplication of the complex conjugate of a direct neighbor mitigates the nonstationary interferometric phase of $\mathbf{g}_m$. Due to the multiplication, the RME is a fourth-order descriptor of the sample statistics. An element-wise square root on $\left|\hat{\mathbf{C}}_{RME}\right|$ should be performed in order to obtain the second-order momentum. It was proven that the element-wise square root of $\left|\hat{\mathbf{C}}_{RME}\right|$ approaches $\left|\hat{\mathbf{C}}_{\mathbf{gg}}\right|$ asymptotically under CCG distribution when calculating the rank using one neighborhood sample [5].

3.2.3. Comparison

We compared the sample covariance matrix, M-Estimator, and the RME under three different scenarios: (1) multivariate CCG, (2) a heavily tailed multivariate distribution (complex t-distribution), and (3) nonstationary multivariate complex t-distribution. For each scenario, 1000 ten-acquisition vectors were simulated according to the distribution and a predefined coherence matrix that has a exponential decay of the coherence w.r.t. the temporal baseline. In the last scenario, linear phase fringes

with ten different fringe frequencies randomly picked within [0 $\pi/10$] were added to the phases of the ten acquisitions.

The results are shown in Figure 11, where each row corresponds to the three scenarios, respectively. The top left subplot can be regarded as the ground truth, because MLE is the optimal estimator under CCG, and is asymptotically unbiased. All three estimators can preserve the correct shape of the covariance matrix under CCG. The MLE fails in the second scenario, where the samples are contaminated by outliers. The coherence is usually overestimated because of the large amplitude of the outliers. In the last scenario, both MLE and M-estimator are not capable of dealing with nonstationary phases. Heavy underestimation occurs because of the summation of the complex numbers with non-constant phases. The estimates of M-estimator are extremely low due to more summation operations caused by the iterative process. Last but not least, RME is invariant to such nonstationary phase, and hence maintains good performance in all conditions.

A quantitative experiment shows that the robust estimator is extremely effective for samples with low coherence. At true coherence of 0.2, M-estimator outperforms the Gaussian MLE by a factor of 1.1 to 2.3, and a factor of 1 to 10, in terms of the accuracy and the bias, respectively, under a wide range of outlier percentages [5].
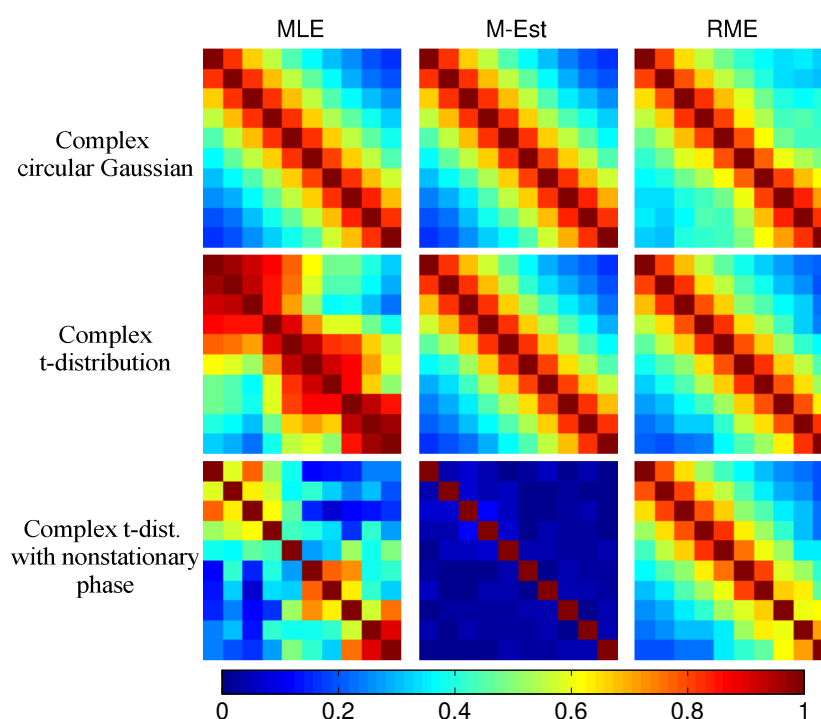
Figure 11. Comparison of three covariance matrix estimators under three different observation cases: first row: complex circular Gaussian, second row: complex t-distribution with one degree of freedom, and third row: nonstationary complex t-distribution with one degree of freedom. First column: MLE (under Gaussian), second column: M-estimator with t-distribution weighting, and third column: rank M-estimator with t-distribution weighting.

3.3. Robust Phase History Parameters Retrieval

A robust covariance matrix estimate alone is not sufficient for a robust estimation of the phase history parameters, i.e., elevation, and motion parameters, because a multi-pass InSAR observation $\mathbf{g} \in \mathbb{C}^N$ may contain an unmodeled phase, e.g., uncompensated atmospheric phase, unmodeled motion phase, etc. The following content provides examples of robust estimators for the retrieval of the phase history parameters of both PS and DS.

3.3.1. Robust PS Estimator

The general form of the MLE of PS phase history parameters can be expressed as follows:

$$\hat{\theta}_{MLE} = \arg\min_{\theta} \|\mathbf{g} - \bar{\mathbf{g}}(\theta)\|_2^2,\tag{19}$$

where $\bar{\mathbf{g}}(\theta)$ is the modeled PS signal. Equation (19) is shown to be equivalent to Equation (6) in [79]. Similar to the robust covariance estimator, it can be robustified by an M-estimator:

$$\hat{\theta}_{M-est} = \arg\min_{\theta} \sum_{i=1}^{N} \rho\left(\text{Re}\left[\varepsilon_i(\theta)\right]/\sigma_R\right) + \rho\left(\text{Im}\left[\varepsilon_i(\theta)\right]/\sigma_I\right),\tag{20}$$

where the residual $\varepsilon_i(\theta)$ equals $g_i - \bar{g}_i(\theta)$, $\text{Re}\left[\cdot\right]$, $\text{Im}\left[\cdot\right]$ are the real and imaginary parts of a complex number, and σ_R and σ_I are the standard deviations of the real and imaginary parts of the residual, respectively. The function $\rho(x)$ is the so-called robust loss function that can be derived from the PDF of the contaminated distribution of $\mathbf{g}$, if it is known. However, it is usually unknown in practice. We shall use stable empirical functions instead, e.g., the Tukey biweight function.

3.3.2. Robust DS Estimator

According to [80], the MLE of DS phase history parameters can be expressed as follows:

$$\hat{\theta}_{MLE} = \arg\min_{\theta} \left\{ \mathbf{g}^H \mathbf{\Phi}(\theta) \left|\mathbf{C}\right|^{-1} \mathbf{\Phi}(\theta)^H \mathbf{g} \right\}\tag{21}$$

where $\mathbf{\Phi}$ is a diagonal matrix of the modeled interferometric phase. If stationarity is assumed for a DS and its neighborhood, one can treat a cluster of DSs as a single PS by averaging them, as proposed in SqueeSAR [4]. Then, the robustified DS estimator is identical to Equation (20).

However, if the objective is a full inversion of individual single-look DS observation (without averaging) without the strict assumption of phase stationarity, the robustified estimator is shown in [5] to be in the following form:

$$\hat{\theta}_{M-est} = \arg\min_{\theta} \left\{ \varepsilon^H(\theta)\, \mathbf{W}(\bar{\varepsilon})\, \varepsilon(\theta) \right\},\tag{22}$$

where the residuals $\varepsilon(\theta)$ is shown in Equation (23). It is whitened by a robust covariance matrix estimate, e.g., $\hat{\mathbf{C}}_{RME}$. The matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$ is a diagonal robust weighting matrix computed from the mean residual $\bar{\varepsilon}$. Because of possible outliers in the residual, $\bar{\varepsilon}$ should also be robustly estimated, for example by a robust weighted averaging of $\varepsilon(\theta)$ of the selected samples.

$$\varepsilon(\theta) = \left|\hat{\mathbf{C}}_{RME}\right|^{-0.5} \mathbf{\Phi}(\theta)^H \mathbf{g}\tag{23}$$

To summarize, Equation (22) is a joint estimation of the phase parameters of individual single-look DS observations in a neighborhood. It is solved iteratively. Its computation should begin with initial estimates of each sample in the neighborhood (assumed to be the same), which jointly determine the initial weighting matrix. The same weighting matrix is used to retrieve the parameters of each single-look DS in the neighborhood, and is updated on the basis of all the estimates upon finishing one iteration.

To demonstrate the robustness of the estimator, Figure 12 shows the linear deformation rate of the volcano Stromboli, Italy, estimated by the robust DS phase history parameter retrieval method. Parameter estimation in active volcanic areas is challenging due to strong decorrelation, and the varying deformation model. In the experiment, only 16 interferograms acquired in 2008 were used. We can see that scatterers over 50% of the surface area were retrieved, although most of them did not undergo any significant deformation. The crater shows an uplift of 10 cm/year, and the southern slope

undergoes a subsidence of up to 20 cm/year. This may suggest certain displacement of the magma underneath the volcano.
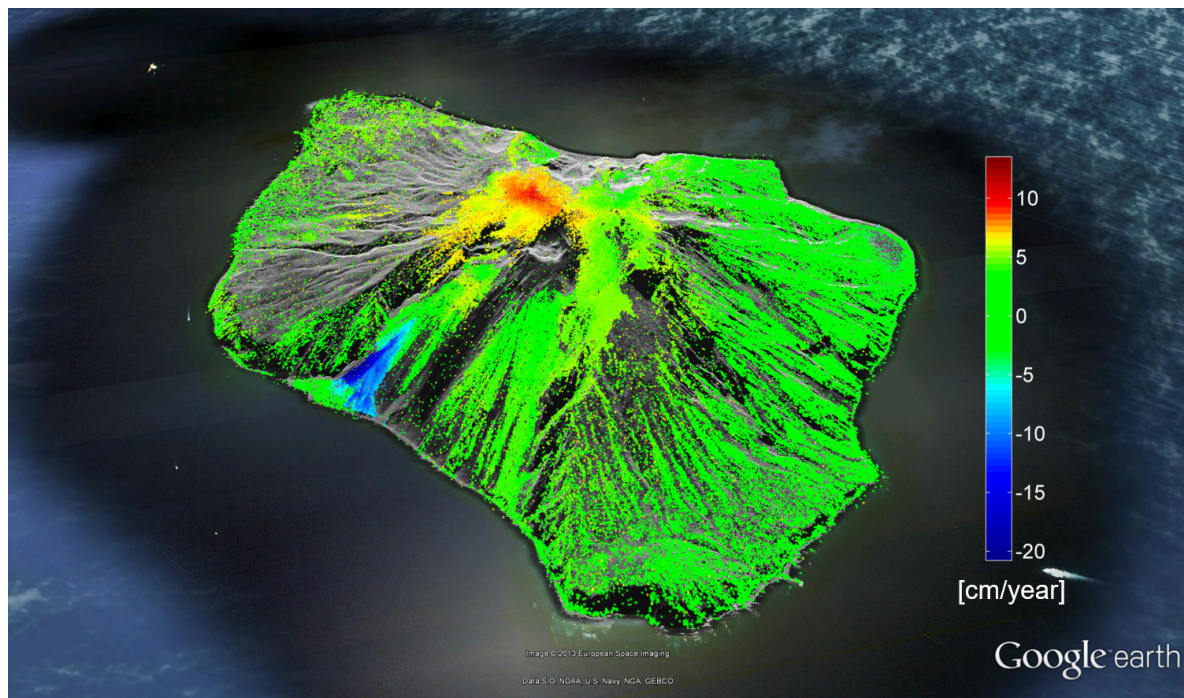

Figure 12. The linear deformation rate of the volcano Stromboli, in Italy, estimated by the robust DS phase history parameter retrieval method. In total, 16 interferograms acquired in 2008 were used. The crater shows an uplift of 10 cm/year, and the southern slope undergoes a subsidence of up to 20 cm/year. This may suggest certain displacement of the magma underneath the volcano. Courtesy: the tropospheric correction was done by Cong et al.

4. Advances in Absolute Positioning

A unique feature of TerraSAR-X is its precise orbit determination and high precision range measurements, which allows for an unprecedented 2D localization accuracy of image pixels below one meter. In recent years, this level of accuracy has been further improved by thorough consideration of the most prominent error factors affecting range and azimuth measurements of SAR, a method termed SAR imaging geodesy [7,81]. SAR imaging geodesy is seen as a great leap in SAR technology, because it extends the applications of SAR to the geodetic positioning domain rather than the imaging domain. Two of the numerous application examples of SAR imaging geodesy are geodetic stereo SAR [82], a method that retrieves the precise 3D absolute position of a target by combining its 2D radar timings from different orbit tracks, and a framework called geodetic InSAR [6], in which multi-baseline InSAR and stereo SAR are combined to achieve absolute 4D InSAR point clouds. A brief introduction to the two methods is given below, and the most recent advances of these techniques and their new applications are described.

