



TECHNISCHE UNIVERSITÄT MÜNCHEN
Institut für Photogrammetrie und Kartographie
Fachgebiet Photogrammetrie und Fernerkundung

Reconstruction of Urban Surface Models from
Multi-Aspect and Multi-Baseline Interferometric SAR

Michael Schmitt

Dissertation

2014



TECHNISCHE UNIVERSITÄT MÜNCHEN
Institut für Photogrammetrie und Kartographie
Fachgebiet Photogrammetrie und Fernerkundung

Reconstruction of Urban Surface Models from Multi-Aspect and Multi-Baseline Interferometric SAR

Michael Schmitt

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

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

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

Prüfer der Dissertation:

1. Univ.-Prof. Dr.-Ing. Uwe Stilla
2. Univ.-Prof. Dr.-Ing. Joachim Ender
Universität Siegen
3. Univ.-Prof. Dr.-Ing. Uwe Sörgel
Technische Universität Darmstadt

Die Dissertation wurde am 15.01.2014 bei der Technischen Universität München eingereicht und durch die Ingenieurfacultät Bau Geo Umwelt am 07.07.2014 angenommen.

Abstract

Synthetic aperture radar interferometry has long been used to derive three-dimensional topographic information independently of cloud coverage or daytime. This has made it a valuable tool for rapid data acquisition in time-critical scenarios such as defense-related reconnaissance missions or disaster response situations. The side-looking SAR imaging geometry, however, leads to disturbing effects like layover and shadowing. In order to cope with these deficiencies, this thesis investigates enhanced InSAR processing strategies for a reconstruction of urban surface models utilizing both multi-aspect and multi-baseline data. Since only sensors carried by aircraft provide the possibility to acquire high-coherent single-pass data from almost arbitrary aspect angles, this thesis is focused on the utilization of airborne SAR.

The first step in a multi-baseline InSAR processing chain usually is the estimation of the complex covariance matrices for all pixels in the stack of co-registered images, since these matrices contain the full interferometric information of the corresponding resolution cell. Therefore, the first contribution of this thesis is the proposition of two new adaptive covariance matrix estimation procedures, which are specifically designed for single-pass InSAR stacks consisting of just few acquisitions.

Secondly, a novel maximum-likelihood-based SAR tomography algorithm aiming at the separation of layovered scatterers and the focusing of sparse three-dimensional SAR images is described. In contrast to most hitherto proposed TomoSAR approaches, it does not rely on an exploitation of repeat-pass data stacks in order to provide a large overall baseline and high number of observations per resolution cell for a sufficient elevation resolution.

Complementary to that, as a third contribution, the fusion of InSAR data acquired from multiple aspect angles is investigated with respect to the goal to fill in information occluded by radar shadowing in single aspects. For this reason, a radargrammetric registration approach for multi-aspect SAR data is proposed, which is used as a pre-processing step for further fusion operations. Building on this pre-requisite, a new maximum-likelihood estimation framework is developed, which is used to fuse multi-baseline InSAR data acquired from multiple aspect angles simultaneously in order to derive comprehensive 2.5D height models. In analogy, a voxel-space-based fusion of 3D point clouds generated by SAR Tomography is proposed.

The applicability of all the methods described in this thesis is analyzed using experimental SAR data acquired by the airborne millimeterwave sensor MEMPHIS, which is able to provide single-pass multi-baseline InSAR stacks containing four co-registered images. The test dataset shows the inner city area of Munich, Germany, and serves as an example for complex urban scenes. It is comprised of dense building blocks, isolated large buildings, streets, and many urban trees.

In a first set of experiments, both efficiency and adaptivity of the covariance matrix estimation methods are evaluated based on image processing techniques. Subsequently, the layover resolution capabilities of the maximum-likelihood TomoSAR algorithm are examined, before the reconstruction results achieved by simultaneous fusion of multi-aspect multi-baseline InSAR data are investigated. Finally, these reconstruction results are compared to the 3D data resulting from a fusion of multi-aspect TomoSAR point clouds. By matching both kinds of reconstruction results to a dense point cloud derived from helicopter-borne multi-aspect LiDAR measurements, it could be shown that a comprehensive reconstruction of 2.5D height maps and 3D point clouds with accuracies of about 1 m are possible from airborne single-pass InSAR data.

Kurzfassung

Schon seit vielen Jahren wird Synthetische Apertur Radar-Interferometrie dazu verwendet, die dreidimensionale Topographie der Erdoberfläche unabhängig von Wolkenbedeckung oder Tageszeit zu erfassen. Sie gilt deshalb als wertvolles Werkzeug für schnelle Datenerfassung vor allem in zeitkritischen Szenarien. Die seitwärtsblickende SAR-Abbildungsgeometrie führt jedoch zu den störenden Effekten Überlagerung und Radarschatten. Mit dem Ziel, diese Nachteile aufzulösen, untersucht die vorliegende Arbeit innovative InSAR-Prozessierungsstrategien zur Rekonstruktion urbaner Oberflächenmodelle unter Verwendung von Daten mehrerer Aspekte sowie Basislinien. Da nur Flugzeuge die Möglichkeit bieten, hochkohärente Single-Pass-Daten von beinahe beliebigen Aspektwinkeln aufzunehmen, liegt der Fokus der Arbeit dabei auf der Verwendung von flugzeuggetragenem SAR.

Der erste Schritt in einer Mehrfachbasislinien-InSAR-Prozesskette ist normalerweise die Schätzung der komplexen Kovarianzmatrizen aller Pixel im Stapel der koregistrierten Bilder, da diese Matrizen die vollständige interferometrische Information der zugehörigen Auflösungszelle beinhalten. Deshalb ist der erste Beitrag dieser Arbeit die Vorstellung zweier neuer adaptiver Verfahren zur Kovarianzmatrix-Schätzung, die speziell für Single-Pass-InSAR-Stapel mit nur wenigen Aufnahmen entworfen wurden.

Als zweites wird ein neuartiger Algorithmus zur SAR-Tomographie beschrieben, der darauf abstellt, überlagerte Streuer zu trennen und die Fokussierung dünnbesetzter dreidimensionaler SAR-Bilder zu ermöglichen. Im Gegensatz zu den meisten bislang vorgeschlagenen TomoSAR-Ansätzen ist er nicht auf Repeat-Pass-Datenstapel, die sich durch eine große Gesamtbasislinie und eine hohe Zahl an Beobachtungen ausweisen, angewiesen, um eine ausreichende Elevationauflösung bereitzustellen.

Ergänzend dazu wird als dritter Beitrag die Fusion von Multi-Aspekt-InSAR-Daten untersucht. Dabei ist das Ziel Informationen dort aufzufüllen, wo sie in einzelnen Aspekten von Radarschatten verdeckt wurden. Zu diesem Zweck wird ein radargrammetrischer Registrierungsansatz, der als Vorprozessierungsschritt für weitere Fusionsoperationen verwendet wird, beschrieben. Auf dieser Voraussetzung aufbauend wird ein neues Maximum-Likelihood-Schätzverfahren entwickelt, das verwendet wird, um InSAR-Daten mehrerer Aspekte und Basislinien simultan zu fusionieren um ein flächendeckendes 2,5D-Höhenmodell zu generieren. Analog dazu wird eine Voxelraum-basierte Fusion von 3D-Punktwolken, welche durch SAR-Tomographie gewonnen wurden, vorgeschlagen.

Alle in der Arbeit beschriebenen Methoden werden mit Hilfe von experimentellen SAR-Daten des flugzeuggetragenen Millimeterwellen-Sensors MEMPHIS untersucht. Sie bestehen aus Single-Pass-Mehrfachbasislinien-InSAR-Stapeln, die vier koregistrierte Aufnahmen enthalten. Der Testdatensatz zeigt die Innenstadt von München und dient als Beispiel für komplexe urbane Szenen. Er ist zusammengesetzt aus dichten Gebäudeblocks, isolierten großen Gebäuden, Straßen und vielen Stadtbäumen.

Zuerst wird in den Experimenten sowohl die Effizienz als auch die Adaptivität der Kovarianzschätzungsmethoden mit Hilfe von Bildverarbeitungstechniken evaluiert. Anschließend wird die Fähigkeit, Überlagerungen aufzulösen, des TomoSAR-Algorithmus untersucht, bevor die Rekonstruktionsergebnisse, die mit der simultanen Fusion von Multi-Aspekt- und Mehrfachbasislinien-Daten erreicht werden, analysiert werden. Zuletzt werden diese Rekonstruktionsergebnisse mit den 3D-Daten, die durch eine Fusion von Multi-Aspekt-TomoSAR-Punktwolken erhalten werden, verglichen. Indem die beiden Rekonstruktionsergebnisse mit einer dichten Laser-Punktwolke abgeglichen werden, kann gezeigt werden, dass eine flächendeckende Rekonstruktion von 2,5D-Höhenkarten und 3D-Punktwolken mit Genauigkeiten im Meterbereich mit flugzeuggetragenem Single-Pass-InSAR-Daten möglich ist.

Contents

Abstract	3
Kurzfassung	5
Contents	7
List of Abbreviations	9
List of Figures	11
List of Tables	13
1 Introduction	15
1.1 Motivation and Objective of the Thesis	15
1.2 Structure of the Thesis	16
2 Synthetic Aperture Radar Interferometry	17
2.1 Synthetic Aperture Radar Basics	17
2.1.1 SAR Principle	17
2.1.2 Statistics of SAR Resolution Cells	18
2.2 Geometrical and Radiometrical Effects in SAR Imagery	21
2.3 SAR Interferometry	23
2.3.1 Types of SAR Interferometry	23
2.3.2 Basic InSAR Principle	24
2.3.3 Interferometric Coherence	25
2.3.4 Phase-to-Height Conversion	26
3 State of the Art in SAR Interferometry of Urban Areas	29
3.1 Interferometric SAR Image Analysis	29
3.2 Markovian Regularization Frameworks	30
3.3 Persistent Scatterer Interferometry	30
3.4 SAR Tomography and Layover Separation	31
3.5 Contribution of this Thesis	32
4 Adaptive Covariance Matrix Estimation	35
4.1 Statistical Properties of Multi-Baseline InSAR Resolution Cells	36
4.2 Principal Component Analysis-Based Thresholding	37
4.2.1 Principal Component Analysis of the InSAR Stack	38
4.2.2 Total Variation Denoising	39
4.2.3 Thresholding	41
4.3 Probabilistic Similarity Determination	42
4.3.1 Robust Estimation of Initial Covariance Matrix	43
4.3.2 Similarity and Connectivity Testing	44
4.3.3 Determination of Optimal Parameter Settings	45

5	Maximum-Likelihood SAR Tomography	47
5.1	Tomographic Imaging Model	48
5.2	Model Order Selection	49
5.3	Maximum-Likelihood-Based TomoSAR Inversion	51
5.3.1	Height Estimation	51
5.3.2	Reflectivity Estimation	53
6	Fusion of Multi-Aspect InSAR Datasets	57
6.1	Radargrammetric Multi-Aspect SAR Image Registration	57
6.1.1	Radargrammetric calibration in a strict Gauss-Helmert model	58
6.1.2	Recapitulation of range-Doppler geometry	59
6.1.3	Parameter estimation in the strict Gauss-Helmert model	59
6.2	Fusion Based on Interferometric Forward Geocoding	63
6.3	Fusion Based on Interferometric Backward Geocoding	66
6.4	Multi-Aspect Multi-Baseline SAR Interferometry	67
6.4.1	Estimation Procedure	67
6.4.2	Reliability Measure	69
7	Utilized Test System and Data	73
7.1	Sensor Description	73
7.2	Characteristics of Millimeterwave SAR	74
7.2.1	Propagation through the Atmosphere	74
7.2.2	Surface Roughness Properties	75
7.3	Measurement Campaign and Test Area	75
7.4	LiDAR Reference Data	78
8	Experimental Results	81
8.1	Efficiency of the Adaptive Covariance Matrix Estimators	81
8.1.1	Competing Approaches for Comparison	81
8.1.2	Theoretical Investigations Based on Simulated InSAR Data	82
8.1.3	Practical Results on Real Multi-Baseline InSAR Data	85
8.1.4	Discussion	89
8.2	Maximum-Likelihood TomoSAR Results	90
8.2.1	Experiments on Simulated Data	91
8.2.2	Experiments on Real Data	91
8.2.3	Discussion	93
8.3	Multi-Aspect Multi-Baseline Interferometry Results	96
8.3.1	Reconstruction Result	97
8.3.2	Discussion	100
8.4	Benefit of Multi-Aspect Data Fusion	103
8.4.1	Analysis of Different Aspect Combinations	103
8.4.2	Discussion	103
8.5	Comparison of 3D-Reconstruction Results	106
8.5.1	Reconstruction Results	106
8.5.2	Discussion	107
9	Conclusions and Perspectives	109
9.1	Summary and Conclusion	109
9.2	Future Work	110
	Bibliography	111
	Acknowledgment	123

List of Abbreviations

Abbreviation	Description	Page
A/D	analog-to-digital	91
AIC	Akaike information criterion	50
Az	azimuth	17
BW	bandwidth	17
dB	Decibel	91
DEM	digital elevation model	100
dGPS	differential GPS	76
DHM	digital height model	100
DSM	digital surface model	30
EDC	efficient detection criterion / criteria	50
GPS	Global Positioning System	76
ICM	iterated conditional modes	30
Im	imaginary part of a complex number	19
IMU	inertial measurement unit	76
INS	inertial navigation system	76
InSAR	interferometric synthetic aperture radar / SAR interferometry	15
ITC	information theoretic criterion / criteria	49
LGMRF	local Gaussian Markov random field	30
LiDAR	light detection and ranging	29
LS	least-squares	108
MAMBIInSAR	multi-aspect multi-baseline SAR interferometry	67
MASAR	multi-aspect SAR	78
MDL	minimum description length	50
ML	maximum-likelihood	53
MLE	maximum-likelihood estimation / estimator	53
MO	model order	92
MOS	model order selection	92
MUSIC	multiple signal classification	53
NMEA	National Marine Electronics Association	76
PC	principal component	36
PCA	principal component analysis	36
pdf	probability density function	19
PSI	persistent scatterer interferometry	30
PPS	pulse per second	76
R	range	17
Re	real part of a complex number	19
RMSE	root-mean-square error	84
Rx	receiving antenna	49
SAR	synthetic aperture radar	15
SLC	single look complex	24
SMPI	speckle suppression and mean preservation index	87
SNR	signal-to-noise ratio	47
SSI	speckle suppression index	86

Abbreviation	Description	Page
STD	standard deviation	46
TomoSAR	tomographic synthetic aperture radar / SAR tomography	15
TUM	Technische Universität München	75
TV	total variation	37
Tx	transmitting antenna	49
UTM	Universal Transverse Mercator	106
WGS84	World Geodetic System 1984	62
WSSF	weighted signal subspace fitting	91

List of Figures

2.1	Side-looking airborne SAR geometry	18
2.2	Coherent summation of independent subscatterers within a SAR resolution cell	19
2.3	Exemplary SAR probability density functions	20
2.4	Probability density function of interferometric phase	21
2.5	Geometrical effects in SAR imaging	22
2.6	Layover and shadowing in an urban environment	23
2.7	Types of SAR interferometry	24
2.8	Basic InSAR geometry	26
3.1	Sketch of possible multi-aspect multi-baseline InSAR acquisition geometries	33
4.1	Flowchart of the PCA-TV-based covariance matrix estimation method	38
4.2	Example of PCA on a stack of four SAR amplitude images	40
4.3	Relative share of eigenvalues corresponding to principal components	41
4.4	Thresholding examples for heterogeneous and homogeneous image patches	42
4.5	Size of homogeneous neighborhood as determined for each pixel	43
4.6	Sketch of the sliding window based probabilistic covariance matrix estimation procedure	44
4.7	Comparison of conventional and robust covariance matrix estimation	45
4.8	Optimal choice of parameters for the probabilistic covariance matrix estimation procedure	46
5.1	TomoSAR imaging geometry	49
5.2	Flowchart of the Maximum-Likelihood-based TomoSAR algorithm	52
5.3	Comparison of likelihood function peaks for covariance matrix and correlation matrix based maximum-likelihood estimation	53
5.4	Two-dimensional ML-TomoSAR objective function for two simulated scatterers	54
6.1	Exemplary illustrations of radargrammetric registration results	58
6.2	Structure of the Jacobian matrix employed in the Gauss-Helmert model	60
6.3	Significance of the individual flight parameters in the radargrammetric calibration	63
6.4	Voxelization of an irregular point cloud	65
6.5	Flowchart of multi-aspect multi-baseline SAR interferometry	68
6.6	Likelihood functions of the MAMBIInSAR estimator for different numbers of antennas	69
6.7	Slice through the probabilistic volume model created by MAMBIInSAR	70
7.1	The MEMPHIS sensor mounted on a C-160 Transall	74
7.2	Optical image of the test area	76
7.3	Amplitude images of the MASAR data acquired during the MEMPHIS campaign in 2011	77
7.4	Flight track configuration of the 2011 MEMPHIS campaign	78
7.5	Enhanced multi-aspect LiDAR point cloud used as reference dataset	79
8.1	Filtering efficiency comparison with respect to noise standard deviation	83
8.2	Filtering efficiency comparison with respect to growing stack size	83
8.3	Filtering efficiency of the adaptive covariance matrix estimation algorithms with respect to noise level	84

8.4	Filtering efficiency of the adaptive covariance matrix estimation algorithms with respect to stack size	84
8.5	Filtering results for a simulated InSAR pair	85
8.6	Quantitative comparison of filtering results for a simulated InSAR pair	85
8.7	Filtering results for a simulated 4-image stack	86
8.8	Quantitative comparison of filtering results for a simulated 4-image stack	86
8.9	Quantitative comparison of filtering results for a 4-image MEMPHIS stack	87
8.10	Filtering results for a 4-image MEMPHIS stack	88
8.11	Coherence maps	89
8.12	Mean and standard deviation of the coherence magnitudes	89
8.13	Scatterer height reconstruction results for simulated data	92
8.14	Test scene for the real data TomoSAR experiments	93
8.15	Model order map	93
8.16	Tomographic processing result	94
8.17	Comparison of WSSF and ML-TomoSAR	95
8.18	Histogram of scatterer height differences	96
8.19	Simulation result for a single scatterer and an overestimated model order	96
8.20	Result of the MAMBIInSAR reconstruction	98
8.21	Height errors of the MAMBIInSAR result	99
8.22	Detailed illustration: Reconstruction result for trees	101
8.23	Detailed illustration: Reconstruction result for metal-roofed building	102
8.24	Detailed illustration: Reconstruction result for concrete-roofed building	102
8.25	Height reconstruction results for different combinations of multi-aspect single-baseline data . .	104
8.26	Share of shadow- or ambiguity-affected grid elements	104
8.27	Oblique-view optical image of the test subse	106
8.28	Comparison of MAMBIInSAR and MA-TomoSAR reconstruction results	107
8.29	Example illustration of different point densities	108

List of Tables

6.1	Observation types in radargrammetric calibration	62
7.1	MEMPHIS sensor parameters	73
7.2	Definition of radar roughness categories	75
7.3	MEMPHIS MASAR campaign 2011 flight track configuration	77
7.4	Parameters of the LiDAR reference data	79
8.1	Accuracy assessment of the maximum-likelihood based MAMBIInSAR approach	99
8.2	Numerical comparison of MAMBIInSAR and MA-TomoSAR reconstruction results	108

1 Introduction

1.1 Motivation and Objective of the Thesis

During the last decades, synthetic aperture radar imaging has become an invaluable tool for the remote sensing of the Earth's surface. Its principle is based on the emission of electromagnetic waves to illuminate the scene of interest, measuring the echo of the backscatter and the following generation of a reflectivity map. Therefore, in comparison to optical and infrared sensors, a different kind of information is acquired: Whereas in the optical domain chemical characteristics cause the reflectivity of the object surfaces, in the microwave domain dielectric and geometrical properties are responsible for the observed backscattering. In addition to that, radar imaging shows some specific advantages with respect to conventional optical and infrared remote sensing systems: First of all, radar sensors are active sensors, providing their own illumination; they can therefore be well operated independent of any daylight. Since microwaves furthermore provide significantly longer wavelengths than optical light, radar imaging techniques can also be applied during adverse weather conditions, e.g. cloud coverage or even rain. Last, but not least, the large frequency diversity in the microwave domain, ranging from about 1 mm to 1 m, enables the measurement of quite different scene characteristics: In longer wavelengths (e.g. L- or P-band) most surfaces show a mirror-like appearance, and volume structures such as vegetation can be penetrated. In shorter wavelengths (e.g. Ka- or Ku-band), instead, even rather smooth surfaces appear rough and volume penetration is less likely.

Like for all imaging techniques, also in SAR remote sensing a mapping of the three-dimensional world to a two-dimensional image takes place. In order to recover the lost third dimension, the coherent SAR signals can be exploited interferometrically. SAR interferometry (InSAR) has been an established tool for the measurement of extended terrain topography for more than two decades. In combination with modern sensors that offer resolutions well in the sub-meter and even decimeter range, the detailed analysis of urban objects like streets, buildings, or even single trees has met growing interest recently. However, SAR remote sensing of urban areas is a difficult, non-trivial task: Due to the side-looking sensor principle, geometrical effects like shadowing and layover appear. Whereas shadowing leads to image areas without any information, layover leads to a superposition of different reflection components, such that several scattering contributions are mixed in one resolution cell.

These drawbacks can be tackled using advanced InSAR techniques based on multi-baseline and multi-aspect acquisition geometries. A special case of multi-baseline SAR interferometry, often also called SAR Tomography (TomoSAR), can be seen as an extension of the conventional InSAR technique, allowing the reconstruction of the three-dimensional scattering distribution within a resolution cell by establishing a second synthetic aperture in elevation direction. In urban areas, this can gainfully be exploited to solve the layover problem. In addition to that, the fusion of multi-aspect InSAR data can help to fill scene parts with information that are occluded in the single aspects due to radar shadowing. Furthermore, redundant measurements can be exploited to improve the overall accuracy.

The topic of this thesis is the *reconstruction of urban surface models from multi-aspect and multi-baseline interferometric SAR* using data acquired with an airborne single-pass sensor. For that purpose, new approaches for the separation of layover contributions by tomographic SAR inversion and the simultaneous fusion of multi-aspect multi-baseline InSAR data in a maximum-likelihood estimation framework are developed. Since both methods rely on a knowledge of the complex covariance matrix for every pixel in the image stack, furthermore two novel procedures for adaptive covariance matrix estimation are proposed.

1.2 Structure of the Thesis

The structure of this thesis can be summarized as follows:

An introduction to the basics of synthetic aperture radar interferometry is given in Chapter 2. First, the synthetic aperture radar principle is outlined. Subsequently, geometrical and radiometrical effects occurring in SAR imaging are explained. Finally, the concept of SAR interferometry is described.

The state-of-the-art of interferometric SAR remote sensing over urban areas is discussed in Chapter 3. From that, the objectives of this thesis are developed.

Chapter 4 then presents techniques for adaptive covariance matrix estimation, a necessary pre-requisite for the exploitation of single-pass InSAR data stacks by the methods developed in this thesis: A maximum-likelihood-based approach for SAR tomography intending to separate layover contributions in urban areas is developed in Chapter 5. Subsequently, Chapter 6 describes ways to fuse InSAR results from multiple aspects, finally leading to an estimation framework for simultaneous fusion of multi-aspect multi-baseline InSAR data proposed in Section 6.4.

In Chapter 7 the utilized Ka-band test system is introduced and the peculiarities of millimeter-wave SAR are discussed.

Chapter 8 contains the experimental results and corresponding discussions for determination of the efficiency of the adaptive covariance matrix estimators (Section 8.1), as well as the applicability of maximum-likelihood SAR tomography (Section 8.2) and multi-aspect multi-baseline SAR interferometry (Section 8.3). In addition, experiments aiming at an assessment of the benefit introduced by multi-aspect data fusion (Section 8.4) and a comparison of the two 3D reconstruction approaches (Section 8.5) are shown.

Chapter 9 finally concludes the thesis and gives perspectives for future research directions.

2 Synthetic Aperture Radar Interferometry

In this chapter, the fundamentals of synthetic aperture radar interferometry are described. Since this technology has been intensively studied and well described in many textbooks during recent decades, only the background necessary for understanding the contents of this thesis are touched. For in-depth discussion, see e.g. [Curlander & McDonough, 1991; Jakowatz et al., 1996; Franceschetti & Lanari, 1999; Hanssen, 2001; Hein, 2004; Oliver & Quegan, 2004; Cumming & Wong, 2005; Massonet & Souyris, 2008; Richards, 2009].

The chapter starts with an explanation of the very basics of the synthetic aperture radar technology, from the process of image formation and the synthetic aperture principle to the critically important statistical attributes of SAR measurements. In the second section, the well-known geometrical and radiometrical effects occurring during SAR imaging are described, before finally the concept of SAR interferometry is developed.

2.1 Synthetic Aperture Radar Basics

2.1.1 SAR Principle

In this first section the basic principle of SAR image acquisition and formation are shortly introduced. For a more detailed discussion, the reader is referred to one of the many textbooks on the topic, e.g. [Curlander & McDonough, 1991] or [Cumming & Wong, 2005].

Radar (**R**adio **D**etection **A**nd **R**anging) is an active sensor technology, which acquires measurement data by illuminating the scene of interest with electromagnetic signals and receiving the backscattered waves. Theoretically, any frequency may be used, but frequencies in the microwave domain, especially X- (2.4-3.8 cm), C- (3.8-7.5 cm) or L-band (15-30 cm), are most common.

Imaging radar systems enable a two-dimensional imaging of the Earth's surface. The basic configuration of a typical side-looking airborne radar (SLAR) can be seen in Fig. 2.1. The antenna axis is usually orthogonal to the velocity vector of the aircraft, creating a two-dimensional image coordinate system defined by *azimuth* (Az), corresponding to the flight direction, and *range* (R), corresponding to the distance from the sensor. The range resolution is a function of the bandwidth BW of the emitted pulses:

$$\rho_R = \frac{c}{2 \cdot BW}, \quad (2.1)$$

where c is the speed of light. Therefore, ρ_R is theoretically independent of the distance to the target. The azimuth resolution is limited by the length of the physical antenna and gets worse with increasing distance. This is overcome by the synthetic aperture principle, which was introduced by Carl A. Wiley of Goodyear Aircraft Company in 1951 [Wiley, 1954]. SAR systems employ a comparably short physical antenna with a broad beamwidth. During the forward movement of the

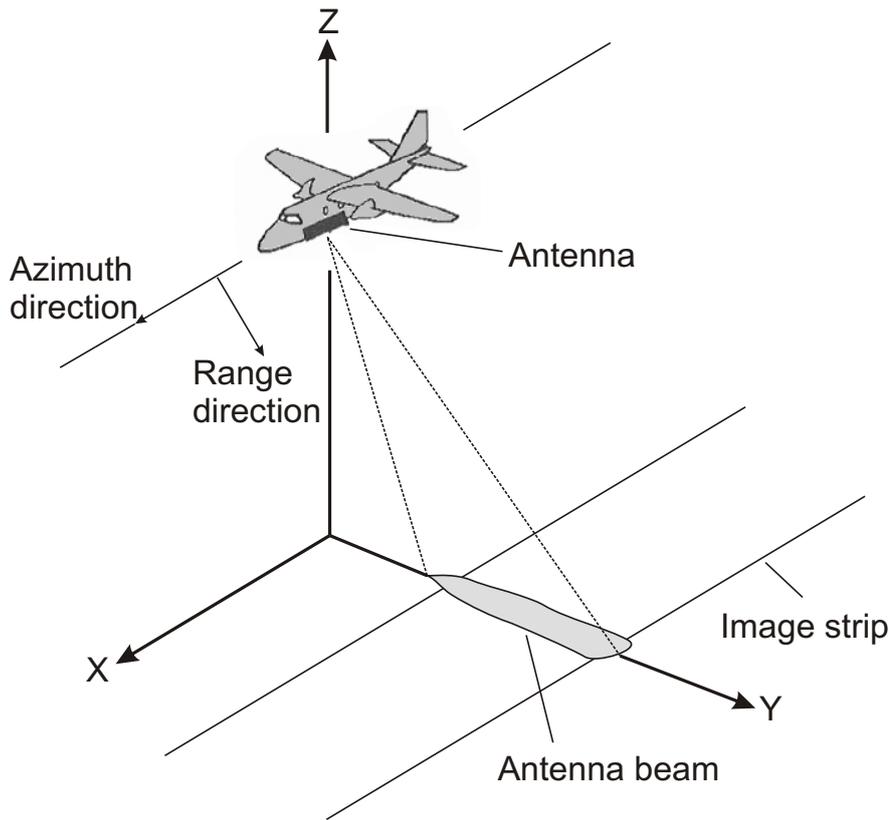


Figure 2.1: Side-looking airborne SAR geometry. The aircraft flies in X-direction and images a strip to the side of the flight trajectory.

carrier platform and coherent processing of the backscattered signals a long, synthetic aperture is achieved. The length of the synthetic aperture depends on the total illumination time of the target. Since the illumination time increases with increasing distance from the target, also the azimuth resolution of SAR sensors is independent from the distance and just related to the length of the physical antenna d_a :

$$\rho_{Az} = \frac{d_a}{2}. \quad (2.2)$$

Detailed information on the signal processing techniques applied during SAR image formation are addressed in [Cumming & Wong, 2005]. For the scope of this thesis, it is only important to mention that the resulting imagery consists of complex-valued resolution cells (or pixels), each describing the amplitude and phase of the focussed radar signal. Whereas the amplitude represents the reflectivity of the scene and is related to roughness, dielectric properties and the geometrical appearance of the surface, the phase carries information about the distance of the target to the sensor. In Section 2.3 we will see that therefore the phase is the key observation for measurement of topography.

2.1.2 Statistics of SAR Resolution Cells

The statistics of SAR images can be explained by looking at the two extreme cases of scattering objects [Bamler & Hartl, 1998]: point scatterers and Gaussian scatterers. While the response of a point scatterer can be seen as a deterministic signal, Gaussian scatterers are the result of a sufficiently high number of random subscatterers within a resolution cell, whose individual responses are not known and cannot be reconstructed from the data. The resulting amplitude

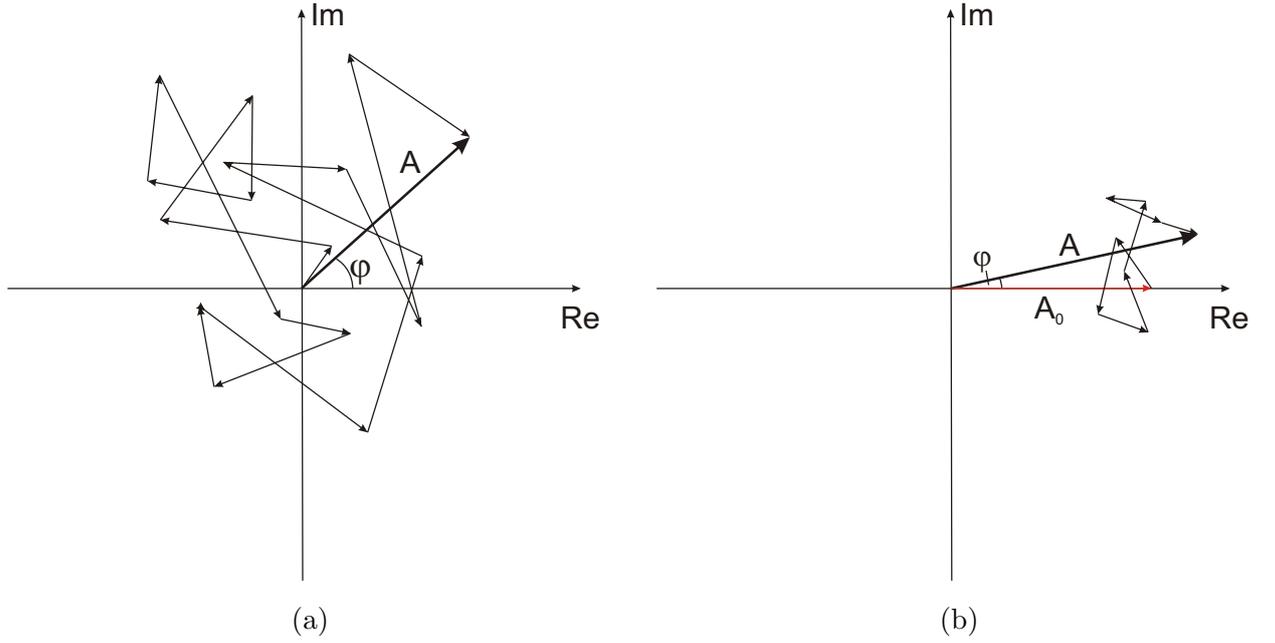


Figure 2.2: (a) Coherent summation of independent subscatterers within one SAR resolution cell. This phenomenon is often referred to as *random walk*. (b) Random walk process considering a dominant scatterer (displayed in red color) in the resolution cell.

and phase of the backscattered signal then results from coherent summation of the subscatterer signals (cf. Fig. 2.2).

If the number of subscatterers in the resolution cell is large enough, the central limit theorem applies, and the SAR image pixel value z can be considered a complex circular Gaussian random variable. In general, this assumption is true for low and medium resolution SAR and rural scenes containing mostly natural scatterers such as soil, rock, vegetation, or agriculture. In contrast to that, the assumption is often violated for high and very high resolution SAR data showing urban scenes, where artificial objects and man-made structures cause only few dominant scatterers to be present in each resolution cell, such that either deterministic point scattering or a Rician distribution has to be applied. The Gaussian assumption can, however, be considered a convenient approximation in most cases, leading to the probability density function (pdf)

$$f(z) = f(\operatorname{Re}\{z\}, \operatorname{Im}\{z\}) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\operatorname{Re}\{z\}^2 + \operatorname{Im}\{z\}^2}{2\sigma^2}\right), \quad (2.3)$$

where $\sigma^2/2 = \sigma_{\operatorname{Re}\{z\}}^2/2 = \sigma_{\operatorname{Im}\{z\}}^2/2$ is the variance of the Gaussian distribution considering that real and imaginary parts are uncorrelated.

Exploiting the relations

$$\begin{aligned} \operatorname{Re}\{z\} &= A \cos \varphi \\ \operatorname{Im}\{z\} &= A \sin \varphi, \end{aligned} \quad (2.4)$$

which lead to

$$\begin{aligned} A &= \sqrt{\operatorname{Re}\{z\}^2 + \operatorname{Im}\{z\}^2} \\ \varphi &= \arctan \frac{\operatorname{Im}\{z\}}{\operatorname{Re}\{z\}}, \end{aligned} \quad (2.5)$$

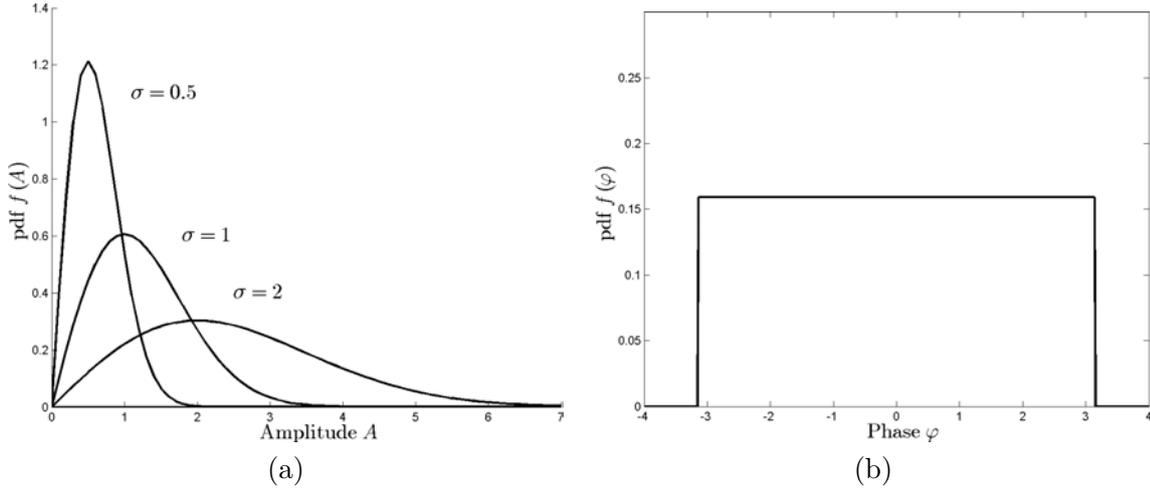


Figure 2.3: Exemplary SAR probability density functions: (a) Rayleigh distributed pdfs of amplitudes for varying σ . (b) Uniformly distributed phase pdf.

the probability density function for amplitude and phase $f(A, \varphi)$ can be deduced. For this, the usual techniques for transformations of random variables have to be employed:

$$f(A, \varphi) = f(\operatorname{Re}\{z\}, \operatorname{Im}\{z\}) \left| \frac{\partial(\operatorname{Re}\{z\}, \operatorname{Im}\{z\})}{\partial(A, \varphi)} \right| = \frac{A}{2\pi\sigma^2} \exp\left(-\frac{A^2}{2\sigma^2}\right), \quad (2.6)$$

where $\left| \frac{\partial(\operatorname{Re}\{z\}, \operatorname{Im}\{z\})}{\partial(A, \varphi)} \right|$ is the Jacobian of the transformation (2.4). The marginal probability density function of the amplitude alone can then be derived as

$$f(A) = \int_{-\pi}^{\pi} f(A, \varphi) d\varphi = \frac{A}{\sigma^2} \exp\left(-\frac{A^2}{2\sigma^2}\right), \quad (2.7)$$

while the marginal pdf of the phase is found to be

$$f(\varphi) = \int_0^{\infty} f(A, \varphi) dA = \frac{1}{2\pi}. \quad (2.8)$$

Both probability density functions are plotted exemplarily in Fig. 2.3. Equation (2.8) shows that the phase of the backscattered signal of a single SAR resolution cell is uniformly distributed, i.e. it does not contain any useful information. The amplitude pdf of the signal instead follows a Rayleigh distribution, which characterizes the so-called speckle effect, which is well known from the context of coherent imaging of rough surfaces [Goodman, 1975]. Although speckle is often referred to as noise, this is rather misleading: the speckle pattern of the imaged object contains information about its subresolution structure and can therefore be exploited beneficially.

In contrast to the uniformly distributed phase of a single resolution cell, which does not contain relevant information, the information content of the phase difference of two SAR signals depends on the degree of correlation between each other. Lee et al. [1994] and Bamler & Hartl [1998] derived the relation between this phase difference and the coherence between the two signals as

$$f(\phi) = \frac{1 - |\gamma|^2}{2\pi \left(1 - |\gamma|^2 \cos^2(\phi - \phi_0)\right)} \left[1 + \frac{|\gamma| \cos(\phi - \phi_0) \arccos(-|\gamma| \cos(\phi - \phi_0))}{\sqrt{1 - |\gamma|^2 \cos^2(\phi - \phi_0)}} \right], \quad (2.9)$$

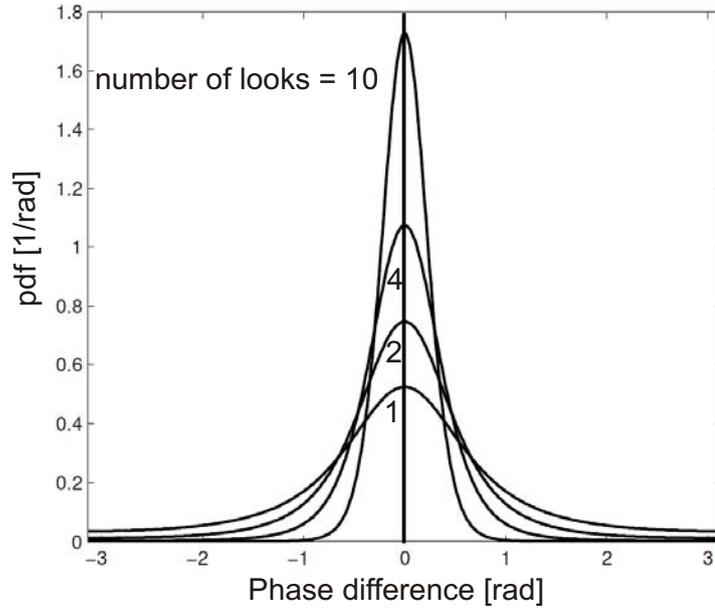


Figure 2.4: Probability density functions of the interferometric phase for $l = 1, 2, 4$ and 10 looks and a coherence magnitude of $|\gamma| = 0.7$ (after [Meyer, 2004]).

where ϕ_0 denotes the expectation of the phase and γ the complex coherence, which will be explained in Section 2.3.3 in greater detail. It has to be noted that (2.9) was derived only for single-look, i.e. unfiltered data. In many cases, it is, however, desirable to denoise the interferometric phase by *multilooking*. This procedure of course changes the probability density function of ϕ , which then becomes

$$f(\phi) = \frac{\Gamma(l + \frac{1}{2}) (1 - |\gamma|^2)^l \beta}{2\sqrt{\pi}\Gamma(l) (1 - \beta^2)^{l + \frac{1}{2}}} + \frac{(1 - |\gamma|^2)^l}{2\pi} \cdot F\left(l; 1; \frac{1}{2}; \beta^2\right), \quad (2.10)$$

with $\beta = |\gamma| \cos(\phi - \phi_0)$, the Gamma function

$$\Gamma(a) = \int_0^{\infty} t^{a-1} \exp(-t) dt \quad (2.11)$$

and the Hypergeometric function

$$F\left(l; 1; \frac{1}{2}; \beta^2\right) = \frac{\Gamma(\frac{1}{2})}{\Gamma(l)\Gamma(1)} \sum_{i=0}^{\infty} \frac{\Gamma(l+i)\Gamma(1+i)}{\Gamma(\frac{1}{2}+i) i!} (\beta^2)^i. \quad (2.12)$$

The probability density functions for interferometric phases with $l = 1, 2, 4$ and 10 looks and a coherence magnitude of $|\gamma| = 0.7$ are exemplarily shown in Fig. 2.4. It is obvious that the number of looks corresponds to an improvement of the phase accuracy.

2.2 Geometrical and Radiometrical Effects in SAR Imagery

In synthetic aperture radar imaging, the 3D objects of the scene are mapped to the two-dimensional azimuth-range image plane. Due to the side-looking imaging geometry and the fact that radar is based on measuring the distances to the real-world objects, certain geometrical effects occur if

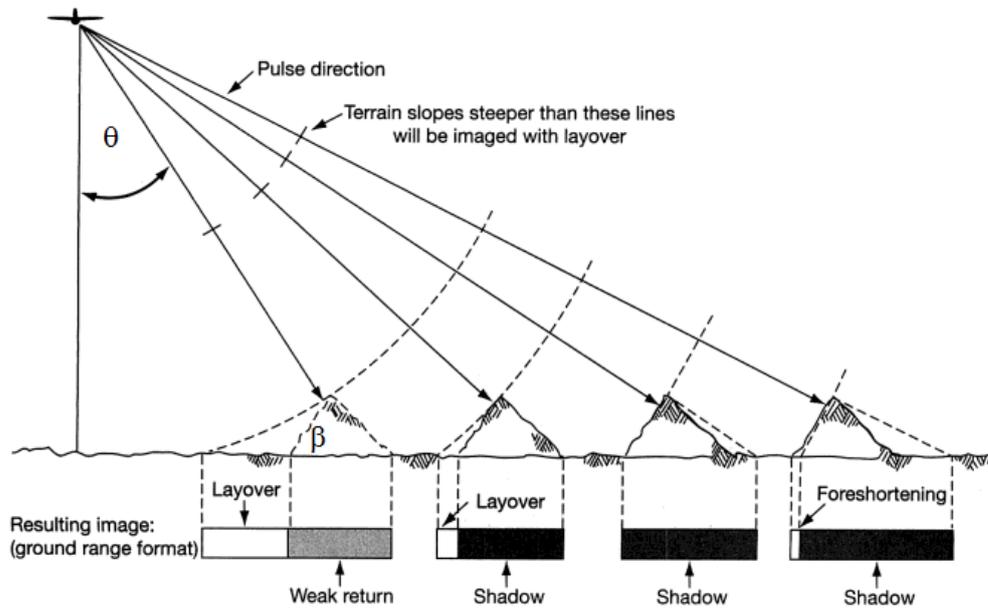


Figure 2.5: Geometrical effects in SAR imaging. After [Lillesand et al., 2004, p. 655].

elevated objects (e.g. mountains, trees or buildings) are illuminated: foreshortening, layover and shadowing are illustrated in Fig. 2.5 and can be explained if the impact of the surface slope angle β is investigated.

