Ingenieurfakultät Bau Geo Umwelt
Technische Universität München

Nonlocal Filtering for Single- and Multibaseline SAR Interferometry

Gerald Norbert Baier

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

Doktor-Ingenieurs (Dr.-Ing.)

genehmigten Dissertation.

Vorsitzender:
Prof. Dr.-Ing. habil. Florian Seitz

Prüfende der Dissertation:
1. Prof. Dr.-Ing. habil. Xiaoxiang Zhu
2. Prof. Dr.-Ing. habil. Richard H. G. Bamler
3. Prof. Dr.-Ing. habil. Andreas Reigber

Die Dissertation wurde am 16.05.2018 bei der Technischen Universität München eingereicht und durch die Ingenieurfakultät Bau Geo Umwelt am 11.07.2018 angenommen.

Abstract

With the launch of several new Synthetic Aperture Radar (SAR) remote sensing satellites, the recent years have seen unprecedented growth of spaceborne SAR data sets in every aspect. From available resolution, where naturally X-Band missions like TerraSAR-X and COSMO-SkyMed set the benchmark with up to 1 m resolution, repeated global mapping capabilities, which the Sentinel-1 satellites do every six days, to polarization, with Japan's ALOS-2 satellite supplying fully polarimetric data on a global scale. The future L-Band missions of the National Aeronautics and Space Administration (NASA), the Deutsches Zentrum für Luft- und Raumfahrt (engl. German aerospace center) (DLR) and the Japan Aerospace Exploration Agency (JAXA) will provide an even greater wealth of global, fully polarimetric, high-resolution data.

Notwithstanding the steadily increasing quality of available data, in virtually every SAR processing chain a denoising filter is still a mandatory processing step. The digital elevation model (DEM) generated from TerraSAR-X and TanDEM-X interferograms relies on moving average filtering and techniques extracting deformation measurements from Interferometric SAR (InSAR) time series employ adaptive multilooking to increase the estimation accuracy. Precisely for these two applications: single- and multibaseline InSAR, advanced filtering techniques have been developed within this thesis. The proposed filter algorithms follow the nonlocal filtering paradigm that has set the standard for denoising methods during the last decade.

The developed *single baseline nonlocal filter* adaptively selects key filtering parameters, making it applicable to a wide range of terrain. It also includes mitigation strategies for the rare-patch effect, which is commonly seen when applying the nonlocal filtering regime without a second thought. Extensive simulations back up the design decisions by analyzing the resolution preservation, noise reduction and bias-variance trade-off. The proposed filter compares favorably with other state of the art or legacy filters.

As a practical application several *high-resolution DEMs* from TerraSAR/TanDEM-X interferograms are generated by plugging the developed filter into DLR's DEM processing toolchain for the globally available, commercial DEM. The performance and the filtering characteristics are evaluated over a multitude of terrain types and by a comparison to light detection and ranging (LiDAR). The proposed filter provides roughly an improvement by a factor of two in terms of both resolution and height accuracy, over the default boxcar averaging filter.

This thesis also presents experiments on *multi-baseline nonlocal filtering*, which exploits the temporal domain in SAR image stacks. Appropriate similarity measures to test the equality of time series are analyzed and a set of experiments evaluates different nonlocal filtering strategies within the newly proposed framework for multi-baseline nonlocal filtering.

At last, as a result of the steadily increasing data volume of SAR missions, one contribution of this dissertation is an investigation about the actual *implementation of nonlocal filters on*

modern compute architectures. This includes an analysis of the computational costs of the various operations of nonlocal filters, their memory access patterns and how these map to contemporary compute architectures.

Zusammenfassung

Die vergangenen Jahre haben durch den Start neuer SAR-Erdbeobachtungssatelliten ein, in jeglicher Hinsicht, noch nie dagewesenes Wachstum an SAR-Datensätzen erlebt. Von der verfügbaren Auflösung, wo naturgemäß X-Band-Systeme wie TerraSAR-X und COSMO-SkyMed mit einer Auflösung von bis zu 1 m die Referenz sind, zu der Fähigkeit die Erde regelmäßig und vollständig zu beobachten, was die Sentinel-1 Satelliten innerhalb von sechs Tagen schaffen, bis hin zu der Polarisation, wo mit dem japanische Satelliten ALOS-2 vollpolarimetrische Aufnahmen der gesamten Erde geliefert werden können. Die geplanten L-Band-Missionen der NASA, des DLR und der JAXA werden einen sogar noch größeren Reichtum an global verfügbaren, vollpolarimetrischen, hoch aufgelösten Daten liefern.

Trotz der immer wachsenden Qualität der verfügbaren Daten, ist nach wie vor ein Rauschunterdrückungsfilter ein existenzieller Bestandteil jedweder SAR-Prozessierungskette. Das mittels TerraSAR-X und TanDEM-X Interferogrammen generierte digitale Höhenmodell (DEM) benutzt einen Mittelwertfilter und Techniken um Deformationsmessungen mittels SAR-Interferometrie (InSAR) zu erhalten verwenden adaptives Multilooking um die Messgenauigkeit zu erhöhen. Genau wegen dieser zwei Anwendungen: InSAR mit einer und mehreren Basislinien, wurden fortschrittliche Filterungstechniken im Rahmen dieser Arbeit entwickelt. Die vorgeschlagenen Filterungsalgorithmen folgen dem nichtlokalen Filterungsansatz, welcher während des letzten Jahrzehnts den Standard für Entrauschungsmethoden setzte.

Das entwickelte *nichtlokale Filter für einzelne Interferogramme* wählt Filterparamter adaptiv, was es für verschiedenste Arten von Terrain einsetzbar macht. Es beinhaltet des Weiteren Strategien um den Rare-Patch-Effekt zu vermeiden, welcher oft bei einem unüberlegten Einsatz von nichtlokalen Filtern eintritt. Umfassende Simulationen belegen die Effektivität der verschiedenen Maßnahmen, indem die resultierende Auflösung, die Rauschunterdrückung und der Ausgleich zwischen Verzerrung und Varianz analysiert werden. Das entwickelte Filter ist vergleichbar oder übertrifft den aktuellen Stand der Technik und herkömmliche Filter.

Als eine praktische Anwendung wurden mehrere *hochauflösende DEMs* von TerraSAR/TanDEM-X Interferogrammen generiert, indem das entwickelte Filter in die Prozessierungskette für das globale, kommerzielle DEM des DLR eingebunden wurde. Die Performanz und Filtercharakteristiken wurden für eine Vielzahl von Terrain evaluiert und mittels eines Vergleiches zu LiDAR. Das entwickelte Filter liefert ungefähr eine Verbesserung um den Faktor zwei im Vergleich zu des standardmäßig verwendeten Boxcar-Mittelwertfilters, sowohl bezüglich der Auflösung, als auch der Höhengenauigkeit.

Diese Dissertation beinhaltet auch Experimente für *nichtlokale Filter für mehrere Basislinien*, welche die zeitliche Achse in Stapeln von SAR-Bildern ausnutzen. Geeignete Ähnlichkeitsmaße für die Gleichheitsmessung von Zeitreihen werden analysiert und mehrere Experimente bewerten unterschiedliche nichtlokale Filterstrategien innerhalb des neu en-

twickelten Rahmengerüstes für nichtlokale Filter für mehrere Basislinien.

Als Antwort auf das stetig steigende Datenvolumen von SAR-Missionen ist ein weiterer Beitrag dieser Dissertation die *Implementierung nichtlokaler Filter auf modernen Rechnerarchitekturen* zu untersuchen. Dies beinhaltet eine Analyse der Rechenkosten der einzelnen Routinen eines nichtlokalen Filters, deren Speicherzugriffsmustern und wie diese auf moderne Rechnerarchitekturen übertragen werden können.

Table of contents

Acronyms and initialisms

AWGN additive white Gaussian noise 27, 42, 43

BM3D block-matching and 3D filtering 34, 41, 45, 47, 85, 114

CoSSC Coregistered Single look Slant range Complex 43, 44, 63, 66–68, 73, 74, 78, 105, 111

CPU central processing unit 82, 106, 108, 111

DEM digital elevation model

DLR Deutsches Zentrum für Luft- und Raumfahrt (engl. German aerospace center) 3, 5, 43, 67, 68

DTM digital terrain model 43, 73, 75, 82, 113, 137

ENL equivalent number of looks 87, 101, 139

EVM error vector magnitude 86, 100, 101, 139

FMA fused multiply-add 108, 110, 112, 139

GLRT generalized likelihood-ratio test 9, 40–42, 87, 88, 90–92, 94, 98, 138

GPU graphics processing unit 82, 106

HDEM high-resolution DEM 68, 69, 82

HEM height error map 78–80, 82, 113, 114

HRE high resolution elevation 114

HRWS high resolution wide swath 115

ICC Intel C/C++ Compiler 110, 111

InSAR Interferometric SAR

ITP integrated processor 68, 70, 71, 78

JAXA Japan Aerospace Exploration Agency 3, 5

LiDAR light detection and ranging

MLE maximum likelihood estimate 88, 90

MSE mean squared error 41

NASA National Aeronautics and Space Administration 3, 5

NSS nonlocal self-similarity 27, 33, 34

PDF probability density function 24–26, 38, 78, 88, 92, 135

PPB probabilistic patch-based 39, 40, 98

PSI persistent scatterer interferometry 41, 42, 86

PSNR peak signal-to-noise ratio 41

RMSLE root mean squared logarithmic error 86, 100, 138, 139

SAR Synthetic Aperture Radar

SBAS small baseline subset 42, 98

SCM sample covariance matrix 95

SIMD single instruction, multiple data 107

SLAR side-looking airborne radar 19, 20, 135

SLC single look complex 21, 39, 48, 57, 100, 102, 104, 138

SRTM shuttle radar topography mission 67

SSIM structure similarity 41

SURE Stein's unbiased risk estimate 33

TGV total generalized variation 80, 81, 137

TV total variation 35, 38, 42, 80

Notation and symbols

a amplitude

arg argument of a complex number

ρ coherence

γ complex coherence $\rho e^{j\varphi}$

$\overline{u}$ complex conjugate

H conjugate transpose

$*$ convolution

$\cdot_{\mathbf{x}}$ $\cdot$ at coordinates defined by $\mathbf{x}$

Δ distance function

$\mathrm{E}\{\cdot\}$ expected value of a random variable

$\circ$ Hadamard product

h_{amb} height of ambiguity

h smoothing parameter of the weighting kernel

z interferogram

φ interferometric phase

$\boldsymbol{U}$ matrices or patch

L number or effective number of looks

ϕ phase

u scalar

$\mathcal{U}$ tensor of order 3 or higher

$\mathrm{Var}\{\cdot\}$ variance of a random variable

$\mathbf{u}$ vectors

1 Introduction

1.1 Motivation

SAR is a remote sensing technology that produces up to decimeter resolution images of the earth from airborne and spaceborne platforms by transmitting electromagnetic waves and measuring the received echo. The defining features of SAR are its all-weather capability, that the observable scattering mechanism is related to physical properties, and the measurement of height and deformations through the interferometric phase between acquisitions.

The recent years have seen unprecedented growth of spaceborne SAR data sets in every aspect. From available resolution, where naturally X-Band missions like TerraSAR-X [WB10] and COSMO-SkyMed [Cal+14] set the benchmark with up to 1 m resolution, repeated global mapping capabilities, which the Sentinel-1 [FA12] satellites do every six days, to polarization, with Japan's ALOS-2 satellite supplying fully polarimetric data on a global scale. The future L-Band missions of NASA [Ros+17], JAXA [Mot+17] and DLR [Mor+15] will provide an even greater wealth of global, fully polarimetric, high-resolution data.

New technologies push these boundaries even further. Digital beamforming, an existential factor for NASA's, JAXA's and DLR's future L-Band missions, eliminates the need to trade resolution for swath width and vice versa. Optical inter-satellite communication links have already proven themselves invaluable to relay the humongous data volume of ESA's copernicus programme and will play an even more vital role in the just mentioned missions' global, fully polarimetric, high-resolution data.

By sophisticated processing this wealth of available data is transformed into information and intelligence. The TanDEM-X mission [Kri+07] produces a globally available DEM with so far unseen resolution and accuracy using bistatic InSAR acquisitions. Country-wide subsidence maps can be generated with stacks of multiple Sentinel-1 images [Goe+16]. Both products rely on the interferometric phase, which is unfortunately corrupted by the inherent coherence loss between SAR acquisitions, to measure the height or the deformation. Techniques to restore some of the phase's lost information have been a topic of research since the beginning of SAR interferometry. Recently the nonlocal filtering paradigm [BCM05b], which shows extremely good results for regular optical images at high computational costs, has been adapted to InSAR [DDT11]. Such advanced filters can significantly increase the amount of information that can be extracted from raw data. As one particular example [Zhu+14] generated a DEM with improved accuracy and resolution from a TerraSAR/TanDEM-X interferogram.

1.2 Objective

The objective of this thesis is to develop widely applicable nonlocal filtering algorithms by extending other previous work in three ways. Geographically, meaning having to deal with a wide array of different terrain, from densely-packed urban, over rural to agricultural and mountainous areas. Temporally, by including the temporal dimension of SAR image stacks. And finally computationally to deal with the considerable data volume of modern and future SAR missions. This thesis addresses four separate areas of research to reach this goal.

Single baseline nonlocal InSAR filtering

Development of a nonlocal filter with an adaptive selection of key filtering routines and parameters is necessary for processing the very diverse scenery of global missions. As far as possible simulations need to confirm the wide applicability of the developed method.

High-resolution DEM generation by nonlocal filtering

The developed single baseline filter is to be evaluated on the TanDEM-X dataset, which provides a wealth of different types of terrain. Such an exhaustive evaluation needs to demonstrate how much the quality of the globally available TanDEM-X DEM can be improved, and whether these benefits only manifest for some types of terrain.

Multibaseline nonlocal InSAR filtering

Current and future SAR missions map the earth multiple times per year, creating time series for every point on earth. One goal it so develop denoising algorithms which can take advantage of this redundancy, but also preserve and highlight changes, which is vital for interpreting and higher-level decision-making.

High performant implementation of nonlocal filters

Wide adaption of nonlocal filters was so far held back by their computational costs compared to legacy filters. New computer architectures offer specialized instructions and circuits to speed up common image processing tasks. How nonlocal filters can make use of these features needs to be investigated to offset to some extent the substantial increase in computational costs.

1.3 Organization of this thesis

Chapter 2 briefly outlines key SAR principles and statistics necessary for the rest of this thesis. An introduction to the nonlocal filtering paradigm and how it can be extended is given by Chapter 3. The state of the art of InSAR denoising is presented in Chapter 4. Chapter 5 to Chapter 8 each cover one of the goals described in the previous section. At last Chapter 9 summarizes the results, draws final conclusions, and details promising future research directions.

As this thesis deals with higher-dimensional data some sentences about the employed notation are in order. This thesis adopts the tensor notation used in [KB09a]. Scalars or single pixels are denoted by lowercase letters such as u. Bold lowercase letters, e.g., $\mathbf{u}$, designate vectors, their use will be obvious from the context. Matrices are denoted by boldface, capital letters like $\boldsymbol{U}$. This also applies to patches of a single image. At last, boldface, Euler script letters like $\mathcal{U}$ designate higher-dimensional data.

2 Interferometric synthetic aperture radar

This chapter serves as a short introduction into the basic operating principle of SAR and InSAR and their imaging modes, acquisition geometries and statistical properties. This includes mechanisms that degrade the received signal, motivating the use of denoising filters to at least partially retrieve the lost information. These topics are covered in far greater depth in [BH98; Ros+00; Mor00; CW05; Mor+13], which the interested reader is advised to consult for a wider grasp of them.

2.1 SAR imaging mode

A conventional side-looking airborne radar (SLAR) measures the backscattering coefficients of targets, which it localizes in range, the direction its antenna points to, and in azimuth, its flight direction. Figure 2.1 illustrates the acquisition geometry.

In range targets are distinguished by measuring the runtime of their reflected echos, and the resolution

$$\delta_{\mathrm{r}} = \frac{c_0}{2B} \tag{2.1}$$

is governed by the bandwidth B of the transmitted signal and the speed of light c_0 [Mor+13].

For azimuth the illuminated area, dictated by the antenna's beam width and the altitude of the airplane, defines the resolution. An antenna's beam width is inversely proportional to its size d_{A} relative to the wavelength λ. Since the antenna needs to be mounted on a platform, this imposes physical constraints on how narrow the beam can be. With r_0 being the distance of the sensing platform to the illuminated ground, the azimuth resolution is given by [Mor+13]

$$\delta_{\mathrm{az,\,SLAR}} = \frac{\lambda}{d_{\mathrm{A}}} \cdot r_0, \tag{2.2}$$

making it immediately apparent that the maximum altitude of such a system is limited.

Synthetic Aperture Radar, the acquisition geometry of which is identical to SLAR, is able to circumvent this drawback by

- increasing the antenna's beam width in flight direction,
- and sending multiple pulses whose echos are coherently recorded.

By doing so a SAR system builds up a synthetic aperture in azimuth direction of length $d_{\mathrm{SA}} = \frac{\lambda}{d_{\mathrm{A}} r_0}$, i.e. exactly the illuminated ground area from Equation (2.2). The resulting azimuth resolution can be obtained by plugging d_{SA} in Equation (2.2) in lieu of the physical antenna size d_{A}, with an additional factor of 2 caused by the phase gradient introduced by the

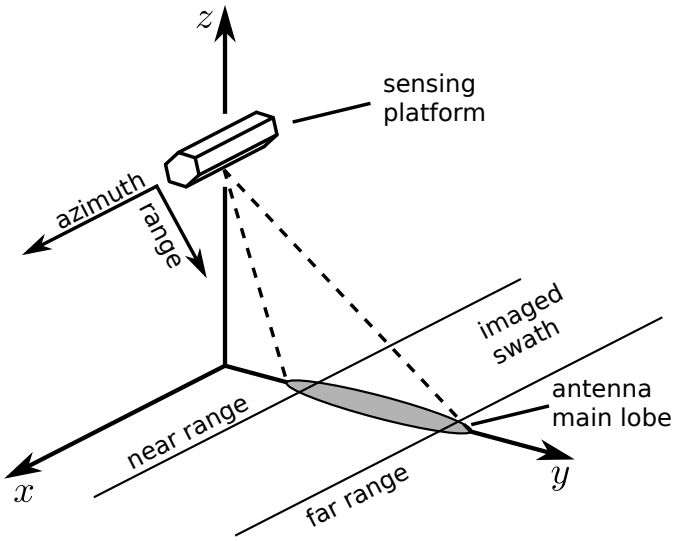


Figure 2.1: SLAR and SAR acquisition geometry. SAR systems have a wider antenna main lobe in azimuth than SLAR systems, observing targets multiple times with subsequently transmitted pulses.

consecutive transmit and receive positions. This leads to an azimuth resolution of [Mor+13]

$$\delta_{\text{az, SAR}} = \frac{d_{\text{A}}}{2} \tag{2.3}$$

that is independent of the sensor's altitude. Consequently not only airborne but also spaceborne SAR sensors are possible.

2.2 SAR and InSAR signal models

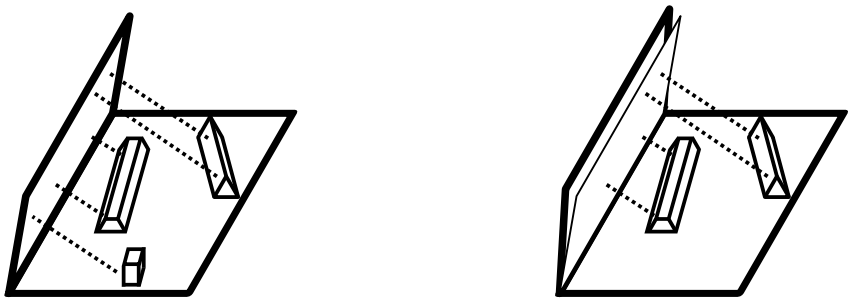
Pixels or resolution cells in a SAR image have properties completely unlike from optical images due to the coherent acquisition mode, which can further be exploited by forming an interferogram of two SAR images. This section elucidates both: the statistics of a SAR pixel and the benefits of SAR interferograms. In the following pixels and images are both referred to interchangeably by small, non-bold characters such as u.

2.2.1 SAR statistics

Within one resolution cell of a SAR image, defined by the resolution in azimuth and range, usually multiple scatterers contribute to the returned signal

$$u = \sum_k a_k^e \mathrm{e}^{\mathrm{j}\phi_k^e}, \tag{2.4}$$

where a_k^e denotes the amplitude of the elemental scatterers and ϕ_k^e their phase. This behaviour is illustrated in Figure 2.2a.



(a) **(b)**

Figure 2.2: Several elemental scatterers within one resolution cell contribute to the returned signal (a) and the total radar return is a random process. Changes of the elemental scatterers or a change of observation angle (depicted in (b) result in a correlation loss with respect to the original signal.

A common assumption is that all elemental scatterers inside a resolution cell have approximately the same backscatter, that is return a signal of the same amplitude, and that their number tends to infinity. Additionally the surface's roughness, which is large compared to the wavelength, results in the various elemental scatterers to interfere randomly with each other. Under these conditions u is a complex Gaussian random variable with a Rayleigh distributed amplitude and a uniformly distributed phase [DR58]. This is generally referred to as fully developed *speckle*, a name given to the visible signal fluctuations in homogeneous areas between resolution cells due to the random nature of the amplitude.

The constraint of equivalent backscatter of all elemental scatterers limits the general applicability of the fully developed speckle model. Plenty other signal models with either theoretical or practical justification exist. The well-known Rice distribution [Ric48] is a generalization of the Rayleigh distribution, when there is a single, dominant, deterministic scatterer present in addition to the many equally strong elemental scatterers. In a similar vein the Nakagami distribution [Nak60] can model the sum of independent and identically distributed Rayleigh random variables.

Other distributions like the Weibull- [SM90], G- [Fre+97], and Fisher- [Tis+04] distributions, due to their heavy-tails, are better suited to model heterogeneous urban areas with many strong reflections. The K-distribution [Jao84; JPW93] can model compositions of different scatterer classes.

2.2.2 SAR interferogram formation

Although the phase of a single SAR image carries no information the interferometric phase φ of two single look complex (SLC) images u_1 and u_2

$$z = u_1 \overline{u}_2 = a_1 a_2 e^{\mathrm{j}(\phi_1 - \phi)} = |z| e^{\mathrm{j}\varphi}, \tag{2.5}$$

where $\overline{u}$ designates complex conjugate, consists of several components

$$\varphi = \varphi_\mathrm{h} + \varphi_\mathrm{d} + \varphi_\mathrm{a} + \varphi_\mathrm{n}, \tag{2.6}$$

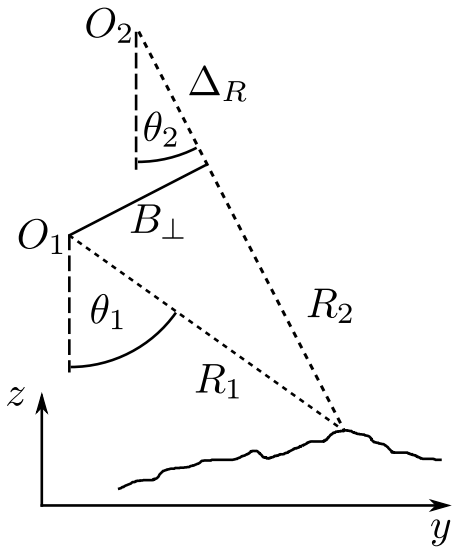


Figure 2.3: In across-track SAR interferometry the same target is observed from different satellite orbits. The change in path length Δ_R is proportional to the interferometric phase and conveys information about the terrain's height or changes between the two acquisitions.

that influence the images' propagation path lengths and can thus be used for measurements. φ_a encompasses ionospheric distortion due to varying total electron counts of the two acquisitions [Mey+06] and also tropospheric differences [Bek+15]. The height component and the deformation, in line of sight direction, component are denoted by φ_h and φ_d. The noise due to correlation loss is described by φ_n.

Several factors contribute to the decorrelation between two SAR acquisitions [ZV92]. The intrinsic thermal noise in radio receivers lowers the correlation. Figure 2.2b illustrates two more decorrelation mechanisms. Due to temporal changes the composition of elemental scatterers may change, resulting in a change of the resolution cell's total reflected signal, which is no longer fully correlated with the originally reflected signal from Figure 2.2a. Another contributing factor to the decorrelation are orbital changes between two acquisitions, with the effect that slightly different Doppler spectra are observed as a result of the different spatial positions of the satellites. This factor also includes volume decorrelation [Alb04], which is especially relevant for vegetated areas such as forests. The total correlation is then obtained as the product of the individual coherence factors [ZV92]

$$\rho_\mathrm{total} = \rho_\mathrm{thermal} \cdot \rho_\mathrm{temporal} \cdot \rho_\mathrm{spatial}. \tag{2.7}$$

2.2.3 InSAR height measurements

With the phase of a SAR interferogram being a function of the observed terrain's elevation, in reverse the height can be determined from the phase. The acquisition geometry for such a measurement is depicted in Figure 2.3, where the two different acquisition orbits are denoted by O_1 and O_2.

For a bistatic acquisition the interferometric phase is given by

$$\varphi = \frac{2\pi}{\lambda}(R_2 - R_1) = \frac{2\pi}{\lambda}\Delta_R, \tag{2.8}$$

with λ denoting the electromagnetic wave's wavelength.

Under the far field approximation $R_1 = R_2 = R$ and $\theta_1 = \theta_2 = \theta$ and the plane wave approximation, the phase's height sensitivity [BH98]

$$\frac{\partial\varphi}{\partial z} = \frac{2\pi}{\lambda}\frac{B_\perp}{R\sin\theta} \tag{2.9}$$

gives the conversion between phase and height. The orthogonal baseline $B_\perp$ defines the height sensitivity, as all other parameters are essentially fixed. Equation (2.9) can further be trivially converted to express height errors as a function of phase errors caused by coherence loss.

2.3 InSAR denoising

The inherent coherence loss between SAR acquisitions degrades the quality of the observed interferometric phase. Depending on how the phase is utilized the application of a denoising filter might be necessary.

2.3.1 Weighted mean

The most common filter is a weighted average of the interferogram

$$\hat{z}_\mathbf{x} = \sum_{\mathbf{y}\in\Omega_\mathbf{x}} w_{\mathbf{x},\mathbf{y}} z_\mathbf{y}, \tag{2.10}$$

where the subscript $\cdot_\mathbf{x}$ denotes the coordinates within the interferogram. The weights w inside the averaging window Ω are normalized to sum to one and play a crucial role in the trade-off between reducing the variance of the phase and the loss of resolution. Often indiscriminate spatial averaging, called *multilooking*, is applied with weights defined as

$$w_{\mathbf{x},\mathbf{y}} = \frac{1}{|\Omega_\mathbf{x}|} \quad \forall\mathbf{y}. \tag{2.11}$$

Here $|.|$ denotes the cardinality of a set, that is the number of pixels inside the averaging window, also referred to as the number of looks L.

The weighted mean for a single interferogram in Equation (2.10) can straightforwardly be extended to a stack of N SAR images $\mathbf{u}$

$$\hat{\mathbf{\Sigma}}_\mathbf{x} = \sum_{\mathbf{y}\in\Omega_\mathbf{x}} w_{\mathbf{x},\mathbf{y}} \mathbf{u}_\mathbf{y}\mathbf{u}_\mathbf{y}^\mathrm{H}, \tag{2.12}$$

where H denotes conjugate transpose, and provides an estimate of the covariance matrix

$$\begin{bmatrix} a_1^2 & a_1 a_2 \rho_{1,2} e^{j(\phi_1-\phi_2)} & \cdots & a_1 a_N \rho_{1,N} e^{j(\phi_1-\phi_N)} \\ a_2 a_1 \rho_{1,2} e^{j(\phi_2-\phi_1)} & a_2^2 & \cdots & a_2 a_N \rho_{2,N} e^{j(\phi_2-\phi_N)} \\ \vdots & & \ddots & \vdots \\ a_N a_1 \rho_{1,N} e^{j(\phi_N-\phi_1)} & \cdots & & a_N^2 \end{bmatrix}. \tag{2.13}$$

The complex coherence

$$\gamma_{m,n} = \rho_{m,n} e^{j(\phi_m-\phi_n)} \tag{2.14}$$

contains all interferometric information and decorrelation effects (Equation (2.7)). The individual off-diagonal elements in Equation (2.12) reduce to Equation (2.10).

The following paragraphs briefly present some of the probability density functions (PDFs) that arise when dealing with multilooked InSAR data. They come in handy when analyzing and evaluating filtered data.

2.3.2 The multilooked interferometric phase

The PDF of the interferometric phase φ_0, multilooked by L and with complex coherence γ is [Lee+94; JB94; BH98]

$$p_\varphi(\gamma, L) = \frac{\Gamma(L+1/2)}{2\sqrt{\pi}\Gamma(L)} \frac{\left(1-|\gamma|^2\right)^L \cos(\varphi-\varphi_0)}{\left(1-|\gamma|^2\cos^2(\varphi-\varphi_0)\right)^{L+1/2}} + \frac{\left(1-|\gamma|^2\right)^L}{2\pi} F\left(L, 1; \frac{1}{2}; |\gamma|^2\cos^2(\varphi-\varphi_0)\right) \tag{2.15}$$

where F is a Gauss hypergeometric function and the subtraction of φ_0 ensures that the PDF is centered at zero. Its dependence on L and γ is highlighted by Figure 2.4.

The figure's left hand side shows the effect of different coherence levels. For zero coherence the phase is uniformly distributed and contains no information. With increasing coherence the peak's width narrows and the phase's standard deviation decreases, yielding more reliable measurements.

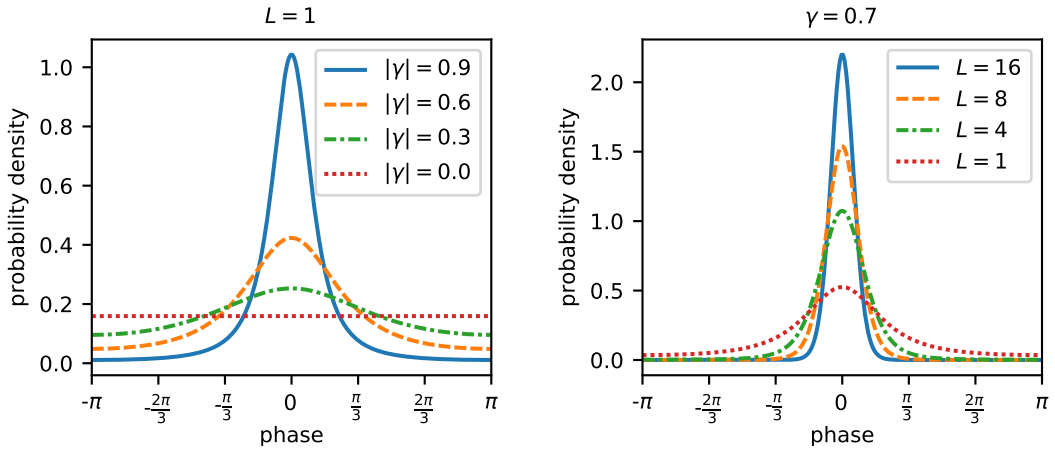
A similar behaviour can be observed for increasing the number of looks L for a fixed coherence level, as depicted on the right side of Figure 2.5. The peak gets narrower for more looks producing a more accurate phase estimate.

For further illustration of this behaviour Figure 2.5 shows the standard deviation of Equation (2.15) as a function of coherence for different number of looks L.

Using Equation (2.9) the phase standard deviation can be converted to a height standard deviation, which can serve as a figure of merit when evaluating InSAR generated digital elevation models. In allusion to later chapters, this means the expected coherence level of a SAR interferogram together with Figure 2.5 defines the accuracy of a derived DEM.

2.3.3 Wishart distribution

If an InSAR stack $\mathbf{u}$ follows a zero mean, multivariate complex normal distribution, meaning fully developed speckle is assumed, and for the case of simple multilooking, with weights defined as in Equation (2.11) the covariance matrix estimate's PDF from Equation (2.12) follows



(a) Varying coherence

(b) Varying number of looks

Figure 2.4: Probability density function of the multilooked interferometric phase (Equation (2.15)). An increase in coherence reduces the phase's standard deviation (a). Increasing the number of looks for a fixed coherence level (b) also narrows the uncertainty of the estimate.

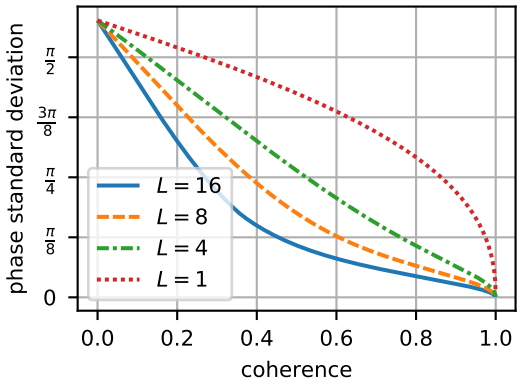


Figure 2.5: Phase standard deviation of a SAR interferogram for different numbers of looks L as a function of coherence.

a complex Wishart distribution [Goo63]. Its PDF is given by

$$p_{\hat{\mathbf{C}}}(L,\boldsymbol{\Sigma}) = \frac{1}{\Gamma_N(L)}\left(\frac{L}{|\boldsymbol{\Sigma}|}\right)^L |\hat{\mathbf{C}}|^{L-N} \exp\left\{-L\,\mathrm{tr}\left(\boldsymbol{\Sigma}^{-1}\hat{\mathbf{C}}\right)\right\} \tag{2.16}$$

with $\boldsymbol{\Sigma}$ being the covariance matrix of $\mathbf{u}$ and Γ_N the multivariate Gamma function defined by

$$\Gamma_N(L) = \pi^{N(N-1)/2} \prod_{k=1}^{N} \Gamma(L-k+1). \tag{2.17}$$

For the case $N = 1$, i.e. working with a multilooked SAR intensity image I_L, the Wishart-distribution reduces to the Gamma-distribution [Goo75]

$$p_{I_L}(L,\sigma) = \frac{1}{\Gamma(L)}\left(\frac{L}{\sigma}\right)^L I_L^{L-1} \exp\left\{-\frac{LI_L}{\sigma}\right\}, \tag{2.18}$$

whose relationship to Equation (2.16) is obvious.

3 Nonlocal filters

Since nonlocal filters were first mentioned in [BCM05b] they have essentially set the standard for image denoising. Their popularity is founded in their superb noise reduction and detail preservation. Some authors even claim that the field of denoising is essentially dead as only minor gains can be accomplished [CM10; LN11].

Fundamental to the understanding of the nonlocal filtering paradigm is nonlocal self-similarity (NSS). In a nutshell, NSS means that in a natural image the same or similar structures or textures are found multiple times. Intuitively this gives different noisy measurements of the same object. The goal of nonlocal filters is to exploit this inherent redundancy without compromising details.

Figure 3.1 illustrates this concept by showing for four selected small patches the eight most similar patches in the image. For edges or homogeneous areas many similar patches are found in the vicinity of the original target patch, which supports the denoising process. On the other hand for isolated details like the eye, finding similar patches seems to be more difficult and even some outliers would have to be considered.

After such an offhand introduction of nonlocal filters Section 3.1 puts them in the correct historical setting and provides more theoretical background plus the appropriate terminology. Some of the original nonlocal filter's shortcomings and how they can be alleviated are addressed in Section 3.2. Section 3.3 introduces more profound extensions to the nonlocal filtering paradigm, which can potentially improve its result or help in some edge cases.

For a general review of image denoising algorithms the interested reader is referred to [BCM05a; Kat+09; Leb+12; Piž17]. Their knowledge can help the understanding of algorithms that exploit NSS or they can be even integrated and combined with nonlocal algorithms.

3.1 Theoretical background

This section introduces the nonlocal filtering concept by starting from simple Gaussian blurring and moving on to bilateral filtering, which is later extended to the nonlocal filtering concept. The whole section assumes that the original image v has been corrupted by additive white Gaussian noise (AWGN) n to form the noisy input image z

$$z = v + n\,. \tag{3.1}$$

To illustrate the characteristics of the various filters Figure 3.2 shows their results for a noisy input image. The original image has pixels with values ranging from zero to one and is corrupted by AWGN with a standard deviation of 0.1.

Figure 3.1: Illustration of nonlocal self-similarity. For four selected small patches the eight most similar patches in the image are plotted. Along edges or for homogeneous areas many patches of the object are found. Details, such es Lena's eye, are more challenging.

3.1.1 The bilateral filter

As the original nonlocal filter [BCM05b] is quite closely related to the bilateral filter [TM98; Par+09] the latter is introduced first. Bilateral filtering itself can be seen as an extension to a simple convolution, which to better preserve edges incorporates a second weighting term depending on the *photometric* or *radiometric* distance between two pixels.