The SAR imaging geodesy method aims at attaining 2D absolute pixel localization [7]. A single pixel in a focused complex SAR image is localized, in across-track, by range τ_{rg} and, in along-track, by azimuth t_{az} times. For a point target inside the mentioned pixel, the following equations read:

$$\tau_{rg} = \frac{2R}{c} + \delta\tau_{SD} + \delta\tau_O + \delta\tau_F + \delta\tau_I + \delta\tau_T + \delta\tau_G$$
$$t_{az} = t + \delta t_{SD} + \delta t_O + \delta t_F + \delta t_G, \tag{24}$$

where R is the geometric distance from the sensor to the center of the pixel in meters and c is the speed of light in vacuum; the other terms are all expressed in seconds. The raw acquisition time is denoted by t and the timing error terms subscripted by SD, O, F, I, T and G represent delays caused by satellite dynamics, orbit inaccuracies, feature localization error, ionospheric delay, tropospheric delay, and geodynamic effects, respectively. The magnitude of individual errors range from a couple of centimeters for the ionospheric effect, if the satellite operates in X-band, followed by decimeter regimes for satellite dynamic effects and geodynamic effects for both components, to up to a few meters for the tropospheric effect, depending on the weather conditions and the average incidence angle of the acquired SAR images. Some of the mentioned errors and their effects on SAR measurements are shown in Figure 13. The curved propagation path shown in Figure 13 is highly exaggerated for visualization purposes only. In order to remove the mentioned timing errors, the imaging geodesy method exploits the highly precise orbit data of TerraSAR-X and Tandem-X [1,2,83], utilizes a highly sophisticated SAR processor to avoid unnecessary approximations [84], precisely extracts targets with sub-pixel sensitivity [85,86], and corrects the path delay and geodynamic errors by global numerical weather data [81,87] and state-of-the-art geodetic models [88].
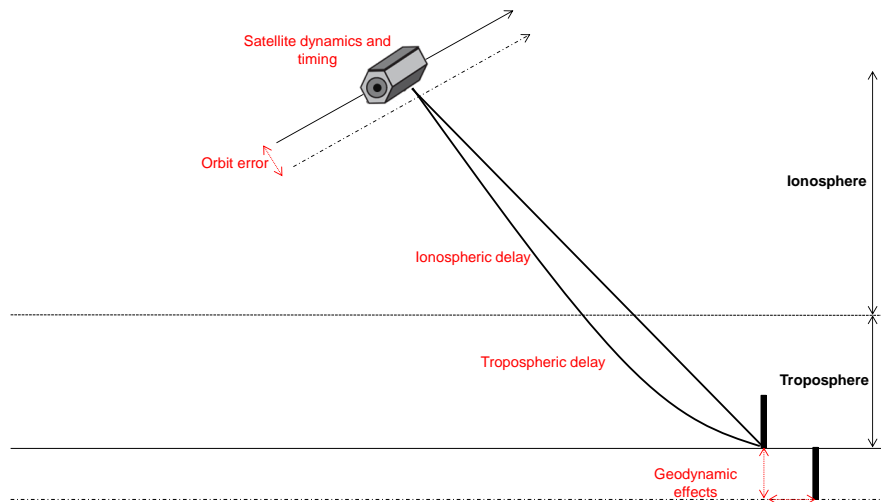
Figure 13. The errors affecting range and azimuth timings of SAR measurements, colorized in red. Orbit errors cause the satellite trajectory to deviate from the true track, while satellite dynamics and atmospheric disturbances cause delays in the timings, which lead to incorrect annotation of τ_{rg} and t_{az}. Geodynamic effects change the position of a target on the ground, which again hampers the accuracy of the timings. Please note that the atmospheric effect shown in the figure is highly exaggerated for visualization purposes only. The main cause for atmospheric delay is the decrease of the speed of light.

By combining the τ_{rg} and t_{az} of the same target visible in SAR images acquired from two or more different viewing geometries, the stereo SAR method determines the 3D position of the target (see Figure 14). The 2D radar timing coordinates of a particular target in the SAR image $x_T = (t_{az}, \tau_{rg})$ are linked to their corresponding 3D coordinates on the surface of the Earth $\mathbf{X}_T = (X, Y, Z)$ by the range-Doppler equation system [85]:

$$
\begin{aligned}
|\mathbf{X}_S - \mathbf{X}_T| - c \cdot \tau_{rg} &= 0 \\
\frac{\dot{\mathbf{X}}_S(\mathbf{X}_S - \mathbf{X}_T)}{|\dot{\mathbf{X}}_S||\mathbf{X}_S - \mathbf{X}_T|} &= 0
\end{aligned}
\tag{25}
$$

with $\mathbf{X}_S$ and $\dot{\mathbf{X}}_S$ denoting the position and the velocity vector of the satellite relative to t_{az}, and τ_{rg} being the calibrated two-way traveled time from the satellite to the target. The variable t_{az} is implicitly included in the second equation relating the state-vector of the satellite to the time of the acquisition using a polynomial model [82]. The estimation of the coordinates is performed by least squares adjustment plus stochastic modeling of timing observations using the variance component estimation (VCE) [82]. The relative accuracy of the estimated coordinates depends on the SCR of the target, the precision of the external atmospheric and geodynamic corrections, the degree of difference in the combined viewing geometries, and the number of SAR acquisitions.

(**a**) Cross-heading   (**b**) Same-heading
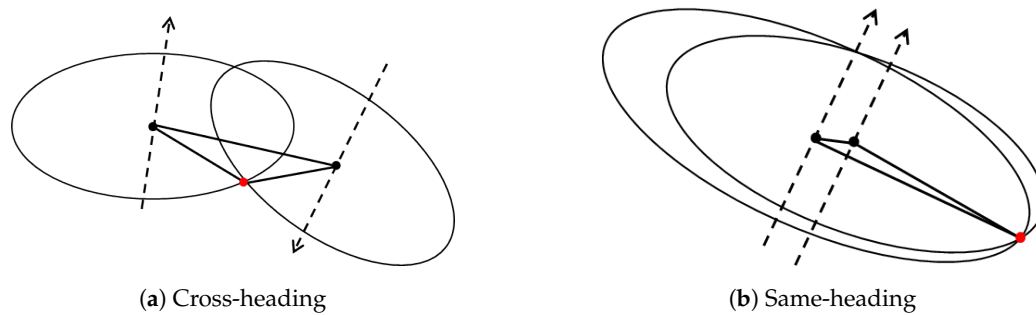
Figure 14. Localization of a point target (red dot) from (**a**) cross-heading and (**b**) same-heading satellite tracks. The satellites are shown by black dots; their trajectories are presented by dashed lines and the baselines are depicted by solid lines between the satellite positions. The black circles are defined by the range-Doppler equations and their intersection leads to the 3D position of the target [89].

4.1. Overview of Advances

SAR imaging geodesy was first named in 2011 by Eineder et al., since the method incorporates correction principles used in geodesy with SAR [7]. Schubert et al. reported on the correction of atmospheric delays by local height dependent models in [90]. Gisinger studied the effect of utilizing different mapping functions for converting the zenith atmospheric delays into the radar line of sight in [91]. These methods used the local GNSS zenith path delays for atmospheric corrections. Cong et al. introduced atmospheric correction through the 3D integration of weather data obtained from the European Center for Medium-Range Weather Forecasts (ECMWF) and using global TEC maps [81]. Apart from atmospheric errors, calibration of internal electronic delays of the SAR sensor was investigated in [7] and the precision of azimuth timing was improved by calibrating the sensor's internal clock rate [92]. The most prominent geodynamic effects, such as solid earth tides, pole tides, and continental drifts, were included in further studies [7,81,86,93]. In order to improve the localization precision into sub-centimeter regimes, Balss et al. further modeled geodynamic effects with smaller magnitudes, such as atmospheric pressure loading, ocean tidal loading, ocean pole tides, and atmospheric tidal loading [94]. In all the studies, the geodynamic effects were considered by the state-of-the-art models of the IERS 2010 convention [88]. The already precise orbit of TerraSAR-X [83] has been further improved by modeling the non-gravitational forces and also solar radiation pressure modeling [1]. The world-wide reproducibility of high precision measurements was demonstrated in [95] and an operational processor called the SAR Geodesy Processor (SGP) was introduced in [87]. Relative to applications, the high precision ranging measurement of TerraSAR-X has been exploited for maritime purposes [96,97]. In terms of achievable accuracy, SAR imaging geodesy is capable of localizing corner reflectors with 1.16 cm and 1.85 cm range and azimuth standard deviations, respectively [98].

The first results on 3D localization of CRs by means of stereo SAR was reported in [99]. Although 3D positioning using multi-aspect TerraSAR-X images had been previously demonstrated in [100–102], the results in [99] were unique in the sense that the stereo processing was carried out on thoroughly

calibrated range and azimuth timings. Gisinger et al. demonstrated the applicability of the geodetic stereo SAR method not only on CRs but on opportunistic non-ideal scatterers such as PS in an urban area [82]. The manually extracted scatterers could be localized with 3D precision better than 10 cm [82], which paved the way for new geodetic applications such as secular ground movement estimation using natural PS [103,104], high precision mapping of road networks (DriveMark) [105], and highly precise automatic SAR Ground Control Point (GCP) generation [89,106–109]. In terms of achievable accuracy, geodetic stereo SAR is able to localize corner reflectors with 3D precision better than 4 cm and an absolute accuracy of 2–3 cm when compared to independently surveyed reference positions [82].

4.2. Geodetic InSAR

The geodetic InSAR approach integrates the capabilities of multi-baseline InSAR with SAR imaging geodesy and stereo SAR techniques. The goal of the framework is to tackle the shortcomings of both methods: the relative estimates of all InSAR approaches and the small number of points that can be absolutely localized by geodetic stereo SAR. Therefore, it tends to achieve absolute positioning of a large number of scatterers by exploiting the advantages of both techniques. In the following, the workflow of the geodetic InSAR technique is described and some example applications are demonstrated.

4.2.1. SAR GCP Generation

The first major part of the procedure is concerned with extraction of GCPs from multi-aspect SAR images. This includes [89]:

Step 1   Detection and matching of identical PS from SAR images acquired from different orbits. In the reference geodetic SAR tomography technique this task was performed manually [6]. At the current state of the framework, the identification of common PS can be carried out using the PSI multi-track fusion algorithm [39] for same-heading tracks and utilizing high resolution optical data [106] or external geospatial road network data [109] for cross-heading tracks. A combination of all the mentioned methods for automatic detection of large number of GCPs was used in [89].

Step 2   Precise timing extraction of PS from stacks of non-coregistered SLC images. This is done by PTA [85,86].

Step 3   PS visibility check and initial outlier removal. The time series of phase noise approximated by SCR of each PS [19] is analyzed and the outliers are robustly removed by the adjusted box plot method [110].

Step 4   Correction of PS timings in the stack of images using imaging geodesy.

Step 5   Absolute 3D positioning of each PS by the stereo SAR method [82]. The posterior quality measures of the observations and the estimates are also reported in this step.

4.2.2. Absolute Localization of InSAR Point Clouds

The main objective of the geodetic InSAR framework is to resolve the DEM error of the reference point with respect to which the topography and deformation parameters are estimated [20,21,40]. The geodetic InSAR approach can overcome this problem, to some extent, in two ways dependent on the number of available GCPs. If only a small number of GCPs are available, the best candidate will be chosen as the reference point during PSI/TomoSAR processing and at the final stage the geocoded coordinates of all points in the point clouds are shifted toward the absolute coordinates of this point [6]. If a large number of GCPs are available, for instance using the GCP generation approaches in [89,107], the DEM error of the reference point is approximated as a post-processing step. Therefore, the difference in ellipsoidal heights of GCPs and their corresponding geocoded PS heights are calculated and a height offset is robustly estimated. The height offset is added to the geocoded PS

heights and an updated geocoding is carried out which results in absolute coordinates of the InSAR point cloud [111].

4.2.3. Applications

To conclude this subsection, a few examples and applications of the geodetic InSAR framework are demonstrated below.