Foreshortening occurs for surfaces that are oriented towards the sensor and show a terrain slope smaller than the radar look angle θ , i.e. $-\theta < \beta < \theta$. In short, it means that the relative distance between two points is shortened due to its projection onto the slant range plane. Image parts affected by foreshortening therefore appear relatively bright in the SAR intensity image, because the backscattering energy of a surface patch is concentrated in a proportionally smaller image patch. If the terrain slope is even steeper than the look angle ($\beta > \theta$), the layover effect appears, i.e. several objects are mapped on the same location in the SAR image and their backscattering signals mix. In Fig. 2.6, for example, the first building's facade is overlaid with the ground in front of the building, and parts of the roof even mix with both facade and ground. Therefore, the positions of elevated objects are inverted and seem to be shifted towards the sensor in SAR imagery. The resulting image parts also appear bright in the intensity image due to the accumulation of several backscattering contributions. For $\beta < \theta - 90^\circ$, radar shadowing occurs. Here, the sensor-scene configuration leads to surfaces invisible to the sensor, such that no data is available for the corresponding areas of the scene. With the exception of thermal noise influences, radar shadow appears black in the SAR intensity images. All three effects depend on the look angle of the radar signal with respect to the geometry of the objects in the scene.

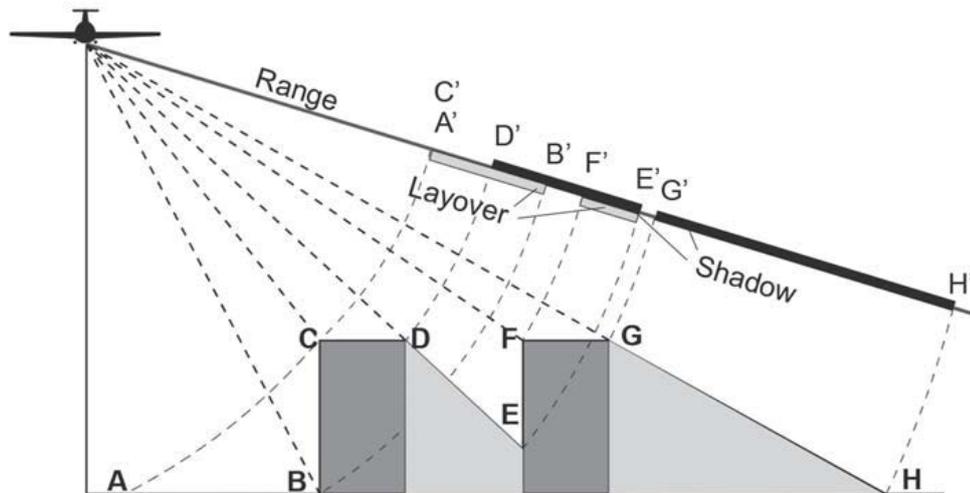


Figure 2.6: Layover and shadowing in an urban environment. The response of the first building mixes with the response of the ground in front, whereas the response of the second building overlays the shadow of the first one.

2.3 SAR Interferometry

2.3.1 Types of SAR Interferometry

Since synthetic aperture radar interferometry has been introduced in the 1970s [Graham, 1974], it has continuously attracted the attention of an interdisciplinary research community. In general, SAR interferometry can be employed in two different modes (cf. Fig. 2.7 and [Klausing & Holpp, 2000]):

- **Across-track interferometry**
The antennas are spatially aligned such that a baseline component in cross-track direction is introduced. This technique is used for derivation of terrain height.
- **Along-track interferometry**
The antennas are arranged along-track, i.e. the baseline is parallel to the flight direction. Along-track interferometry is usually employed for the detection of slow target movements (e.g. ocean currents or glacier movements) with respect to the radar line-of-sight.

Especially the derivation of the topography of extended areas by cross-track interferometry has received great attention and been used operational for quite some years now [Bamler & Hartl, 1998; Rosen et al., 2000; Hanssen, 2001; Richards, 2007].

Besides the classification in across-track and along-track methods, SAR interferometry can also be distinguished with respect to the number of antennas on the carrier platform (cf. Richards [2009]):

- **Single-pass interferometry**
Two (or more) antennas are mounted on the same carrier platform. An interferometric acquisition therefore needs just a single pass over the scene.
- **Repeat-pass interferometry**
Just one antenna is mounted on the carrier platform. For an interferometric analysis multiple passes over the scene are required.

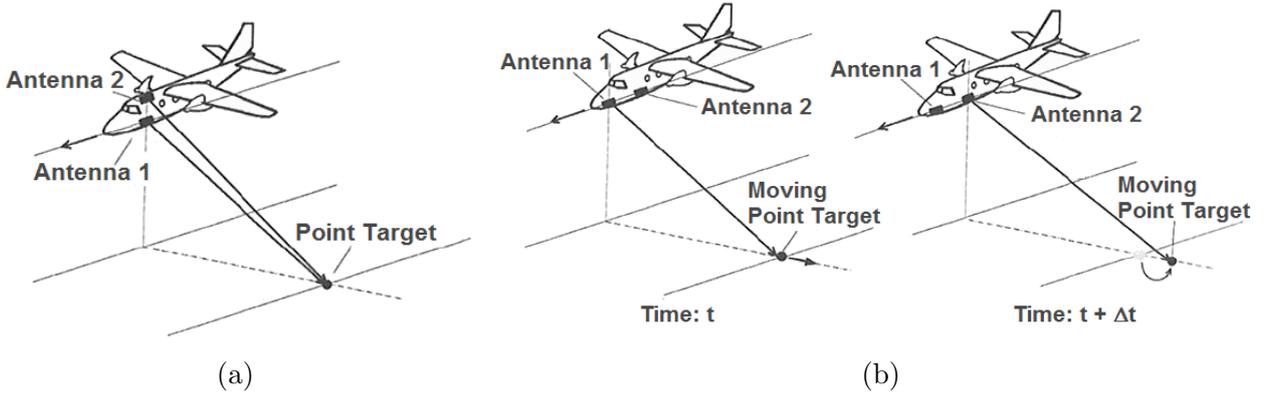


Figure 2.7: Types of SAR interferometry (after [Klausing & Holpp, 2000]): (a) Across-track InSAR geometry. (b) Along-track InSAR geometry.

Besides the additional time that is needed in order to acquire a repeat-pass InSAR dataset, another disadvantage is caused by the fact that repeat-pass data are affected by temporal decorrelation (see Section 2.3.3 for more detailed explanation). On the other hand, repeat-pass data can be used for determination of movements in addition to topography reconstruction.

2.3.2 Basic InSAR Principle

In general, SAR interferometry is based on the main characteristic that sets radar aside from optical imaging technologies: In radar imaging, instead of just intensity, both the amplitude and the phase of the signal backscattered from the landscape are received and stored in the form of a complex number for each resolution cell:

$$z = A \exp(j\varphi), \quad (2.13)$$

where A is the amplitude of the signal and

$$\varphi = -\frac{2\pi}{\lambda} (R_{fw} + R_{bw}) + \varphi_{scatt} \quad (2.14)$$

the signal phase resulting from the two-way distance between the sensor and the target. λ denotes the wavelength of the radar signal, R_{fw} and R_{bw} the radial distances between the emitting horn and the target, and the target and the receiving antenna, respectively. φ_{scatt} denotes an offset mainly caused by the physical properties of the target, which therefore is individual for each backscatterer. The offset can be eliminated by calculation of the phase difference (or interferometric phase) of two SAR signals. This is usually established on a pixel by pixel basis employing two precisely coregistered single look complex (SLC) SAR images s_1 and s_2 . The interferogram is calculated by

$$s_1 s_2^* = A_1 A_2 \exp\left(-j \frac{2\pi}{\lambda} (\varphi_1 - \varphi_2)\right). \quad (2.15)$$

From (2.15) it can be seen that under the pre-condition $\varphi_{scatt,1} = \varphi_{scatt,2}$ and the utilization of the same emitting horn for both images leading to $R_{fw,1} = R_{fw,2}$, the interferometric phase is just related to the range difference of the two antennas:

$$\phi = \Delta\varphi = \varphi_1 - \varphi_2 = -\frac{2\pi}{\lambda} (R_{bw,1} - R_{bw,2}) = -\frac{2\pi}{\lambda} \Delta R. \quad (2.16)$$

Unfortunately, by definition, the phase values are restricted to the interval $[-\pi, \pi[$, corresponding to one so-called fringe. Therefore, the interferometric phase is ambiguous and usually

has to be unwrapped prior to terrain reconstruction. The relationship between the wrapped and the unwrapped, absolute phase is given as follows:

$$\phi_{abs} = \phi + 2\pi k. \quad (2.17)$$

Solving for the integer value k generally is a non-trivial task and has thoroughly been studied in the literature for decades [Goldstein et al., 1988; Ghiglia & Romero, 1994; Fornaro et al., 1996a,b; Flynn, 1997; Costantini, 1998; Ghiglia & Pritt, 1998; Zebker & Lu, 1998; Davidson & Bamler, 1999; Xu & Cumming, 1999; Gens, 2003].

2.3.3 Interferometric Coherence

One of the key parameters in SAR interferometry is the interferometric coherence, whose magnitude is providing a valuable measure for the quality of interferometric phase measurements. As described in Section 2.3.2, the interferometric exploitability of two SAR acquisitions is based on the assumption that the scatterer phases $\varphi_{scatt,i}$ are identical in both images, thus yielding a perfectly defined phase difference. In reality, however, several phenomena disturb this assumption and lead to different kinds of decorrelation:

- Change of the backscattering properties of the scene between the acquisition times. This is especially critical for vegetation or moving surfaces such as glaciers, and for repeat-pass systems.
- Differences in the information content of the images caused by different viewing angles or different sensors.
- Thermal noise of the radar system hardware.
- Errors introduced during processing.

The degree of correlation, or *coherence*, between two co-registered images of a SAR interferometer is measured by

$$\gamma = \frac{E\{z_1 z_2^*\}}{\sqrt{E\{|z_1|^2\} \cdot E\{|z_2|^2\}}} = |\gamma| \cdot \exp(j\phi), \quad (2.18)$$

where $E\{\cdot\}$ denotes the expectation of the complex SAR signal. In practice, it is substituted by the spatial average over L adjacent looks assuming local stationarity. The coherence magnitude $|\gamma| \in [0; 1]$ resembles the correlation coefficient between the images, i.e. 0 indicates complete decorrelation, while 1 means full coherence. Since the phase standard deviation is related to the coherence magnitude by the Cramer Rao bound [Rodriguez & Martin, 1992]

$$\sigma_\varphi = \frac{1}{\sqrt{2L}} \frac{\sqrt{1 - |\gamma|^2}}{|\gamma|}, \quad (2.19)$$

where L is the number of looks utilized for coherence (and therefore also phase) estimation, it is obvious that a high coherence is a necessary prerequisite for SAR interferometry.

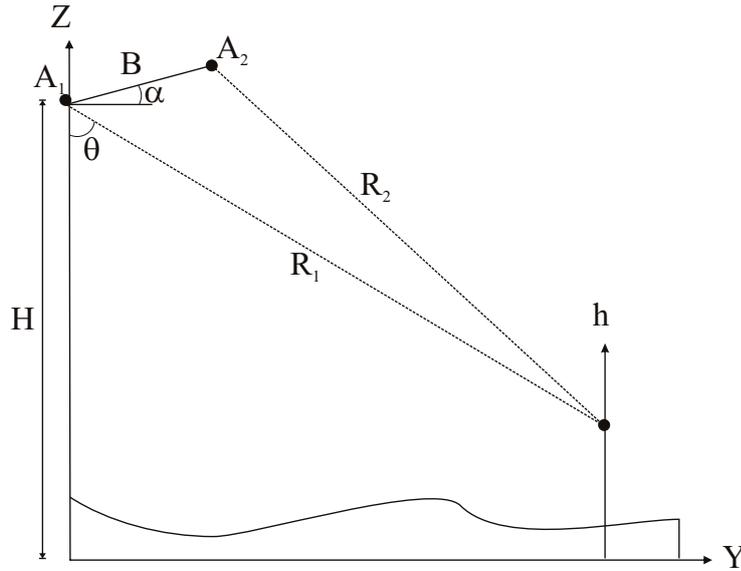


Figure 2.8: Basic interferometric SAR geometry.

2.3.4 Phase-to-Height Conversion

The geometry of a cross-track SAR interferometer is shown in Fig. 2.8. The two radar antennas are separated by the baseline B , and antenna 1 be both emitting and receiving, antenna 2 be only receiving. In principle, the imaging geometry of cross-track SAR interferometry is similar to stereogrammetric techniques as known from photogrammetry. The main difference is that in stereoscopic systems the parallax is determined by a direct measurement of the observation angles, whereas a SAR interferometer exploits the phase difference between the received signals of both antennas. As we have seen from (2.16), the phase difference ϕ is corresponding to the range distance of the antennas ΔR . Therefore, the interferometric phase is a very precise measure of the range difference between the antennas with an accuracy down to the fraction of a wavelength. In analogy to stereogrammetry, this equals an enhanced angular resolution. Applying the cosine law to the triangle defined by the antennas and the measured object, the distance R_2 of the second antenna can be written as

$$R_2^2 = R_1^2 + B^2 + 2R_1B \cos(90^\circ - \theta + \alpha). \quad (2.20)$$

Considering that $R_2 = R_1 + \Delta R$ and $\cos(90^\circ - \theta + \alpha) = \sin(\theta - \alpha)$ we receive

$$\Delta R = \sqrt{R_1^2 + B^2 + 2R_1B \sin(\theta - \alpha)} - R_1 \quad (2.21)$$

for the range difference. Combining (2.21) with (2.16), we see that the interferometric phase at a certain range distance $R = R_1$ and for a known baseline described by its length B and its inclination α only depends on the sensor look angle θ :

$$\phi = -\frac{2\pi}{\lambda} \left(\sqrt{R^2 + B^2 + 2RB \sin(\theta - \alpha)} - R \right). \quad (2.22)$$

If the fact that the look angle θ is range and height dependent, i.e.

$$\theta = \arccos\left(\frac{H - h}{R}\right), \quad (2.23)$$

is considered, we can deduce the direct relationship between the topographic height h and the measured interferometric phase

$$\phi = -\frac{2\pi}{\lambda} \left(\sqrt{R^2 + B^2 + 2RB \sin \left(\arccos \left(\frac{H-h}{R} \right) - \alpha \right)} - R \right), \quad (2.24)$$

which can be solved for h employing non-linear optimization techniques. Usually, however, one is not interested in the full interferometric phase, but only its topographically induced share. For this, we have to reformulate (2.21) as

$$B \sin(\theta - \alpha) = \Delta R + \frac{\Delta R^2}{2R} - \frac{B^2}{2R}. \quad (2.25)$$

Since the baseline B and the range difference ΔR are usually small in comparison to the range distance R , the two last terms in (2.25) can be neglected, such that (2.16) becomes

$$\phi = -\frac{2\pi}{\lambda} \Delta R \approx -\frac{2\pi}{\lambda} B \sin(\theta - \alpha) \quad (2.26)$$

for a scatterer at reference height. If then another scatterer with the same range distance but with a topographic height difference Δh to the first scatterer is measured, its interferometric phase becomes

$$\phi(\Delta h) = -\frac{2\pi}{\lambda} B \sin(\theta + \Delta\theta(\Delta h) - \alpha) \quad (2.27)$$

because of the slightly changing look angle $\theta + \Delta\theta$. Exploiting the approximation $\Delta\theta(\Delta h) \approx \frac{\Delta h}{R \sin(\theta)}$, the observed phase difference between both scatterers can be expressed as

$$\begin{aligned} \Delta\phi &= -\frac{2\pi}{\lambda} B (\sin(\theta + \Delta\theta(\Delta h) - \alpha) - \sin(\theta - \alpha)) \\ &\approx -\frac{2\pi}{\lambda} B \cos(\theta - \alpha) \Delta\theta(\Delta h) \\ &\approx -\frac{2\pi}{\lambda} \frac{B \cos(\theta - \alpha)}{R \sin(\theta)} \Delta h, \end{aligned} \quad (2.28)$$

with $B_{\perp} = B \cos(\theta - \alpha)$ being the orthogonal or normal baseline. Often,

$$k_z = \frac{2\pi}{\lambda} \frac{B_{\perp}}{R \sin \theta} \quad (2.29)$$

is referred to as vertical wavenumber, which can be used to introduce a linear relation between phase and height:

$$\phi = -k_z \cdot h. \quad (2.30)$$

Note that application of (2.30) requires that the reference height h_0 and the corresponding phase $\phi(h_0)$ have to be calibrated to 0 first.

As mentioned before, the interferometric phase is only known modulo 2π , i.e. in the interval $[-\pi; +\pi[$. Therefore, the so-called height of ambiguity

$$h_{2\pi} = \frac{\lambda R \sin(\theta)}{B \cos(\theta - \alpha)}. \quad (2.31)$$

can be deduced by inversion of (2.30). It describes the height leading to a phase change of 2π . Before applying any of the phase-to-height conversion formulas described in this section, it is necessary to reconstruct the absolute phase from the observed wrapped phase. Although a vast number of phase unwrapping algorithms have been published during the last decades, no fully satisfactory and reliable solution has been found yet. In general, the ability to resolve phase ambiguities mainly depends on the local terrain slope and the phase noise level, making it a particularly challenging task for urban areas with many strong height jumps.

3 State of the Art in SAR Interferometry of Urban Areas

The analysis of urban areas by SAR interferometry belongs to the great challenges in remote sensing and has been studied extensively since the first sensors providing sufficient resolution have been introduced. Following the developments in this field of research and in spite of potential overlaps, most of the hitherto published approaches can be classified in four main categories: interferometric SAR image analysis, Markovian regularization frameworks, persistent scatterer interferometry, and SAR tomography. This section gives a short survey over the state-of-the-art in SAR interferometry of urban areas and elaborates the contribution of this thesis with respect to this background. More comprehensive information on radar remote sensing of urban areas, not limited to SAR interferometry, can be found in [Soergel, 2010].

3.1 Interferometric SAR Image Analysis

Many of the early approaches in urban SAR remote sensing exploit the signatures of buildings in the InSAR data by methods from the field of image analysis. Instead of analyzing the data on a pixel-by-pixel basis or by reconstructing the interferometric range-Doppler imaging geometry in a comprehensive way, geometric approximations are utilized. Consequently, both the shadowing and the layover effect are either exploited within geometrical relationships or considered to be disturbances, which have to be coped with by the proposed building reconstruction algorithms.

Some of the first considerations towards building extraction from InSAR data were presented in [Burkhart et al., 1996]; they were largely based on the hypothesis that gradients in an interferometrically derived elevation map directly correspond to building edges. The main focus of the paper was put on the filtering of elevation maps, and an exploitation of the so-called “front porch” effect caused by building layover.

[Gamba et al., 2000] provides an extension to this by proposing a complete procedure for the extraction of building structures from InSAR elevation data segmented by a region-growing method and following local approximation of the 3D data by means of best-fitting planes. In [Gamba & Houshmand, 2000], the authors state that LiDAR-derived height data is better suitable for the shape characterization of buildings, because layover and shadowing can only partially be corrected by means of segmentation procedures.

More segmentation-based approaches were proposed by Soergel et al. [2000, 2001] and Stilla et al. [2003], with the extension that now the elevation data were combined with the corresponding intensity and coherence maps in order to improve the segmentation result. Apart from that, again SAR peculiarities caused by the side-looking imaging geometry were discussed.

Although the work published up until then already showed the general potential of airborne high-resolution InSAR data for urban area reconstruction, the missing information caused by

radar shadowing still remained an open problem. Therefore, in parallel, the first solutions utilizing InSAR data acquired from multiple viewing directions were suggested [Bolter & Leberl, 2000; Bolter, 2001]. The idea was readily seized on and fused with the segmentation-based approaches [Soergel et al., 2003]. Thiele et al. [2007a, 2010a,b] further investigated in this direction by considering the signatures of flat- and gable-roofed buildings in multi-aspect InSAR data. Their approach is centered around the extraction of geometrical primitives such as double bounce lines in slant range amplitude data and an analysis of phase profiles in building layover areas [Thiele et al., 2007b, 2008]. After geocoding, the primitives are grouped and expanded to building footprints. Then, in combination with the InSAR-derived heights, 3D hypotheses are generated and compared to simulated phase profiles in order to determine the correct building model.

3.2 Markovian Regularization Frameworks

Whereas the approaches described in the last section were based on an exploitation of geometrical approximations, building signatures in InSAR data and image analysis techniques, another group of publications utilizes statistical regularization frameworks in order to extract comprehensive surface models rather than building models from single- or multi-baseline InSAR data. Ferraiuolo et al. [2004], for example, proposed a Gaussian Markov random field for recovering topographic profiles affected by strong height discontinuities. In a similar manner, Denis et al. [2009] presented a Markov random field in combination with total variation regularization for a joint denoising of phase and amplitude data, aiming at the 3D reconstruction of urban areas. The approach proposed in [Tison et al., 2007] provides a framework for simultaneous reconstruction of height model and classification map, again based on a Markovian optimization. Starting from previously extracted high-level features, a digital surface model and the object classes are estimated and afterwards improved by comparison with a layover/shadow map derived from the estimated DSM.

Ferraioli [2008] was among the first to extend the idea of the hitherto published global regularization approaches to multi-channel (i.e. multi-baseline or multi-frequency) InSAR data. Starting from the need to solve the phase-unwrapping problem, which has already been tackled by multi-baseline approaches before [e.g. Corsini et al., 1999; Essen et al., 2007], a local Gaussian Markov random field (LGMRF) is developed and optimized using the iterated conditional modes (ICM) algorithm. The method also allows the integration of auxiliary input data such as optical imagery [Baselice et al., 2009b], and provides interesting perspectives for edge detection [Baselice & Ferraioli, 2012].

The latest extension to this group of approaches was proposed by Shabou et al. [2012]. The algorithm is designed to simultaneously unwrap and regularize the complex InSAR data. In particular, the exploitation of amplitude data within the unwrapping process helps to preserve sharp discontinuities as they are frequently encountered in urban areas. Again, the estimation framework is a Markov random field, which is optimized using a graph-cut-based optimization. Although the authors state the algorithm has yet to be extended in order to be able to consider geometrical distortions such as layover, it is able to compute highly accurate height maps of areas containing large, isolated buildings.

3.3 Persistent Scatterer Interferometry

One of the most important developments in the field of SAR remote sensing certainly was the introduction of persistent scatterer interferometry (PSI) by Ferretti et al. [2001]. The main advantage of this technique is that only quasi-deterministic scatterers, whose reflectivities remain stable during a set of multi-temporal acquisitions, are considered for the analysis. Thus, high

quality 3D point clouds can be derived, even containing information about the movement of the points [Kampes, 2006]. Since persistent scatterers rely on stable backscattering processes as they usually only occur on non-vegetated areas, the high potential for urban area remote sensing has been recognized early [Perissin & Rocca, 2006; Crosetto et al., 2010]. The breakthrough, however, came with the advent of very high resolution spaceborne SAR missions, which now even allowed for a 4-dimensional monitoring of single buildings [Gernhardt et al., 2010; Gernhardt, 2012; Schunert & Soergel, 2012].

In parallel, Ferretti et al. [2011] extended the PSI concept to distributed scatterers, i.e. in addition to resolution cells containing just point scatterers now also groups of resolution cells caused by the same backscattering phenomenon can be considered now. This way, the number of scatterers – and also the number of reconstructed 3D points – can greatly be enlarged [Wang et al., 2012].

However, as with all other hitherto mentioned approaches, PSI itself does not provide a solution to the missing information caused by radar shadowing. Therefore, a fusion of PSI point clouds processed from data acquired from ascending and descending orbits has been proposed in analogy to the previously mentioned multi-aspect fusion in the image analysis context [Gernhardt et al., 2012].

3.4 SAR Tomography and Layover Separation

Among the most recent developments in urban SAR remote sensing is the extension of PSI by an adaption of SAR tomography (TomoSAR) to the discrete scattering profiles, which are frequently encountered in man-made environments.

Originally, the desire to extend the synthetic aperture radar (SAR) principle to a fully three-dimensional imaging technique has led to the highly investigated field of TomoSAR. Since its first practical introduction around the turn of the century [Homer et al., 1996; She et al., 1999; Reigber & Moreira, 2000], a rapid development of more sophisticated processing algorithms – providing, for example, super-resolution [Röbbing & Ender, 2001] – as well as the opening of a variety of application fields has taken place: At the beginning, the most important research direction was the analysis of volume structures such as forests. Using sensors with relatively low frequencies (e.g. L- or P-band), the coherent combination of multi-baseline SAR images from several flight tracks enables a full reconstruction of the continuous volume between canopy and ground [Frey et al., 2008; Frey & Meier, 2011; Tebaldini & Rocca, 2012; Aguilera et al., 2013], even allowing for the detection of objects hidden below foliage [Nannini et al., 2008; Huang et al., 2012].

Besides this reconstruction of continuous reflectivity profiles, SAR tomography also makes the determination of multiple (discrete) scattering contributions within a single resolution cell possible. This is particularly interesting for the separation of layover contributions in mountainous terrain and urban areas and can also be seen as a multi-baseline extension to conventional single-baseline SAR interferometry (InSAR) [Gini et al., 2002; Lombardini et al., 2003]. With the newest generation of sub-meter- and even decimeter-resolution spaceborne SAR sensors, the interest in this topic was newly stimulated as now a detailed 3D analysis of densely built-up inner city areas has become feasible. The most recent development in this field is based on the rationale that discrete scatterers cause sparse reflectivity profiles, which eventually led to the adaption of sparse reconstruction techniques and compressive sensing theory to the TomoSAR context [Budillon et al., 2011; Zhu & Bamler, 2010a]. This way, even higher tomographic focusing resolutions have become possible. In addition to that, TomoSAR methods can be integrated into the persistent scatterer framework, which enhances the potential of PSI for urban point cloud reconstruction even more [Ferretti et al., 2005; Fornaro & Serafino, 2006; Zhu et al., 2012]. The most recent

step in this field of research is the fitting of facades to the point cloud data in order to enable the reconstruction of building models in a bottom-up approach [Zhu & Shahzad, 2014].

3.5 Contribution of this Thesis

While the methods summarized in Section 3.1 all start from conventionally derived InSAR height maps and try to cope with the SAR inherent geometric effects (layover and shadowing) by image analysis techniques, the described PSI and TomoSAR approaches share the need for large multi-temporal data stacks acquired by repeat-pass platforms. However, collecting this kind of repeat-pass data is expensive and cumbersome, especially if the main advantage of SAR remote sensing, namely its applicability in time-critical scenarios, is considered. In case, for example, a 3D model of a disaster-affected city is to be produced in order to provide orientation information for response teams, it is not possible to wait for a whole PSI stack to be delivered.

Apart from that, current satellite missions only use ascending and descending orbits, which means that only two aspects can be combined for filling up image gaps caused by shadowing. In contrast, airborne single-pass multi-baseline SAR interferometers provide small InSAR stacks from just a single flight over the scene of interest, and are able to fly along arbitrarily defined flight trajectories in order to realize a large variety of many different multi-aspect configurations. Figure 3.1 illustrates the wide range of possible acquisition campaigns: While *aspects* are defined by the heading angle of the flight, *paths* relate to differing flying altitudes, which lead to severely dissimilar viewing angles such that no interferograms can be created between acquisitions of two different ones. In contrast to that, multiple *passes* with only slightly different trajectories can be used to simulate a multi-*antenna* array if the carrier platform is equipped with only one receiving antenna. It is important to mention, however, that in the context of the methods described in this thesis, paths are treated like aspects ($Q = Q_1 + Q_2$), whereas for each aspect/path either multi-antenna or multi-pass data can be employed to form multi-baseline interferograms ($N = N_1$ or $N = N_2$).

Therefore, with this kind of airborne sensor setup, the full timely flexibility of SAR remote sensing is kept, while still multi-baseline data acquired from multiple aspects can be exploited for comprehensive urban area reconstruction.

In this context, the contribution of this thesis is to provide new processing chains, which enable the comprehensive reconstruction of urban area surface models, where the layover and shadowing effects are to be coped with inherently. That is, conventional SAR interferometry is to be extended by multi-baseline techniques and multi-aspect data fusion in such a way that the resulting height data is as correct and as precise as possible and without having to collect large amounts of multi-temporal data first.

In order to reach this main goal, the intermediate goals are three-fold:

- 1) Development of adaptive estimation techniques for covariance matrices of single-pass multi-baseline InSAR resolution cells. This is a necessary prerequisite for a statistical exploitation of coherent SAR observations.
- 2) Development of a workflow for tomographic layover separation in single-pass multi-baseline InSAR stacks with just few images.
- 3) Development of methods for the fusion of multi-aspect multi-baseline InSAR data both after tomographic height reconstruction as well as during the height reconstruction process.

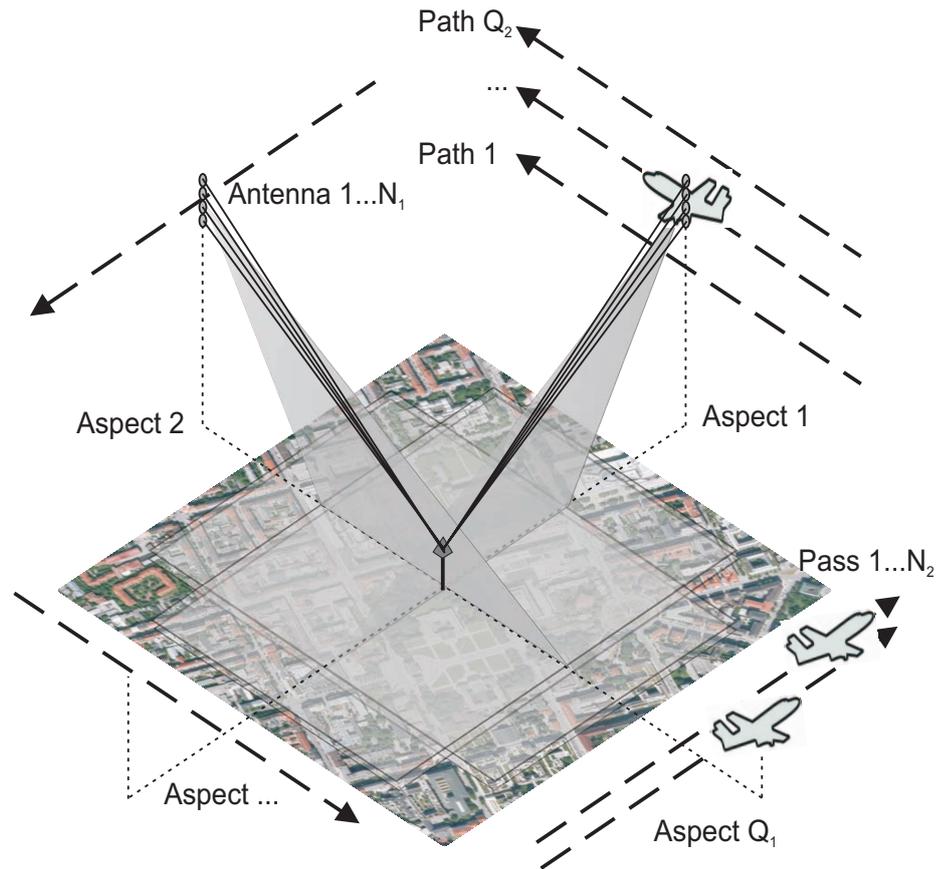


Figure 3.1: Sketch illustrating the variety of the possible multi-aspect multi-baseline InSAR acquisition geometries. The acquisition configuration can be chosen almost arbitrarily.

Although the goals of this thesis aim at the peculiarities of airborne single-pass multi-baseline SAR interferometers, all developed methods can be applied to any kind of SAR data straight-forwardly, no matter if spaceborne or airborne, repeat-pass or single-pass – basically even to imagery acquired by circular SAR as described e.g. by Oriot & Cantalloube [2008] or Palm et al. [2012]. The generalization usually is straight-forward and consists mostly of an adaption of the flight geometry from linearized tracks to, for example, polynomial orbits, or an addition of more available observations.

4 Adaptive Covariance Matrix Estimation

Apart from multi-looking based SAR tomography as, for example, presented by Gini & Lombardini [2005] or Baselice et al. [2009a] the estimation of the covariance matrix for each resolution cell is a critical processing step for many multi-dimensional applications of synthetic aperture radar imaging [Just & Bamler, 1994; Tough et al., 1995; Cloude & Papathanassiou, 1998; Lopez-Martinez & Fabregas, 2003].

In the low- and medium-resolution beginnings of InSAR mapping, fixed-size boxcar windows were commonly used for both the estimation of the complex coherence and the denoising of the interferometric phase as the hypothesis of statistical homogeneity of neighboring pixels could be considered valid for rural scenes, which were in the center of interest then. For high-resolution images of complex scenes such as urban areas, however, this approach is not suitable anymore [Touzi et al., 1996; Gao, 2010]. Therefore, more sophisticated adaptive methods have to be found in order to ensure an unbiased estimation of the covariance matrices. In the literature, many papers can be found on the topic of adaptive filtering of different kinds of SAR data, all based on the exploitation of homogeneous pixel neighborhoods. Among the first problems to be addressed was the simple speckle filtering of amplitude or intensity imagery [Lee, 1980; Frost et al., 1982; Kuan et al., 1985]; a comprehensive survey of the methods published during the 1980s and 1990s can be found in [Touzi, 2002]. Via the application to multi-temporal amplitude imagery [Ciuc et al., 2001], the general idea of these filters was gradually extended to interferometric and polarimetric SAR data, now aiming at an unbiased estimation of phase, coherence, and polarimetric scattering information [Lee et al., 1998; Vasile et al., 2004, 2006]. One of the most efficient filters in this context was proposed by Deledalle et al. [2011]: This algorithm, which utilizes a non-local estimation framework [Buades et al., 2005; Yang & Clausi, 2009], allows for the simultaneous extraction of all relevant information of a pair of co-registered InSAR images. Unfortunately, NL-InSAR was designed under the assumption of bivariate datasets; it can not be applied to stacks of multi-baseline imagery*.

Therefore, more recently some first papers have been published about the adaptive filtering of multi-dimensional SAR stacks. Ferretti et al. [2011] proposed DespecKS, an algorithm embedded in their SqueeSAR framework that uses a two-sample Kolmogorov-Smirnov test in order to evaluate if two stack pixels within a pre-defined search window belong to the same statistical distribution. Parizzi & Brcic [2011] further investigated this approach with respect to different goodness-of-fit tests such as Kullback-Leibler divergence, Anderson-Darling test, or generalized likelihood ratio test. Although all these formulations show promising results, they all suffer from one certain disadvantage: They typically work only for stacks of at least eight images and up [Stephens, 1970]. If, however, just standard InSAR pairs or stacks with a limited number of

*Only during finalization of this thesis, an extension of NL-InSAR to multi-dimensional data – now called NL-SAR to express its generalized nature – was proposed in a pre-print [Deledalle et al., 2013].

images, e.g. acquired by single-pass multi-baseline systems as they are frequently equipped on airborne platforms, are to be processed, alternative methods are necessary. Aiming to overcome the limitation on the large sample number inherent to [Ferretti et al., 2011] and [Parizzi & Brcic, 2011], in this thesis two novel algorithms for the adaptive multilooking of airborne single-pass multi-baseline InSAR stacks that typically consist only of a low number of simultaneously acquired images (three to six) are described. The first method was published in [Schmitt & Stilla, 2014b] and intends to concentrate the information of the stack – no matter of how many acquisitions it consists – via principal component (PC) analysis (PCA). A filtered version of the first PC is then used to determine homogeneous pixels around the resolution cell of interest using just a simple thresholding. From these homogeneous pixels then the complex covariance matrix of the resolution cell is estimated. The second approach is described in [Schmitt et al., 2014a]. In contrast to all other mentioned filters, which rely on the assumption that the amplitude values of the pixels can be used as a hint for changes in their phase values as well, it exploits the whole complex information for homogeneity determination. Usually, the determination of statistical similarity is solely based on the pixel amplitudes, while phase and coherence information is neglected. If an object in the scene causes a backscattering amplitude similar to the background, its phase values will be joined with the phase of this background during filtering, potentially causing blurring or even the loss of the object’s phase information. In comparison, the general idea of this new method is based on an analysis whether resolution cells belong to the same statistical distribution as the currently investigated center pixel by thresholding of the respective probability density function.

4.1 Statistical Properties of Multi-Baseline InSAR Resolution Cells

This whole thesis is centered around the assumption that each single look complex (SLC) observation z agrees at least approximately with Goodman’s model [Goodman, 1975], where both the real and imaginary parts follow a zero-mean Gaussian distribution, i.e.

$$\operatorname{Re}\{z\}, \operatorname{Im}\{z\} \sim \mathcal{N}\left(0, \frac{\sigma^2}{2}\right), \quad (4.1)$$

and are statistically independent (cf. Section 2.1.2). Although De Zan [2008] has investigated this assumption in comparison to a Constant-plus-Gaussian model, which might be considered more appropriate for heterogeneous scenes such as urban areas, where often (quasi-)deterministic point scattering can occur, he came to the conclusion that it is a sufficient approximation. This view is also seconded by other authors working with X-band SAR [Baselice et al., 2009a; Wang et al., 2012], while for smaller wavelengths (e.g. millimeterwave SAR), the validity of the assumption should be even better (cf. Section 7.2).

Therefore, it follows that each pixel vector \mathbf{z} in a stack of coregistered SAR images can be considered a sample from the multi-variate probability density function

$$f(\mathbf{z}) = \frac{1}{\pi^N \det(\mathbf{C})} \exp(-\mathbf{z}^H \mathbf{C}^{-1} \mathbf{z}), \quad (4.2)$$

which is fully characterized by its complex covariance matrix

$$\mathbf{C} = \begin{bmatrix} I_1 & \gamma_{12}\sqrt{I_1 I_2} & \dots & \gamma_{1N}\sqrt{I_1 I_N} \\ \gamma_{12}^*\sqrt{I_1 I_2} & I_2 & \dots & \gamma_{2N}\sqrt{I_2 I_N} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{1N}^*\sqrt{I_1 I_N} & \gamma_{2N}^*\sqrt{I_2 I_N} & \dots & I_N \end{bmatrix} \quad (4.3)$$

The I_i denote the intensities of all N acquisitions in the stack, whereas

$$\gamma_{ij} = \frac{E \{ z_i z_j^* \}}{\sqrt{I_i I_j}} = |\gamma_{ij}| \exp(j\phi_{ij}) \quad (4.4)$$

is defined as the complex coherence, which itself is composed of the magnitude of coherence (or correlation) $|\gamma_{ij}|$ and the interferometric phase ϕ_{ij} between acquisitions i and j , where z_i and z_j are the related complex observations.

Applying the normalization of (4.4) to (4.3) such that the main diagonal of the covariance matrix becomes a unity vector, we receive the so-called coherence (or correlation) matrix

$$\mathbf{\Gamma} = \begin{bmatrix} 1 & |\gamma_{12}| \exp(j\phi_{12}) & \dots & |\gamma_{1N}| \exp(j\phi_{1N}) \\ |\gamma_{21}| \exp(j\phi_{21}) & 1 & \dots & |\gamma_{2N}| \exp(j\phi_{2N}) \\ \vdots & \vdots & \ddots & \vdots \\ |\gamma_{N1}| \exp(j\phi_{N1}) & |\gamma_{N2}| \exp(j\phi_{N2}) & \dots & 1 \end{bmatrix}. \quad (4.5)$$

The most interesting characteristic of the complex covariance matrix of InSAR stack pixels therefore is that the covariance information is directly related to its interferometric information content. After estimating

$$\mathbf{C} = E\{\mathbf{z}\mathbf{z}^H\} \in \mathbb{C}^{N \times N} \quad (4.6)$$

for each pixel by

$$\hat{\mathbf{C}} = \frac{1}{L} \sum_{l=1}^L \mathbf{z}_l \mathbf{z}_l^H \quad (4.7)$$

from L statistically homogeneous pixels, all the denoised interferometric measurements can be extracted: despeckled intensity images, all possible multi-looked interferograms, and finally all related coherence maps. Two important facts have to be noted in this context: First, in addition to ensuring backscattering homogeneity, usually also the local phase fringe frequency has to be eliminated for a fully unbiased estimate of the covariance matrix [Trouve et al., 1996; Wu et al., 2006; Vasile et al., 2008; Cai et al., 2008]. For small patch sizes and a relatively large ambiguity height, however, a simple subtraction of the phase corresponding to the terrain surface of the scene can be considered a valid approximation [Richards, 2007]. Second, whenever possible, it should be ensured that $L > N$ in order to receive a non-singular covariance matrix for which its often needed inverse exists.

Two approaches for adaptive covariance matrix estimation with an emphasis on small interferometric stacks (i.e. a comparably low number of images, usually between three and six) are described in the following sections. Boiled down to their essence, they are supposed to answer the question: Which pixels should be used for covariance matrix estimation?

4.2 Principal Component Analysis-Based Thresholding

In this section, the thresholding based procedure carried out on the first principal component of the stack of logarithmic amplitude images of the dataset is described. The PCA is applied in order to get a more reliable indicator of backscattering homogeneity as would be possible from one raw amplitude image alone (see Section 4.2.1.) After the PCA, as described in Section 4.2.2, a total variation (TV) norm based denoising algorithm [Getreuer, 2012] is used in order to receive a more reliable indicator of backscattering energy on which then a sliding window is moved across.



Figure 4.1: A flowchart of the PCA-TV-based covariance matrix estimation method. Note that after the thresholding an 8-connectivity check is applied.

Within the sliding window, a thresholding is applied that yields the pixels of the window that are considered to show homogeneous backscattering.

A flowchart of the proposed procedure is shown in Fig. 4.1. Inspired by [Ferretti et al., 2011], a sliding search window is defined that moves from pixel to pixel, such that each pixel is considered as a center pixel of the search window once.

4.2.1 Principal Component Analysis of the InSAR Stack

Principal Component Analysis is a well-known mathematical method that uses an orthogonal transformation in order to convert a dataset into a set of linearly uncorrelated variables, called principal components [Shlens, 2009]. The idea is to identify linear combinations of the original variables that contain most of the information present in the data. This is based on the assumption that useful information is proportional to the variance of the data. Since the number of principal components is less than or equal to the number of original variables, the method is often employed for dimension reduction.

Considering a stack of N coregistered SAR amplitude images, the first step is a transformation of the stack onto its principal components. In order to ensure that the Gaussian assumption that is a prerequisite for a correct application of PCA is met at least approximately, the decadic logarithm to the approximately Rayleigh distributed amplitude images is applied [Gao, 2010]. Afterwards, all the amplitude values of the stack are put into a two-dimensional data matrix $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_N]^T$ with \mathbf{x}_i being the pixel values of image i put into a row vector. This means, the rows of \mathbf{X} describe the images of the stack as variables, whereas the columns denote the logarithmic amplitude values

as observations of these variables. After mean-centering of the dataset, i.e. subtraction of the row means, the data covariance matrix is estimated by

$$\mathbf{K} = \frac{1}{P-1} \mathbf{X}\mathbf{X}^T \quad (4.8)$$

P denotes the number of pixels in one image, \mathbf{K} quantifies the correlations between all possible pairs of observed amplitudes. Since we want to reduce any redundancies, i.e. exploit redundant observations in order to remove noise, we seek a transformation of the data such that the covariances between separate measurements become zero.

Therefore, \mathbf{K} is decomposed by eigenvalue decomposition, i.e.

$$\mathbf{K} = \mathbf{E}\mathbf{D}\mathbf{E}^T, \quad (4.9)$$

where \mathbf{E} is a matrix containing the eigenvectors of \mathbf{K} , and \mathbf{D} is a diagonal matrix containing the related eigenvalues. The transformation matrix we are looking for is then simply defined by $\mathbf{A} = \mathbf{E}^T$, if the eigenvectors contained in \mathbf{E} are ordered with respect to their related eigenvalues (from the largest eigenvalue to the smallest one).

Finally, we are able to project our stack of amplitude images onto its principal components using

$$\tilde{\mathbf{X}} = \mathbf{A}\mathbf{X}. \quad (4.10)$$

$\tilde{\mathbf{X}}$ then consists of N rows and P columns where each row represents one principal component of the dataset. Since we assume that the first principal component (PC1) contains most relevant signal information, while the other principal components contain only noise (cf. Fig. 4.2), we just keep PC1, i.e. the first row of $\tilde{\mathbf{X}}$ as input to the neighborhood homogeneity test. Exemplary relations between the first PC and the following ones for stacks of different size can be found in Fig. 4.3. It can be seen that a distinction between signal and noise information is of course not possible with one image alone, whereas the relative share of signal energy contained in the first principal component becomes less for larger stacks until at some point convergence will be reached. Note that the relationship between the single principal components depends on the overall noise level.