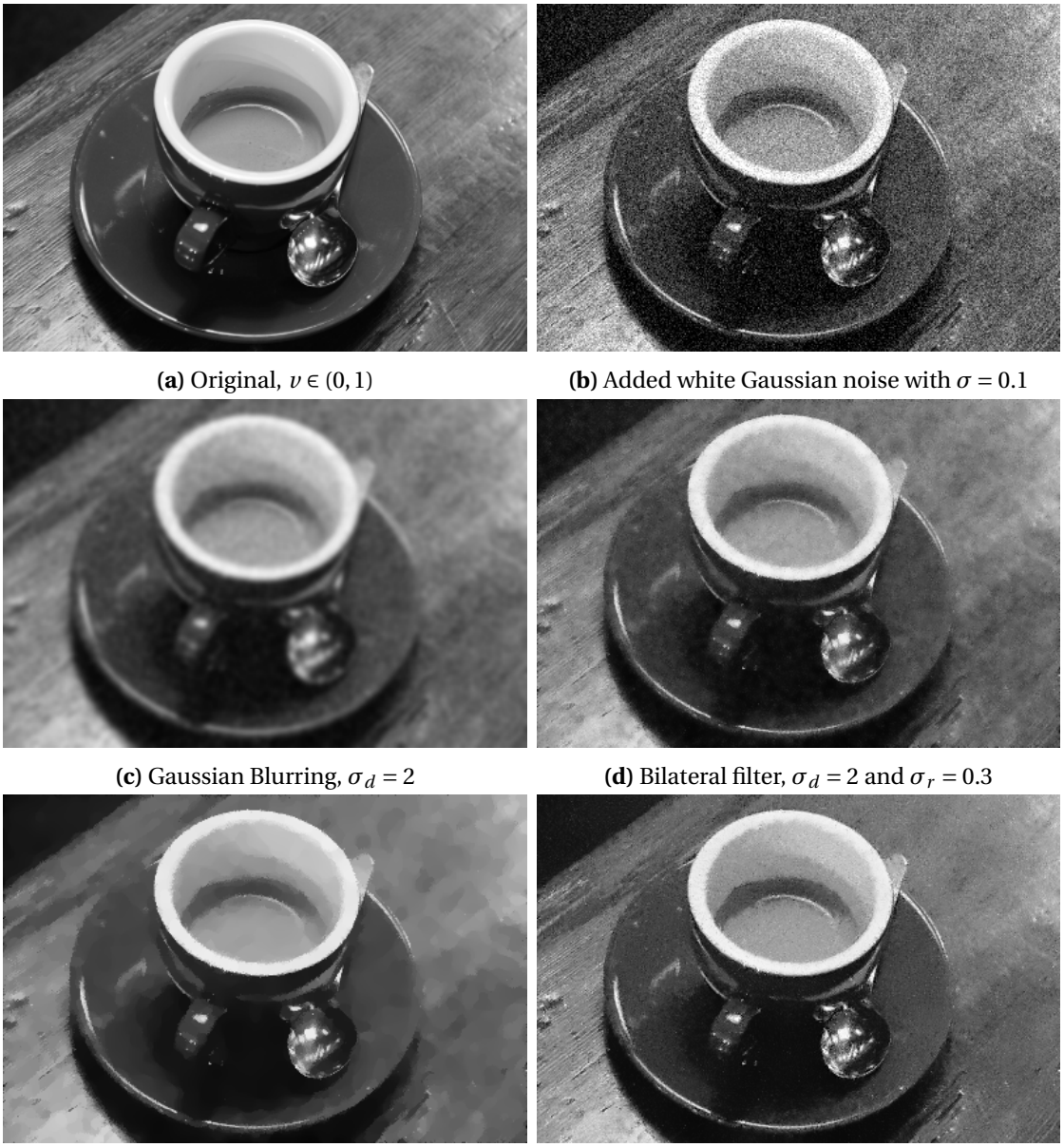
The most common form of the bilateral filter employs two Gaussian kernels. One depending on the Euclidean distance $\|\mathbf{x}-\mathbf{y}\|$ of two pixel coordinates $\mathbf{x}$ and $\mathbf{y}$, and a second measuring their photometric similarity $\|z_\mathbf{x} - z_\mathbf{y}\|$. The resulting filter obtains an estimate $\hat{v}$ of v by

$$\hat{v}_\mathbf{x} = \sum_\mathbf{y} \mathrm{e}^{-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2\sigma_d^2}} \mathrm{e}^{-\frac{\|z_\mathbf{x}-z_\mathbf{y}\|^2}{2\sigma_r^2}} z_\mathbf{y} \tag{3.2}$$

where σ_d and σ_r set how discriminative the filter is, i.e. its preference for smoothing over detail preservation. In case the photometric term would be discarded Equation (3.2) reduces to a simple Gaussian blurring operation. The spatial and photometric distances need not be Euclidean with a Gaussian kernel, although this is the most common case. Indeed [TM98] alludes to different metrics for color images for example.

The effect that the photometric distance has on the filtering result is highlighted by Figure 3.2c and Figure 3.2d, which shows for a noisy image the filtering results obtained with Gaussian blurring and the Bilateral filtering given by Equation (3.2). Clearly edges are no longer smoothed due to the large jumps in luminance.

Although the bilateral filter was initially a completely intuitive solution for preserving edges, it can actually be connected to anisotropic diffusion and weighted least squares [Ela02] for a more theoretical foundation.



(a) Original, $v \in (0,1)$

(b) Added white Gaussian noise with $\sigma = 0.1$

(c) Gaussian Blurring, $\sigma_d = 2$

(d) Bilateral filter, $\sigma_d = 2$ and $\sigma_r = 0.3$

(e) Bilateral filter with multiple iterations

(f) Nonlocal means filter, $h = 0.7$ and the patch size is 7×7

Figure 3.2: Demonstration of various denoising filters. The bilateral filters and the nonlocal filter avoid the smoothing of edges. Running the bilateral filter for several iterations removes the mottling effect but introduces piecewise constant areas. The nonlocal filter does not exhibit both artifacts. Test image taken from [Wal+14].

Figure 3.2 also shows the filtering result obtained with the original nonlocal means filter of [BCM05b], which clearly provides an improvement over the bilateral filter. The following section will explain why this is the case and also explain some of the design decisions.

3.1.2 The case for image patches

In Equation (3.2) the photometric distance is a function of two pixel values alone. As both are corrupted by noise their resulting distance will be noisy as well, leading to mottling of smooth surfaces as in Figure 3.2d. One solution is an iterative application of the bilateral filter, using the filtered output of one iteration as input for the next, the result of which is depicted in Figure 3.2e. This technique successfully decreases the noise level at the cost of a cartoonification of the image as already noted in [TM98], also commonly referred to as *staircasing*. Staircasing is not only the result of a neighborhood filter's iterative application, but can already manifest itself after one iteration. A solution was presented in [BCM06].

Nonlocal filters employ a different strategy to obtain a more robust estimate of the distance between two pixels. They compute the distance on the pixels surrounding areas, called patches. Since this includes more pixels the resulting estimate of their distance are to a lesser degree affected by noise. This approach is valid, assuming that similar patches share similar center pixels, also referred to as *patch regularity* [DAG11].

For the original nonlocal filter [BCM05b] the patch distance is computed by the Euclidean distance, a direct extension of the bilateral filter's *photometric* distance.

3.1.3 The search for similar pixels

The next modification is that nonlocal filters, true to their name, do not take into account the location of various patches when filtering. The benefit of this is visible in Figure 3.1, where for example all similar blue patches are stretched out along the edge. In this particular case limiting the spatial extent of a filter would discard useful information, degrading the quality of its output.

For computational reasons the search for similar patches is usually restricted to a *search window*. Common sizes are 21×21 up to 35×35, which still provides plenty of redundancy for a filter to make use of, especially for homogeneous areas.

3.1.4 Connection to bilateral filtering

The original nonlocal filter proposed in [BCM05b], just like the bilateral filter, uses an exponential kernel to map the patch similarities into weights, resulting in the following formula for estimating a pixel

$$\hat{z}_{\mathbf{x}} = \sum_{\mathbf{y} \in \partial_{\mathbf{x}}} \mathrm{e}^{-\frac{\|\boldsymbol{G} \circ (\boldsymbol{Z}_{\mathbf{x}} - \boldsymbol{Z}_{\mathbf{y}})\|^2}{h}} z_{\mathbf{y}}, \tag{3.3}$$

where ∂_x is a set that contains all pixel indexes in the search window, $\boldsymbol{G}$ a Gaussian kernel, $\circ$ the Hadamard product and $\boldsymbol{Z}_{\mathbf{x}}$ denotes a patch centered at $\mathbf{x}$. The size of the patch is set to a somewhat reasonable size with regard to the Gaussian kernel width. The Gaussian kernel $\boldsymbol{G}$

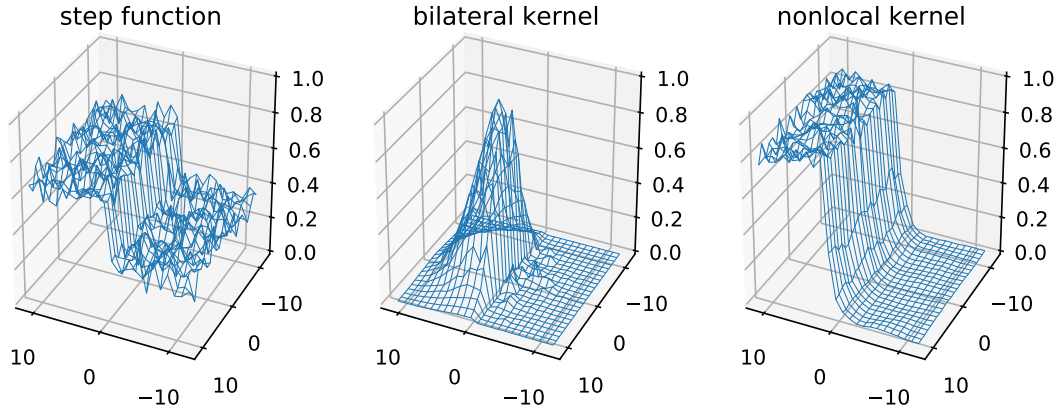


Figure 3.3: Comparison of the resulting filtering kernels of a bilateral and a nonlocal filter for a step function (inspired by [TM98]). The target pixel is located such that the nonlocal filter's patch does not include the step. The nonlocal filter offers better suppression of pixels of different values.

gives more weight to pixels at the center of the patch than at the edge, but can also be omitted without any noticeable differences [Dar+08]. The parameter h has the same impact has σ_d and σ_r of the bilateral filter. The difficulty in selecting it is elaborated in a subsequent section.

Contrasting Equation (3.3) to Equation (3.2) shows the close connection that bilateral and nonlocal filters share. Their difference are

- the lack of spatial discrimination of the nonlocal filter, and
- the use of patches instead of pixels for computing the photometric distance.

If one regards the search window as a hard-thresholded spatial distance function the relationship between nonlocal and bilateral filters is even more evident, as their only difference is the use of patches instead of pixels for computing the weights.

Figure 3.3 elucidates the difference of nonlocal and bilateral filters by comparing their respective filtering kernels for a step function. In this example the nonlocal filter uses a patch size of 5×5 and the target pixel is located two pixels to the left of the edge, meaning the patch does not contain the edge. Comparing the nonlocal to the bilateral filtering kernel two observation can be made. The nonlocal filtering kernel has greater support, yielding a more accurate estimate and also better suppresses pixels with a different photometric distance, since its similarity measure is more robust to noise.

3.2 Refinements to the original nonlocal filter

Although the nonlocal filter in its most basic form [BCM05b] already provides substantial improvements over the bilateral filter (see Figure 3.2) it has some deficiencies, which subsequent publications tried to address. The following paragraphs will list the shortcomings and refer to the original publications.

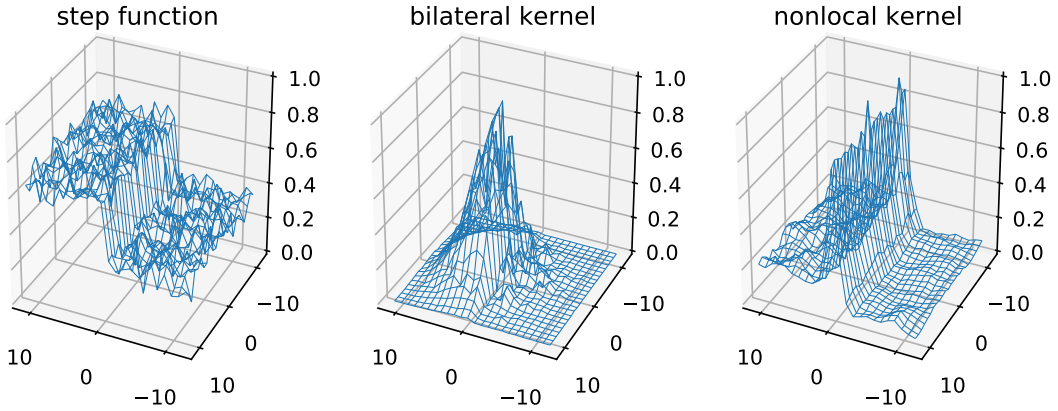


Figure 3.4: Comparison of the resulting filtering kernels of a bilateral and nonlocal filter for a step function. The target pixel is located at the edge, which compels the nonlocal filter to only find similar patches along the edge, drastically reducing its ability to reduce the variance. For the bilateral filter the kernel shares the same characteristics to the one in Figure 3.3.

3.2.1 Avoiding the rare patch effect

The rare patch effect crops up at and around distinct features. For patches that contain said feature only few similar patches are found, resulting in a halo of high noise in its vicinity.

By repeating the experiment of Figure 3.3, but this time the target pixel is located on the edge, the rare patch effect makes its appearance. Whereas the bilateral filter is unfazed, the nonlocal filtering kernel only gives large weights to pixels that also contain the edge. The result would be a filtering result with comparatively large variance.

Common solutions are varying the patch size locally [DAG11] or using shape adaptive patches [DDS12].

The rational for locally adapting the patch size is the trade-off between robustness to noise, the original argument for employing patches in the first place, and the *rare patch effect*. In homogeneous areas large patches counteract the noise whereas for heterogeneous areas large patches render the search for similar patches more difficult. The solution is to define the patch size as a function of local scene heterogeneity or the risk of the nonlocal estimator.

The argument for shape adaptive patches is similar. If patches are aligned to features, for example by moving them off-center with respect to the target patch, features will no longer hamper the search for similar patches.

Some of the more elaborated methods that Section 3.3 covers also help to mitigate the rare patch effect.

3.2.2 Search window size

Just like the patch size needs to be optimized (see Section 3.2.1) the dimension of the search window also has an effect on a filter's estimates. Somewhat contradicting the nonlocal filtering paradigm experiments in [DAG11] show that limiting the search for similar patches

from the whole image to the search window actually improves the denoising result. The case against a search of the whole image, or a large search window for that matter, is that also many dissimilar patches will be found. With the result, that even though their individual contributions in the weighted mean are limited, by their sheer number they corrupt the estimate.

There are again several solutions. [KB08] adaptively sets the search window size by iteratively growing it, with the *risk*, i.e., the bias and variance, as a stopping criterion. Another approach, advocated for in [DAG11], is the use of a thresholding weighting kernel that discards very dissimilar patches. Both methods make the nonlocal means filter more robust since they limit the number of pixels considered in the weighted mean.

3.2.3 Weighting kernel

The weighting kernel of nonlocal filters transforms the measured patch distances into weights, assigning large weights to similar patches and conversely low weights to dissimilar ones. From Equation (3.3) the weights of the original filter are computed as

$$w_{\mathbf{x},\mathbf{y}} = \mathrm{e}^{-\frac{\|Y_\mathbf{x} - Y_\mathbf{y}\|^2}{h}}. \tag{3.4}$$

The choice of h is critical. It should be large enough to provide a sufficient amount of denoising and small enough not to bias the estimate or smooth out isolated details. A typical *bias-variance trade-off* prohibiting a fixed global selection of h.

Both [VK09] and [DAG11] derived solution for the nonlocal mean to select h based on Stein's unbiased risk estimate (SURE)[Ste81]. Yet SURE is only applicable for white Gaussian noise, making it impracticable for SAR.

Another amelioration for the weighting kernel was proposed in [DAG11], which argues for an exponential kernel with thresholding. With the argument, that this makes the nonlocal filter more robust with regard to dissimilar patches, having the same effect as picking an appropriate search window size in Section 3.2.2.

How to define the center patch's weight, i.e. $w_{\mathbf{x},\mathbf{x}}$ is also of importance. A separate treatment is mandatory as its weight would otherwise dominate the weighted means. Common solutions are setting it simply to zero or to the second largest weight.

3.3 Extending the nonlocal filtering paradigm

The simple nonlocal means filter of [BCM05b] showed the power of NSS and sparked research into extending it. This section introduces techniques that further exploit NSS hereby improving the original filter's capabilities. In particular they address some of the aforementioned shortcomings in Section 3.2.

3.3.1 Transformations and sparsity

As the name nonlocal means already implies a weighted mean is computed over similar patches by the original nonlocal filter. In 2007 Dabov et al. extended this practice with their

block-matching and 3D filtering (BM3D)[Dab+07] approach, which to this day is still considered to provide state of the art denoising capabilities for regular optical images.

BM3D creates a 3D volume by grouping similar patches, called blocks in BM3D lingo, in a third dimension. The 3D volume is assumed to be sparse in a transformed domain, where coefficient thresholding and subsequent inverse transforming denoises the 3D volume. BM3D employs wavelets as a sparsifying basis, which have been in use for a long time for denoising by means of wavelet shrinkage [DJ94]. The final 3D wavelet transform is composed of a 2D patch transform and a 1D group transform.

The allure of BM3D is that by collectively denoising the 3D volume it not only exploits inter-patch redundancies, as the classic nonlocal means, but also intra-patch redundancies. The authors also confirmed that indeed the intra-patch redundancies are of importance by showing that the performance suffers if the 2D patch transform is not applied. The 2D patch transform can also to some extent circumvent the rare patch effect, since even if no similar patches are found, denoising is still performed on a single patch basis by conventional wavelet shrinkage. Maybe surprisingly the selection of the underlying wavelet basis has only marginal effect on the filtering output [Dab+07].

BM3D demonstrated the benefit of augmenting NSS with sparsifying transforms. Several works were inspired by BM3D and adapted it to other applications besides image denoising, new data types or extended it. Danielyan, Katkovnik, and Egiazarian cast BM3D as a variational problem in [DKE12] to deblur an image. A despeckling filter for SAR images, based on sparsifying transform, was proposed in [Par+12]. [Liu+15] extends BM3D by replacing the fixed transform with a data-driven approach. To denoise a group of similar patches their covariance matrix's eigenvectors are used as a sparsifying basis. By making use of such a locally adaptive transform the proposed method was able to improve upon BM3D.

3.3.2 Dictionary learning

As Section 3.3.1 already alluded to there is an inherent redundancy in a collection of similar patches which suggests further exploitation. Learned dictionaries have started replacing simple wavelet shrinkage methods. Their appeal, in contrast to a fixed transform, is that they adapt to the data at hand allowing for a sparser representation.

For example [EA06] uses the K-SVD algorithm [AEB06] for denoising patches with a learned dictionary. In a similar vein one can see the approach of [Mai+09a] as an extension to BM3D. [Mai+09a] denoises a set of similar patches by exploiting their joint sparsity using an overcomplete learned dictionary, which in this specific case was learned using the approach in [Mai+09b].

Just like BM3D the dictionary based approach helps avoiding the rare patch effect, if fewer similar patches are found just the level of joint sparsity decreases, but denoising still takes place using the sparse dictionary approximation.

3.3.3 Regularization

Other proposed methods combine nonlocal filtering with variational approaches.[GO08; PBC08; ELB08] extend the commonly used total variation (TV) approach [ROF92] by defining

the derivative not locally but on a weighted graph. This way pixels are not only regularized with their direct neighbours, but also with disconnected pixels with similar statistics, improving the performance for common image restoration tasks, such as denoising and inpainting.

An elegant solution for avoiding the rare patch effect is presented in [SDA14]. In brief the proposed approach provides a trade-of between the solutions given by the nonlocal means and TV regularization. If the solution of the nonlocal means is judged to be unreliable, indicated by a low number of looks of the weighted mean, more importance is given to the regularization term.

4 InSAR filtering: history and state of the art

As already described in Chapter 2 SAR interferograms suffer from an inherent coherence loss resulting in phase measurements corrupted by noise. Since many applications depend on accurate phase information, noise-removal algorithms have been an active topic of research for more than two decades. Often they are adapted or inspired from denoising filters for regular optical images, which have always been one or two steps ahead due to the greater research interest and community. This chapter gives an overview on the history and current state of the art of denoising filters, both for single and multi baseline InSAR. Special emphasis is put on nonlocal filtering techniques.

4.1 Single baseline InSAR filters

As simple spatial averaging, while reducing the variance of the phase estimate (see Figure 2.5), essentially decreases the resolution of the image by mixing pixels of different nature at edges or other structures, more sophisticated methods have been a topic of research since more than two decades.

4.1.1 Adaptive filtering

With regard to the weighted mean, as defined in Equation (2.10), a relatively straightforward extension of simple multilooking is the use of more discriminant averaging windows. One way is employing directional filtering kernels based on the interferometric phase's gradient [Lee+98], which is especially useful for dense fringe areas. Other approaches employ statistical tests to discard pixels that might belong to other scattering classes such as [Vas+06], which relies on intensity-driven region growing for an adaptive neighborhood selection of homogeneous pixels instead of a fixed filtering window.

4.1.2 Transform domain filtering

Transform domain filtering, which was briefly touched on in Section 3.3.1, had also found its way to the field of SAR interferometry. The widely employed Goldstein filter [GW98] essentially employs coefficient rooting or alpha-rooting [McC80] on overlapping blocks of the interferogram to suppress noise at non-dominant frequencies. It works in the frequency domain, filtering a block's Fourier transform $\mathscr{F}\{\boldsymbol{Z}\}$ by a smoothed version of its magnitude to the power of α

$$\mathscr{F}\left\{\hat{\boldsymbol{Z}}_{\mathbf{x}}\right\} = \left(\boldsymbol{K} * \left|\mathscr{F}\left\{\boldsymbol{Z}_{\mathbf{x}}\right\}\right|\right)^{\alpha} \cdot \mathscr{F}\left\{\boldsymbol{Z}_{\mathbf{x}}\right\}, \tag{4.1}$$

where K denotes a smoothing kernel and $*$ convolution. The filtering parameter α regulates how much dominant frequency bins, assumed to contain the signal, are amplified with respect to other frequencies.

Later extensions of the filter [Bar+03] improved its performance by locally selecting α based on the coherence, a measure for the interferograms noise level, or even more parameters [Suo+16].

Wavelet shrinkage [DJ94], a very popular denoising scheme for optical images, can also be adapted to InSAR [LF02; BM11]. After transformation with wavelets thresholding or shrinkage of wavelet coefficients with small magnitude removes noise but leaves the contained signal mostly untouched. Although at first glance their denoising principle seems similar to the Goldstein filter, these methods usually operate on the complete interferogram alone and not on small overlapping blocks.

4.1.3 Variational approaches

Variational based approaches have long been employed for the restoration of regular optical images. Their goal is to minimize a cost function consisting of a data fidelity D and a regularization term R modelling prior knowledge of the signal, such as piecewise smoothness

$$\arg\min_{v} \mathrm{D}(z, v) + \mathrm{R}(v)\,. \tag{4.2}$$

Especially total variation (TV) first proposed in [ROF92] has practically become synonymous with the whole category, as a result of its applicability to many problem domains [CCN11] and computational tractability [CP11]. TV penalizes perturbations of the image u by limiting its gradient

$$\mathrm{TV}(u) = \int_{\Omega} |\nabla u| \mathrm{d}\, x\,. \tag{4.3}$$

Naturally, these types of algorithms have found their way to the field of InSAR as well [Den+09; SDT11]. The interferometric phase's probability density function (Equation (2.15)) usually serves as the data fidelity term and a regularization term such as TV on the phase, possibly together with the amplitude, or the unwrapped result, leads to a preference for smooth solutions. However, finding a solution is not as computationally efficient as for denoising non-cyclic data.

Extensions to TV exist for cyclic data [SC11]. [Ber+14] considers even higher derivatives for denoising images in cyclic color spaces or SAR interferograms within a variational framework.

4.1.4 Others

This subsection lists two methods that do not really fit any of the previous sections. Neural network based approaches, that recently experienced a surge in popularity for many domains, can also be employed to filter SAR interferograms [IH17]. As an example for a less exotic method, [SW18] proposes an efficient algorithm for computing the median of cyclic data, citing InSAR denoising as one possible application. The appeal of median filter is their

ability to preserve edges, which is of special interest for SAR images with their widely varying amplitude.

4.1.5 Nonlocal filters

As already mentioned in Chapter 3, in the past years nonlocal filters have gained traction and several nonlocal InSAR filters were proposed. The previous chapter explained the motivation for using patches in lieu of pixels for computing similarities. The original nonlocal filter relied on the Euclidean distance, which is an adequate and suitable choice for additive white Gaussian noise, but since the noise characteristics of SAR and InSAR fundamentally deviate from Gaussian noise it cannot be used for these applications. The next paragraphs introduce several nonlocal InSAR filters and their similarity criteria. For an overview on various similarity measures and how to assess them [DDT12] can be consulted.

An InSAR-specific problem is the joint influence of phase and reflectivity when computing similarities. Since the quantity of interest is the phase, taking also the reflectivity into account might distort the outcome, if both do not change concurrently. Yet often phase changes are accompanied by reflectivity changes, which can support the similarity computation. In fact the filter introduced in [HST13] proposes a technique which makes use of a so called *guidance image* on which the weights are computed. In the InSAR case this can well be the reflectivity.

NL-InSAR

NL-InSAR [DDT11] is an extension to InSAR from the probabilistic patch-based (PPB) approach introduced for despeckling SAR images in [DDT09] and is the first nonlocal InSAR filter.

PPB defines the similarity of two pixels $u_{\mathbf{x}}$ and $u_{\mathbf{y}}$ as the probability that their underlying parameters $\theta_{\mathbf{x}}$ and $\theta_{\mathbf{y}}$ are identical

$$p(\theta_{\mathbf{x}} = \theta_{\mathbf{y}} | u_{\mathbf{x}}, u_{\mathbf{y}}), \tag{4.4}$$

which using Bayes' Theorem can be written as

$$p(\theta_{\mathbf{x}} = \theta_{\mathbf{y}} | u_{\mathbf{x}}, u_{\mathbf{y}}) \propto p(u_{\mathbf{x}}, u_{\mathbf{y}} | \theta_{\mathbf{x}} = \theta_{\mathbf{y}}) p(\theta_{\mathbf{x}} = \theta_{\mathbf{y}}) \,. \tag{4.5}$$

The likelihood $p(u_{\mathbf{x}}, u_{\mathbf{y}} | \theta_{\mathbf{x}} = \theta_{\mathbf{y}})$ can be computed from the joint distribution $u_{\mathbf{x}}$ and $u_{\mathbf{y}}$ assuming that they both rely on a single parameter θ. For the prior term $p(\theta_{\mathbf{x}} = \theta_{\mathbf{y}})$ the symmetric Kullback-Leibler divergence is computed over an exponential decay function [DDT09; DDT11]. Since the prior requires estimates of $\theta_{\mathbf{x}}$ and $\theta_{\mathbf{y}}$ PPB is an iterative algorithm, which updates the prior using the estimates of the previous iteration.

The following paragraphs describe the similarity criterion for the InSAR case from [DDT11] in greater detail. The likelihood of observing the two SLC images at position $\mathbf{x}$ and $\mathbf{y}$, that is $u_{i,\mathbf{x}}$ and $u_{i,\mathbf{y}}$ ($i = 1,2$) making up the interferogram (see Equation (2.5)), given that the true parameters, the coherence ρ, the intensity I and the interferometric phase φ are identical is [DDT11]

$$p\left(u_{1,\mathbf{x}}, u_{2,\mathbf{x}}, u_{1,\mathbf{y}}, u_{2,\mathbf{y}} | I_\mathbf{x} = I_\mathbf{y}, \theta_\mathbf{x} = \theta_\mathbf{y}, \gamma_\mathbf{x} = \gamma_\mathbf{y}\right) = \sqrt{\frac{B}{C}}^3 \left(\frac{A+C}{A}\sqrt{\frac{C}{A-C}} - \arcsin\sqrt{\frac{C}{A}}\right), \quad (4.6)$$

where

$$A = \left(a_{1,\mathbf{x}}^2 + a_{2,\mathbf{x}}^2 + a_{1,\mathbf{y}}^2 + a_{2,\mathbf{y}}^2\right)^2,$$
$$B = a_{1,\mathbf{x}} a_{2,\mathbf{x}} a_{1,\mathbf{y}} a_{2,\mathbf{y}}, \text{ and}$$
$$C = 4\left(a_{1,\mathbf{x}}^2 a_{2,\mathbf{x}}^2 + a_{1,\mathbf{y}}^2 a_{2,\mathbf{y}}^2 + 2B\cos\left(\varphi_\mathbf{x} - \varphi_\mathbf{y}\right)\right).$$

The probability that the true intensity, interferometric phase and coherence are identical is gauged by the prior term $p(\theta_\mathbf{x} = \theta_\mathbf{y})$. For two zero-mean complex circular Gaussian distributions, the underlying joint distribution of $\hat{\rho}, \hat{I}$ and $\hat{\varphi}$, the symmetric Kullback-Leibler divergence is given by [DDT11]

$$p(\theta_\mathbf{x} = \theta_\mathbf{y}) \propto \frac{4}{\pi}\left[\frac{\hat{I}_\mathbf{x}}{\hat{I}_\mathbf{y}}\frac{1-\hat{\rho}_\mathbf{x}\hat{\rho}_\mathbf{y}\cos(\hat{\varphi}_\mathbf{x} - \hat{\varphi}_\mathbf{y})}{1-\hat{\rho}_\mathbf{y}^2} + \frac{\hat{I}_\mathbf{y}}{\hat{I}_\mathbf{x}}\frac{1-\hat{\rho}_\mathbf{y}\hat{\rho}_\mathbf{x}\cos(\hat{\varphi}_\mathbf{y} - \hat{\varphi}_\mathbf{x})}{1-\hat{\rho}_\mathbf{x}^2} - 2\right]. \quad (4.7)$$

Equation (4.6) and Equation (4.7) together define the similarity of two interferogram pixels. For PPB all pixel probabilities are assumed to be independent and the patch probability can simply be computed as the product over the contained pixel probabilities, or alternatively summing up their logarithms.

Besides the similarity computation and the iterative weight computation [DDT09; DDT11] are largely identical to the original nonlocal algorithm proposed by Buades, Coll, and Morel in [BCM05b]. Both use fixed patch sizes, making them susceptible to the rare patch effect, and also apply a straightforward exponential weighting kernel.

NL-SAR

NL-SAR [Del+15], proposed by the same author as NL-InSAR, addresses several of its shortcomings. It employs a similarity criterion derived for Wishart-distributed matrices (see Equation (2.16)), and can therefore operate on simple SAR, InSAR, PolSAR and PolInSAR data, making it suitable for a wider array of applications than NL-InSAR.

The similarity criterion was originally proposed by [Con+03] for PolSAR change detection. It is based on a generalized likelihood-ratio test (GLRT) for two complex Wishart-distributed matrices $\boldsymbol{X}$ and $\boldsymbol{Y}$, obtained from zero mean complex normal random variables with covariance matrices $\boldsymbol{C_X}$ and $\boldsymbol{C_Y}$. The GLRT considers two hypotheses

1. $H_0 : \boldsymbol{C_X} = \boldsymbol{C_Y}$
2. $H_1 : \boldsymbol{C_X} \neq \boldsymbol{C_Y}$

Following the derivation in [Con+03] one arrives at the following GLRT test statistic

$$Q = \frac{(n+m)^{p(n+m)}}{n^{pn}m^{pm}}\frac{|\boldsymbol{X}|^n|\boldsymbol{Y}|^m}{|\boldsymbol{X}+\boldsymbol{Y}|^{n+m}} \quad (4.8)$$

where p is the dimension of data, i.e. one for SAR, two for InSAR, ...and m and n are the number of looks of $\boldsymbol{X}$ and $\boldsymbol{Y}$. In fact the Wishart-based GLRT was first used as a similarity criterion for nonlocal filtering in [Che+11], but solely for denoising PolSAR images.

NL-SAR automatically selects several filtering parameters to produce the estimate with the lowest variance. As Equation (4.8) requires covariance matrices of full rank, as a preprocessing step the data is convolved with a Gaussian kernel to obtain an initial covariance estimate. The width of the Gaussian kernel is one of the automatically selected parameters. Others are the patch size, which combats the rare patch effect (Section 3.2.1), and the search window size (Section 3.2.2), improving the robustness of the filtering estimate. NL-SAR employs a data-driven weighting kernel, using preselected homogeneous training data, that takes pixel correlations into account. Furthermore, NL-SAR includes a bias-reduction step to better preserve point targets or heterogeneous areas.

All of these additional measures compared to a basic nonlocal filter make NL-SAR more widely-applicable and able to deal with a lot of edge-cases. Consequently, to this day NL-SAR can be considered as the most advanced nonlocal InSAR filter and can serve as a reference.

Nonlocal InSAR filtering by transform shrinkage

For the sake of completeness three more nonlocal InSAR filters should be mentioned. The first two are not covered in depth as they ignore the reflectivity of an interferogram, potentially discarding useful information, and their similarity criteria are not derived from InSAR statistics.

The nonlocal filter proposed in [Lin+15] relies on the structure similarity (SSIM) index [Wan+04] as a similarity criterion. SSIM is a metric to quantify the quality of image denoising or compression algorithms and seems to be in better accordance to human perception than peak signal-to-noise ratio (PSNR) or mean squared error (MSE). The actual denoising is performed by hard thresholding after a higher order tensor singular value decomposition [KB09a].

[Fan+16] filters the inphase and quadrature component of the interferometric phase separately by means of wavelet shrinking and uses the Euclidean distance to search for similar patches.

The recently published InSAR-BM3D [Sic+18] follows the main filtering steps of BM3D: grouping, collaborative filtering, and aggregation to denoise interferograms. Unfortunately the publication came too late to influence the work of this thesis.

4.2 Multibaseline InSAR filters

In recent years InSAR stack filtering became more relevant since it can serve as an extension to persistent scatterer interferometry (PSI) [FPR01]. Whereas PSI only exploits high quality persistent scatterers, also time series of consistent, distributed scatterers provide useful information. As their coherence is usually lower denoising is mandatory in many cases to extract accurate measurements. Multilooking is often chosen with the same obvious drawbacks as for single baseline denoising.

4.2.1 Adaptive filtering

Multibaseline adaptive filters provide an estimate of the covariance matrix as in Equation (2.12), but exclude pixels in the averaging window they deem to be coming from different distributions than the target pixel.

They attracted more interest when in 2011 [Fer+11] showed with their SqueeSAR algorithm that the phase of distributed scatterers can be used as an extension to PSI, given it is accurate enough. To preserve more of the original resolution they proposed an adaptive filter based on the Kolmogorov-Smirnov test to identify similar amplitude time-series. Alternative tests for detecting statistically homogeneous pixels are proposed and evaluated in [PB11] and [Jia+15], again with adaptive multilooking as an application.

The line between these methods and nonlocal filtering is not clear-cut but blurred. They do impose that statistically homogeneous pixels must be connected for further robustness, but this restriction is not essential to the filtering process and can be lifted. Using patches, the other defining quality of nonlocal filters, is equivalent to the use of time-series for identifying similar pixels. Both measures increase the accuracy of the estimate by including multiple observations. As the authors of these methods abstained from referring to them as *nonlocal* this dissertation follows their lead.

4.2.2 Nonlocal filters

Just like there are single baseline nonlocal filters, the nonlocal filtering paradigm has been extended to multibaseline filtering. The already mentioned NL-SAR algorithm [Del+15] can directly work on SAR stacks. Although, because of its preestimation step, it is only applicable to comparatively small stacks.

[Lv+14], extends on the Kolmogorov-Smirnov statistical test from SqueeSAR, by performing a scattering analysis to select an either Rayleigh-, Weibull-, lognormal-, or K-distribution based GLRT similarity measure for nonlocal adaptive multilooking. Other recent publications applied the nonlocal filtering paradigm to SAR stacks in the fields of differential SAR interferometry [Sic+15] or 3D reconstruction using SAR tomography [DHo+18]. [Sic+15] investigates different amplitude similarity measures and the effects prefiltering with despeckling algorithms for nonlocal InSAR stack filtering. In a similar vein [DHo+18] proposes a nonlocal TomoSAR filter, which due to the employed similarity measure is however limited to comparatively small stacks, rendering it unsuitable for large stacks, as the Sentinel-1 mission generates.

[Del+17] proposes an ingenious extension of the *homomorphic* transform approach, regularly used in the field of speckle reduction for single SAR images, to stacks of SAR images, which are assumed to follow a multivariate complex Gaussian distribution, by employing the matrix logarithm and matrix exponential. Although technically not a nonlocal approach their method permits to apply any regular denoising algorithm designed for AWGN to stacks of SAR images, making numerous nonlocal algorithms available for multibaseline filtering.

Multibaseline phase unwrapping is robustified in [Fer+18] by combining an estimate of NL-SAR with a TV regularization. The only algorithm making direct use of the phase information is [Wan+18], which uses small baseline subset (SBAS) to obtain a subset of interferograms, on

these interferograms statistically similar temporal vectors are identified using the similarity criterion of NL-InSAR.

4.3 Contributions of this thesis

For both, the single- and the multibaseline case, earlier publications have already shown that nonlocal methods outperform conventional neighbourhood or transform domain filters. But new nonlocal algorithms in the field of denoising optical images with AWGN have shown the advantage of locally adjusting parameters and filtering strategies to the scenes characteristics in that specific area. Albeit NL-SAR already has such an adaptive parameter selection it might be constrained by having to be applicable to SAR, PolSAR, InSAR and PolInSAR. A specialized filter for interferograms has more prior information to work with, which could enable enhanced image restoration capabilities.