Figure 15 shows the city of Oulu in Finland overlaid by 2049 GCPs obtained from four stacks of TerraSAR-X high resolution spotlight images.
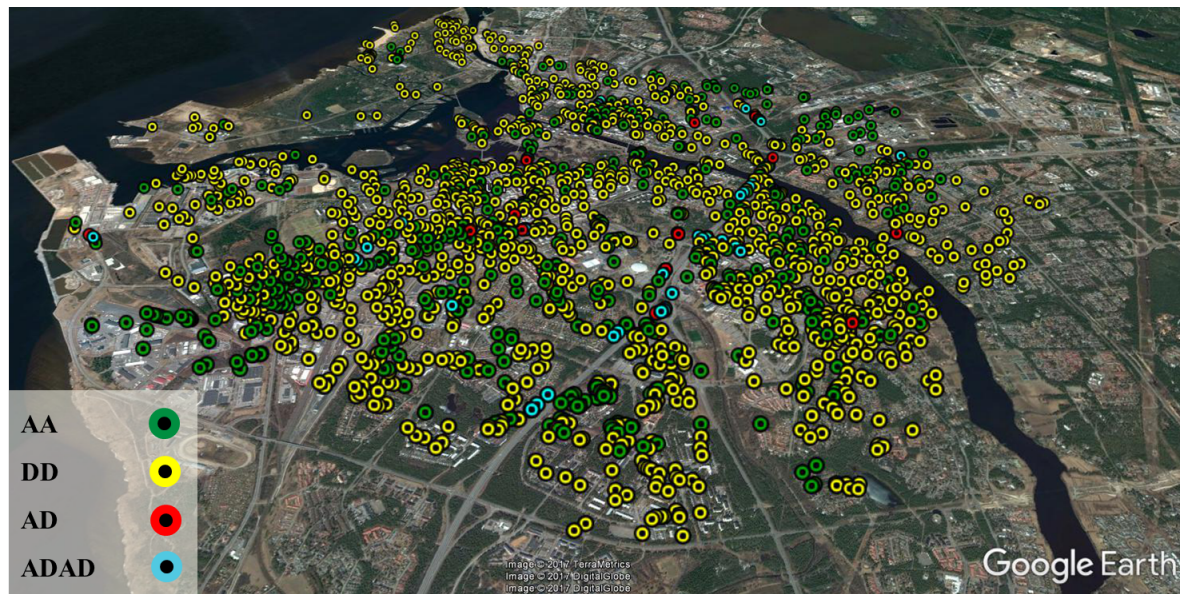

Figure 15. Total number of 2049 GCPs in Oulu, color-coded based on the geometry configuration used for their positioning (AA: ascending-ascending, DD: descending-descending, AD: ascending-descending and ADAD: quad geometry) [89]. The underlying optical image is taken from Google Earth.

The GCPs are color-coded based on the underlying geometry configuration used for their localization, where AA, DD, and AD stand for ascending-ascending, descending-descending and ascending-descending orbits, respectively; ADAD means that scatterers were localized from all the four viewing geometries. It is observed that the entire central area of Oulu is covered with the generated GCPs. The candidates from the same-heading geometries stem from built areas, while the ones from cross-heading orbits include the bases of lamp poles, street lights, and traffic lights. The statistics of the generated GCPs are reported in Table 3, which demonstrate the extremely high potential of TerraSAR-X for precise 3D positioning.

Comparison with a reference LiDAR point cloud shows that we can achieve a horizontal absolute accuracy of 20 cm using just a single GCP to correct the geocoding of an InSAR point cloud [6,55]. Therefore, employing over one thousand GCPs, as shown previously, can achieve extremely high absolute accuracy, presumably in the order of centimeter. In order to demonstrate this, a close comparison of two cross-heading InSAR point clouds before and after height correction is shown in Figure 16, where the red and green points represent the PS of descending and ascending tracks, respectively. It can be seen that after the calibration of the height of the reference point using the GCPs, the endpoint of building facades correctly match.

Table 3. Averaged statistics based on the stereo SAR least squares estimated 3D coordinate standard deviations in Oulu. The letters A and D stand for ascending and descending geometries, respectively. The sample mean and standard deviation are denoted by μ and σ and $S_{[ENH]}$ represents the local coordinates standard deviations within a 95% confidence level [89].

Geometry	Number of Scatterers	μ_{S_E} (cm)	μ_{S_N} (cm)	μ_{S_H} (cm)	σ_{S_E} (cm)	σ_{S_N} (cm)	σ_{S_H} (cm)
AA	565	17.73	5.04	15.87	11.98	2.63	11.09
DD	1417	15.08	3.80	16.71	10.38	2.10	11.30
AD	24	2.26	2.50	1.75	0.99	1.11	0.83
ADAD	43	1.17	1.40	1.12	0.42	0.55	0.37

Figure 16. Demonstration of absolute localization of PSI point clouds obtained from an ascending and a descending orbit track of Oulu. The endpoints of buildings visible from each geometry match correctly with the endpoints from the opposing geometry.

To give an impression of the fused TomoSAR point cloud of a large area, Figure 17 shows a result obtained by fusing four TomoSAR point clouds of Berlin obtained from two pairs of cross-heading high resolution TerraSAR-X spotlight images that are fused by selecting an identical GCP as the reference point of all point clouds. The point cloud has in total 63 million scatterers in an area of 50 km^2. Such shadow-free highly detailed TomoSAR point clouds can be further utilized to reconstruct dynamic 3D and 4D city models [44–46,112].
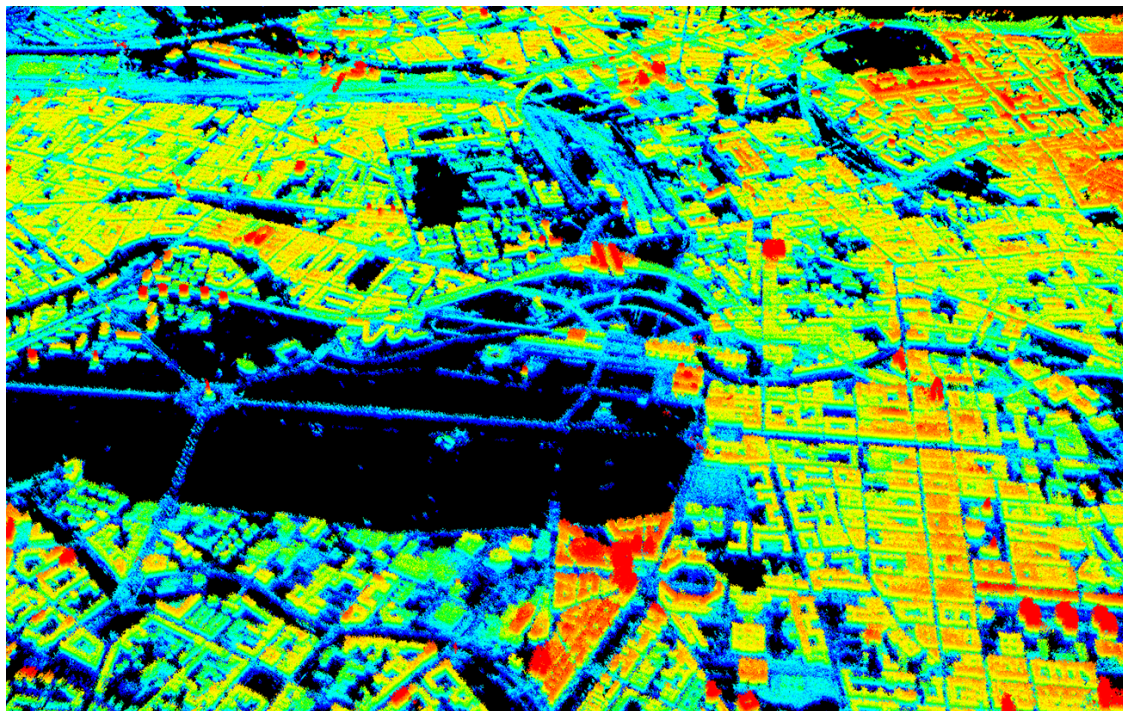
Figure 17. 3D view of central Berlin after geodetic registration of four TomoSAR point clouds obtained from a pair of cross-heading high resolution TerraSAR-X spotlight data. The height is color-coded and ranges between 70 m and 110 m [6].

5. Conclusions and Outlook

This paper provides a review of the multi-baseline InSAR techniques in the scope of TerraSAR-X data. It covers the evolution of multi-baseline InSAR techniques, particularly with respect to improving the relative estimation accuracy, introducing robustness to the estimators, and achieving accurate absolute positioning of scatterers, which includes bridging the absolutely located scatterers with the relative measures obtained from multi-baseline techniques. Particular focus was placed on our own development work, specifically SL1MMER, M-SL1MMER, Tomo-GENESIS (TomoSAR), RIO (robust estimation), RoMIO (object-based InSAR), and geodetic InSAR (absolute positioning).

Looking into the future, the next generation spaceborne SAR missions, including high resolution wide swath (HRWS) and Tandem-L, will simultaneously possess high resolution and global coverage, which would enable novel applications such as monitoring global changes. Retrieving geo-parameters from these data will require not only new technological approaches to manage large amounts of data, but also new analysis methods. In the following, we would like to point out some promising future directions:

- Big data management technologies: So far, besides big missions, such as global TanDEM-X DEM generation, scientists are dealing with SAR data in the order of up to terabytes. However, this is about to change. Already today, petabytes of Sentinel-1 data are openly accessible to the public. Yet, only very limited groups are capable of national-scale InSAR data processing, to say nothing about global. To be prepared for the future, novel big geo-data management technologies are of high relevance.

- Fast and accurate parameter inversion algorithms: The development of inversion algorithms should keep up with the pace of data growth. For example, as a pre-study of Tandem-L, sequential interferometric phase estimators are proposed instead of full covariance matrix inversion to tackle the challenge of big InSAR data [72]. Fast solvers are demanded for many advanced parameter inversion models that often involve non-convex, nonlinear, and complex-valued

optimization problem, such as CS-based tomographic inversion, or low rank complex tensor decomposition. Besides aforementioned model-based inversion methods, recently, data-driven machine learning/deep learning methods have boosted the baseline performances in many remote sensing problems [113], mostly in classification and detection tasks, yet its potential in InSAR processing or more generally in geoparamater estimation is not yet exploited at all. This deserves more attention of the community.

- Complicated motion: Up to now, only limited motion models, such as linear, seasonal or a combination of several basic models, are considered for deformation estimations of InSAR. There are also studies using model order selection to detect different types of motion either being embedded in the estimation [114] or considered as a post-processing [115]. However, the actual motion can be far more complex than any model can describe. The weekly repeat cycle and long-term monitoring capability of future sensors will enable retrieving much more complex motion models, and even allow performing classification of different types of motions and detecting anomalies. This calls for more sophisticated algorithms.

- Data assimilation: At present, the interferometric stack is usually a static cube of interferograms. As Sentinel-1 provides global coverage every six days, new stacking and multi-pass InSAR concepts should be able to include new images without excessive computational burden. This requires development of the data assimilation strategy, as well as novel inversion algorithms that only require the new measurements and the previous estimates for updating the parameters of interest.

- Multi-sensor data fusion: In the Copernicus era, it is standard that more than one data source, such as SAR and optical, is available at any test site. Intelligent use of the complementary peculiarities of the ever-increasing number of diverse remote sensing sensors and other geo-data sources has become the natural choice for many applications [116]. Some preliminary work in the community demonstrated that introducing the geometric prior or semantic prior to InSAR or TomoSAR reconstruction could significantly reduced the number of required SAR data while retaining the estimation accuracy [28,63]. This is definitely a promising future direction.

Abbreviations

The following abbreviations are used in this manuscript:

ADI	Amplitude dispersion index
CCG	Complex circular Gaussian
CS	Compressive sensing
DEM	Digital elevation model

DS   Distributed scatterer
D-TomoSAR   Differential TomoSAR
ECMWF   European Center for Medium-Range Weather Forecasts
GCP   Ground control point
HRWS   High resolution wide swath
ICP   Iterative closest point
IERS   International Earth Rotation and Reference Systems Service
InSAR   SAR interferometry
LOS   Line of sight
MAP   Maximum a posteriori
MLE   Maximum likelihood estimator
M-SL1MMER   Multiple-snapshot SL1MMER
OSM   OpenStreetMap
PDF   Probability density function
PSI   Persistent scatter interferometry
PS   Point/Persistent Scatterer
RIO   Robust InSAR optimization
RME   Rank M-estimator
ROI   Region of interest
ROMIO   Robust multi-pass InSAR via object-based low rank decomposition
SAR   Synthetic aperture radar
SCR   Signal-to-clutter ratio
SL1MMER   Scale-down by L1 norm minimization, model selection, and estimation reconstruction
SNR   Signal-to-noise ratio
TEC   Total electron content
TomoSAR   SAR tomography
VHR   Very high resolution

References

1. Hackel, S.; Montenbruck, O.; Steigenberger, P.; Balss, U.; Gisinger, C.; Eineder, M. Model improvements and validation of TerraSAR-X precise orbit determination. *J. Geodesy* **2017**, *91*, 547–562. [CrossRef]

2. Hackel, S.; Gisinger, C.; Balss, U.; Wermuth, M.; Montenbruck, O. Long-Term Validation of TerraSAR-X Orbit Solutions with Laser and Radar Measurements. *Remote Sens.* **2018**, *10*, 762. [CrossRef]

3. Zhu, X.X. Very High Resolution Tomographic SAR Inversion for Urban Infrastructure Monitoring—A Sparse and Nonlinear Tour. Ph.D. Thesis, Technische Universität München, München, Germany, 2011.