4.2.2 Total Variation Denoising

After the data have been projected onto the first principal component, a TV norm based image denoising algorithm is used to receive an even more reliable indicator of the backscattering characteristics of the scene [Getreuer, 2012]. The algorithm combines the Split Bregman method for L_1 norm regularization problems [Goldstein & Osher, 2009] with the Rudin-Osher-Fatemi problem [Rudin et al., 1992], which basically consists of estimating a denoised image u as the solution of the minimization

$$\hat{u} = \arg \min \|u\|_{TV} + \frac{\mu}{2} \|f - u\|_2^2, \quad (4.11)$$

where μ is a positive regularization parameter that can be determined by methods like generalized cross-validation (GCV) or L-Curve [Batu & Cetin, 2011]. $f = u + n$ describes the relation between the noisy image f , the denoised image u and the additive white Gaussian noise n . The term

$$\|u\|_{TV} \approx \sum_{i=1}^{rows} \sum_{j=1}^{cols} |\nabla u_{i,j}| \quad (4.12)$$

describes the TV norm of u , which sums the vector magnitude over all pixels $u_{i,j}$ and basically is the L_1 -norm of the gradient image. By looking for the solution with the sparsest gradient image

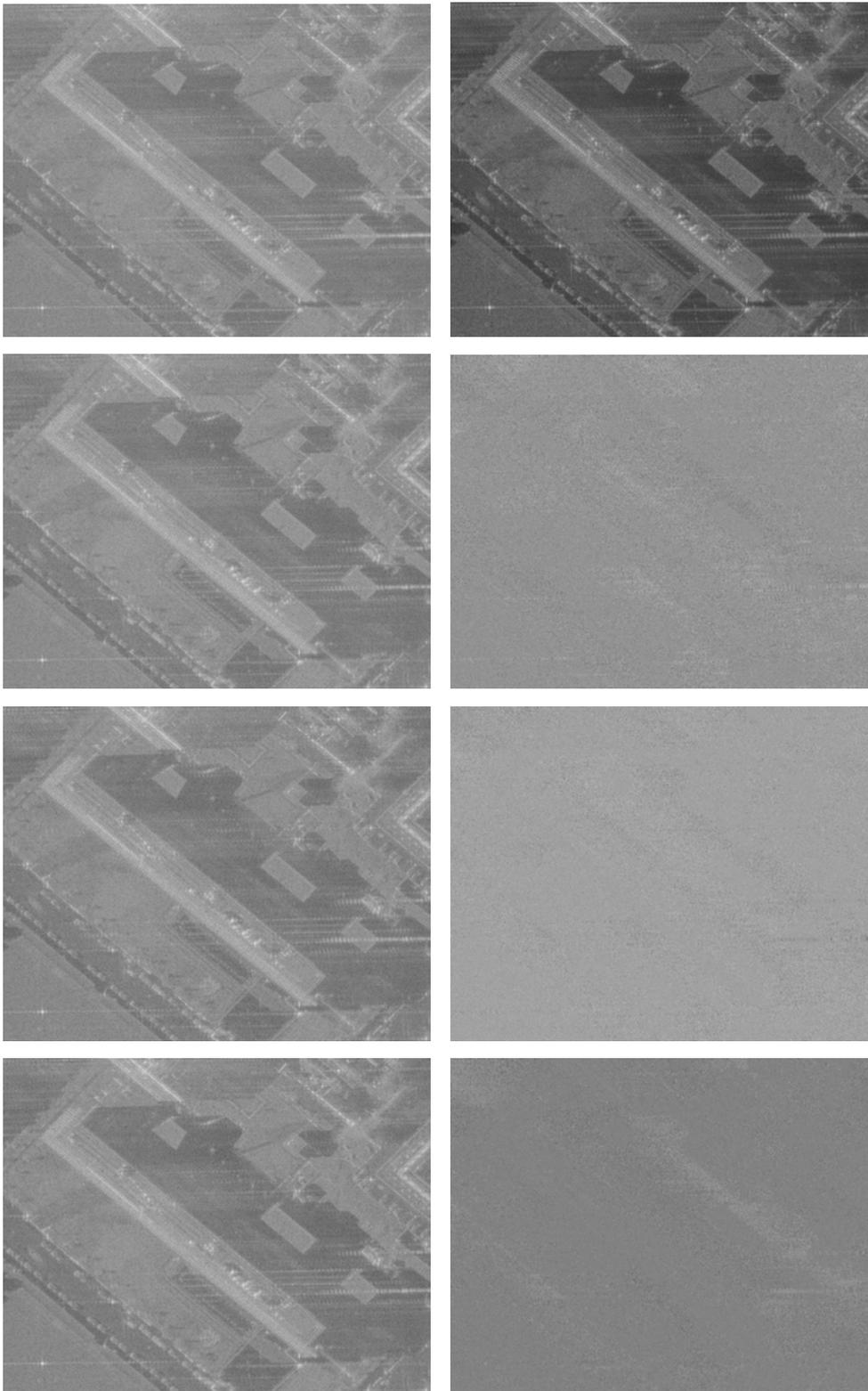


Figure 4.2: Example of PCA on a stack of four logarithmic amplitude images (left column). The corresponding principal components are shown in the right column, from top to bottom sorted by their corresponding eigenvalues. In this case, the first principal component accounted for 82% of the signal energy.

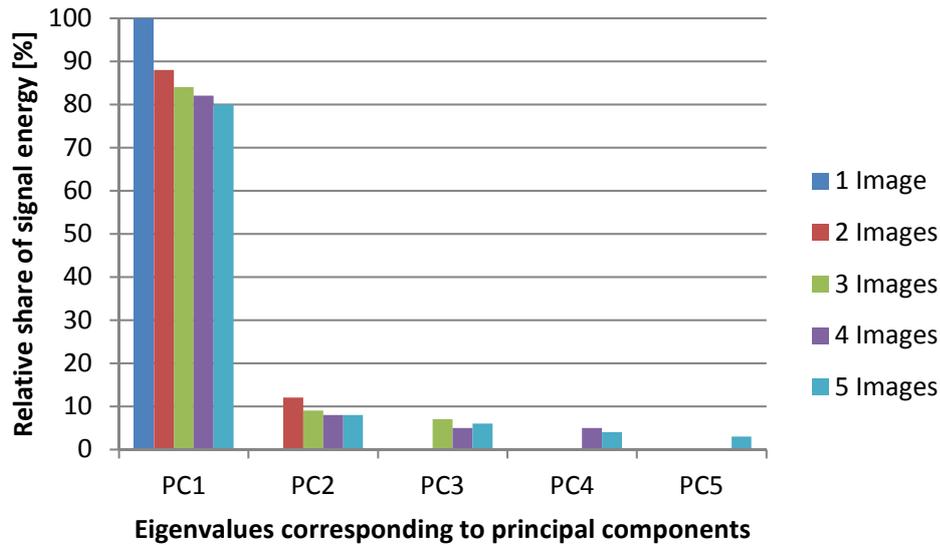


Figure 4.3: The relative share of the eigenvalues corresponding to the principal components of data stacks with different number of images is plotted exemplarily. The sum of all eigenvalues always gives the entire energy content in the signal or image, respectively. Note that the number of principal components is always less than or equal to the number of available images.

discontinuity-preserving filtering is ensured. The second term in (4.11) ensures that the solution of the optimization is as close as possible to the observed image f . The resulting image then is a filtered version of the main principal component with a high rate of detail and edge preservation that serves as a very good index for the following homogeneity investigation.

4.2.3 Thresholding

Within the search window, which is moved across the denoised PC1 image, all pixels are checked for similarity to the respective center pixel by a simple thresholding:

$$|g_p - g_c|^2 \leq \sigma \quad (4.13)$$

g_c denotes the PC1 value of the center pixel, g_p the PC1 value of the tested pixel, and σ the standard deviation of the whole first principal component image. The reason the distance between g_p and g_c is taken to the power of 2 is motivated by the consideration that small differences (< 1) are mellowed whereas larger differences (> 1) are even emphasized. From empirical trials, it was found that this way also outlier pixels that sometimes occur within homogeneous areas are mitigated.

The main advantages of this approach over region-growing based algorithms are the speed and a natural upper limit of the number of tested pixels given by the search window size. This limited extent enhances the chance that the homogeneous patch as detected by the algorithm really just corresponds to one certain backscattering phenomenon of the scene. If on the contrary the emphasis is put on larger neighborhoods and stronger filtering, the search window size can easily be enlarged by the cost of computational speed and robustness of the thresholding step.

Finally, after the thresholding, an 8-connectivity check is carried out in order to ensure that only pixels connected to the central pixel are considered to belong to the same backscattering area. An example of the thresholding for a heterogeneous as well as a homogeneous image patch can be seen in Fig. 4.4. The main advantage of this inspection is that the search window can

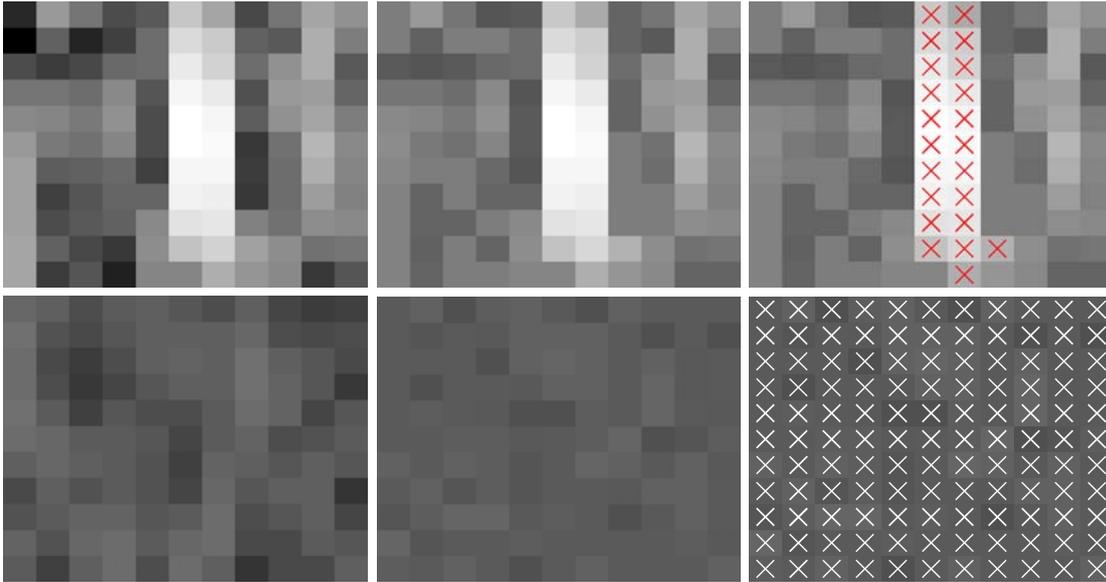


Figure 4.4: Thresholding on filtered main principal component for a heterogeneous (top row) as well as a homogeneous (bottom row) image patch. From left to right the search window extracted from the main principal component, the TV-filtered window and the thresholding result are shown: pixels that passed the thresholding test and are included in the homogeneous neighborhood are marked.

in principle be chosen arbitrarily large, because spatial similarity, i.e. all pixels really belonging to the same backscattering phenomenon, is ensured by the pixel connectivity. In a final step, it would be possible to check the detected neighborhood for a minimum size, e.g. in order to ensure a minimum amount of smoothing or the estimation of a non-singular covariance matrix. By default, the minimum neighborhood size is set to 1 pixel due to the intention of showing the natural behaviour of the proposed method. An example of the final neighborhood size for every pixel can be seen in Fig. 4.5. After having determined the homogeneous neighborhood, the sample covariance matrix is estimated by (4.7).

4.3 Probabilistic Similarity Determination

In this section, another approach for adaptive InSAR covariance matrix estimation is described. In analogy to the PCA-based thresholding procedure described in the last section, also this method is centered around a sliding window operation (see Fig. 4.6). Again, the *search window* defines the maximum size of the homogeneous pixel neighborhood. In addition to that, a smaller *central window*, which need not be rectangular necessarily, is defined. Using the pixels in this central window, a robust covariance estimator (see next section) is applied in order to generate an initial guess for the covariance matrix of the respective center pixel from its direct, enclosing neighbors. Afterwards, the circularly-symmetric complex Gaussian probability density function for each pixel in the search window is evaluated with respect to this initial covariance matrix, and a threshold is applied such that pixels belonging to the distribution of the center pixel are separated from pixels belonging to a different distribution. Again, the adaptivity test is finalized by a connectivity check in order to ensure that only pixels connected to the center pixel are considered for the neighborhood. Finally, the desired covariance matrix is estimated from all *connected inliers*, i.e. from all pixels that were detected to be part of the homogeneous neighborhood.

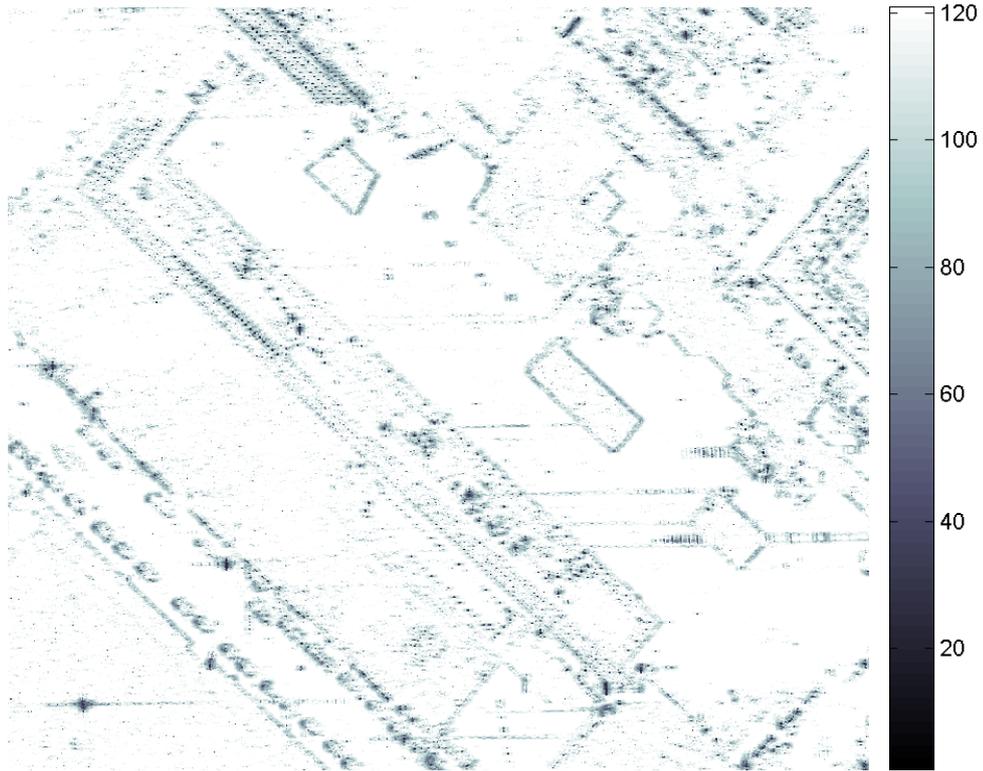


Figure 4.5: Size of the homogeneous neighborhood as detected for each pixel. Note how especially strong scatterers are generally just comprised of very few pixels, while homogeneous areas use almost the whole search window most of the time.

4.3.1 Robust Estimation of Initial Covariance Matrix

The initial guess for the covariance matrix of the center pixel's distribution is calculated from its enclosing 4- or 8-neighborhood. However, it is possible – and in heterogeneous areas such as urban scenes even probable – that not all resolution cells in this neighborhood really belong to a statistically homogeneous population of backscattering observations. Hence, a robust estimation of $\hat{\mathbf{C}}_{init}$ is employed in order to minimize the potential bias. An additional advantage of this robust initialization is that the effect of single pixels affected by strong speckle is mitigated. The utilized M-estimator, which has been proposed by Ollila & Koivunen [2003] and was recently discussed in further detail by Zoubir et al. [2012] is defined as

$$\hat{\mathbf{C}}_{init,k+1} = \frac{1}{L_{init}} \sum_{l=1}^{L_{init}} w \left(\mathbf{z}_l^H \hat{\mathbf{C}}_{init,k}^{-1} \mathbf{z}_l \right) \mathbf{z}_l^H \mathbf{z}_l, \quad (4.14)$$

where L_{init} is the initial number of looks, i.e. the size of the initial neighborhood, and $\hat{\mathbf{C}}_{init,0}$ is the standard sample covariance matrix estimated from the classic 3×3 -boxcar window. Since L_{init} needs to be larger than N in order to ensure that $\hat{\mathbf{C}}_{init,0}$ is non-singular and can be inverted in (4.14), a larger initial neighborhood has to be employed for stacks with more than 9 images. $w(x)$ is a robust weighting function, which is descending to zero, such that a highly deviating observation \mathbf{z}_l with large $\|\hat{\mathbf{C}}_{init}^{-\frac{1}{2}} \mathbf{z}_l\|^2 = \mathbf{z}_l^H \hat{\mathbf{C}}_{init}^{-1} \mathbf{z}_l$ receives smaller weights in the estimation. The general idea behind this robust estimation procedure is the assumption that the random vectors \mathbf{z}

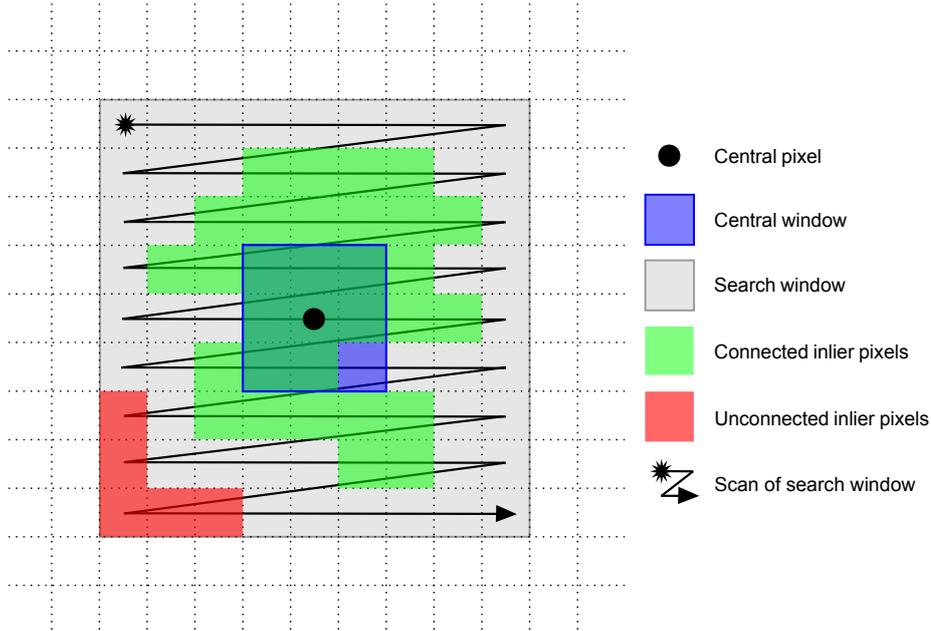


Figure 4.6: Sketch of the sliding window based adaptive filter. Within an arbitrarily large *search window* around the pixel of interest, all pixels are thresholded whether they belong to the probability distribution defined by the robust covariance matrix estimated initially from the pixels in the *central window*. A subsequent connectivity check then separates *connected inliers* from *unconnected inliers*.

in the local window are assumed to follow a complex multivariate t -distribution $Ct_{k,\nu}(\mu, \mathbf{C})$ with $\nu > 0$ degrees of freedom:

$$f(\mathbf{z}|\mu, \mathbf{C}) = c|\mathbf{C}|^{-1} \left(1 + 2\frac{s}{\nu}\right)^{-\frac{2k+\nu}{2}}, \quad (4.15)$$

where $s = (\mathbf{z} - \mu)^H \mathbf{C}^{-1} (\mathbf{z} - \mu)$ and c is a normalizing constant. With $\nu = 1$, this distribution is called the multivariate complex Cauchy distribution, which is a prominent robust heavy-tailed alternative for the Gaussian distribution, that itself is obtained for $\nu \rightarrow \infty$. Due to this reason, the complex multivariate t -distributions are very useful for analyzing the robustness of multivariate statistics, since a decrement of ν yields a distribution with an increased heaviness of the tails. Therefore, the weighting function is chosen by

$$w(x) = w_\nu(x) = \frac{2N + \nu}{\nu + 2x}. \quad (4.16)$$

Usually, the parameter is chosen in the range $1 \leq \nu \leq 5$ if the goal is to ensure a certain robustness against outliers and noisy data.

As a stopping criterion for the iterative estimation of (4.14)

$$\|\hat{\mathbf{C}}_{init,k+1} - \hat{\mathbf{C}}_{init,k}\| < \varepsilon \quad (4.17)$$

is chosen, where $\varepsilon > 0$ is a small number. The benefit achieved by relying on a robust estimation of the initial covariance matrix instead of just the standard sample estimate is illustrated in Fig. 4.7.

4.3.2 Similarity and Connectivity Testing

After a robust initial guess of the distribution's covariance matrix has been attained, the statistical similarity to the corresponding central population is determined for each pixel within the larger

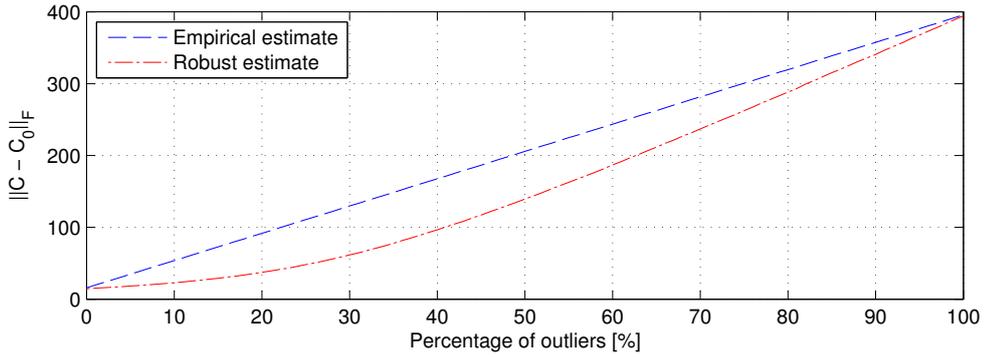


Figure 4.7: Comparison of conventional and robust covariance matrix estimation. The Frobenius norm of the difference between estimated covariance matrix and theoretical covariance matrix is plotted with respect to a growing percentage of outlier observations. The benefit of robust covariance matrix estimation is clearly visible.

search window. A stack pixel \mathbf{z} is assumed to be part of the distribution $\mathcal{CN}(\mathbf{0}, \hat{\mathbf{C}}_{init})$ of the central pixel if a certain threshold on the probability density is exceeded:

$$f(\mathbf{z} | \hat{\mathbf{C}}_{init}) > \varepsilon_{\text{pdf}}. \quad (4.18)$$

The choice of this threshold is further investigated in Section 4.3.3.

After applying the thresholding, a final 4- or 8-connectivity test is carried in analogy to the previously described PCA-TV-based approach. Then, the sample covariance matrix is estimated from the set of connected inliers by (4.7).

4.3.3 Determination of Optimal Parameter Settings

From Sections 4.3.1 and 4.3.2 it is obvious that the proposed algorithm depends on two main parameters, which have to be tuned manually: The degrees of freedom ν for robust estimation of the initial covariance matrix, and the probability density threshold ε_{pdf} for determination whether a pixel belongs to the same distribution as the center pixel. Fig. 4.8 shows two plots analyzing the optimal parameter choice with respect to the noise standard deviation and the number of images in the stack. For this investigation, monte carlo simulations were carried out. The simulated data is synthesized from a set of true amplitude and phase images according to the multiplicative speckle noise model discussed e.g. in [Richards, 2009]. The primary noise level is assumed to be equal for the real and imaginary part, resulting in Rayleigh distributed amplitude and Gaussian distributed interferometric phase (cf. Section 2.1.2). Finally, in addition to this SAR inherent speckle, the signal is augmented with Gaussian distributed thermal noise of the SAR sensor. Both speckle and thermal noise are summarized to describe the overall noise level shown in the evaluation figures.

Obviously, ν does not depend on the stack size but only on the noise level and is optimally chosen between 2.9 and 3.8, which corresponds to the statement of Ollila & Koivunen [2003], who suggest $1 \leq \nu \leq 5$. Due to the small variations, the choice of ν seems not very critical; if the noise level is not known a priori, $\nu = 3$ is suggested as a good compromise.

The decadic logarithm of ε_{pdf} stays at approximately -10 independently of the noise level. For larger stacks, however, it slowly decreases to -12 . However, also the choice of the threshold is not very critical. Therefore, in the remainder of this thesis, a constant threshold of $\varepsilon_{\text{pdf}} = 10^{-10}$ is

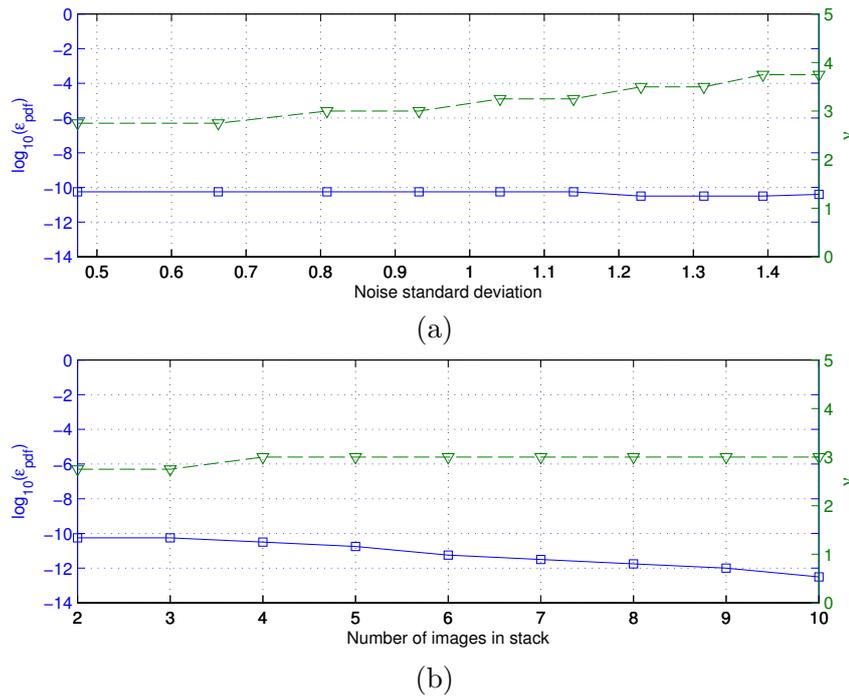


Figure 4.8: Optimal choice of degrees of freedom ν for the robust initial covariance matrix estimation and optimal choice of probability density threshold ϵ_{pdf} for similarity checking: (a) Plot for fixed stack size (2 images) with respect to growing noise level. (b) Plot for fixed noise level (STD 0.5) with respect to growing number of images in the stack.

employed. If large stacks, such as typically encountered in the context of repeat-pass spaceborne SAR, serve as input to the presented approach, the threshold should be lowered accordingly.

Considering these empirical findings, the proposed method does not necessarily require any manual settings or parameter tunings and can be applied to different kinds of InSAR data with pre-defined default parameters. Only if highest precision is required and, e.g., very large stacks are to be processed or very large noise levels are present, a refinement of the parameters can be advisable.

5 Maximum-Likelihood SAR Tomography

As already explained in Chapter 3, the extension of SAR interferometry to multi-baseline data and eventually to SAR tomography was one of the most important developments in SAR remote sensing so far – as now real three-dimensional imaging is enabled. This is especially true for the analysis of urban areas and other scenes rich of man-made structures, as the observations are frequently affected by layover in these cases. Although many promising methods for the separation of scattering signals contributing to a layover observation have been published during the last years, most of them rely on large data stacks as they can only be acquired in the context of repeat-pass campaigns. Whereas airborne sensors usually enable an almost arbitrary flight planning, yielding the possibility to acquire the repeat-pass data still in a relatively short time, this is an even more severe drawback for spaceborne platforms that follow pre-defined orbits and therefore usually experience a temporal baseline of at least several days. This repeat-pass configuration eventually leads to a) a loss in coherence depending on wavelength and temporal baseline, b) the need to consider displacements even of stable scatterers for larger temporal baselines, and c) the need to wait for data delivery, which can be unfavorable for time-critical applications (e.g. disaster response scenarios).

The reason for this drawback is that most TomoSAR methods rely on spectral estimation algorithms designed for array signal processing [Krim & Viberg, 1996], which means they are physically limited by the Rayleigh resolution of the available sensor geometry. This resolution in elevation direction can only be enhanced by providing a sufficient tomographic aperture and a sufficient number of receiving antennas. Although there are many super-resolution procedures available, even the most sophisticated ones are not able to resolve two arbitrarily close scattering contributions [Zhu & Bamler, 2011]. Due to the fact that the number of antennas and the overall antenna length is typically limited on single carrier platforms, the system attributes are then often artificially enhanced by simulating a multi-antenna sensor via multiple passes over the scene of interest.

In contrast, a maximum-likelihood-based estimation framework was developed in [Schmitt & Stilla, 2014a], which allows the separation of multiple scattering contributions in a single resolution cell without being limited by the Rayleigh resolution limit, since no conventional inversion of the spectrum is employed. As for other methods in this vein, apart from the signal-to-noise ratio (SNR) the achievable resolution only depends on a priori knowledge of the model order.

To the author’s knowledge, two different alternative maximum-likelihood-based TomoSAR approaches have been proposed so far: The first is an adaption of [Wax, 1991] and was e.g. described in [Sauer et al., 2009]. It is part of the class of spectral estimation algorithms, aiming at simultaneously solving the detection and localization (i.e. reconstruction) problems. The second one was published in [Baselice et al., 2009a] and is more similar to the approach proposed in this paper. However, the approach proposed in this thesis does not make use of any assumptions

concerning the coherence of the scatterers, nor does it aim at estimating scatterer heights and reflectivities in a single step. This reduces the number of unknowns that have to be estimated simultaneously significantly, and clearly separates detection from localization.

Since the proposed method allows a precise reconstruction of 3D scatterers and their reflectivities even for single-pass data, this contribution will help to explore the potentials and limits of SAR tomography with respect to sparse datasets. The findings will aid the optimization of future sensor developments as well as the design of orbit configurations for future tandem-like multi-satellite missions [Krieger & Moreira, 2005]. For some of the rare literature about single-pass SAR tomography experiments, see for example [Röbbing & Ender, 2001; Lombardini et al., 2004; Zhang et al., 2012].

In the remainder of this chapter, first the general TomoSAR imaging model will be explained, before a new two-step procedure for the realization of SAR tomography on single-pass stacks consisting of just few SLC images is described. After the complex covariance matrices of every resolution cell have been estimated, a model order selection for determination of the expected number of scatterers, which have contributed to the SLC observations is carried out. With the thus determined prior knowledge, tomographic SAR inversion is realized based on a maximum-likelihood estimation framework.

5.1 Tomographic Imaging Model

As already mentioned, TomoSAR aims at creating a synthetic aperture not only in azimuth but also in elevation direction. However, in opposition to the high number of densely and regularly spaced azimuth samples, for the elevation aperture comparably few observations per resolution cell are available in the form of a stack of N coregistered SLC SAR images acquired from slightly different antenna positions (cf. Fig. 5.1). As, for example, described by Zhu & Bamler [2010b], the complex measurement stored in a pixel of the n^{th} acquisition with a baseline B_n between the respective slave antenna n and the master antenna is the ingegral of the reflected signal weighted by a linear phase term:

$$z_n = \int_{s_{min}}^{s_{max}} x(s) \exp(j \cdot \phi_n(s)) ds, \quad (5.1)$$

where $x(s)$ is the reflectivity function along elevation, and $[s_{min}; s_{max}]$ defines the relevant part of the elevation profile. ϕ_n can be calculated via the vertical wavenumber $k_{z,n}$ as described in Section 2.3.4 (2.29), if the elevation s is converted to height h for compability with conventional interferometry by

$$h = s \cdot \sin(\theta). \quad (5.2)$$

In particular for sparse reflectivity profiles, as they are usually encountered in urban areas or mountaneous terrain, it is advisable to approximate (5.1) by discretizing and sparsifying it. The measurement vector is then formulated as

$$\mathbf{z} = \sum_{k=1}^K x_k \mathbf{a}(h) + \mathbf{n} = \mathbf{A}(\mathbf{h}) \mathbf{x} + \mathbf{n}. \quad (5.3)$$

In this notation, $\mathbf{x} = [x_1, \dots, x_K]^T$ is the source signal vector containing the complex reflectivities of the K discrete scattering contributions, and \mathbf{n} represents complex circularly symmetric Gaussian noise. $\mathbf{A}(\mathbf{h}) = [\mathbf{a}(h_1), \dots, \mathbf{a}(h_K)]$ is the $N \times K$ steering matrix containing K steering vectors each of which corresponds to one backscattering source:

$$\mathbf{a}(h) = [\exp(j \cdot k_{z,1} \cdot h), \dots, \exp(j \cdot k_{z,N} \cdot h)]^T. \quad (5.4)$$

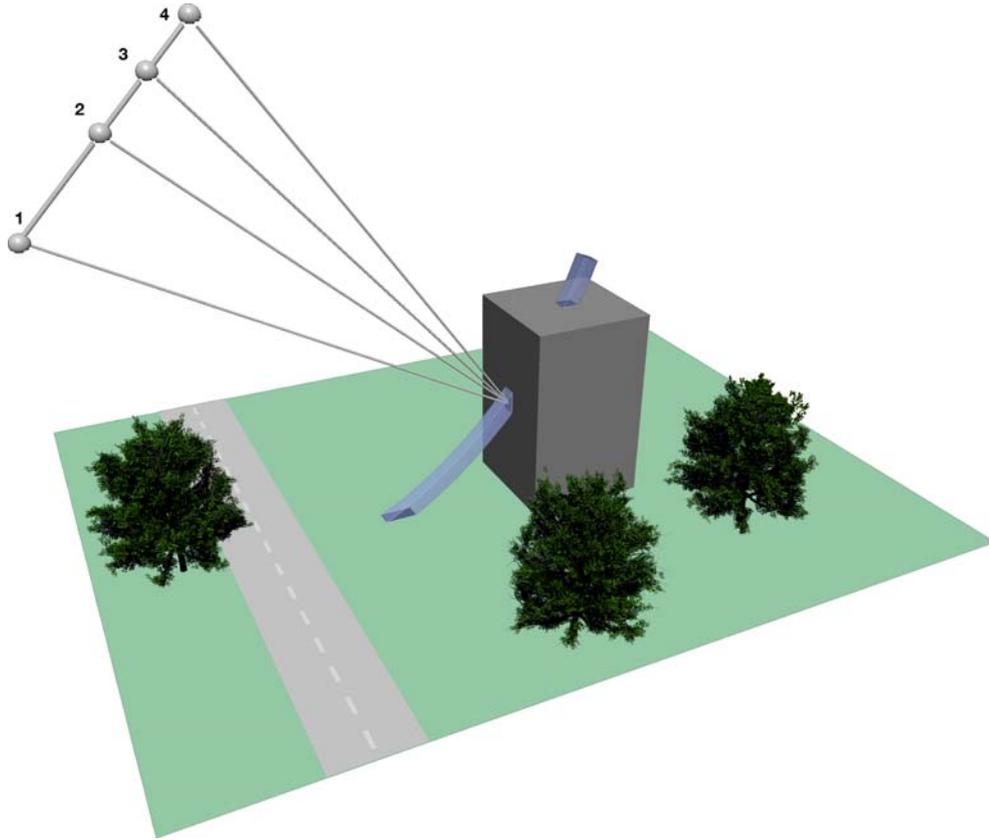


Figure 5.1: Typical TomoSAR situation for a single-pass antenna array consisting of four receiving antennas. The three backscattering contributions (lawn, facade, roof) are mixed in the resulting signal measured by the corresponding resolution cell. Antenna 1 is assumed to act both as transmitter (Tx) and receiver (Rx).

It has to be noted that the utilization of k_z in the presented linear manner is only justified if the sensor is sufficiently far from the target such that range- or height-dependent changes in the off-nadir angle (and therefore the phase) can be neglected. In the remainder of this thesis, approximation-free modelling of the phase term will be employed due to the focus on airborne SAR typically flown at comparably low altitudes.

In summary, the task of urban SAR tomography is now to invert the imaging model described in this section in order to reconstruct the K scatterers that contributed to a certain resolution cell both with respect to their elevation positions (corresponding to certain heights above the reference surface) and their reflectivities.

5.2 Model Order Selection

From the descriptions in the following section, it will become clear that the proposed multi-baseline InSAR method relies on a preliminary knowledge of the model order, i.e. the actual number of scattering contributions in each resolution cell. Very promising model order selection methods for multi-snapshot scenarios based on information theoretic criteria (ITC) have been proposed by Wax & Kailath [1985]. They have been further investigated in the context of multi-baseline SAR interferometry by Gini & Bordonni [2003] and Lombardini & Gini [2005]. Their basic idea is to minimize a criterion over the hypothesized number of signals k . All these approaches need a family of probability density functions, which describe the model that generated the data

while being a function of k . In addition to that, different penalty functions are needed, because otherwise the minimum would always be found for the highest possible model dimension. The general form of each ITC is then the penalized log-likelihood function

$$\text{ITC}(k) = -\ln f(\mathbf{g}|\hat{\vartheta}(k)) + p(\eta(k)), \quad (5.5)$$

where $\hat{\vartheta}(k)$ is the maximum-likelihood estimate of the vector of parameters, which describe the model that generated the data \mathbf{g} . $p(\eta(k))$ is the penalty function depending on the degrees of freedom $\eta(k)$. The main difference between the different ITC is the formulation of this penalty term: The Akaike information criterion [Akaike, 1974], for example, is then given by

$$\text{AIC}(k) = -\ln f(\mathbf{g}|\hat{\vartheta}(k)) + \eta(k), \quad (5.6)$$

while the minimum description length [Schwarz, 1978; Rissanen, 1978] is defined by

$$\text{MDL}(k) = -\ln f(\mathbf{g}|\hat{\vartheta}(k)) + \frac{1}{2}\eta(k) \log L. \quad (5.7)$$

The MDL is a special case of the family of efficient detection criteria developed by Zhao et al. [1986]. They are all consistently based on the formulation

$$\text{EDC}(k) = -\ln f(\mathbf{g}|\hat{\vartheta}(k)) + \eta(k) C_L, \quad (5.8)$$

where C_L can be any function of L such that

$$\lim_{L \rightarrow \infty} \frac{C_L}{L} = 0 \quad \text{and} \quad \lim_{L \rightarrow \infty} \frac{C_L}{\ln(\ln L)} = \infty. \quad (5.9)$$

In [Lombardini & Gini, 2005], two EDC were implemented by choosing $C_L = \log L$ (EDC₁) and $C_L = \sqrt{L \log L}$ (EDC₂), respectively.

As shown in [Wax & Kailath, 1985], the log-likelihood function, which serves as the first term of all the ITC can be expressed by

$$\ln f(\mathbf{g}|\hat{\vartheta}(k)) = L(N-k) \ln \left(\frac{\sqrt{\prod_{i=k+1}^N \hat{\lambda}_i}}{\frac{1}{N-k} \sum_{i=k+1}^N \hat{\lambda}_i} \right), \quad k = 0, 1, \dots, N-1 \quad (5.10)$$

for sensor array data affected by additive white Gaussian noise. The $\hat{\lambda}_i$ ($i = 1 \dots N$) are the eigenvalues in descending order of the previously estimated complex covariance matrix. In analogy, the expression for the degrees of freedom can be written as

$$\eta(k) = k(2N-k), \quad k = 0, 1, \dots, N-1. \quad (5.11)$$

Concluding from the available literature and empirical findings, in this thesis EDC₂ is utilized, which eventually transforms (5.5) to

$$\text{EDC}_2(k) = L(N-k) \ln \left(\frac{\sqrt{\prod_{i=k+1}^N \hat{\lambda}_i}}{\frac{1}{N-k} \sum_{i=k+1}^N \hat{\lambda}_i} \right) + k(2N-k) \sqrt{L \log L}, \quad (5.12)$$

where again L is the number of looks, i.e. the number of pixels that were used to estimate the sample covariance matrix. Looking at (5.12), it becomes obvious that a sensor with N antennas can only resolve $N-1$ scatterers, since the data term is always 0 for EDC₂ ($k = N$).

Accordingly, the number of scatterers in the resolution cell can finally be determined by

$$\hat{K} = \arg \min_{k \in [0; N-1]} \text{EDC}_2(k). \quad (5.13)$$

A more thorough explanation and discussion of ITC methods is beyond the scope of this thesis and has been carried out in the literature extensively. The reader is, for example, referred to [Wax & Kailath, 1985; Gini & Bordoni, 2003; Lombardini & Gini, 2005; Sauer, 2008] for additional details.

5.3 Maximum-Likelihood-Based TomoSAR Inversion

After the necessary pre-processing steps of adaptive covariance matrix estimation and model order selection, this section describes the core method for single-pass SAR tomography based on maximum-likelihood estimation. In a first step, the heights of the \hat{K} scatterers assumed to be contributors to the resolution cell are estimated using a \hat{K} -dimensional grid search. Afterwards, the reflectivities of the scatterers are reconstructed from the eigenvalues of the covariance matrix. A flowchart of the procedure is shown in Fig. 5.2.

5.3.1 Height Estimation

As all other methods developed in the framework of this thesis, also the proposed TomoSAR inversion algorithm is based on the assumption of Gaussian scattering as described in Section 2.1.2. Therefore, the modelling of the problem employs the circularly complex Gaussian probability density function introduced in (4.2):

$$f(\mathbf{z}) = \frac{1}{\pi^N \det(\mathbf{C})} \exp(-\mathbf{z}^H \mathbf{C}^{-1} \mathbf{z}). \quad (5.14)$$

However, for the algorithm presented in this thesis, things are done back to front: First of all, the sample covariance matrix $\hat{\mathbf{C}}$ is considered the observation, whereas \mathbf{z} is replaced by a model signal vector composed of the model signals related to N receiving antennas and \hat{K} unknown scatterer heights $\mathbf{h} = [h_1, \dots, h_{\hat{K}}]^T$:

$$\tilde{\mathbf{z}}(\mathbf{h}) = [\tilde{z}_1(\mathbf{h}), \dots, \tilde{z}_N(\mathbf{h})]^T. \quad (5.15)$$

Based on the a priori estimate of the model order K , the individual model signals $\tilde{z}_n(\mathbf{h})$ can be calculated by

$$\tilde{z}_n(\mathbf{h}) = \sum_{k=1}^{\hat{K}} \exp(j\tilde{\varphi}_{n,k}), \quad (5.16)$$

where

$$\tilde{\varphi}_{n,k} = \varphi_n(h_k) - \varphi_n(0) \quad (5.17)$$

is the absolute signal phase caused by scatterer k at height h_k minus the “flat earth phase” defined by the reference surface at height $h = 0$. Both can be calculated without any approximation by

$$\varphi_n(h) = -\frac{2\pi}{\lambda} \left(\sqrt{(R \sin \theta)^2 + (h - H)^2} + \sqrt{(R \sin \theta - B_n \cos \alpha)^2 + (h - (H + B_n \sin \alpha))^2} \right), \quad (5.18)$$

where λ is the wavelength, R the slant range distance, θ the off-nadir angle, H the altitude of the master antenna, B_n the baseline between the master antenna (which is assumed to be emitting) and the receiving antenna n , and α the baseline inclination.