With regard to multibaseline filtering, so far filters mostly ignore temporal changes. Taking inspiration from video denoising filters, which often incorporate motion estimation and compensation, by also specifically addressing the temporal dimension in InSAR stacks further filtering enhancements should be possible.

At last, although the denoising performance of nonlocal methods are indisputable, their computational costs of have often been cited as one reason against their widespread adoption or use in production systems outside of academia.

In light of these challenges this thesis makes several contributions to the field of InSAR denoising.

1. Design of a nonlocal filter for SAR interferograms with adaptive parameter selection. The filter avoids some of the shortcomings of a basic nonlocal filter, such as the rare patch effect. It also accounts for deterministic phase components, which would otherwise hamper the search for similar pixels.

2. A systematic evaluation of the proposed single baseline filter, using real data and easily repeatable simulations. TanDEM-X Coregistered Single look Slant range Complex (CoSSC) interferograms serve as input to generate DEMs using DLR's processing chain, which are then compared to aerial photographs or LiDAR digital terrain models (DTMs). These different experiments can serve as a general testing framework to evaluate denoising algorithms for interferograms and compare their filtering characteristics.

3. Analysis and development of similarity criteria for InSAR stacks of varying temporal dimensions and coherence profiles. These serve as a prerequisite for a basic nonlocal multibaseline filtering framework, which subdivides the search volume into matching clusters. The subsequent cluster-wise filtering is not compromised by outliers or temporal changes.

4. Implementation guidelines for nonlocal filtering algorithms to make use of modern compute architectures. These provide substantial speed-ups and make large-scale processing feasible.

5. As a byproduct of the evaluation and the efficient implementation guidelines this thesis shows that it would be possible to generate a global DEM of 6 m resolution by reprocessing existing TanDEM-X CoSSC interferograms, which were acquired for the generation of the global, commercially available DEM.

5 Single baseline SAR interferometry nonlocal filtering

Section 3.2 already raised several shortcomings of the original filter [BCM05b]. This chapter presents a nonlocal filter that addresses them in the context of InSAR. It also contrasts the presented filter with NL-InSAR and NL-SAR (see Section 4.1.5 and the references therein). The presented method was also published in [Bai+18].

Figure 5.1 shows a high-level flow graph of the proposed method called NL-SWAG. The reasoning behind its naming will become apparent in Chapter 6. The following sections describe in greater detail the individual operations and how they affect the filtering performance and outcome. Blocks highlighted in gray are explicitly explained in the respectively named sections, which also cover other adjacent blocks.

5.1 Aggregation of estimates

Figure 3.3 and Figure 3.4 have shown the effect that a simple feature, like an edge, can have on different filters. For the nonlocal filter an aggregation of estimates, by averaging all possible pixel estimates from overlapping patches, can mitigate the *rare patch effect*, enhancing the filtering estimate near edges and other features.

Already back in 1992 [NS92] proposed a nonlinear filter that adapts the kernel shape to image gradients and also displaces the kernel off-center, improving the filter estimate near features and structures. Although this does not in fact add up to an aggregation of multiple estimates it indicates that a simple shift of the estimation kernel can ameliorate the result. This was then actually done by [CD95], which showed that transform domain denoising also produces artifacts near singularities, such as edges and corners, similar to the well-known Gibbs phenomenon of the Fourier transform. Observing that shifts or translations of the input signal diminishes such artifacts the proposed method averages multiple denoised estimates of different shifts. To some extent lapped transforms [Mal92] work in a similar manner. By computing the transform on overlapping blocks the commonly observed blocking artifacts in traditional transform coding schemes can be avoided. The already several times mentioned BM3D filter [Dab+07] adds up estimates of overlapping blocks and serves as an example of a nonlocal filter, that combines and exploits multiple denoised estimates.

Such an aggregation of estimates is also performed by the proposed nonlocal InSAR filter. To do so it computes a patch-wise weighted mean in lieu of the conventional pixel-wise nonlocal means filter as in Equation (2.10)

$$\hat{\boldsymbol{Z}}_{\mathbf{x}} = \sum_{\mathbf{y} \in \partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} \boldsymbol{Z}_{\mathbf{y}} \,, \tag{5.1}$$

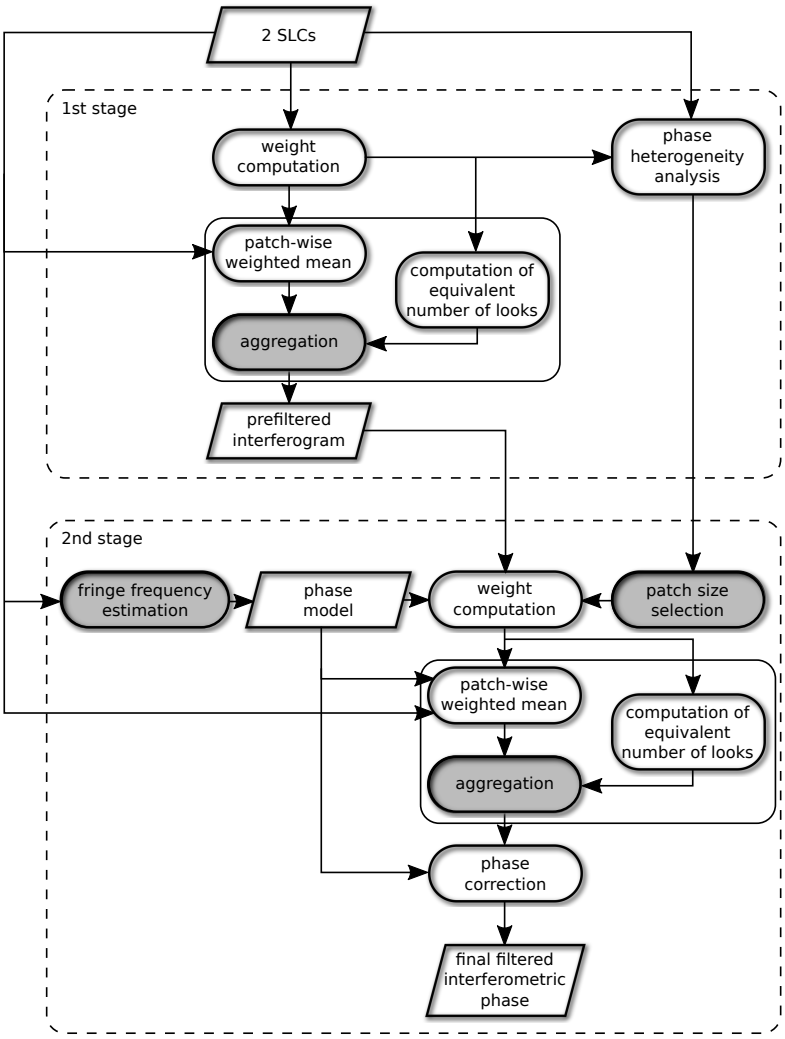


Figure 5.1: Flow graph of the proposed filter. Blocks highlighted in gray have their own respective sections, which also covers other related operations. The second stage uses the prefiltered output of the first stage for computing a new, more reliable set of weights.

where z is a SAR interferogram as defined in Equation (2.5).

Each pixel is part of multiple, overlapping patch estimates, of which the effective number of looks L is an indicator for the reliability. Consequently weighting of the individual patch estimates by their respective effective number of looks

$$\hat{z}_{\mathbf{x}} = \frac{\sum_{\mathbf{y}\in\mathscr{P}_{\mathbf{x}}} L_{\mathbf{y}} \hat{\boldsymbol{Z}}_{\mathbf{y},\mathbf{x}-\mathbf{y}}}{\sum_{\mathbf{y}\in\mathscr{P}_{\mathbf{x}}} L_{\mathbf{y}}}, \tag{5.2}$$

ensures that patch estimates with a higher number of looks, and therefore a smaller variance, have a larger impact on the final estimate. In Equation (5.2) $\mathscr{P}_{\mathbf{x}}$ denotes the set of all pixel indices within a patch centered at $\mathbf{x}$. So individual pixels inside a patch can be designated by the relative index $\mathbf{x}-\mathbf{y}$, i.e. $\boldsymbol{Z}_{\mathbf{y},\mathbf{x}-\mathbf{y}} = z_{\mathbf{x}}$.

The effective number of looks, i.e., the variance reduction of the weighted mean, can directly be computed from the weight map [DDT11]

$$L_{\mathbf{x}} = \frac{\left(\sum_{\mathbf{y}\in\partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}}\right)^2}{\sum_{\mathbf{y}\in\partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}}^2}. \tag{5.3}$$

Aggregation mitigates the rare patch effect as it also properly denoises pixels near features, such as edges, as long as they also belong to patches which do not contain said features.

5.2 Two stage filtering

The argument of nonlocal filters to use patches instead of single pixels was to improve the accuracy of the similarity measurement, as multiple noisy "estimates" are averaged. An alternate approach is the repeated application of a filter, the benefit of which was already depicted in Figure 3.2. Since the noise level decreases with every iteration, the resulting similarities are less and less affected by noise. The downside of this approach is the introduction of piecewise constant plateaus[1] in the resulting image.

The proposed filter in fact combines both approaches in its two stages. The first stage prefilters the image to reduce the noise level and computes image statistics. Subsequently the second stage computes the final denoised output with weights that depend on the prefiltered output of the first stage.

Such two- or even multi-stage approaches are actually quite common. NL-InSAR [DDT11], as previously mentioned, iteratively refines its weights and NL-SAR [Del+15] also incorporates a preestimation step for the covariance matrices. A lot of emphasis on a two stage approach is put on by BM3D, where the prefiltered output of the first stage supports the grouping and Wiener-filtering in the second stage. Another two stage filter is proposed in [SWA12],

[1] This artifact can be explained by a small digression to diffusion filters. In fact there is a direct link between diffusion and neighborhood filters, such as bilateral or nonlocal filters [BC04]. Diffusion filters have the interesting property that their repeated application steadily decreases the noise level and produces piecewise constant approximations of the original data [Wei97]. This can actually be a desired result for image segmentation or generating abstractions [WOG06], for example bilateral filters are often used to cartoonify photographs, but since physical properties seldom change abruptly in the case of InSAR this is unwelcome.

which in the first stage employs wavelet or curvelet thresholding for prefiltering and in the second stage the Yaroslavsky filter [Yar85] to obtain the final estimate. The Yaroslavsky filter is closely related to the bilateral filter, the only difference being that it has a fixed spatial kernel: the indicator function. All these filters have in common that the search for similar pixels is to be performed on an already denoised image.

The two stages, which are also depicted in the filters flowgraph in Figure 5.1, employ the two similarity criteria derived in [DDT11] for two SLC images in the first stage and a filtered interferogram in the second stage. The similarity criteria have already been covered in detail in Section 4.1.5.

5.2.1 First stage

The patch similarity in the first stage Δ^1 is computed as the product over all contained pixel likelihood as defined in Equation (4.6), or equivalently the sum over their logarithms

$$\Delta^1_{\mathbf{x},\mathbf{y}} = \sum_{\mathbf{o}\in\mathcal{O}} \log p(u_{1,\mathbf{x}+\mathbf{o}}, u_{2,\mathbf{x}+\mathbf{o}}, u_{1,\mathbf{y}+\mathbf{o}}, u_{2,\mathbf{y}+\mathbf{o}} | I_{\mathbf{x}+\mathbf{o}} = I_{\mathbf{y}+\mathbf{o}}, \theta_{\mathbf{x}+\mathbf{o}} = \theta_{\mathbf{y}+\mathbf{o}}, \rho_{\mathbf{x}+\mathbf{o}} = \rho_{\mathbf{y}+\mathbf{o}}), \tag{5.4}$$

where $\mathcal{O}$ denotes the set of all relative index offsets in the patch..

The similarities are mapped into weights by an exponential kernel as in Equation (3.4). As the purpose is only to reduce the noise level and remove outliers without introducing severe filtering artifacts, before computing the similarities in the second step, h is set to a comparatively small value. Except for the aggregation step the first stage is identical to the non-iterative version of NL-InSAR and its guidelines for picking h can be used. Estimates of the phase and intensity are obtained via Equation (2.10) by taking the argument and the absolute value. The coherence can be computed using

$$\hat{\rho} = \frac{\sum_{\mathbf{y}\in\partial_\mathbf{x}} w_{\mathbf{x},\mathbf{y}} u_1 \bar{u}_2}{\sum_{\mathbf{y}\in\partial_\mathbf{x}} w_{\mathbf{x},\mathbf{y}} \frac{1}{2}\left(|u_1|^2 + |u_2|^2\right)}. \tag{5.5}$$

5.2.2 Second stage

The second stage computes the similarities as a function of the coherence $\hat{\rho}$, intensity $\hat{I}$ and interferometric phase $\hat{\varphi}$ estimates, produced by the first stage, and uses NL-InSAR's prior term Equation (4.7) as a similarity criterion. Instead of a fixed patch size the second stage changes the patch size adaptively based on the local heterogeneity. The exact patch similarity and weight computation is covered in the following two sections since, as visible from Figure 5.1, it is based on other operations.

Even though the two-step approach alleviates the problems caused by the high noise level in SAR images, it is important to stress that a repeated application of any filter can potentially introduce staircase-like artifacts in the filtered output. These artifacts are for example very noticeable in some of the later experiments for the iterative NL-InSAR algorithm.

5.3 Adaptive patch size selection

Image patches contain information about the local texture and structure. These features play a crucial role in distinguishing between patches inside the search window and determine whether to discard from or include one in the weighted mean. Which raises the question: how to select the best patch size? In [DAG11] the authors demonstrate that a global selection is suboptimal and that the patch size should depend on the local neighborhood. The following section repeats their reasoning in the context of remote sensing, i.e. observations of large areas of very different nature.

As an example, for the original nonlocal filter, the patch similarity, just like Equation (5.4), is essentially the sum of all contained pixel similarities. Naturally large patches reduce the variance and provide the most robust estimate of the patch similarity. This is indeed the best strategy for plains, agricultural fields or other slowly varying terrain.

The situation is quite different for more complex terrain, for instance urban sites or mountain ridges. In these areas a large patch sizes leads to the rare patch effect, since for every patch that contains some local structure only patches with similar features will have a significant impact on the averaging process. The likelihood of finding such patches decreases with increasing patch size.

NL-SWAG's solution is to adaptively select the patch size as a function of local scene heterogeneity. This way a more robust patch similarity can be computed in flat regions or moderately hilly areas, due to the larger patch size, while at the same time the rare patch effect is alleviated in areas with many features and details. Although small patches come at the cost of less reliable patch similarity estimates.

The argument for an adaptive patch size selection to avoid the rare patch effect is actually identical to the one for aggregation. Both measures favor patches that exclude local structures by either shrinking the patch or including estimates where the patch is moved off-center with respect to the pixel that is to be denoised. This is somewhat contrary to the initial argument that patch-based methods perform so well because they take textures and details into account. Texture and details are indeed an effective means for discarding patches of different classes. But to maximize the number of patches that are classified as similar both techniques also try using the just mentioned patch modification schemes.

To identify heterogeneous pixels and select the patch size accordingly NL-SWAG applies the local phase heterogeneity measure derived in [Lee+98]

$$\eta_{\mathbf{x}} = \frac{\operatorname{Var}\{\varphi\}_{\mathbf{x}} - \sigma^2_{0,\mathbf{x}}}{\operatorname{Var}\{\varphi\}_{\mathbf{x}}}, \tag{5.6}$$

which lies in the interval $[0,1)$. $\operatorname{Var}\{\cdot\}$ denotes the variance of a random variable, so that $\operatorname{Var}\{\varphi\}$ is the estimated variance of the phase in the search window. σ_0^2, on the other hand, is the variance one would expected from the coherence [BH98]. For non-heterogeneous terrain $\operatorname{Var}\{\varphi\}$ is comparable in magnitude to σ_0^2, as only phase noise causes changes of the phase and Equation (5.6) is close to 0. The situation changes when the search window contains structures, such as buildings. Their distinct phase profiles increase $\operatorname{Var}\{\varphi\}$ resulting in larger phase heterogeneity values.

As the phase is wrapped, the filter first performs local unwrapping as in [Lee+98] to obtain the locally unwrapped phase $\tilde{\varphi}$ with respect to the average of the 5×5 pixels in the center. The phase variance is then estimated inside the search window, weighted by the respective weight map computed in the first stage

$$\begin{aligned} \operatorname{Var}\{\varphi\}_{\mathbf{x}} &= \mathrm{E}\{\varphi_{\mathbf{x}}^2\} - \mathrm{E}\{\varphi_{\mathbf{x}}\}^2 \\ &\approx \sum_{\mathbf{y}\in\partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} \tilde{\varphi}_{\mathbf{y}}^2 - \left(\sum_{\mathbf{y}\in\partial_{\mathbf{x}}} w_{\mathbf{x},\mathbf{y}} \tilde{\varphi}_{\mathbf{y}} \right)^2 , \end{aligned} \tag{5.7}$$

where $\mathrm{E}\{\cdot\}$ denotes the expected value. As $\operatorname{Var}\{\varphi\}$ is estimated in a local window due to insufficient sample size Equation (5.6) might be negative. In this case the heterogeneity measure is set to zero.

To yield a more reliable estimate of σ_0 the coherence is estimated following the methodology in [GP97] as

$$\gamma = \frac{\mathrm{E}\{|u_1|^2 \cdot |u_2|^2\}}{\sqrt{\mathrm{E}\{|u_1|^4\}\,\mathrm{E}\{|u_2|^4\}}} . \tag{5.8}$$

This way the coherence is estimated from the speckle pattern and is not influenced by the topographic phase, which would yield an underestimation of the coherence if the common coherence estimator is used. Just like in Equation (5.7) the expected value is replaced by the weighted mean over the respective quantities.

An example of the heterogeneity measure is depicted in Figure 5.2. The urban area is clearly detected as being heterogeneous, the grassland is classified as the most homogeneous site and the forested areas are identified as moderately heterogeneous regions.

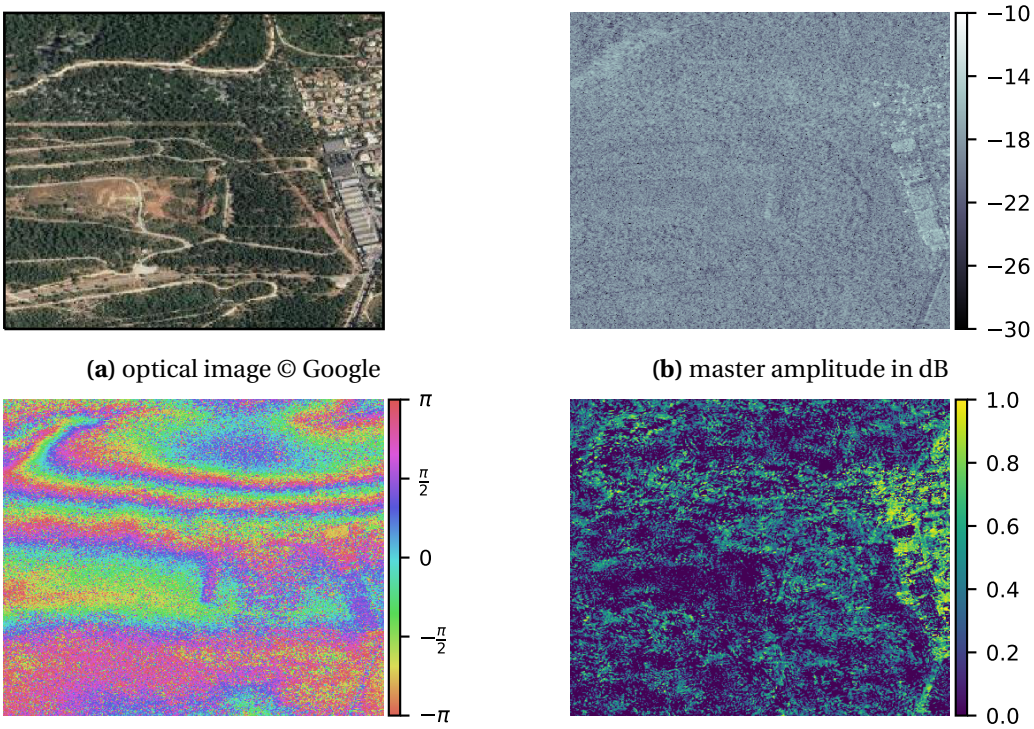
Instead of selecting a fixed patch size from a predefined set, depending on the local heterogeneity, NL-SWAG employs Gaussian windows of variable width. A possible mapping of the phase heterogeneity index into Gaussian window widths could be

$$\sigma_{\mathrm{Gauss}} = 2 \cdot (1 - \eta) + 1 , \tag{5.9}$$

which gives strict lower and upper bounds for the window widths. Other mappings would also be possible as long as they result in wide Gaussian windows for homogeneous areas and the reverse for heterogeneous areas.

As an alternative approach for selecting the best effective patch size the phase variance in Equation (5.6) could be computed in Gaussian windows of successively increasing widths. This process is halted as soon as the heterogeneity level exceeds a predefined threshold, i.e. when significant phase changes, which most likely are the result of heterogeneous structures inside the patch, are detected. A similar approach was presented in [KB08] for adaptively selecting the search window size.

For Gaussian blurring the reduction in variance is related to σ_{Gauss} by approximately $4\pi\sigma_{\mathrm{Gauss}}^2$. So with Equation (5.9) the variance of the patch similarity estimation is reduced by a factor ranging from 4π to 36π, roughly equivalent to 3×3 up to 11×11 patches.



(a) optical image © Google

(b) master amplitude in dB

(c) phase

(d) phase heterogeneity

Figure 5.2: Phase heterogeneity computed in the first stage. Urban areas, forests and grassland show different level of heterogeneity.

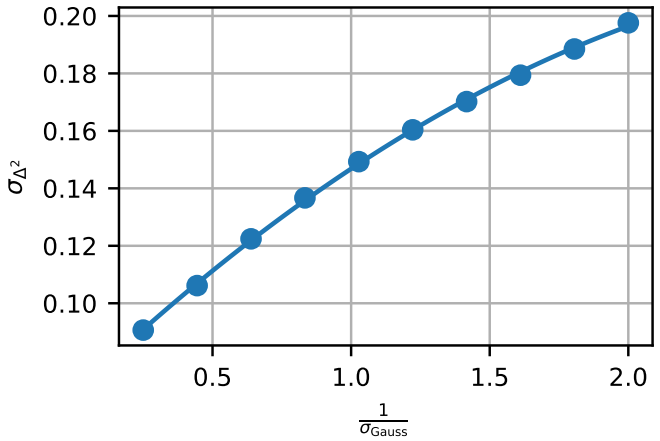


Figure 5.3: Relation between the width of the Gaussian window σ_{Gauss} and the standard deviation of the resulting patch similarities σ_{Δ^2}. Due to correlation of the pixel similarities there is no linear mapping.

Correspondingly to Equation (5.4) the adaptive patch similarities in the second stage are computed as the sum over the pixel similarities weighted by a Gaussian window $\boldsymbol{G}_{\mathbf{x}}$

$$\Delta^2_{\mathbf{x},\mathbf{y}} = \frac{\sum_{\mathbf{o}\in\mathcal{O}} \boldsymbol{G}_{\mathbf{x},\mathbf{o}}\,\delta^2_{\mathbf{x}+\mathbf{o},\mathbf{y}+\mathbf{o}}}{\sum_{\mathbf{o}\in\mathcal{O}} \boldsymbol{G}_{\mathbf{x},\mathbf{o}}}. \tag{5.10}$$

where δ^2 denotes the symmetric Kullback-Leibler divergence in Equation (4.7).

The patch dissimilarities still need to be mapped into weights, which in the second stage is also done by an exponential kernel. The normalization factor h directly impacts the trade-off between bias and variance.

The standard deviation of Δ^2 is inversersely proportional to σ_{Gauss}, which effectively governs the patch size. Consequently a fixed h for all heterogeneity levels will be insufficient and a method is needed that accounts for varying patch sizes. For this purpose, on a homogeneous training area one analyzes how the patch similarity's standard deviation σ_{Δ^2} changes with varying $\frac{1}{\sigma_{\text{Gauss}}}$. Figure 5.3 shows the relationship for a fixed set of Gaussian window widths in an homogeneous test site without any topography. Clearly the relationship is non-linear, due to correlation between pixel similarities, but a second order polynomial, also depicted, is a good fit.

The weights are computed as

$$w_{\mathbf{x},\mathbf{y}} = \exp\left\{-\frac{\Delta^2_{\mathbf{x},\mathbf{y}}}{h\cdot\xi\left(\sigma^{-1}_{\text{Gauss},\mathbf{x}}\right)}\right\}, \tag{5.11}$$

where ξ is the second order polynomial that accounts for the varying effective patch sizes and h provides a fixed compromise between detail preservation and noise reduction. In general, the interval $[1 \le h \le 2]$ has proven to offer the best trade-off between the contradictory requirements.

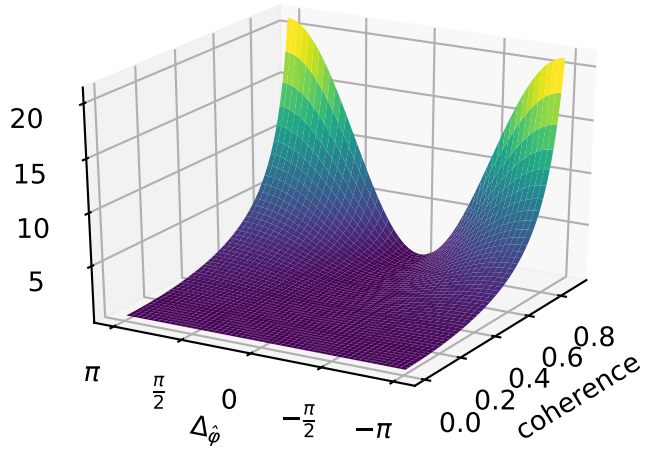


Figure 5.4: Symmetric Kullback-Leibler divergence from Equation (4.7) for two pixels with identical reflectivity and coherence, dependent on their phase difference.

To account for the fact that, due to the Gaussian window, not every pixel in the patch estimate contributed equally to the similarity computation in contrast to Equation (5.2) the respective pixels are additionally weighted by their Gaussian weight in the final aggregation step

$$\hat{z}_{\mathbf{x}} = \frac{\sum_{\mathbf{y}\in\mathscr{P}_{\mathbf{x}}} L_{\mathbf{y}} \boldsymbol{G}_{\mathbf{y},\mathbf{x}-\mathbf{y}} \hat{\boldsymbol{Z}}_{\mathbf{y},\mathbf{x}-\mathbf{y}}}{\sum_{\mathbf{y}\in\mathscr{P}_{\mathbf{x}}} L_{\mathbf{y}} \boldsymbol{G}_{\mathbf{y},\mathbf{x}-\mathbf{y}}} \,. \tag{5.12}$$

5.4 Fringe frequency compensation

Atmosphere, height and deformation all contribute to the phase term in Equation (2.6) and as such hinder the search for similar pixels, by affecting the similarity criterion, which is supposed to only take noise into account. Figure 5.4 shows the symmetric Kullback-Leibler divergence from Equation (4.7) with $\hat{I}_{\mathbf{x}} = \hat{I}_{\mathbf{y}}$ and $\hat{\gamma}_{\mathbf{x}} = \hat{\gamma}_{\mathbf{y}}$ as a function of the coherence and the phase difference $\Delta_{\hat{\varphi}} = \hat{\varphi}_{\mathbf{x}} - \hat{\varphi}_{\mathbf{y}}$, that is used in the second stage as the similarity criterion. Evidently the similarity quickly drops off with increasing $\Delta_{\hat{\varphi}}$ and the higher the coherence the more dramatic the decline. Consequently the denoising performance suffers if the atmosphere's, height's or deformation's contributions are not accounted for.

This analysis is not exclusive to the Kullback-Leibler divergence. Similar arguments can be made for different similarity criteria, i.e. the one employed in Equation (4.6) and Equation (4.8).

Inside the search window in first approximation the phase component of the atmosphere, the height and the deformation can be considered to have a linear trend. NL-SWAG takes them into account by incorporating a linear fringe model as in [SLB10] when computing the similarities and the weighted mean. The linear fringe model compensates to some degree for deterministic phase components such that the search for similar pixels yields more results

and the denoising performance increases. This approach is distantly related to [FB17], which employs affine transforms to find more similar patches.

For every pixel this fringe compensation algorithm obtains an estimate of the fringe frequencies in azimuth and range $\mathbf{f} = \left[f_\mathrm{r}, f_\mathrm{az}\right]^T$ using the 2D Fourier transform. To circumvent abrupt changes a Gaussian kernel smoothes the fringe frequency estimates $\mathbf{f}$.

Without loss of generality Equation (4.7) ca be considered as a function of only the phase difference between two pixels

$$\delta^2_{\mathbf{x},\mathbf{y}}(\hat{\varphi}_\mathbf{x} - \hat{\varphi}_\mathbf{y})\,. \tag{5.13}$$

The fringe compensation takes the fringe frequencies at $\mathbf{x}$ into account by changing the pixel similarity function to

$$\delta^2_{\mathbf{x},\mathbf{y}}(\hat{\varphi}_\mathbf{x} - (\hat{\varphi}_\mathbf{y} - (\mathbf{x}-\mathbf{y})^T\mathbf{f}_\mathbf{x}) \bmod 2\pi), \tag{5.14}$$

so that the phase component caused by the fringe frequency in azimuth and range is removed.

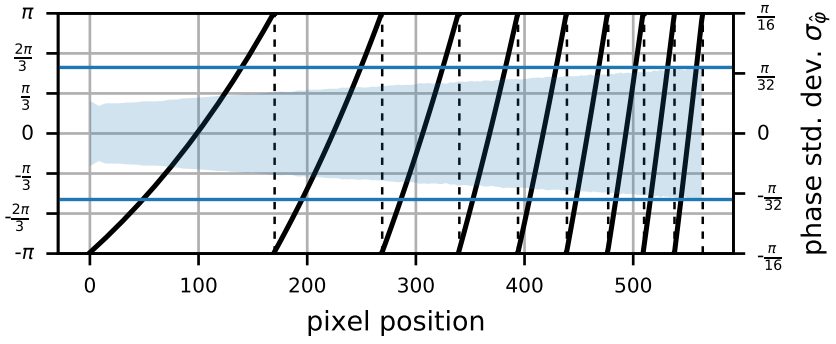
The computation of the patch-wise weighted mean of the interferogram has to account for the phase model

$$\hat{\boldsymbol{Z}}_\mathbf{x} = \sum_{\mathbf{y}\in\partial_\mathbf{x}} w_{\mathbf{x},\mathbf{y}} \boldsymbol{Z}_\mathbf{y} \circ e^{-\mathrm{j}\mathcal{F}_x \bar{\times}_1 (\mathbf{x}-\mathbf{y})}\,. \tag{5.15}$$

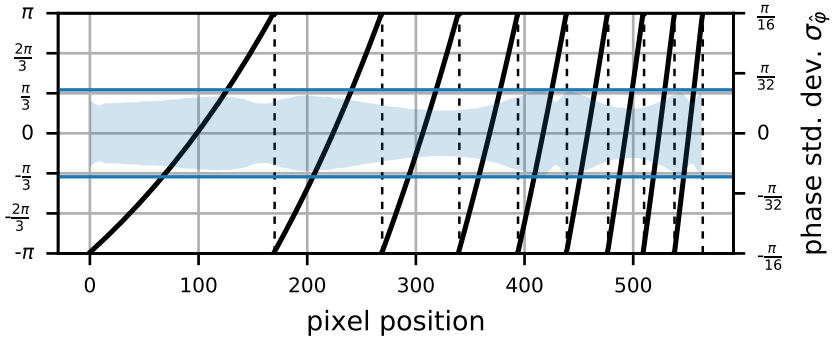
Here $\circ$ denotes element wise multiplication, $\mathcal{F} \in \mathbb{R}^{2\times p\times p}$ is a third order tensor that contains all fringe frequencies of the pixels inside the $p \times p$ patch centered at $\mathbf{x}$ and $\bar{\times}_1$ is the n-mode product as defined in [KB09b].

Figure 5.5 shows the effect that fringe frequency compensation has on the noise reduction. Denoising of a nonlinear phase ramp with constantly increasing frequency is performed using NL-SWAG with and without fringe frequency compensation. If the fringe frequency is not accounted for the phase estimate's standard deviation increases steadily with increasing frequency. With fringe frequency compensation the standard deviation is limited. Due to the discrete nature of the frequency estimation by fast Fourier transform in the implementation, the frequency is not perfectly estimated and the performance is not entirely frequency independent, which results in the wave-like pattern of the standard deviation. A more sophisticated frequency estimation algorithm would certainly alleviate this problem.

As a final note the difference between the fringe frequency compensation and the local phase heterogeneity based adaptive patch size selection needs to be emphasized. Both approaches address deterministic phase changes which can hamper the search for similar patches. But whereas the fringe frequency compensation strictly deals with large-scale phase changes due to topography by a linear compensation, the role of the phase heterogeneity is more to take care of arbitrary small-scale phase changes, which would not be necessarily captured by a simple linear approximation.



(a) without fringe compensation



(b) with fringe compensation

Figure 5.5: Standard deviation (shaded blue area) of NL-SWAG's estimate of a nonlinear phase profile (in black) with and without compensating for the fringe frequency. The maximum value of the standard deviation are marked with a horizontal blue line. If the filter does not account for the deterministic phase change inside the search window the denoising performance decreases substantially with increasing frequency.

5.5 Validation

This section presents experiments that justify the proposed filter's design decisions and contrasts its performance to other filters. Namely

- a simple 5×5 Boxcar filter,
- the original Goldstein filter with tile sizes of 32×32, 75% overlap, $\alpha = 0.5$ and a 3×3 mean spectrum smoothing kernel,
- NL-InSAR (Section 4.1.5) with a search window size of 21×21, 7×7 patches and in total five iterations, and
- NL-SAR (Section 4.1.5) uses the same parameter space as in the original paper: a circular search window with a radius of 12, patch sizes ranging from 3×3 to 11×11 with increments of two and different Gaussian kernels for the preestimation step.

The Boxcar and Goldstein filters are representative for conventional InSAR filters, whereas the two nonlocal filters can be considered as *state-of-the-art*.

Experiments on TanDEM-X stripmap interferograms and simulated data are the basis of the evaluation. Synthetic SLCs with the desired amplitude A, interferometric phase φ and coherence γ profile can be obtained from regular independent complex normal distributed samples, assuming fully developed speckle. Let $\boldsymbol{C} = \mathbf{L}\mathbf{L}^{\dagger}$ be the Cholesky decomposition, where $\dagger$ denotes conjugate transpose, of the covariance matrix of the two SLCs [Goo63]

$$\boldsymbol{C} = \begin{bmatrix} a^2 & a^2\rho e^{\mathrm{j}\varphi} \\ a^2\rho e^{-\mathrm{j}\varphi} & a^2 \end{bmatrix}. \tag{5.16}$$

A multiplication by $\mathbf{L}$ transforms two independent complex normal distributed samples r_1 and r_2 of zero mean and unit variance

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \boldsymbol{L} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} = a \begin{bmatrix} 1 & 0 \\ \rho e^{-\mathrm{j}\varphi} & \sqrt{1-\rho^2} \end{bmatrix} \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \tag{5.17}$$

into samples with the desired correlation properties, amplitude and phase defined by the covariance matrix.

5.5.1 Noise reduction

In order to get an understanding for the noise reduction a TanDEM-X interferogram of Salar de Uyuni, a salt flat in Bolivia, is processed by all filters. Salar de Uyuni is almost perfectly flat on an area of 10 582 km^2 with height variations below 1 m, making it well-suited for calibration purposes. Figure 5.6 shows normalized histograms of the amplitude and the interferometric phase of a $4{,}000 \times 4{,}000$ pixel sized area of Salar de Uyuni. For comparison the theoretical distributions introduced in Chapter 2 are also plotted, which almost perfectly fit the histograms. This is testament to how well this test site is suited for calibration and evaluation, since it features no discernible elevation and intensity changes,

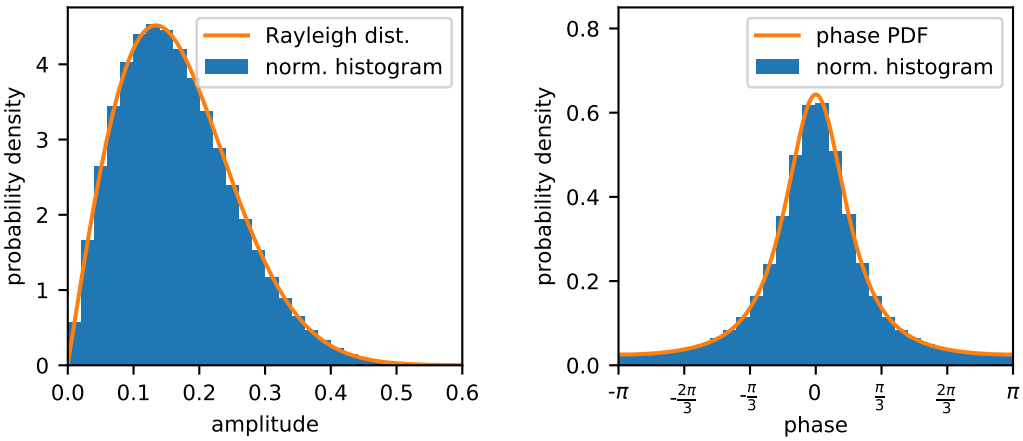


Figure 5.6: Normalized histograms of the amplitude and the interferometric phase for the Salar de Uyuni test site. Their theoretical distributions were fitted to the data and are in almost perfect agreement.