4. Ferretti, A.; Fumagalli, A.; Novali, F.; Prati, C.; Rocca, F.; Rucci, A. A New Algorithm for Processing Interferometric Data-Stacks: SqueeSAR. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3460–3470. [CrossRef]

5. Wang, Y.; Zhu, X.X. Robust Estimators for Multipass SAR Interferometry. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 968–980. [CrossRef]

6. Zhu, X.X.; Montazeri, S.; Gisinger, C.; Hanssen, R.F.; Bamler, R. Geodetic SAR Tomography. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 18–35. [CrossRef]

7. Eineder, M.; Minet, C.; Steigenberger, P.; Cong, X.Y.; Fritz, T. Imaging Geodesy—Toward Centimeter-Level Ranging Accuracy with TerraSAR-X. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 661–671. [CrossRef]

8. Zhu, X.X.; Bamler, R. Superresolving SAR Tomography for Multidimensional Imaging of Urban Areas: Compressive sensing-based TomoSAR inversion. *IEEE Signal Process. Mag.* **2014**, *31*, 51–58. [CrossRef]

9. Bamler, R.; Hartl, P. Synthetic aperture radar interferometry. *Inverse Probl.* **1998**, *14*, R1. [CrossRef]

10. Reigber, A.; Moreira, A. First demonstration of airborne SAR tomography using multibaseline L-band data. *IEEE Trans. Geosci. Remote Sens.* **2000**, *38*, 2142–2152. [CrossRef]

11. Lombardini, F.; Montanari, M.; Gini, F. Reflectivity estimation for multibaseline interferometric radar imaging of layover extended sources. *IEEE Trans. Signal Process.* **2003**, *51*, 1508–1519. [CrossRef]

12. Fornaro, G.; Serafino, F.; Soldovieri, F. Three-dimensional focusing with multipass SAR data. *IEEE Trans. Geosci. Remote Sens.* **2003**, *41*, 507–517. [CrossRef]

13. Zhu, X. Spectral Estimation for Synthetic Aperture Radar Tomography. Master's Thesis, Lehrstuhl für Methodik der Fernerkundung, TU München, Munich, Germany, 2008.

14. Fornaro, G.; Reale, D.; Serafino, F. Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 224–237. [CrossRef]

15. Zhu, X.X.; Bamler, R. Very high resolution spaceborne SAR tomography in urban environment. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 4296–4308. [CrossRef]

16. Zhu, X.X.; Bamler, R. Let's do the time warp: Multicomponent nonlinear motion estimation in differential SAR tomography. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 735–739. [CrossRef]

17. Crosetto, M.; Monserrat, O.; Cuevas-González, M.; Devanthéry, N.; Crippa, B. Persistent Scatterer Interferometry: A review. *ISPRS J. Photogramm. Remote Sens.* **2016**, *115*, 78–89. [CrossRef]

18. Ferretti, A.; Prati, C.; Rocca, F. Permanent scatterers in SAR interferometry. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 8–20. [CrossRef]

19. Adam, N.; Kampes, B.M.; Eineder, M. The development of a scientific persistent scatterer system: Modifications for mixed ERS/ENVISAT time series. In Proceedings of the Envisat and ERS Symposium, Salzburg, Austria, 6–10 September 2004; pp. 1–9.

20. Kampes, B.M. *Radar Interferometry: Persistent Scatterer Technique*; Remote Sensing and Digital Image Processing; Springer: Dordrecht, The Netherlands, 2006.

21. Van Leijen, F. Persistent Scatterer Interferometry Based on Geodetic Estimation Theory. Ph.D. Thesis, Delft University of Technology, Delft, The Netherlands, 2014.

22. Wang, Y.; Zhu, X.X.; Bamler, R. An Efficient Tomographic Inversion Approach for Urban Mapping Using Meter Resolution SAR Image Stacks. *IEEE Geosci. Remote Sens. Lett.* **2014**, *11*, 1250–1254. [CrossRef]

23. Rodriguez Gonzalez, F.; Bhutani, A.; Adam, N. L1 network inversion for robust outlier rejection in persistent Scatterer Interferometry. In Proceedings of the 2011 IEEE International Geoscience and Remote Sensing Symposium, Vancouver, BC, Canada, 24–29 July 2011; pp. 75–78. [CrossRef]

24. Adam, N.; Gonzalez, F.R.; Parizzi, A.; Brcic, R. Wide area Persistent Scatterer Interferometry: Current developments, algorithms and examples. In Proceedings of the 2013 IEEE International Geoscience and Remote Sensing Symposium, Melbourne, VIC, Australia, 21–26 July 2013; pp. 1857–1860. [CrossRef]

25. Rodriguez Gonzalez, F.; Adam, N.; Parizzi, A.; Brcic, R. The Integrated Wide Area Processor (IWAP): A Processor For Wide Area Persistent Scatterer Interferometry. In Proceedings of the ESA Living Planet Symposium 2013, Edinburgh, UK, 9–13 September 2013; pp. 1–4.

26. Zhu, X.; Bamler, R. Very high Resolution SAR tomography via Compressive Sensing. In Proceedings of the Fringe 2009 Workshop, Frascati, Italy, 30 November–4 December 2009; pp. 1–7.

27. Zhu, X.; Bamler, R. Tomographic SAR Inversion by L1-Norm Regularization—The Compressive Sensing Approach. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 3839–3846. [CrossRef]

28. Zhu, X.X.; Ge, N.; Shahzad, M. Joint sparsity in SAR tomography for urban mapping. *IEEE J. Sel. Top. Signal Process.* **2015**, *9*, 1498–1509. [CrossRef]

29. Pauciullo, A.; Reale, D.; De Maio, A.; Fornaro, G. Detection of Double Scatterers in SAR Tomography. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3567–3586. [CrossRef]

30. Zhu, X.X.; Bamler, R. Demonstration of Super-Resolution for Tomographic SAR Imaging in Urban Environment. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 3150–3157. [CrossRef]

31. Zhu, X.; Bamler, R. Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation With Application to Spaceborne Tomographic SAR. *IEEE Trans. Geosci. Remote Sens.* **2012**, *50*, 247–258. [CrossRef]

32. Budillon, A.; Johnsy, A.C.; Schirinzi, G. A Fast Support Detector for Superresolution Localization of Multiple Scatterers in SAR Tomography. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2017**, *10*, 1–12. [CrossRef]

33. Ma, P.; Lin, H. Robust Detection of Single and Double Persistent Scatterers in Urban Built Environments. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 2124–2139. [CrossRef]

34. Siddique, M.A.; Wegmüller, U.; Hajnsek, I.; Frey, O. Single-Look SAR Tomography as an Add-On to PSI for Improved Deformation Analysis in Urban Areas. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6119–6137. [CrossRef]

35. Shi, Y.; Zhu, X.X.; Yin, W.; Bamler, R. A fast and accurate basis pursuit denoising algorithm with application to super-resolving tomographic SAR. *IEEE Trans. Geosci. Remote Sens.* **2018**, in press. [CrossRef]

36. Gernhardt, S.; Bamler, R. Deformation monitoring of single buildings using meter-resolution SAR data in PSI. *ISPRS J. Photogramm. Remote Sens.* **2012**, *73*, 68–79. [CrossRef]

37. Gernhardt, S.; Adam, N.; Hinz, S.; Bamler, R. Appearance of Persistent Scatterers for Different TerraSAR-X Acquisition Modes. In Proceedings of the ISPRS Workshop, Hannover, Germany, 2–5 June 2009; Volume XXXVIII-1-4-7/W5, pp. 1–5.

38. Gernhardt, S.; Adam, N.; Eineder, M.; Bamler, R. Potential of very high resolution SAR for persistent scatterer interferometry in urban areas. *Ann. GIS* **2010**, *16*, 103–111. [CrossRef]

39. Gernhardt, S.; Cong, X.Y.; Eineder, M.; Hinz, S.; Bamler, R. Geometrical Fusion of Multitrack PS Point Clouds. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 38–42. [CrossRef]

40. Gernhardt, S. High Precision 3D Localization and Motion Analysis of Persistent Scatterers using Meter-Resolution Radar Satellite Data. Ph.D. Thesis, Technische Universität München, München, Germany, 2012.

41. Kalia, A.C. Value-added Products in the Framework of the German Ground Motion Service. In Proceedings of the FRINGE 2017: 10th International Workshop on Advances in the Science and Applications of SAR Interferometry and Sentinel-1 InSAR, Helsinki, Finland, 5–9 June 2017.

42. Zhu, X.X.; Wang, Y.; Gernhardt, S.; Bamler, R. Tomo-GENESIS: DLR's tomographic SAR processing system. In Proceedings of the Joint Urban Remote Sensing (JURSE), Sao Paulo, Brazil, 2013; pp. 159–162. [CrossRef]

43. Wang, Y.; Zhu, X.X. Automatic Feature-Based Geometric Fusion of Multiview TomoSAR Point Clouds in Urban Area. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2015**, *8*, 953–965. [CrossRef]

44. Zhu, X.X.; Shahzad, M. Facade Reconstruction Using Multiview Spaceborne TomoSAR Point Clouds. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 3541–3552. [CrossRef]

45. Shahzad, M.; Zhu, X. Robust Reconstruction of Building Facades for Large Areas Using Spaceborne TomoSAR Point Clouds. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 752–769. [CrossRef]

46. Shahzad, M.; Zhu, X.X. Automatic Detection and Reconstruction of 2-D/3-D Building Shapes from Spaceborne TomoSAR Point Clouds. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1292–1310. [CrossRef]

47. Wang, Y.; Zhu, X.X.; Bamler, R.; Gernhardt, S. Towards TerraSAR-X Street View: Creating City Point Cloud from Multi-Aspect Data Stacks. In Proceedings of the 2013 Joint Urban Remote Sensing Event (JURSE), Sao Paulo, Brazil, 21–23 April 2013; pp. 198–201. [CrossRef]

48. Ge, N.; Rodriguez Gonzalez, F.; Wang, Y.; Zhu, X.X. Spaceborne Staring Spotlight SAR Tomography— A First Demonstration with TerraSAR-X. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, in press.

49. Eineder, M.; Adam, N.; Bamler, R.; Yague-Martinez, N.; Breit, H. Spaceborne spotlight SAR interferometry with TerraSAR-X. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1524–1535. [CrossRef]

50. Breit, H.; Fritz, T.; Balss, U.; Lachaise, M.; Niedermeier, A.; Vonavka, M. TerraSAR-X SAR processing and products. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 727–740. [CrossRef]

51. Duque, S.; Breit, H.; Balss, U.; Parizzi, A. Absolute Height Estimation Using a Single TerraSAR-X Staring Spotlight Acquisition. *IEEE Geosci. Remote Sens. Lett.* **2015**, *12*, 1735–1739. [CrossRef]

52. Bamler, R.; Eineder, M.; Adam, N.; Zhu, X.; Gernhardt, S. Interferometric Potential of High Resolution Spaceborne SAR. *Photogramm. Fernerkund. Geoinf.* **2009**, *2009*, 407–419. [CrossRef]

53. Zhu, X.; Montazeri, S.; Gisinger, C.; Hanssen, R.; Bamler, R. Geodetic TomoSAR—Fusion of SAR Imaging Geodesy and TomoSAR for 3D Absolute Scatterer Positioning. In Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Quebec City, QC, Canada, 13–18 July 2014.