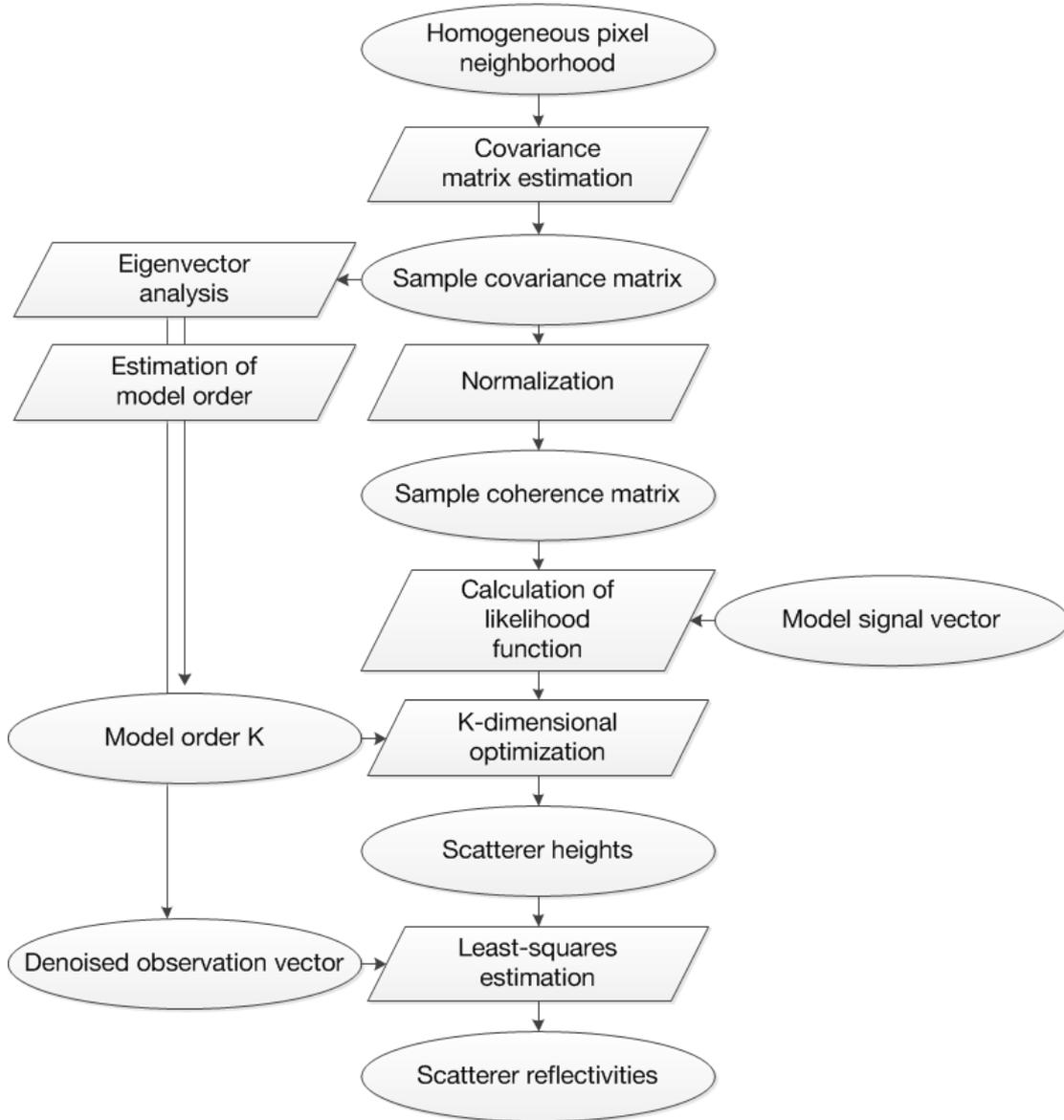


Figure 5.2: Flowchart of the Maximum-Likelihood-based TomoSAR algorithm.

As can be seen from (5.16), the model vector $\tilde{\mathbf{z}}(\mathbf{h})$ resembles the steering vector from (5.3) and (5.4), respectively. The main difference is that it contains a mixture of \hat{K} phase contributions as they appear when different scatterers collapse in a layover resolution cell. Based on this formulation, the likelihood function

$$\mathcal{L}(\hat{\mathbf{\Gamma}}; \tilde{\mathbf{z}}(\mathbf{h})) = \frac{1}{\pi^N \det(\hat{\mathbf{\Gamma}})} \exp\left(-\tilde{\mathbf{z}}^H(\mathbf{h}) \hat{\mathbf{\Gamma}}^{-1} \tilde{\mathbf{z}}(\mathbf{h})\right) \quad (5.19)$$

is created, where

$$\hat{\mathbf{\Gamma}} = \begin{bmatrix} 1 & \hat{\gamma}_{12} & \dots & \hat{\gamma}_{1N} \\ \hat{\gamma}_{12}^* & 1 & \dots & \hat{\gamma}_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{\gamma}_{1N}^* & \hat{\gamma}_{2N}^* & \dots & 1 \end{bmatrix} \quad (5.20)$$

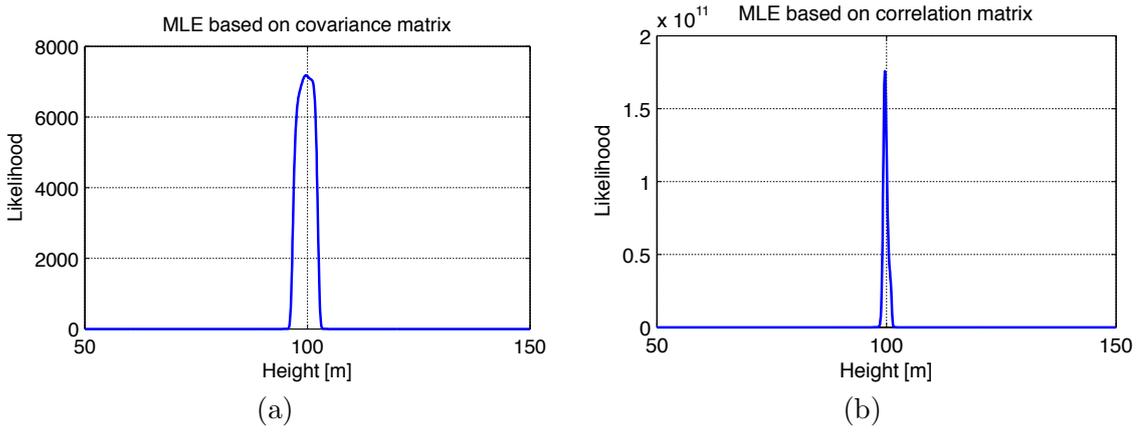


Figure 5.3: Simulated peak of the likelihood function given in (5.19) for one scatterer at $h = 100$ m if (a) the covariance matrix and (b) the correlation matrix is used in the maximum-likelihood estimation.

is the estimated version of (4.5), i.e. the sample coherence matrix. This normalization is necessary since $\tilde{\mathbf{z}}$ does not contain any information about the intensity values of the observations. The benefit of this normalization is illustrated in Fig. 5.3.

The principle of the proposed method is to maximize (5.19) with respect to the \hat{K} unknown scatterer heights h_k . However, De Zan [2008] has shown in a different context that the determinant in (5.19) is constant during the iterations of the maximum search, wherefore the term can be discarded. The simplified objective function to be maximized then becomes:

$$\hat{\mathbf{h}} = \underset{\mathbf{h} \in [h_{min}, h_{max}]}{\arg \max} \exp \left(-\tilde{\mathbf{z}}^H(\mathbf{h}) \hat{\mathbf{\Gamma}}^{-1} \tilde{\mathbf{z}}(\mathbf{h}) \right), \quad (5.21)$$

where $\hat{\mathbf{h}} = [\hat{h}_1, \dots, \hat{h}_K]^T$.

Instead of the one-dimensional search with following peak detection that is employed for conventional spectral estimation methods such as Capon beamforming [Capon, 1969] or MUSIC [Schmidt, 1986], here a \hat{K} -dimensional optimization needs to be solved. As no analytical solution to this optimization has been found yet, it can be carried out by a simple grid search: For this, the objective function in (5.21) is calculated for \hat{K} unknown heights, each hypothesized with respect to a discretized search interval. Then the maximum is found in the \hat{K} -dimensional search space aggregated from the \hat{K} search intervals. An example for $\hat{K} = 2$ is illustrated in Fig. 5.4. The dependence of the optimization problem on prior knowledge of the model order (i.e. the number of scatterers in the resolution cell under investigation) introduces the need to automatic model order selection (see Section 5.2).

5.3.2 Reflectivity Estimation

If SAR tomography is not only seen as an extension of SAR interferometry aiming at the separation of scatterers that have been projected into a common resolution cell, but as a means to carry out three-dimensional focusing in order to create a real 3D image, not only the scatterer heights but also their reflectivities have to be estimated. This can be achieved by further exploitation of the sample covariance matrix: First, the eigen-decomposition of $\hat{\mathbf{C}}$ has to be applied, such that

$$\hat{\mathbf{C}} = \hat{\mathbf{E}} \hat{\mathbf{\Lambda}} \hat{\mathbf{E}}^{-1}, \quad (5.22)$$

where $\hat{\mathbf{E}} = [\hat{\mathbf{E}}_1, \dots, \hat{\mathbf{E}}_N]$ is the matrix of the N eigenvectors $\hat{\mathbf{E}}_n$ and $\hat{\mathbf{\Lambda}} = \text{diag}(\hat{\lambda}_1, \dots, \hat{\lambda}_N)$ is the matrix containing the corresponding eigenvalues on the main diagonal, sorted in descending

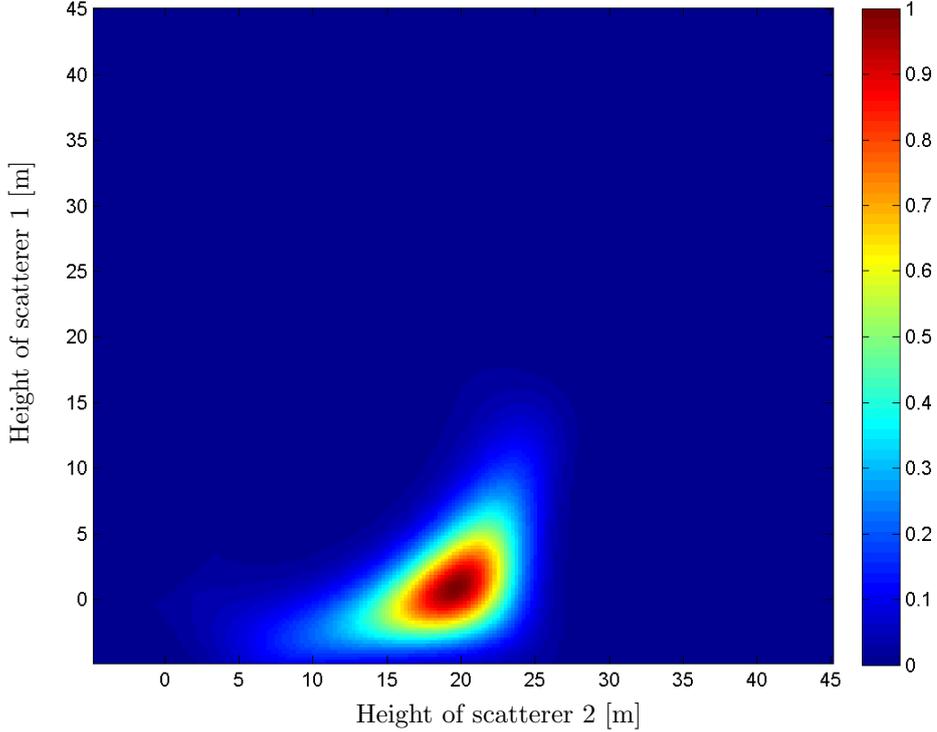


Figure 5.4: Two-dimensional objective function for two simulated scatterers at $h_1 = 0$ m and $h_2 = 20$ m. The function is normalized to its maximum.

order. Then, we add up the signal space components – which are defined by the number of scattering contributions \hat{K} – in order to create a denoised version of the SLC resolution cell under investigation by

$$\hat{\mathbf{z}} = \sum_{k=1}^K \sqrt{\hat{\lambda}_k} \cdot \hat{\mathbf{E}}_k. \quad (5.23)$$

In this way, a denoised version of the actual observation vector \mathbf{z} is received, which is supposed to contain only signal information and no noise contribution anymore. A similar concept is also employed in the context of principal component analysis [Shlens, 2009].

The complex reflectivities of the \hat{K} scatterers can then be estimated by solving the least-squares problem

$$\|\mathbf{A}\mathbf{x} - \hat{\mathbf{z}}\|_2 \rightarrow \min \quad (5.24)$$

by

$$\hat{\mathbf{x}} = (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \hat{\mathbf{z}}, \quad (5.25)$$

where

$$\mathbf{A} = [\exp(j\varphi_1), \dots, \exp(j\varphi_K)], \quad (5.26)$$

and

$$\varphi_k = [\varphi_{1,k}, \dots, \varphi_{N,k}]^T \quad (5.27)$$

is calculated like described in (5.17).

Since the reflectivities are only estimated for \hat{K} discrete scatterers, it is obvious that only sparsely populated reflectivity profiles can be reconstructed with this algorithm at reasonable computational cost. Similar to the popular compressive-sensing based techniques [Schmitt & Stilla, 2013], this favors especially the application to urban areas.

6 Fusion of Multi-Aspect InSAR Datasets

While Chapter 5 describes a maximum-likelihood-based formulation for layover separation by single-pass SAR tomography, this chapter aims at the fusion of (multi-baseline) InSAR data from multiple viewing directions, i.e. aspects. First, a new method for multi-aspect SAR image registration exploiting the radargrammetric imaging geometry and geodetic least-squares adjustment is proposed, which can both be employed for amplitude image fusion and as a pre-processing step for the remaining height reconstruction tasks. Then the fusion of InSAR or TomoSAR observations by conventional forward geocoding is described, before a statistical estimation procedure for the simultaneous fusion of multi-baseline data acquired from multiple aspects is developed exploiting a backward geocoding framework.

6.1 Radargrammetric Multi-Aspect SAR Image Registration

For multi-aspect SAR data fusion, a rigid link between the individual datasets has to be established. This holds independently of the specific nature of the SAR data – amplitude/intensity imagery, InSAR data, etc. In all cases, the side-looking imaging geometry leads to severely different appearances of the imaged objects for every viewing angle, making multi-aspect registration a non-trivial task especially for densely built-up inner city areas.

Building upon the insight that a fusion of decimeter-resolution multi-aspect SAR data of urban areas is only reasonably possible in object space, in [Schmitt et al., 2013c] a method is presented that combines the well-known calibration of the SAR flight geometry using one or more known ground control points [Wu & Lin, 2000; Sohn et al., 2005; Leberl, 1990] with a total least squares (TLS) approach formulated as a strict Gauss-Helmert model [Neitzel & Petrovic, 2008; Neitzel, 2010; Lenzmann & Lenzmann, 2004] and the simultaneous adjustment of an arbitrary number of acquisitions of the same urban scene taken from arbitrarily oriented flight tracks. The proposed approach does not use any approximations or assumptions with respect to the projection of the imagery to be registered; also it does not require a preliminary orthoprojection or similar pre-processing step. Although the mathematical model was derived for linear flight tracks and zero-Doppler processed data, the algorithm can straight-forwardly be extended to more general imaging geometries. Even if the method was originally designed for the co-registration and fusion of SAR intensity images with respect to a reference plane (see Fig. 6.1 for an example), it can also be used as a pre-processing step for multi-aspect InSAR data fusion: First, interferometric or tomographic point clouds are calculated in the master images' slant range geometries, then the registration algorithm is applied on the corresponding trajectories. Using the resulting registered trajectory data, each aspect can be geocoded to a common reference frame.

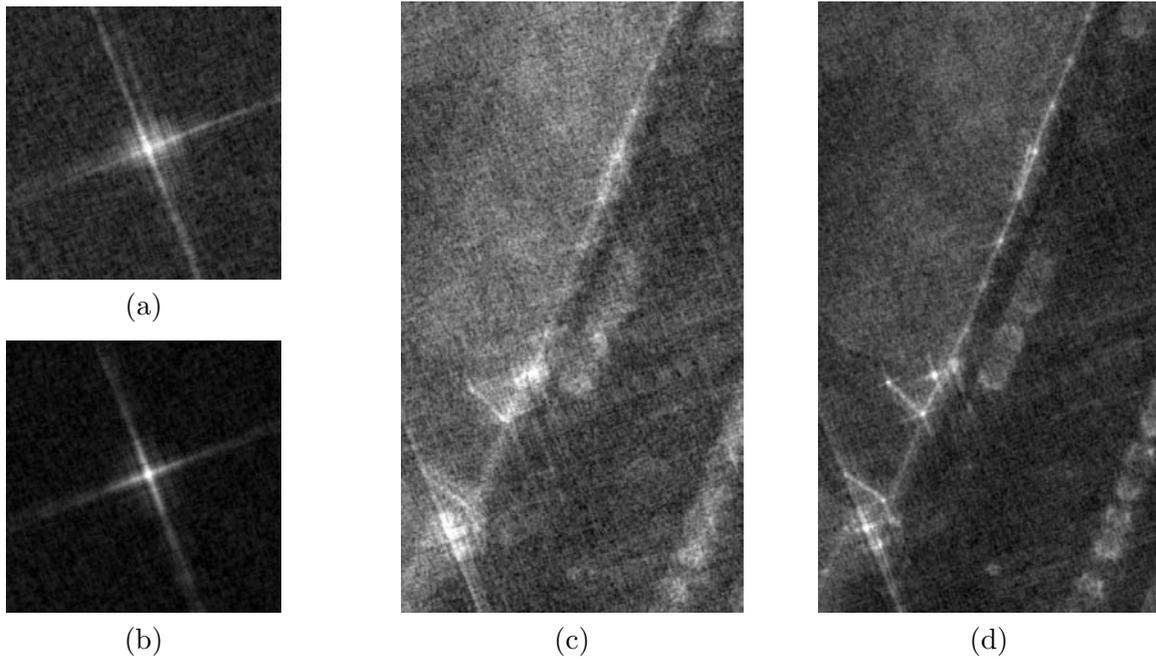


Figure 6.1: Exemplary illustrations of radargrammetric registration results: Comparison of incoherent mean maps of a corner reflector (a) before and (b) after radargrammetric registration. Obviously the reflector appears much more focused in the registered case. Image (c) shows a part of a road with some vehicles and a hoarding along the left side of the road. It can be seen that the level of detail is increased significantly in the registered case (d) making it easier to recognize the different objects.

6.1.1 Radargrammetric calibration in a strict Gauss-Helmert model

Papers explaining the radargrammetric calibration of SAR imaging geometries using one or more ground control points have been published for decades [Schreier, 1993]. While it is not uncommon that least-squares based approaches are employed, e.g. [Hellwich & Ebner, 2000; Raggam & Gutjahr, 2000; Yue et al., 2008], to the author’s knowledge [Schmitt et al., 2013c] was the first publication to propose a solution based on adjusting only observations in a strict Gauss-Helmert model. The formulation of the well-known range-Doppler approach in the framework of this sophisticated model is shown. The main advantage of this estimation method is that it does not depend on the introduction of any unknown parameters and just seeks to estimate corrections that adjust the observations in a least squares sense [Mikhail, 1976, pp. 137 ff.]. One of the core advantages of this implementation is the fact that an arbitrary number of flight tracks can be calibrated simultaneously without any loss in redundancy: Usually, the parameters describing the flight track are considered as unknowns in the calibration, such that each track would introduce new unknowns and therefore reduce the available redundancy. Due to that, the number of necessary ground control points and the number of tie points that link the images one to another increases. In contrast to that, in the proposed formulation for every flight track that is included in the estimation, the redundancy is even enhanced, because only additional conditional equations without any unknowns are introduced. An additional benefit is that in a simultaneous adjustment, all flight tracks will be linked via the ground control points that have jointly been used for the calibration. Thus, a precise combination of multi-aspect SAR and InSAR data of urban areas is enabled, an otherwise non-trivial task.

6.1.2 Recapitulation of range-Doppler geometry

As mentioned before, SAR imaging geometry calibration routines are usually based on the well-known range-Doppler equations [Leberl, 1990], which – in the airborne case – can be combined with the assumption of linear flight trajectories and zero-Doppler processed data. This is justified for airborne SAR data with limited swath length and an appropriate focusing of the raw radar data [Magnard et al., 2012].

First, the range and azimuth coordinates of any ground control point measured in the scene can be derived from its pixel position using the linear relations

$$t = t_0 + dt(r - 1) \quad (6.1)$$

and

$$R = R_0 + dR(c - 1), \quad (6.2)$$

respectively. Here, t_0 denotes the time at which the first azimuth bin was imaged, whereas R_0 is the near-range value of the system given in slant range geometry. dt and dR denote the pixel spacing in azimuth (time) and range direction, and r and c are the row (azimuth) and column (range) coordinates of the control point, counted from the center of the top left image pixel.

Using (6.1), the three-dimensional sensor position during each azimuth bin can be derived:

$$\mathbf{s}(t) = \mathbf{s}_0 + \mathbf{v} \cdot t \quad (6.3)$$

Analogous to t_0 , \mathbf{s}_0 describes the sensor position during the acquisition of the first azimuth bin, and \mathbf{v} describes the sensor velocity. Equations (6.2) and (6.3) are then put into the range and Doppler equations

$$R - \|\mathbf{p} - \mathbf{s}(t)\| = 0 \quad (6.4)$$

$$\mathbf{v}(\mathbf{p} - \mathbf{s}(t)) = 0, \quad (6.5)$$

where \mathbf{p} is the 3D position of the ground control point in world coordinates. Again it is important to note that (6.5) is a simplified Doppler equation that only holds for zero-Doppler processed data. It can, however, easily be extended to other types of SAR processing without loss of generality for the reasoning of this description.

If (6.4) and (6.5) are set up for every ground control point in every scene, a non-linear system of conditional equations is constructed. Since all included parameters are typically known from the flight navigation control and the SAR focusing, respectively, it is convenient to model the correction of the sensor geometry by a Gauss-Helmert model without unknowns [Mikhail, 1976, pp. 137 ff.].

6.1.3 Parameter estimation in the strict Gauss-Helmert model

For the parameter estimation, the system of conditions consists of the the range and Doppler equations (6.4) and (6.5). They depend not only on the flight track, but also on the positions of the ground control points and their corresponding image positions. Therefore, the goal is to achieve

$$\mathbf{f}(\hat{\mathbf{v}}) = \begin{pmatrix} \mathbf{f}_1(\mathbf{b} + \hat{\mathbf{v}}) \\ \vdots \\ \mathbf{f}_Q(\mathbf{b} + \hat{\mathbf{v}}) \end{pmatrix} = \mathbf{0}, \quad (6.6)$$

		s_{0X}	s_{0Y}	s_{0Z}	v_X	v_Y	v_Z	t_0	dt	R_0	dR	r_{p1}	c_{p1}	r_{p2}	c_{p2}	p_{1X}	p_{1Y}	p_{1Z}	p_{2X}	p_{2Y}	p_{2Z}
\mathbf{p}_1	Range																				
	Doppler																				
\mathbf{p}_2	Range																				
	Doppler																				

Figure 6.2: Exemplary structure of the Jacobian matrix for the case of just a single flight track with two ground control points (\mathbf{p}_1 , \mathbf{p}_2) imaged. The magnitude of the partial derivatives is shown from white (no derivative) to black (very large derivative).

where Q is the number of flight tracks and $\hat{\mathbf{b}} = \mathbf{b} + \hat{\boldsymbol{\nu}}$ are the corrected observations. The vector of conditional equations for a single flight track $q \in \{1, \dots, Q\}$ is

$$\mathbf{f}_q(\hat{\boldsymbol{\nu}}) = \begin{pmatrix} f_{q,r}^{(1)}(\mathbf{b} + \hat{\boldsymbol{\nu}}) \\ f_{q,d}^{(1)}(\mathbf{b} + \hat{\boldsymbol{\nu}}) \\ \vdots \\ f_{q,r}^{(P)}(\mathbf{b} + \hat{\boldsymbol{\nu}}) \\ f_{q,d}^{(P)}(\mathbf{b} + \hat{\boldsymbol{\nu}}) \end{pmatrix}, \quad (6.7)$$

with P being the number of ground control points and $f_{q,r}^{(p)}$ and $f_{q,d}^{(p)}$ representing (6.4) and (6.5) for flight track q and ground control point $p \in \{1, \dots, P\}$. Equation (6.6) is the objective that is to be achieved by calculating the correction vector $\hat{\boldsymbol{\nu}}$.

In general, for the application of a Gauss-Helmert model, the conditional equations have to be linearized, i.e. a Jacobian matrix containing the partial derivatives of the conditional equations with respect to all the corrections is needed:

$$\mathbf{B} = \frac{\partial \mathbf{f}(\boldsymbol{\nu})}{\partial \boldsymbol{\nu}}. \quad (6.8)$$

\mathbf{f} denotes the vector of conditional equations, $\boldsymbol{\nu}$ the vector of corrections. A visualization of the structure of the Jacobian matrix for a typical configuration is illustrated in Fig. 6.2.

These corrections $\boldsymbol{\nu}$ need to be estimated from the linearized set of conditional equations

$$\mathbf{B}\hat{\boldsymbol{\nu}} + \mathbf{w} = \mathbf{0} \quad (6.9)$$

based on the least squares principle. In most of the common literature on least squares estimation, the linearization is only carried out at the position $\boldsymbol{\nu}_0 = \mathbf{0}$, leading to the vector of contradictions

$$\mathbf{w} = \mathbf{f}(\boldsymbol{\nu}_0). \quad (6.10)$$

The proposed method, however, follows the approach of Lenzmann & Lenzmann [2004] and thus renews the linearization for every update of $\boldsymbol{\nu}_0$, which yields the actual vector of contradictions

$$\mathbf{w} = -\mathbf{B}\boldsymbol{\nu}_0 + \mathbf{f}(\boldsymbol{\nu}_0) \quad (6.11)$$

that has to be recalculated in every iteration of the minimization of the objective function. If (6.11) is used instead of (6.10), this is called ‘‘strict’’ Gauss-Helmert model in the rare literature about the topic and was shown to equal common Total Least Squares solutions [Neitzel & Petrovic, 2008], [Neitzel, 2010].

The afore-mentioned objective function is in both cases defined as

$$\Omega = \hat{\boldsymbol{\nu}}^T \mathbf{Q}_{bb}^{-1} \hat{\boldsymbol{\nu}} - 2\hat{\mathbf{k}}^T (\mathbf{B}\hat{\boldsymbol{\nu}} + \mathbf{w}), \quad (6.12)$$

where \mathbf{Q}_{bb} is the cofactor matrix of the observations that can be derived from their covariance matrix and is used for the weighting of the observations. If \mathbf{Q}_{bb} is the unity matrix, an unweighted least squares solution is found. $\hat{\mathbf{k}}$ is the vector of correlates and calculated by

$$\hat{\mathbf{k}} = -(\mathbf{B}\mathbf{Q}_{bb}\mathbf{B}^T)^{-1} \mathbf{w}. \quad (6.13)$$

If (6.12) is minimized with respect to $\hat{\boldsymbol{\nu}}$ and $\hat{\mathbf{k}}$, this finally leads to the least squares estimator for the vector of corrections:

$$\hat{\boldsymbol{\nu}} = \mathbf{Q}_{bb}\mathbf{B}^T\hat{\mathbf{k}}. \quad (6.14)$$

If $\mathbf{f}(\hat{\boldsymbol{\nu}})$ is fulfilled to a satisfying precision (e.g. 10^{-6}), the iteration can be aborted and the final solution is found. Otherwise, $\boldsymbol{\nu}_0 = \hat{\boldsymbol{\nu}}$ is used for the next iteration and a new set-up of (6.8) and (6.11).

With the estimated vector of corrections $\hat{\boldsymbol{\nu}}$, the adjusted observation vector

$$\hat{\mathbf{b}} = \mathbf{b} + \hat{\boldsymbol{\nu}} \quad (6.15)$$

can be derived easily. The main benefit of the strict formulation of the Gauss-Helmert model is that the conditional equations will be fulfilled to almost arbitrary precision after the adjustment.

Implementation of a step size parameter

If the conditional equations (6.9) are fulfilled, the solution of the objective function (6.12) is given by

$$\Omega_{\min} = \hat{\boldsymbol{\nu}}^T \mathbf{Q}^{-1} \hat{\boldsymbol{\nu}} = -\mathbf{k}^T \hat{\mathbf{w}} \quad (6.16)$$

Obviously the solution depends on the vector of contradictions. From experience, however, with the applied mathematical model and the described optimization, chances are high that the solution is oscillating. Therefore, the utilization of a suitable step size parameter is proposed: $\boldsymbol{\nu}_0$ in (6.11) is replaced by $\alpha\boldsymbol{\nu}_0$ constrained to $\alpha \geq 0$. The objective function to determine the optimal step size parameter α for updating the contradictions in every iteration is given by

$$\|\mathbf{w}\|_2^2 \rightarrow \min. \quad (6.17)$$

This objective enables to construct a solution in which the norm of the iteratively updated contradiction vector is a monotonically decreasing sequence in the number of iterations. Objective (6.17) is a simple one-dimensional optimization problem with the analytical solution

$$\alpha = \frac{\mathbf{f}^T \mathbf{p}}{\mathbf{p}^T \mathbf{p}} \quad (6.18)$$

where $\mathbf{p} = \mathbf{B}\boldsymbol{\nu}_0$ given that the Jacobian \mathbf{B} and the conditional equations \mathbf{f} are evaluated at $\boldsymbol{\nu}_0$ and assumed to be independent of $\boldsymbol{\nu}$ and therefore constant during one iteration. The latter is only fulfilled if \mathbf{B} and \mathbf{f} are linear in $\boldsymbol{\nu}$ but in practice this procedure works very well leading to a solution in only a few iterations. The reason for the introduction of the step size parameter α constraining the iterative updates is that the linear system (6.11) shows a significant sensitivity according to small perturbations in \mathbf{B} and $\boldsymbol{\nu}$. As stated in [Golub & Van Loan, 1996, pp. 80 ff.] the condition number $\kappa(\mathbf{B}) = \lambda_{\max}/\lambda_{\min}$, where λ are the singular values of \mathbf{B} , is a measure for this kind of sensitivity. It indicates the amount of amplification of \mathbf{w} for a small perturbation of $\boldsymbol{\nu}$. In our case it is especially the derivation of the Doppler equation with respect to the azimuth (time) pixel spacing dt which introduces significant imbalances in the singular values of the Jacobian matrix (cf. Fig. 6.2) leading to $\kappa(\mathbf{B}) \gg 10^6$.

Observation	Unit	Order of magnitude	Standard deviation
<i>Sensor parameters</i>			
Sensor starting position, $\mathbf{s}_0 = [s_{0X}, s_{0Y}, s_{0Z}]^T$	<i>m</i>	10^6	0.1
Sensor velocity, $\mathbf{v} = [v_X, v_Y, v_Z]^T$	$\frac{m}{s}$	10^1	0.02
Starting time, t_0	<i>s</i>	10^{-3}	0
Azimuth pixel spacing (time), dt	<i>s</i>	10^{-4}	0
Near range, R_0	<i>m</i>	10^3	0.02
Range pixel spacing (distance), dR	<i>m</i>	10^{-1}	0
<i>Ground control point measurements</i>			
GCP position in WGS84, $\mathbf{p} = [p_X, p_Y, p_Z]^T$	<i>m</i>	10^6	0.03
GCP image measurements, r, c	pixel	10^4	0.5

Table 6.1: Observation Types in radargrammetric calibration

Covariance matrix of observations

As mentioned before, it is possible to accomplish a weighted least squares solution using the cofactor matrix \mathbf{Q}_{bb} derived from the covariance matrix \mathbf{C}_{bb} of the observations by

$$\mathbf{Q}_{bb} = \frac{1}{\sigma_0^2} \mathbf{C}_{bb}. \quad (6.19)$$

A priori, the so-called variance factor σ_0^2 is usually set to 1. Its actual value is estimated after the adjustment. In general, a weighting of the observations can always be considered reasonable, especially if they are very heterogeneous in their nature. In the functional model employed for this study, observations of different units and different orders of magnitude are combined, e.g. pixel coordinates, world coordinates, time measurements, and velocities (see Table 6.1). Furthermore, the stochastic model can also be used to exclude certain observations from the least squares adjustment if they are to be considered constant and their variance is set to 0 (or rather a small number $\epsilon > 0$ for computational reasons).

Posterior statistical evaluation

Based on variance-covariance propagation, it is possible to deduce the covariance matrix of the corrections and the adjusted observations after the parameter estimation. The covariance matrix of the corrections is found to be

$$\mathbf{C}_{\hat{\nu}} = \hat{\sigma}_0^2 \cdot \mathbf{Q}_{bb} \mathbf{B}^T (\mathbf{B} \mathbf{Q}_{bb} \mathbf{B}^T)^{-1} \mathbf{B} \mathbf{Q}_{bb}, \quad (6.20)$$

the covariance matrix of the adjusted observations to be

$$\mathbf{C}_{\hat{b}} = \hat{\sigma}_0^2 \cdot \left(\mathbf{Q}_{bb} - \mathbf{Q}_{bb} \mathbf{B}^T (\mathbf{B} \mathbf{Q}_{bb} \mathbf{B}^T)^{-1} \mathbf{B} \mathbf{Q}_{bb} \right) \quad (6.21)$$

with the posterior variance factor

$$\hat{\sigma}_0^2 = \frac{\hat{\nu}^T \mathbf{Q}_{bb}^{-1} \hat{\nu}}{r}. \quad (6.22)$$

r is defined as the redundancy of the equation system, which equals the number of conditional equations for the Gauss-Helmert model without unknowns. It has to be noted that the variance factor also contains an assertion about the a priori stochastic model that has been employed by the relation given already in (6.19): Using (6.22), the “true” covariance matrix of original observations can be calculated by

$$\hat{\mathbf{C}}_{bb} = \hat{\sigma}_0^2 \cdot \mathbf{Q}_{bb}. \quad (6.23)$$

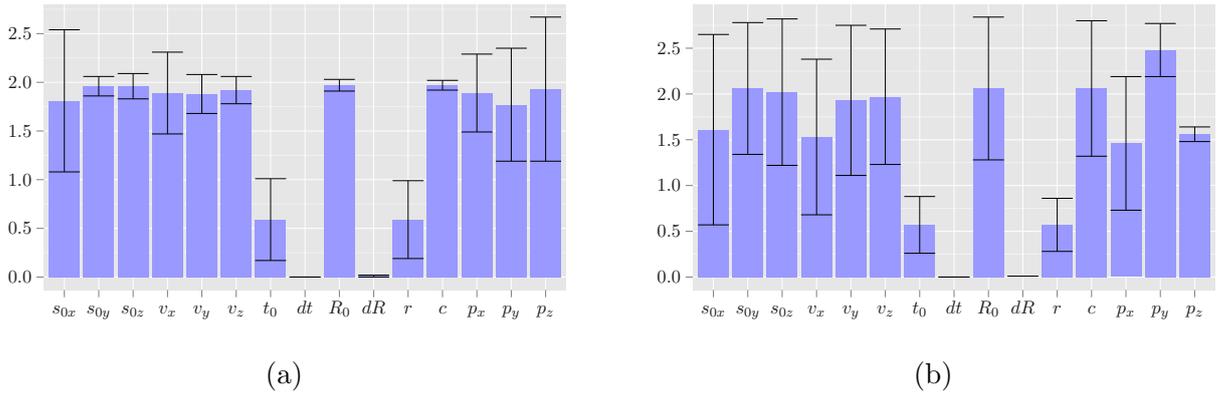


Figure 6.3: Significance of the individual parameters used in the radargrammetric calibration: Median t -values for the unweighted (a) separate and (b) joint calibration of 5 datasets acquired from different viewing angles. The black error bars indicate the standard deviations of the t -values.

This means: The closer $\hat{\sigma}_0^2$ is to 1, the better were the stochastic assumptions about the original observations.

Based on (6.20) and (6.21) the relation between corrections and the respective standard deviations can be described by the so-called t -value

$$t = \left| \frac{\hat{v}}{\sigma_{\hat{v}}} \right|, \quad (6.24)$$

for every parameter with the standard deviation having been extracted from the covariance matrix of corrections. The higher the t -value is, the more significant is the corresponding correction. Using this evaluation strategy and additional prior knowledge on SAR processing standards, the experiments in [Schmitt et al., 2013c] came to the conclusion that t_0 , dt and dR should be considered as constant parameters, whose values should not be adjusted in the registration process (cf. Fig. 6.3). The suggested weights for the remaining parameters can be found in Table 6.1.

After the simultaneous calibration of data acquired from multiple aspect angles via a common ground control point, all aspects are registered by their adjusted flight navigation parameters. If these updated parameters are then used during geocoding procedures, all data will be fused in the same reference system showing only minimal residual errors, if any.

6.2 Fusion Based on Interferometric Forward Geocoding

The most conventional approach to generate 3D point clouds from InSAR data is by interferometric geocoding of single-baseline data. This is done by solving the non-linear interferometric range-Doppler equations (IRDE) that directly relate the InSAR measurements to geocoded three-dimensional object points [Hellwich & Ebner, 2000]. In case of a linear flight track and zero-Doppler processed data, the IRDE can simply be written as:

$$R = \|\mathbf{p} - \mathbf{s}_M(t)\| \quad (6.25)$$

$$\mathbf{v}_M(\mathbf{p} - \mathbf{s}_M(t)) = 0 \quad (6.26)$$

$$\phi = -\frac{2\pi}{\lambda} (\|\mathbf{p} - \mathbf{s}_M(t)\| - \|\mathbf{p} - \mathbf{s}_{SI}(t)\|) \quad (6.27)$$

R is the slant range distance between the object point $\mathbf{p} = [p_X, p_Y, p_Z]^T$ and the master antenna position $\mathbf{s}_M(t) = [s_{MX}(t), s_{MY}(t), s_{MZ}(t)]^T$ at time t . $\mathbf{v}_M = [v_{MX}, v_{MY}, v_{MZ}]^T$ denotes the

velocity of the master antenna, which is assumed to be constant. $\mathbf{s}_{Sl}(t)$ symbolizes the slave antenna position analogue to $\mathbf{s}_M(t)$, while ϕ denotes the absolute interferometric phase and λ the wavelength of the radar system.

If (6.25)-(6.27) are solved for the unknown object coordinates \mathbf{p} for each pixel in the interferogram, the result is an irregular 3D point cloud that suffers from both layover, which causes the so-called front-porch effect [Burkhart et al., 1996; Thiele et al., 2007b], as well as shadowing in the form of interferometric phase values that contain only noise, leading to random 3D points. Furthermore, the phase-unwrapping operation, which is needed in order to extract the absolute phase ϕ from the wrapped phase values of the complex interferogram is still a difficult, error-prone task [Ghiglia & Pritt, 1998].

Besides this IRDE-solution, another approach for this task is to first carry out phase-to-height conversion in slant range geometry, before the resulting height data are transformed from the slant range plane to a geodetic reference frame [Small et al., 1996]. This can also be applied for TomoSAR point clouds as generated by the method described in Section 5.3.1. In this case, equation (6.27) is replaced by

$$\frac{p_X^2 + p_Y^2}{(a+h)^2} + \frac{p_Z^2}{(b+h)^2} - 1 = 0, \quad (6.28)$$

where a and b are the semi-major and semi-minor axes of the reference Earth ellipsoid, respectively, and h is the previously reconstructed height [Schwäbisch, 1998]. Again this system of three equations can be solved for $\mathbf{p} = [p_X, p_Y, p_Z]^T$ using an iterative optimization algorithm.

If registered (cf. Section 6.1) multi-baseline or multi-aspect data are available, several separate point clouds can be generated based on this proceeding, all of which are geocoded to the same common reference system. Due to the irregular structure of these point clouds and their rather noisy nature, however, a beneficial fusion of possible redundant measurements of the same real-world object points is usually not considered.

Point Cloud Fusion by Voxelization

After having geocoded point clouds resulting from interferometric or tomographic height reconstruction to a common world coordinate system, it can be desirable to fuse the individual point clouds. Reasons for that desire might be e.g. the need to reduce the number of points (e.g. due to memory requirements) or the goal to fuse redundantly reconstructed points belonging to the same scene feature in order to receive a more robust 3D result. Another motivation might be the wish to store the point cloud in a regular raster instead of an unstructured point list.

In the framework of this thesis, all these goals are achieved by a voxelization of the point cloud using a combination of a k -d tree and a range query based on the Chebyshev distance. A k -d tree is a multidimensional binary search tree, which serves as a space-partitioning data structure for storage of information to be retrieved by associative searches in a k -dimensional space [Bentley, 1975]: If data is represented as a k -d tree, then each data point is stored as a node in the tree. Every non-leaf node implicitly generates a splitting hyperplane that divides the space into two parts (i.e. half-spaces). Points to the left of this hyperplane are then represented by the left subtree of that node, and points on the right are represented by the right subtree. Since every node in the tree is associated with one of the k dimensions the hyperplane is chosen perpendicular to that dimension's axis. Each split can be denoted using the dimension number and split value, whereas the splits are arranged in order to balance the tree, i.e. its maximum depth is kept as small as possible. If points are queried, the k -d tree search first locates the respective point in its appropriate node and then searches nearby leaves in the tree until it can guarantee that the correct

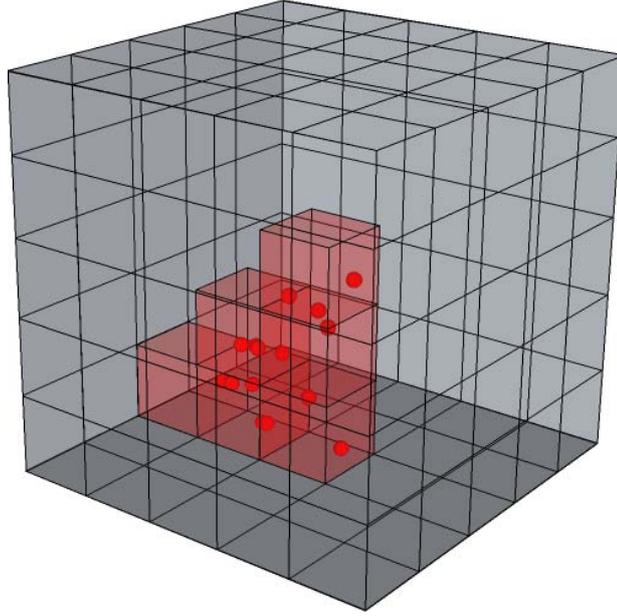


Figure 6.4: Voxelization of an irregular point cloud: A regular voxel space is created, which is imposed on the 3D point cloud. The red voxels symbolize voxels that contain at least one 3D point.

point has been found. Due to their efficient storage structure, k -d trees are particularly useful for multi-dimensional search operations such as range searches or nearest neighbor searches. Besides, an abundance of ready-to-use implementations in different programming languages is available online, many of them open source.

For the task of voxelizing unstructured 3D point clouds, first a regular point cloud describing the centers of the voxel space has to be created. These voxel centers are stored in a k -d tree of dimension $k = 3$. Then, the irregularly sampled InSAR or TomoSAR point cloud is used with this k -d tree in order to apply a range search based on the Chebyshev distance. This distance metric is defined for two points $\mathbf{p} = [p_1, \dots, p_n]^T$ and $\mathbf{q} = [q_1, \dots, q_n]^T$ as

$$\lim_{k \rightarrow \infty} \left(\sum_{i=1}^n |p_i - q_i|^k \right)^{\frac{1}{k}} = \max_{i \in \{1; \dots; n\}} (|p_i - q_i|). \quad (6.29)$$

This means, it defines the distance between \mathbf{p} and \mathbf{q} as the greatest of their differences along any coordinate dimension. For three-dimensional vectors and a Chebyshev distance of $\frac{d}{2}$, this corresponds to a cube with an edge length of d . Therefore, the distance query results in a list of voxels and the corresponding points from the unstructured point cloud contained in that voxel (see Fig. 6.4). With this information, two potential strategies become possible:

- 1) Simply activate each voxel containing at least a certain minimum number of points (e.g. one). This results in a thinned out regularly spaced point cloud, since each returned point has to equal one of the voxel centers.
- 2) Calculate the mean or median value of all points contained in each voxel. Also in this case a thinned out point cloud is returned, however, not regularly gridded.

Both strategies can straight-forwardly be used to fuse different point clouds from different sources. Especially for the fusion of InSAR and TomoSAR point clouds reconstructed from different aspects, a certain redundance can be expected in scene parts that were acquired by more

than one flight. Therefore, a possible extension of the two basic voxelization strategies for multi-aspect data is proposed: In this case, it is additionally ensured that only voxels are considered that contain points from at least a minimum number of aspects (e.g. two). Although this proceeding is not advisable in the the case of two opposing aspects, as the correspondingly opposing shadow areas would be lost, it provides a promising solution for full multi-aspect configurations that ensure the scene was illuminated from at least four complementary viewing directions.

6.3 Fusion Based on Interferometric Backward Geocoding

As already investigated by Schwäbisch [1998] and Eineder [2003], the geocoding of interferometric SAR data can also be carried out starting from a pre-defined surface grid in the reference world coordinate system and some prior knowledge about the height extension of the investigated scene. Using this information, for each grid cell a number of height hypotheses $\tilde{p}_Z \in [h_{min}; h_{max}]$ can be defined following a step size fitting to the expected height resolution of the SAR interferometer. Each hypothetical 3D point $\tilde{\mathbf{p}} = [p_X, p_Y, \tilde{p}_Z]^T$ that is created this way is then projected into the SAR image by directly solving the inverted Doppler equation

$$t = \frac{(\tilde{\mathbf{p}} - \mathbf{s}_M(t_0)) \mathbf{v}_M}{\|\mathbf{v}_M\|^2}, \quad (6.30)$$

for the azimuth time t , and straight-forward calculation of the respective range and phase values by (6.25) and (6.27). In this context, $\mathbf{s}_M(t_0)$ denotes the 3D position of the master antenna during acquisition of the first azimuth bin.