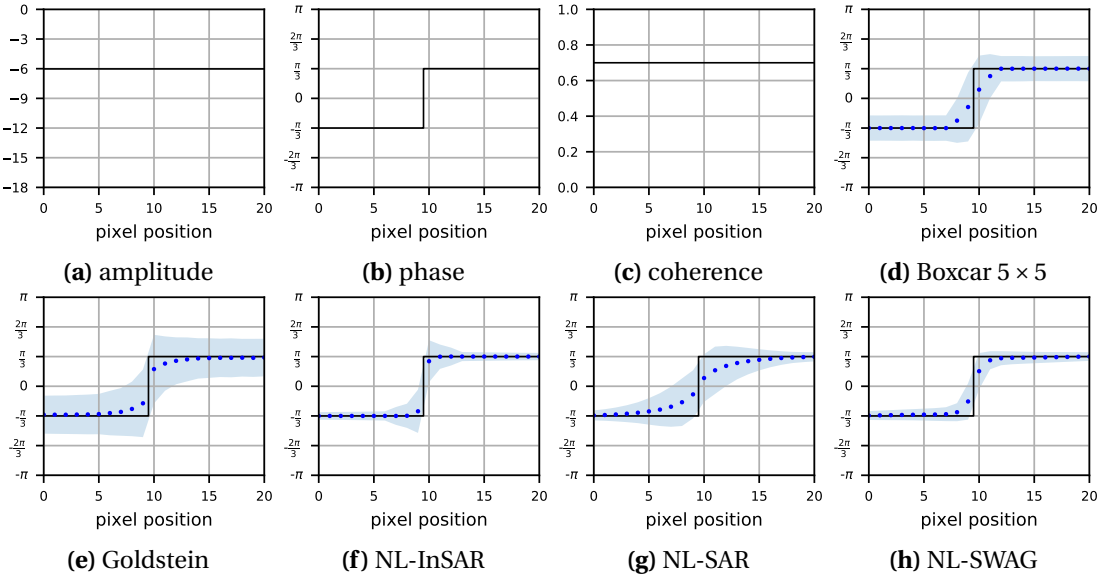
Table 5.1: Phase standard deviation $\sigma_{\hat{\varphi}}$ in radians and the equivalent number of looks L, for a flat and homogeneous area of the salt flat Salar de Uyuni. The analyzed nonlocal filters provide roughly a gain of three compared to the boxcar filter and slightly better performing Goldstein filter.

	unfiltered phase	5 × 5 Boxcar	Goldstein	NL-InSAR	NL-SAR	NL-SWAG
$\sigma_{\hat{\varphi}}$	0.9791	0.2183	0.1822	0.0812	**0.0656**	0.0675
L	1	22.12	28.88	145.39	**222.76**	210.40

The phase standard deviation of the aforementioned filters' estimates are reported in Table 5.1 together with the equivalent number of looks. The 5 × 5 Boxcar filter provides an improvement close to a factor of five, falling short of the theoretical value due to correlation between neighbouring pixels. The correlation is a result of slight oversampling and windowing for sidelobe suppression to weaken azimuth ambiguities. The Goldstein filter's performance is comparable with the Boxcar filter. But all nonlocal filters improve on that by a factor of approximately three. A result of the test site's extreme homogeneity, which enables the nonlocal filters to take perfect advantage of their large search windows.

5.5.2 Detail preservation

Simulation of a phase jump from $-\frac{\pi}{3}$ to $\frac{\pi}{3}$ accompanied by various reflectivity and coherence profiles shows their influence on the estimation and can be used to judge how well InSAR filters preserve resolution. The results of Monte-Carlo simulations with 10000 repetitions are reported in Figure 5.7, Figure 5.8, and Figure 5.9, which show the true reflectivity, phase and coherence and the different filters' phase estimates. The estimates' expected values are plotted as blue dots ± three times their standard deviations as shaded blue areas.



(a) amplitude **(b)** phase **(c)** coherence **(d)** Boxcar 5×5

(e) Goldstein **(f)** NL-InSAR **(g)** NL-SAR **(h)** NL-SWAG

Figure 5.7: Expected values of various filters' phase estimate's of a step function with constant amplitude and coherence of 0.7. The shaded blue area delineates $\pm$ three times the estimate's standard deviation. 10000 simulations were made to obtain the statistics.

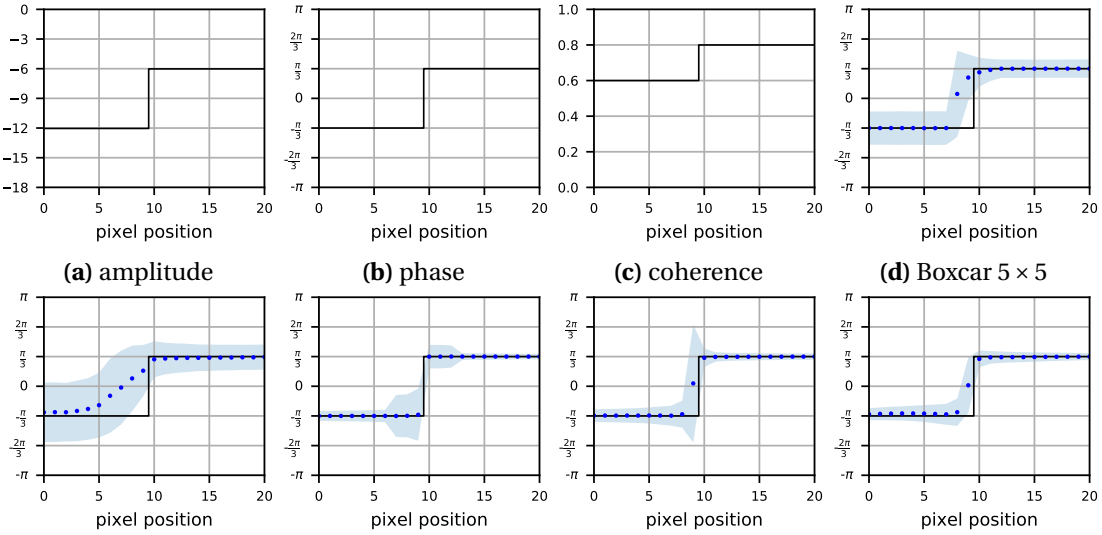
In the first example, depicted in Figure 5.7, both the reflectivity and the coherence are constant. The Boxcar filter gives the expected smoothing directly related to its width. Compared to the Boxcar filter's estimate the Goldstein filter smoothes the edge to a lesser degree but its estimate is of higher variance. This bias-variance trade-off can of course be tuned by different values of α. In general the Boxcar and Goldstein filter produces results of comparable quality.

The nonlocal filters merit a more verbose discussion. Pixels remotely located from the jump give a bias free estimate of low variance. Closer to the jump several factors come into play for all filters.

1. The estimates are increasingly biased since the search window contains more and more pixels of different statistics, which the nonlocal filters are unable to perfectly suppress.

2. The rare-patch effect at the edge increases the variance of the estimates.

The rare-patch effect is especially pronounced in NL-InSAR's estimate, since it employs a fixed patch size. Indeed the noisy strip's width at the jump matches the employed patch size. All patches in this area include the edge and consequently suffer from the rare patch effect. Yet the expected values of NL-InSAR's phase estimate delineate a sharp transition, which is a result of the iterative refinement of the computed weights. NL-SWAG in contrast smoothes the edge, but due to its variable patch width and the aggregation step it can mostly avoid the rare-patch effect. The reason for NL-SAR's considerable oversmoothing is its approach of intentionally producing a preliminary oversmoothed result and later applying a bias-reduction step based on a heterogeneity test, which however only considers the intensity and therefore breaks down when only the phase changes.



(a) amplitude **(b)** phase **(c)** coherence **(d)** Boxcar 5 × 5

(e) Goldstein **(f)** NL-InSAR **(g)** NL-SAR **(h)** NL-SWAG

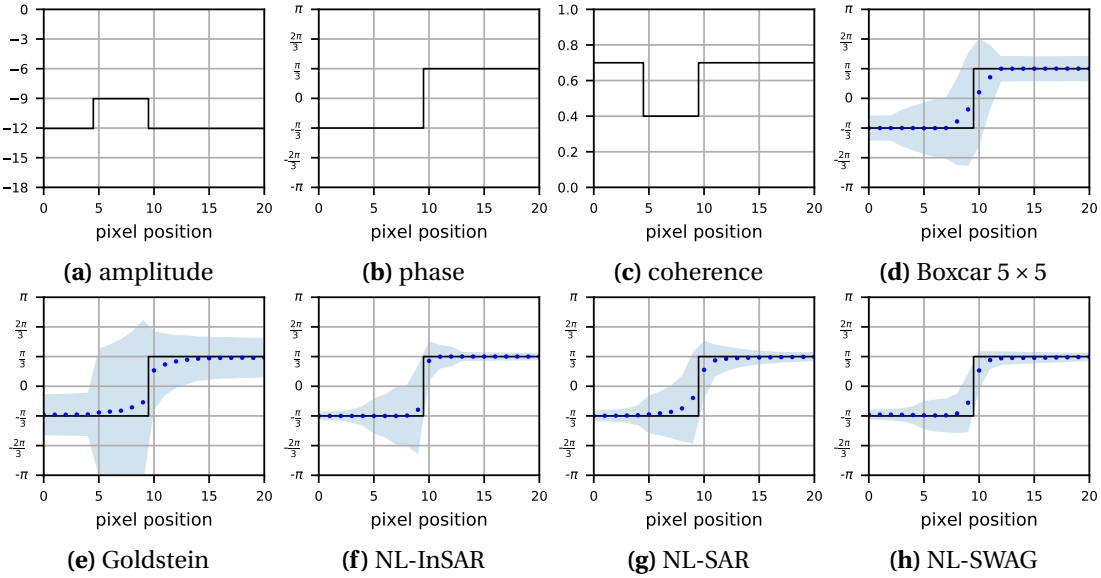
Figure 5.8: Expected value of a step function's phase estimate, constant amplitude and coherence of 0.7. The shaded blue area delineates ± three times the estimate's standard deviation. 10 000 simulations were made to obtain the statistics.

Since the similarity criteria are not only a function or phase but also of intensity or coherence (see Section 4.1.5) the situation changes when the phase jump is accompanied by changes of other variables. Such is the situation for the experiment in Figure 5.8, where also the intensity and coherence change. Compared to Figure 5.7 the intensity change aids nonlocal filters in discriminating between similar pixels, resulting in sharper transitions. The higher coherence also plays a role in biasing the estimates as the pixels on the right hand side have a larger impact on the average, due to their higher coherence, than the left hand sided pixels. The effect of the intensity change, which helps the nonlocal filters, biases the filtering result of the indiscriminate Boxcar and Goldstein filter.

The rare patch effect is now more pronounced for NL-InSAR than in Figure 5.7. The benefit of setting the patch size adaptively is highlighted by NL-SAR and NL-SWAG, which do not exhibit a halo of high variance at the discontinuity. NL-SWAG additionally benefits from the aggregation step, which further reduces the variance along the edge. Although even with these measures in place the variance near the edge is increased.

A more complex scenario is depicted in Figure 5.9, which models an overlay area of higher amplitude but low coherence. Here again all nonlocal filters provide a sharper transition than the legacy Boxcar and Goldstein filter.

Overall the results shown Figure 5.7, Figure 5.8, and Figure 5.9 are proof of the advantages of nonlocal filters. With the exception of NL-SAR in one single edge case, they provide significantly less biased and noisy estimates than legacy filters. Although the conducted experiments also highlight some of their limits: a non-perfect reconstruction of the edge and the susceptibility to the rare-patch effect, even with countermeasures in place. The following pages present experiments that further elucidate the characteristics of nonlocal filters.



(a) amplitude **(b)** phase **(c)** coherence **(d)** Boxcar 5×5

(e) Goldstein **(f)** NL-InSAR **(g)** NL-SAR **(h)** NL-SWAG

Figure 5.9: Expected value of a step function's phase estimate, with amplitude and coherence profiles mimicking layover. The shaded blue area delineates ± three times the estimate's standard deviation. 10000 simulations were made to obtain the statistics.

5.5.3 Fringe frequency dependent noise suppression

Section 5.4 introduced the motivation for an integration of fringe frequency compensation in the proposed nonlocal filter. In this section experiments try to gauge its efficiency.

A first analysis is carried out by denoising phase ramps of different inclinations. In the simulations the intensity is set to a constant value and the coherence is set to 0.7. Figure 5.10 shows the standard deviation of the various filters' phase estimates for different fringe frequencies, given as phase change per pixel in radians.

The nonlocal filters are more sensitive to higher fringe frequencies compared to the Boxcar and Goldstein filters as a result of their large search windows. Especially NL-InSAR and NL-SAR, since they do not compensate for the deterministic phase component. The reason for the Boxcar filter's degradation in performance with increasing fringe frequency is that the phasors in its window add up less and less coherently. The Goldstein filter is more affected by spectral leakage at higher frequencies, which reduces the effectiveness of the spectral filtering window. As mentioned earlier the fringe estimation of NL-SWAG is not perfect due to the discrete nature of the fast Fourier transform used in the implementation and hence is still slope dependent. Overall NL-SWAG provides roughly an improvement by a factor of three compared to the Boxcar estimate over all frequencies, which matches the analysis of the Salar de Uyuni test site Table 5.1.

A more involved phase profile can reveal more details about the characteristics of nonlocal filters. The diamond-square algorithm [Mil86] is a fractal algorithm for generating synthetic terrain. Exactly due to this fractal nature it can generate terrain of different sizes but of the same profile. Within this thesis the diamond-square algorithm is used to directly generate

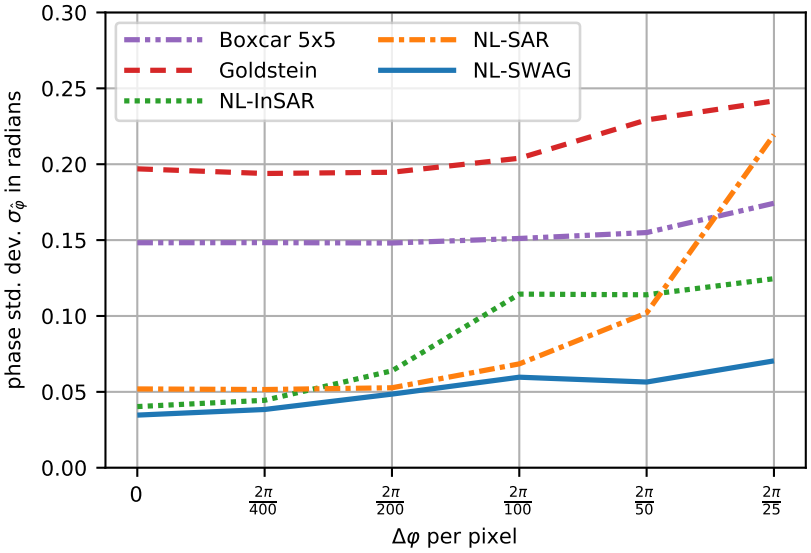


Figure 5.10: Standard deviation of the phase estimate as a function of different fringe frequencies, given as phase change per pixel in radians. The higher the frequencies the higher the standard deviation. The fringe frequency estimation of NL-SWAG alleviates this problem.

unwrapped phase images, for simplicity reasons without taking into account geometric distortions like foreshortening, layover and shadowing. From the corresponding wrapped phase images, noisy interferograms with a coherence of 0.7 are generated.

Figure 5.11 shows four images of different sizes, but identical phase profiles, denoised by NL-SWAG and their difference to the true phase. From the difference maps it is obvious that lower fringe frequencies, corresponding to larger image sizes, help to produce more noise free and less biased estimates.

5.5.4 Variance and bias analysis

An extension of the experiment using fractal terrain of Figure 5.11 illustrates the propensity of the filters to produce the earlier introduced terrace-like features and other biasing artifacts. Again a noisy interferogram with constant reflectivity and a coherence of 0.7, this time of fixed size, helps to understand the various filters' characteristics. Figure 5.12 shows in the top row the simulated noisy interferogram and the filters' denoised results. The second row shows the true simulated phase and its difference to the filters' output.

A TerraSAR- TanDEM-X interferogram, obtained from a CoSSC data product, whose phase resembles the simulation in the analysis exemplifies how these filtering characteristics affect real data. In order to do so shaded reliefs of DEMs generated from it by the various filters are shown in the last row together with the original phase.

Generally the noise level of all nonlocal filters compared to the Goldstein and Boxcar filter is lower, especially in areas with low fringe frequencies. The difference plots identify one further characteristic of nonlocal filters. Whereas their denoised phase estimates look almost

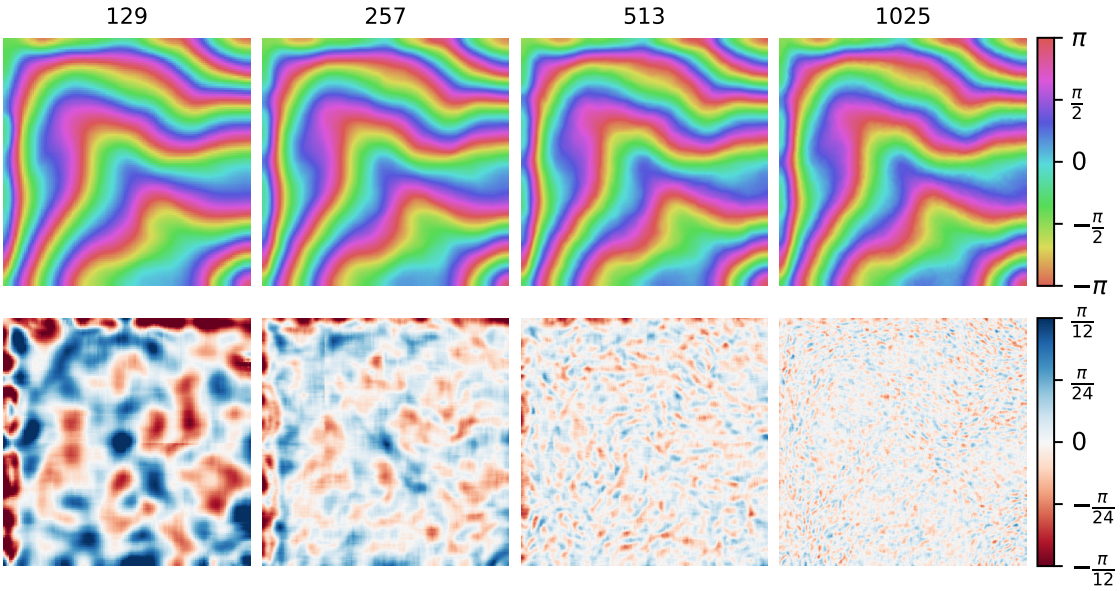


Figure 5.11: Influence of the fringe frequency on the noise reduction and bias. Synthetic interferograms with the same profile but different sizes ($\{s \times s : s \in 129, 257, 513, 1025\}$) are generated. The top row shows the by NL-SWAG denoised results and the bottom row their differences to the true phase. From the difference plots a reduction of bias and noise level with increasing size can be observed.

perfectly noise-free, by studying the difference plots it becomes apparent that the nonlocal filters introduce slowly varying undulations of spatially correlated strips. Basically the filters manage to significantly suppress the noise's high-frequency components but create biased regions, which are later studied in greater detail.

NL-InSAR merits further investigation as it produces a distinct pattern along isophase lines visible in the difference plot. The reason for their formation are akin to the cartoonification by a repeated application of the bilateral filter in Section 3.1.1. In essence NL-InSAR also produces piecewise constant solutions as a result of its many iterations. These filtering artifacts create plateaus of equal heights clearly visible in the DEM generated from the TerraSAR-TanDEM-X CoSSC.

The bias-variance trade-off of nonlocal filter that the previous paragraph alluded to warrants a closer inspection. A Monte-Carlo simulation using the same data as in Figure 5.12 sheds some more light on the bias-variance trade-off and other filtering characteristics. Figure 5.13 shows the expected value and standard deviation of the filters' phase estimates.

All nonlocal filters and the Goldstein filter bias the estimate along the ridge at the interferogram's diagonal. These filters have a higher propensity to bias the estimate due to their comparatively large area of effect, i.e. the search window for nonlocal filters and the tile size of the Goldstein filter. The Boxcar filter is only hypothetically affected at minima or maxima, since the bias introduced by its comparatively small window is usually negligible with regard to the remaining noise.

The standard deviation plots show the fringe frequency dependent noise suppression of

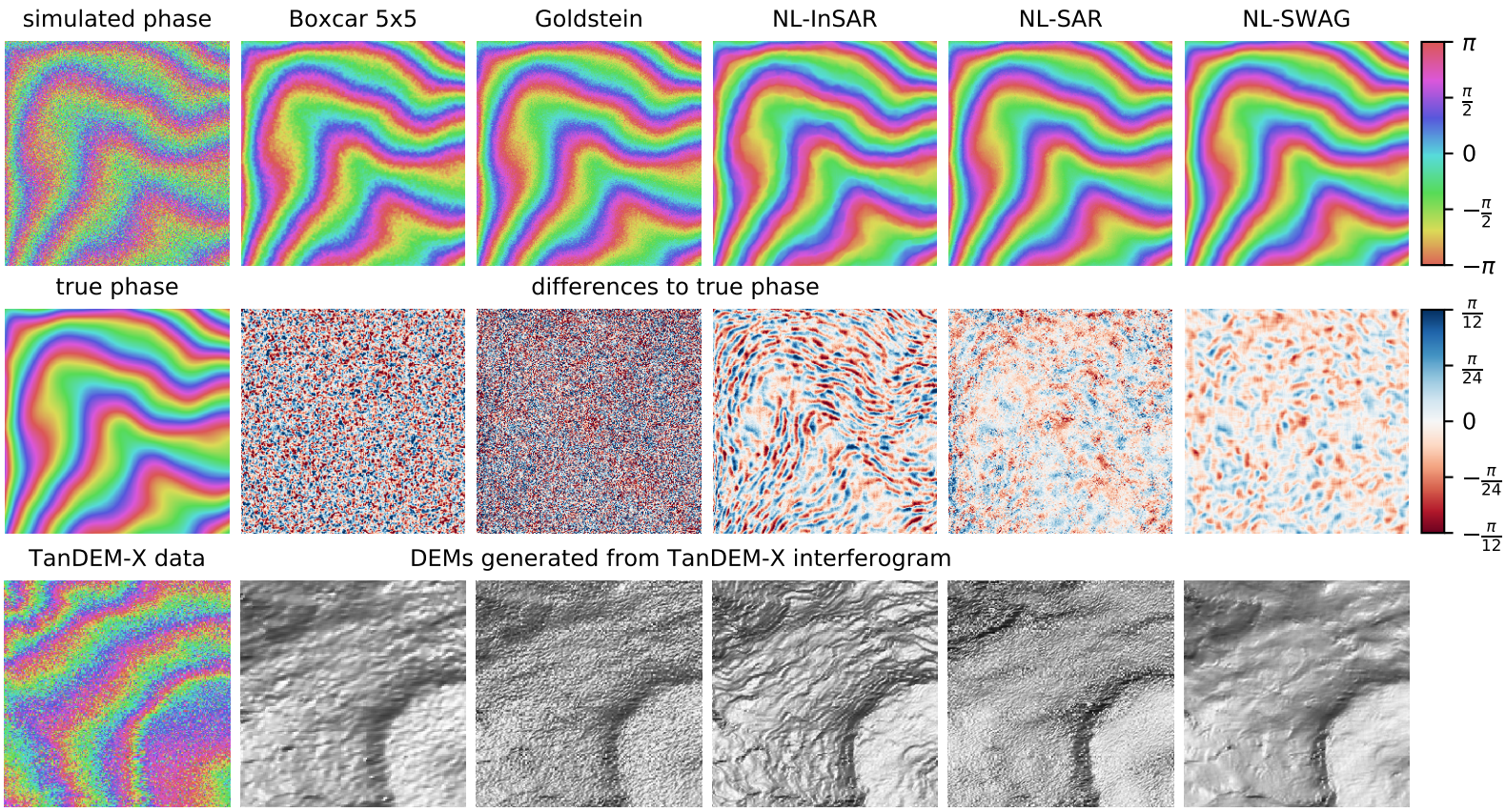


Figure 5.12: Phase estimates of several filters for a synthetically generated interferogram and their differences to the true phase are shown together with the noisy interferogram (the coherence was set to 0.7) and the true phase in the first two rows. The last row shows a comparable TanDEM-X strip map interferogram and shaded relief of DEMs generated from it by the corresponding filter. The phase estimate of NL-InSAR shows a distinct staircase-like pattern, which is also clearly visible in the shaded relief plot. All nonlocal filters suppress the noise's high-frequency component but produce low frequency undulations in the estimate.

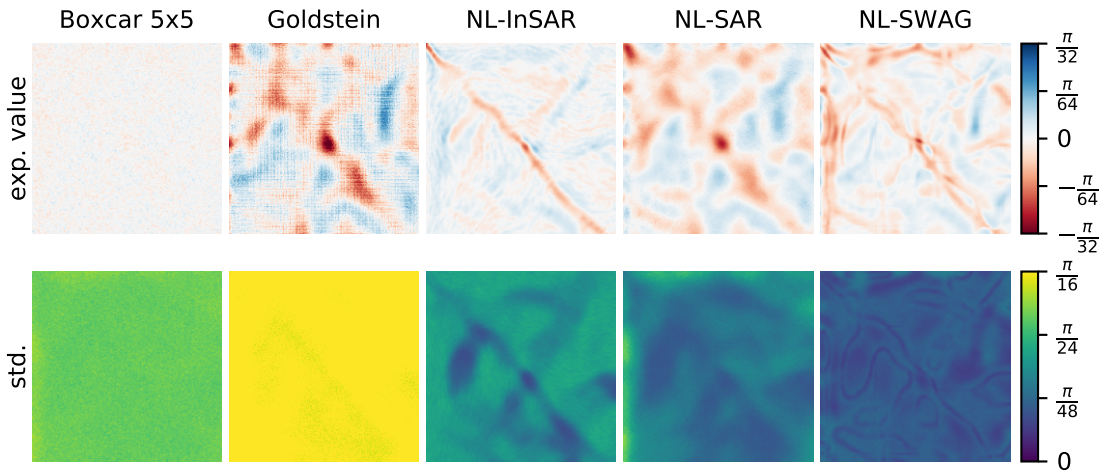


Figure 5.13: Expected values (top) and standard deviation (bottom) for a Monte-Carlo simulation of the simulated phase in Figure 5.12. Minor biases are present in the phase estimates. The slope dependent denoising performance of nonlocal filters is evident in the standard deviation plots.

Table 5.2: Standard deviation in radians and average equivalent number of looks, rounded to the nearest integer, for the Monte-Carlo simulation in Figure 5.13. The gain that NL-SWAG provides over the Boxcar filter approximately match the results in Table 5.1 for the Salar de Uyuni test site.

	Boxcar 5×5	Goldstein	NL-InSAR	NL-SAR	NL-SWAG
$\sigma_{\hat{\varphi}}$ in rad	0.1513	0.2103	0.0990	0.0785	**0.0548**
looks	25	13	58	93	**190**

NL-InSAR and NL-SAR. NL-SWAG is to a far lesser degree affected by this aspect, although it is also not completely immune as noted earlier.

Table 5.2 lists the mean standard deviations and the average equivalent number of looks, rounded to the nearest integer, over the whole image and all simulation runs. In accordance with previous experiments it is considerably lower for nonlocal filters. Contrasting Table 5.2 to Table 5.1 reveals that the simulation results are roughly in line with real data. The nonlocal filters provide an improvement by a factor between two and three in both cases over the boxcar filter. NL-SWAG can take advantage of its fringe frequency compensation routine, and therefore provides the most gain.

NL-SWAG's proposed modifications can rather straightforwardly be applied to existing nonlocal filters. The fringe compensation requires only a minor adaption of the similarity criterion. Changing the patch size adaptively is an isolated modification and could also be performed based on the intensity heterogeneity criterion derived in [Lee81], for example in a nonlocal SAR despeckling filter. At last the aggregation step is an extension of the pixel-wise weighted mean and can be treated separately from all other adjustments.

The next chapter analyzes in greater detail the proposed filter's performance on real TanDEM-X CoSSC data with high-resolution DEM generation as a potential application.

6 High-resolution DEM generation by nonlocal filtering

Digital elevation model (DEM) have numerous applications, such as water flow modeling for flood prevention, line-of-sight sight analysis for mobile cell coverage estimation or autonomous flight planning. Naturally, DEMs of improved resolution, accuracy or coverage can aid these use-cases, which is one of the main reasons why during the past two decades several missions were launched and provided globally available DEMs of increasingly better quality.

The shuttle radar topography mission (SRTM) [Rab+03] used bistatic X- and C-Band InSAR for generating an almost globally available (between −58° and 60° latitude) DEM. Initially, in 2003 only data with a 90 m pixel spacing was made publicly available. But in 2015 the United States government lifted its restrictions and all data of the full 30 m resolution were released to the public. In 2009 a global DEM generated from advanced spaceborne thermal emission and reflection radiometer (ASTER) data [Fuj+05] was made available to the public, covering the earth from −83° to 83° latitude and having 30 m pixel spacing. A global, 5 m high-resolution DEM is commercially available from Japan's second advanced land observing satellite (ALOS-2) [TTT14] since 2016, and a DEM with a reduced resolution of 30 m to the general public.

All these missions demonstrate the demand for DEMs with global coverage and high-resolution, which is also the primary goal of the TanDEM-X InSAR satellite mission. This chapter describes the benefits that the nonlocal filter from Chapter 5 has on the resulting DEM. To lift the secret from the previous chapter behind its name, NL-SWAG is short for **N**on**L**ocal-**S**AR interferogram filter for **w**ell-performing **A**ltitude Map **G**eneration, since it indeed provides an improvement when generating DEMs. A short introduction into the TanDEM-X mission in Section 6.1 describes in brief the mission, its parameters and the processing steps from InSAR acquisition to the final DEM. The advantages that nonlocal filters and the proposed method in particular bring to the table when it comes to generating DEMs from TanDEM-x CoSSCs are listed in Section 6.2. Section 6.3 presents the evaluation and verification using multiple test sites. Finally, Section 6.4 briefly touches on possible post-processing steps to further enhance the generated DEM and Section 6.5 elaborates on the desired goal of generating a global DEM of increased resolution and accuracy from the TanDEM-X mission data using nonlocal filters.

6.1 The TanDEM-X mission

The TanDEM-X twin satellite mission [Kri+07] was conceived by DLR and Airbus Defence and Space, with the primary goal of producing a global DEM of 12 m resolution [Riz+17]. Its two satellites, TerraSAR-X and TanDEM-X, fly in close formation and acquire bistatic SAR

Table 6.1: Resolution and accuracy requirements of the standard global TanDEM-X DEM and the locally available HDEM [Hof+16].

	pixel spacing	absolute horizontal / vertical accuracies	relative vertical accuracy
global DEM	≈ 12 m	< 10 m	2 m (slope ≤ 20 %) 4 m (slope > 20 %)
local HDEM	≈ 6 m	< 10 m	goal: 0.8 m

stripmap interferograms with approximately 3 m by 3 m resolution. Because of this acquisition mode temporal and atmospheric decorrelations are largely not present in the interferograms. To cope with other decorrelation effects, and arrive at the desired height accuracy, spatial multilooking by a Boxcar averaging filter is applied reducing the effective resolution to 12 m by 12 m. The exact relative and absolute accuracies that the global TanDEM-X DEM has to adhere to, according to its specification, are listed in Table 6.1.

The table also lists the specification of the high-resolution DEM (HDEM) with increased horizontal resolution and vertical accuracy, compared to the default TanDEM-X DEM product [LF16]. HDEMs rely on several new acquisitions with larger baselines resulting in smaller height errors from phase noise. For comparison the heights of ambiguity for HDEMs range from 10 m to 20 m, whereas the values for the regular DEMs start at 35 m and go up to 50 m. Thus a boxcar averaging phase filter with a smaller spatial extent compared to the default processing toolchain suffices to fulfill the vertical accuracy goal and more of the original spatial resolution can be retained. As new acquisitions are needed, the HDEM is only available for selected areas.

The bistatic interferograms are processed into raw DEMs by DLR's integrated processor (ITP) [Bre+10; Fri+11; Ros+12], which performs several processing steps. Common spectrum filtering improves the coherence at a slight loss of resolution. Removing the phase's flat earth component makes boxcar filtering more efficient, provides the foundation for a less biased coherence estimate and eases phase unwrapping at a later stage. The just mentioned Boxcar InSAR phase filter provides a phase estimate good enough to meet the accuracy requirements in Table 6.1. Dual-baseline phase unwrapping [LFB14; LFB18] increases the robustness of phase unwrapping. Finally, geocoding is performed. For every region TerraSAR- and TanDEM-X acquire at least two CoSSCs. Difficult terrain, such as forests or mountains, gets up to seven coverages in order to provide an accurate height estimate, prevent phase unwrapping errors, or fill voids. The various raw DEMs of the same area are combined using a weighted average that takes into account the expected heigh error.

The allure of employing a nonlocal filter is that it can simply be plugged in instead of the boxcar averaging filter, requiring no further modifications to the processing chain. The following pages detail whether the proposed nonlocal filter can generate DEMs from the globally available data that adhere to the HDEM specification requirements.

Table 6.2: TanDEM-X stripmap acquisition parameters for the global DEM [Fri+13; Kri+07; Riz+17]

parameter	value
center frequency	9.65 GHz
chirp bandwidth	100 MHz
azimuth resolution	3.3 m
ground range resolution	1.70 m to 3.49 m
swath width	30 km
height of ambiguity (1st coverage)	≈ 50 m
height of ambiguity (2nd coverage)	≈ 35 m

6.2 The case for nonlocal filters

To fulfill the HDEM requirements coming from the global TanDEM-X DEM, a factor of 2.5 in height accuracy, would require, as a rough estimate, $2.5^2 = 6.25$ as many looks, equivalent to a 13×13 boxcar averaging filter. As a reminder Figure 2.5 shows the phase standard deviation, and consequently also the height estimate standard deviation, as a function of coherence and number of looks.

The argument for nonlocal filters is actually well founded. Due to their large search windows they can theoretically reach the required number of looks and their highly discriminate search for similar pixels avoids the smoothing of details. At least in flat terrain, such as the salt flat Salar de Uyuni, nonlocal filters indeed provide the desired improvement (see Table 5.1). The question is whether this advantage of nonlocal filters also exists for more complex terrain.

In addition the design decisions taken for the proposed nonlocal filter in Chapter 5 directly benefit DEM generation with TanDEM-X stripmap data. The following paragraphs provides the appropriate context.

6.2.1 Topographic phase component

As TanDEM-X and TerraSAR-X both acquire their respective images at virtually the same instance the atmospheric and deformation components of the interferometric phase in Equation (2.6) can be neglected, so that only the height components and the superimposed noise remain. With a pixel size of 3 m × 3 m there can be height changes of several meters within the search window of a nonlocal filter. Since the bistatic interferogram's height of ambiguity ranges only from 35 m to 50 m, as reported in Table 6.2, this translates to significant phase changes. Due to this topographic phase component it is considerably harder to detect statistically homogeneous pixels in regions with non-negligible height differences, that is pixels with identical noise distribution but different heights, and the denoising performance suffers. This effect is even more pronounced for bistatic TanDEM-X data due to their generally high coherence (compare Figure 5.4).

As a result of its fringe frequency compensation, introduced in Section 5.4, the proposed nonlocal filter is able to account for this deterministic phase component and ensures proper

denoising of terrain with topography changes.

6.2.2 Scene heterogeneity

Since TanDEM-X is a global mission, by definition its satellites acquire all possible types of terrain from sand dunes, over rain forests to tundra and mountains and many more. These different types of terrain naturally exhibit a vast array of different features and other characteristics for which traditional filters with their fixed filtering strategy are ill-suited for.

By adapting their kernels to the current image region bilateral and nonlocal filters are already applicable to a wider range of terrain types. In addition Section 3.2.1 presented the rationale for further locally adapting filtering parameters. The proposed nonlocal filter also employs this technique, which renders it a feasible choice for globally generating DEMs for any type of terrain.

6.3 Evaluation and verification

The assumption that nonlocal filters provide a gain compared to boxcar filters in terms of height accuracy and resolution needs to be verified. To some extent this was already achieved in the various simulations of Chapter 5. This section extends this analysis to real data and makes the connection to the characteristics that have been observed in the previously described simulations.

Quantifying the height accuracy and resolution of InSAR DEMs is an intricate problem. A comparison to LiDAR is inherently flawed due to the different sensor types. Whereas LiDAR DEMs usually are derived from the last return pulse, the measurements that InSAR provides depend on many factors. Electromagnetic waves at X-Band do not fully penetrate canopies or other types of vegetation so that the resulting phase center is located somewhere in the volume occupied by plants. Conversely, especially at lower frequencies, radar partially penetrates the ground [NF03]. In other words InSAR does not provide a measurement of the ground's elevation but depending on vegetation, soil moisture, wavelength and other factors, either an over- or underestimation. Quantifying the resolution is also not straightforward as abrupt height changes for InSAR result in layover or shadowing.

In the following subsections several experiments try nonetheless to paint a complete picture of the performance gain provided by the proposed nonlocal filter. Table 6.3 lists the type of experiment and what conclusion can be drawn from them.