54. Montazeri, S.; Zhu, X.X.; Eineder, M.; Bamler, R. Three-Dimensional Deformation Monitoring of Urban Infrastructure by Tomographic SAR Using Multitrack TerraSAR-X Data Stacks. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 6868–6878. [CrossRef]

55. Montazeri, S. The Fusion of SAR Tomography and Stereo-SAR for 3D Absolute Scatterer Positioning. Master's Thesis, Delft University of Technology, Delft, The Netherlands, 2014.

56. Sun, Y.; Shahzad, M.; Zhu, X. First Prismatic Building Model Reconstruction from TomoSAR Point Clouds. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2016**, *XLI-B3*, 381–386. [CrossRef]

57. Parizzi, A.; Brcic, R. Adaptive InSAR Stack Multilooking Exploiting Amplitude Statistics: A Comparison Between Different Techniques and Practical Results. *IEEE Geosci. Remote Sens. Lett.* **2011**, *8*, 441–445. [CrossRef]

58. Wang, Y.; Zhu, X.; Bamler, R. Retrieval of Phase History Parameters from Distributed Scatterers in Urban Areas Using Very High Resolution SAR Data. *ISPRS J. Photogramm. Remote Sens.* **2012**, *73*, 89–99. [CrossRef]

59. Deledalle, C.A.; Denis, L.; Tupin, F. NL-InSar: Nonlocal interferogram estimation. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1441–1452. [CrossRef]

60. Zhu, X.X.; Bamler, R.; Lachaise, M.; Adam, F.; Shi, Y.; Eineder, M. Improving TanDEM-X DEMs by Non-local InSAR Filtering. In Proceedings of the 10th European Conference on Synthetic Aperture Radar, Berlin, Germany, 2–6 June 2014; pp. 1–4.

61. Baier, G.; Rossi, C.; Lachaise, M.; Zhu, X.X.; Bamler, R. Nonlocal InSAR filtering for high resolution DEM generation from TanDEM-X interferograms. *IEEE Trans. Geosci. Remote Sens.* **2018**, in press. [CrossRef]

62. Kang, J.; Wang, Y.; Körner, M.; Zhu, X. Robust Object-based Multi-pass InSAR Deformation Reconstruction. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 4239–4251. [CrossRef]

63. Kang, J.; Wang, Y.; Schmitt, M.; Zhu, X.X. Object-based Multipass InSAR via Robust Low Rank Tensor Decomposition. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 3062–3077. [CrossRef]

64. Aguilera, E.; Nannini, M.; Reigber, A. Multisignal compressed sensing for polarimetric SAR tomography. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 871–875. [CrossRef]

65. Schmitt, M.; Stilla, U. Compressive sensing based layover separation in airborne single-pass multi-baseline InSAR data. *IEEE Geosci. Remote Sens. Lett.* **2013**, *10*, 313–317. [CrossRef]

66. Fornaro, G.; Verde, S.; Reale, D.; Pauciullo, A. CAESAR: An approach based on covariance matrix decomposition to improve multibaseline–multitemporal interferometric SAR processing. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 2050–2065. [CrossRef]

67. OpenStreetMap Contributors. Planet Dump. 2017. Available online: https://www.openstreetmap.org (accessed on 16 April 2018).

68. Schmitt, M.; Schonberger, J.L.; Stilla, U. Adaptive Covariance Matrix Estimation for Multi-Baseline InSAR Data Stacks. *IEEE Trans. Geosci. Remote Sens.* **2014**, *52*, 1–11. [CrossRef]

69. Samiei-Esfahany, S.; Martins, J.E.; van Leijen, F.; Hanssen, R.F. Phase Estimation for Distributed Scatterers in InSAR Stacks Using Integer Least Squares Estimation. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 5671–5687. [CrossRef]

70. Cao, N.; Lee, H.; Jung, H.C. A Phase-Decomposition-Based PSInSAR Processing Method. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 1074–1090. [CrossRef]

71. Even, M. A Study on Spatio-Temporal Filtering in the Spirit of SqueeSAR. In Proceeding of the 11th European Conference on Synthetic Aperture Radar (EUSAR), Hamburg, Germany, 6–9 June 2016; pp. 462–465.

72. Ansari, H.; De Zan, F.; Bamler, R. Sequential Estimator: Toward Efficient InSAR Time Series Analysis. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 5637–5652. [CrossRef]

73. Verde, S.; Reale, D.; Pauciullo, A.; Fornaro, G. Improved Small Baseline processing by means of CAESAR eigen-interferograms decomposition. *ISPRS J. Photogramm. Remote Sens.* **2018**, *139*, 1–13. [CrossRef]

74. Even, M.; Schulz, K. InSAR Deformation Analysis with Distributed Scatterers: A Review Complemented by New Advances. *Remote Sens.* **2018**, *10*, 744. [CrossRef]

75. Ollila, E.; Koivunen, V. Influence functions for array covariance matrix estimators. In Proceedings of the 2003 IEEE Workshop on Statistical Signal Processing, St. Louis, MO, USA, 28 September–1 October 2003; pp. 462–465.

76. Zoubir, A.M.; Koivunen, V.; Chakhchoukh, Y.; Muma, M. Robust Estimation in Signal Processing: A Tutorial-Style Treatment of Fundamental Concepts. *IEEE Signal Process. Mag.* **2012**, *29*, 61–80. [CrossRef]

77. Visuri, S.; Koivunen, V.; Oja, H. Sign and rank covariance matrices. *J. Stat. Plan. Inference* **2000**, *91*, 557–575. [CrossRef]

78. Croux, C.; Ollila, E.; Oja, H. Sign and rank covariance matrices: Statistical properties and application to principal components analysis. In *Statistical Data Analysis Based on the L1-Norm and Related Methods*; Springer: Basel, Switzerland, 2002; pp. 257–269.

79. Rife, D.; Boorstyn, R.R. Single tone parameter estimation from discrete-time observations. *IEEE Trans. Inf. Theory* **1974**, *20*, 591–598. [CrossRef]

80. De Zan, F. Optimizing SAR Interferometry for Decorrelating Scatterers. Ph.D. Thesis, Politecnico di Milano, Milan, Italy, 2008.

81. Cong, X.Y.; Balss, U.; Eineder, M.; Fritz, T. Imaging Geodesy—Centimeter-Level Ranging Accuracy with TerraSAR-X: An Update. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 948–952. [CrossRef]

82. Gisinger, C.; Balss, U.; Pail, R.; Zhu, X.X.; Montazeri, S.; Gernhardt, S.; Eineder, M. Precise Three-Dimensional Stereo Localization of Corner Reflectors and Persistent Scatterers with TerraSAR-X. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 1782–1802. [CrossRef]

83. Yoon, Y.; Eineder, M.; Yague-Martinez, N.; Montenbruck, O. TerraSAR-X Precise Trajectory Estimation and Quality Assessment. *IEEE Trans. Geosci. Remote Sens.* **2009**, *47*, 1859–1868. [CrossRef]

84. Breit, H.; Börner, E.; Mittermayer, J.; Holzner, J.; Eineder, M. The TerraSAR-X Multi-Mode SAR Processor—Algorithms and Design. In Proceedings of the 5th European Conference on Synthetic Aperture Radar, Ulm, Germany, 25–27 May 2004.

85. Cumming, I.G.; Wong, F.H.C. *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*; Artech House Remote Sensing Library: Boston, MA, USA, 2005.

86. Balss, U.; Cong, X.Y.; Brcic, R.; Rexer, M.; Minet, C.; Breit, H.; Eineder, M.; Fritz, T. High precision measurement on the absolute localization accuracy of TerraSAR-X. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012; pp. 1625–1628. [CrossRef]

87. Eineder, M.; Balss, U.; Suchandt, S.; Gisinger, C.; Cong, X.; Runge, H. A definition of next-generation SAR products for geodetic applications. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 1638–1641. [CrossRef]

88. Petit, G.; Luzum, B. *IERS Conventions*; IERS Technical Note No. 36; Verlag des Bundesamts für Kartographie und Geodäsie: Frankfurt am Main, Germany, 2010.

89. Montazeri, S.; Gisinger, C.; Eineder, M.; Zhu, X.X. Automatic Detection and Positioning of Ground Control Points Using TerraSAR-X Multiaspect Acquisitions. *IEEE Trans. Geosci. Remote Sens.* **2018**, *56*, 2613–2632. [CrossRef]

90. Schubert, A.; Jehle, M.; Small, D.; Meier, E. Influence of Atmospheric Path Delay on the Absolute Geolocation Accuracy of TerraSAR-X High-Resolution Products. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 751–758. [CrossRef]

91. Gisinger, C. Atmospheric Corrections for TerraSAR-X Derived from GNSS Observations. Master's Thesis, Technische Universität München, München, Germany, 2012.

92. Balss, U.; Breit, H.; Fritz, T.; Steinbrecher, U.; Gisinger, C.; Eineder, M. Analysis of internal timings and clock rates of TerraSAR-X. In Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Quebec City, QC, Canada, 13–18 July 2014; pp. 2671–2674, doi:10.1109/IGARSS.2014.6947024. [CrossRef]

93. Schubert, A.; Jehle, M.; Small, D.; Meier, E. Mitigation of atmospheric perturbations and solid Earth movements in a TerraSAR-X time-series. *J. Geodesy* **2012**, *86*, 257–270. [CrossRef]

94. Balss, U.; Gisinger, C.; Cong, X.; Eineder, M.; Fritz, T.; Breit, H.; Brcic, R. GNSS-Based Signal Path Delay and Geodynamic Corrections for Centimeter Level Pixel Localization with TerraSAR-X. In Proceedings of the 5th TerraSAR-X Science Team Meeting, Oberpfaffenhofen, Germany, 10–14 June 2013; pp. 1–4.

95. Balss, U.; Gisinger, C.; Cong, X.Y.; Brcic, R.; Hackel, S.; Eineder, M. Precise measurements on the absolute localization accuracy of TerraSAR-X on the base of far-distributed test sites. In Proceeding of the EUSAR 2014: 10th European Conference on Synthetic Aperture Radar, Berlin, Germany, 2–4 June 2014.

96. Eineder, M.; Balss, U.; Biarge, S.D. Water level measurement by controlled radar reflection and TerraSAR-X imaging geodesy. In Proceedings of the 2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Quebec City, QC, Canada, 13–18 July 2014; pp. 5141–5143. [CrossRef]

97. Duque, S.; Balss, U.; Cong, X.Y.; Yague-Martinez, N.; Fritz, T. Absolute Ranging for Maritime Applications using TerraSAR-X and TanDEM-X Data. In Proceedings of the EUSAR 2014: 10th European Conference on Synthetic Aperture Radar, Berlin, Germany, 2–4 June 2014; pp. 1–4.

98. Eineder, M.; Gisinger, C.; Balss, U.; Cong, X.Y.; Montazeri, S.; Hackel, S.; Rodriguez Gonzalez, F.; Runge, H. SAR Imaging Geodesy—Recent Results for TerraSAR-X and for Sentinel-1. In Proceedings of the FRINGE 2017: 10th International Workshop on Advances in the Science and Applications of SAR Interferometry and Sentinel-1 InSAR, Helsinki, Finland, 5–9 June 2017.

99. Balss, U.; Gisinger, C.; Cong, X.; Eineder, M.; Brcic, R. Precise 2-D and 3-D Ground Target Localization with TerraSAR-X. *ISPRS Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2013**, *XL-1/W*, 23–28. [CrossRef]

100. Eldhuset, K.; Weydahl, D.J. Geolocation and Stereo Height Estimation Using TerraSAR-X Spotlight Image Data. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 3574–3581. [CrossRef]

101. Raggam, H.; Perko, R.; Gutjahr, K.; Kiefl, N.; Koppe, W.; Hennig, S. Accuracy Assessment of 3D Point Retrieval from TerraSAR-X Data Sets. In Proceedings of the 8th European Conference on Synthetic Aperture Radar, Aachen, Germany, 7–10 June 2010; pp. 1–4.

102. Koppe, W.; Wenzel, R.; Hennig, S.; Janoth, J.; Hummel, P.; Raggam, H. Quality assessment of TerraSAR-X derived ground control points. In Proceedings of the 2012 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Munich, Germany, 22–27 July 2012; pp. 3580–3583. [CrossRef]

103. Gisinger, C.; Gernhardt, S.; Auer, S.; Balss, U.; Hackel, S.; Pail, R.; Eineder, M. Absolute 4-D positioning of persistent scatterers with TerraSAR-X by applying geodetic stereo SAR. In Proceedings of the 2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Milan, Italy, 26–31 July 2015; pp. 2991–2994. [CrossRef]

104. Gisinger, C.; Montazeri, S.; Balss, U.; Cong, X.Y.; Hackel, S.; Zhu, X.X.; Pail, R.; Eineder, M. Applying geodetic SAR with TerraSAR-X and TanDEM-X. In Proceedings of the TerraSAR-X/TanDEM-X Science Team Meeting, Oberpfaffenhofen, Germany, 17–20 October 2016.