Then, the observed interferometric phase corresponding to the (t, R) SAR image coordinate is measured in the pre-processed interferogram, so that the phase error between the ideal (i.e. simulated) and the measured phase can be calculated by

$$\Delta\phi(\tilde{\mathbf{p}}) = \|\phi_{sim}(\tilde{\mathbf{p}}) - \phi_{meas}\|. \quad (6.31)$$

Finally, the height hypothesis with the minimum phase error $\Delta\phi(\tilde{\mathbf{p}})$ is chosen as height estimate of the respective grid element:

$$\hat{p}_Z = \arg \min_{p_Z \in [h_{min}; h_{max}]} \Delta\phi(\tilde{\mathbf{p}}). \quad (6.32)$$

In [Schmitt & Stilla, 2011], this backward geocoding procedure was extended to be applied for the fusion of InSAR data acquired from multiple aspect angles. This is accomplished by weighting the phase error with the corresponding squared coherence magnitude $|\gamma|$ as a measure for the quality of the interferometric phase observation and yields the weighted phase error

$$\Delta\phi_w(\tilde{\mathbf{p}}) = \frac{\sum_{q=1}^Q |\gamma_q|^2 \cdot \Delta\phi_q(\tilde{\mathbf{p}})}{\sum_{q=1}^Q |\gamma_q|^2}. \quad (6.33)$$

It has to be noted that this is a very general framework, which means that the $q = 1 \dots Q$ phase values and coherence magnitudes can address both multiple aspects as well as multiple baselines within each single aspect at the same time. In any case, the final height estimate is chosen by

$$\hat{p}_Z = \arg \min_{p_Z \in [h_{min}; h_{max}]} \Delta\phi_w(\tilde{\mathbf{p}}). \quad (6.34)$$

in analogy to (6.32).

Although a fusion of multi-aspect InSAR data by this method already leads to a significant reduction of surface model parts affected by radar shadowing [Schmitt et al., 2011], still some patches can be expected to remain impinged by this effect. This leads to surface grid cells in which the estimated height values are pure noise. Therefore, it is advisable to detect these unreliable height estimations and exclude them from the resulting point cloud. A reconstructed height \hat{p}_Z is considered a reliable estimate if the combined coherence magnitude of the single interferometric phase measurements exceeds a certain threshold:

$$\frac{1}{Q} \sum_{q=1}^Q |\gamma_q| > \varepsilon \quad (6.35)$$

The value of ε depends on the number of phase measurements Q and a trade-off between the resulting point density and point reliability.

6.4 Multi-Aspect Multi-Baseline SAR Interferometry

The backward geocoding based 3D reconstruction from InSAR data can also be cast in the framework of maximum-likelihood estimation. In [Schmitt & Stilla, 2014c], the basic idea, which was first proposed by Eineder & Adam [2005], was enhanced to a more general formulation based on the joint likelihood function of both multi-baseline and multi-aspect data that is also able to take correlations between several images acquired from the same aspect into account. This is especially necessary if the available data consists of single-pass multi-baseline InSAR stacks as they are frequently delivered by airborne sensors. In this case, a neglect of correlation between the observations of the different antennas is an invalid simplification [Meglio et al., 2006]. Apart from that, also this approach is based on the statistical assumptions described in Section 4.1. Besides the height reconstruction algorithm, a quality measure that can be used for the dismissal of unreliable height estimations in analogy to (6.35) was derived. Since the method enables the simultaneous fusion of arbitrary kinds of multi-aspect and multi-baseline InSAR data, it is called MAMBIN SAR (multi-aspect multi-baseline SAR interferometry) in the following.

6.4.1 Estimation Procedure

The estimation procedure is illustrated in Fig. 6.5: Analogue to the proceeding described in Section 6.3, the starting situation is a pre-defined surface grid in a world coordinate system. Again, a column of height hypotheses is run through for each grid element, and by means of the backward geocoding equations (6.25) and (6.30) each hypothesized grid point $\tilde{\mathbf{p}}$ is projected into one of the available SAR stacks, such that the corresponding sample covariance matrix $\hat{\mathbf{C}}$ can be obtained by bilinear interpolation of the sample covariance matrices of the neighboring pixels. In addition to that, in analogy to the TomoSAR procedure described in Section 5.3.1, a complex observation vector $\tilde{\mathbf{z}} = [\tilde{z}_1, \dots, \tilde{z}_N]^T$ is simulated for every height hypothesis \tilde{p}_Z by

$$\tilde{z}_n = \exp(j\varphi_{sim,n}), \quad (6.36)$$

where

$$\varphi_{sim,n} = -\frac{2\pi}{\lambda} (\|\tilde{\mathbf{p}} - \mathbf{s}_M\| + \|\tilde{\mathbf{p}} - \mathbf{s}_{Sl,n}\|), \quad (6.37)$$

and $n = 1 \dots N$ indexes the slave antennas while the master antenna $\mathbf{s}_M = \mathbf{s}_{Sl,1}$ is assumed to double-act as both emitter and receiver.

Also in this case, the sample covariance matrix must be normalized to achieve the sample coherence matrix described in (5.20), which does not contain any information about the backscattering intensities anymore.

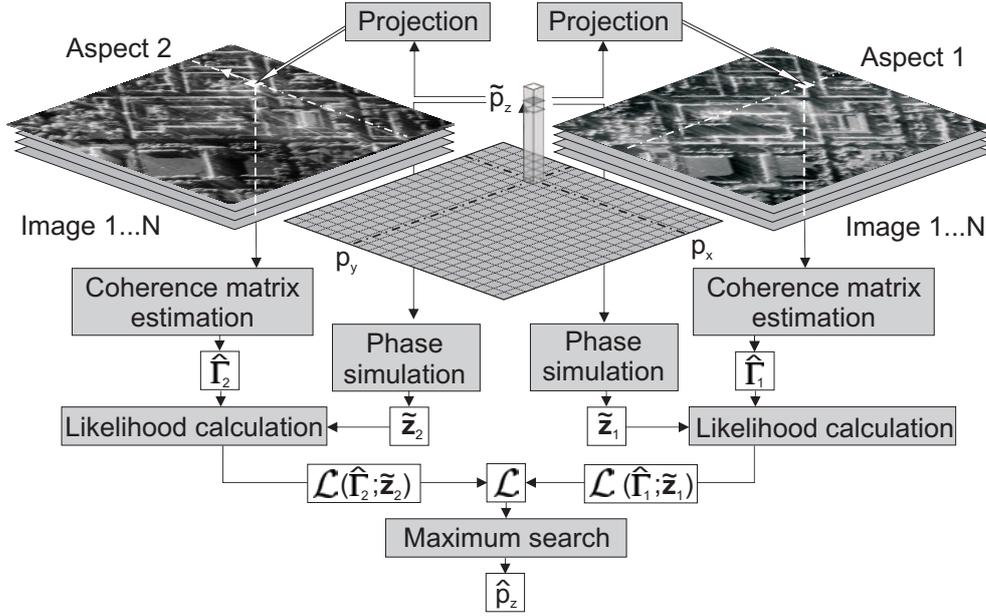


Figure 6.5: Flowchart of the proposed maximum-likelihood estimation procedure exploiting both multi-aspect as well as multi-baseline data. As an example, the case of $Q = 2$ aspects with stacks of $N = 4$ images per aspect is shown.

Eventually, this leads again to the likelihood function

$$\mathcal{L}(\hat{\mathbf{\Gamma}}; \tilde{\mathbf{z}}(\tilde{\mathbf{p}})) = \frac{1}{\pi^N \det(\hat{\mathbf{\Gamma}})} \exp\left(-\tilde{\mathbf{z}}^H(\tilde{\mathbf{p}}) \hat{\mathbf{\Gamma}}^{-1} \tilde{\mathbf{z}}(\tilde{\mathbf{p}})\right), \quad (6.38)$$

which only exploits complex coherence information. While this likelihood function resembles the one described in Section 6.4.1 (5.19), their main difference is the way $\tilde{\mathbf{z}}$ is created: For ML-TomoSAR, the model signal vector contains a mixture of several hypothetical scatterers, whereas the backward-geocoding-based nature of MAMBIInSAR, which is centered around a data fusion in object space, allows to consider only a single scatterer hypothesis per surface grid element. In addition, in the TomoSAR inversion algorithm always the same coherence matrix is used during the optimization as only one certain resolution cell is processed at a time. In contrast to that, the utilized coherence matrix may change during the MAMBIInSAR optimization due to the iterative projection of object space height hypotheses into the SAR image data, which usually leads to different image positions and thus different covariance matrices changing for each height iteration step.

Anyway, again (6.38) is evaluated by a simple grid search – albeit always one-dimensional in this case – for all height hypotheses $\tilde{p}_Z \in [h_{min}; h_{max}]$, such that the maximum-likelihood estimate of the desired height in a grid cell is found as

$$\hat{p}_Z = \arg \max_{p_Z \in [h_{min}; h_{max}]} \mathcal{L}(\hat{\mathbf{\Gamma}}; \tilde{\mathbf{z}}(\tilde{\mathbf{p}})). \quad (6.39)$$

This estimator can be applied to an arbitrary configuration of multi-baseline data. If individual baselines are to be neglected in the estimation, simply the respective columns and rows of $\tilde{\mathbf{z}}$ and $\hat{\mathbf{\Gamma}}$ can be eliminated, while N in (6.38) has to be reduced correspondingly. In this context, it should be mentioned that Corsini et al. [1999] and Ashok & Wilkinson [2001] have already shown that a combined evaluation of InSAR data from multiple baselines significantly extends the unambiguous height interval of the SAR interferometer (see Fig. 6.6). In addition to that, redundant phase

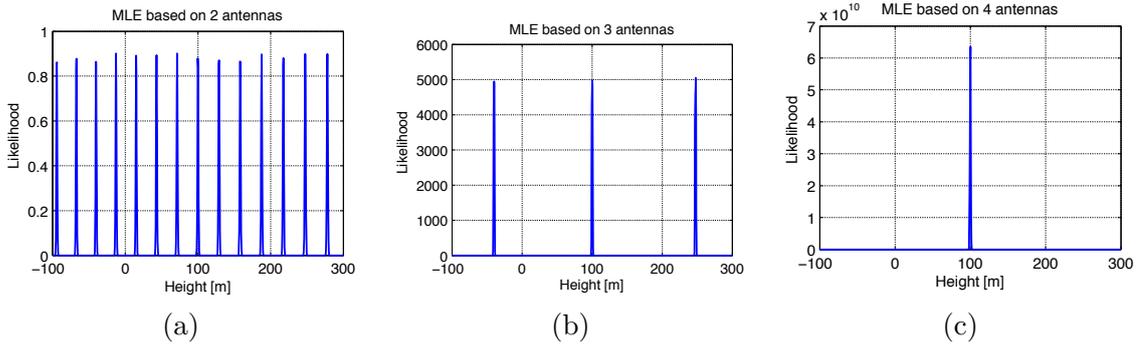


Figure 6.6: Likelihood functions of the MAMBIInSAR maximum-likelihood estimator based on (a) two, (b) three, and (c) four receiving antennas for a simulated scatterer located at a height of 100 m. Depending on the height search interval, the utilization of more than two antennas (i.e. more than one baseline) greatly helps to reduce phase ambiguities.

measurements help to reduce phase noise that is usually propagated to the reconstructed height values.

The core of this work, however, is the interferometric data fusion with respect not only to multi-baseline, but also to multi-aspect datasets; Fig. 6.5 exemplarily shows a case where $Q = 2$ aspects, each consisting of $N = 4$ images are utilized. For this purpose, the proposed framework can easily be extended: If multi-baseline stacks acquired from Q viewing directions are available, their joint likelihood function is found to be

$$\mathcal{L}(\hat{\mathbf{\Gamma}}_1, \dots, \hat{\mathbf{\Gamma}}_Q; \tilde{\mathbf{z}}_1(\tilde{\mathbf{p}}), \dots, \tilde{\mathbf{z}}_Q(\tilde{\mathbf{p}})) = \prod_{q=1}^Q \mathcal{L}(\hat{\mathbf{\Gamma}}_q; \tilde{\mathbf{z}}_q(\tilde{\mathbf{p}})), \quad (6.40)$$

as the acquisitions from the different viewing angles are considered as independent variables [Koch, 2007]. The maximum-likelihood estimator for a full multi-aspect multi-baseline InSAR dataset is then:

$$\hat{p}_Z = \arg \max_{p_Z \in [h_{min}; h_{max}]} \mathcal{L}(\hat{\mathbf{\Gamma}}_1, \dots, \hat{\mathbf{\Gamma}}_Q; \tilde{\mathbf{z}}_1(\tilde{\mathbf{p}}), \dots, \tilde{\mathbf{z}}_Q(\tilde{\mathbf{p}})). \quad (6.41)$$

This estimator allows a maximum of flexibility and can be applied to any kind of interferometric SAR data, no matter how many baselines per aspect or how many aspects are available. Beyond that, a densely populated probabilistic volume model (PVM) bearing resemblance to the work of Restrepo et al. [2012] is achieved as an interesting intermediate result (i.e. before application of (6.41)). An exemplary slice through such a PVM is shown in Fig. 6.7: High likelihood values around surfaces can be seen just like low values for the shadowed street canyons. Furthermore, the vertical “smearing” of the likelihoods at the facade parts of the large isolated building in the left half of the PVM slice could provide another promising investigation topic in the future.

6.4.2 Reliability Measure

In analogy to the standard backward geocoding technique, the exclusion of unreliable height estimates for surface grid cells that are affected by radar shadowing in all available datasets, is a necessary post-processing step. Since the likelihood values are not limited to a fixed interval, they cannot be employed as a universal reliability measure with a constant threshold for the entire scene of interest. Therefore, a more suitable measure following [Zandona Schneider & Fernandes,

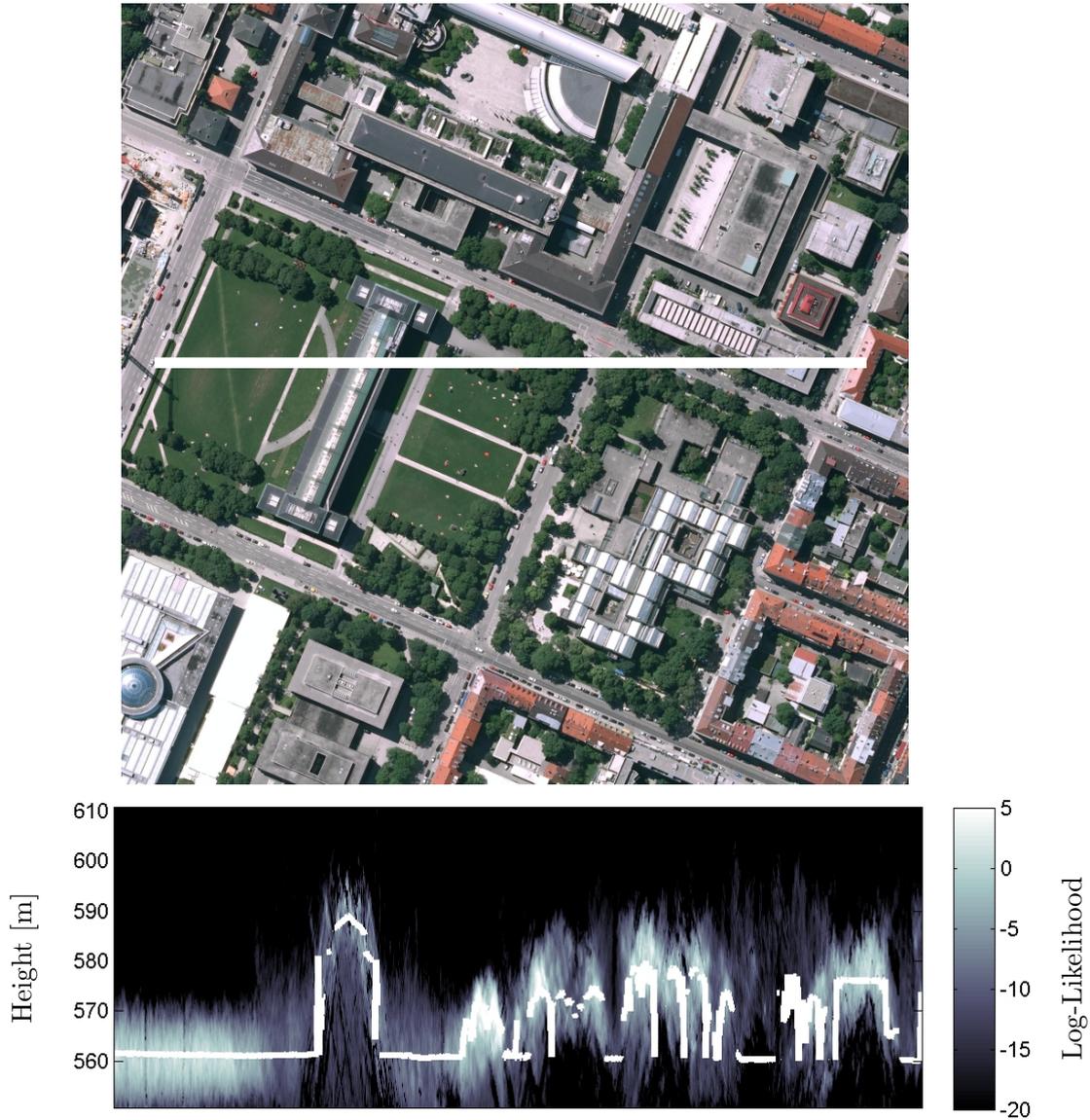


Figure 6.7: Slice through the probabilistic volume model created by the multi-aspect multi-baseline ML estimator. The likelihoods are shown in a logarithmic scale, and the white graph depicts the LiDAR reference data.

2002] is proposed, which exploits the concept of entropy between several SAR images: First of all, the entropy of a resolution cell among N co-registered SAR images is defined to be

$$H = - \sum_{n=1}^N \bar{\lambda}_n \log_N (\bar{\lambda}_n), \quad (6.42)$$

where the normalized eigenvalues $\bar{\lambda}_n$ are given by

$$\bar{\lambda}_n = \frac{\lambda_n}{\sum_{n=1}^N \lambda_n}. \quad (6.43)$$

$\lambda_n, n = 1 \dots N$ denotes the eigenvalues of the complex covariance matrix of the respective stack pixel. As Zandona Schneider & Fernandes [2002] have shown, a value with a behavior comparable to the coherence magnitude can then be calculated by

$$\zeta = 1 - H. \quad (6.44)$$

If the height estimate is generated from multiple aspects, ζ_q needs to be calculated for each pixel of each aspect q separately. A global threshold can now be applied to exclude estimates which are considered as unreliable:

$$\frac{1}{Q} \sum_{q=1}^Q \zeta_q < \varepsilon. \quad (6.45)$$

The main advantage of this reliability measure is that it is directly comparable to the mean coherence as employed in (6.35).

The maximum-likelihood estimation framework proposed in this Section allows for a statistically sound exploitation of arbitrary amounts of available InSAR data. In combination with the derived reliability measure, urban topography can be reconstructed comprehensively without the need to collect multi-temporal stacks or to just focus on point scattering.

7 Utilized Test System and Data

The potential of the methods developed in the context of this thesis is investigated based on experimental data acquired on a flight campaign of the airborne millimeterwave SAR MEMPHIS in May 2011. The sensor is described in Section 7.1, and the special characteristics of millimeterwave SAR data are discussed in Section 7.2. The descriptions were originally published in [Schmitt et al., 2013b].

7.1 Sensor Description

MEMPHIS (Millimeterwave Experimental Multifrequency Polarimetric High-resolution Interferometric System) was developed by the Fraunhofer Institute for High Frequency Physics and Radar Technology FHR in 1998 and was first described by Schimpf et al. [2002]. The radar system uses two front-ends of identical architecture, and operates at 35 GHz and 94 GHz (Ka-band and W-band), respectively. The primary frequencies of 25 GHz and 85 GHz are generated by successive multiplication and filtering of the reference frequency of 100 MHz. For both subsystems the waveform is modulated onto an auxiliary signal at 9.4 GHz, which is upconverted into the respective frequency band together with the primary signal. Depending on the application, the sensor can either be used with polarimetric monopulse feeds or an interferometric set of four receiving antennas. The elevation-azimuth asymmetry of the beam that is necessary for SAR applications is achieved by aspheric lenses in front of the feed horns. Being an experimental, modular and removable system, MEMPHIS is typically mounted on a C-160 Transall airplane of the German Armed Forces (see Fig. 7.1). Due to various possible antenna shapes and configurations, data can be acquired in many different SAR modes: single-pass multi-baseline cross-track interferometry with four receiving antennas, dual-pol circular or linear polarimetry and even monopulse for moving target indication (MTI). The corresponding system parameters are shown in Table 7.1. More information about the raw data processing can be found in [Magnard et al., 2012].

Sensor	MEMPHIS
Carrier frequency	35 GHz (Ka-band)
Wavelength	8.55 mm
Nominal depression angle	25°
Resolution	
<i>azimuth</i>	10.2 cm
<i>range</i>	16.7 cm
Pixel spacing	
<i>azimuth</i>	5.1 cm
<i>range</i>	16.7 cm
Available baselines	5.5 cm, 11 cm, 16.5 cm, 22 cm, 27.5 cm
Approximate ambiguity heights	225 m, 112 m, 75 m, 56 m, 45 m

Table 7.1: MEMPHIS sensor parameters used during the 2011 measurement campaign.

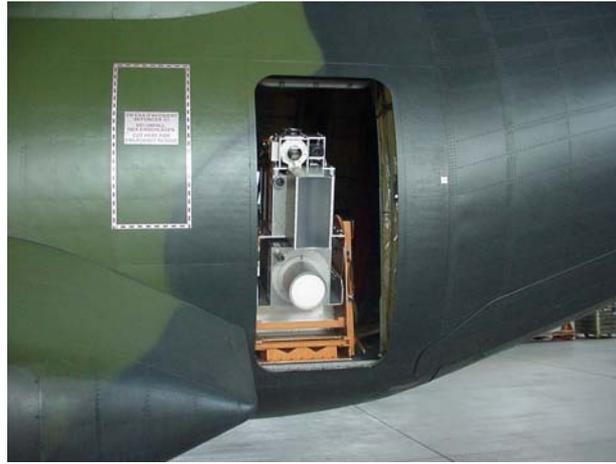


Figure 7.1: The MEMPHIS sensor mounted in the paratrooper door of a C-160 Transall aircraft of the German Armed Forces.

7.2 Characteristics of Millimeterwave SAR

Due to the fact that typical wavelengths of millimeterwave frequencies differ from more common radar remote sensing bands (e.g. L, C, X) in about one order of magnitude, several peculiarities have to be considered; some of them can be exploited advantageously. The main advantages of millimeterwave systems certainly are two-fold: First of all, they allow for a significant miniaturization of the hardware, thus enabling the use on unmanned aerial vehicles (UAVs) and other small-scale carrier platforms [Palm et al., 2013, 2014]. Second, very high resolutions may be achieved with comparably short synthetic apertures. One of the advantages resulting from a short synthetic aperture is that images of vegetation will be better focused, because blurring caused by movements of leaves and branches etc. is reduced.

Additional peculiarities of millimeterwaves in comparison to conventional microwave regions occur in the fields of atmospheric propagation and surface roughness and are explained in the following. A more detailed summary of millimeterwave specifics can be found in [Essen, 2010].

7.2.1 Propagation through the Atmosphere

For millimeterwave radar applications, mainly the transmission windows around 35 GHz and 94 GHz are employed, although high propagation losses prohibit long range applications (>10 km). The millimeterwave region is nevertheless an interesting alternative to the more common X-band due to considerably different propagation properties [Skolnik, 1980], which are caused by resonance absorption at these frequencies related to energy levels of vibration and rotation states of molecules in the atmosphere, e.g. water vapor or oxygen.

In remote sensing, the propagation through snow, fog, haze or clouds is one of the most important reasons why SAR sensors are used. While in optical remote sensing the drop size within fog and clouds is in an order of magnitude where interactions with the electromagnetic radiation of the visible spectrum is most likely, these effects are of much minor importance for millimeterwaves. As long as the density of droplets is not too high, and as long as the liquid water content of snow is not excessively high, millimeterwave signals are able to penetrate most weather phenomena. Only hydrometeors with high density of large drop sizes in the order of the electromagnetic wavelength can severely influence the propagation of the signal and thus prevent the desired imaging of the Earth surface [Danklmayer & Chandra, 2009]. This is caused by the fact

	Ka-band ($\lambda = 0.86$ cm)	X-band ($\lambda = 3.2$ cm)	L-band ($\lambda = 23.5$ cm)
Smooth	< 0.05	< 0.18	< 1.33
Intermediate	0.05 ... 0.28	0.18 ... 1.03	1.33 ... 7.55
Rough	> 0.28	> 1.03	> 7.55

Table 7.2: Definition of radar roughness categories. The RMS surface variations [cm] at a local incidence angle of 45° are shown (after [Lillesand et al., 2004]).

that the drops act as antennas in this case, absorbing the energy of the resonant electromagnetic wave.

For the case of smoke first experiments show a low attenuation for millimeterwaves, due to the small particle size of smoke in comparison to e.g. sand or dust. In these latter cases, experimental results can be used for an estimation of the expected propagation loss. These attributes make millimeterwaves almost just as interesting for any kind of mapping or reconnaissance mission during disaster scenarios, be it floodings (mostly in concurrence with clouds and rainfall), dust storms or fires.

It has to be mentioned however, that even within the millimeterwave domain differences between the different frequencies appear. For example, in W-band the attenuation significantly increases with high temperatures and humidity; it therefore is often used in weather radars [Liebe, 1985]. For that reason, the choice of the band eventually depends on the mission goal.

7.2.2 Surface Roughness Properties

In the millimeterwave region, the wavelength is naturally very short in comparison with classical radar bands, i.e. the relating phase reacts very sensitive on movements of objects or the radar itself. While this might seem disadvantageous for a signal processing based imaging system that relies on the evaluation of the phase of the backscattered signal, it can be utilized beneficially instead. The reason is the specific scattering mechanism, which is dominated by a comparably much rougher surface (factor of 10 in comparison to X-band), making millimeterwave SAR more robust against uncontrolled movements of the carrier aircraft. In general, the roughness of surfaces causes diffuse scattering, whereas smooth surfaces result in specular reflections. At millimeterwave frequencies, most surfaces appear rough, and diffuse scattering dominates the images (see Table 7.2). Diffuse scattering leads to coherent averaging, an effect similar to multilook processing. Therefore, the inherent speckle effect within scenes of homogeneous surface structure is lower at millimeterwave frequencies than at X-band for an equal amount of multilook processing. Besides this primary advantage of higher roughness sensitivity, another one is the larger extent of rough appearing surfaces in often rather smooth urban environments. This provides a convenient benefit to the analysis of backscattering characteristics, which is often based on the assumption of Gaussian scattering. Since this assumption only holds for so-called distributed scatterers and not for frequently occurring point scatterers, millimeterwaves are favourable over longer wavelengths, where tendentially surfaces appear less rough and therefore more point scattering behaviors or specular reflections are observed.

7.3 Measurement Campaign and Test Area

The study area for the experiments in this thesis is located in the Maxvorstadt district in the city of Munich, Germany. It is centered around the TUM main campus (target coordinates: $48^\circ 08' 56''$ N, $11^\circ 34' 02''$ E) and shows mainly dense building blocks, but also some larger buildings



Figure 7.2: Optical image of the test area composed from four orthophotos provided by the Bavarian Administration for Surveying (LVG).

surrounded by patches of concrete or lawn, respectively, and many urban trees. An optical image of the area is displayed in Fig. 7.2, one logarithmically scaled amplitude image for each aspect is shown in Fig. 7.3.

The measurement campaign was carried out in May 2011 in cooperation with the Remote Sensing Laboratories of the University of Zurich and the Fraunhofer Institute for High Frequency Physics and Radar Technology. As illustrated in Fig. 7.4, the test scene was illuminated from a full multi-aspect configuration consisting of five orthogonal and anti-parallel flight tracks (cf. Table 7.3). The goal of this campaign was to create an exemplary dataset of airborne multi-aspect multi-baseline InSAR data using a millimeterwave sensor in order to promote the development of advanced processing strategies necessary for urban area analysis. Detailed information about the sensor setup and the raw data processing can be found in [Schmitt et al., 2013b]: Since the precision of the navigation data provided by the aircraft is not sufficient for high-precision SAR processing, the system was complemented with a differential GPS system composed of a GPS L1/L2 antenna (Aero Antenna AT2775-41) coupled to a receiver running at 20 Hz sampling rate (Trimble R7) and a precise INS working at 500 Hz sampling rate (iNAV-RQH from the company IMAR). The GPS, INS and SAR systems were synchronised through event markers and secondary markers with the GPS time. The realization of time synchronisation for the IMU was carried out by the pulse per second (PPS) signal and NMEA information of the GPS receiver. The navigation solution of the GPS and IMU data was then processed with the commercial software Inertial Explorer using dGPS data from reference base stations. The navigation data were finally smoothed with a Kalman filter to avoid small variations in the millimeter range, which would introduce artefacts in the focused SAR data. The lever arms between the dGPS antenna, the INS and the SAR antennas fixed in operating position were measured using terrestrial surveying methods with a few centimeters accuracy.

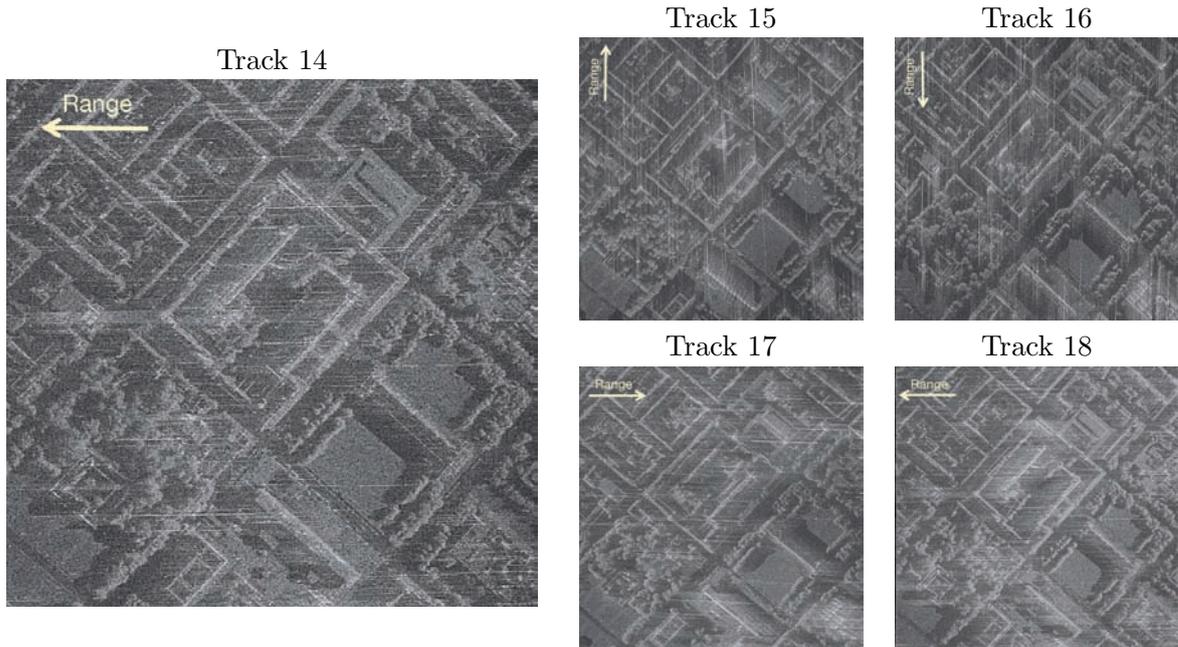


Figure 7.3: Logarithmically scaled amplitude images of the MASAR data (displayed in slant range geometry, pixels approximately squared) acquired during the MEMPHIS campaign in 2011. The images show the area around Technische Universität München and Alte Pinakothek in Munich, Germany. The different viewing directions are indicated by the arrow in the upper left corner.

	Heading angle	Flying altitude	Off-nadir angle
Track 14	340°	768 m	60°
Track 15	70°	714 m	65°
Track 16	250°	709 m	65°
Track 17	160°	714 m	65°
Track 18	340°	712 m	65°

Table 7.3: MEMPHIS MASAR campaign 2011 flight track configuration.

For the SAR raw data focusing a stepped-frequency approach is employed: In high-resolution mode, MEMPHIS successively transmits 8 chirps of 200 MHz bandwidth with a 100 MHz frequency shift between each other, thus building together a 900 MHz full bandwidth, resulting in a range resolution of about 16.7 cm. As described in [Magnard et al., 2012], the raw data from each chirp are first focused in range using a chirp replica with the conventional matched filtering technique. The full bandwidth is then reconstructed in the frequency domain through an algorithm based on [Lord, 2000] and [Wilkinson et al., 1998]. The azimuth compression is performed with the Extended Omega-K algorithm [Reigber et al., 2006], resulting in a zero-Doppler slant range geometry. During the focusing one of the most critical steps for reaching high focusing quality and geolocation accuracy is the motion compensation, which is carried out in a two-step procedure [Schmitt et al., 2013b]. It depends on a precise knowledge of the reference surface and works best for objects located at this reference surface. The more the objects are above or below this surface or the larger the difference between the real and the linearized flight path is, the larger will the geolocation errors and the focusing degradation be. In the context of urban area reconstruction, however, often a precise digital surface model is not available. Since it is furthermore the original goal of this thesis to enable a reconstruction of urban surface models from InSAR data in the first place, the data have been processed using only a coarse digital terrain model, which does not



Figure 7.4: Flight track configuration and overlapping image strips of the 2011 MEMPHIS MASAR campaign (optical image of downtown Munich ©2012 Google).

include any information about above-surface elements such as trees, buildings or other man-made structures.

The result of this pre-processing are multi-baseline InSAR stacks consisting of four co-registered SLC SAR images corresponding to a linearized flight path with constant aircraft velocity. These data serve as input for the experiments conducted in the frame of this thesis.

7.4 LiDAR Reference Data

In order to evaluate the 2.5D height maps and 3D point clouds that are reconstructed from the MEMPHIS data by utilization of the ML-TomoSAR approach (Chapter 5) or the MAMBIInSAR framework (Chapter 6), a reliable reference dataset is needed. In this thesis, a point cloud derived from helicopter-borne forward-looking LiDAR measurements has been used. This LiDAR point cloud was provided by Dr. Marcus Hebel of Fraunhofer Institute of Optronics, System Technologies and Image Exploitation. Its acquisition and preparation is described in [Hebel, 2012]: As explained in [Hebel & Stilla, 2007], the point cloud is the result of a fusion of four individual multi-aspect point clouds, leading to a comprehensive 3D dataset of the scene, including also building facades. The corresponding parameters are summarized in Table 7.4.

LiDAR Reference Data	
Date of acquisition	02.09.2009
Aspects	4
Number of points	4,400,000
Point density	5.2 pts/m ²

Table 7.4: Parameters of the LiDAR reference data.

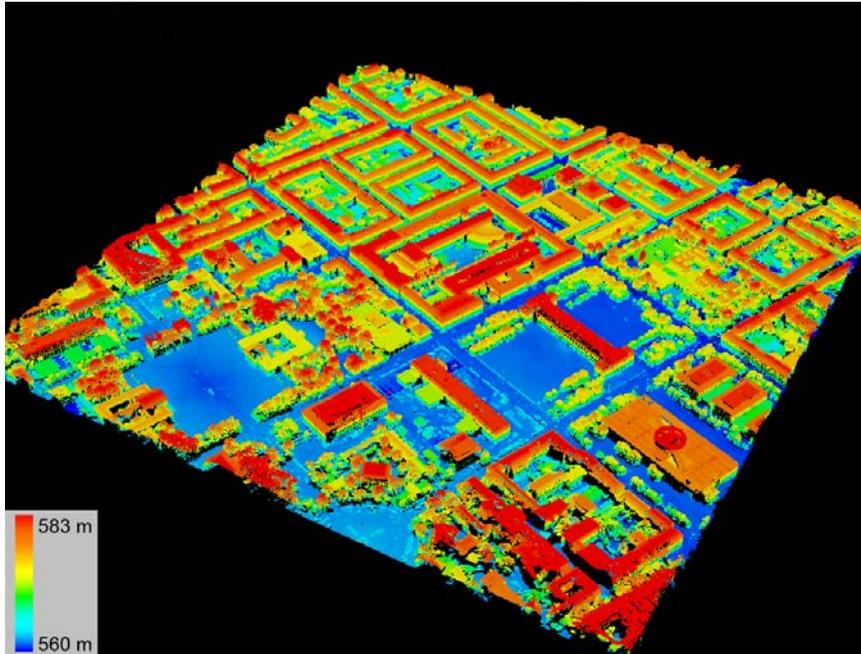


Figure 7.5: Enhanced multi-aspect LiDAR point cloud used as reference dataset.

Since, however, also multi-aspect LiDAR data does not provide complete coverage of the entire city surface, the reference data were further densified by carrying out the following post-processing steps:

- 1) Reduce 3D point cloud to a 2.5D height map by imposing a spatial grid on the scene and choosing the maximum point height within an imaginary column established above each grid cell.
- 2) Fill in holes in the resulting DEM with the median value of the 3×3 -neighborhood of each empty grid cell in a region-growing manner.
- 3) Fuse the original 3D point cloud with the resulting, extended 2.5D point cloud.

In this context, it is important to note that depending on the desired evaluation task, step 3 is optional; e.g. in Section 8.3 only the enhanced 2.5D height map resulting from step 2 is used as reference.

Figure 7.5 shows an image of the enhanced LiDAR point cloud. Although it was processed using high-precision inertial navigation data and dGPS measurements, the sensor positions still contain residual errors of about 1-20 cm, which propagates into the resulting point cloud.

8 Experimental Results

In this chapter, the experimental results achieved in order to validate the methodical developments in the context of this thesis are presented. In the first section, the results proving the efficiency of the proposed adaptive covariance matrix estimation procedures are shown. The second section contains experiments concerning the ML-TomoSAR inversion algorithm, while the third section discusses the results achieved by employing the MAMBIN SAR framework. Section 8.4 discusses the benefit of multi-aspect data fusion, before finally Section 8.5 compares MAMBIN SAR results to 3D reconstruction data generated by a fusion of multi-aspect TomoSAR point clouds.

8.1 Efficiency of the Adaptive Covariance Matrix Estimators

For the evaluation of the adaptive covariance matrix estimation techniques described in this thesis, experiments on both simulated and real test data were carried out. In order to get an impression about both filtering efficiency and adaptivity, the amplitude map of the stack's master image as well as the interferogram corresponding to the longest available baseline combination are extracted from the covariance matrices of all pixels in the stack. In this way, evaluation methods known from the field of conventional image denoising can be employed for evaluation. The tests which utilize simulated data aim at the determination of the theoretical filtering efficiency with respect to the overall noise level and the number of images in the stack. Besides, a comparison to state-of-the-art InSAR filters both under laboratory and real life conditions is to be realized.

8.1.1 Competing Approaches for Comparison

If image filters are developed (and basically, adaptive InSAR covariance matrix estimation is an image filtering problem), it is a good idea to evaluate their capabilities with respect to competitive methods. For this task, three specifically chosen state-of-the-art interferometric multilooking procedures for (multi-baseline) InSAR data have been chosen for comparison to the filters developed in this thesis:

- **Boxcar multilooking**

Boxcar filtering is the common reference procedure that has efficiently been used for low- and medium-resolution SAR and InSAR data for decades [Hanssen, 2001]. Its application to multi-baseline stacks is straight-forward. For the experiments in this thesis, a 15×15 window was employed.

- **DespeckS**

Being the first adaptive multilooking approach specifically designed for stacks of multi-baseline SAR imagery, this method can be considered a direct benchmark [Ferretti et al., 2011]. However, its filtering capacities break down for stacks smaller than about 8 acquisitions [Stephens, 1970], since it is based on a goodness-of-fit test evaluating the statistical

similarity between the amplitude values of two neighboring stack pixels. Also for DespeckKS, a 15×15 window was used.

- **NL-InSAR**

Although NL-InSAR was designed for conventional InSAR pairs only*, it can be considered the most efficient adaptive multilooking algorithm at the moment [Deledalle et al., 2011]. The MATLAB implementation provided at <http://www.math.u-bordeaux1.fr/~cdeledal/nlinsar> is therefore used for comparison. Please note that, although NL-InSAR is not meant to be a window-based algorithm (instead, all pixels are supposed to be considered in a nonlocal manner), in the provided implementation, a 21×21 window is used due to computational reasons.

8.1.2 Theoretical Investigations Based on Simulated InSAR Data

As already described in Section 4.3.3, the simulated data is created based on an optical image and according to the multiplicative speckle noise model as discussed in [Richards, 2009]. The original noise level of the real and imaginary parts is assumed equal, which results in a Rayleigh distributed amplitude value and a Gaussian distributed interferometric phase, whereas the phase of the individual signals is uniformly distributed. In addition to this multiplicative “noise” (i.e. speckle), the signal is augmented with Gaussian distributed thermal noise as it would occur during the sensing process. Both speckle and thermal noise are then summarized to describe the overall noise level mentioned in the following evaluations.

First, the theoretical estimation efficiency was investigated by monte carlo experiments on the covariance matrices of simulated InSAR data pairs. For the probabilistic similarity determination approach the optimal parameter settings as determined in Section 4.3.3 ($\varepsilon_{\text{PDF}} = 10^{-10}$, $\nu = 3$) were used.

The mean normalized bias

$$\|\hat{\mathbf{C}} - \mathbf{C}_0\|_F / N \quad (8.1)$$

of the estimated covariance matrices $\hat{\mathbf{C}}$ to the true covariance matrices \mathbf{C}_0 with respect to a growing noise level is shown in Fig. 8.1. In this context, $\|\cdot\|_F$ denotes the Frobenius norm

$$\|\mathbf{A}\|_F := \sqrt{\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2}, \quad (8.2)$$

where m and n are the number of rows and columns of $\mathbf{A} = [a_{ij}]$ and N is the number of images in the stack.

While the bias reduction of both PCA-TV and probabilistic similarity determination is close the bias reduction offered by NL-InSAR, Boxcar filtering and DespeckKS show a comparably lower efficiency. This is caused by the lack of adaptivity in the Boxcar case, and the fact that DespeckKS requires large sample numbers for the goodness-of-fit testing.

In order to prove the efficiency also for larger stacks with more than two images, additional tests have been carried out on simulated stacks of different size, using a constant noise level. The corresponding plot is shown in Fig. 8.2. Note that in this case, NL-InSAR is not included due to its restriction to the two-image case. Again, PCA-TV and probabilistic similarity determination show a comparably high efficiency (i.e. strong bias reduction), whereas Boxcar filtering and DespeckKS are again limited by non-adaptivity and reliance on large sample numbers, respectively.

*Only during finalization of this thesis a pre-print was published describing an extension of the NL-InSAR principle to multi-dimensional SAR data [Deledalle et al., 2013].

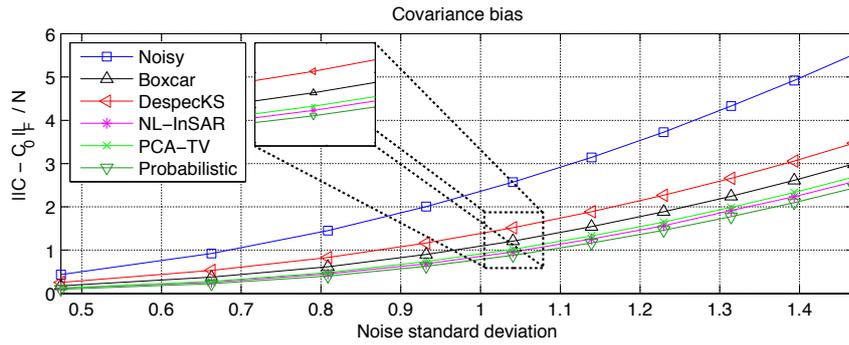


Figure 8.1: Filtering efficiency comparison with respect to noise standard deviation. The normalized bias between estimated and true covariance matrix averaged over all pixels is plotted. The number of images in the stack was set to $N = 2$.