6.3.1 Shaded reliefs of DEMs

Although this is mostly a visual inspection some of the filtering characteristics can be gleaned from such an analysis. Figure 6.1 shows for a mixed agricultural, industrial and mountainous area, near the French city of Marseille, shaded reliefs of DEMs generated with ITP. For better interpretation the figure also shows an optical image and a plot of the unfiltered phase and amplitude. The same filters with identical parameters as in Section 5.5 replace the standard boxcar filter in ITP. The resolution of the generated DEMs is 6 m for longitude and latitude with the exception of the one produced by the 5×5 Boxcar filter. The reason being that ITP

Table 6.3: Different types of analysis conducted to evaluate the proposed filter. Every experiments provides other insights into the improvements it provides.

type of analysis	observed properties
shaded reliefs of DEMs	general filtering characteristics
difference to LiDAR	filtering bias and systematic differences
shaded reliefs of DEMs for crop fields	resolution
difference to LiDAR without vegetation	noise reduction
height error maps	large scale noise reduction and filtering characteristics

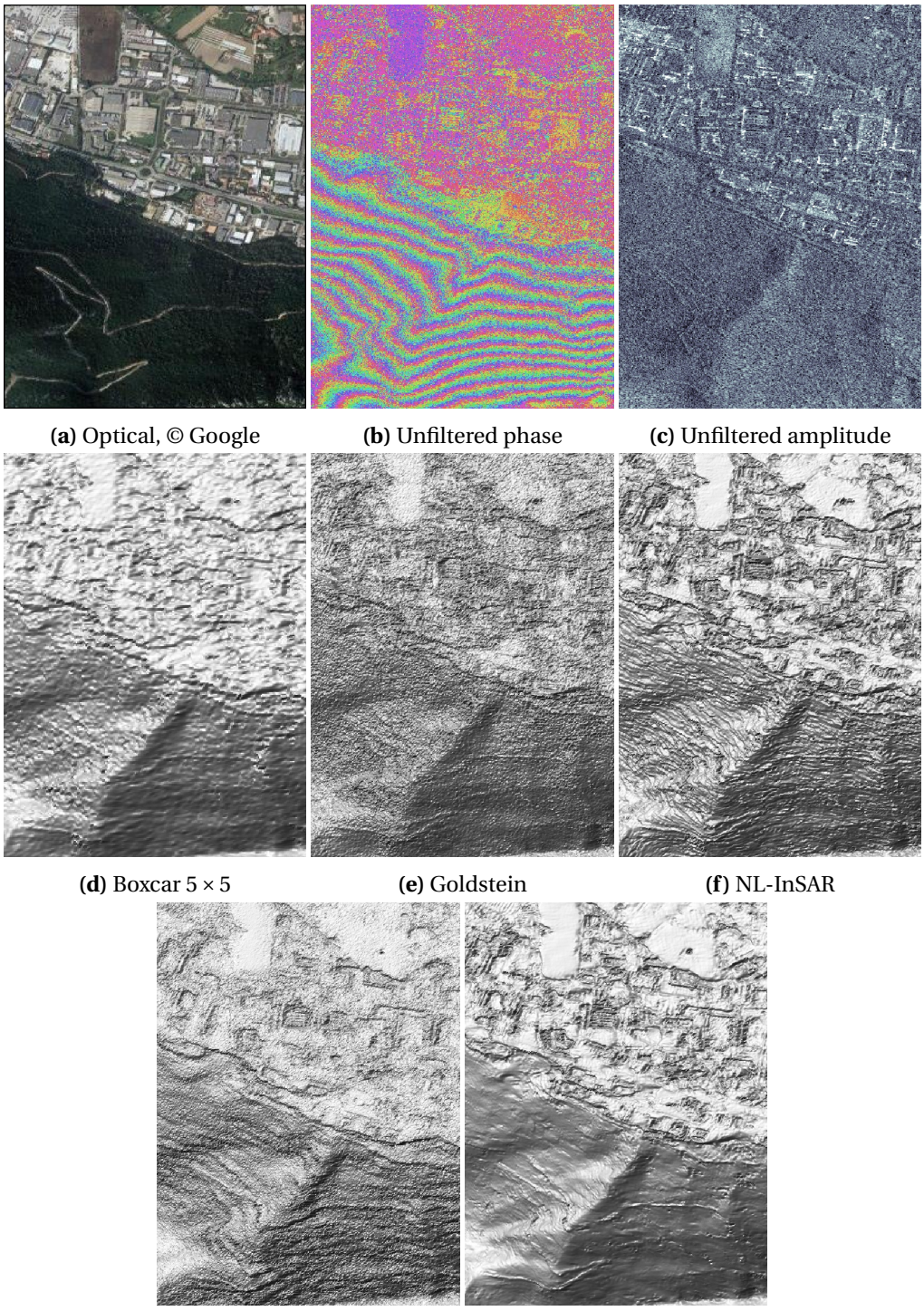
also employs Boxcar filters for generating the global DEM with a resolution of 12 m. Their dimensions depend on range resolution, incidence angle and imaging mode, but for strip map data are close enough to 5×5. Consequently the selected boxcar filter with a 12 m spacing serves as a point of reference for what improvements can be achieved with other filters over the default ITP processing chain. Since the final DEM product is only a weighted average of several raw DEMs the conclusion drawn from the analysis on a single raw DEM is also applicable.

Figure 6.1 confirms lower noise levels and more visible details of all nonlocally generated DEMs compared to the Boxcar and Goldstein filters. The staircasing effect that affects NL-InSAR manifests itself at the hill in the lower half. As a consequence of its fixed patch size NL-InSAR also suffers from the rare patch effect, which materializes as noisy halos near buildings in the industrial site. In this region both NL-SAR and NL-SWAG can produce a less perturbed, cleaner estimate. Admittedly, the shaded reliefs exaggerates some of the remaining noise in NL-SAR's estimate, especially at the hill in the test site's lower half. Yet NL-SAR tends to oversmooth some details, such as the road in the test site's lower part, which is hardly distinguishable from its surrounding. NL-SWAG's estimate generally exhibits low noise over the whole image and the most details.

Figure 6.2 sheds some more light on NL-SWAG's filtering characteristics by showing the Gaussian window width used for computing the patch similarities and the final equivalent number of looks after the aggregation step. Both show that homogeneous areas benefit from wide Gaussian windows, resulting in accurate patch similarity estimates, and a large number of similar pixels within the search window, leading to low-noise estimates. The reverse is true for the industrial site, where narrow Gaussian windows are employed, due to the regions heterogeneity, resulting in a comparatively low number of looks. The impact that the fringe frequency estimation and compensation has on the estimate can be inferred from the equivalent number of looks for the hilly terrain to the south, which is virtually unaffected by the trend of the phase.

6.3.2 Difference to LiDAR

Generating DEMs and examining their deviations to a LiDAR DEM reveals the earlier described, systematic differences between InSAR and LiDAR, but can also be used to judge the

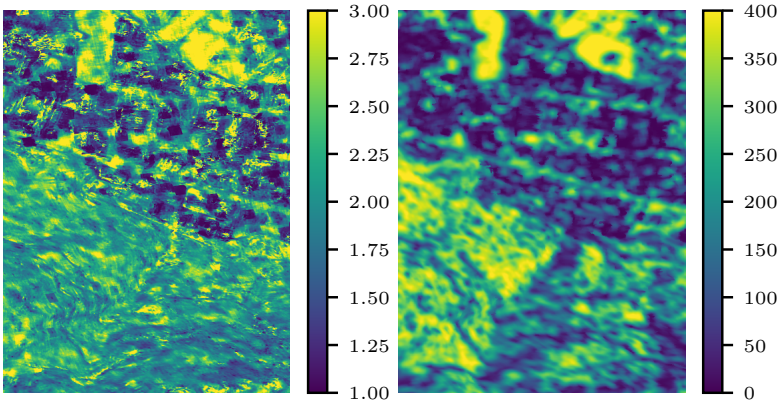


(a) Optical, © Google **(b)** Unfiltered phase **(c)** Unfiltered amplitude

(d) Boxcar 5×5 **(e)** Goldstein **(f)** NL-InSAR

(g) NL-SAR **(h)** NL-SWAG

Figure 6.1: Shaded reliefs of DEMs generated with the various InSAR filters under analysis.



(a) Sigma of Gaussian windows **(b)** Equivalent number of looks

Figure 6.2: Width of the Gaussian windows used for computing the patch similarities and the equivalent number of looks for the test site from Figure 6.1.

noise reduction and the introduction of filtering artifacts if all limitations are taken properly into account. The test site is a hamlet surrounded by hilly terrain near the German city of Munich. Its CoSSC was acquired on August 19th, 2011. The reference LiDAR DTM, whose purpose is to more closely evaluate the qualities of each filter, was acquired in the months of March and April in 2012 and was generously supplied by the honorable Bayerische Vermessungsverwaltung. With a 1 m by 1 m spacing and height errors below 20cm its accuracy and resolution far surpass the requirements for the InSAR DEM, justifying its use as a reference. For the LiDAR DTM man-made structures and the effects of vegetation, like agricultural fields and forests, were removed.

Figure 6.3 shows in the top row an optical image, the LiDAR digital terrain model and the unfiltered amplitude and interferometric phase. The other rows show the filters' generated DEMs and their height difference to the LiDAR digital terrain model.

For the village in the image center and the forested area to the right large height differences to the LiDAR DTM are observable, as the X-band radar of TanDEM-X is reflected from roofs and only partially penetrates the canopies of trees. This behaviour shows the difficulty of a direct comparison of LiDAR to InSAR without considering the respective sensor characteristics and DEM generation routines. The same phenomenon occurs for some of the agricultural fields, whose outline can be deduced from the optical image. X-Band radar only marginally penetrates even low vegetation [RE15] and the measured phase scattering center will have a different height than the LiDAR DTM, where the effects of vegetation have been removed. Consequently as the CoSSC was acquired in August when some of the crops were in full blossom, these areas will exhibit systematic height differences.

Yet a more nuanced comparison is still possible. The height difference maps clearly show a higher variance of the height estimate for the Goldstein and the Boxcar filters in comparison to the nonlocal filters. NL-InSAR's estimate features very pronounced terraces, which to a far lesser extent are also visible in NL-SWAG's estimate. However some of the terraces can also be discovered in the Boxcar filter's estimate, which shows that they are not entirely

artificially created filtering artifacts. NL-SAR overall provides the best result, yet admittedly for this particular test site its tendency for smooth solutions works in its favor.

6.3.3 Shaded reliefs of DEMs for crop fields

In the previous example of Figure 6.3 the outlines of crop fields are visible as a result of the phase scattering center of electromagnetic waves at X-Band. This behaviour can actually be beneficial to get an impression of the generated DEM's resolutions. Figure 6.4 shows DEMs for another test site, an agricultural area near Munich, Germany. The resolution is the same as in the previous example: 6 m for the Goldstein and nonlocal DEMs and 12 m for the Boxcar filter. The CoSSC is the same as in Section 6.3.2 and since some of the fields had already been harvested and as a result of the small penetration depth at X-Band height differences are visible. To see the trade-off of nonlocal filters between detail preservation and noise reduction NL-SWAG uses several values for h.

The shaded reliefs confirm the simulation results in Figure 5.7, Figure 5.8 and Figure 5.9 that NL-InSAR provides the best result for this particular scenario, as it favors piecewise constant solutions and sharp edges. But this propensity is also the source of the highly unwelcome staircasing for regions with a more interesting topographic profile. The effect that a change of h has on the filtering result is illustrated by Figure 6.4f to Figure 6.4i. A lower value of h produces sharper transitions at the fields' edges but leads to lower noise reduction in flat terrain. NL-SAR's DEM exhibits oversmoothing compared to the other DEMs.

6.3.4 Difference to LiDAR without vegetation

Section 6.3.2 showed systematic differences between DEMs derived from LiDAR and InSAR data, requiring that areas exhibiting such behaviour are excluded from the analysis. This experiment again compares DEMs generated with NL-SWAG and a 5×5 boxcar filter to a high-resolution LiDAR DEM, which serves as a gold standard. The goal is twofold:

1. Quantify what level of improvement in noise reduction can be achieved,
2. and give an impression of how many more details are preserved with the proposed nonlocal filter.

The test site is the town Terrassa close to Barcelona in Spain. The top row in Figure 6.5 shows an optical image from Google maps and the LiDAR DEM with 5 m spacing, plus DEMs generated from a single CoSSC by the respective filters, which are resampled to the same grid as the LiDAR DEM.

To remove areas with systematic errors, such as urban areas suffering from layover and shadowing or vegetation, the LiDAR DEM is compared to the final global TanDEM-X DEM. Regions with a height difference larger than 2 m are excluded. The result is depicted in the bottom row of Figure 6.5 and a cleaned mask, using morphological operations, right next to it. The height differences are the remaining two pictures annotated with the standard deviation of the height difference computed over the masked area.

Concerning the experiment's first point to provide a quantification of the noise reduction on the not masked-out, moderately hilly, homogeneous terrain NL-SWAG improves the noise

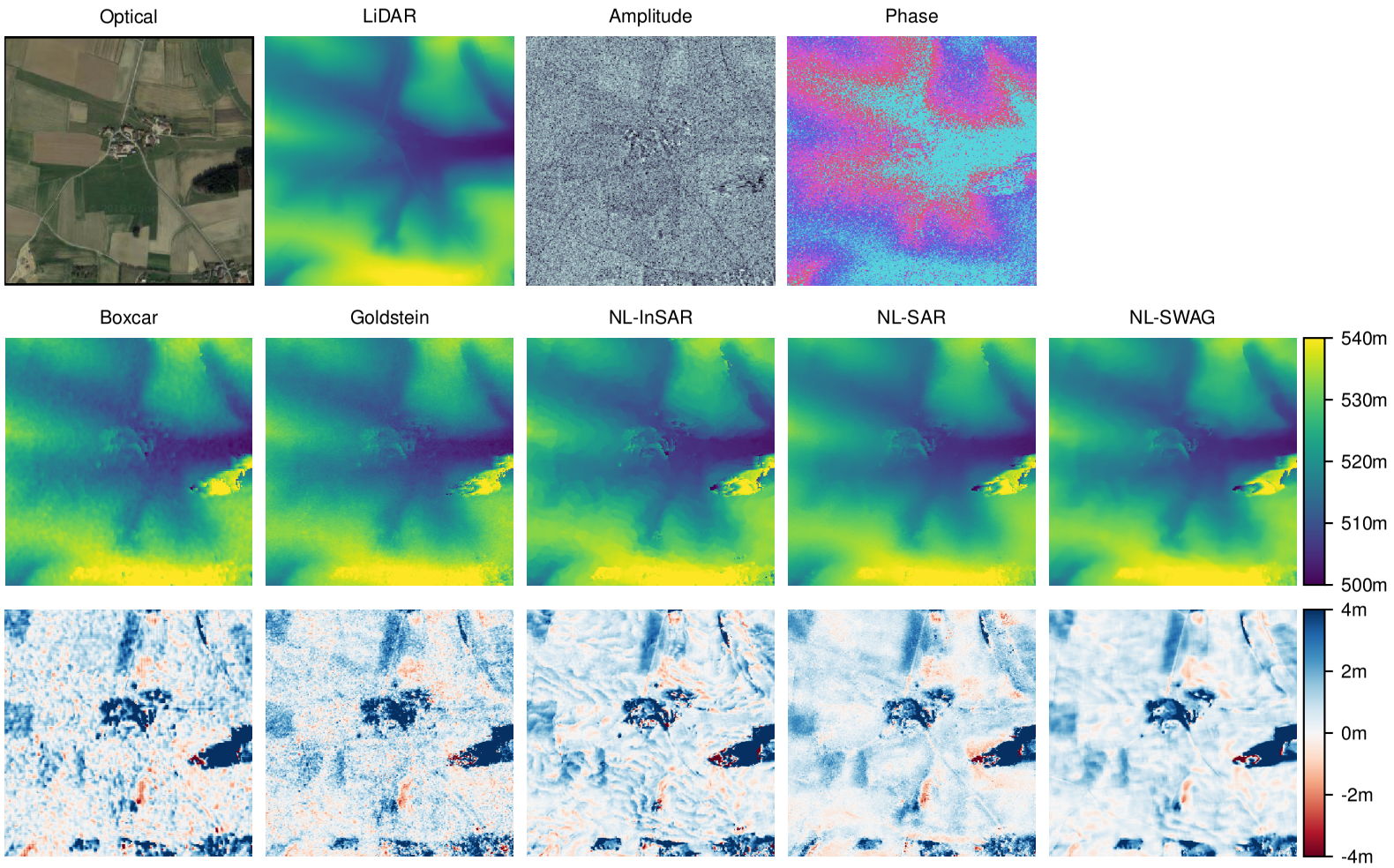
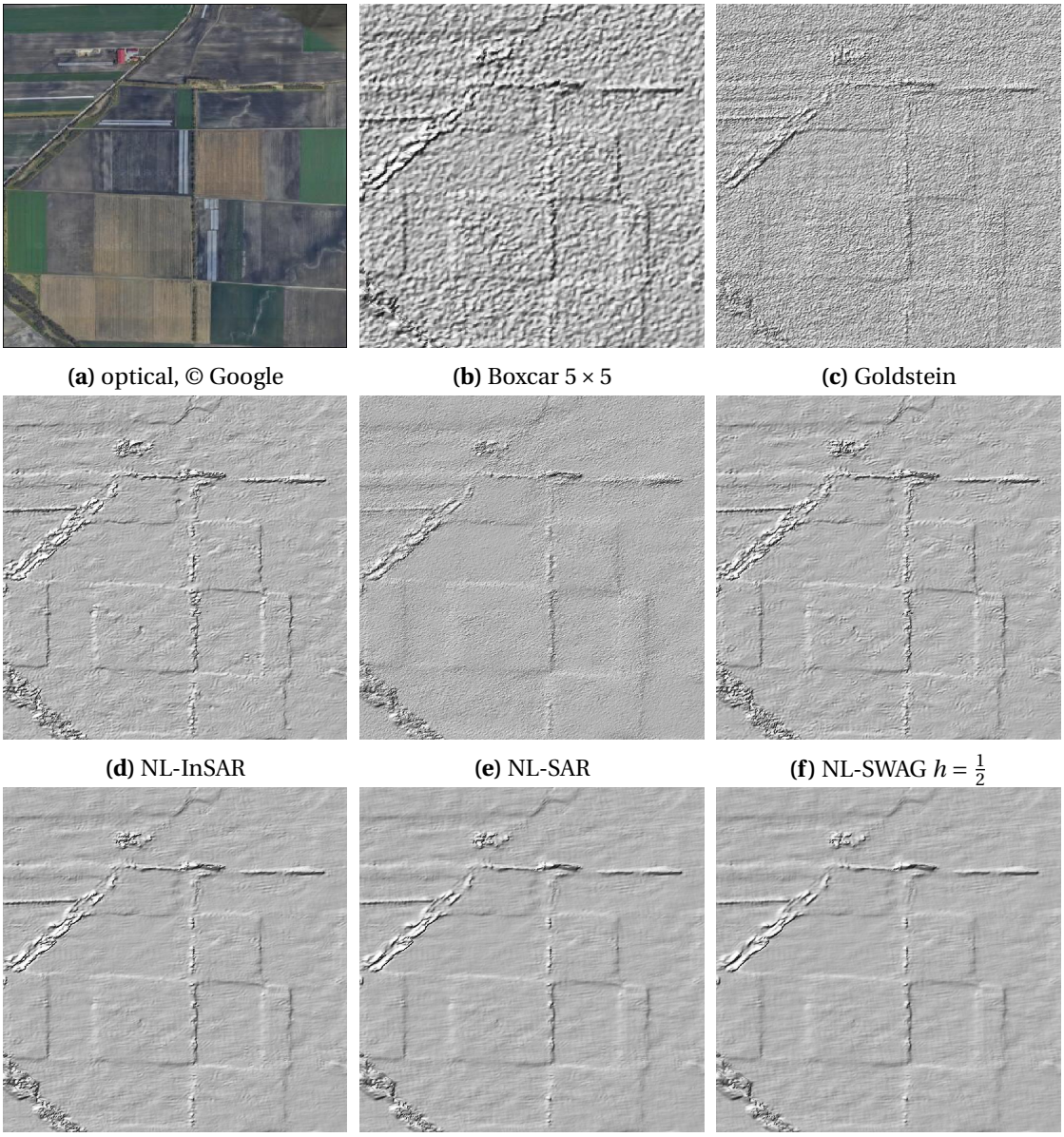


Figure 6.3: Rural area near Munich, Germany. Top row: optical image, LiDAR DTM and SAR interferogram. Second and third row: DEMs generated with the different denoising filters and their respective height difference to the LiDAR DTM. Again NL-InSAR's difference map shows distinct terraces. NL-SAR provides a result of low variance but tends to oversmooth small details such as single trees.



(a) optical, © Google **(b)** Boxcar 5×5 **(c)** Goldstein

(d) NL-InSAR **(e)** NL-SAR **(f)** NL-SWAG $h = \frac{1}{2}$

(g) NL-SWAG $h = 1$ **(h)** NL-SWAG $h = 1.5$ **(i)** NL-SWAG $h = 2$

Figure 6.4: Shaded reliefs of DEMs of an agricultural site. Clearly visible are height changes between fields. The bottom row shows the effect that changing h in the second stage of NL-SWAG has on detail preservation and noise reduction.

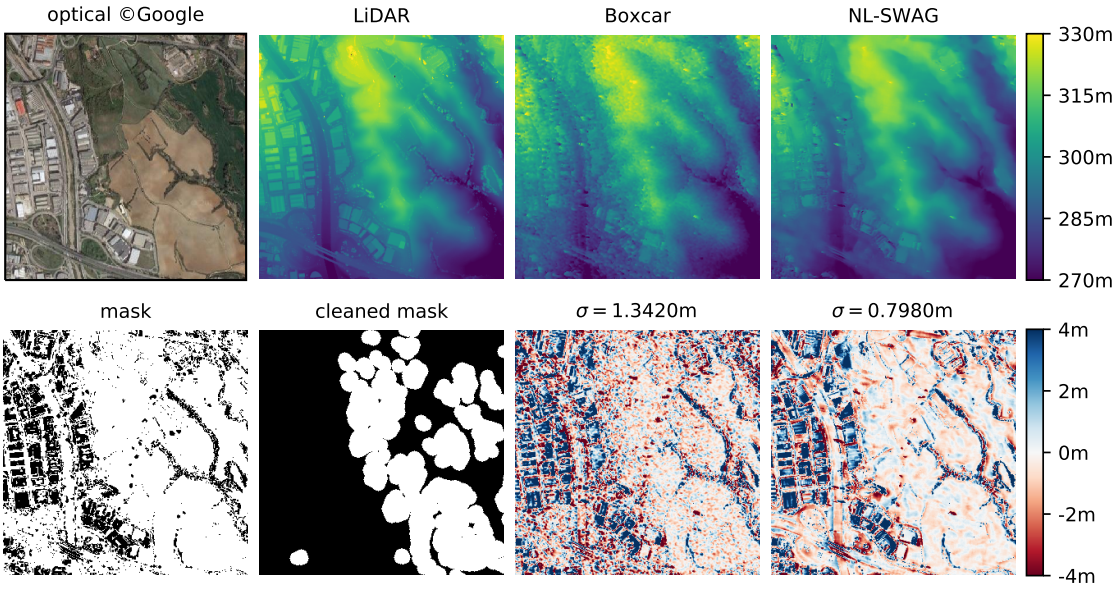


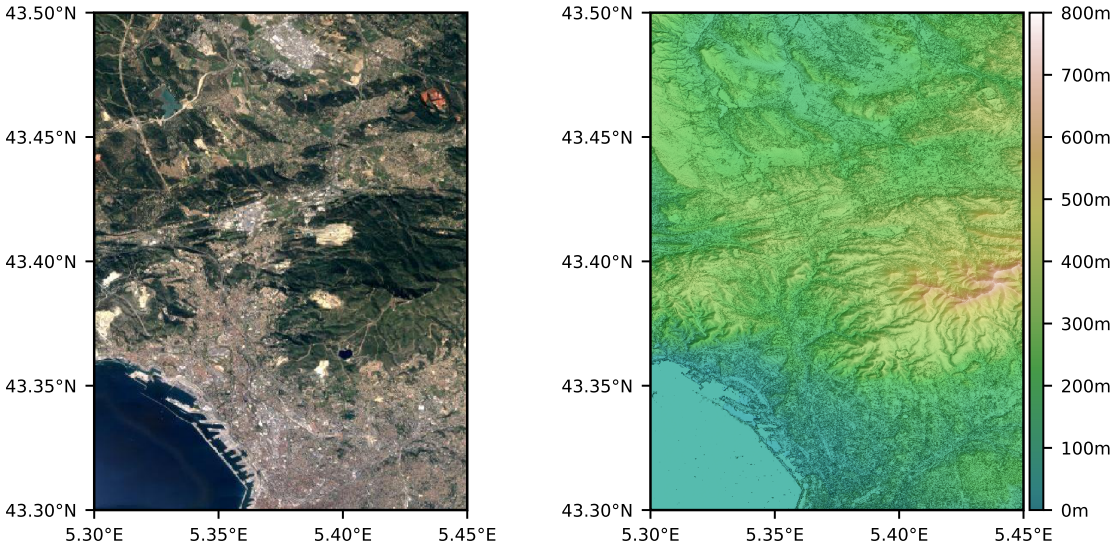
Figure 6.5: DEMs generated by NL-SWAG and a 5×5 Boxcar filter from a TanDEM-X interferogram are compared to a LiDAR DEM. The bottom row shows the height differences to the LiDAR DEM. For the masked out area the height differences' standard deviations are computed.

level roughly by a factor of $1.3420\,\text{m}/0.7980\,\text{m} \approx 1.6817$. At first glance, this improvement in noise reduction contradicts the findings reported in Table 5.2, where NL-SWAG provides an improvement by a factor of 2.7 on the simulated interferogram; or the results for the Salar de Uyuni test site with an improvement of about three. Systematic height differences due to the different physical properties of LiDAR and SAR can be excluded as the penetration depth of electromagnetic waves at X-Band is negligible as an error source [NF03]. Coregistration errors of the LiDAR and InSAR DEMs might also be a contributing factor for height differences but for the moderately hilly terrain they would only play a minor role. Such error sources would equally increase the difference to the LiDAR DEM, leading to a misrepresented noise level reduction. The true reason for this discrepancy is the resampling from approximately 3 m pixel spacing in range and azimuth to the 5 m LiDAR pixel spacing, which essentially increases the footprint of the Boxcar filter. For NL-SWAG this effect is imperceptible due to its comparatively large search window.

Studying the industrial site in the leftmost part shows that more buildings are discernible in the DEM generated by the nonlocal filter. Although due to layover and shadowing large height errors at the building's footprints can be observed which make a quantitative analysis impossible.

6.3.5 Height error maps

The official TanDEM-X DEM product specification document [Hof+16] defines the height error map (HEM) to give customers an impression of the quality of the delivered DEM. HEMs are computed by calculating for each pixel its phase standard deviation following Equa-



(a) Optical, © Google **(b)** Digital elevation model

Figure 6.6: Optical image and DEM for the city of Marseille at the Mediterranean Sea. The corresponding height error maps are depicted in Figure 6.8.

tion (2.15) using a 11 × 11 boxcar coherence estimate and converting it to a height standard deviation by multiplication with the height of ambiguity [Kri+07]

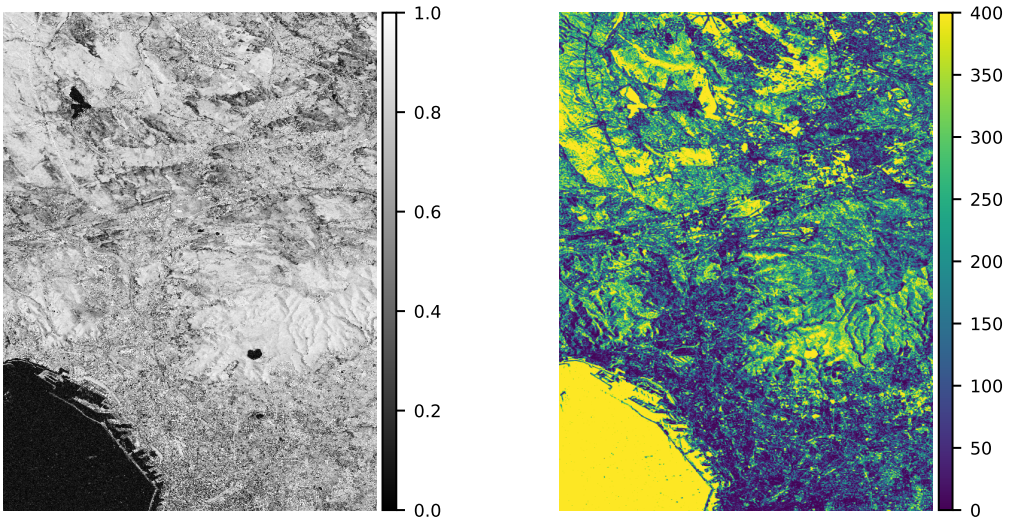
$$\sigma_h = h_{\text{amb}} \frac{\sigma_\varphi}{2\pi}, \tag{6.1}$$

where h_{amb} is the height of ambiguity provided by ITP, which slowly increases with range. As the delivered DEM is a weighted mean of several raw DEMs the final HEM is computed in an identical fashion from the raw DEMs' corresponding HEMs.

The following paragraphs compare the HEMs derived for a 5 × 5 phase boxcar filter and NL-SWAG from a single CoSSC, i.e., raw DEM HEMs. Figure 6.6 shows an optical image and a DEM of the test site of which the HEMs are computed. The test site covers the area around Marseille at the Mediterranean Sea and is a nice mix of urban, rural, agricultural and mountainous terrain.

The computation of the HEM slightly differs from the one in [Hof+16] as both HEMs are computed using the coherence estimate provided by NL-SWAG. Using Equation (2.15) the interferometric phase's PDF is defined for every pixel, where Equation (5.12) together with Equation (5.1) and Equation (5.3) provides the number of looks of the nonlocal filter and the boxcar filter obviously has 25 looks. Admittedly the nonlocal filter will slightly overestimate the coherence [Tou+99] and also the number of looks L, as Equation (5.12) assumes that all pixels are independent, which is not the case due to antenna sidelobes. Still, since both HEMs are computed using the same coherence estimate and the boxcar filter also suffers from correlated pixels one can juxtapose their HEMs.

To better understand the final HEMs Figure 6.7 depicts NL-SWAG's coherence estimate and its effective number of looks. The coherence has the expected behaviour of almost zero co-



(a) Coherence

(b) Effective number of looks

Figure 6.7: NL-SWAG's coherence estimate and effective number of looks for the test site in Figure 6.6. For homogeneous areas NL-SWAG reaches a high number of looks, whereas the low numbers for Marseille itself indicate fine details are preserved.

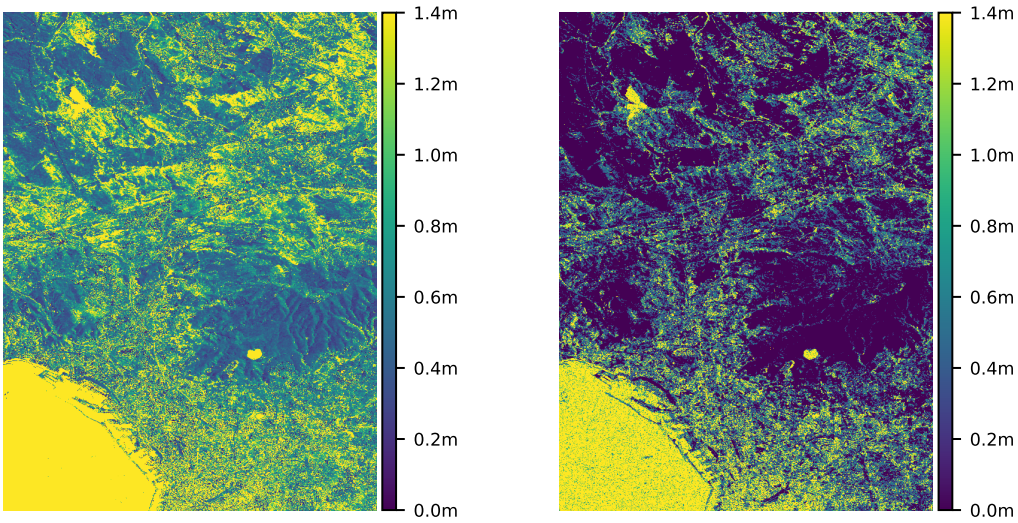
herence for water bodies, a mix of high and low coherence for urban areas due to bright scatterers but also shadowing, and comparatively high levels for other terrain due to the bistatic acquisition mode. The effective number of looks also has few surprises, heterogeneous areas such as Marseille itself have rather few looks, whereas for the sea or agricultural fields the search within the search window yields several hundred similar pixels. Also visible is that for the mountain in the test site's center (compare Figure 6.6b) the number of looks does not drop significantly as a result of NL-SWAG's fringe frequency compensation.

The resulting HEMs of both filters in Figure 6.8 show as expected that InSAR derived DEMs exhibit large errors in urban areas due to shadowing. The same holds of course for water bodies due to low coherence. For other types of terrain the DEMs provide reliable height estimates, but due to its higher number of looks NL-SWAG's HEM shows a distinctly lower height error. This is in line with other experiments, as for example Figure 6.5.

6.4 DEM void filling

The HEMs shown in Figure 6.8 indicate that for example all water bodies suffer from non-negligible height errors. This is actually accounted for in the TanDEM-X DEM generation process by detecting water bodies through joint analysis of the amplitude and coherence and consequently flattening the DEM at these locations [Wen+13]. Similar arguments can be made for layover or generally low coherence areas.

In fact these artifacts can be addressed by regularization, i.e., incorporating prior knowledge. The earlier introduced TV regularization can lead to staircasing artifacts, as it favors



(a) HEM Boxcar **(b)** HEM NL-SWAG

Figure 6.8: Height error maps corresponding to the DEMs in Figure 6.6. Especially in homogeneous areas NL-SWAG produces a more accurate DEM.

piecewise constant solutions. A very undesirable property when generating DEMs, since it results in unnatural terrace-like artifacts. Total generalized variation (TGV) [BKP10] circumvents this problem by also considering higher order derivatives in the regularization term and can be simply used in lieu of TV. The second order TGV

$$\mathrm{TGV}_\alpha^2(u) = \min_v \alpha_1 \int_\Omega |\nabla u - v| \mathrm{d}\, x + \frac{\alpha_0}{2} \int_\Omega |\nabla v + \nabla v^T| \mathrm{d}\, x \tag{6.2}$$

already rids the result of any staircasing artifacts.

The approach proposed in [SDA14], which combines nonlocal filtering with TV regularization, can readily be extended with TGV. The following optimization problem lends itself to regularize the DEM H

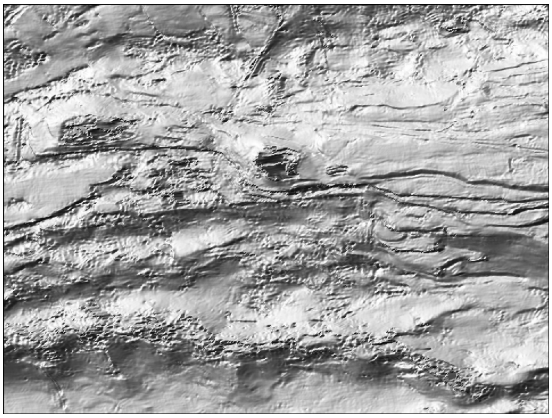
$$\operatorname*{argmin}_H \sum_{\mathbf{x}\in\Omega} \lambda_{\mathbf{x}} \left(H_{\mathbf{x}} - H_{\mathbf{x}}^{\mathrm{NL}}\right)^2 + \mathrm{TGV}(H), \tag{6.3}$$

where H^{NL} is the nonlocal estimate and the locally adaptive regularization number λ controls the balance between the nonlocal estimate and the regularization term.

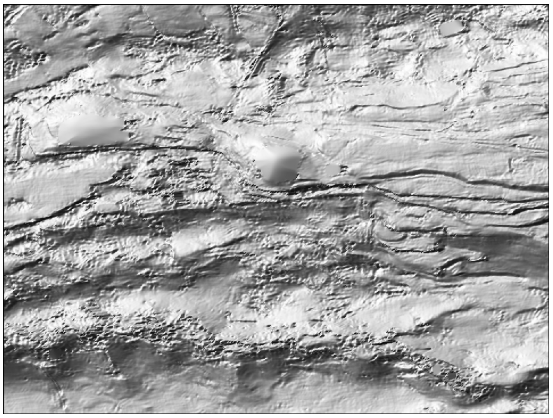
Good choices for the regularization parameter can be the inverse of the HEM or for example the aforementioned water indication masks. Figure 6.9 shows such a simple example, where TGV flattens the perturbation caused by random phase unwrapping results on water.

Solving Equation (6.3) can be realized by the primal-dual algorithm of Chambolle and Pock [CP11], which can be efficiently implemented on central processing units (CPUs) or graphics processing units (GPUs) [Kno+11].

(a) Optical, © Google



(b) Original DEM



(c) DEM with inpainted water bodies

Figure 6.9: DEM void filling by total generalized variation. Waterbodies are identified using the coherence and subsequently inpainted. This method could also be extended to terrain affected by shadowing.

6.5 Discussion

The initial goal of this chapter was to ascertain whether nonlocal filters are suitable for generating a DEM close to the HDEM standard (see Table 6.1) from the globally available TanDEM-X data. The following paragraphs detail how NL-SWAG holds up to these challenges.