105. Runge, H.; Balss, U.; Suchandt, S.; Klarner, R.; Cong, X. DriveMark – Generation of High Resolution Road Maps with Radar Satellites. In Proceedings of the 11th ITS European Congress, Glasgow, Scotland, 6–9 June 2016; pp. 1–6.

106. Montazeri, S.; Zhu, X.X.; Balss, U.; Gisinger, C.; Wang, Y.; Eineder, M.; Bamler, R. SAR ground control point identification with the aid of high resolution optical data. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 3205–3208. [CrossRef]

107. Balss, U.; Runge, H.; Suchandt, S.; Cong, X.Y. Automated extraction of 3-D Ground Control Points from SAR images—An upcoming novel data product. In Proceedings of the 2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China, 10–15 July 2016; pp. 5023–5026. [CrossRef]

108. Nitti, D.O.; Morea, A.; Nutricato, R.; Chiaradia, M.T.; La Mantia, C.; Agrimano, L.; Samarelli, S. Automatic GCP extraction with high resolution COSMO-SkyMed products. In Proceedings of the SPIE Remote Sensing 2016, Edinburgh, UK, 26–29 September 2016; Volume 10003. [CrossRef]

109. Montazeri, S.; Gisinger, C.; Zhu, X.X.; Eineder, M.; Bamler, R. Automatic positioning of SAR ground control points from multi-aspect TerraSAR-X acquisitions. In Proceedings of the 2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), Fort Worth, TX, USA, 23–28 July 2017; pp. 961–964. [CrossRef]

110. Hubert, M.; Vandervieren, E. An adjusted boxplot for skewed distributions. *Comput. Stat. Data Anal.* **2008**, *52*, 5186–5201. [CrossRef]

111. Montazeri, S.; Zhu, X.X.; Gisinger, C.; Rodriguez Gonzalez, F.; Eineder, M.; Bamler, R. Towards Absolute Positioning of InSAR Point Clouds. In Proceedings of the FRINGE 2017: 10th International Workshop on Advances in the Science and Applications of SAR Interferometry and Sentinel-1 InSAR, Helsinki, Finland, 5–9 June 2017.

112. Wang, Y.; Zhu, X.X.; Zeisl, B.; Pollefeys, M. Fusing Meter-Resolution 4-D InSAR Point Clouds and Optical Images for Semantic Urban Infrastructure Monitoring. *IEEE Trans. Geosci. Remote Sens.* **2017**, *55*, 14–26. [CrossRef]

113. Zhu, X.X.; Tuia, D.; Mou, L.; Xia, G.S.; Zhang, L.; Xu, F.; Fraundorfer, F. Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geosci. Remote Sens. Mag.* **2017**, *5*, 8–36. [CrossRef]

114. Ge, N.; Zhu, X. Bistatic-like Differential SAR Tomography—A Preliminary Framework for Tandem-L. *IEEE Trans. Geosci. Remote Sens.* **2018**, submitted.

115. Chang, L.; Hanssen, R.F. A Probabilistic Approach for InSAR Time-Series Postprocessing. *IEEE Trans. Geosci. Remote Sens.* **2016**, *54*, 421–430. [CrossRef]

116. Schmitt, M.; Zhu, X.X. Data Fusion and Remote Sensing: An ever-growing relationship. *IEEE Geosci. Remote Sens. Mag.* **2016**, *4*, 6–23. [CrossRef]

Bibliography

[1] H. A. Zebker and R. M. Goldstein, "Topographic mapping from interferometric synthetic aperture radar observations," *Journal of Geophysical Research*, vol. 91, no. B5, p. 4993, 1986.

[2] R. Bamler and P. Hartl, "Synthetic aperture radar interferometry," *Inverse problems*, vol. 14, no. 4, R1, 1998.

[3] P. Rosen, S. Hensley, I. Joughin, F. Li, S. Madsen, E. Rodriguez, and R. Goldstein, "Synthetic aperture radar interferometry," *Proceedings of the IEEE*, vol. 88, no. 3, pp. 333–382, Mar. 2000.

[4] R. F. Hanssen, *Radar interferometry: data interpretation and error analysis*, ser. Remote sensing and digital image processing v. 2. Dordrecht ; Boston: Kluwer Academic, 2001.

[5] A. K. Gabriel, R. M. Goldstein, and H. A. Zebker, "Mapping small elevation changes over large areas: Differential radar interferometry," *Journal of Geophysical Research*, vol. 94, no. B7, p. 9183, 1989.

[6] A. Ferretti, C. Prati, and F. Rocca, "Permanent scatterers in SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 1, pp. 8–20, Jan. 2001.

[7] B. M. Kampes, *Radar interferometry: persistent scatterer technique*, ser. Remote sensing and digital image processing v. 12. Dordrecht, the Netherlands: Springer, 2006.

[8] A. Reigber and A. Moreira, "First demonstration of airborne SAR tomography using multibaseline L-band data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 38, no. 5, pp. 2142–2152, Sep. 2000.

[9] G. Fornaro, F. Serafino, and F. Soldovieri, "Three-dimensional focusing with multipass SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 3, pp. 507–517, Mar. 2003.

[10] F. Lombardini, "Differential tomography: A new framework for SAR interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 1, pp. 37–44, Jan. 2005.

[11] S. Gernhardt and R. Bamler, "Deformation monitoring of single buildings using meter-resolution SAR data in PSI," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 73, pp. 68–79, Sep. 2012.

[12] X. X. Zhu and R. Bamler, "Very high resolution spaceborne SAR tomography in urban environment," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 48, no. 12, pp. 4296–4308, 2010.

[13] M. Eineder, C. Minet, P. Steigenberger, X. Y. Cong, and T. Fritz, "Imaging Geodesy - Toward Centimeter-Level Ranging Accuracy With TerraSAR-X," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 2, pp. 661–671, Feb. 2011.

[14] A. Schubert, M. Jehle, D. Small, and E. Meier, "Mitigation of atmospheric perturbations and solid Earth movements in a TerraSAR-X time-series," *Journal of Geodesy*, vol. 86, no. 4, pp. 257–270, Apr. 2012.

[15] Y. Yoon, M. Eineder, N. Yague-Martinez, and O. Montenbruck, "TerraSAR-X Precise Trajectory Estimation and Quality Assessment," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 6, pp. 1859–1868, Jun. 2009.

[16] S. Hackel, O. Montenbruck, P. Steigenberger, U. Balss, C. Gisinger, and M. Eineder, "Model improvements and validation of TerraSAR-X precise orbit determination," *Journal of Geodesy*, vol. 91, no. 5, pp. 547–562, May 2017.

[17] X. Y. Cong, U. Balss, M. Eineder, and T. Fritz, "Imaging Geodesy - Centimeter-Level Ranging Accuracy With TerraSAR-X: An Update," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 5, pp. 948–952, Sep. 2012.

[18] U. Balss, C. Gisinger, and M. Eineder, "Measurements on the Absolute 2-D and 3-D Localization Accuracy of TerraSAR-X," *Remote Sensing*, vol. 10, no. 4, p. 656, Apr. 2018.

[19] C. Gisinger, U. Balss, R. Pail, X. X. Zhu, S. Montazeri, S. Gernhardt, and M. Eineder, "Precise Three-Dimensional Stereo Localization of Corner Reflectors and Persistent Scatterers With TerraSAR-X," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 1782–1802, Apr. 2015.

[20] I. G. Cumming and F. H.-c. Wong, *Digital processing of synthetic aperture radar data: algorithms and implementation*, ser. Artech House remote sensing library. Boston: Artech House, 2005.

[21] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou, "A tutorial on synthetic aperture radar," *IEEE Geoscience and Remote Sensing Magazine*, vol. 1, no. 1, pp. 6–43, Mar. 2013.

[22] S. Buckreuss, W. Balzer, P. Muhlbauer, R. Werninghaus, and W. Pitz, "The TerraSAR-X satellite project," vol. 5, IEEE, 2003, pp. 3096–3098.

[23] F. De Zan and A. Monti Guarnieri, "TOPSAR: Terrain Observation by Progressive Scans," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 9, pp. 2352–2360, Sep. 2006.

[24] R. Bamler, M. Eineder, N. Adam, X. Zhu, and S. Gernhardt, "Interferometric Potential of High Resolution Spaceborne SAR," *Photogrammetrie - Fernerkundung - Geoinformation*, vol. 2009, no. 5, pp. 407–419, Nov. 2009.

[25] S. Gernhardt, X. Y. Cong, M. Eineder, S. Hinz, and R. Bamler, "Geometrical Fusion of Multitrack PS Point Clouds," *IEEE Geoscience and Remote Sensing Letters*, vol. 9, no. 1, pp. 38–42, Jan. 2012.

[26] Y. Wang and X. X. Zhu, "Automatic Feature-Based Geometric Fusion of Multiview TomoSAR Point Clouds in Urban Area," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, pp. 1–1, 2014.

[27] M. Schwabisch, "A fast and efficient technique for SAR interferogram geocoding," in *Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, 1998, 1100–1102 vol.2.

[28] M. Eineder, "Efficient simulation of sar interferograms of large areas and of rugged terrain," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 6, pp. 1415–1427, Jun. 2003.

[29] F. van Leijen, "Persistent Scatterer Interferometry based on geodetic estimation theory," OCLC: 905870982, PhD thesis, Delft University of Technology, Delft, 2014.

[30] H. Zebker and J. Villasenor, "Decorrelation in interferometric radar echoes," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 30, no. 5, pp. 950–959, Sep. 1992.

[31] C. Rossi, F. Rodriguez Gonzalez, T. Fritz, N. Yague-Martinez, and M. Eineder, "TanDEM-X calibrated Raw DEM generation," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 73, pp. 12–20, Sep. 2012.

[32] E. Erten, A. Reigber, and O. Hellwich, "Generation of three-dimensional deformation maps from InSAR data using spectral diversity techniques," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 65, no. 4, pp. 388–394, Jul. 2010.

[33] L. Gray, "Using multiple RADARSAT InSAR pairs to estimate a full three-dimensional solution for glacial ice movement," *Geophysical Research Letters*, vol. 38, no. 5, pp. 1–6, Mar. 2011.

[34] J. Hu, Z. Li, X. Ding, J. Zhu, L. Zhang, and Q. Sun, "Resolving three-dimensional surface displacements from InSAR measurements: A review," *Earth-Science Reviews*, vol. 133, pp. 1–17, Jun. 2014.

[35] P. Teunissen, D. Simons, and C. Tiberius, *Probability and Observation Theory*. Delft: Delft University of Technology, 2009.

[36] E. M. Mikhail and F. E. Ackermann, *Observations and least squares*. Washington, D.C: University Press of America, 1982.

[37] S. C. Narula and J. F. Wellington, "The Minimum Sum of Absolute Errors Regression: A State of the Art Survey," *International Statistical Review / Revue Internationale de Statistique*, vol. 50, no. 3, p. 317, Dec. 1982.

[38] A. Amiri-Simkooei, "Formulation of L1 Norm Minimization in Gauss-Markov Models," *Journal of Surveying Engineering*, vol. 129, no. 1, pp. 37–43, Feb. 2003.

[39] U. Balss, X. Y. Cong, R. Brcic, M. Rexer, C. Minet, H. Breit, M. Eineder, and T. Fritz, "High precision measurement on the absolute localization accuracy of TerraSAR-X," in *Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Jul. 2012, pp. 1625–1628.

[40] S. Hackel, O. Montenbruck, P. Steigenberger, U. Balss, C. Gisinger, and M. Eineder, "Model improvements and validation of TerraSAR-X precise orbit determination," *Journal of Geodesy*, Dec. 2016.

[41] S. Hackel, C. Gisinger, U. Balss, M. Wermuth, and O. Montenbruck, "Long-Term Validation of TerraSAR-X Orbit Solutions with Laser and Radar Measurements," *Remote Sensing*, no. Special Issue Ten Years of TerraSAR-X Scientific Results, 2018.

[42] M. Jehle, D. Perler, D. Small, A. Schubert, and E. Meier, "Estimation of Atmospheric Path Delays in TerraSAR-X Data using Models vs. Measurements," *Sensors*, vol. 8, no. 12, pp. 8479–8491, Dec. 2008.