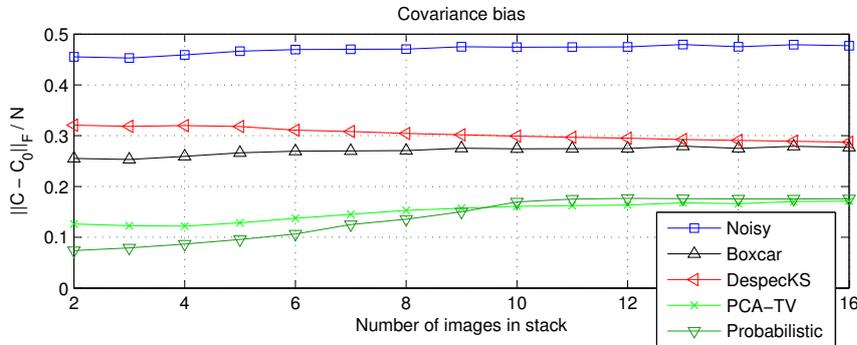


Figure 8.2: Filtering efficiency comparison at a constant noise level, and with respect to growing stack size. Again, the normalized bias between estimated and true covariance matrix averaged over all pixels is plotted.

Image-Based Evaluation

The remaining evaluations are carried out on the image data that can be composed if the corresponding variables (i.e. the amplitude of the master image and the interferometric phase of the longest baseline) are extracted for every pixel in the stack. Figure 8.3 shows the filtering efficiency with respect to these image data of both the PCA-TV-based and the probabilistic covariance matrix estimation algorithms for a co-registered pair of simulated interferometric SAR images with respect to a growing noise level.

Again also an additional experiment has been carried out on simulated stacks of different sizes, using a constant noise level. The resulting filtering efficiencies are displayed in Fig. 8.4.

Qualitative results of the experimental comparison to the competing approaches for simulated InSAR image pairs are visualized in Fig. 8.5, quantitative results can be found in Fig. 8.6. The results of Boxcar filtering, DespeckKS and NL-InSAR are as expected: The Boxcar approach provides high filtering efficiency, but completely blurs all image details. DespeckKS is not able to filter the data, since two samples per pixel are too few for statistical goodness-of-fit testing. Eventually, NL-InSAR provides the benchmark result. Both the PCA-TV based filter and the probabilistic filter provide strong filtering efficiency and detail preservation. However, the PCA-TV result is disturbed by single dark spots and an unsatisfactory result for the interferometric phase of letter 3, which is probably caused by the fact that two images do not allow for a sufficient

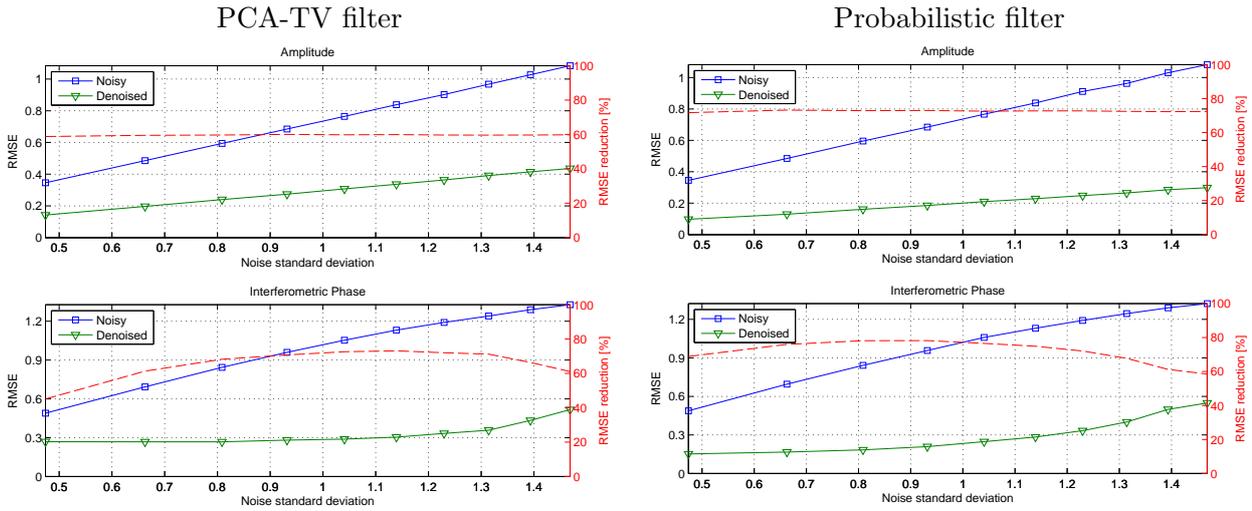


Figure 8.3: Filtering efficiency of the proposed adaptive covariance matrix estimation algorithms with respect to the noise standard deviation. The left column shows the results of the PCA-TV based procedure, the right column shows the results of the probabilistic procedure. For this experiment, standard InSAR image pairs have been simulated.

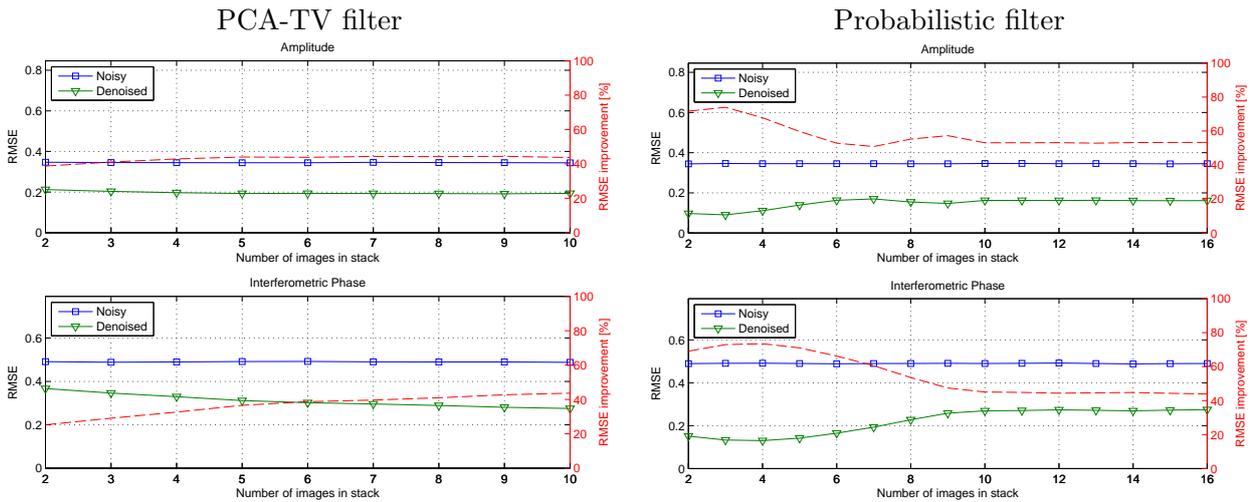


Figure 8.4: Filtering efficiency of the proposed adaptive covariance matrix estimation algorithms with respect to a growing number of images per stack. The left column shows the results of the PCA-TV based procedure, the right column shows the results of the probabilistic procedure. In this experiment, the noise level was kept constant.

separation between signal and noise components. In contrast, the probabilistic approach leads to a result comparable to NL-InSAR in quality.

For stacks consisting of four simulated multi-baseline images, the results are shown in Figs. 8.7 and 8.8, respectively. Note that NL-InSAR is not part of the experiment anymore since it is restricted to conventional InSAR image pairs. Apart from that, the results lead to an impression similar to the previously investigated two-image case: While Boxcar filtering blurs the results, DespeckKS still suffers from the low number of images. The PCA-TV result shows fewer dark spot outliers, and the probabilistic filter still creates the best results considering the desired trade-off between filtering efficiency and detail preservation.

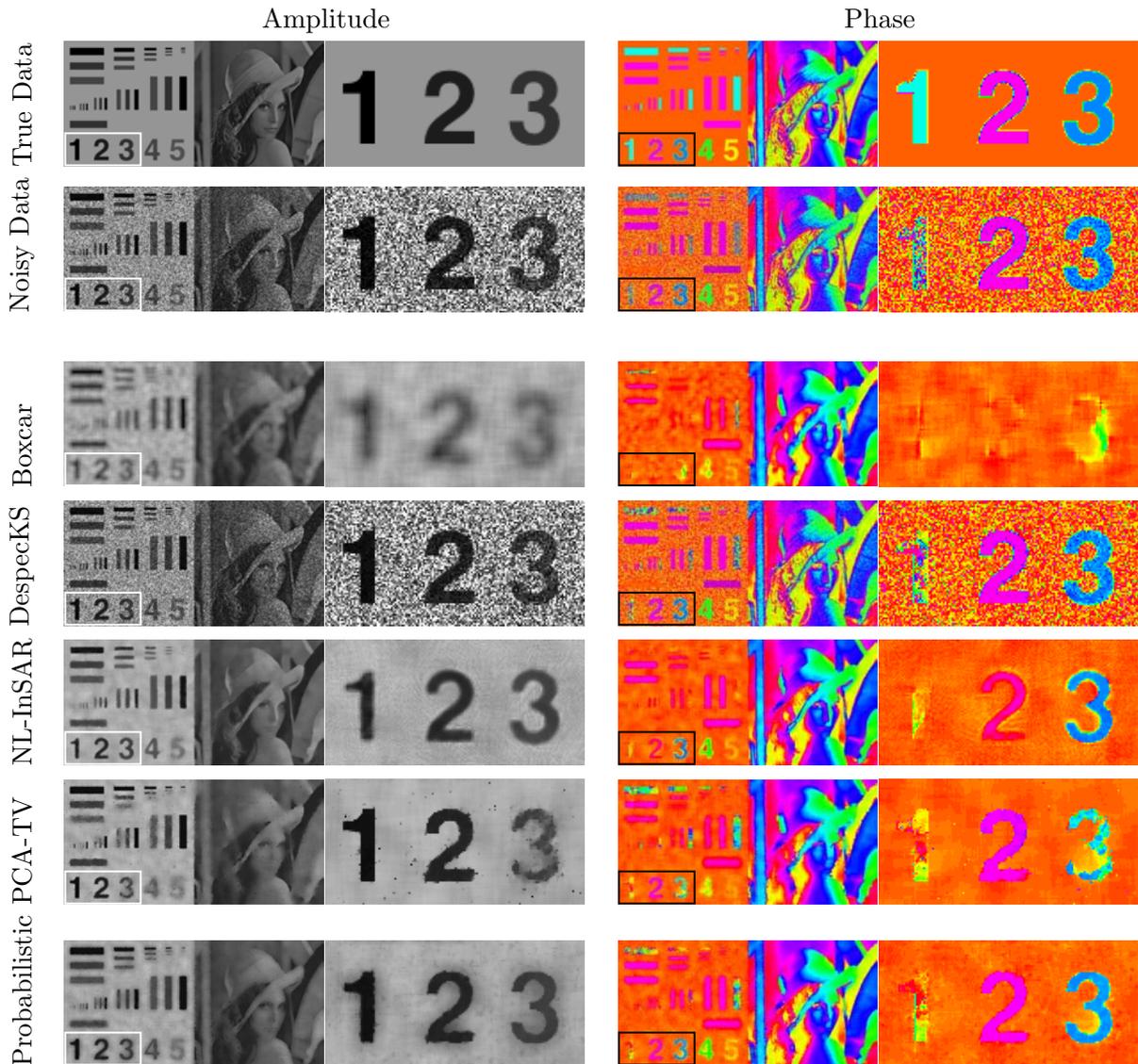


Figure 8.5: Filtering results for a simulated InSAR pair. The boxes around numbers 1, 2, 3 are enlarged for better interpretability.

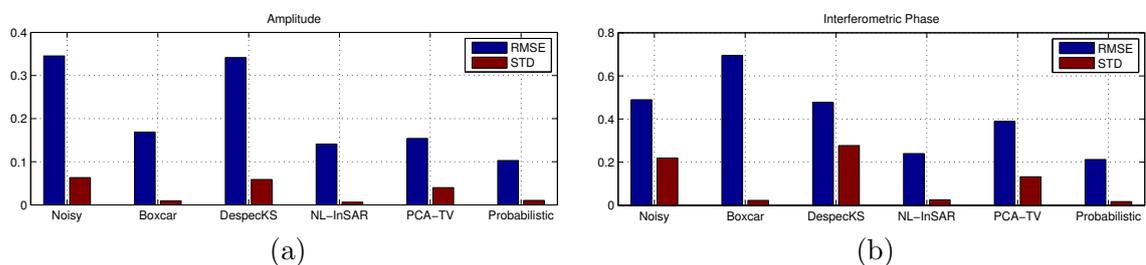


Figure 8.6: Quantitative comparison of filtering results on a simulated InSAR pair. (a) Amplitude and (b) interferometric phase data. For the evaluation of the phase, shadow regions have been excluded.

8.1.3 Practical Results on Real Multi-Baseline InSAR Data

In spite of all simulated test results, in the end, the quality of a filter can only be evaluated based on real data experiments. Since the expectation value, i.e. a noise-free “ground truth” of the

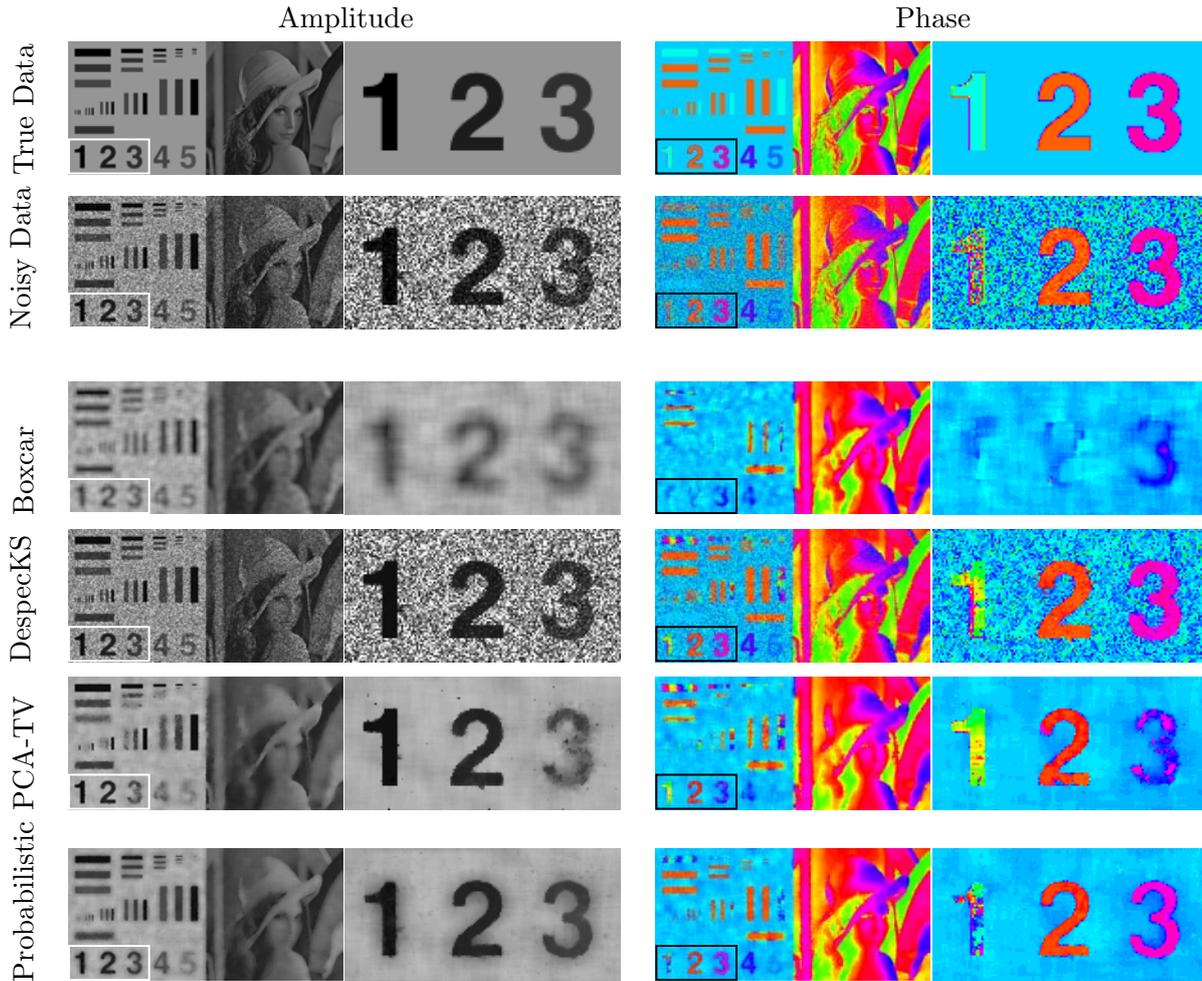


Figure 8.7: Filtering results for a simulated multi-baseline stack with 4 images. The boxes around numbers 1, 2, 3 are enlarged for better interpretability.

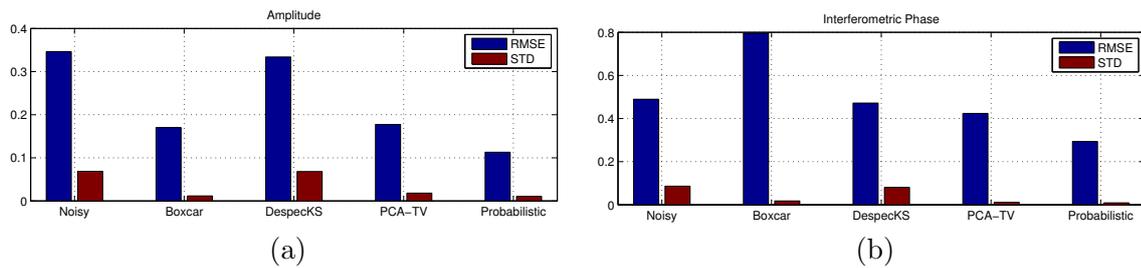


Figure 8.8: Quantitative comparison of filtering results on a simulated stack consisting of 4 images. (a) Amplitude and (b) interferometric phase data. For the evaluation of the phase, shadow regions have been excluded.

real SAR imagery is not known, it is necessary to provide alternative measures to the RMSE for assessing the efficiency of the filters. In this context, efficiency is defined as a combination of variance reduction and mean preservation, which are commonly estimated on homogeneous image patches. Unfortunately, these estimates tend to be biased in certain cases, such that this thesis refers to the speckle suppression index (SSI) and the speckle suppression and mean preservation

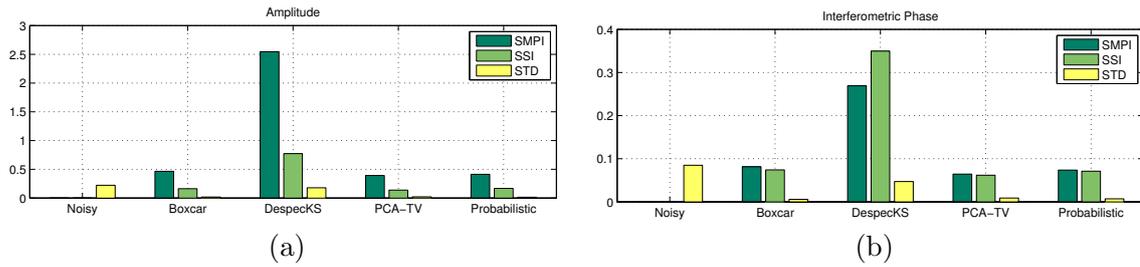


Figure 8.9: Quantitative comparison of filtering results for a 4-image MEMPHIS stack. (a) Amplitude and (b) interferometric phase data.

index (SMPI) as proposed by Shamsoddini & Trinder [2010] in order to get a more reliable evaluation. The SSI is calculated by

$$\text{SSI} = \frac{\sigma_f}{\mu_f} \cdot \frac{\mu_0}{\sigma_0}, \quad (8.3)$$

where μ_0 , μ_f , σ_0 and σ_f denote the means and the standard deviations of the original and the filtered images, respectively. If the filter is efficient in speckle reduction, the SSI is usually less than one.

However, also the SSI tends to be unreliable if the filter overestimates the mean value. Therefore, the more sophisticated SMPI, which is calculated by

$$\text{SMPI} = (1 + |\mu_0 - \mu_f|) \cdot \frac{\sigma_f}{\sigma_0} \quad (8.4)$$

is used additionally. The lower the SMPI values are, the more efficient the filter performs. Figure 8.9 shows standard deviation (STD), SSI and SMPI for amplitude and phase data extracted from the complex covariance matrices that were estimated for each pixel in the MEMPHIS test data stack. Again, NL-InSAR was not considered because the stack consists of four co-registered images. It can easily be observed that the real data results agree with the results achieved with simulated data: Both the PCA-TV filter and the probabilistic filter achieve a filtering efficiency comparable to the optimal Boxcar method. The relatively bad result for DespeckKS is again caused by the low number of images in the stack, providing only four samples per pixel for the goodness-of-fit test.

In addition to the filtering efficiency, a core requirement for the proposed covariance matrix estimation algorithms is their adaptivity: In order to ensure local stationarity for the estimation of the covariance matrix of a pixel in an interferometric SAR stack, all sample pixels must belong to the same statistical distribution. Therefore, adaptive sample selection is needed. Unfortunately, this adaptivity can hardly be evaluated quantitatively. Due to this reason, a visual comparison of both amplitude and phase images was chosen for the assessment. In this way, the adaptivity of the proposed filtering method can be analyzed by its capability to preserve fine image details such as edges or strong point scatterers. Figure 8.10 shows the comparison for the different filtering approaches. As expected, it can be seen that the Boxcar approach completely blurs the image and destroys fine details, which proofs the lack of adaptivity. DespeckKS, the PCA-TV based and the probabilistic filter all preserve the exemplarily chosen facade scatterers in a satisfying manner. Looking at the phase images, however, it becomes obvious that the PCA-TV approach tends to stronger filtering and less detail preservation in comparison to the probabilistic filter.

Finally, a very important task for interferometric SAR applications or change detection efforts is the estimation of the coherence between the individual acquisitions. As shown in Section 4.1, also this measure can easily be extracted from the complex covariance matrix of each pixel. In

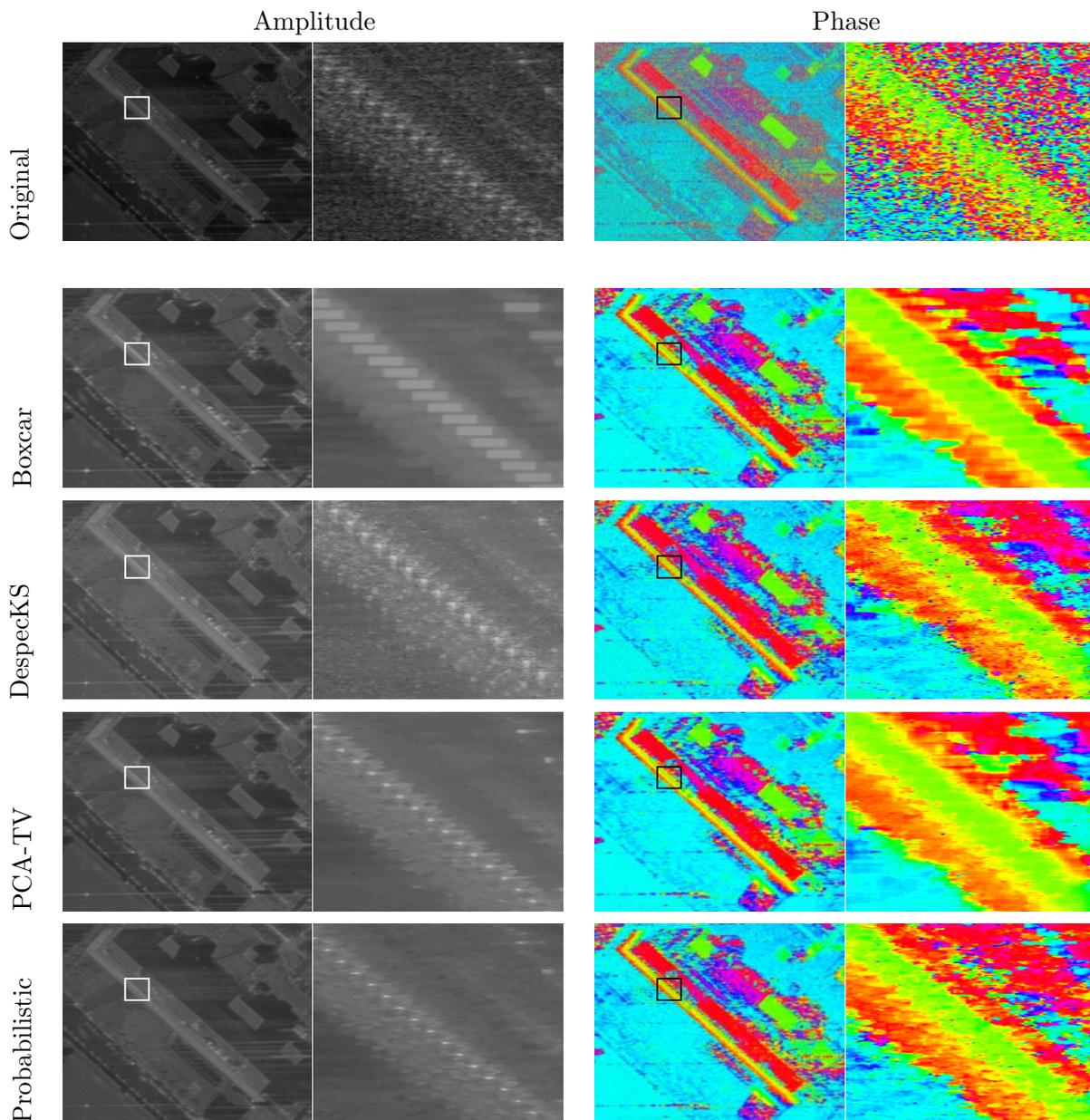


Figure 8.10: Filtering results for a multi-baseline stack with 4 images, acquired by the airborne MEMPHIS sensor. The interferometric phase data of the longest baseline is shown.

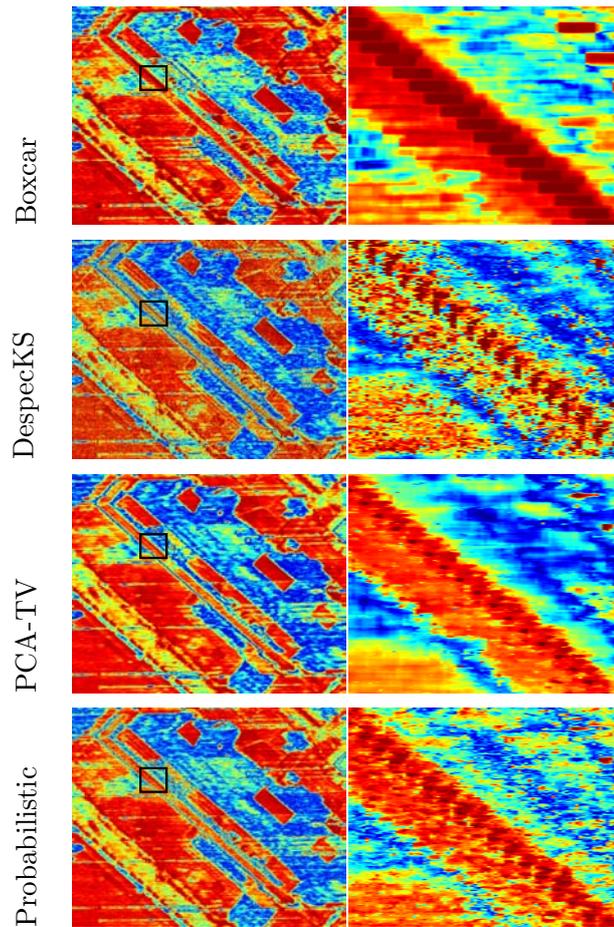


Figure 8.11: Coherence maps estimated for the longest baseline of the real test dataset.

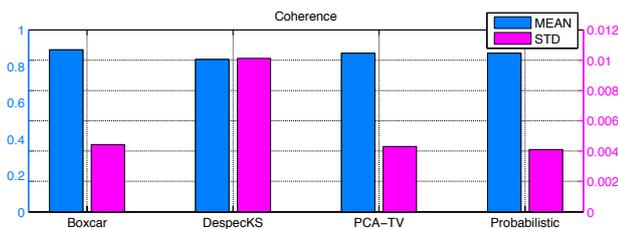


Figure 8.12: Mean and standard deviation of the coherence magnitudes within a small homogeneous window on the lawn area in the real test dataset.

order to prove the effect of the proposed adaptive estimation procedures on coherence estimation, the results for the longest baseline available in the real test dataset are again compared to the coherence maps extracted by state-of-the-art methods (see Fig. 8.11). The mean and standard deviations of the coherence magnitudes within the same homogeneous window already used for amplitude and phase evaluation are displayed in Fig. 8.12. It can be seen that PCA-TV and the probabilistic filter achieve almost the same coherence quality as the Boxcar filter, while still preserving fine-structured details in the data.

8.1.4 Discussion

From both the results of the simulated data and real data experiments, several insights can be acquired: It is obvious that most of the filtering methods provide advantages and disadvantages

at the same time. While Boxcar filtering generally shows a very high filtering efficiency, which can be considered as a benchmark for homogeneous areas, it naturally destroys fine details and is not adaptive in any way. NL-InSAR is probably the most adaptive and efficient filter known at the moment, but it is not able to process multi-baseline InSAR stacks with more than two images. In analogy to that, DespecKS is a very promising approach for stack filtering, but requires a rather high number of acquisitions to make the goodness-of-fit testing work.

Both PCA-TV filtering and the probabilistically motivated approach for adaptive covariance matrix estimation close the gap between NL-InSAR and DespecKS, providing promising techniques for small-sized stacks with about 3 to 5 images. However, the PCA-TV results are weaker in the two-image case; this is caused by the nature of the method, which is based on a compression of signal components in the first principal component. In contrast, the probabilistic approach provides high filtering efficiency and strong detail preservation for independent stack sizes while still being optimally suitable for small-stack-cases. Again the reason lies in the filter principle, which is centered around the multi-dimensional probability density function of each pixel, making the general design independent of the number of samples per pixel. Apart from that, the two necessary parameters (ν for the robust initial covariance matrix estimation and ε_{pdf} for the probability density thresholding) can be set globally, making the algorithm a powerful and still easy-to-use automatic tool for different data configurations. However, the main drawback of the method must not be neglected: Since it is based on a sliding window operation and a three-stage covariance matrix estimation process, it is computationally more expensive than PCA-TV.

8.2 Maximum-Likelihood TomoSAR Results

The utilization of MEMPHIS data for the evaluation of the proposed TomoSAR algorithm implies two important attributes: First, MEMPHIS is a good example of a single-pass multi-baseline system with a very limited overall baseline $\Delta B = 27.5$ cm and just four receiving antennas. From a TomoSAR point of view, this means a critically short elevation aperture with an additionally critically low number of samples per resolution cell. Therefore, this experiment can also be seen as a study exploring the limits of SAR tomography with respect to potential future mission design. The second point is the employment of millimeterwave SAR, which is supposed to follow the Gaussian scattering assumption even in urban areas due to the high sensitivity to surface roughness as already explained in Section 7.2. The height resolution of the utilized interferometric configuration of the MEMPHIS sensor can be approximated by

$$\rho_h = \frac{\lambda R \sin(\theta)}{\Delta B} \approx 42 \text{ m} \quad (8.5)$$

with its mean off-nadir angle $\theta = 60^\circ$ at a slant range distance of $R = 1545$ m. The elevation-to-height conversion factor $\sin(\theta)$ is needed in order to receive heights above the reference surface instead of just elevations. Therefore, the utilization of super-resolution techniques would be necessary if classical spectral estimation algorithms were meant to separate scattering contributions that differ less than 42 m in height – a common occurrence in urban areas. Since, however, the method proposed in this thesis is not based on spectral estimation methods for array signal processing, the height resolution is only limited by the question how good the model signal values fit the measured coherence matrix and by the discretization of the search space. (cf. Section 5.3.1).

For all experiments in this section, the height search interval was chosen as $[h_{\min}; h_{\max}] = [-5 \text{ m}; 45 \text{ m}]$, with a spacing of $\Delta h = 0.5$ m. After covariance matrix estimation for every stack pixel, additionally a pre-summing in azimuth direction was carried out in order to achieve approximately square slant range pixels.

8.2.1 Experiments on Simulated Data

In order to evaluate the proposed ML-TomoSAR method, a comparison with the long-known Capon beamformer [Capon, 1969] and the Weighted Signal Subspace Fitting (WSSF) approach was chosen. WSSF was first presented by Viberg et al. [1991] and transferred to the TomoSAR context by Huang et al. [2012]. Although the Capon algorithm looks quite similar to the TomoSAR approach developed in this thesis, the main differences are that the non-normalized complex covariance matrix is used and only a one-dimensional objective function is evaluated in the maxima detection:

$$\hat{\mathbf{h}} = \arg \max_{\mathbf{h} \in [h_{min}; h_{max}]} \frac{1}{\mathbf{z}^H(\mathbf{h}) \hat{\mathbf{C}}^{-1} \mathbf{z}(\mathbf{h})}. \quad (8.6)$$

In contrast, ML-TomoSAR and WSSF are quite similar in their philosophy: Just like Capon, both are multi-looking-based, i.e. they rely on covariance matrix estimates for every resolution cell, and they make use of a multi-dimensional grid search for the optimization step. For WSSF, the estimator is derived as

$$\hat{\mathbf{h}} = \arg \max_{\mathbf{h} \in [h_{min}; h_{max}]} \text{trace} \left(\mathbf{P}_{\mathbf{A}}^{\perp} \hat{\mathbf{E}}_s \left(\hat{\mathbf{\Lambda}}_s - \hat{\sigma}_n^2 \mathbf{I}_K \right) \hat{\mathbf{E}}_s^H \right)^2 \hat{\mathbf{\Lambda}}_s^{-1}, \quad (8.7)$$

where

$$\mathbf{P}_{\mathbf{A}}^{\perp} = \left(\mathbf{I}_N - \mathbf{A} (\mathbf{A}^H \mathbf{A})^{-1} \mathbf{A}^H \right) \quad (8.8)$$

represents the orthogonal projector on the null space of \mathbf{A}^H , and \mathbf{A} is set up as described in (5.26). $\hat{\mathbf{E}}_s$ is the matrix containing the K eigenvectors defining the signal subspace, while $\hat{\mathbf{\Lambda}}_s$ is a diagonal matrix containing the corresponding eigenvalues. $\hat{\sigma}_n^2$ is an estimate of the noise variance.

For a first theoretical comparison, experiments based on simulated resolution cells, utilizing the MEMPHIS system parameters, were carried out. Each resolution cell was simulated to contain two Gaussian scatterers with equal backscattering power. While the height of one scatterer was fixed to ground level ($h_1 = 0$ m), the second scatterer was assigned a growing height. For every incrementation of h_2 , 100 monte carlo simulations were carried out at an SNR of 10 dB. The results of this experiment can be seen in Fig. 8.13. In order to emphasize the core methodology proposed in this thesis, the analyses have been carried out both with automatic model order selection as described in Section 5.2 and with a fixed model order of $\hat{K} = 2$.

8.2.2 Experiments on Real Data

For the real-data experiments, a cut-out from aspect 14 has been chosen as test data. Figure 8.14 displays the test area located in the heart of the museum district of the city of Munich, Germany, as well as a corresponding SAR intensity image. An analysis of the signal-to-noise ratio of the SLC observations resulted in an estimated median SNR of about 4 dB for non-shadow pixels, whereas potential layover resolution cells can show an SNR of 10 dB and up. The strong range sidelobes occurring in some parts of the SAR image are caused by saturation effects at the A/D converter of the sensor. As a pre-processing step, the phase data were calibrated using the deterministic signal of the corner reflector in the bottom left part of the scene. After flat earth correction, the covariance matrices of the complex multi-baseline observations were estimated by the PCA-TV method proposed in this thesis.

The model order map needed as prior knowledge for the TomoSAR algorithm was estimated by the EDC₂ method described in Section 5.2 (see Fig. 8.15). Although the sidelobes cause some wrong model order estimates, it can be seen that most parts of the scene correctly contain just one scattering contribution. As expected, two scatterers are found mainly on the facade of the

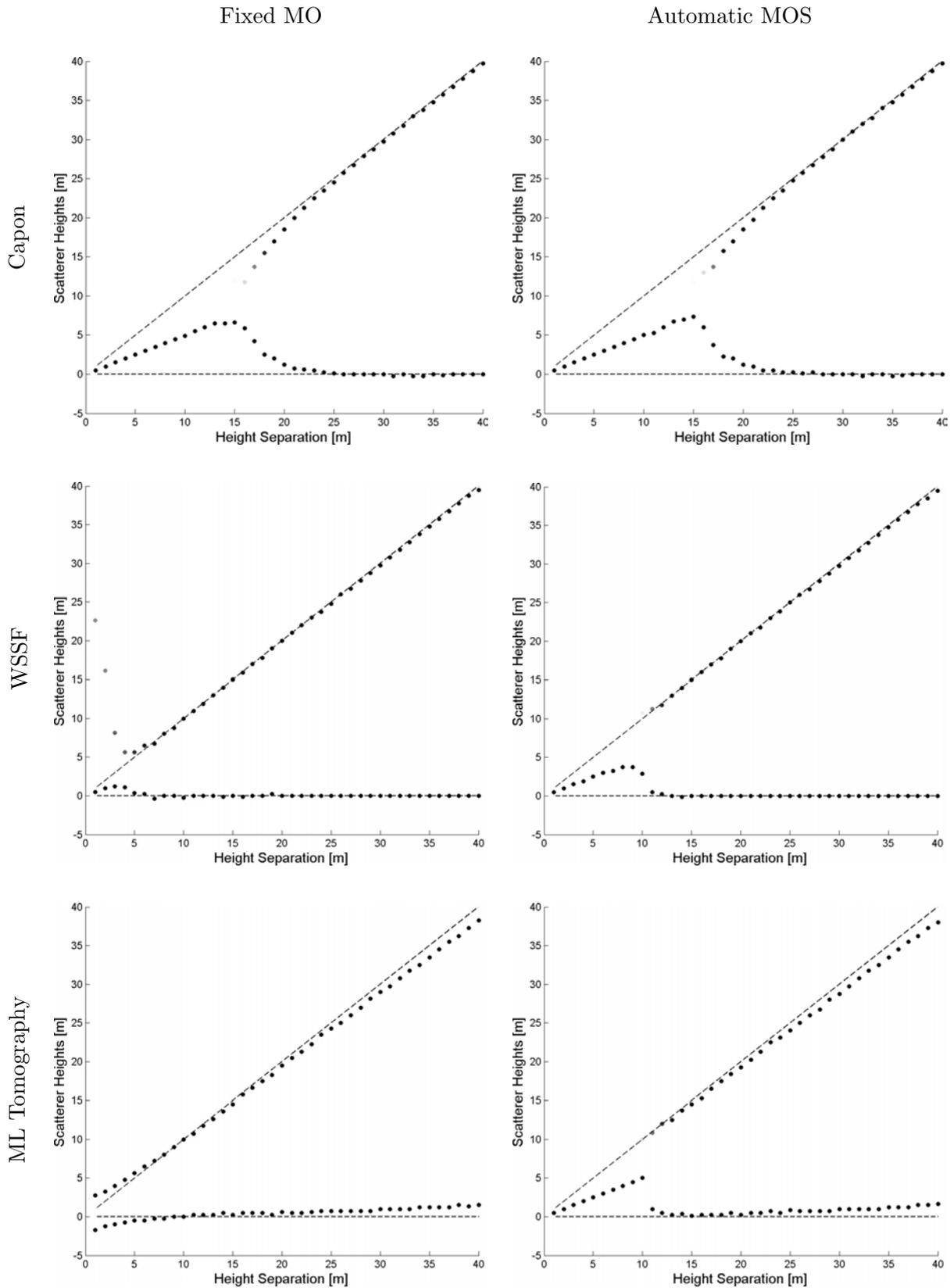


Figure 8.13: Scatterer height reconstruction results for simulated data at an SNR of 10 dB. Left column: model order (MO) fixed to $\hat{K} = 2$, right column: automatic model order selection (MOS)

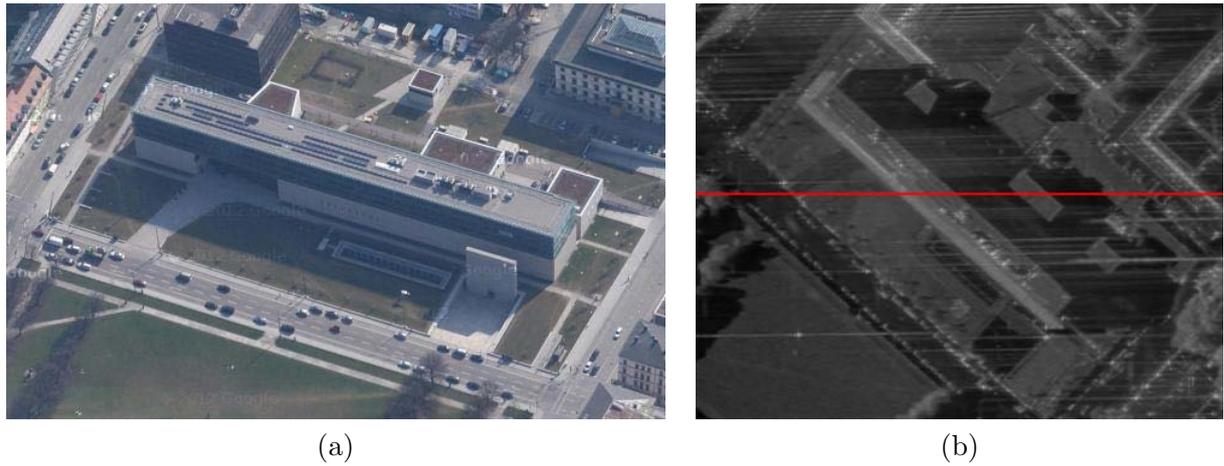


Figure 8.14: Test scene for the real data TomoSAR experiments: (a) Oblique view aerial image ©Google 2013, (b) despeckled intensity image, where the red line indicates an azimuth bin, which is used for further analysis.

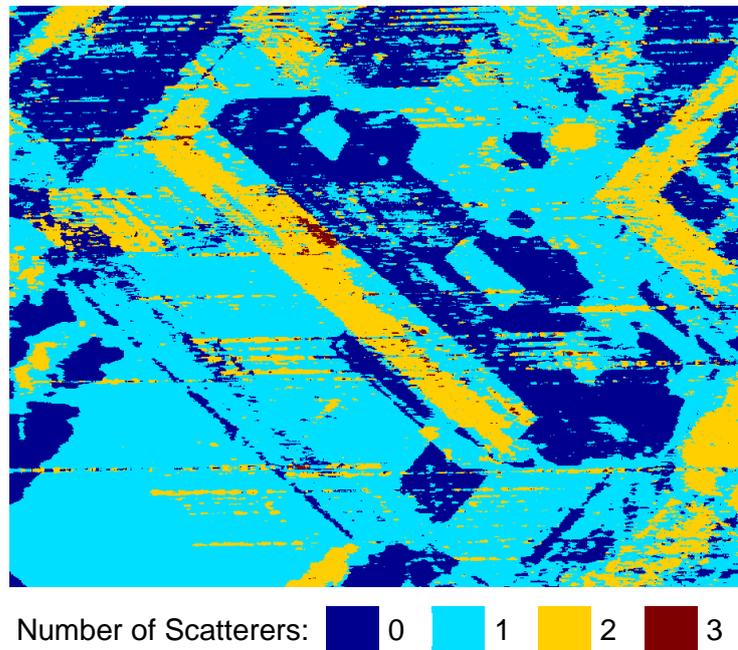
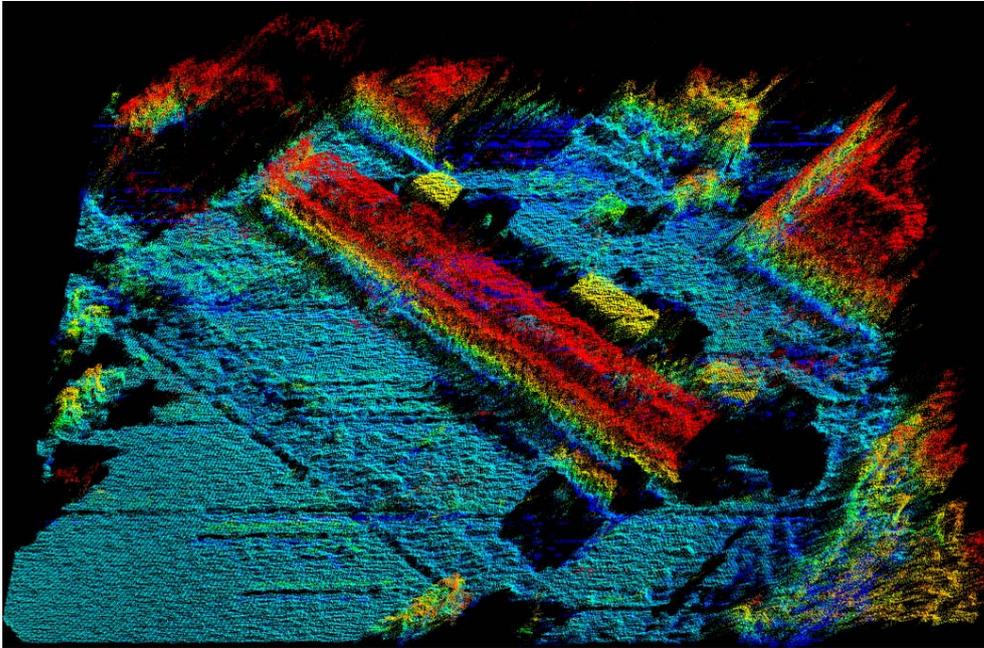


Figure 8.15: Model order map created by application of the EDC_2 criterion.