All conducted experiments confirm that nonlocal filters are able to deliver a vastly improved noise reduction over the exemplary conventional Boxcar or Goldstein filters. The reason is that due to their large search windows nonlocal filters find a multitude of pixels for the averaging process, even for comparatively heterogeneous terrain. To further quantify this improvement for the real data set of Figure 6.5 on moderately complex terrain the height deviation to a LiDAR DTM is reduced approximately by a factor of 1.7. Relating this to the level of noise reduction aimed for in Table 6.1 the proposed filter falls short of reaching the target of 2.5 roughly by a factor of $\sqrt{2}$. Dependent on the type of terrain this might still be sufficient to obtain a DEM that fulfills the requirements of the HDEM, as already the globally available TanDEM-X DEM often overfulfills its accuracy requirements. In any case having twice as many acquisitions available would also satisfy the specification.

The proposed filter implements several techniques to reach this level of noise reduction. It reduces the detrimental effect of topography by its fringe frequency compensation accounting for the deterministic topographic phase component, which is confirmed to be working by Figure 6.2 and Figure 6.7. Furthermore even on flat, homogeneous terrain the high inherent noise level of InSAR can hamper denoising, which is countered by the two-step approach.

The analysis whose results are reported in Figure 5.13, which was conducted in the last chapter, shows that nonlocal filters bias the estimate for nonlinear phase profiles. The bias is limited by approximately $\pm\frac{\pi}{100}$. With a height of ambiguity of 40 m, which is a typical value for TanDEM-X interferograms, this translates to deterministic height errors of $\pm$ 20 cm, well within the HDEM specifications.

The HEMs in Figure 6.8 highlight that for nonlocal filters it is far easier to denoise homogeneous terrain than heterogeneous targets, as more similar pixels are found. Nonetheless nonlocal filters are well suited for preserving heterogeneous targets as shown by last chapter's simulation results in Figure 5.7, but also the shaded reliefs in Figure 6.1, where the adaptive patch size and the aggregation step play a significant role to avoid the rare patch effect near the edge.

For filtering SAR interferograms of urban areas or terrain with man-made structures nonlocal filters are particularly appealing as such heterogeneous targets exhibit a very high radar cross section compared to their surrounding. These high intensity variations aid nonlocal filters to preserve details, as they render the resulting weight maps more discriminant. The gain in resolution, compared to simple boxcar averaging, is again evident for real data in Figure 6.1 or Figure 6.5.

In conclusion, employing the proposed nonlocal filters permits to move up approximately one level in the quality criteria defined in the high resolution elevation product specifications [HRE].

7 Multibaseline nonlocal filtering

The previous two chapters proposed a single baseline nonlocal InSAR filter and highlighted its effectiveness by generating DEMs of higher resolution and accuracy than with conventional filters. Section 4.2 already introduced several multibaseline nonlocal InSAR filters and showed that the advantages of nonlocal filtering also transfer to time series. Other sources of inspiration are nonlocal video denoising filters, which share the existence of an additional dimension, the temporal one, with multibaseline methods.

Indeed several nonlocal algorithms for video denoising exist. [DFE07] is an adaption of BM3D [Dab+07] to videos and simply increases the search window to also include previous and later frames. The original nonlocal filter can also straightforwardly be extended in the same manner [BCM08]. Actual 3d patches that also include the temporal domain are used by [PE09] together with a dictionary to denoise videos. The dictionary is continuously updated to adapt to changing scenes. The approach proposed in [Ji+10] assumes that the collection of patches have low rank. Single pixels with outliers are removed improving the robustness of the denoising algorithm. These void pixel locations are then filled via matrix completion. [Mag+12] uses 4d sparsifying transforms instead of the 3d ones in BM3D for regular optical images and also employs 3d instead of 2d patches. Introducing motion estimation in the filtering process is done by [MSF14] where a sequence of 2d patches is denoised using sparsifying transforms.

What sets these filters apart on a conceptually level from the nonlocal multibaseline filters introduced in Section 4.2 is their special treatment of the temporal dimension. Either by considering outliers or gradual and sudden changes in the scene. The filtering framework that this chapter proposes shares this core idea.

The remainder of this chapter is organized as follows. Section 7.1 introduces several figures of merit to evaluate stack filtering algorithms. As a precondition for a meaningful weighted average Section 7.2 analyzes different similarity criteria. The description of a general multibaseline filtering framework finds itself in Section 7.3. For some straightforward filters within this framework Section 7.4 reports the evaluation's results with the previously described figures of merits and on a real data stack.

7.1 Measuring the effectiveness of stack filtering

Judging the effectiveness of any spatial filter is already an intricate problem, the introduction of the temporal dimension makes this endeavour even more challenging. The following figures of merit can give an impression of the performance of multibaseline InSAR filtering methods. The purpose of starting with the evaluation criteria, even though admittedly unconventional, is to reflect on what behaviour to expect from a SAR image stack, how this

affects the design of an appropriate filter and at last how to quantify the effectiveness of the undertaken measures. In later experiments these figures of merits will, in conjunction with simulated InSAR stacks with different features, assess the noise reduction, the introduction of filtering artifacts and other undesired by-products.

7.1.1 Phase noise reduction

Large stacks of SAR images are mostly the result of repeated spaceborne acquisitions such that temporal decorrelation dominates the overall coherence loss. Models for SAR coherence loss usually assume an exponential decay, possibly with some residual coherence [PCE09]. Going back to Figure 2.5, which plots the relation between coherence loss and the resulting phase standard deviation, this leads to high phase standard deviations for large temporal baselines.

Evidently the most straightforward way to judge any denoising filter is to measure the reduction in phase noise it provides for a stack, ideally by one single quantity. To prevent the noise of low coherence interferograms to dominate the root mean squared logarithmic error (RMSLE)

$$\mathrm{RMSLE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(\log\left(\arg\left\{e^{\mathrm{j}(\hat{\varphi}-\varphi)}\right\}+\pi\right)-\log(\pi)\right)^2} \tag{7.1}$$

is used as a figure of merit in lieu of the regular root mean squared error. In Equation (7.1) $\hat{\varphi}$ and φ represent the estimated and true phase.

In case of a completely homogeneous stack evidently multilooking with an as large as possible footprint gives the best result, yet with the same draw backs as in the single baseline case. Moving to test data sets with more sophisticated phase and amplitude patterns can give further insights into the filters' noise reduction capabilities. The following subsections introduce other figures of merit which can assess filters under such circumstances.

7.1.2 Point target preservation

Point target preservation of InSAR filters is an important factor for advanced stack processing techniques such as PSI [FPR01], which rely on time series of extremely coherent scatterers. Smoothing of such targets would lower the quality of the estimated phase history resulting in a lower estimation accuracy of the height and deformation parameters. Simulated stacks with persistent scatterers provide one way to gauge the filters' capabilities to preserve point targets.

The error vector magnitude (EVM) [Sch+12] of a persistent scatterer's estimated phase history indicates how much of the true phase information is lost by filtering. EVM is a widely used metric in communications technology to asses the quality of a communication channel. It is defined as

$$\mathrm{EVM} = \sqrt{\frac{\sum_k |\hat{s}_k - s_k|^2}{\sum_k |s_k|^2}}, \tag{7.2}$$

where $\hat{s}_k$ is the received signal, distorted by various error sources, and s_k the true signal. For assessing the detail preservation of any SAR stack filter, the estimated phasor and the true phasor of a persistent scatterer would be used instead.

7.1.3 Robustness to temporal outliers

Processing InSAR stacks can prove difficult by unmodelled phase contributions or the presence of outliers [WZ16]. With regard to the nonlocal means filter ideally an outlier does neither affect other pixels nor their search for similar pixels, which can be hampered if an outlier is contained in their respective patches. This already hints at separate filtering procedure for the outlier itself and the stack excluding the outlier.

The susceptibility to temporal outliers can to some extent be measured by the equivalent number of looks (ENL). Excluding the outlier from a temporal vector in the search for similar vectors results in a high ENL. Conversely a low ENL for the image of the stack which contains the outlier signifies that the outlier does not influence the weighted mean of other pixels. The ratio of the minimum and maximum in a time series of the ENL

$$\frac{\min_t \mathrm{ENL}_{t,\mathbf{x}}}{\max_t \mathrm{ENL}_{t,\mathbf{x}}} \tag{7.3}$$

can thus serve as a measure for the filters' robustness to outliers.

7.2 Stack similarity measures

Although Equation (4.8) measures the similarity of Wishart distributed covariance matrices and is therefore applicable to InSAR stacks under the fully developed speckle assumption, it requires an initial estimate of the covariance matrices, rendering it unusable for large stacks. A commonly used statistical tests for detecting amplitude vectors of similar characteristics is the Kolmogorov-Smirnov test employed in [Fer+11] for adaptive multilooking. For a comparison of several amplitude statistical tests [PB11] can be consulted. GLRT based similarity criteria can be derived for Gamma distributed samples, as in [Con+03]. With X and Y being gamma distributed random variables as in Equation (2.18) and n, m their respective number of looks, their GLRT is

$$Q = \frac{(n+m)^{(n+m)}}{n^n m^m} \frac{X^n Y^m}{(X+Y)^{n+m}}, \tag{7.4}$$

which is the one-dimensional case of Equation (4.8). The temporal dimension can be incorporated by either averaging all GLRTs along the temporal axis or by temporal multilooking.

The idea for a new similarity criterion is to try to take the sample correlation into account. Staying with the GLRT approach the remainder of the section derives a similarity criterion based on correlated Rayleigh samples, the underlying distribution of amplitude measurements assuming fully developed speckle. Another approach is to stay with the Gamma GLRT but estimate the sample correlation and adjust the effective number of looks.

The aforementioned Kolmogorov-Smirnov test and the Gamma distribution based GLRT similarity criteria will serve as a baseline for evaluating these stack similarity measures. All of the similarity criteria have the pleasant property that they lie in the interval $[0, 1]$.

7.2.1 Multivariate Rayleigh generalized likelihood-ratio test

In [Mal03] Mallik derived the PDF of a multivariate Rayleigh random variable and some of its special cases. In particular for the case of exponential coherence loss between the observations, that is the covariance matrix $\mathbf{C}$ of the underlying circular complex Gaussian has the form

$$\boldsymbol{C}_{ij} = \begin{cases} \sigma^2 & \text{for } i = j \\ \rho^{i-j} \mathrm{e}^{\mathrm{j}\varphi_{i,j}} \sigma^2 & \text{for } i > j \\ \rho^{j-i} \mathrm{e}^{-\mathrm{j}\varphi_{i,j}} \sigma^2 & \text{for } j > i, \end{cases} \tag{7.5}$$

where $\gamma = \rho \mathrm{e}^{\mathrm{j}\varphi_{i,j}}$ is the complex coherence and σ^2 the intensity, the log likelihood of the PDF is given by [Mal03]

$$\begin{aligned} \ln\left(f(\mathbf{a};\sigma,\rho)\right) = &\sum_{i=1}^{L} \ln(a_i) - 2L\ln(\sigma) - (L-1)\ln(1-\rho^2) \\ &- \frac{a_1^2 + a_L^2 + (1+\rho^2)\sum_{i=2}^{L-1} a_i^2}{2(1-\rho^2)\sigma^2} \\ &+ \sum_{i=1}^{L-1} \ln\left(\mathrm{I}_0\left(\frac{\rho}{(1-\rho^2)\sigma^2} a_i a_{i+1}\right)\right). \end{aligned} \tag{7.6}$$

Here I_0 denotes the modified Bessel function of the first kind.

Exponential coherence loss models temporal decorrelation and is together with seasonal decorrelation commonly used as an InSAR coherence model [Fer+11; Sam+16] motivating its use for InSAR stacks. More complex coherence models with a residual coherence [PCE09] can also in first approximation be covered by Equation (7.5).

The GLRT based approach used in Equation (4.8) can be used for arbitrary probability distributions [DDT12]. The first step is to derive a maximum likelihood estimate (MLE) of Equation (7.6), which requires its derivation with respect to σ and ρ. The most intricate part is the last term with the modified Bessel function. With

$$\frac{\partial}{\partial x} \mathrm{I}_0(x) = \mathrm{I}_1(x) \tag{7.7}$$

and consequently

$$\frac{\partial}{\partial x} \ln \mathrm{I}_0(x) = \frac{\mathrm{I}_0(x)}{\mathrm{I}_1(x)} \tag{7.8}$$

the ratio of the modified Bessel functions of the first kind of the zeroth and first order is of interest. What makes a solution tractable are approximations. For large x [DLMF]

$$\mathrm{I}_0(x) \approx \frac{\mathrm{e}^x}{\sqrt{2\pi x}} \tag{7.9}$$

and

$$\frac{\mathrm{I}_0(x)}{\mathrm{I}_1(x)} \approx 1 - \frac{1}{2x}. \tag{7.10}$$

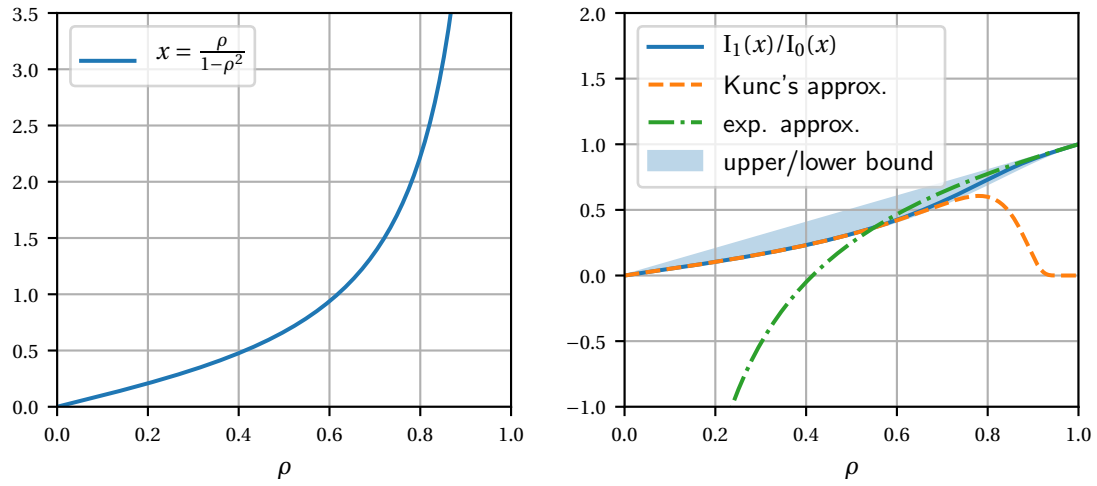


Figure 7.1: Approximations of the zero and first order modified Bessel functions' ratio as a function of ρ. Different approximations are needed for small and large arguments.

Conversely for small x using Kunc's approximation [Kun83]

$$\mathrm{I}_k(x) \approx \mathrm{I}_0(x) \left(\frac{x^k}{2^k k!} \mathrm{e}^{-\frac{k}{k+1}\frac{x}{4}} \right) \tag{7.11}$$

and

$$\frac{\mathrm{I}_0(x)}{\mathrm{I}_1(x)} \approx \frac{x}{2} \mathrm{e}^{-\frac{x^2}{8}}. \tag{7.12}$$

Figure 7.1 shows the two approximations of the modified Bessel function's ratio as a function of $\frac{\rho}{1-\rho^2}$. Their accuracy can to some extent be judged by computing lower and upper bounds of the ratio of modified Bessel functions [Amo74]

$$\frac{x}{\frac{1}{2} + \sqrt{x^2 + \frac{9}{4}}} \leq \frac{\mathrm{I}_1(x)}{\mathrm{I}_0(x)} \leq \frac{x}{\frac{1}{2} + \sqrt{x^2 + \frac{1}{4}}}, \tag{7.13}$$

which are also plotted in Figure 7.1.

Deriving Equation (7.6) with respect to σ, setting the result equal to zero and plugging in Equation (7.10) or Equation (7.12) gives

$$\sigma^2_{\text{large }\rho} = \frac{1}{L+1} \frac{1}{1-\rho^2} \left[a_1^2 + a_L^2 + (1+\rho^2) \sum_{i=2}^{L-1} a_i^2 - 2\rho \sum_{i=1}^{L-1} a_i a_{i+1} \right] \tag{7.14}$$

and

$$\sigma^2_{\text{small }\rho} = \frac{a_1^2 + a_L^2 + (1-\rho^2) \sum_{i=2}^{L-1} a_i^2}{2L(1-\rho^2) + 4\rho^2}. \tag{7.15}$$

Setting ρ to zero in Equation (7.15) reduces it to the maximum likelihood estimator for uncorrelated samples [Per73].

The same procedure for finding the MLE for ρ does not lead to a closed form solution. To still quickly judge whether a similarity criterion based on the multivariate Rayleigh distribution adequately separates samples of different distributions some necessary simplifications need to be accepted. ρ is not estimated in an MLE way but by fitting an exponential curve to the samples autocorrelation function. Additionally the joint estimation of σ and ρ is obtained by a simple average of the individual estimates. It needs to be stressed that this procedure no longer adheres to the GLRT criteria. The goal is simply to gauge whether this approach could make sense at all.

Analysis

A comparison of the multivariate Rayleigh GLRT based similarity criterion with other commonly used tests to detect statistically homogeneous pixels gives an impression of the expected performance and suitability. The other tests are the Kolmogorov-Smirnov statistical test and the two Gamma distributed GLRT based similarity criteria for spatial-temporal patches, with and without temporal averaging, which was already mentioned at the beginning of this section. For all GLRT based similarity criteria the individual pixel similarities inside a 3d patch are aggregated by taking their logarithm first and multiplying by minus one, akin to the spatial patch similarity measures for the two nonlocal filters [Che+11; Del+15]. The logarithm destroys however the nice property that all similarity criteria lie in the interval $[0,1]$ and their different scales need to be kept in mind.

The subsequent analysis shows the impact of the stack size and the patch size. Multivariate Rayleigh distributed samples with exponential coherence loss $\rho = 0.8$ are generated for two different values (1 and 2) for σ, meaning a difference of 3 dB.

Figure 7.2 shows for each similarity criterion normalized histograms if the two samples belong to the same (blue) or different (orange) distributions for a stack of size 64. The region where the two histograms overlap would lead to a misidentification and corruption of the filtering result, for example the smoothing of edges. Surprisingly there are only minor differences between the multivariate Rayleigh and the temporal mean Gamma based similarity criteria, except for their scale. The sometimes discredited Kolmogorov-Smirnov test performs distinctly better than the gamma log GLRT, which is hardly able to distinguish the two amplitude time-series.

Just like the extension of the bilateral filter to the nonlocal filter by moving from single pixel similarities to patch similarities the robustness of the single vector similarity estimate can be increased by computing its average over small patches. The effect can be seen in Figure 7.3 where the vector similarities from Figure 7.2 are averaged over 3×3 spatial patches. Due to the central limit theorem the histograms' shapes converge to a Gaussian and they are more concentrated with respect to the original histograms.

In conclusion Figure 7.2 and Figure 7.3 show that the multivariate Rayleigh GLRT, the Gamma GLRT with temporal mean and the Kolmogorov-Smirnov test perform comparably in this scenario. As the multivariate Rayleigh GLRT and the Kolmogorov-Smirnov test are much harder to compute their usage is difficult to justify compared to the Gamma GLRT with

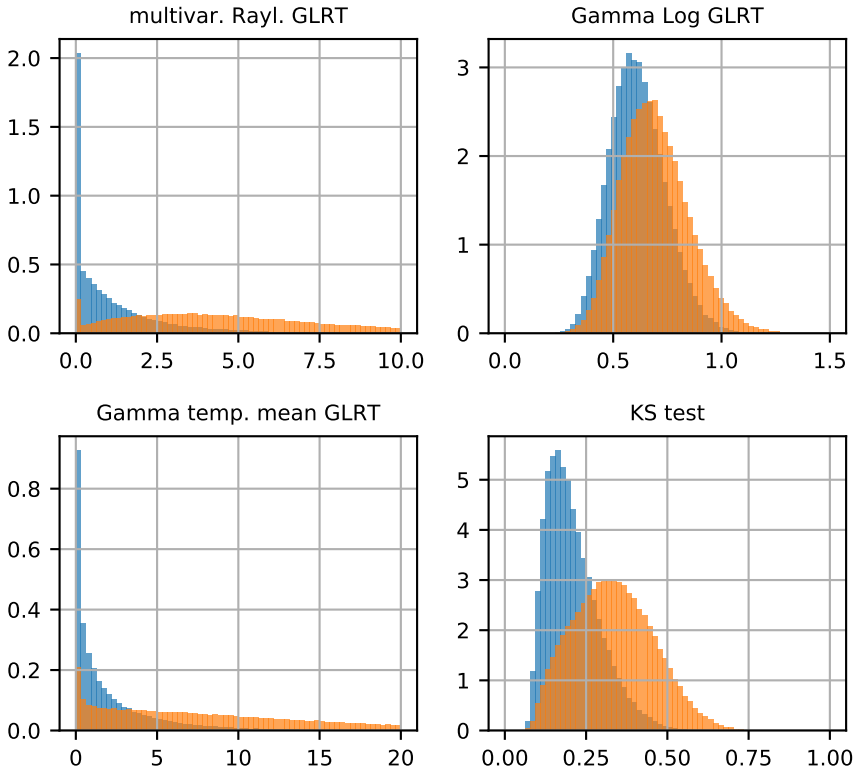


Figure 7.2: Normalized histograms of different similarity criteria for a multivariate Rayleigh distribution with exponential coherence loss $\rho = 0.8$ of size 64. The blue histograms are the result if samples of the Rayleigh distribution share identical parameters. A change by a factor of two for one of the multivariate Rayleigh distribution's σ, i.e. from 1 to 2, produces the orange histograms.

temporal mean.

Figure 7.4 further shows the impact of the stack size. The depicted histograms are the result of the same kind of experiment as before, with 3×3 patches but this time a stack size of 8. Evidently compared to the results with a temporal size of 64 Figure 7.3 the two distributions are not as well separated, indicating that just like an increase of spatial footprint an increase of temporal size helps identifying statistically homogeneous pixels.

7.2.2 Influence of coherence

The analysis of the previous subsection showed that the multivariate Rayleigh GLRT and the Gamma temporal mean GLRT perform similarly. Figure 7.4 also showed the impact that the size of the stack has on being able to separate samples from different distributions. As the coherence determines the effective size of the stack a more verbose investigation on its influence is appropriate.

Figure 7.5 shows for various stack sizes from 1 to 16 normalized histograms of Gamma temporal mean GLRT similarities for a 3×3 spatial patch. The randomly generated time series

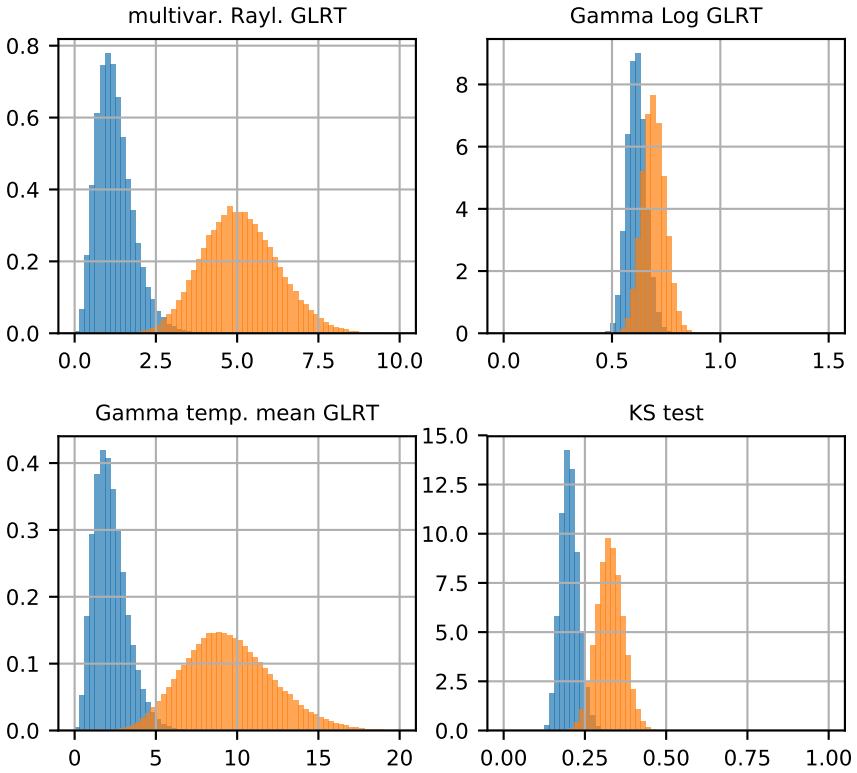


Figure 7.3: The same experiment as in Figure 7.2 but this time the average over 3×3 patches is taken, improving the robustness of the similarity estimate. As proved by the central limit theorem the histograms shape converge to a Gaussian distribution.

have the same underlying intensity but differ with respect to the exponential coherence loss, where ρ is defined as in Equation (7.5). Evidently the similarity values PDFs are a function of coherence and stack size. Any adaptive filter needs to account for these characteristics to make it applicable for a wide range of terrain, with different levels of coherence decay and also stack sizes.

Let $\mathbf{u} \sim \mathcal{CN}(0, \mathbf{\Sigma}) \in \mathbb{C}^N$ be the circular complex normal distributed of the stack. The temporal mean required for the gamma based GLRT is

$$\mu_t = \frac{\mathbf{u}^\dagger \mathbf{u}}{N} \tag{7.16}$$

and the covariance matrix's of $|\mathbf{u}|^2$, that is the exponentially distributed image intensity is [Mal03]

$$\mathbf{\Sigma}^I = 4\Re\{\mathbf{\Sigma}\}^2 + 4\Im\{\mathbf{\Sigma}\}^2 \tag{7.17}$$

under the assumption that $\mathbf{u}$ is circularly distributed, where $\Re$ and $\Im$ denote the real and

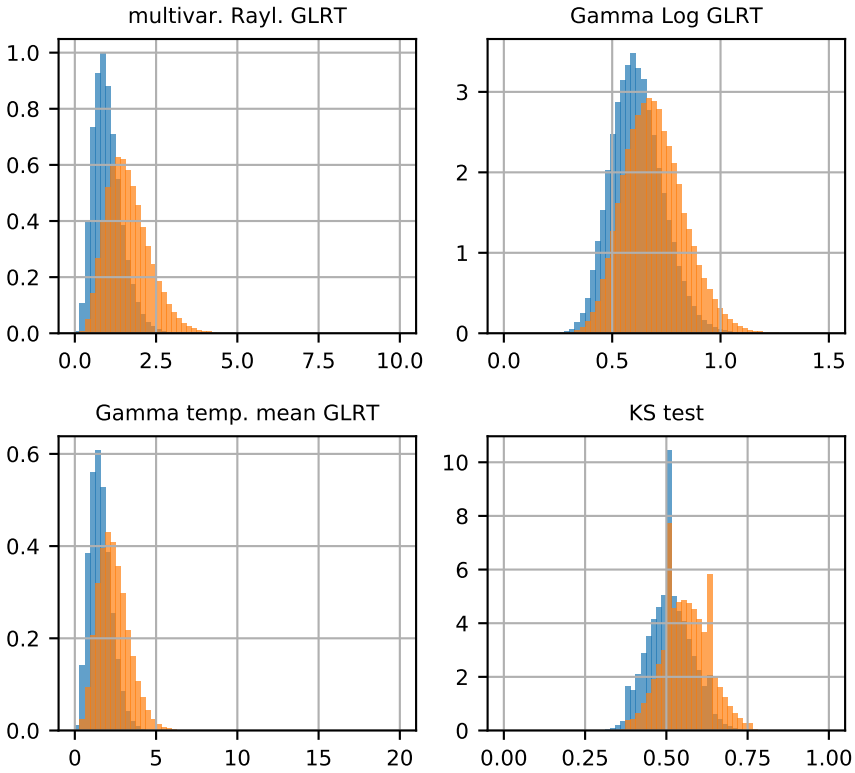


Figure 7.4: Identical experiment to Figure 7.3 but with the exception of a stack size of only 8. Compared to a stack size of 64 the histograms for identically and differently distributed samples overlap more, making effective filtering more difficult.

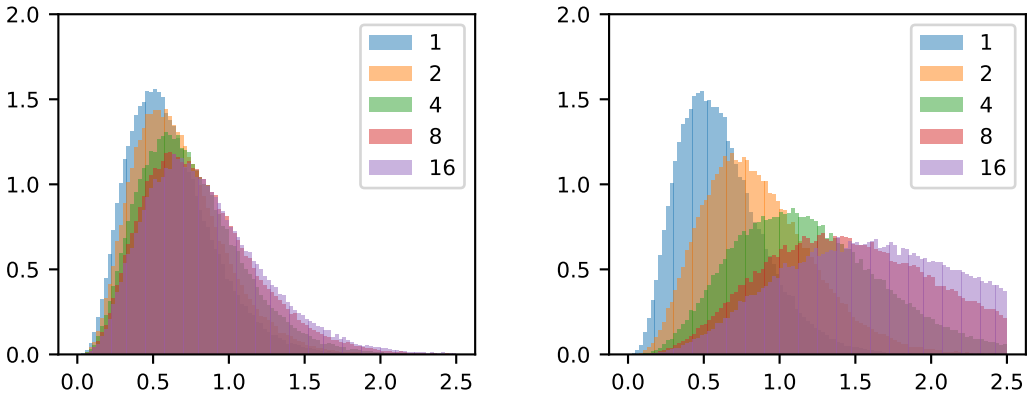
imaginary part. The temporal mean's variance is then given by

$$\sigma_t^2 = \frac{1}{N^2}\left(\mathrm{tr}\left(\mathbf{\Sigma}^I\right) + \sum_{k\neq l}|\mathbf{\Sigma}^I_{k,l}|\right). \tag{7.18}$$

The second term in Equation (7.18) accounts for correlation between acquisitions and increases the variance. The effective number of looks can then be calculated by the reduction in variance

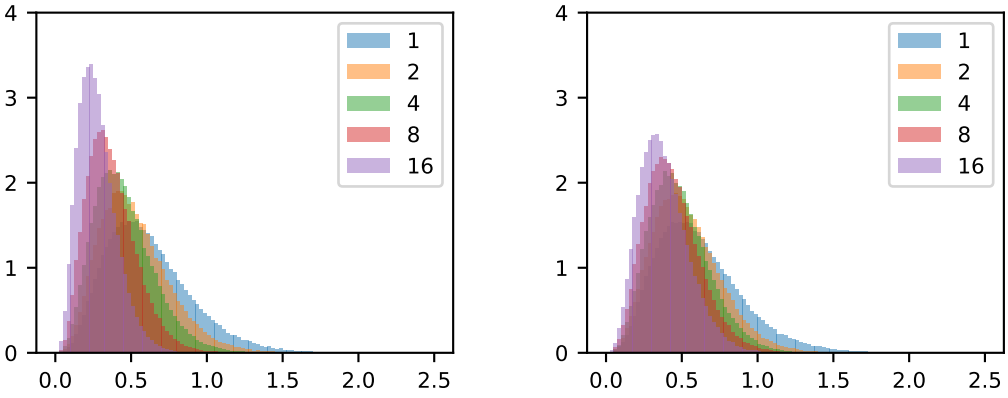
$$L = \frac{\mathrm{tr}\left(\mathbf{\Sigma}^I\right)}{N\sigma_t^2} = \frac{N\,\mathrm{tr}\left(\mathbf{\Sigma}^I\right)}{\mathrm{tr}\left(\mathbf{\Sigma}^I\right) + \sum_{k\neq l}|\mathbf{\Sigma}^I_{k,l}|}. \tag{7.19}$$

Using the effective number of looks L instead of N in Equation (7.4) makes the similarity computation more coherence independent. Figure 7.6 shows the result for the same experiment as in Figure 7.5, where the covariance matrix was estimated by spatial averaging inside the patch. An alternative would be an iterative filtering algorithm where the covariance matrix estimate of the previous iterations is used to compute the effective number of looks. In any case already for the simple employed estimation scheme the histograms are evidently to a lesser degree influenced by changes of the correlation coefficient, which makes this similarity criterion applicable to a wider range of terrain and decorrelation processes.



(a) $\rho = 0.5$ **(b)** $\rho = 0.8$

Figure 7.5: Histograms for the log gamma GLRT based similarity criterion Equation (7.4) with different number of looks, i.e. stack size with temporal mean, with exponential coherence loss and 3×3 patches. The histograms are strongly dependent on the correlation and the stack size.



(a) $\rho = 0.5$ **(b)** $\rho = 0.8$

Figure 7.6: Identical experiment as for Figure 7.5 but the effective number of looks estimated by Equation (7.19) is used instead. The similarities histograms are almost independent of stack size and the decorrelation factor.

7.3 Proposed stack filtering framework

The previous two sections introduced two potential similarity measures for InSAR stacks and figures of merit for evaluating a nonlocal stack filter. The remainder of this section describes a new nonlocal stack filtering framework that makes use of these lessons. The general procedure is identical to the conventional nonlocal filter, with a search for statistically homogeneous pixels by comparing their surrounding patches and a subsequent collective estimate.

7.3.1 InSAR covariance matrix estimation

In contrast to the single baseline nonlocal filter introduced in Chapter 5, which performs filtering patch-wise, SAR stacks introduce time as a third dimension. Consequently a multi-baseline nonlocal filter will operate on temporal-spatial patches $\mathcal{U} \in \mathbb{C}^{t\times p\times p}$ with t denoting the temporal extension of the stack and $p \times p$ being the quadratic spatial patch dimensions. There exist two ways how to obtain InSAR covariance matrix estimates using a collection of patches.

Weighted mean

The nonlocal patch wise weighted mean from Equation (5.1) can directly by extended to deal with spatial-temporal patches $\mathcal{U} \in \mathbb{C}^{t\times p\times p}$ and produce a 4D $\hat{\mathcal{C}} \in \mathbb{C}^{t\times t\times p\times p}$ patch wise covariance matrix estimate. The only difference being that instead of a singular weight per patch there is now a matrix $\boldsymbol{W} \in \mathbb{R}^{t\times t}$ for every possible interferogram. With a slight abuse of notation the weighted mean can be written as

$$\hat{\mathcal{C}}_{\mathbf{x}} = \sum_{\mathbf{y}\in\partial_{\mathbf{x}}} \boldsymbol{W}_{\mathbf{x},\mathbf{y}} \circ \mathcal{U}_{\mathbf{y}}\mathcal{U}_{\mathbf{y}}^{\dagger}. \tag{7.20}$$

Depending on how the weight map is constructed individual interferograms can be filtered independently or also as substacks in a joint fashion.

Covariance matrix estimators

An alternative to the weighted mean is to directly estimate the covariance matrix. In the scope of this thesis the sample covariance matrix (SCM)

$$\hat{\boldsymbol{\Sigma}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{u}_i\mathbf{u}_i^{\dagger} \tag{7.21}$$

is compared to Tyler's iterative M estimator [Oll+12]

$$\hat{\boldsymbol{\Sigma}}_{k+1} = \frac{m}{n}\sum_{i=1}^{n} \frac{\mathbf{u}_i\mathbf{u}_i^{\dagger}}{\mathbf{u}_i^{\dagger}\hat{\boldsymbol{\Sigma}}_k^{-1}\mathbf{u}_i} \tag{7.22}$$

and a robust shrinkage based, iterative estimator [CWH11]

$$\begin{aligned}\tilde{\boldsymbol{\Sigma}}_{k+1} &= (1-\beta)\frac{m}{n}\sum_{i=1}^{n}\frac{\mathbf{u}_i\mathbf{u}_i^\dagger}{\mathbf{u}_i^\dagger\hat{\boldsymbol{\Sigma}}_k^{-1}\mathbf{u}_i} + \beta\boldsymbol{I}\\ \hat{\boldsymbol{\Sigma}}_{k+1} &= \frac{\tilde{\boldsymbol{\Sigma}}_{k+1}}{\operatorname{tr}\left(\tilde{\boldsymbol{\Sigma}}_{k+1}\right)/m},\end{aligned} \tag{7.23}$$

where β is the shrinkage coefficient and $\boldsymbol{I}$ the identity matrix. A formula on how to choose β to minimize the mean-squared error can be found in the original publication.

For all covariance matrix estimation methods the set of samples can simply be obtained by hard-thresholding the patch similarities, which is comparable to tuning the h parameter of the conventional nonlocal mean's exponential kernel.

Going to a patch-wise estimate can simply be accomplished by using the abovementioned formulae individually on the contained pixels. Just like the weighted mean Equation (7.20) the covariance matrix estimates can be performed on the complete stack or also separate substacks.

Aggregation

The individual patch-wise estimates of the previous section naturally overlap in space and, depending on how the substacks are formed, possibly in time as well. Aggregating them into a final estimate is done identically to the single baseline filter in Chapter 5 by weighting them by their equivalent number of looks, which mitigates the rare patch effect. The actual formation of, possibly overlapping, substacks is described in the following section.

7.3.2 Clustering

As a result of existing satellite missions' repeat cycles[1], about 12 days for Sentinel-1, 11 days for TerraSAR-X, 14 days for ALOS-2, and 16 days for COSMO-SkyMed InSAR stacks span multiple weeks, months or years even. With these long temporal baselines gradual as well as sudden changes can occur in the observed area, which can inhibit the search for similar patches. On a high level of abstraction the proposed framework circumvents this limitation by splitting up the search volume along the temporal dimension into homogeneous, not necessarily disjunct parts.