[43] A. Schubert, M. Jehle, D. Small, and E. Meier, "Influence of Atmospheric Path Delay on the Absolute Geolocation Accuracy of TerraSAR-X High-Resolution Products," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 2, pp. 751–758, Feb. 2010.

[44] P. Misra and P. Enge, *Global positioning system: signals, measurements, and performance*, 2. ed. Lincoln, Mass: Ganga-Jamuna Pr, 2006.

[45] C. Gisinger, "Atmospheric corrections for TerraSAR-X derived from GNSS observations," Master's thesis, Technische Universität München, München, 2012.

[46] U. Balss, C. Gisinger, X. Y. Cong, R. Brcic, P. Steigenberger, M. Eineder, R. Pail, and U. Hugentobler, "High resolution geodetic earth observation with TerraSAR-X: Correction schemes and validation," in *2013 IEEE International Geoscience and Remote Sensing Symposium - IGARSS*, Melbourne, Australia: IEEE, Jul. 2013, pp. 4499–4502.

[47] X. Cong, "Sar interferometry for volcano monitoring: 3d-psi analysis and mitigation of atmospheric refractivity.," PhD thesis, Technische Universität München, Jun. 2014, pp. 1–174.

[48] G. Petit and B. Luzum, "IERS Conventions," Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main, IERS Technical Note No. 36, 2010, p. 179.

[49] U. Balss, C. Gisinger, X. Cong, M. Eineder, T. Fritz, H. Breit, and R. Brcic, "Gnss-based signal path delay and geodynamic corrections for centimeter level pixel localization with terrasar-x," in *5th TerraSAR-X Science Team Meeting*, 2013, pp. 1–4.

[50] U. Balss, H. Breit, T. Fritz, U. Steinbrecher, C. Gisinger, and M. Eineder, "Analysis of internal timings and clock rates of TerraSAR-X," in *Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Jul. 2014, pp. 2671–2674.

[51] H. Breit, T. Fritz, U. Balss, M. Lachaise, A. Niedermeier, and M. Vonavka, "TerraSAR-X SAR processing and products," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 2, pp. 727–740, 2010.

[52] S. Duque, U. Balss, X. Y. Cong, N. Yague-Martinez, and T. Fritz, "Absolute ranging for maritime applications using terrasar-x and tandem-x data," in *EUSAR 2014; 10th European Conference on Synthetic Aperture Radar*, Jun. 2014, pp. 1–4.

[53] M. Eineder, U. Balss, and S. D. Biarge, "Water level measurement by controlled radar reflection and TerraSAR-X imaging geodesy," IEEE, Jul. 2014, pp. 5141–5143.

[54] U. Balss, C. Gisinger, X. Y. Cong, R. Brcic, S. Hackel, and M. Eineder, "Precise measurements on the absolute localization accuracy of TerraSAR-X on the base of far-distributed test sites," in *In Proceeding of the 10th European Conference on Synthetic Aperture Radar (EUSAR)*, 2014.

[55] X. Cong, U. Balss, F. Rodriguez Gonzalez, and M. Eineder, "Mitigation of Tropospheric Delay in SAR and InSAR Using NWP Data: Its Validation and Application Examples," *Remote Sensing*, vol. 10, no. 10, p. 1515, Sep. 2018.

[56] M. Eineder, U. Balss, S. Suchandt, C. Gisinger, X. Cong, and H. Runge, "A definition of next-generation SAR products for geodetic applications," in *Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Jul. 2015, pp. 1638–1641.

[57] A. Schubert, D. Small, N. Miranda, D. Geudtner, and E. Meier, "Sentinel-1a Product Geolocation Accuracy: Commissioning Phase Results," *Remote Sensing*, vol. 7, no. 7, pp. 9431–9449, Jul. 2015.

[58] A. Schubert, N. Miranda, D. Geudtner, and D. Small, "Sentinel-1a/B Combined Product Geolocation Accuracy," *Remote Sensing*, vol. 9, no. 6, p. 607, Jun. 2017.

[59] G. Gisinger, U. Balss, H. Breit, A. Schubert, M. Garthwaite, D. Small, T. Gruber, M. Eineder, T. Fritz, and N. Miranda, "Recent findings on the sentinel-1 geolocation accuracy using the australian corner reflector array," in *Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Jul. 2018, pp. 6360–6363.

[60] P. Dheenathayalan, D. Small, A. Schubert, and R. F. Hanssen, "High-precision positioning of radar scatterers," *Journal of Geodesy*, vol. 90, no. 5, pp. 403–422, May 2016.

[61] P. Dheenathayalan, D. Small, and R. F. Hanssen, "3-D Positioning and Target Association for Medium-Resolution SAR Sensors," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–13, 2018.

[62] F. W. Leberl, *Radargrammetric image processing*, ser. Artech House remote sensing library. Norwood, MA: Artech House, 1990.

[63] Pu-Huai Chen and I. Dowman, "A weighted least squares solution for space intersection of spaceborne stereo SAR data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 39, no. 2, pp. 233–240, Feb. 2001.

[64] E. Sansosti, "A simple and exact solution for the interferometric and stereo SAR geolocation problem," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 8, pp. 1625–1634, Aug. 2004.

[65] K. Eldhuset and D. J. Weydahl, "Geolocation and Stereo Height Estimation Using TerraSAR-X Spotlight Image Data," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3574–3581, Oct. 2011.

[66] U. Balss, C. Gisinger, X. Cong, M. Eineder, and R. Brcic, "Precise 2-d and 3-d ground target localization with terrasar-x," in *ISPRS Hannover Workshop*, C. Heipke, K. Jacobsen, F. Rottensteiner, and U. Sörgel, Eds., ser. ISPRS Archives, vol. XL-1/W, Copernicus Publications, May 2013, pp. 23–28.

[67] H. Raggam, R. Perko, K. Gutjahr, N. Kiefl, W. Koppe, and S. Hennig, "Accuracy assessment of 3d point retrieval from terrasar-x data sets," in *8th European Conference on Synthetic Aperture Radar*, VDE, Jun. 2010, pp. 1–4.

[68] W. Koppe, R. Wenzel, S. Hennig, J. Janoth, P. Hummel, and H. Raggam, "Quality assessment of TerraSAR-X derived ground control points," IEEE, Jul. 2012, pp. 3580–3583.

[69] M. Willberg, C. Gisinger, U. Balss, T. Fritz, and M. Eineder, "Geodetic stereo SAR with small multi-directional radar reflectors," in *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Beijing, China: IEEE, Jul. 2016, pp. 5011–5014.

[70] C. Gisinger, M. Willberg, U. Balss, T. Klügel, S. Mähler, R. Pail, and M. Eineder, "Differential geodetic stereo SAR with TerraSAR-X by exploiting small multi-directional radar reflectors," *Journal of Geodesy*, vol. 91, no. 1, pp. 53–67, Jan. 2017.

[71] C. Gisinger, S. Gernhardt, S. Auer, U. Balss, S. Hackel, R. Pail, and M. Eineder, "Absolute 4-D positioning of persistent scatterers with TerraSAR-X by applying geodetic stereo SAR," IEEE, Jul. 2015, pp. 2991–2994.

[72] C. Gisinger, S. Montazeri, U. Balss, X. Y. Cong, S. Hackel, X. X. Zhu, R. Pail, and M. Eineder, "Applying geodetic sar with terrasar-x and tandem-x," in *TerraSAR-X / TanDEM-X Science Team Meeting*, Oct. 2016.

[73] H. Runge, U. Balss, S. Suchandt, R. Klarner, and X. Cong, "Drivemark – generation of high resolution road maps with radar satellites," in *11th ITS European Congress*, ITS Europe, Jun. 2016, pp. 1–6.

[74] S. Montazeri, X. X. Zhu, U. Balss, C. Gisinger, Y. Wang, M. Eineder, and R. Bamler, "SAR ground control point identification with the aid of high resolution optical data," IEEE, Jul. 2016, pp. 3205–3208.

[75] U. Balss, H. Runge, S. Suchandt, and X. Y. Cong, "Automated extraction of 3-D Ground Control Points from SAR images - an upcoming novel data product," in *Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Jul. 2016, pp. 5023–5026.

[76] S. Montazeri, C. Gisinger, X. X. Zhu, M. Eineder, and R. Bamler, "Automatic positioning of SAR ground control points from multi-aspect TerraSAR-X acquisitions," IEEE, Jul. 2017, pp. 961–964.

[77] S. Montazeri, C. Gisinger, M. Eineder, and X. X. Zhu, "Automatic Detection and Positioning of Ground Control Points Using TerraSAR-X Multiaspect Acquisitions," *IEEE Transactions on Geoscience and Remote Sensing*, pp. 1–20, 2018.

[78] D. O. Nitti, A. Morea, R. Nutricato, M. T. Chiaradia, C. La Mantia, L. Agrimano, and S. Samarelli, "Automatic GCP extraction with high resolution COSMO-SkyMed products," C. Notarnicola, S. Paloscia, N. Pierdicca, and E. Mitchard, Eds., Oct. 2016, p. 1 000 302.

[79] K. Eldhuset, "Combination of stereo SAR and InSAR for DEM generation using TanDEM-X spotlight data," *International Journal of Remote Sensing*, vol. 38, no. 15, pp. 4362–4378, Aug. 2017.

[80] S. Gernhardt, N. Adam, M. Eineder, and R. Bamler, "Potential of very high resolution SAR for persistent scatterer interferometry in urban areas," *Annals of GIS*, vol. 16, no. 2, pp. 103–111, Aug. 2010.

[81] M. Crosetto, O. Monserrat, M. Cuevas-González, N. Devanthéry, and B. Crippa, "Persistent Scatterer Interferometry: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 115, pp. 78–89, May 2016.

[82] N. Adam, B. M. Kampes, and M. Eineder, "The development of a scientific persistent scatterer system: Modifications for mixed ERS/ENVISAT time series," in *Proceedings of the Envisat and ERS Symposium*, Salzburg, Austria, 2004, pp. 1–9.

[83] N. Adam, B. M. Kampes, M. Eineder, J. Worawattanamateekul, and M. Kircher, "The Development of a Scientific Permanent Scatterer System," in *Proceedings of the Joint ISPRS/EARSeL Workshop on High Resolution Mapping from Space 2003*, Hannover, 2003.

[84] N. Adam, F. R. Gonzalez, A. Parizzi, and R. Brcic, "Wide area Persistent Scatterer Interferometry: Current developments, algorithms and examples," in *Geoscience and Remote Sensing Symposium (IGARSS)*, IEEE, Jul. 2013, pp. 1857–1860.

[85] F. Rodriguez Gonzalez, N. Adam, A. Parizzi, and R. Brcic, "The Integrated Wide Area Processor (IWAP): A Processor For Wide Area Persistent Scatterer Interferometry," in *Proceedings of ESA Living Planet Symposium 2013*, Edinburgh, UK: ESA, Sep. 2013, pp. 1–4.

[86] C. Werner, U. Wegmuller, T. Strozzi, and A. Wiesmann, "Interferometric point target analysis for deformation mapping," in *Geoscience and Remote Sensing Symposium (IGARSS)*, vol. 7, IEEE, 2003, pp. 4362–4364.

[87] A. Hooper, H. Zebker, P. Segall, and B. Kampes, "A new method for measuring deformation on volcanoes and other natural terrains using InSAR persistent scatterers: A NEW PERSISTENT SCATTERERS METHOD," *Geophysical Research Letters*, vol. 31, no. 23, Dec. 2004.

[88] Y. Wang, X. X. Zhu, and R. Bamler, "An Efficient Tomographic Inversion Approach for Urban Mapping Using Meter Resolution SAR Image Stacks," *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 7, pp. 1250–1254, Jul. 2014.

[89] P. J. G. Teunissen, "The least-squares ambiguity decorrelation adjustment: A method for fast GPS integer ambiguity estimation," *Journal of Geodesy*, vol. 70, no. 1-2, pp. 65–82, Nov. 1995.

[90] B. Kampes and R. Hanssen, "Ambiguity resolution for permanent scatterer interferometry," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 42, no. 11, pp. 2446–2453, Nov. 2004.

[91] F. Rodriguez Gonzalez, A. Bhutani, and N. Adam, "L1 network inversion for robust outlier rejection in persistent Scatterer Interferometry," IEEE, Jul. 2011, pp. 75–78.

[92] C. Colesanti, A. Ferretti, F. Novali, C. Prati, and F. Rocca, "Sar monitoring of progressive and seasonal ground deformation using the permanent scatterers technique," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 41, no. 7, pp. 1685–1701, Jul. 2003.

[93] D. Perissin and F. Rocca, "High-Accuracy Urban DEM Using Permanent Scatterers," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 11, pp. 3338–3347, Nov. 2006.