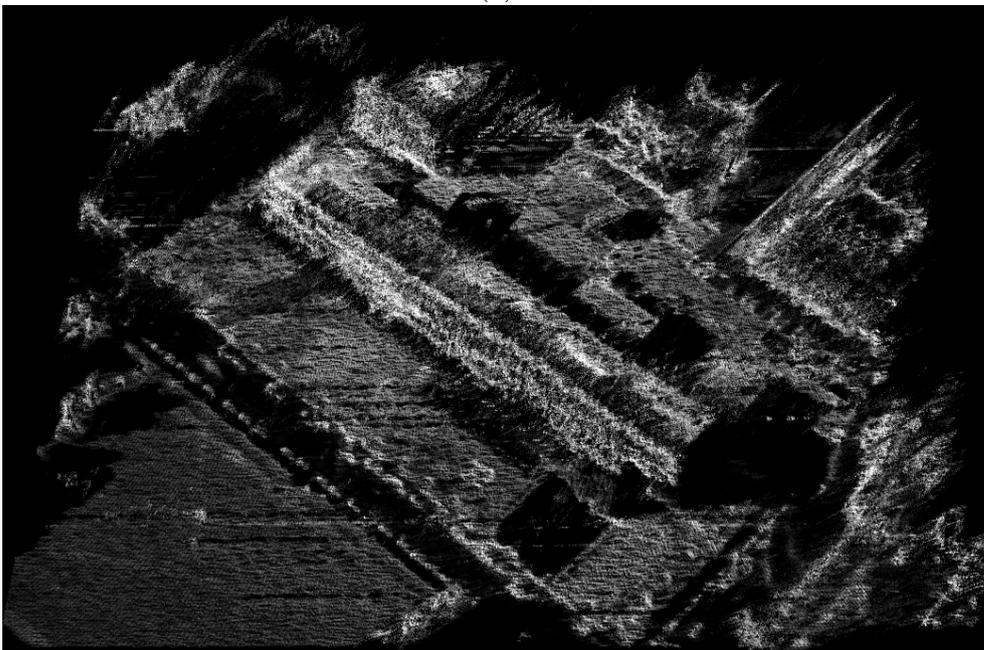
large building in the scene center. Of all resolution cells containing scatterers, 23% were found to contain double scatterers, whereas the share of triple scatterers is only at about 0.5% and can therefore be neglected. The final tomographic result for the entire test scene is shown in Fig. 8.16. Note especially how the cars on the road are clearly separated from the low reflectivity asphalt in the 3D tomogram, while not really being distinguishable in the pure height data. For a detailed comparison of the proposed ML-TomoSAR approach and WSSF as a benchmark, Fig. 8.17 shows the two-dimensional tomographic slice for the azimuth bin marked in Fig. 8.14 (b).

8.2.3 Discussion

The results of both the simulated and real data experiments show the feasibility of maximum-likelihood based SAR tomographic reconstruction. From the simulation analysis illustrated in



(a)



(b)

Figure 8.16: Tomographic processing result by the proposed maximum-likelihood approach: (a) Height colored scatterer point cloud, (b) sparse 3D tomogram showing the logarithmically scaled amplitudes of each scatterer.

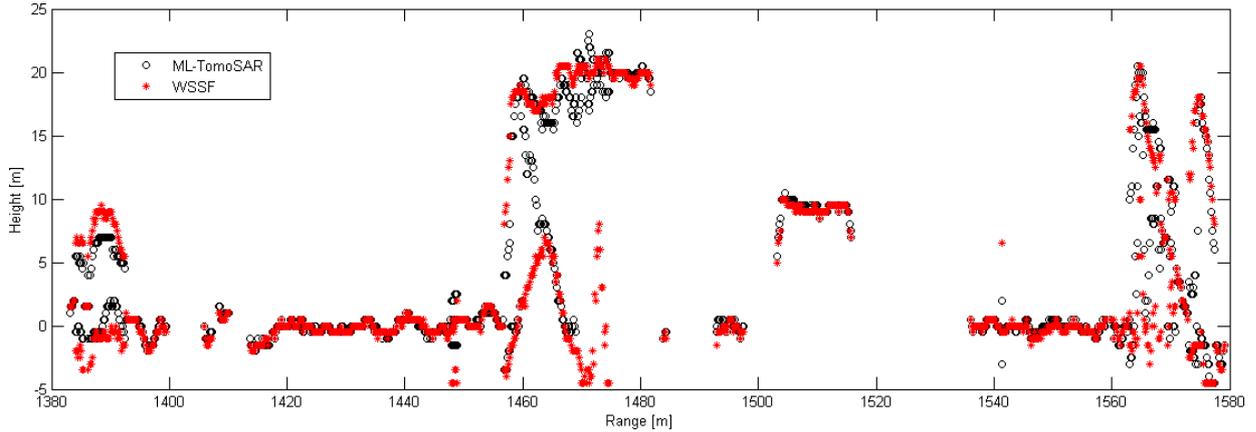


Figure 8.17: Comparison of WSSF and ML-based TomoSAR results in the slant range-height plane for the azimuth bin marked in Fig. 8.14 (b).

Fig. 8.13 and the real data comparison shown in Fig. 8.17 it can be seen that the ML-TomoSAR results are quite similar to the WSSF results, whereas the layover separation capability even exceeds the one of WSSF for very closely spaced scatterers: While both methods are limited by the resolution of the a priori model order selection step (approx. 10 m), the reconstruction accuracy of WSSF would break down for a height difference smaller than 5 m. In contrast, ML tomography would enable a separate reconstruction of two scatterers even for height differences below 5 m, although with increasing bias. In exchange, WSSF allows a seemingly perfect reconstruction of the scatterers with sufficient height difference, whereas ML tomography shows slightly biased height values. The main advantage of ML-TomoSAR therefore is that its computational cost is lower than for WSSF since less matrix operations are involved. In comparison to these multi-dimensional optimization methods, Capon's super-resolution capability is significantly worse.

A reflection of the real data results also shows the high potential of the proposed method. Again, it becomes visible that the pre-processing step of model order selection is a critical part of the procedure: Looking at Fig. 8.15, one can see that only very few pixels are found to contain three scatterers, although one might expect more mixtures of roof, facade and ground. Figure 8.18 shows a plot of the double scatterer differences of the test scene reconstruction. Since the Rayleigh resolution limit is at about 42 m, it proves the super-resolving capabilities of ML-TomoSAR also for real SAR data. Obviously, the majority of the double scatterers show a height difference of about 5 m to 10 m, which fits the considerations of Zhu & Bamler [2012], who state that most of them are rather closely spaced. The outlying peak of the plot at 0 m to 1 m is caused by the combination of an a priori model order selection and the two-dimensional grid search based optimization: In these cases, the MOS step over-estimates the number of scatterers in the respective resolution cell, while the ML-estimator finds both to be at very close heights below the actual resolution limit of the model order selection step. Figure 8.19 shows a plot containing the results of another simulation: This time, the resolution cell was supposed to contain only one scatterer at $h = 15$ m, but the model order was wrongly set to $\hat{K} = 2$. It can be seen that, with growing SNR, the two reconstructed scatterers are found closer to each other, while their mean value is always at the correct height. Therefore, an over-estimation of the number of scatterers is not necessarily problematic from a methodological point of view, since a threshold could be used to fuse two close pseudo-scatterers. In the presented case, e.g. a height difference of $\Delta h = 5$ m would be well suited. In contrast, an under-estimation of the model order usually leads to a wrongly determined 3D scatterer. Due to the high computational cost of multi-dimensional grid search optimization, a GPU-based, parallelized implementation of the algorithm would be advisable. In

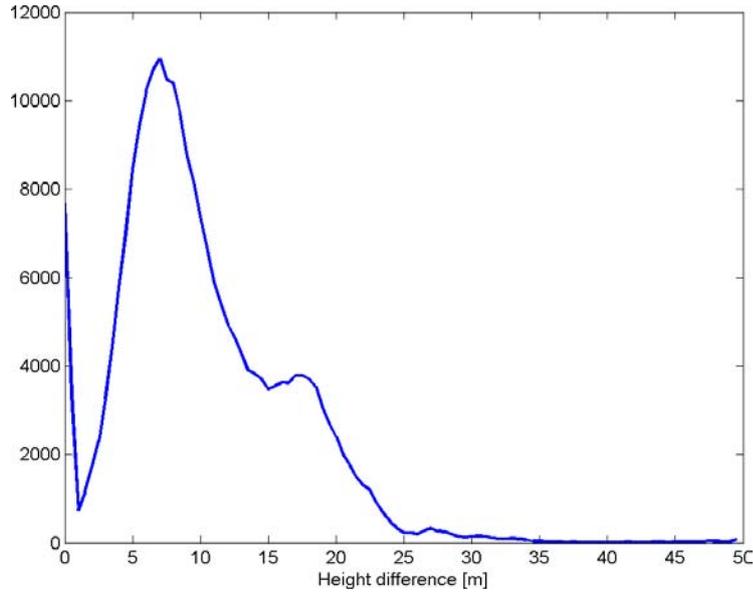


Figure 8.18: Histogram of the height differences between the detected double scatterers in the real data test case. The Rayleigh resolution limit is at about 42 m.

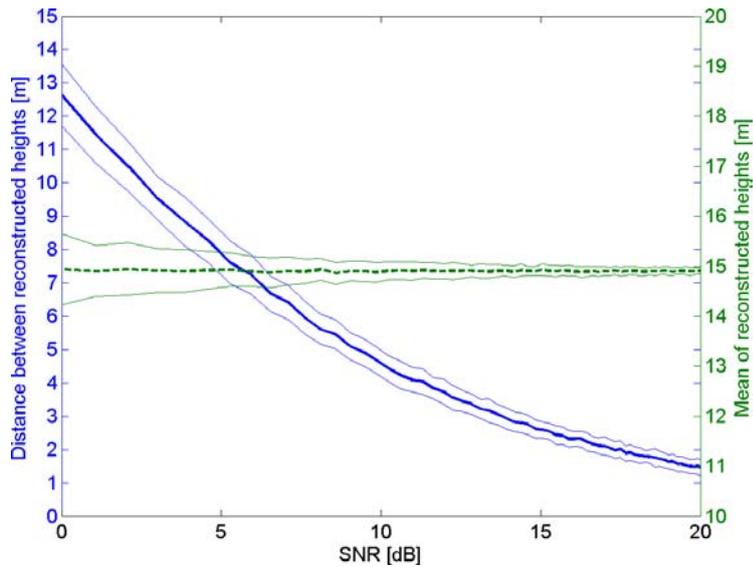


Figure 8.19: Experimental result of a simulation of a single scatterer at $h = 15$ m and an overestimated model order of $\hat{K} = 2$. It can be seen that the mean height of the two wrongly reconstructed pseudo-scatterers always equals the desired correct height of the actual scatterer, whereas the height difference between the two pseudo-scatterers gets smaller with growing SNR.

this case, it would be possible to systematically employ a larger number of scatterers than actually estimated. In combination with a thresholding operation that fuses two closely spaced scatterer estimates to a single scatterer, a significantly enhanced comprehensiveness and quality could be expected.

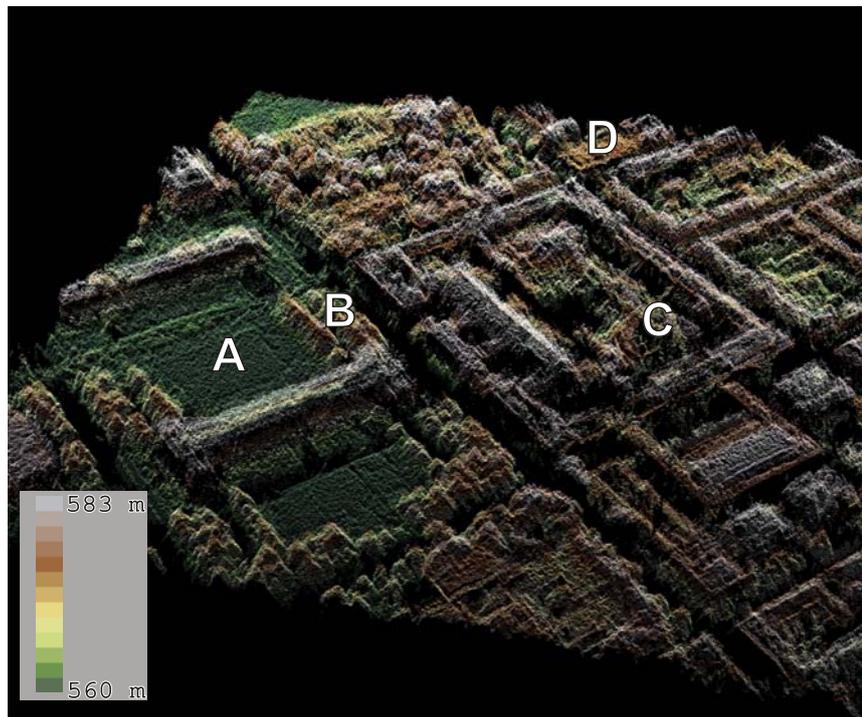
8.3 Multi-Aspect Multi-Baseline Interferometry Results

For the experiments described in this section, the height search interval for the MAMBIInSAR estimator was defined to be $[h_{min}; h_{max}] = [h_0 - 10 \text{ m}; h_0 + 50 \text{ m}]$, where $h_0 = 560.7$ m is the

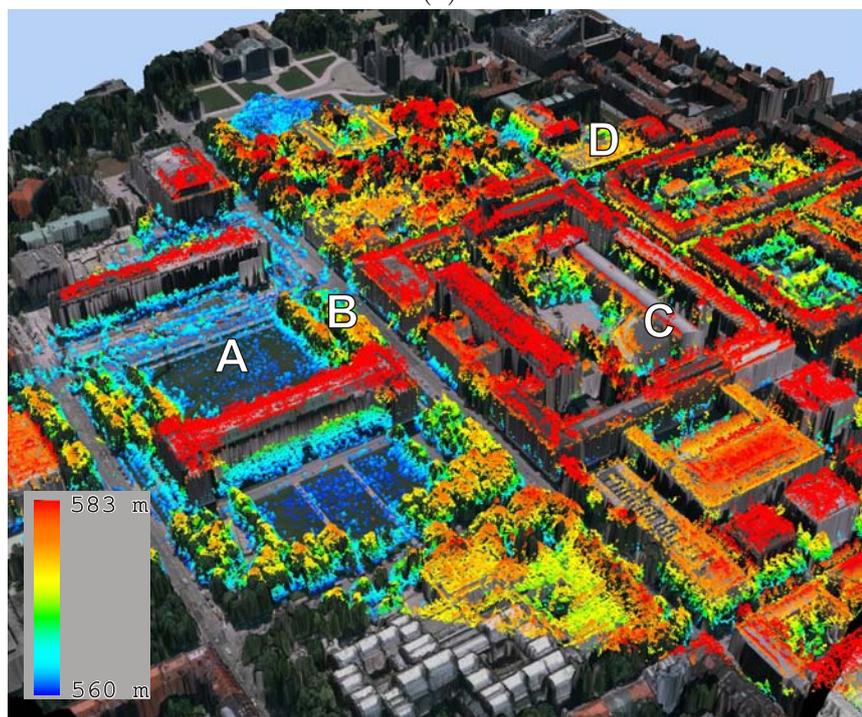
reference terrain height above the WGS84 ellipsoid. The height spacing Δh was 5 cm, the spacing of a surface grid element was 50 cm \times 50 cm. After the reconstruction of one height value for every grid element was completed, unreliable height estimates have been discarded using the threshold $\varepsilon = 0.35$. Afterwards, the resulting 2.5D height grid was 3×3 median filtered.

8.3.1 Reconstruction Result

The MAMBIN SAR reconstruction result can be seen in Fig. 8.20; the reconstruction accuracy is summarized and compared to the conventional backward geocoding based InSAR data fusion in Table 8.1. An overall improvement of the reconstruction quality can be confirmed. Especially the evaluation over all occupied grid elements, gives better results for the MAMBIN SAR method than for non-statistical InSAR data fusion. Figure 8.21 displays the corresponding height error map. In this illustration, several phenomena can be noticed at first glance: First of all, most of the ground areas, which are not affected by shadowing from all radar viewing angles, are well reconstructed – as are most of the roof areas. Second, in spite of the full multi-aspect data availability, street canyons as well as backyards are still mostly discarded. Third, the largest errors occur at the building edges. The particular areas marked by the letters A, B, C, and D will be discussed in further detail in Section 8.3.2.



(a)



(b)

Figure 8.20: Result of the MAMBIInSAR reconstruction. (a) 2.5D surface model displayed as a point cloud. (b) Point cloud overlaid onto a textured height model of the city for better visual comparison of experimental result and reference data. The patches marked by A, B, C, and D are discussed in further detail.

	Global accuracy	Accuracy on homogeneous lawn area (A)	Accuracy on homogeneous roof area (D)
Backward geocoding based multi-aspect InSAR fusion	3.48 m	0.79 m	1.15 m
Multi-aspect multi-baseline SAR interferometry	2.97 m	0.81 m	1.09 m

Table 8.1: Accuracy assessment of the maximum-likelihood based MAMBIInSAR approach developed in this thesis in comparison to non-statistical multi-aspect InSAR data fusion by backward geocoding. The mean absolute deviation was calculated for all occupied grid elements (global accuracy) and for homogeneous patches on the lawn area (A) as well as the roof area (D), respectively.

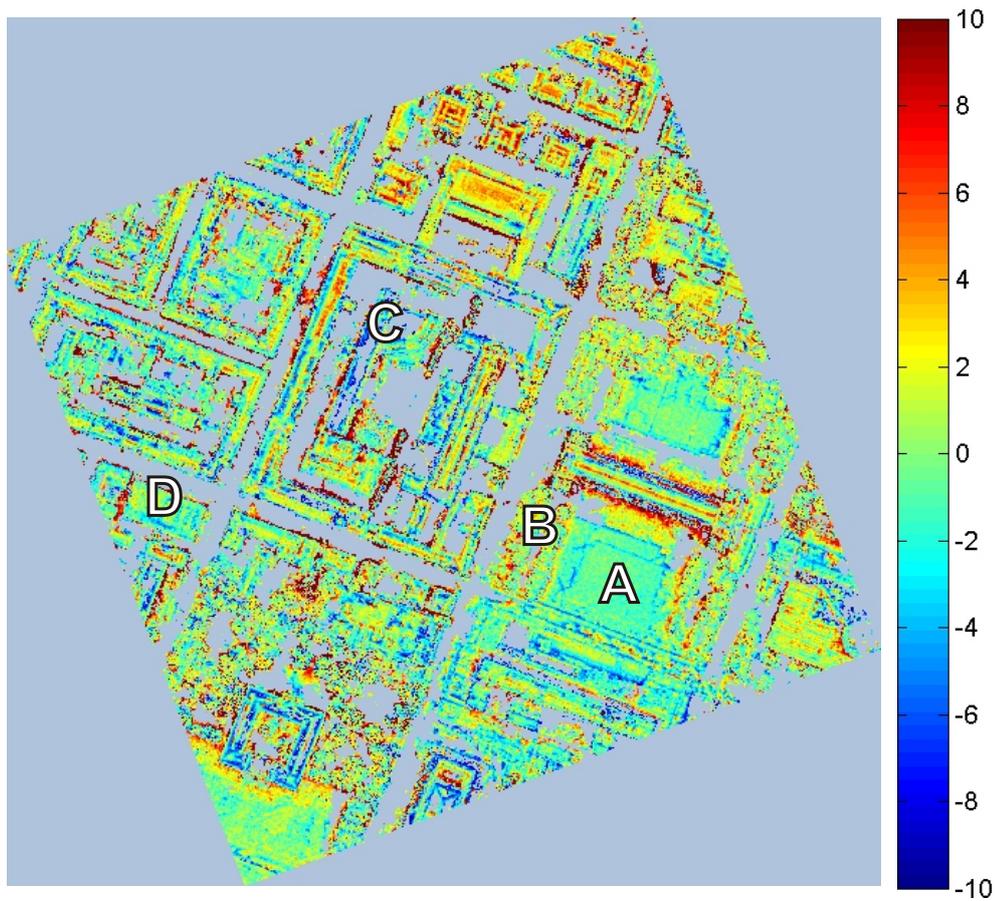


Figure 8.21: Height errors in [m] of the maximum-likelihood based MAMBIInSAR reconstruction.

8.3.2 Discussion

Comments on the overall accuracy

As can be seen from the accuracies of the lawn (A) as well as the roof area (D) listed in Table 8.1, a height reconstruction in complex urban environments based on multi-aspect multi-baseline SAR data leads to decent accuracies for objects of different kinds. For comparison: the mean height error of non-urban DEMs reconstructed from single-pass TanDEM-X imagery using sophisticated operational processing strategies is stated to be 2.76 m by Rossi et al. [2012]. Still, the global mean absolute difference gives the impression that the overall reconstruction accuracy is rather bad. In this regard, several important factors have to be discussed:

First of all, MEMPHIS is an experimental system, where the focusing and calibration of the raw SAR data is still under investigation [Magnard et al., 2012]. Although the necessary motion compensation has been further improved and a phase correction depending on the depression angle has been added for upcoming campaigns [C. Magnard, personal communication, June 10, 2013], the processing would benefit from a precisely known a priori digital height model (DHM) of the scene, which is supposed not to be available in the context of this investigation. Even more important, however, is the fact that both the reconstructed data as well as the LiDAR derived reference data might show slight georeferencing errors leading to height errors at the edges of buildings or other elevated structures (cf. Fig. 8.24). In addition to that, layover pixels may cause the multi-aspect likelihood profiles to blur (cf. Section 6.4.1), also adding to these deviations from the reference data. Last but not least, some of the comparably large errors seem to be connected to certain roof materials: Fig. 8.23 for example shows a zoom into the area denoted by C where unreliable measurements (and therefore discarded heights) and large reconstruction errors appear next to each other. It is assumed that this effect is interrelated with the utilized wavelength and the electromagnetic properties of the imaged surfaces. Apart from the possibilities to optimize the data by exploitation of different radar wavelengths, this problem is inherent to this sensor technology. Therefore, a posteriori measures for outlier elimination and height interpolation have to be applied if DSMs are to be derived operationally.

Detail Discussion

For a detailed discussion of the advantages and shortcomings of the proposed reconstruction procedure, we refer to the areas labeled A, B, C, and D in Figs. 8.20 and 8.21:

A Topographically flat area covered by lawn

Generally, lawn appears highly coherent in single-pass InSAR data leading to a comprehensive reconstruction. However, a certain variance of the reconstructed height values leads to a number of underestimated heights, wherefore only few points are displayed on the lawn in Fig. 8.20 (b).

B Urban Trees

Although a comparison of reconstruction results and reference data acquired at different points in time is always questionable for vegetation, the tree structures seem to be perfectly recovered (Fig. 8.22). Basically, this application is a good example for showing the main advantages of airborne single-pass millimeterwave SAR interferometry in comparison to more conventional spaceborne repeat-pass X-band SAR [Schmitt et al., 2013a].

C Metallic roof structures

In this case, a curved metallic roof structure is shown, which was not reconstructed properly at all (see Fig. 8.23): Either the height values have been discarded by the reliability

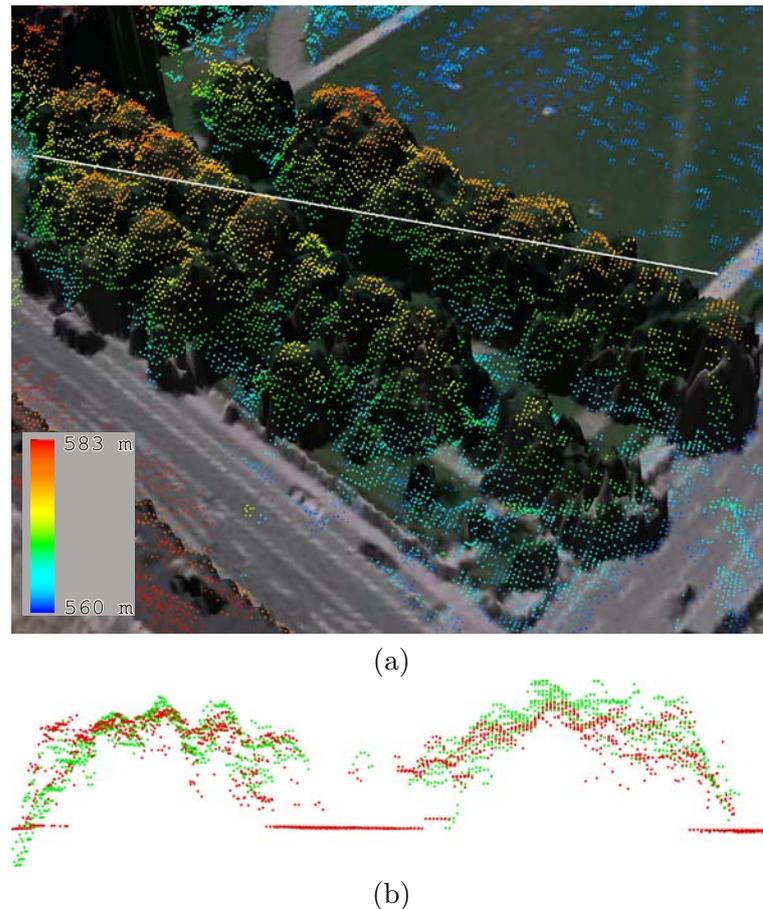


Figure 8.22: Detailed illustration of the trees marked with B in Figs. 8.20 and 8.21, showing that the tree shapes and heights have been perfectly reconstructed. (a) 2.5D point cloud overlaid onto the texturized reference data. (b) Profile corresponding to the white line. The reference LiDAR points are colored red, the MAMBIInSAR reconstruction result is colored green.

thresholding or – where their coherence has exceeded the threshold – they show a severe under- or overestimation. This is a shortage inherent to the wavelength domain employed in radar remote sensing and is assumed to get even worse for longer wavelengths.

D Concrete roof structures

In contrast to some metallic roofs such as the one marked by C, concrete roofs can be reconstructed with better results. Not only are there no discarded height values caused by non-reliable phase measurements; the mean absolute deviation is also comparable to the one achieved over open lawn spaces (A). An enlarged 2D illustration of the reconstruction and the corresponding error map can be seen in Fig. 8.24. It also shows the above-mentioned errors at the building edges.



Figure 8.23: Detailed illustration of the building marked with C in Figs. 8.20 and 8.21. (a) Surface points reconstructed for the Audimax building on TUM main campus. (b) Corresponding optical image acquired during a follow-up flight campaign. Note how the curved metallic roof structures have not been reconstructed well.

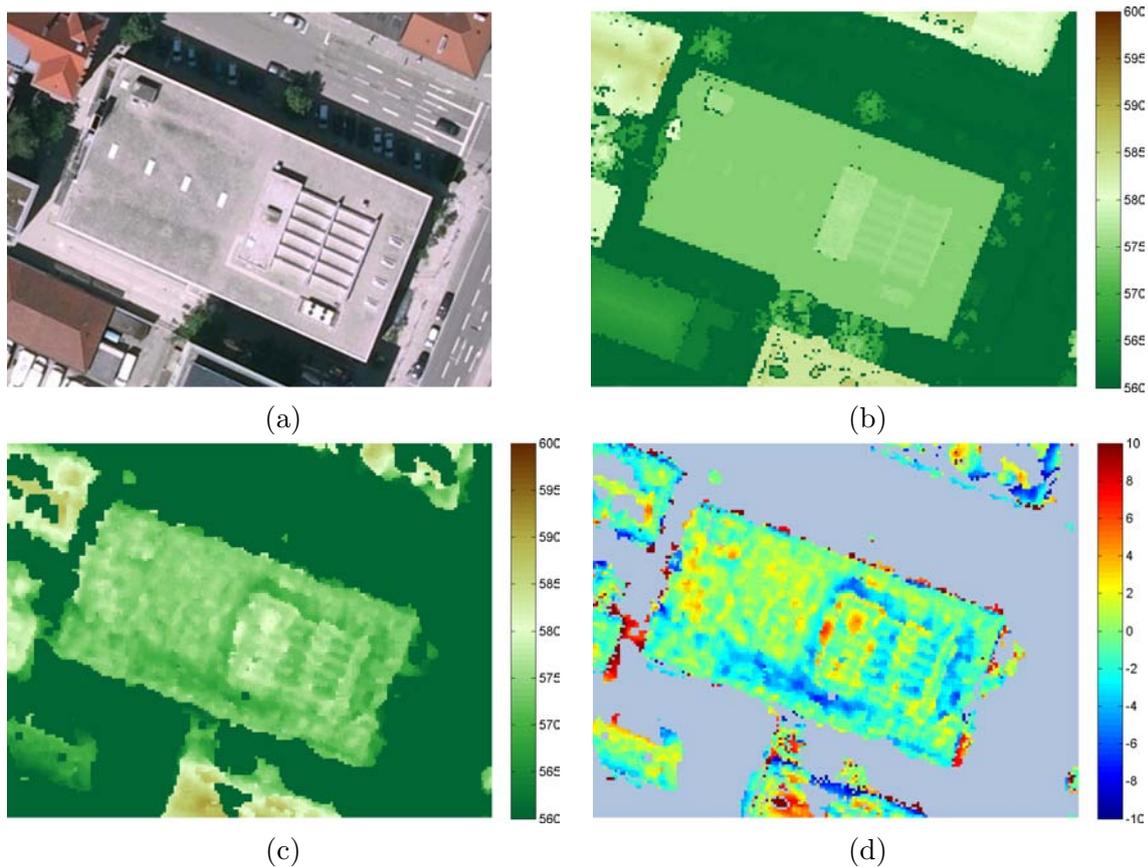


Figure 8.24: Detailed illustration of the flat, concrete-roofed building marked with D in Figs. 8.20 and 8.21. (a) Optical image. (b) LiDAR reference data. (c) MAMBIInSAR reconstruction result. (d) Error image [m]. Many of the outliers appear at the edges of the buildings and are caused by mis-registration between the reference data and the reconstructed data or blurring of the likelihood profiles by layover pixels.

Summary

In summary, a look at the results shows that most parts of the scene can be reconstructed with satisfying precision. Also, the improvements achieved by the statistically sound maximum-likelihood estimation procedure in comparison to conventional backward geocoding based InSAR data fusion become evident, especially if one considers the possibility to include additional prior knowledge (e.g. available rough terrain models) in form of the corresponding probability density functions in the estimation process. The mathematical extension is straight-forward and easy to be implemented. In addition to that: Although the only post-processing applied to the results presented in this paper is a 3×3 median filter, thanks to its regular nature in form of a 2.5D height grid, the data could readily be enhanced by further filtering, outlier removal methods or gap filling procedures in order to receive an accurate, comprehensive DSM.

8.4 Benefit of Multi-Aspect Data Fusion

8.4.1 Analysis of Different Aspect Combinations

The 2.5D nature of the multi-aspect multi-baseline SAR interferometry output allows a straight-forward analysis of the benefit induced by multi-aspect data fusion, especially in combination with the reliability threshold introduced in Section 6.4.2: If a grid element is discarded because it shows a low reliability measure, for high-coherent single-pass data it can well be expected that this is caused by radar shadowing. For the experiment in this section, only a single baseline (i.e. the longest available one with an ambiguity height of 45 m) has been used in order to emphasize the effect of multi-aspect SAR. Figure 8.25 shows the height maps reconstructed by different aspect combinations. The resolution cells colored in gray have been found unreliable ($\varepsilon = 0.35$) and can be considered affected by shadowing therefore. The dark brown pixels, which are frequently found on the lawn patch shown in the zooms for the single-aspect case are caused by choosing the wrong maximum from the ambiguous single-baseline likelihood function (cf. Fig. 6.6) and can be detected due to their large height difference to the LiDAR ground truth.

Figure 8.26 contains bar plots showing the number of shadow pixels and the number of ambiguous grid elements for the same aspect combinations. A more thorough analysis of the benefit provided by multi-aspect as well as multi-baseline data can be found in [Schmitt et al., 2014b].

8.4.2 Discussion

From the analysis of different multi-aspect combinations, several insights can be drawn: First of all, the number of shadowed grid elements leading to erroneous or discarded height values can be significantly reduced by inclusion of complementary aspects. While orthogonal viewing angles already show a certain amount of improvement, it is – as probably expected – better to employ opposing views. The best results, however, can be achieved by a full multi-aspect configuration. It is found that while the combination of two complementary aspects (no matter if orthogonal or from opposing viewing angles) already reduces the amount of shadow pixels for about 20%, a full multi-aspect configuration consisting of four complementary aspects reduces the shadow effect for almost 50%.

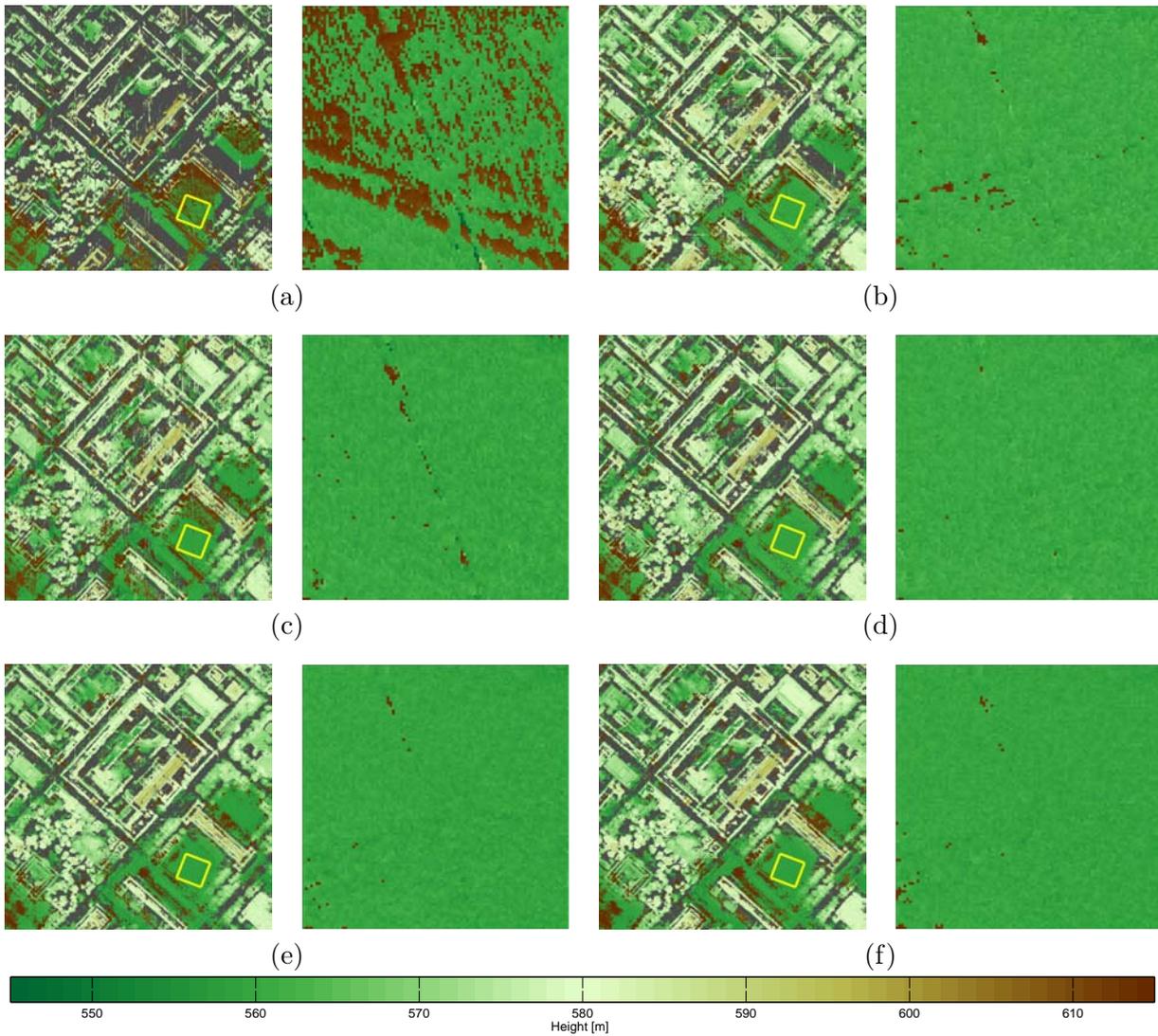


Figure 8.25: Height reconstruction results for different combinations of multi-aspect single-baseline data: (a) Single aspect 15 as a reference, (b) orthogonal aspects 15 and 17, (c) opposing aspects 15 and 16, (d) aspects 15-17, (e) full multi-aspect configuration (15-18), (f) full multi-aspect configuration plus additional aspect acquired from different altitude (14-18).

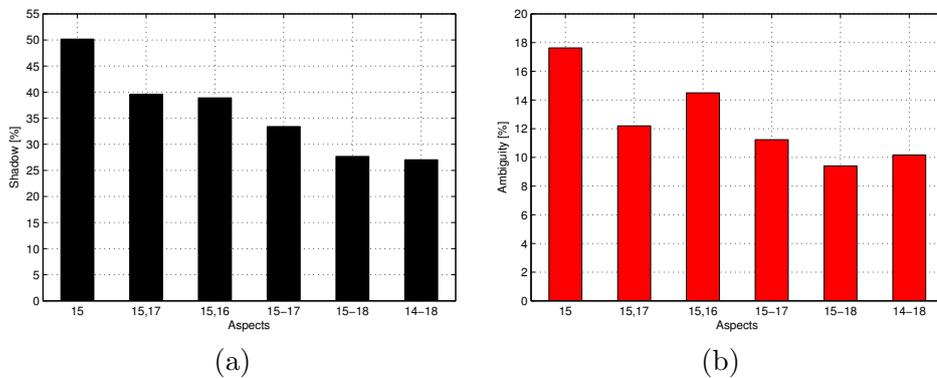


Figure 8.26: Bar plots summarizing (a) the number of shadow-affected grid elements and (b) the number of ambiguous height reconstruction values.

A similar outcome is achieved concerning the ambiguously reconstructed grid cells. Even with respect to this problem, multi-aspect InSAR data can help to reduce ambiguity effects: A full multi-aspect configuration helps to reduce the ambiguities for about 50% – still using just a single baseline per aspect. Another noticeable fact can be drawn from Fig. 8.26 (b): Due to the strongly differing viewing angles, the fusion of two opposing aspects obviously does not lead to many redundant observations of individual grid cells, wherefore the ambiguity reduction is the weakest in this case. Interestingly, however, for both shadowing and ambiguity reduction the addition of data from a track with the same viewing angle but a different height as an already used track, does not bring any further improvements.



Figure 8.27: Oblique-view optical image of the test subset (©Google 2013). Note the free-standing wall in front of the main entrance of the State Museum of Egyptian Art.

8.5 Comparison of 3D-Reconstruction Results

The final experiment is intended to compare the reconstruction results of the MAMBIInSAR framework and the ML-TomoSAR algorithm. For this task, a subset of the TUM test area, which is located around the State Museum of Egyptian Art in Munich, Germany, is investigated. This is the same area that was already used for the TomoSAR experiments described in Section 8.2. Since the results are shown from a different perspective to enhance the visual impression, another oblique-view optical image is shown in Fig. 8.27. In order to create a fair comparison, both methods were applied to the same data, i.e. tracks 15-18. Since they were flown at a lower altitude and therefore with a more shallow off-nadir angle than track 14 (which was used in Section 8.2), the Rayleigh resolution converted to height is only about 46 m in this case, i.e. almost 10% worse than before. In addition, the height search intervals Δh were set to 0.5 m for both approaches.

The point clouds derived by ML-tomography from the individual aspects are fused as described in Section 6.2: First, they are geocoded to the UTM coordinate system. Afterwards, they are fused using an imposed voxel space with voxels of $0.5 \times 0.5 \times 0.5 \text{ m}^3$, and the constraint that each voxel to be considered for further processing has to contain points from at least two different aspects. For each of these voxels then the mean of all contained 3D points is returned. Since also the MAMBIInSAR reconstruction started from a pre-defined surface grid with a cell spacing of $0.5 \text{ m} \times 0.5 \text{ m}$, this parameter set-up ensures a high degree of comparability and prevents a biased analysis that might have been caused by different discretization of the data owed to different computational costs.

8.5.1 Reconstruction Results

The reconstruction results for both methods are visualized in Fig. 8.28. For numerical evaluations, the distance of each reconstructed 3D point to the 3D reference data has to be calculated. This can be achieved in several ways: A first intuitive assessment is based on an assignment of each point to its nearest neighbor in the LiDAR reference point cloud, which again can be realized by

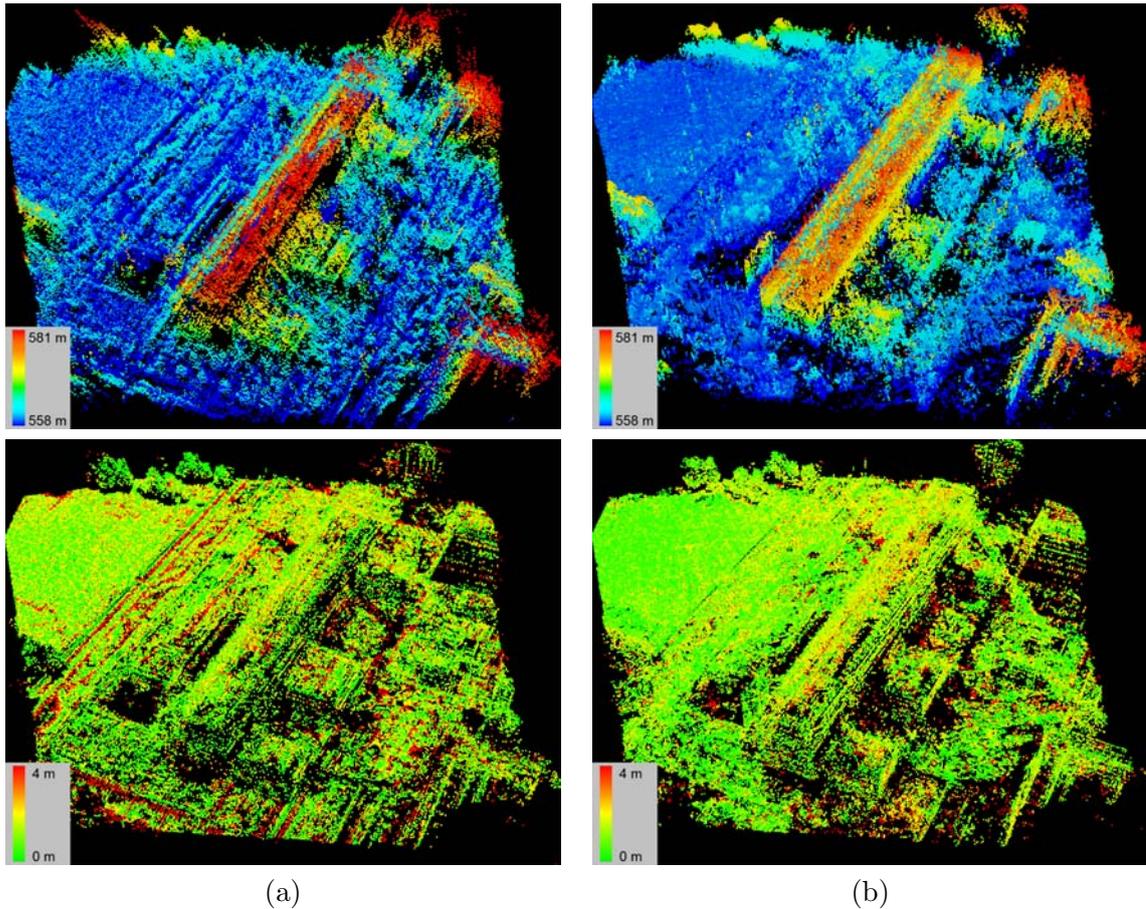


Figure 8.28: Comparison of (a) MAMBIInSAR and (b) MA-TomoSAR reconstruction results. The top row shows the point clouds, the bottom row shows the LiDAR points used in the k -d tree evaluation, colored by the errors their corresponding points showed.

a k -d tree search (cf. Section 6.2). Afterwards, the euclidian distance to this nearest neighbor is calculated.