This clustering procedure first computes for every spatial coordinate of the stack a distance matrix measuring the similarity of spatial patches along the temporal axis

$$\boldsymbol{D}_{k,l} = \Delta\left(\mathfrak{U}_{t_k}, \mathfrak{U}_{t_l}\right), \tag{7.24}$$

where Δ denotes the distance function.

The distance matrix can directly be used by various clustering methods to split up the search volume. Either into disjoint sets, so-called *hard* clustering, or allowing overlapping

[1] Repeat cycles are given for single satellites only. For Sentinel-1 and COSMO-SkyMed where the constellation consists of two or respectively four distributed satellites the effective repeat cycle is significantly shorter

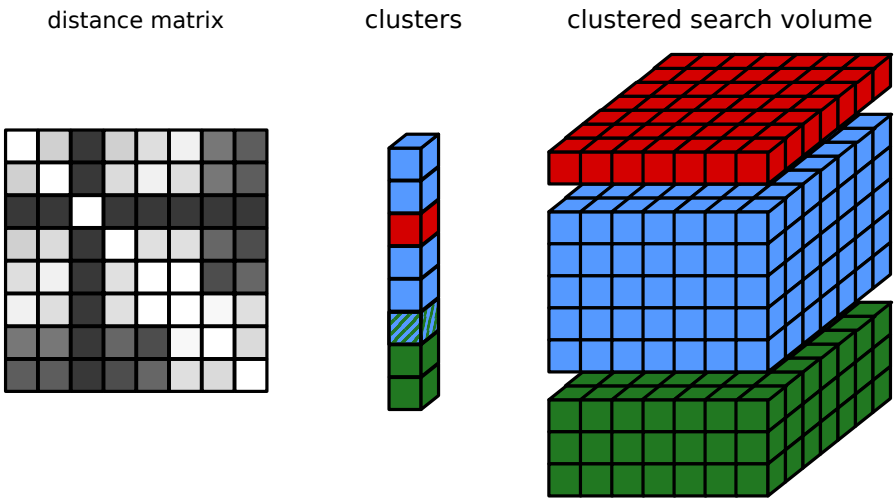


Figure 7.7: Clustering of the search volume in possibly overlapping clusters by the distance matrix of a time-series.

search windows by *fuzzy* clustering [Yan93]. Within each cluster $\mathscr{C}_k$ the search volume can be considered as homogeneous in the temporal dimension and a weight map $\boldsymbol{W}^{\mathscr{C}_k} \in \mathbb{R}^{s\times s}$ is computed.

Figure 7.7 illustrates the clustering procedure. The symmetric distance matrix assigns each image of the stack to a set of clusters, which is then use to divide up the search volume into separate entities. The number of clusters and their members vary in space since the computed distance matrices are naturally data dependent.

The weighted mean as well as the various covariance matrix estimators introduced in Section 7.3.1, can work on the full stack or a set of substacks, which can either be disjunct or overlapping. This set of substacks does not necessarily cover all possible combinations of interferograms. In other words, to get a covariance matrix estimate of the complete stack the weight map $\boldsymbol{W}$ must not only include the actual clusters' weight maps, i.e. intra-cluster weights, but also inter-cluster weights between clusters.

There exist two possibilities how to handle this set of interferograms. They can either be dealt with individually by conventional single baseline filtering, or the missing weights of $\boldsymbol{W}$ are computed as a function of the clusters' weight maps $\boldsymbol{W}^{\mathscr{C}_k}$ the corresponding images are assigned to.

Let $\mathbf{1}_k : T \rightarrow 0,1$ be the indicator function defined as

$$\mathbf{1}_k(t) := \begin{cases} 1 & \text{if } t \in \mathscr{C}_k \\ 0 & \text{if } t \notin \mathscr{C}_k \end{cases}. \tag{7.25}$$

The set of weights that were computed for clusters that either include t_1 or t_2 can then be written as

$$W_{t_1,t_2} = \left\{ \boldsymbol{W}^{\mathscr{C}_k}_{\mathbf{x},\mathbf{y}} : \mathbf{1}_k(t) = 1 \text{ where } t \in \{1,2\} \right\} \tag{7.26}$$

and the individual missing weights can be computed as

$$w_{\mathbf{x},\mathbf{y},t_1,t_2} = f\left(W_{t_1,t_2}\right), \tag{7.27}$$

where f is some sensible function. Suitable choices are taking the geometric or arithmetic mean or the maximum or minimum of the individual cluster weights.

A more exhaustive discussion of the clustering procedure is appropriate. There exist a wealth of clustering algorithms [Sch07]. Together with the distance function the most critical part is setting the number of clusters. In the most common case one deals with temporal homogeneous scenes so that the clustering method's parameters should be tuned such that one single cluster remains. Other scenarios are singular outliers or sudden but permanent changes so that it makes sense to limit the maximum number of clusters to two or three. Both considerations define an interval for a suitable set of clustering parameters.

The uninitiated reader may liken the clustering approach to the SBAS technique [Ber+02], which works on subsets of stacks with small spatial and temporal baseline and hence intrinsically higher coherence. Mutual coherence is not the motivation for the clustering, finding similar images to prevent the rare-patch effect in the temporal domain is.

There also exist some similarities to the multitemporal despeckling filter proposed in [Lê+14], which also performs clustering based on the distance matrix.

7.3.3 Exemplary filters

This section presents a couple of straightforward nonlocal stack filters implemented within the previously described framework. The remaining paragraphs describe the three major choices that have to be made: which similarity criteria to use, how clustering is performed and the actual estimator of the stack's covariance matrix. Section 7.4 evaluates their performance.

Similarity criteria

Two similarity criteria are needed

1. The first one for computing the distance matrix to cluster the search window,
2. and a second one to compute the spatial-temporal patch similarities within each cluster.

The first one must be strictly reflectivity-based and Equation (7.4) on single images without neither temporal nor spatial multilooking is used as essentially a change detection test. Other possible similarity measures such as PPB [DDT09], could also be employed.

The actual spatial-temporal patch similarities are then computed using the coherence-compensated Gamma GLRT proposed in Section 7.2.2. This has the advantage that the computed similarities are rather independent of the varying stack sizes that naturally arises due to clustering.

For both similarity measures comparatively small 3×3 patches are used for the whole image. The idea is that temporal outliers are well localized in space and hence benefit from small

patch sizes. For homogeneous areas, which require larger spatial dimensions to increase the accuracy of the computed area, one can assume that they remain largely unchanged over the observation duration and the temporal domain will provide enough information to compute accurate patch similarity estimates, so that small spatial dimensions suffice.

Clustering

The proposed filters rely on hierarchical clustering [DE84] which builds up a structured, tree-like class hierarchy, called *dendogram*, by iteratively merging clusters with the shortest distance, starting with all spatial patches and the original distance matrix $\boldsymbol{D}$. The actual clustering is done by *single-linkage*, which defines the distances between two clusters $\mathscr{C}_{k\,:\,k\in\{1,2\}}$ as the minimum distance between all patches $\boldsymbol{U}$ that are contained in the clusters

$$\Delta_{\mathscr{C}}\left(\mathscr{C}_1,\mathscr{C}_2\right)=\min_{\boldsymbol{U}_1\in\mathscr{C}_1,\boldsymbol{U}_2\in\mathscr{C}_2}\Delta\left(\boldsymbol{U}_1,\boldsymbol{U}_2\right). \tag{7.28}$$

The subsequent flattening of the hierarchical cluster structure subdivides the search volume into clusters of similar nature, such that all cophenetic distances inside the remaining clusters are lower than a preset threshold $\Delta_{\max}$

$$\Delta_{\mathscr{C}}\left(\mathscr{C}_k,\mathscr{C}_l\right)<\Delta_{\max}. \tag{7.29}$$

Two parameters are at play and define the final clustering: the hierarchical clustering's linkage method and the choice of flattening procedure. For both many choices exist [DE84]. As mentioned earlier $\Delta_{\max}$ must be tuned such that in most cases only a single cluster remains, which translates to no temporal changes. If the number of clusters is kept low or even to one, with the exception of the occasional outlier, the choice of the linkage algorithm plays no major role and confining oneself to the computational efficient single-linkage algorithm is not critical.

Stack filtering

The proposed filters include the weighted means and all previously described (see Section 7.3.1) covariance matrix estimators. The weighted mean relies on weights computed by an exponential kernel, whereas the covariance matrix estimators employ simple hard thresholding to discard dissimilar pixels. Obviously the weighting kernel's smoothing parameter h and the threshold can be either tuned for noise reduction or detail preservation.

7.4 Evaluation

In total the previous section described eight different filters within the proposed framework. They filters use either the previously described clustering procedure in Section 7.3.3 or a fixed temporal size extending over the complete stack, i.e. no clustering. The other degree of freedom is the actual estimation of the stack covariance matrix by one of the four methods from Section 7.3.1, which results in eight filtering combinations.

Estimator	Fixed	Clustering
weighted mean	0.1433	0.1434
SCM	**0.0298**	0.0310
Tyler	0.0332	0.0341
Chen	0.0322	0.0333

Table 7.1: RMSLE for a stack of size eight.

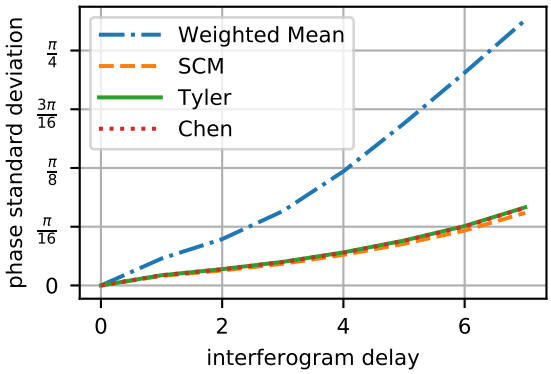


Figure 7.8: Phase standard deviation for the experiment from Table 7.1 without clustering. The plot is in good agreement with the reported RMSLEs.

7.4.1 Simulations

Table 7.1 lists the RMSLE values for a completely homogeneous stack of size $8 \times 512 \times 512$ with exponential coherence loss described by $\rho = 0.8^{|i-j|}$, where i and j are the integer indices of the simulated SLCs. For all methods clustering slightly decreases the filtering efficiency, as expected. Furthermore all covariance matrix filters perform comparably, the big outlier being the weighted mean. Whether its poor performance is remedied by a better detail preservation is covered later. Obviously adjusting the weighting parameter h can have a major influence on the detail preservation.

To better justify the use of the RMSLE Figure 7.8 shows the phase standard deviation of the various filters with no clustering as a function for the same experiment. Overall the RMSLE is able to capture the slight differences between the various covariance matrix estimators and also roughly reflects the difference to the weighted mean.

To get a better grasp of the methods' detail preservation capabilities Table 7.2 lists the error vector magnitude in dB for a noise-free persistent scatterer. The numbers are the result of a Monte-Carlo simulation with 10000 repetitions, for a completely homogeneous stack with temporal size 8 and a single 3×3 persistent scatterer, which is 9 dB brighter than its surroundings. The persistent scatterer stays constant over the observation duration, whereas the remaining pixels decorrelate as before with $\rho = 0.8^{|i-j|}$. Again clustering lowers the performance, slightly for the weighted means and significantly for all covariance matrix estimators. Chen's regularized covariance estimator also degrades the detail preservation compared to

Estimator	Fixed	Clustering
weighted mean	-4.082	-4.049
SCM	**-8.608**	-6.663
Tyler	-8.391	-6.590
Chen	-5.513	-4.356

Table 7.2: Error vector magnitude in dB.

weighted mean	SCM
0.595	**0.402**

Table 7.3: Ratio of the equivalent number of looks' minimum and maximum along the temporal domain with a single outlier. The ratios are the result of a Monte-Carlo simulation with 10000 runs.

the other estimators.

As a last example Table 7.3 shows the ratio of the minimum to the maximum of the ENL, again for a homogeneous stack with temporal size 8 and a single temporal outlier of size 5×5 and an increase in reflectivity by a factor of 4. With no clustering the ratio would obviously be 1 and a lower number indicates that the outlier does to a lesser degree affect the search for similar pixels.

In one aspect one has to take the results reported in Table 7.1, Table 7.2, and Table 7.3 with a grain of salt. The sample covariance matrix estimator seems to offer the best performance. However with these rather simplistic simulations it might well be that the binary decision of its hard thresholding does give it an unfair advantage. A real data set might provide the missing insights.

7.4.2 Real data

A stack of Sentinel-1 interferometric wide swath acquisitions featuring the coastal village Santa Marina de Salina on the Italian island Salina assesses some of the aforementioned filters. The data set was generously provided by the authors of [AZB17], who performed stacking and removed the phase's topographic component, presents a mix of different scenes. The village and its harbor are examples of heterogeneous terrain and their shipping traffic produces temporal outliers. Speckle reduction for homogeneous areas can be evaluated on the Mediterranean sea itself.

In particular, the weighted mean is evaluated with and without clustering and is compared to the SCM estimator, also with clustering. The filters are further compared to a wavelet-shrinkage based despeckling filter, which is identical to SAR-BM3D [Par+12], with the exception that it does not perform block matching, but directly performs wavelet shrinkage on fixed size spatial-temporal groups, which extend over the complete time axis.

Figure 7.9 shows the despeckling results for a substack of eight images, which also allows to draw conclusions for the denoising of the interferometric phase. The wavelet-shrinkage

based filter introduces *ghosting* artifacts for outliers and provides the overall worst performance, indicating that adapting spatial filters to the temporal dimension is not straightforward. Filtering artifacts are also generated by the weighted mean filter without clustering, since outliers affect the weights of the whole stack, which is especially visible for the big ship in the third image. Both filters with clustering, the weighted mean and the SCM estimator, avoid these artifacts. Their difference in noise reduction and detail preservation can be attributed to their filtering parameters, which can be tuned for one or the other.

For the same test site, Figure 7.10 shows all raw interferograms and the result obtained by the previously mentioned nonlocal means algorithm with clustering, again for a sub stack of eight images. The lower left triangular matrix shows the raw input data, whereas the corresponding filtered interferograms are depicted in the upper right triangular matrix. Pixels with a coherence lower than 0.3 have been masked out. Artifacts observable in interferograms containing outliers, i.e. ships, are a result of combining weights from both SLCs.

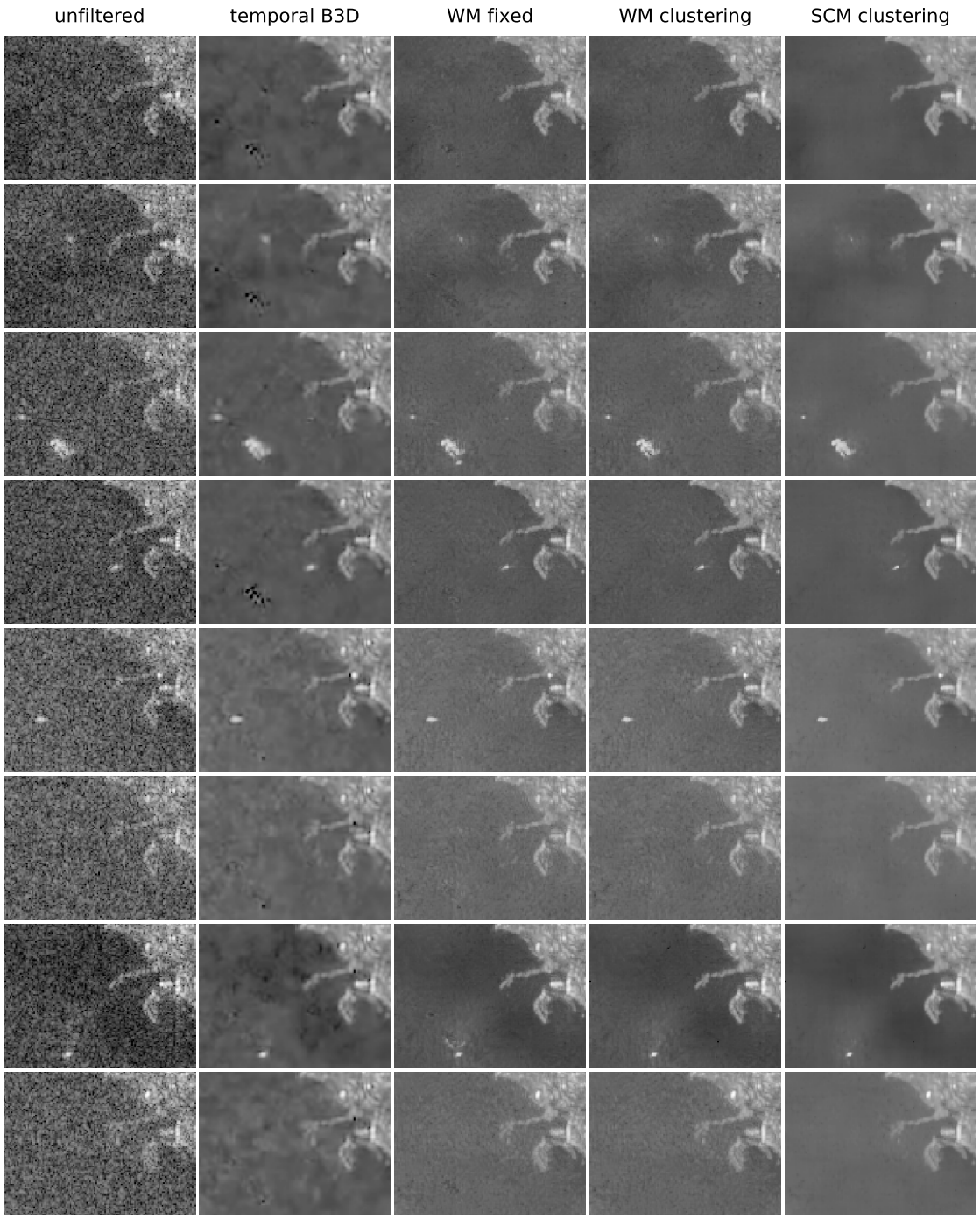


Figure 7.9: Comparison of different stack filtering methods. The test site shows a time series of a coastal region with shipping traffic. Evaluated are three of the previously introduced filters: weighted mean with and without clustering and the SCM estimator with clustering. The result of a despeckling filter using fixed spatial-temporal blocks and following the wavelet shrinkage procedure in [Par+12] is plotted as well.

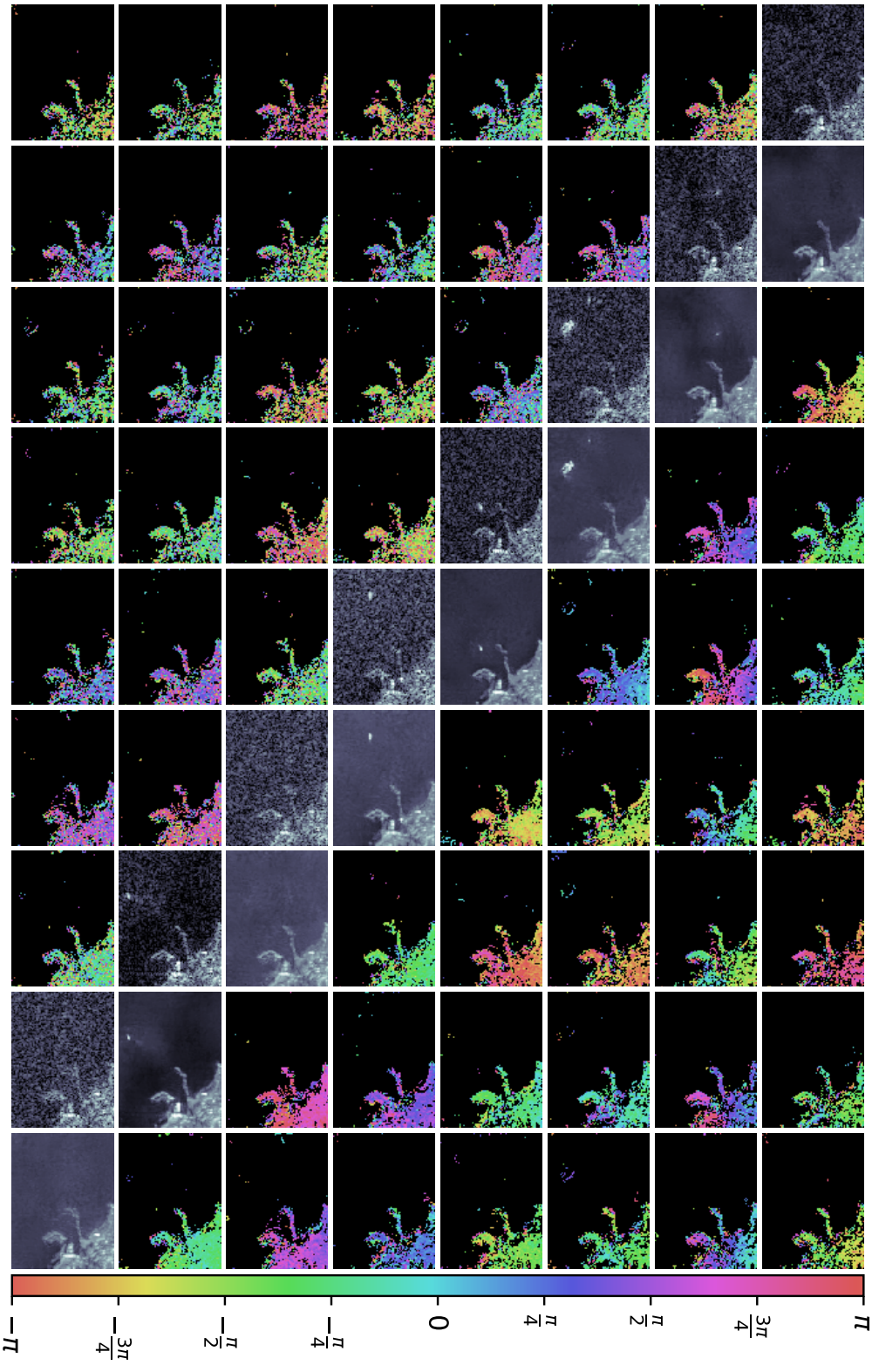


Figure 7.10: Raw interferograms (lower left triangular matrix) and nonlocal means with clustering filtered interferograms (upper right triangular matrix) for the test size of Figure 7.9. Artifacts observable in interferograms containing outliers, i.e. ships, are a result of combining weights from both SLCs.

8 Efficiently implementing nonlocal filters

The significant improvement in denoising performance of the nonlocal means over legacy filters comes at the cost of far higher computational costs. This is especially pertinent for the data volume of TanDEM-X and particularly Sentinel-1. For example the generation of the global TanDEM-X DEM required about 500,000 CoSSC acquisitions, occasionally up to seven for the same spot depending on the terrain's difficulty [Riz+17]. Such humongous data sets require efficient processing algorithms, which has long been recognized as one of the pillars required for big data processing [Chi+16].

There exist several modifications of the original nonlocal filter to make it computationally more tractable. In [BKC08] the authors propose that instead of an exhaustive search over the whole image or search window for similar patches, clustering of similar patches in a binary tree is performed as a preprocessing step. The actual denoising is then only performed using patches belonging to the same cluster. Another approach is to preclassify similar patches using easily computable metrics such as the patch mean value or greyscale gradient and discard non-matching patches early on from the nonlocal means [MS05]. In a similar vein the method proposed in [VOK10] terminates the patch similarity computation, which is comprised of individual pixel similarities, as soon as it is unlikely that the final patch similarity will be high. The approach chosen in [Dar+08] exploits redundancies in the nonlocal filtering algorithm and alludes to efficient computation on modern computer architectures.

This chapter is structured as follows. Section 8.1 provides some theoretical justification for taking a closer look at the algorithm and the actual hardware it is running on. The algorithmic complexity of the nonlocal means and some redundancies that can be exploited are highlighted in Section 8.2. An overview on which mechanisms of contemporary processor architectures contributed in the past years to increases in numerical performance is given in Section 8.3. How nonlocal filters can take advantage of them is described in Section 8.4 and Section 8.5. Finally Section 8.6 shows benchmark results for the NL-InSAR filter [DDT11], which can be generalized to other nonlocal filters, and Section 8.7 introduces some possible further optimizations.

8.1 Theoretical considerations

[Pat07] postulates that efficiently solving any problem that runs on microchips requires a multifaceted understanding of several layers of abstraction starting from the problem itself down to the actual microchip's electronic circuit. Table 8.1 lists the different layers where a transformation from one domain to another takes place.

In more detail, to solve any kind of problem scientists and engineers come up with an appropriate algorithm that is implemented as a program, which at some point is translated to

Table 8.1: Transformation hierarchy [Pat07]

Problem
Algorithm
Program
Instruction Set Architecture
Microarchitecture
Circuits
Electrons

a more low-level instruction set. The program's instructions are then carried out by a processor, the architecture of which defines the actual execution. For example special processing blocks for common functions, such as trigonometric functions, or fast memory blocks for quickly storing or caching intermediate results may differ for different architectures. That is the actual electronic circuits, which themselves are an arrangement of transistors, where at the last layer even electrons have to be considered.

Completely understanding this whole domain can lead to efficient implementations of algorithms on modern CPU or GPU architectures. Such a study was for example conducted in [Pla+11], which evaluates different compute architectures for common hyperspectral imaging algorithms. The following sections dissect the nonlocal filtering algorithm highlighting the relevant aspects in Table 8.1.

8.2 Algorithmic complexity

A naïvely implemented nonlocal filter has an algorithmic complexity of

$$\mathcal{O}(SPN), \tag{8.1}$$

where S denotes the number of pixels in the search window, P the number of pixels in a patch and N the number of pixels in the image.

Upon closer inspection and as mentioned in [Dar+08], there exist redundancies when computing the similarities between pixels, which are later summed up to patch similarities. Figure 8.1 illustrates how this can be exploited. Some of the pixel similarities that where computed for the patch similarity between the blue patches can be reused for calculating the red patches' similarity.

Keeping this in mind the computational cost of a nonlocal filter can be split up into separate terms for the pixel similarities, which no longer depends on the patch size, the patch similarities, the mapping of similarities into weights by the weighting kernel and the actual weighted mean

$$\underbrace{\mathcal{O}(SN)}_{\text{pixel similarities}} + \underbrace{\mathcal{O}(SPN)}_{\text{patch similarities}} + \underbrace{\mathcal{O}(SN)}_{\text{weighting kernel}} + \underbrace{\mathcal{O}(SN)}_{\text{weighted mean}} . \tag{8.2}$$

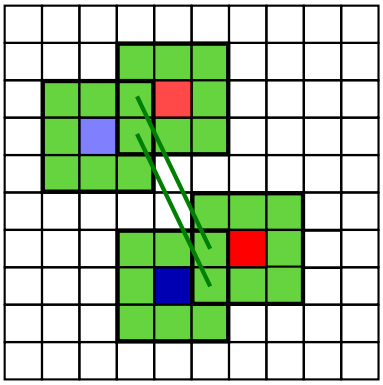


Figure 8.1: Redundancies between patch similarities. As overlapping patches share the same pixel similarities, these do not have to be recomputed, but can be computed only once for all patches.

A more thorough analysis of the individual terms is in order. Computing the pixel similarities involves evaluating complex expressions such as Equation (4.6), Equation (4.7) or Equation (4.8). In particular the calculation of trigonometric functions or matrix determinants incurs significant costs. As a result, even though the term is independent of the patch size P, it might well dominate or in the very least play a major role in the overall computational costs. On the contrary, although dependent on P, the computation of the actual patch similarities requires only summing up pixel similarities. Essentially this can be accomplished by S convolutions, for each offset of the pixel similarities, by a filter with footprint size P and is therefore relatively cheap to compute. The algorithmic complexities of the weighting kernel and the weighted means hold no surprises. Mapping similarities into weights requires usually just the evaluation of an exponential function and the weighted means can be regarded as a series of multiplications and additions.

8.3 Modern CPU architectures

With the previous analysis of the nonlocal filtering algorithm the only thing that remains is understanding how the actual silicon performs computation to efficiently implement the filter. Table 8.2 lists several top of the line processors from Intel of the last years and their relevant specifications. The base frequency is the default operating or clock frequency of the processor and each of the cores can perform arithmetic calculations or other instructions. The vector width deserves a more verbose introduction. Modern processors can not only work on scalars alone but also on short vectors of data, which is widely known as single instruction, multiple data (SIMD). The length of the vector together with the datatype defines how many math operations are performed within one cycle. For example four single precision floating point numbers of 32 bits each fit into a 128 bit SIMD register.

To calculate the theoretical peak performance of a processor in Table 8.2 in floating point operations per second it is best to start at the instruction level. Each core can execute two SIMD instructions per cycle. So with the previous example of a 128 bit SIMD register and

Table 8.2: Evolution of clock frequency, number of cores and vector width and resulting theoretical peak performance over the years for top-of-the-line Intel Xeon processors. Processors marked by * support fused multiply-add (FMA) instructions, doubling their peak performance.

Processor	year	Base frequency	# cores	Vector width	Peak performance
E7–8870	2011 Q2	2.4 GHz	10	128 bits	192 GFLOPS
E7–8890 v2	2014 Q1	2.8 GHz	15	256 bits	672 GFLOPS
E7–8890 v3*	2015 Q2	2.5 GHz	18	256 bits	1440 GFLOPS
E7–8890 v4*	2016 Q2	2.2 GHz	24	256 bits	1690 GFLOPS
Plat. 8180*	2017 Q3	2.5 GHz	28	512 bits	4480 GFLOPS

single precision floating point numbers

$$2\,\frac{\text{instr.}}{\text{cycle}}\,\frac{1}{\text{core}} \times 4\,\frac{\text{FLOP}}{\text{instr.}} = 8\,\frac{\text{FLOP}}{\text{cycle}}\,\frac{1}{\text{core}}. \tag{8.3}$$

Multiplying these with the number of cores and the clock frequency gives

$$8\,\frac{\text{FLOP}}{\text{cycle}}\,\frac{1}{\text{core}} \times 10\,\text{cores} \times 2.4\,\text{GHz} = 192\,\text{GFLOPS}, \tag{8.4}$$

where the values of the E7–8870 were plugged in.

With regard to Table 8.2, Figure 8.2 shows how the clock frequency, the number of cores and the vector width have evolved over time.

Evidently to make efficient use of the resources on a CPU the code must be parallelized to take advantage of the increasing number of cores but also make use of vector instructions. The following sections will briefly describe how this affects the implementation of a nonlocal filter.

8.4 Vectorization

The previous sections outlined the different parts of the nonlocal filtering algorithm and the importance of vectorization. Fortunately exploiting the redundancies when computing pixel similarities (see Figure 8.1) lends itself to a vectorized implementation, which requires

1. operating on short vectors instead of scalars,
2. whose elements are located in contiguous memory addresses, i.e. coalesced memory access.

Computing all pixel similarities $\delta_{\mathbf{x},\mathbf{y}}$ for a search window defined by the offsets $\mathbf{o} \in \mathcal{O}$ of pixels to the center pixel, is equivalent to shifting the whole image I by $\mathbf{o}$ and applying the similarity function f on I and its shifted version

$$\delta_{\mathbf{x},\mathbf{y}} = \delta_{\mathbf{x},\mathbf{x}+\mathbf{o}} = f\left(I_{\mathbf{x}}, I_{\mathbf{x}+\mathbf{o}}\right). \tag{8.5}$$

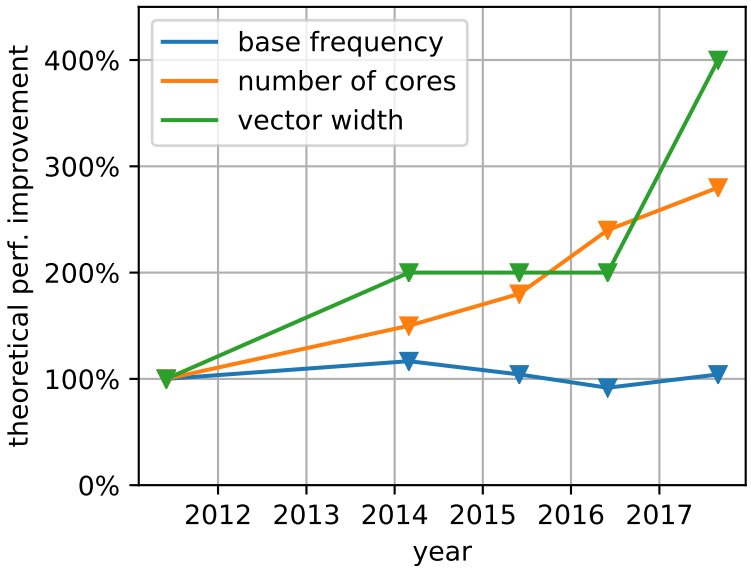
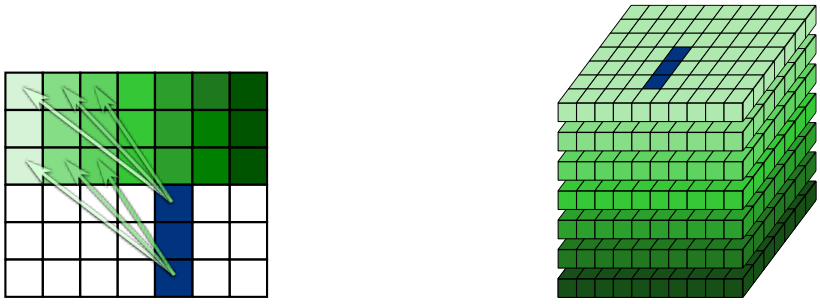


Figure 8.2: Evolution of clock frequency, number of cores and vector width over the years for the processors listed in Table 8.2

The crux is, that instead of computing all pixel similarities in a search window to a single pixel and moving to the next pixel, all pixel similarities of the same offset are computed for the whole image and then the procedure is repeated for the next offset. As I is stored contiguously in memory f can be computed in a vectorized manner.

The concept is illustrated by Figure 8.3, which on the left shows how similarities of the same offset, indicated by their different shades of green, are computed in a vectorized form and stored contiguously in memory on the right side.

Moving from pixel similarities to patch similarities is only a matter of performing a convolution with the appropriate kernel on the pixel similarities, as thanks to the previous step pixel similarities with the same offset are stored contiguously in memory. 2D convolutions



(a) pixel similarity computation **(b)** pixel similarity memory layout

Figure 8.3: Vectorized computation of pixel similarities with the same offset and their storage in memory as contiguous arrays with the image's dimension.

Table 8.3: Different instruction sets and their vector widths

Instruction Set	Vector Width
SSE4.2	128 bits
AVX	256 bits
AVX2	256 bits + FMA

with a boxcar or Gaussian kernel are further separable into two 1D convolution, providing additional speed ups to the algorithm.

The application of the weighting kernel and the weighted mean itself are not covered in detail, as both are rather straightforward. Essentially the weighting kernel is only a mapping of the pixel similarities into weights without any spatial considerations and restrictions. The weighted means is basically a 2D convolution with a spatially varying filtering kernel.

8.5 Parallelization

The actual parallelization of nonlocal filters is rather straightforward. Either the image itself can be split up into overlapping tiles, which are then processed separately or the individual operations can be performed block-wise, where each block is assigned to a core.

8.6 Benchmarks

As a case study the NL-InSAR filter [DDT11] was implemented. The filter is both rather close to the original nonlocal filter [BCM05b], but on the other hand shares some characteristics to the proposed nonlocal InSAR filter in Chapter 5, so that some conclusions can be drawn for that filter as well. Single precision floating point numbers are used to maximize the benefit of vectorization and since the numerical accuracy that they provide is more than suffices

Figure 8.4 shows how long the computation of the pixel likelihoods and Kullback-Leibler divergences, the patch similarities, the weighting kernel and the weighted mean take. The measurements were taken for a search window size of 21×21 and a patch size of 5×5 for an image of size 124×124. The code was compiled using Intel C/C++ Compiler (ICC) version 18.0.1; making use of different vector instruction sets by the compiler's auto-vectorization functionality. The various instruction sets are listed in Table 8.3 together with their respective vector widths. Calls to transcendental functions, that occur when computing the pixel similarities and the weights, make use of Intel's short vector math library, which provides optimized, vectorized implementations.

Benchmarks were run on an Intel i7–4770 on a single core with disabled frequency scaling. Their runtimes are depicted in Figure 8.4.

In theory with 32 bits single precision floating point numbers SSE4.2 should provide a speed up of 4 over the scalar implementation and AVX a factor of 2 over SSE4.2. The advantage of AVX2 over AVX due to FMA is more difficult to gauge beforehand. In practice with increasing computational efficiency the memory bandwidth of a system becomes a limiting

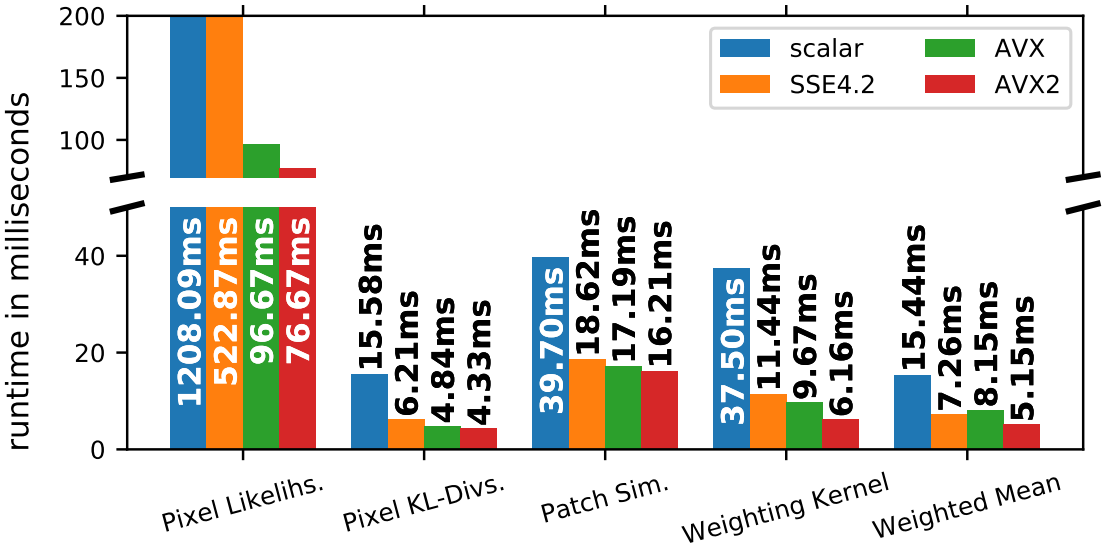


Figure 8.4: Runtimes of NL-InSAR's most relevant routines. The code was compiled with varying compiler options such that vectors of different length are used (see Table 8.3). Longer vectors speed up the computation.

factor. Additionally, due to the increased power requirements of vector instructions, processors lower their operating frequency to stay withing their thermal and power envelope. Both factors contribute to a lower speed up than the theoretical value.