[94] D. Perissin, "Validation of the Submetric Accuracy of Vertical Positioning of PSs in C-Band," *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 3, pp. 502–506, Jul. 2008.

[95] D. Perissin and A. Ferretti, "Urban-Target Recognition by Means of Repeated Spaceborne SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 45, no. 12, pp. 4043–4058, Dec. 2007.

[96] L. Chang and R. F. Hanssen, "Detection of cavity migration and sinkhole risk using radar interferometric time series," *Remote Sensing of Environment*, vol. 147, pp. 56–64, May 2014.

[97] S. Gernhardt, "High Precision 3D Localization and Motion Analysis of Persistent Scatterers using Meter-Resolution Radar Satellite Data," Doctoral Thesis, Technische Universität München, München, 2012.

[98] S. Gernhardt, S. Auer, and K. Eder, "Persistent scatterers at building facades – Evaluation of appearance and localization accuracy," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 100, pp. 92–105, Feb. 2015.

[99] O. Frey, F. Morsdorf, and E. Meier, "Tomographic Imaging of a Forested Area By Airborne Multi-Baseline P-Band SAR," *Sensors*, vol. 8, no. 9, pp. 5884–5896, Sep. 2008.

[100] S. Tebaldini and F. Rocca, "Multibaseline Polarimetric SAR Tomography of a Boreal Forest at P- and L-Bands," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 232–246, Jan. 2012.

[101] E. Aguilera, M. Nannini, and A. Reigber, "A Data-Adaptive Compressed Sensing Approach to Polarimetric SAR Tomography of Forested Areas," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 3, pp. 543–547, May 2013.

[102] X. X. Zhu and R. Bamler, "Let's do the time warp: Multicomponent nonlinear motion estimation in differential SAR tomography," *Geoscience and Remote Sensing Letters, IEEE*, vol. 8, no. 4, pp. 735–739, 2011.

[103] X. X. Zhu, "Very High Resolution Tomographic SAR Inversion for Urban Infrastructure Monitoring - A Sparse and Nonlinear Tour," Doctoral Thesis, Technische Universität München, München, 2011.

[104] F. Lombardini, M. Montanari, and F. Gini, "Reflectivity estimation for multibaseline interferometric radar imaging of layover extended sources," *IEEE Transactions on Signal Processing*, vol. 51, no. 6, pp. 1508–1519, Jun. 2003.

[105] G. Fornaro, D. Reale, and F. Serafino, "Four-dimensional SAR imaging for height estimation and monitoring of single and double scatterers," *Geoscience and Remote Sensing, IEEE Transactions on*, vol. 47, no. 1, pp. 224–237, 2009.

[106] X. Zhu and R. Bamler, "Tomographic SAR Inversion by L1-Norm Regularization – The Compressive Sensing Approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 10, pp. 3839–3846, 2010.

[107] X. X. Zhu and R. Bamler, "Demonstration of Super-Resolution for Tomographic SAR Imaging in Urban Environment," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 8, pp. 3150–3157, Aug. 2012.

[108] X. Zhu and R. Bamler, "Super-Resolution Power and Robustness of Compressive Sensing for Spectral Estimation With Application to Spaceborne Tomographic SAR," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 1, pp. 247–258, 2012.

[109] X. X. Zhu and R. Bamler, "Superresolving sar tomography for multidimensional imaging of urban areas: Compressive sensing-based tomosar inversion," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 51–58, Jul. 2014.

[110] Y. Shi, X. X. Zhu, W. Yin, and R. Bamler, "A fast and accurate basis pursuit denoising algorithm with application to super-resolving tomographic SAR," *IEEE Transactions on Geoscience and Remote Sensing*, May 2018, in press.

[111] X. Zhu, N. Ge, and M. Shahzad, "Group sparsity in sar tomography for urban mapping," *IEEE Journal of Selected Topics in Signal Processing*, 2015, accepted.

[112] M. A. Siddique, U. Wegmüller, I. Hajnsek, and O. Frey, "Single-Look SAR Tomography as an Add-On to PSI for Improved Deformation Analysis in Urban Areas," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 10, pp. 6119–6137, Oct. 2016.

[113] E. Sansosti, R. Lanari, G. Fornaro, G. Franceschetti, M. Tesauro, G. Puglisi, and M. Coltelli, "Digital elevation model generation using ascending and descending ERS-1/ERS-2 tandem data," *International Journal of Remote Sensing*, vol. 20, no. 8, pp. 1527–1547, Jan. 1999.

[114] R. Deo, C. Rossi, M. Eineder, T. Fritz, and Y. S. Rao, "Framework for Fusion of Ascending and Descending Pass TanDEM-X Raw DEMs," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 7, pp. 3347–3355, Jul. 2015.

[115] A. Gruber, B. Wessel, M. Martone, and A. Roth, "The TanDEM-X DEM Mosaicking: Fusion of Multiple Acquisitions Using InSAR Quality Parameters," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 3, pp. 1047–1057, Mar. 2016.

[116] M. Eineder and N. Adam, "A maximum-likelihood estimator to simultaneously unwrap, geocode, and fuse SAR interferograms from different viewing geometries into one digital elevation model," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 1, pp. 24–36, Jan. 2005.

[117] Antonio Pepe and Fabiana Calò, "A Review of Interferometric Synthetic Aperture RADAR (InSAR) Multi-Track Approaches for the Retrieval of Earth's Surface Displacements," *Applied Sciences*, vol. 7, no. 12, p. 1264, Dec. 2017.

[118] C. Colesanti, A. Ferretti, C. Prati, and F. Rocca, "Full exploitation of the ERS archive: Multi data set permanent scatterers analysis," in *IEEE International Geoscience and Remote Sensing Symposium*, vol. 2, Toronto, Ont., Canada: IEEE, 2002, pp. 1234–1236.

[119] G. Ketelaar, F. van Leijen, P. Marinkovic, and R. Hanssen, "Multi-track PS-InSAR datum connection," in *2007 IEEE International Geoscience and Remote Sensing Symposium*, Barcelona, Spain: IEEE, 2007, pp. 2481–2484.

[120] M. Schmitt, O. Maksymiuk, C. Magnard, and U. Stilla, "Radargrammetric registration of airborne multi-aspect SAR data of urban areas," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 86, pp. 11–20, Dec. 2013.

[121] M. Schmitt and U. Stilla, "Maximum-likelihood estimation for multi-aspect multi-baseline SAR interferometry of urban areas," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, pp. 68–77, Jan. 2014.

[122] F. Rocca, "3D motion recovery with multi-angle and/or left right interferometry," in *Proceedings of FRINGE 2003 Workshop*, Frascati, Italy: ESA, 2003, pp. 1–5.

[123] H. Ansari, F. De Zan, A. Parizzi, M. Eineder, K. Goel, and N. Adam, "Measuring 3-D Surface Motion With Future SAR Systems Based on Reflector Antennae," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 2, pp. 272–276, Feb. 2016.

[124] T. J. Wright, B. E. Parsons, and Z. Lu, "Toward mapping surface deformation in three dimensions using InSAR," *Geophysical Research Letters*, vol. 31, no. 1, 2004.

[125] X. X. Zhu, Y. Wang, S. Gernhardt, and R. Bamler, "Tomo-GENESIS: DLR's tomographic SAR processing system," in *Joint Urban Remote Sensing (JURSE)*, IEEE, Apr. 2013, pp. 159–162.

[126] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[127] S. Suri, P. Schwind, J. Uhl, and P. Reinartz, "Modifications in the SIFT operator for effective SAR image matching," *International Journal of Image and Data Fusion*, vol. 1, no. 3, pp. 243–256, Sep. 2010.

[128] F. Dellinger, J. Delon, Y. Gousseau, J. Michel, and F. Tupin, "SAR-SIFT: A SIFT-Like Algorithm for SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 1, pp. 453–466, Jan. 2015.

[129] R. C. Gonzalez and R. E. Woods, *Digital image processing*, 2nd ed. Upper Saddle River, N.J: Prentice Hall, 2002.

[130] K. Briechle and U. D. Hanebeck, "Template matching using fast normalized cross correlation," D. P. Casasent and T.-H. Chao, Eds., Mar. 2001, pp. 95–102.

[131] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 5, pp. 603–619, May 2002.

[132] P. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, Feb. 1992.

[133] OpenStreetMap contributors, *Planet dump retrieved from https://planet.osm.org*, `https://www.openstreetmap.org`, 2017.

[134] M. Hubert and E. Vandervieren, "An adjusted boxplot for skewed distributions," *Computational Statistics & Data Analysis*, vol. 52, no. 12, pp. 5186–5201, Aug. 2008.

[135] J. W. Tukey, *Exploratory data analysis*, ser. Addison-Wesley series in behavioral science. Reading, Mass: Addison-Wesley Pub. Co, 1977.

[136] G Brys, M Hubert, and A Struyf, "A Robust Measure of Skewness," *Journal of Computational and Graphical Statistics*, vol. 13, no. 4, pp. 996–1017, Dec. 2004.

[137] P. J. Rousseeuw and M. Hubert, "Anomaly detection by robust statistics," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 2, pp. 1–14, Mar. 2018.

[138] S. Montazeri, X. X. Zhu, M. Eineder, and R. Bamler, "Three-Dimensional Deformation Monitoring of Urban Infrastructure by Tomographic SAR Using Multi-track TerraSAR-X Data Stacks," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 12, pp. 6868–6878, Dec. 2016.

[139] S. Auer, S. Gernhardt, and R. Bamler, "Ghost Persistent Scatterers Related to Multiple Signal Reflections," *IEEE Geoscience and Remote Sensing Letters*, vol. 8, no. 5, pp. 919–923, Sep. 2011.

[140] S. Montazeri, "The fusion of SAR tomography and Stereo-SAR for 3d absolute scatterer positioning," Master's thesis, Delft University of Technology, Delft, 2014.

[141] Y. Wang, X. Zhu, S. Montazeri, J. Kang, L. Mou, and M. Schmitt, "Potential of the "sarptical" system," in *Proc. 10th Int. Workshop Adv. Sci. Appl. SAR Interferometry Sentinel-1 InSAR*, 2017.

[142] C. Qiu, M. Schmitt, and X. X. Zhu, "Towards automatic SAR-optical stereogrammetry over urban areas using very high resolution imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 138, pp. 218–231, Apr. 2018.

[143] H. Bagheri, M. Schmitt, P. d'Angelo, and X. X. Zhu, "A framework for SAR-optical stereogrammetry over urban areas," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 146, pp. 389–408, Dec. 2018.

[144] Y. Wang, X. X. Zhu, B. Zeisl, and M. Pollefeys, "Fusing Meter-Resolution 4-D InSAR Point Clouds and Optical Images for Semantic Urban Infrastructure Monitoring," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 1, pp. 14–26, Jan. 2017.

[145] L. H. Hughes, M. Schmitt, L. Mou, Y. Wang, and X. X. Zhu, "Identifying Corresponding Patches in SAR and Optical Images With a Pseudo-Siamese CNN," *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 784–788, May 2018.

[146] N. Merkle, S. Auer, R. Muller, and P. Reinartz, "Exploring the Potential of Conditional Adversarial Networks for Optical and SAR Image Matching," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 6, pp. 1811–1820, Jun. 2018.

[147] P. Mahapatra, H. v. der Marel, F. van Leijen, S. Samiei-Esfahany, R. Klees, and R. Hanssen, "InSAR datum connection using GNSS-augmented radar transponders," *Journal of Geodesy*, vol. 92, no. 1, pp. 21–32, Jan. 2018.

[148] C. Rossi and S. Gernhardt, "Urban DEM generation, analysis and enhancements using TanDEM-X," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 85, pp. 120–131, Nov. 2013.

[149] C. Dubois, A. Nascetti, A. Thiele, M. Crespi, and S. Hinz, "SAR-SIFT for Matching Multiple SAR Images and Radargrammetry," *PFG – Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, vol. 85, no. 3, pp. 149–158, Aug. 2017.

[150] S. Auer, S. Hinz, and R. Bamler, "Ray-Tracing Simulation Techniques for Understanding High-Resolution SAR Images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1445–1456, Mar. 2010.

[151] A. M. Zoubir, V. Koivunen, Y. Chakhchoukh, and M. Muma, "Robust Estimation in Signal Processing: A Tutorial-Style Treatment of Fundamental Concepts," *IEEE Signal Processing Magazine*, vol. 29, no. 4, pp. 61–80, Jul. 2012.

List of Figures

List of Tables