This, however, can lead to a rather pessimistic estimate of the accuracies if the point density in the reference data is lower than in the reconstructed data (see 8.29). In this case, a point belonging to the facade, for example, is not evaluated by its normal distance to the facade plane, but by its (larger) distance to the nearest point on the facade. Therefore, a more sophisticated accuracy assessment can be realized by selecting the three nearest neighbors of each point by the k -d tree and calculating the point's distance to the corresponding plane.

Unfortunately, also this evaluation strategy might lead to biased results, especially at building edges, where semantically unrealistic planes might lead to favorable distances. Therefore, in a third approach, the ten nearest neighbors of each reconstructed point are chosen and a least squares plane is fitted to these points. The median distances of all points in the reconstruction results for all three evaluations are summarized in Table 8.2.

8.5.2 Discussion

Looking at the reconstruction results of both the MAMBIInSAR and the MA-TomoSAR approaches one can see that the MA-TomoSAR point cloud is both more accurate (with a median error well below 1 m for all three evaluation strategies in comparison to more than 1 m) and more comprehensive (with a point density of 8.6 pts/m² in comparison to 3.7 pts/m²) than the MAM-

	MAMBIInSAR	MA-TomoSAR
point density	3.7 pts/m ²	8.6 pts/m ²
number of points	ca. 128,000	a. 290,000
<i>median distance</i>		
nearest point	1.44 m	0.76 m
nearest plane	1.02 m	0.49 m
nearest LS plane	1.32 m	0.64 m

Table 8.2: Numerical comparison of MAMBIInSAR and MA-TomoSAR reconstruction results.

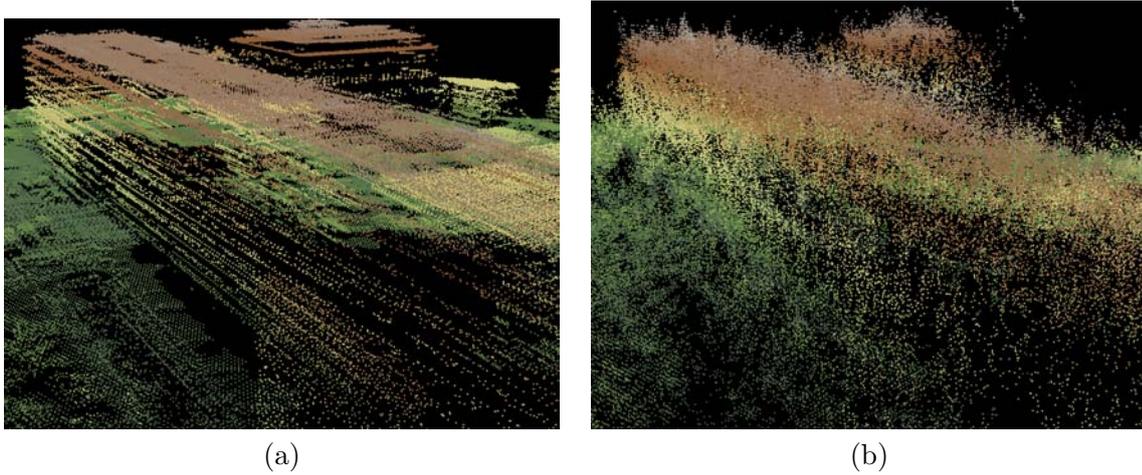


Figure 8.29: Example illustration of different point densities in (a) the LiDAR reference data and (b) the point cloud reconstructed by MA-TomoSAR.

BInSAR point cloud. This indicates two main insights: First, the TomoSAR reconstruction seems to be more precise, which is also confirmed by Fig. 8.28, where the MAMBIInSAR result shows larger errors especially on the road and the scene parts that have been covered in the shadows of the trees in at least some of the aspects. Second, the fact that TomoSAR enables a real 3D reconstruction instead of just creating a 2.5D height map obviously greatly enhances the number of reconstructed points. This is particularly impressive for the free-standing wall in front of the main entrance of the State Museum of Egyptian Art, which is not visible in the MAMBIInSAR point cloud at all, but was nicely reconstructed by MA-TomoSAR.

9 Conclusions and Perspectives

9.1 Summary and Conclusion

This dissertation has investigated the benefits and possibilities that arise from the utilization of interferometric synthetic aperture radar data acquired from multiple viewing angles, and using multiple receiving antennas. Within this context, special emphasis has been put on the potential and limits of single-pass multi-baseline sensor configurations that offer the advantage of simultaneous acquisition of full InSAR stacks with highly coherent phase observations. In order to exploit this kind of data as beneficially as possible for urban area reconstruction, three main methodical contributions were described:

For most applications of multi-baseline SAR interferometry and its extensions, the complex covariance matrices of all the pixels (i.e. resolution cells) in the stack of coregistered SAR images have to be estimated. This, however, is a non-trivial task especially in heterogeneous urban areas, where the local stationarity of the necessary samples cannot be presupposed. In this thesis, two approaches for spatially adaptive covariance matrix estimation have been proposed and evaluated with respect to their adaptiveness and filtering efficiency.

A second contribution is the description of a maximum-likelihood-based algorithm for SAR tomography. This procedure is a necessary tool for the separation of scatterers layovered in a single resolution cell due to the side-looking SAR imaging geometry. Using this method, it has been demonstrated that SAR tomography is possible with single-pass data providing only a short overall baseline (i.e. elevation aperture) and a limited number of observations per individual pixel.

Finally, driven by considerations of how to fuse InSAR data acquired from multiple viewing angles, another maximum-likelihood-based estimation framework for the simultaneous fusion of multi-aspect and multi-baseline InSAR data was developed. With this formulation, interferometric SAR observations of almost arbitrarily configured campaigns can be fused exploiting potential redundancies beneficially. In addition, a voxel-space-based fusion of 3D point clouds reconstructed from multi-aspect InSAR or TomoSAR was proposed. It has been shown that both MA-Tomography and the MAMBIInSAR framework lead to comparable results with the main difference that TomoSAR enables the reconstruction of real 3D point clouds including facade information, whereas MAMBIInSAR only creates 2.5D height maps due to its backward-geocoding-based nature. Using these processing strategies, comprehensive urban surface models can be created with accuracies below 1 m. Although 3D reconstruction by optical stereogrammetry or airborne laserscanning still provides more accurate results, this thesis was able to prove that urban areas can also well be analyzed three-dimensionally by single-pass synthetic aperture radar interferometry, which provides a valuable benefit if bad-weather conditions are faced, while data acquisition is yet time-critical. This might well be the case during defense-related reconnaissance missions or in disaster scenarios, where a quick assessment of the affected areas is necessary.

Although the presented work has focused on airborne SAR and although the future can never be forecast with certainty, the findings of this thesis might also be of interest for future multi-

satellite missions offering a flexibility in data acquisition that is currently available only to airborne sensors.

9.2 Future Work

The experiments carried out for this thesis have revealed a number of potential future research directions: Among the most interesting certainly is the sophistication of the maximum-likelihood SAR tomography method: Since the algorithm is based on a multi-dimensional grid search, which has to be carried out for each resolution cell separately, a GPU-based, parallelized implementation could greatly reduce processing times. In addition, more research effort could be invested in a priori model order selection techniques similar to the ones based on information theoretic criteria. In the end, this model order selection step, which is a prerequisite not only for ML-TomoSAR but also for competing approaches like weighted subspace fitting, is the critical bound in terms of elevation (or height, respectively) resolution for these methods. Furthermore, it could be shown in this thesis that ML-TomoSAR would potentially be able to deal with an overestimated number of scatterers, if an additional thresholding step were employed, which would re-join two wrongly estimated pseudo-scatterers. In combination with the before-mentioned efficient implementation, this proceeding is expected to provide a robust solution for layover separation in urban areas even for closely spaced scattering contributions.

Besides these potential amendments to the proposed TomoSAR algorithm, also an important improvement to the MAMBIInSAR framework can be identified: Here, one of the major drawbacks is that layover is not explicitly modelled, but expected to be mitigated by building the joint likelihood function of multi-baseline InSAR data from multiple aspects. An interesting approach to solve this deficiency came to the author's knowledge during finalization of this thesis: Fornaro et al. [2013] propose a method based on principal component analysis for reducing the covariance matrix of a resolution-cell containing layover to the dominant scattering contribution. Maybe this could serve as a valuable pre-processing step leading to even clearer joint likelihood functions in layover areas. Apart from that, a more thorough investigation of the potential provided by the probabilistic volume model, that is implicitly created during MAMBIInSAR application, could be a valuable perspective for future research.

Last, but not least, it has to be mentioned that all methods described in this thesis rely on the assumption of circularly-symmetric complex Gaussian scattering mechanisms, which is supposed to be met quite well for millimeterwave SAR data. However, it could well be worthwhile to investigate a generalized Gaussian distribution or the special cases of constant-plus-Gaussian or Rician scattering, as these might frequently occur in very high resolution data of urban areas especially for longer wavelengths [Davis et al., 2007].

Bibliography

- Aguilera E, Nannini M, Reigber A (2013) Wavelet-based compressed sensing for SAR tomography of forested areas. *IEEE Transactions on Geoscience and Remote Sensing*, 51 (12): 5283–5295.
- Akaike H (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19 (6): 716–723.
- Ashok A, Wilkinson A (2001) Topographic mapping with multiple antenna SAR interferometry: a Bayesian model-based approach. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 2058–2060.
- Bamler R, Hartl P (1998) Synthetic aperture radar interferometry. *Inverse Problems*, 14 (4): R1–R54.
- Baselice F, Budillon A, Ferraioli G, Pascazio V (2009a) Layover solution in SAR imaging: A statistical approach. *IEEE Geoscience and Remote Sensing Letters*, 6 (3): 577–581.
- Baselice F, Ferraioli G (2012) Statistical edge detection in urban areas exploiting SAR complex data. *IEEE Geoscience and Remote Sensing Letters*, 9 (2): 185–189.
- Baselice F, Ferraioli G, Pascazio V (2009b) DEM Reconstruction in Layover Areas From SAR and Auxiliary Input Data. *IEEE Geoscience and Remote Sensing Letters*, 6 (2): 253–257.
- Batu O, Cetin M (2011) Parameter selection in sparsity-driven SAR imaging. *IEEE Transactions on Aerospace and Electronic Systems*, 47 (4): 3040–3050.
- Bentley JL (1975) Multidimensional binary search trees used for associative searching. *Communications of the ACM*, 18 (9): 509–517.
- Bolter R (2001) Buildings from SAR: Detection and Reconstruction of Buildings from Multiple View High Resolution Interferometric SAR Data. *PhD thesis*, Graz University of Technology.
- Bolter R, Leberl F (2000) Phenomenology-based and interferometry-guided building reconstruction from multiple SAR images. In: *Proceedings of 3rd European Conference on Synthetic Aperture Radar*: 687–690.
- Buades A, Coll B, Morel JM (2005) A non-local algorithm for image denoising. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*: 60–65.
- Budillon A, Evangelista A, Schirinzi G (2011) Three-dimensional SAR focusing from multipass signals using compressive sampling. *IEEE Transactions on Geoscience and Remote Sensing*, 49 (1): 488–499.
- Burkhart G, Bergen Z, Carande R, Hensley W, Bickel D, Fellerhoff J (1996) Elevation correction and building extraction from interferometric SAR imagery. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 659–661.
- Cai B, Liang D, Dong Z (2008) A new adaptive multiresolution noise-filtering approach for SAR interferometric phase images. *IEEE Geoscience and Remote Sensing Letters*, 5 (2): 266–270.
- Capon J (1969) High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57 (8): 1408–1418.

- Ciuc M, Bolon P, Trouve E, Buzuloiu V, Rudant J (2001) Adaptive neighborhood speckle removal in multi-temporal synthetic aperture radar images. *Applied Optics*, 40 (32): 5954–5966.
- Cloude S, Papathanassiou K (1998) Polarimetric SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 36 (5): 1551–1565.
- Corsini G, Diani M, Lombardini F, Pinelli G (1999) Simulated analysis and optimization of a three-antenna airborne InSAR system for topographic mapping. *IEEE Transactions on Geoscience and Remote Sensing*, 37 (5): 2518–2529.
- Costantini M (1998) A novel phase unwrapping method based on network programming. *IEEE Transactions on Geoscience and Remote Sensing*, 36 (3): 813–821.
- Crosetto M, Monserrat O, Herrera G (2010) Urban applications of persistent scatterer interferometry. In: Soergel U (ed) *Radar Remote Sensing of Urban Areas*. Dordrecht, Netherlands: Springer Science+Business Media.
- Cumming IG, Wong FHC (2005) *Digital signal processing of synthetic aperture radar data: algorithms and implementation*. Norwood, MA, USA: Artech House.
- Curlander JC, McDonough RN (1991) *Synthetic Aperture Radar – Systems and Signal Processing*. New York, NY, USA: John Wiley & Sons.
- Danklmayer A, Chandra M (2009) Precipitation Effects for Ka-band SAR. In: *Proceedings of Advanced RF Sensors for Earth Observation*
- Davidson GW, Bamler R (1999) Multiresolution phase unwrapping for SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 37 (1): 163–174.
- Davis M, Bidigare P, Chang D (2007) Statistical modeling and ML parameter estimation of complex SAR imagery. In: *Proceedings of Asilomar Conference on Signals, Systems and Computers*: 500–502.
- De Zan F (2008) Optimizing SAR interferometry for decorrelating scatterers. *PhD thesis*, Politecnico di Milano.
- Deledalle CA, Denis L, Tupin F (2011) NL-InSAR: Nonlocal interferogram estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 49 (4): 1441–1452.
- Deledalle CA, Denis L, Tupin F, Reigber A, Jäger M (2013) NL-SAR: a unified non-local framework for resolution-preserving (Pol)(In)SAR denoising. *Technical report*, hal-00844118.
- Denis L, Tupin F, Darbon J, Sigelle M (2009) Joint regularization of phase and amplitude of InSAR data: application to 3-D reconstruction. *IEEE Transactions on Geoscience and Remote Sensing*, 47 (11): 3774–3785.
- Eineder M (2003) Efficient simulation of SAR interferograms of large areas and of rugged terrain. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (6): 1415–1427.
- Eineder M, Adam N (2005) 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*, 43 (1): 24–36.
- Essen H (2010) Airborne remote sensing at millimeter wave frequencies. In: Soergel U (ed) *Radar Remote Sensing of Urban Areas*. Dordrecht, Netherlands: Springer Science+Business Media.
- Essen H, Brehm T, Boehmsdorff S, Stilla U (2007) Multibaseline interferometric SAR at millimeterwaves – test of an algorithm on real data and a synthetic scene. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36 (3/W49B): 35–39.

- Ferraioli G (2008) Multichannel SAR interferometry based on statistical signal processing. *PhD thesis*, Universita degli Studi di Napoli Parthenope.
- Ferraiuolo G, Pascazio V, Schirinzi G (2004) Maximum a posteriori estimation of height profiles in InSAR imaging. *IEEE Geoscience and Remote Sensing Letters*, 1 (2): 66–70.
- Ferretti A, Bianchi M, Prati C, Rocca F (2005) Higher-order permanent scatterers analysis. *EURASIP Journal on Applied Signal Processing*, 2005 (20): 3231–3242.
- Ferretti A, Fumagalli A, Novali F, Prati C, Rocca F, Rucci A (2011) A new algorithm for processing interferometric data-stacks: SqueeSAR. *IEEE Transactions on Geoscience and Remote Sensing*, 49 (9): 3460–3470.
- Ferretti A, Prati C, Rocca F (2001) Permanent scatterers in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, 39 (1): 8–20.
- Flynn TJ (1997) Two-dimensional phase unwrapping with minimum weighted discontinuity. *Journal of the Optical Society of America A – Optics, Image Science, And Vision*, 14 (10): 2692–2701.
- Fornaro G, Franceschetti G, Lanari R (1996a) Interferometric SAR phase unwrapping using Green’s formulation. *IEEE Transactions on Geoscience and Remote Sensing*, 34 (3): 720–727.
- Fornaro G, Franceschetti G, Lanari R, Sansosti E (1996b) Robust phase-unwrapping techniques: a comparison. *Journal of the Optical Society of America A – Optics, Image Science, And Vision*, 13 (12): 2355–2366.
- Fornaro G, Pauciuolo A, Reale D, Verde S (2013) SAR coherence tomography: A new approach for coherent analysis of urban areas. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 73–76.
- Fornaro G, Serafino F (2006) Imaging of single and double scatterers in urban areas via SAR tomography. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (12): 3497–3505.
- Franceschetti G, Lanari R (1999) *Synthetic Aperture Radar Processing*. Boca Raton, FL, USA: CRC Press.
- Frey O, Meier E (2011) 3-D time-domain SAR imaging of a forest using airborne multibaseline data at L- and P-bands. *IEEE Transactions on Geoscience and Remote Sensing*, 49 (10): 3660–3664.
- Frey O, Morsdorf F, Meier E (2008) Tomographic imaging of a forested area by airborne multi-baseline P-band SAR. *Sensors*, 8 (9): 5884–5896.
- Frost V, Stiles J, Shanmugan K, Holtzman J (1982) A model for radar images and its application to adaptive digital filtering of multiplicative noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-4 (2): 157–166.
- Gamba P, Houshmand B (2000) Digital surface models and building extraction: a comparison of IFSAR and LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 38 (4): 1959–1968.
- Gamba P, Houshmand B, Saccani M (2000) Detection and extraction of buildings from interferometric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 38 (1): 611–618.
- Gao G (2010) Statistical modeling of SAR images: a survey. *Sensors*, 10: 775–795.
- Gens R (2003) Two-dimensional phase unwrapping for radar interferometry: Developments and new challenges. *International Journal of Remote Sensing*, 24 (4): 703–710.
- Gernhardt S (2012) Deformation monitoring of single buildings using meter-resolution SAR data in PSI. *ISPRS Journal of Photogrammetry and Remote Sensing*, 73: 68–79.
- Gernhardt S, Adam N, Eineder M, Bamler R (2010) Potential of very high resolution SAR for persistent scatterer interferometry in urban areas. *Annals of GIS*, 16 (2): 103–111.

- Gernhardt S, Cong X, Eineder M, Hinz S, Bamler R (2012) Geometrical fusion of multitrack PS point clouds. *IEEE Geoscience and Remote Sensing Letters*, 9 (1): 38–42.
- Getreuer P (2012) Rudin-Osher-Fatemi total variation denoising using split Bregman. *Image Processing On Line*, 2012: available online.
- Ghiglia DC, Pritt MD (1998) *Two-dimensional phase unwrapping - theory, algorithms and software*. Hoboken, NJ, USA: John Wiley & Sons.
- Ghiglia DC, Romero LA (1994) Robust two-dimensional weighted and unweighted phase unwrapping that uses fast transforms and iterative methods. *Journal of the Optical Society of America A – Optics, Image Science, And Vision*, 11 (1): 107–117.
- Gini F, Bordoni F (2003) On the behavior of information theoretic criteria for model order selection of InSAR signals corrupted by multiplicative noise. *Signal Processing*, 83 (5): 1047–1063.
- Gini F, Lombardini F (2005) Multibaseline cross-track SAR interferometry: a signal processing perspective. *IEEE Aerospace and Electronics Systems Magazine*, 20 (8): 71–93.
- Gini F, Lombardini F, Montanari M (2002) Layover solution in multi-baseline SAR interferometry. *IEEE Transactions on Aerospace and Electronic Systems*, 38 (4): 1344–1356.
- Goldstein RM, Zebker HA, Werner CL (1988) Satellite radar interferometry: two-dimensional phase unwrapping. *Radio Science*, 23 (4): 713–720.
- Goldstein T, Osher S (2009) The split Bregman method for L1-regularized problems. *SIAM Journal on Imaging Sciences*, 2 (2): 323–343.
- Golub GH, Van Loan CF (1996) *Matrix Computations*. Baltimore, MD, USA: John Hopkins University Press.
- Goodman J (1975) Statistical properties of laser speckle patterns. In: Dainty J, Ennos A, Françon M, Goodman J, McKechnie T, Parry G, Goodman J (eds) *Laser Speckle and Related Phenomena*, volume 9 of *Topics in Applied Physics* (pp. 9–75). Berlin, Germany: Springer-Verlag.
- Graham L (1974) Synthetic interferometer radar for topographic mapping. *Proceedings of the IEEE*, 62 (6): 763–768.
- Hanssen RF (2001) *Radar interferometry – Data Interpretation and Error Analysis*. New York, NY, USA: Kluwer Academic Publishers.
- Hebel M (2012) Änderungsdetektion in urbanen Gebieten durch objektbasierte Analyse und schritthaltenen Vergleich von Multi-Aspekt ALS-Daten [Change detection in urban areas by object-oriented analysis and inline comparison of multi-aspect ALS data]. *PhD thesis*, Technische Universität München.
- Hebel M, Stilla U (2007) Automatic registration of laser point clouds of urban areas. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36 (3/W49A): 13–18.
- Hein A (2004) *Processing of SAR Data – Fundamentals, Signal Processing, Interferometry*. Berlin: Springer-Verlag.
- Hellwich O, Ebner H (2000) Geocoding SAR interferograms by least squares adjustment. *ISPRS Journal of Photogrammetry and Remote Sensing*, 55 (4): 277–288.
- Homer J, Longstaff I, Callaghan G (1996) High resolution 3-D SAR via multi-baseline interferometry. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 796–798.

- Huang Y, Ferro-Famil L, Reigber A (2012) Under-foilage object imaging using SAR tomography and polarimetric spectral estimators. *IEEE Transactions on Geoscience and Remote Sensing*, 50 (6): 2213–2225.
- Jakowatz CV, Wahl DE, Eichel PH, Ghiglia DC, Thompson PA (1996) *Spotlight-Mode Synthetic Aperture Radar: A Signal Processing Approach*. New York: Springer Science+Business Media.
- Just D, Bamler R (1994) Phase statistics of interferograms with applications to synthetic aperture radar. *Applied Optics*, 33 (20): 4361–4368.
- Kampes BM (2006) *Radar Interferometry – Persistent Scatterer Technique*. Dordrecht, The Netherlands: Springer.
- Klausing H, Holpp W, eds (2000) *Radar mit realer und synthetischer Apertur [Radar with real and synthetic aperture]*. Munich, Germany: Oldenbourg Wissenschaftsverlag.
- Koch KR (2007) *Introduction to Bayesian Statistics*. Berlin: Springer-Verlag, 2nd edition.
- Krieger G, Moreira A (2005) Multistatic SAR satellite formations: Potentials and challenges. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 2680–2684.
- Krim H, Viberg M (1996) Two decades of array signal processing research: the parametric approach. *IEEE Signal Processing Magazine*, 13 (4): 67–94.
- Kuan D, Sawchuk A, Strand T, Chavel P (1985) Adaptive noise smoothing filter for images with signal-dependent noise. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-7 (2): 165–177.
- Leberl FW (1990) *Radargrammetric Image Processing*. Boston, MA, USA: Artech House.
- Lee JS (1980) Digital image enhancement and noise filtering by use of local statistics. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-2 (2): 165–168.
- Lee JS, Hoppel K, Mango S, Miller A (1994) Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 32 (5): 1017–1028.
- Lee JS, Papathanassiou K, Ainsworth T, Grunes M, Reigber A (1998) A new technique for noise filtering of SAR interferometric phase images. *IEEE Transactions on Geoscience and Remote Sensing*, 36 (5): 1456–1465.
- Lenzmann L, Lenzmann E (2004) Strenge Auswertung des nichtlinearen Gauss-Helmert-Modells [Strict evaluation of the nonlinear Gauss-Helmert model]. *Allgemeine Vermessungsnachrichten*, 111 (2): 68–73.
- Liebe HJ (1985) An updated model for millimeter wave propagation in moist air. *Radio Science*, 20 (5): 1069–1089.
- Lillesand TM, Kiefer RW, Chipman JW (2004) *Remote Sensing and Image Interpretation*. New York, NY, USA: John Wiley & Sons, 5 edition.
- Lombardini F, Ender J, Rößing L, Galletto M, Verrazzani L (2004) Experiments of interferometric layover solution with the three-antenna airborne AER-II SAR system. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 3341–3344.
- Lombardini F, Gini F (2005) Model order selection in multi-baseline interferometric radar systems. *EURASIP Journal on Applied Signal Processing*, 2005 (20): 3206–3219.
- Lombardini F, Montanari M, Gini F (2003) Reflectivity estimation for multibaseline interferometric radar imaging of layover extended sources. *IEEE Transactions on Signal Processing*, 51 (6): 1508–1519.
- Lopez-Martinez C, Fabregas X (2003) Polarimetric SAR speckle noise model. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (10): 2232–2242.

- Lord R (2000) Aspects of stepped-frequency processing for low-frequency SAR systems. *PhD thesis*, University of Cape Town.
- Magnard C, Brehm T, Essen H, Meier E (2012) High resolution MEMPHIS SAR data processing and applications. In: *PIERS Proceedings*: 328–332.
- Massonet D, Souyris JC (2008) *Imaging with Synthetic Aperture Radar*. Lausanne: EPFL Press.
- Meglio F, Pascazio V, Schirinzi G (2006) Joint statistical distribution of multi-baseline SAR interferograms. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 2549–2552.
- Meyer FJ (2004) Simultane Schätzung von Topographie und Dynamik polarer Gletscher aus multi-temporalen SAR Interferogrammen [Simultaneous estimation of topography and dynamics of polar glaciers from multi-temporal SAR interferograms]. *PhD thesis*, Technische Universität München.
- Mikhail EM (1976) *Observations and Least Squares*. New York: IEP.
- Nannini M, Scheiber R, Horn R (2008) Imaging of targets beneath foliage with SAR tomography. In: *Proceedings of 7th European Conference on Synthetic Aperture Radar*: on CD.
- Neitzel F (2010) Generalization of total least-squares on example of unweighted and weighted 2D similarity transformation. *Journal of Geodesy*, 84 (12): 751–762.
- Neitzel F, Petrovic S (2008) Total Least Squares (TLS) im Kontext der Ausgleichung nach kleinsten Quadraten am Beispiel der ausgleichenden Geraden [Total least squares (TLS) in the context of least-squares adjustment on the example of line adjustment]. *Zeitschrift für Geodäsie, Geoinformation und Landmanagement (zfv)*, 133 (3): 141–148.
- Oliver C, Quegan S (2004) *Understanding Synthetic Aperture Radar Images*. Raleigh, NC, USA: SciTech Publishing.
- Ollila E, Koivunen V (2003) Robust antenna array processing using M-estimators of pseudo-covariance. In: *14th IEEE Proceedings on Personal, Indoor and Mobile Radio Communications*: 2659–2663.
- Oriot H, Cantalloube H (2008) Circular SAR imagery for urban remote sensing. In: *Proceedings of 7th European Conference on Synthetic Aperture Radar*: 205–208.
- Palm S, Maresch A, Stilla U (2013) Investigation on circular mapping by FMCW-SAR on small airplanes. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XL-1 (W1): 281–286.
- Palm S, Oriot H, Cantalloube HM (2012) Radargrammetric DEM extraction over urban area using circular SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 50 (11): 4720–4725.
- Palm S, Wahlen A, Stanko S, Pohl N, Wellig P, Stilla U (2014) Real-time on-board processing and ground based monitoring of FMCW-SAR videos. In: *Proceedings of 10th European Conference on Synthetic Aperture Radar*: 89–92.
- Parizzi A, Brcic R (2011) Adaptive InSAR stack multilooking exploiting amplitude statistics: a comparison between different techniques and practical results. *IEEE Geoscience and Remote Sensing Letters*, 8 (3): 441–445.
- Perissin D, Rocca F (2006) High-accuracy urban DEM using permanent scatterers. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (11): 3338–3347.
- Raggam H, Gutjahr K (2000) InSAR block parameter adjustment. In: *Proceedings of 3rd European Conference on Synthetic Aperture Radar*: 493–496.

- Reigber A, Alivizatos E, Potsis A, Moreira A (2006) Extended wavenumber-domain synthetic aperture radar focusing with integrated motion compensation. *IEE Proceedings – Radar, Sonar and Navigation*, 153 (3): 301–310.
- Reigber A, Moreira A (2000) First demonstration of airborne SAR tomography using multibaseline L-band data. *IEEE Transactions on Geoscience and Remote Sensing*, 38 (5): 2142–2152.
- Restrepo MI, Mayer BA, Ulusoy AO, Mundy JL (2012) Characterization of 3D volumetric probabilistic scenes for object recognition. *IEEE Journal of Selected Topics in Signal Processing*, 6 (5): 522–537.
- Richards JA (2009) *Remote sensing with imaging radar*. Berlin: Springer-Verlag.
- Richards MA (2007) A beginner’s guide to interferometric SAR concepts and signal processing. *IEEE Aerospace and Electronics Systems Magazine*, 22 (9): 5–29.
- Rissanen J (1978) Modeling by shortest data description. *Automatica*, 14 (5): 465–471.
- Rodriguez E, Martin J (1992) Theory and design of interferometric synthetic aperture radars. *IEE Proceedings F – Radar and Signal Processing*, 139 (2): 147–159.
- Rosen PA, Hensley S, Joughin IR, Li FK, Madsen SN, Rodriguez E, Goldstein RM (2000) Synthetic Aperture Radar Interferometry. *Proceedings of the IEEE*, 88 (3): 333–382.
- Rossi C, Gonzalez F, Fritz T, Yague-Martinez N, Eineder M (2012) TanDEM-X calibrated raw DEM generation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 73: 12–20.
- Rößing L, Ender J (2001) Multi-antenna SAR tomography using super resolution techniques. *Frequenz*, 55 (3-4): 123–128.
- Rudin LI, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60 (1-4): 259–268.
- Sauer S (2008) Interferometric SAR Remote Sensing of Urban Areas at L-Band Using Multibaseline and Polarimetric Spectral Analysis Techniques. *PhD thesis*, Université de Rennes I.
- Sauer S, Ferro-Famil L, Reigber A, Pottier E (2009) Polarimetric dual-baseline InSAR height building estimation at L-band. *IEEE Geoscience and Remote Sensing Letters*, 6 (3): 408–412.
- Schimpf H, Essen H, Boehmsdorff S, Brehm T (2002) MEMPHIS – a fully polarimetric experimental radar. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 1714–1716.
- Schmidt R (1986) Multiple emitter location and signal parameter estimation. *IEEE Transactions on Antennas and Propagation*, 34 (3): 276–280.
- Schmitt M, Brück A, Schönberger J, Stilla U (2013a) Potential of airborne single-pass millimeterwave InSAR data for individual tree recognition. In: *Proceedings of 33. Wissenschaftlich-Technische Jahrestagung der DGPF*: 427–436.
- Schmitt M, Magnard C, Brehm T, Stilla U (2011) Towards airborne single pass decimeter resolution SAR interferometry over urban areas. In: Stilla U, Rottensteiner F, Mayer H, Jutzi B, Butenuth M (eds) *Photogrammetric Image Analysis. Vol. 6952 of Lecture Notes in Computer Science*: 197–208.
- Schmitt M, Magnard C, Stanko S, Ackermann C, Stilla U (2013b) Advanced high resolution SAR interferometry of urban areas with airborne millimetrewave radar. *Photogrammetrie Fernerkundung Geoinformation*, 2013 (6): 603–617.
- Schmitt M, Maksymiuk O, Magnard C, Stilla U (2013c) Radargrammetric registration of airborne multi-aspect SAR data of urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 86: 11–20.
- Schmitt M, Schönberger J, Stilla U (2014a) Adaptive covariance matrix estimation for multi-baseline InSAR data stacks. *IEEE Transactions on Geoscience and Remote Sensing*, 52 (11): 6807–6817.

- Schmitt M, Schönberger JL, Stilla U (2014b) Benefit of using multiple baselines and multiple aspects for SAR interferometry of urban areas. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, in press.
- Schmitt M, Stilla U (2011) Fusion of airborne multi-aspect InSAR data by simultaneous backward geocoding. In: *Proceedings of Joint Urban Remote Sensing Event*: 53–56.
- Schmitt M, Stilla U (2013) Compressive sensing based layover separation in airborne single pass multi-baseline InSAR data. *IEEE Geoscience and Remote Sensing Letters*, 10 (2): 313–317.
- Schmitt M, Stilla U (2014a) A maximum-likelihood based approach for single-pass tomography over urban areas. *IET Radar, Sonar & Navigation*, in press.
- Schmitt M, Stilla U (2014b) Adaptive multilooking of airborne single-pass multi-baseline InSAR stacks. *IEEE Transactions on Geoscience and Remote Sensing*, 52 (1): 305–312.
- Schmitt M, Stilla U (2014c) Maximum-likelihood estimation for multi-aspect multi-baseline SAR interferometry of urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87: 68–77.
- Schreier G, ed (1993) *SAR Geocoding: Data and Systems*. Berlin, Germany: Wichmann.
- Schunert A, Soergel U (2012) Grouping of persistent scatterers in high-resolution SAR data of urban scenes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 73: 80–88.
- Schwäbisch M (1998) A fast and efficient technique for SAR interferogram geocoding. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 1100–1102.
- Schwarz G (1978) Estimating the dimension of a model. *Annals of Statistics*, 6 (2): 461–464.
- Shabou A, Baselice F, Ferraioli G (2012) Urban digital elevation model reconstruction using very high resolution multichannel InSAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 50 (11): 4748–4758.
- Shamsoddini A, Trinder J (2010) Image texture preservation in speckle noise suppression. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38 (7A): 239–244.
- She Z, Gray D, Bogner R, Homer J (1999) Three-dimensional SAR imaging via multiple pass processing. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 2389–2391.
- Shlens J (2009) *A tutorial on principal component analysis*. Center for Neural Science, New York University, available online: <http://www.sn1.salk.edu/~shlens/pca.pdf>.
- Skolnik MI (1980) *Introduction to Radar Systems*. New York, NY, USA: McGraw-Hill.
- Small D, Pasquali P, Fuglistaler S (1996) A comparison of phase to height conversion methods for SAR interferometry. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 342–344.
- Soergel U, ed (2010) *Radar remote sensing of urban areas*. Dordrecht, Netherlands: Springer Science+Business Media.
- Soergel U, Schulz K, Thoennesen U (2001) Phenomenology-based segmentation of InSAR data for building detection. In: Radig B, Florczyk S (eds) *Pattern Recognition. Volume 2191 of Lecture Notes in Computer Science*
- Soergel U, Thoennesen U, Gross H, Stilla U (2000) Segmentation of interferometric SAR data for building detection. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 33 (B1): 328–335.

- Soergel U, Thoennesen U, Stilla U (2003) Iterative building reconstruction from multi-aspect InSAR data. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 34 (3W13): on CD.
- Sohn HG, Song YS, Kim GH (2005) Radargrammetry for DEM generation using minimal control points. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 1162–1164.
- Stephens M (1970) Use of the Kolmogorov-Smirnov, Cramer-Von Mises and related statistics without extensive tables. *Journal of the Royal Statistical Society. Series B (Methodological)*, 32 (1): 115–122.
- Stilla U, Soergel U, Thoennesen U (2003) Potential and limits of InSAR data for building reconstruction in built-up areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58 (1-2): 113–123.
- Tebaldini S, Rocca F (2012) Multibaseline polarimetric SAR tomography of a boreal forest at P- and L-bands. *IEEE Transactions on Geoscience and Remote Sensing*, 50 (1): 232–246.
- Thiele A, Cadario E, Schulz K, Soergel U (2010a) Analysis of gable-roofed building signature in multispect InSAR data. *IEEE Geoscience and Remote Sensing Letters*, 7 (1): 83–87.
- Thiele A, Cadario E, Schulz K, Thoennesen U, Soergel U (2007a) Building recognition from multi-aspect high-resolution InSAR data in urban areas. *IEEE Transactions on Geoscience and Remote Sensing*, 45 (11): 3583–3593.
- Thiele A, Cadario E, Schulz K, Thoennesen U, Soergel U (2007b) InSAR phase profiles at building locations. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 36 (3/W49A): 203–208.
- Thiele A, Cadario E, Schulz K, Thoennesen U, Soergel U (2008) Building reconstruction from InSAR data by detail analysis of phase profiles. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37 (B3A): 191–196.
- Thiele A, Wegner JD, Soergel U (2010b) Building reconstruction from multi-aspect InSAR data. In: Soergel U (ed) *Radar Remote Sensing of Urban Areas*. Dordrecht, Netherlands: Springer Science+Business Media.
- Tison C, Tupin F, Maitre H (2007) A fusion scheme for joint retrieval of urban height map and classification from high-resolution interferometric SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 45 (2): 496–505.
- Tough R, Blacknell D, Quegan S (1995) A statistical description of polarimetric and interferometric synthetic aperture radar data. *Proceedings: Mathematical and Physical Sciences*, 449 (1937): 567–589.
- Touzi R (2002) A review of speckle filtering in the context of estimation theory. *IEEE Transactions on Geoscience and Remote Sensing*, 40 (11): 2392–2404.
- Touzi R, Lopes A, Bruniquel J, Vachon P (1996) Unbiased estimation of the coherence from multi-look SAR data. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 662–664.
- Trouve E, Caramma M, Maitre H (1996) Fringe detection in noisy complex interferograms. *Applied Optics*, 35 (20): 3799–3806.
- Vasile G, Trouve E, Ciuc M, Buzuloiu V (2004) General adaptive-neighborhood technique for improving synthetic aperture radar interferometric coherence estimation. *Journal of the Optical Society of America A – Optics, Image Science, And Vision*, 21 (8): 1455–1464.
- Vasile G, Trouve E, Lee JS, Buzuloiu V (2006) Intensity-driven adaptive neighborhood technique for polarimetric and interferometric SAR parameters estimation. *IEEE Transactions on Geoscience and Remote Sensing*, 44 (6): 1609–1621.

- Vasile G, Trouve E, Petillot I, Bolon P, Nicolas JM, Gay M, Chanussot J, Landes T, Grussenmeyer P, Buzuloiu V, Hajnsek I, Andres C, Keller M, Horn R (2008) High-resolution SAR interferometry: estimation of local frequencies in the context of alpine glaciers. *IEEE Transactions on Geoscience and Remote Sensing*, 46 (4): 1079–1090.
- Viberg M, Ottersen B, Kailath T (1991) Detection and estimation in sensor arrays using weighted subspace fitting. *IEEE Transactions on Signal Processing*, 39 (11): 2436–2449.
- Wang Y, Zhu XX, Bamler R (2012) Retrieval of phase history parameters from distributed scatterers in urban areas using very high resolution SAR data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 73: 89–99.
- Wax M (1991) Detection and localization of multiple sources via the stochastic signals model. *IEEE Transactions on Signal Processing*, 39 (11): 2450–2456.
- Wax M, Kailath T (1985) Detection of signals by information theoretic criteria. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 33 (2): 387–392.
- Wiley CA (1954) Pulsed doppler radar methods and apparatus. *U.S. Patent*, No. 3,196,436.
- Wilkinson A, Lord R, Inggis M (1998) Stepped-frequency processing by reconstruction of target reflectivity spectrum. In: *Proceedings of 1998 South African Symposium on Communications and Signal Processing*: 101–104.
- Wu J, Lin DC (2000) Radargrammetric parameter evaluation of an airborne SAR image. *Photogrammetric Engineering & Remote Sensing*, 66 (1): 41–47.
- Wu N, Feng DZ, Li J (2006) A locally adaptive filter of interferometric phase images. *IEEE Geoscience and Remote Sensing Letters*, 3 (1): 73–77.
- Xu W, Cumming I (1999) A region-growing algorithm for InSAR phase unwrapping. *IEEE Transactions on Geoscience and Remote Sensing*, 37 (1): 124–134.
- Yang X, Clausi D (2009) Structure-preserving speckle reduction of SAR images using nonlocal means filters. In: *Proceedings of 16th IEEE International Conference on Image Processing*: 2985–2988.
- Yue XJ, Huang GM, Zhang Y, Zhao Z, Pang L (2008) Multi-photo combined adjustment with airborne SAR images based on F. Leberl orthorectification model. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 37 (B1): 357–360.
- Zandona Schneider R, Fernandes D (2002) Entropy concept for change detection in multitemporal SAR images. In: *Proceedings of 4th European Conference on Synthetic Aperture Radar*: 221–224.
- Zebker HA, Lu Y (1998) Phase unwrapping algorithms for radar interferometry: residue-cut, least-squares, and synthesis algorithms. *Journal of the Optical Society of America A – Optics, Image Science, And Vision*, 15 (3): 586–598.
- Zhang Q, Huang Y, Schwäbisch M, Mercer B, Wei M (2012) Forest height estimation using single-pass dual-baseline L-band PolInSAR data. In: *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*: 7055–7058.
- Zhao L, Krishnaiah P, Bai Z (1986) On detection of the number of signals in presence of white noise. *Journal of Multivariate Analysis*, 20 (1): 1–25.
- Zhu XX, Bamler R (2010a) Tomographic SAR inversion by L1-norm regularization - The compressive sensing approach. *IEEE Transactions on Geoscience and Remote Sensing*, 48 (10): 3839–3846.
- Zhu XX, Bamler R (2010b) Very high resolution spaceborne SAR tomography in urban environment. *IEEE Transactions on Geoscience and Remote Sensing*, 48 (12): 4296–4308.

- Zhu XX, Bamler R (2011) A fundamental bound for super-resolution - with application to 3D SAR imaging. In: *Proceedings of Joint Urban Remote Sensing Event*: 181–184.
- Zhu XX, Bamler R (2012) Demonstration of super-resolution for tomographic SAR imaging in urban environment. *IEEE Transactions on Geoscience and Remote Sensing*, 50 (8): 3150–3157.
- Zhu XX, Shahzad M (2014) Facade reconstruction using multi-view spaceborne TomoSAR point clouds. *IEEE Transactions on Geoscience and Remote Sensing*, 52 (6): 3541–3552.
- Zhu XX, Wang Y, Bamler R (2012) Integration of tomographic SAR inversion and PSI for operational use. In: *Proceedings of 9th European Conference on Synthetic Aperture Radar*: 151–154.
- Zoubir AM, Koivunen V, Chakhchoukh Y, Muma M (2012) Robust estimation in signal processing. A tutorial-style treatment of fundamental concepts. *IEEE Signal Processing Magazine*, 29 (4): 61–80.

Acknowledgment

First of all I would like to thank Prof. Uwe Stilla, who provided me with the scientific freedom to choose and work on the topic of this thesis. This gave me the opportunity to experience significant personal and professional growth. He also taught me many important soft-skills such as how to write scientific reports or how to optimally structure presentations, for which I am forever grateful.

Second, I want to express my gratitude to Prof. Joachim Ender of University of Siegen for being the second reviewer of my thesis. Due to his double function as head of the Fraunhofer Institute for High Frequency Physics and Radar Techniques FHR, which built the MEMPHIS test sensor, eventually he also provided the test data, without which this thesis would not have been possible.

Also many thanks go to Prof. Uwe Sörgel from Technische Universität Darmstadt for being the third reviewer. The discussions with him towards finalization of the thesis were very fruitful and led to further improvements of both the manuscript and the defense presentation.

In addition, I want to thank Prof. Richard Bamler not only for agreeing to be the chairman of the examination committee, but also for several discussions that helped me to gain a deeper understanding of the statistics behind SAR observations.

Apart from the members of the examination committee, I need to express my sincere thanks to a number of additional persons I have worked with during my PhD studies: Without Thorsten Brehm (Fraunhofer FHR) and Christophe Magnard (University of Zurich), I simply wouldn't have had any single-pass multi-baseline test data. Especially Christophe showed incredible patience everytime I asked him to provide us focused SAR imagery of the Munich test scene. He also helped me very much during the beginning of my research work in the SAR field by introducing me to the data format and explaining important InSAR basics to me.

Also, I want to thank all my current and former colleagues at both the Department of Photogrammetry and Remote Sensing and the Chair of Remote Sensing Technology here at TUM for creating a pleasant working atmosphere. Although this applies to all colleagues equally, a special mention needs to go to Florian Burkert, who has become a good friend during the last 10 years we shared at TUM, and to Oliver Maksymiuk, who is an unbelievable source of knowledge when it comes to SAR signal processing, image processing and programming. I have benefitted greatly from discussions with him.

Last, but not least, I want to thank my whole family for their ever-present support. In direct connection with this thesis, that especially holds for my brother, Thomas Schmitt, who helped me proof-reading the text, and my long-term life companion Sandra König, who helped me designing some of the sketches I used for illustration.