Looking at the numbers two anomalies are striking: The scalar and SSE implementations of computing the pixel likelihoods are much slower than what one would expect. Also surprisingly the weighted mean implementation is slightly slower with AVX than SSE4.2. The exact reasons are difficult to pin down. When ICC compiles with a specific instruction set it also takes the underlying CPU architecture into account, which might affect the results. So the benchmark results not only show the benefits of vectorization but also how much the architecture between processor generations has improved, even though they are all executed on the same processor.

Nonetheless the vast majority of the numbers reported in Figure 8.4 fit the earlier theoretical considerations, if the non-ideal speed up is taken into account. This is also true for the algorithmic complexity analyzed in Equation (8.2). A second opportunity is the computation of the patch similarities does not dominate the overall computational cost, even though it is the only term with $\mathcal{O}(SPN)$.

Filtering of a complete TanDEM-X CoSSC, the basic input of the globally available DEM, serves as a more practical example. Running five iterations of NL-InSAR with the same parameter set used for the benchmark, using the AVX2 vectorized implementation, on all four cores of the i7–4770 took approximately 161 min for a 27418×15220 pixel interferogram. This shows that even large scale nonlocal filtering is feasible.

8.7 Further optimizations

Further optimizations are still possible. The weights of nonlocal filters are symmetric

$$w_{\mathbf{x},\mathbf{y}} = w_{\mathbf{y},\mathbf{x}}$$

so that for the most part of the nonlocal filtering algorithm only half of the computations are necessary [Dau+08].

The computation of the similarities may require the evaluation of transcendental functions. Their implementations in mathematics libraries have to trade-off between speed and accuracy usually leaning towards accuracy. As pixel values are corrupted by noise, concerning InSAR from decorrelation, speckle but also quantization, the transcendental functions' implementations could be tuned to an accuracy level for which numerical errors are negligible with respect to the noise level. Such a trade-off can be achieved by using approximation polynomials of a smaller degree[1] or using floating point representations of lower accuracy. Techniques for implementing such approximations can be found for example in [Mul16].

[1] As maybe an interesting remark, polynomials can be very efficiently implemented using the aforementioned FMA instructions

9 Conclusion

9.1 Summary

The earlier stated objective of this thesis was to develop nonlocal filtering algorithms that can cope with virtually any data set, regardless of the topography and heterogeneity of the imaged area, its temporal extension or characteristics, or the sheer size of the data in terms of volume.

- The proposed single baseline filter deals adaptively with scenes with different degrees of heterogeneity and varying height gradients. This is achieved by locally estimating the heterogeneity level and the fringe frequency and setting the filtering parameters accordingly. Several simulations show the effectiveness of these measures and help to analyze the attainable noise reduction and detail preservation, also in comparison to other state of the art nonlocal filters.
- Integrating the single baseline filter into DLR's processing chain for the global TanDEM-X DEM yields an improvement over the default configuration. The evaluation, conducted on several test sites, consisted of comparisons to aerial photographs to visually judge the preservation of details, analyzing the differences to high-resolution LiDAR DTMs and computing HEMs, an error-metric of the globally available TanDEM-X DEM.
- The chapter on multibaseline nonlocal filtering highlighted that similarity criteria need to account for different stack sizes and coherence levels. By taking these into consideration the developed nonlocal filtering framework can deal with SAR image stacks of various sizes and decorrelation trends. Sudden temporal changes are addressed by either hard or fuzzy clustering of the search volume, for which efficient routines and implementations exist. The proposed figures of merit help to quickly analyze stack filtering routines and discard unpromising research directions.
- So far the computational costs of nonlocal filters rendered them prohibitive for large-scale processing. Exploiting redundancies by removing unnecessary calculations and taking advantage of the features of modern computer architectures, as described in the last chapter, leads to substantial speedups and make nonlocal filters applicable for large data volumes.

9.2 Outlook

The results of this dissertation can have some impact on real-world applications.

By employing the developed single baseline filter and reprocessing the TanDEM-X mission's data, used to generate the globally available DEM, one can essentially move up one level in the high resolution elevation (HRE) specifications [HRE]. On a related note, some of the filter's employed techniques can readily be integrated into nonlocal filtering algorithms for SAR amplitude and PolSAR images, boosting their typical applications, such as change detection and classification.

With an optimized implementation of such filters, following the guidelines laid out as part of this thesis, all on a large or even global scale. Regarding high-resolution DEMs using the TanDEM-X data set, this would however require an even more exhaustive analysis of the proposed filter's characteristics and properties.

For SAR stacks, existing adaptive filtering methods perform already remarkably well, since temporal vectors often provide enough robustness for an accurate similarity computation. Still, employing actual 3d patches, particularly when moving to smaller stack sizes, or dealing with temporal changes can be worthwhile. In this regard the analysis on multibaseline similarity criteria and the proposed filtering framework showed promising initial results, but must also be regarded as only a first step.

9.3 Future research

There exist several avenues to further the research conducted as part of this thesis.

Adopting nonlocal filtering extensions to InSAR

The introduction of nonlocal filters already mentioned several techniques to improve the original nonlocal filter's performance, which so far have only been applied to optical and maybe SAR amplitude images. Adapting them also to InSAR will certainly be beneficial.

Sparsifying transforms, an integral part of the popular BM3D denoising filter, can counter the rare patch effect and further boost the denoising performance. Such transforms have already found their way into some InSAR related fields. For example [For+15] showed that only a few eigenvalues are necessary for accurately representing an InSAR covariance matrix. Other ways are combining nonlocal filtering with a regularization term, which this thesis alluded to in the section about DEM void filling. This technique can directly be used with HEMs for DEM generation. Learning patch priors has also been shown to be advantageous for regular optical images and in a similar fashion could help in the field of InSAR.

Filtering of dynamic SAR image stacks

Exploiting the temporal dimension is still in its infancy. Fully integrated filters with atmosphere, height and deformation estimation would further enhance the denoising performance. More sophisticated methods to deal with temporal changes, than the rudimentary approach proposed in this thesis, are also of interest. With current and future SAR satellite missions, that continuously accumulate data of long periods, of particular interest are algorithms that can deal with incrementally growing stacks without the need to reprocess all of the data.

Heterogeneity-aware similarity criteria

SAR data of very high resolution is already a reality for airborne missions and selectively for spaceborne missions [Mit+14]. With the advent of digital beamforming and consequently high resolution wide swath (HRWS) SAR systems, these types of resolution will be commonplace and the fully developed speckle assumption may no longer justified. This is already the case for heterogeneous areas, such as cities, even at lower resolutions. To ensure proper denoising in these cases more elaborate similarity criteria are needed that can deal with various radar clutter distributions.

Simulation of SAR images

The actual evaluation of denoising algorithm, just like for optical images, is a major hurdle. Acquiring ground truth data to compare against is inherently difficult as so many factors affect backscatter and decorrelation. A possible solution could be the simulation of SAR images using raytracing or computational electromagnetics, which provide the original data and corrupted versions of it. Together with appropriate quality indicators, such data sets would make a sytematic, automatic, and repeatable analysis of filtering algorithms possible.

Bibliography

[DLMF] *NIST Digital Library of Mathematical Functions*. http://dlmf.nist.gov/, Release 1.0.17 of 2017-12-22. F. W. J. Olver, A. B. Olde Daalhuis, D. W. Lozier, B. I. Schneider, R. F. Boisvert, C. W. Clark, B. R. Miller and B. V. Saunders, eds. URL: http://dlmf.nist.gov/ (cit. on p. 86).

[HRE] *Implementation Profile for High Resolution Elevation (HRE) Producs*. NGA.IP.0002_1.0. National Geospatial-Intelligence Agency. Oct. 2009 (cit. on pp. 82, 112).

[AEB06] M. Aharon, M. Elad, and A. Bruckstein. "rmK-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation". In: *IEEE Transactions on Signal Processing* 54.11 (Nov. 2006), pp. 4311–4322. ISSN: 1053-587X. DOI: 10.1109/TSP.2006.881199 (cit. on p. 34).

[Alb04] V. Alberga. "Volume decorrelation effects in polarimetric SAR interferometry". In: *IEEE Transactions on Geoscience and Remote Sensing* 42.11 (Nov. 2004), pp. 2467–2478. ISSN: 0196-2892. DOI: 10.1109/TGRS.2004.837330 (cit. on p. 22).

[Amo74] D. E. Amos. "Computation of Modified Bessel Functions and Their Ratios". In: *Mathematics of Computation* 28.125 (1974), pp. 239–251. ISSN: 00255718, 10886842. URL: http://www.jstor.org/stable/2005830 (cit. on p. 87).

[AZB17] H. Ansari, F. De Zan, and R. Bamler. "Sequential Estimator: Toward Efficient InSAR Time Series Analysis". In: *IEEE Transactions on Geoscience and Remote Sensing* 55.10 (Oct. 2017), pp. 5637–5652. ISSN: 0196-2892. DOI: 10.1109/TGRS.2017.2711037 (cit. on p. 99).

[Bai+18] G. Baier, C. Rossi, M. Lachaise, X. X. Zhu, and R. Bamler. "A Nonlocal InSAR Filter for High-Resolution DEM Generation from TanDEM-X Interferograms". In: *IEEE Transactions on Geoscience and Remote Sensing* (2018), accepted (cit. on p. 45).

[Bar+03] I. Baran, M. P. Stewart, B. M. Kampes, Z. Perski, and P. Lilly. "A modification to the Goldstein radar interferogram filter". In: *IEEE Transactions on Geoscience and Remote Sensing* 41.9 (Sept. 2003), pp. 2114–2118. ISSN: 0196-2892. DOI: 10.1109/TGRS.2003.817212 (cit. on p. 38).

[BC04] Danny Barash and Dorin Comaniciu. "A common framework for nonlinear diffusion, adaptive smoothing, bilateral filtering and mean shift". In: *Image and Vision Computing* 22.1 (2004), pp. 73–81. ISSN: 0262-8856. DOI: http://dx.doi.org/10.1016/j.imavis.2003.08.005. URL: http://www.sciencedirect.com/science/article/pii/S0262885603001732 (cit. on p. 47).

[BCM05a] A. Buades, B. Coll, and J. M. Morel. “A Review of Image Denoising Algorithms, with a New One”. In: *Multiscale Modeling & Simulation* 4.2 (2005), pp. 490–530. DOI: 10.1137/040616024. URL: https://doi.org/10.1137/040616024 (cit. on p. 27).

[BCM05b] A. Buades, B. Coll, and J.-M. Morel. “A non-local algorithm for image denoising”. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005. CVPR 2005.* Vol. 2. June 2005, pp. 60–65. DOI: 10.1109/CVPR.2005.38 (cit. on pp. 15, 27, 28, 30, 31, 33, 40, 45, 108).

[BCM06] A. Buades, B. Coll, and J.-M. Morel. “The staircasing effect in neighborhood filters and its solution”. In: *IEEE Transactions on Image Processing* 15.6 (June 2006), pp. 1499–1505. ISSN: 1057-7149. DOI: 10.1109/TIP.2006.871137 (cit. on p. 30).

[BCM08] Antoni Buades, Bartomeu Coll, and Jean-Michel Morel. “Nonlocal Image and Movie Denoising”. In: *International Journal of Computer Vision* 76.2 (Feb. 2008), pp. 123–139. ISSN: 1573-1405. DOI: 10.1007/s11263-007-0052-1. URL: https://doi.org/10.1007/s11263-007-0052-1 (cit. on p. 83).

[Bek+15] D.P.S. Bekaert, R.J. Walters, T.J. Wright, A.J. Hooper, and D.J. Parker. “Statistical comparison of InSAR tropospheric correction techniques”. In: *Remote Sensing of Environment* 170 (2015), pp. 40–47. ISSN: 0034-4257. DOI: https://doi.org/10.1016/j.rse.2015.08.035. URL: http://www.sciencedirect.com/science/article/pii/S0034425715301231 (cit. on p. 22).

[Ber+02] P. Berardino, G. Fornaro, R. Lanari, and E. Sansosti. “A New Algorithm for Surface Deformation Monitoring Based on Small Baseline Differential SAR Interferograms”. In: *Geoscience and Remote Sensing, IEEE Transactions on* 40.11 (Nov. 2002), pp. 2375–2383. ISSN: 0196-2892. DOI: 10.1109/TGRS.2002.803792 (cit. on p. 96).

[Ber+14] Ronny Bergmann, Friederike Laus, Gabriele Steidl, and Andreas Weinmann. “Second Order Differences of Cyclic Data and Applications in Variational Denoising”. In: *SIAM J. Imaging Sciences* 7 (2014), pp. 2916–2953 (cit. on p. 38).

[BH98] R. Bamler and P. Hartl. “Synthetic aperture radar interferometry”. In: *Inverse Problems* 14 (1998), R1–54 (cit. on pp. 19, 23, 24, 49).

[BKC08] T. Brox, O. Kleinschmidt, and D. Cremers. “Efficient Nonlocal Means for Denoising of Textural Patterns”. In: *IEEE Transactions on Image Processing* 17.7 (July 2008), pp. 1083–1092. ISSN: 1057-7149. DOI: 10.1109/TIP.2008.924281 (cit. on p. 103).

[BKP10] Kristian Bredies, Karl Kunisch, and Thomas Pock. “Total Generalized Variation”. In: *SIAM Journal on Imaging Sciences* 3.3 (2010), pp. 492–526. DOI: 10.1137/090769521. URL: http://dx.doi.org/10.1137/090769521 (cit. on p. 80).

[BM11] Y. Bian and B. Mercer. “Interferometric SAR Phase Filtering in the Wavelet Domain Using Simultaneous Detection and Estimation”. In: *IEEE Transactions on Geoscience and Remote Sensing* 49.4 (Apr. 2011), pp. 1396–1416. ISSN: 0196-2892. DOI: 10.1109/TGRS.2010.2076286 (cit. on p. 38).

[Bre+10] H. Breit, T. Fritz, U. Balss, A. Niedermeier, M. Eineder, N. Yague-Martinez, and C. Rossi. “Processing of bistatic TanDEM-X data”. In: *Geoscience and Remote Sensing Symposium (IGARSS), 2010 IEEE International*. July 2010, pp. 2640–2643. DOI: 10.1109/IGARSS.2010.5653602 (cit. on p. 68).

[Cal+14] F. Caltagirone, A. Capuzi, A. Coletta, G. F. De Luca, E. Scorzafava, R. Leonardi, S. Rivola, S. Fagioli, G. Angino, M. L’Abbate, M. Piemontese, E. Zampolini Faustini, A. Torre, C. De Libero, and P. G. Esposito. “The COSMO-SkyMed Dual Use Earth Observation Program: Development, Qualification, and Results of the Commissioning of the Overall Constellation”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 7.7 (July 2014), pp. 2754–2762. ISSN: 1939-1404. DOI: 10.1109/JSTARS.2014.2317287 (cit. on p. 15).

[CCN11] V. Caselles, A. Chambolle, and M. Novaga. “Total Variation in Imaging”. In: *Handbook of Mathematical Methods in Imaging*. Ed. by Otmar Scherzer. New York, NY: Springer New York, 2011, pp. 1016–1057. ISBN: 978-0-387-92920-0. DOI: 10.1007/978-0-387-92920-0_23. URL: https://doi.org/10.1007/978-0-387-92920-0_23 (cit. on p. 38).

[CD95] R. R. Coifman and D. L. Donoho. “Translation-Invariant De-Noising”. In: *Wavelets and Statistics*. Ed. by Anestis Antoniadis and Georges Oppenheim. New York, NY: Springer New York, 1995, pp. 125–150. ISBN: 978-1-4612-2544-7. DOI: 10.1007/978-1-4612-2544-7_9. URL: https://doi.org/10.1007/978-1-4612-2544-7_9 (cit. on p. 45).

[Che+11] J. Chen, Y. Chen, W. An, Y. Cui, and J. Yang. “Nonlocal Filtering for Polarimetric SAR Data: A Pretest Approach”. In: *IEEE Transactions on Geoscience and Remote Sensing* 49.5 (May 2011), pp. 1744–1754. ISSN: 0196-2892. DOI: 10.1109/TGRS.2010.2087763 (cit. on pp. 41, 88).

[Chi+16] M. Chi, A. Plaza, J. A. Benediktsson, Z. Sun, J. Shen, and Y. Zhu. “Big Data for Remote Sensing: Challenges and Opportunities”. In: *Proceedings of the IEEE* 104.11 (Nov. 2016), pp. 2207–2219. ISSN: 0018-9219. DOI: 10.1109/JPROC.2016.2598228 (cit. on p. 103).

[CM10] P. Chatterjee and P. Milanfar. “Is Denoising Dead?” In: *IEEE Transactions on Image Processing* 19.4 (Apr. 2010), pp. 895–911. ISSN: 1057-7149. DOI: 10.1109/TIP.2009.2037087 (cit. on p. 27).

[Con+03] K. Conradsen, A. A. Nielsen, J. Schou, and H. Skriver. “A test statistic in the complex Wishart distribution and its application to change detection in polarimetric SAR data”. In: *IEEE Transactions on Geoscience and Remote Sensing* 41.1 (Jan. 2003), pp. 4–19. ISSN: 0196-2892. DOI: 10.1109/TGRS.2002.808066 (cit. on pp. 40, 85).

[CP11] Antonin Chambolle and Thomas Pock. “A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging”. In: *Journal of Mathematical Imaging and Vision* 40.1 (May 2011), pp. 120–145. ISSN: 1573-7683. DOI: 10.

1007/s10851-010-0251-1. URL: https://doi.org/10.1007/s10851-010-0251-1 (cit. on pp. 38, 80).

[CW05] Ian G. Cumming and Frank H. Wong. *Digital Processing of Synthetic Aperture Radar Data: Algorithms and Implementation*. Boston, London: Artech House Inc., 2005. URL: http://www.artechhouse.com (cit. on p. 19).

[CWH11] Y. Chen, A. Wiesel, and A. O. Hero. "Robust Shrinkage Estimation of High-Dimensional Covariance Matrices". In: *IEEE Transactions on Signal Processing* 59.9 (Sept. 2011), pp. 4097–4107. ISSN: 1053-587X. DOI: 10.1109/TSP.2011.2138698 (cit. on p. 94).

[Dab+07] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering". In: *IEEE Transactions on Image Processing* 16.8 (Aug. 2007), pp. 2080–2095. ISSN: 1057-7149. DOI: 10.1109/TIP.2007.901238 (cit. on pp. 33, 34, 45, 83).

[DAG11] Vincent Duval, Jean-François Aujol, and Yann Gousseau. "A Bias-Variance Approach for the Nonlocal Means". In: *SIAM Journal on Imaging Sciences* 4.2 (2011), pp. 760–788. DOI: 10.1137/100790902. URL: https://doi.org/10.1137/100790902 (cit. on pp. 30, 32, 33, 49).

[Dar+08] J. Darbon, A. Cunha, T. F. Chan, S. Osher, and G. J. Jensen. "Fast nonlocal filtering applied to electron cryomicroscopy". In: *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. May 2008, pp. 1331–1334. DOI: 10.1109/ISBI.2008.4541250 (cit. on pp. 31, 103, 104).

[Dau+08] A Dauwe, Bart Goossens, Hiep Luong, and W Philips. "A fast non-local image denoising algorithm". In: 6812 (Jan. 2008) (cit. on p. 110).

[DDS12] Charles-Alban Deledalle, Vincent Duval, and Joseph Salmon. "Non-local Methods with Shape-Adaptive Patches (NLM-SAP)". In: *Journal of Mathematical Imaging and Vision* 43.2 (June 2012), pp. 103–120. ISSN: 1573-7683. DOI: 10.1007/s10851-011-0294-y. URL: https://doi.org/10.1007/s10851-011-0294-y (cit. on p. 32).

[DDT09] C.A. Deledalle, L. Denis, and F. Tupin. "Iterative Weighted Maximum Likelihood Denoising with Probabilistic Patch-Based Weights". In: *IEEE Transactions on Image Processing* 18.12 (2009), pp. 2661–2672. DOI: 10.1109/TIP.2009.2029593 (cit. on pp. 39, 40, 96).

[DDT11] C.-A. Deledalle, L. Denis, and F. Tupin. "NL-InSAR: Nonlocal Interferogram Estimation". In: *IEEE Transactions on Geoscience and Remote Sensing* 49.4 (Apr. 2011), pp. 1441–1452. ISSN: 0196-2892. DOI: 10.1109/TGRS.2010.2076376 (cit. on pp. 15, 39, 40, 47, 48, 103, 108).

[DDT12] Charles-Alban Deledalle, Loïc Denis, and Florence Tupin. "How to Compare Noisy Patches? Patch Similarity Beyond Gaussian Noise". In: *International Journal of Computer Vision* 99.1 (Aug. 2012), pp. 86–102. ISSN: 1573-1405. DOI: 10.1007/s11263-012-0519-6. URL: https://doi.org/10.1007/s11263-012-0519-6 (cit. on pp. 39, 86).

[DE84] William H. E. Day and Herbert Edelsbrunner. “Efficient algorithms for agglomerative hierarchical clustering methods”. In: *Journal of Classification* 1.1 (Dec. 1984), pp. 7–24. ISSN: 1432-1343. DOI: 10.1007/BF01890115. URL: https://doi.org/10.1007/BF01890115 (cit. on p. 97).

[Del+15] C.-A. Deledalle, L. Denis, F. Tupin, A. Reigber, and M. Jäger. “NL-SAR: A Unified Nonlocal Framework for Resolution-Preserving (Pol)(In)SAR Denoising”. In: *IEEE Transactions on Geoscience and Remote Sensing* 53.4 (Apr. 2015), pp. 2021–2038. ISSN: 0196-2892. DOI: 10.1109/TGRS.2014.2352555 (cit. on pp. 40, 42, 47, 88).

[Del+17] C. A. Deledalle, L. Denis, S. Tabti, and F. Tupin. “MuLoG, or How to Apply Gaussian Denoisers to Multi-Channel SAR Speckle Reduction?” In: *IEEE Transactions on Image Processing* 26.9 (Sept. 2017), pp. 4389–4403. ISSN: 1057-7149. DOI: 10.1109/TIP.2017.2713946 (cit. on p. 42).

[Den+09] L. Denis, F. Tupin, J. Darbon, and M. Sigelle. “Joint Regularization of Phase and Amplitude of InSAR Data: Application to 3-D Reconstruction”. In: *IEEE Transactions on Geoscience and Remote Sensing* 47.11 (Nov. 2009), pp. 3774–3785. ISSN: 0196-2892. DOI: 10.1109/TGRS.2009.2023668 (cit. on p. 38).

[DFE07] K. Dabov, A. Foi, and K. Egiazarian. “Video denoising by sparse 3D transform-domain collaborative filtering”. In: *2007 15th European Signal Processing Conference*. Sept. 2007, pp. 145–149 (cit. on p. 83).

[DHo+18] O. D’Hondt, C. López-Martínez, S. Guillaso, and O. Hellwich. “Nonlocal Filtering Applied to 3-D Reconstruction of Tomographic SAR Data”. In: *IEEE Transactions on Geoscience and Remote Sensing* 56.1 (Jan. 2018), pp. 272–285. ISSN: 0196-2892. DOI: 10.1109/TGRS.2017.2746420 (cit. on p. 42).

[DJ94] David L. Donoho and Iain M. Johnstone. “Ideal Spatial Adaptation by Wavelet Shrinkage”. In: *Biometrika* 81.3 (1994), pp. 425–455. ISSN: 00063444. DOI: 10.2307/2337118. URL: http://www.jstor.org/stable/2337118 (cit. on pp. 34, 38).

[DKE12] A. Danielyan, V. Katkovnik, and K. Egiazarian. “BM3D Frames and Variational Image Deblurring”. In: *IEEE Transactions on Image Processing* 21.4 (Apr. 2012), pp. 1715–1728. ISSN: 1057-7149. DOI: 10.1109/TIP.2011.2176954 (cit. on p. 34).

[DR58] W.B. Davenport and W.L. Root. *An introduction to the theory of random signals and noise*. Lincoln Laboratory publications. McGraw-Hill, 1958 (cit. on p. 21).

[EA06] M. Elad and M. Aharon. “Image Denoising Via Sparse and Redundant Representations Over Learned Dictionaries”. In: *IEEE Transactions on Image Processing* 15.12 (Dec. 2006), pp. 3736–3745. ISSN: 1057-7149. DOI: 10.1109/TIP.2006.881969 (cit. on p. 34).

[Ela02] M. Elad. “On the origin of the bilateral filter and ways to improve it”. In: *IEEE Transactions on Image Processing* 11.10 (Oct. 2002), pp. 1141–1151. ISSN: 1057-7149. DOI: 10.1109/TIP.2002.801126 (cit. on p. 28).

[ELB08] A. Elmoataz, O. Lezoray, and S. Bougleux. “Nonlocal Discrete Regularization on Weighted Graphs: A Framework for Image and Manifold Processing”. In: *IEEE Transactions on Image Processing* 17.7 (July 2008), pp. 1047–1060. ISSN: 1057-7149. DOI: 10.1109/TIP.2008.924284 (cit. on p. 34).

[FA12] K. Fletcher and European Space Agency. *Sentinel-1: ESA's Radar Observatory Mission for GMES Operational Services*. ESA SP. ESA Communications, 2012. ISBN: 978-92-9221-418-0. URL: https://earth.esa.int/web/guest/document-library/-/asset_publisher/2Xho/content/sentinel-1-esa-s-radar-observatory-mission-for-gmes-operational-services (cit. on p. 15).

[Fan+16] Dongsheng Fang, Xiaolei Lv, Yong Wang, Xue Lin, and Jiang Qian. “A Sparsity-Based InSAR Phase Denoising Algorithm Using Nonlocal Wavelet Shrinkage”. In: *Remote Sensing* 8.10 (2016). ISSN: 2072-4292. DOI: 10.3390/rs8100830. URL: http://www.mdpi.com/2072-4292/8/10/830 (cit. on p. 41).

[FB17] V. Fedorov and C. Ballester. “Affine Non-Local Means Image Denoising”. In: *IEEE Transactions on Image Processing* 26.5 (May 2017), pp. 2137–2148. ISSN: 1057-7149. DOI: 10.1109/TIP.2017.2681421 (cit. on p. 54).

[Fer+11] A. Ferretti, A. Fumagalli, F. Novali, C. Prati, F. Rocca, and A. Rucci. “A New Algorithm for Processing Interferometric Data-Stacks: SqueeSAR”. In: *IEEE Transactions on Geoscience and Remote Sensing* 49.9 (Sept. 2011), pp. 3460–3470. ISSN: 0196-2892. DOI: 10.1109/TGRS.2011.2124465 (cit. on pp. 42, 85, 86).

[Fer+18] G. Ferraioli, C. A. Deledalle, L. Denis, and F. Tupin. “Parisar: Patch-Based Estimation and Regularized Inversion for Multibaseline SAR Interferometry”. In: *IEEE Transactions on Geoscience and Remote Sensing* 56.3 (Mar. 2018), pp. 1626–1636. ISSN: 0196-2892. DOI: 10.1109/TGRS.2017.2765761 (cit. on p. 42).

[For+15] G. Fornaro, S. Verde, D. Reale, and A. Pauciullo. “CAESAR: An Approach Based on Covariance Matrix Decomposition to Improve Multibaseline-Multitemporal Interferometric SAR Processing”. In: *IEEE Transactions on Geoscience and Remote Sensing* 53.4 (Apr. 2015), pp. 2050–2065. ISSN: 0196-2892. DOI: 10.1109/TGRS.2014.2352853 (cit. on p. 112).

[FPR01] A. Ferretti, C. Prati, and F. Rocca. “Permanent scatterers in SAR interferometry”. In: *IEEE Transactions on Geoscience and Remote Sensing* 39.1 (Jan. 2001), pp. 8–20. ISSN: 0196-2892. DOI: 10.1109/36.898661 (cit. on pp. 41, 84).

[Fre+97] A. C. Frery, H. J. Muller, C. C. F. Yanasse, and S. J. S. Sant'Anna. “A model for extremely heterogeneous clutter”. In: *IEEE Transactions on Geoscience and Remote Sensing* 35.3 (May 1997), pp. 648–659. ISSN: 0196-2892. DOI: 10.1109/36.581981 (cit. on p. 21).

[Fri+11] T. Fritz, C. Rossi, N. Yague-Martinez, F. Rodriguez-Gonzalez, M. Lachaise, and H. Breit. “Interferometric processing of TanDEM-X data”. In: *Geoscience and Remote Sensing Symposium (IGARSS), 2011 IEEE International*. July 2011, pp. 2428–2431. DOI: 10.1109/IGARSS.2011.6049701 (cit. on p. 68).

[Fri+13] T. Fritz, M. Eineder, J. Mittermayer, A. Roth, E. Börner, H. Breit, B. Braeutigam, and T. Kraus. *TerraSAR-X Ground Segment Basic Product Specification Document*. Tech. rep. 1.9. German Aerospace Center (DLR), Oct. 2013. URL: https://tandemx-science.dlr.de/pdfs/TX-GS-DD-3302_Basic-Products-Specification-Document_V1.9.pdf (cit. on p. 69).

[Fuj+05] H. Fujisada, G. B. Bailey, G. G. Kelly, S. Hara, and M. J. Abrams. "ASTER DEM performance". In: *IEEE Transactions on Geoscience and Remote Sensing* 43.12 (Dec. 2005), pp. 2707–2714. ISSN: 0196-2892. DOI: 10.1109/TGRS.2005.847924 (cit. on p. 67).

[GO08] Guy Gilboa and Stanley Osher. "Nonlocal Operators with Applications to Image Processing." In: *Multiscale Modeling & Simulation* 7.3 (2008), pp. 1005–1028. URL: http://dblp.uni-trier.de/db/journals/mmas/mmas7.html#GilboaO08 (cit. on p. 34).

[Goe+16] K. Goel, N. Adam, R. Shau, and F. Rodriguez-Gonzalez. "Improving the reference network in wide-area Persistent Scatterer Interferometry for non-urban areas". In: *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. July 2016, pp. 1448–1451. DOI: 10.1109/IGARSS.2016.7729370 (cit. on p. 15).

[Goo63] N. R. Goodman. "Statistical Analysis Based on a Certain Multivariate Complex Gaussian Distribution (An Introduction)". In: *Ann. Math. Statist.* 34.1 (Mar. 1963), pp. 152–177. DOI: 10.1214/aoms/1177704250. URL: https://doi.org/10.1214/aoms/1177704250 (cit. on pp. 26, 56).

[Goo75] J. W. Goodman. "Statistical properties of laser speckle patterns". In: *Laser Speckle and Related Phenomena*. Berlin, Heidelberg: Springer Berlin Heidelberg, 1975, pp. 9–75. ISBN: 978-3-540-38013-9. DOI: 10.1007/BFb0111436 (cit. on p. 26).

[GP97] A. M. Guarnieri and C. Prati. "SAR interferometry: a "Quick and dirty" coherence estimator for data browsing". In: *IEEE Transactions on Geoscience and Remote Sensing* 35.3 (May 1997), pp. 660–669. ISSN: 0196-2892. DOI: 10.1109/36.581984 (cit. on p. 50).

[GW98] Richard M. Goldstein and Charles L. Werner. "Radar interferogram filtering for geophysical applications". In: *Geophysical Research Letters* 25.21 (1998), pp. 4035–4038. ISSN: 1944-8007. DOI: 10.1029/1998GL900033. URL: http://dx.doi.org/10.1029/1998GL900033 (cit. on p. 37).

[Hof+16] J. Hoffmann, M. Huber, U. Marschalk, A. Wendleder, B. Wessel, M. Bachmann, B. Bräutigam, T. Busche, J. Hueso González, G. Krieger, P. Rizzoli, M. Eineder, and T. Fritz. *TanDEM-X Ground Segment DEM Products Specification Document*. Tech. rep. German Aerospace Center, 2016. URL: https://tandemx-science.dlr.de/pdfs/TD-GS-PS-0021_DEM-Product-Specification_v3.1.pdf (cit. on pp. 68, 77, 78).

[HST13] K. He, J. Sun, and X. Tang. "Guided Image Filtering". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.6 (June 2013), pp. 1397–1409. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2012.213 (cit. on p. 39).

[IH17] K. Ichikawa and A. Hirose. "Singular Unit Restoration in InSAR Using Complex-Valued Neural Networks in the Spectral Domain". In: *IEEE Transactions on Geoscience and Remote Sensing* 55.3 (Mar. 2017), pp. 1717–1723. ISSN: 0196-2892. DOI: 10.1109/TGRS.2016.2630719 (cit. on p. 38).

[Jao84] Jen Jao. "Amplitude distribution of composite terrain radar clutter and the *K*-Distribution". In: *IEEE Transactions on Antennas and Propagation* 32.10 (Oct. 1984), pp. 1049–1062. ISSN: 0018-926X. DOI: 10.1109/TAP.1984.1143200 (cit. on p. 21).

[JB94] Dieter Just and Richard Bamler. "Phase statistics of interferograms with applications to synthetic aperture radar". In: *Applied optics* 33.20 (1994), pp. 4361–4368 (cit. on p. 24).

[Ji+10] Hui Ji, Chaoqiang Liu, Zuowei Shen, and Yuhong Xu. "Y.: Robust video denoising using low rank matrix completion". In: *In: CVPR, IEEE*. 2010, pp. 1791–1798 (cit. on p. 83).

[Jia+15] M. Jiang, X. Ding, R. F. Hanssen, R. Malhotra, and L. Chang. "Fast Statistically Homogeneous Pixel Selection for Covariance Matrix Estimation for Multitemporal InSAR". In: *IEEE Transactions on Geoscience and Remote Sensing* 53.3 (Mar. 2015), pp. 1213–1224. ISSN: 0196-2892. DOI: 10.1109/TGRS.2014.2336237 (cit. on p. 42).

[JPW93] I. R. Joughin, D. B. Percival, and D. P. Winebrenner. "Maximum likelihood estimation of K distribution parameters for SAR data". In: *IEEE Transactions on Geoscience and Remote Sensing* 31.5 (Sept. 1993), pp. 989–999. ISSN: 0196-2892. DOI: 10.1109/36.263769 (cit. on p. 21).

[Kat+09] Vladimir Katkovnik, Alessandro Foi, Karen Egiazarian, and Jaakko Astola. "From Local Kernel to Nonlocal Multiple-Model Image Denoising". In: *International Journal of Computer Vision* 86.1 (July 2009), p. 1. ISSN: 1573-1405. DOI: 10.1007/s11263-009-0272-7. URL: https://doi.org/10.1007/s11263-009-0272-7 (cit. on p. 27).

[KB08] Ch. Kervrann and J. Boulanger. "Local adaptivity to variable smoothness for exemplar-based image denoising and representation". In: *International Journal of Computer Vision* 79.1 (Aug. 2008), pp. 45–69 (cit. on pp. 33, 50).

[KB09a] Tamara G. Kolda and Brett W. Bader. "Tensor Decompositions and Applications". In: *SIAM Review* 51.3 (2009), pp. 455–500. DOI: 10.1137/07070111X. URL: https://doi.org/10.1137/07070111X (cit. on pp. 17, 41).

[KB09b] Tamara G. Kolda and Brett W. Bader. "Tensor Decompositions and Applications". In: *SIAM REVIEW* 51.3 (2009), pp. 455–500 (cit. on p. 54).

[Kno+11] Florian Knoll, Kristian Bredies, Thomas Pock, and Rudolf Stollberger. “Second order total generalized variation (TGV) for MRI”. In: *Magnetic Resonance in Medicine* 65.2 (2011), pp. 480–491. ISSN: 1522-2594. DOI: 10.1002/mrm.22595. URL: http://dx.doi.org/10.1002/mrm.22595 (cit. on p. 80).

[Kri+07] G. Krieger, A. Moreira, H. Fiedler, I. Hajnsek, M. Werner, M. Younis, and M. Zink. “TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry”. In: *IEEE Transactions on Geoscience and Remote Sensing* 45.11 (Nov. 2007), pp. 3317–3341. ISSN: 0196-2892. DOI: 10.1109/TGRS.2007.900693 (cit. on pp. 15, 67, 69, 78).

[Kun83] Joseph A. Kunc. “Exponential approximation of the modified Bessel function”. In: *Appl. Opt.* 22.3 (Feb. 1983), pp. 382–382. DOI: 10.1364/AO.22.000382. URL: http://ao.osa.org/abstract.cfm?URI=ao-22-3-382 (cit. on p. 87).

[Lê+14] T. T. Lê, A. M. Atto, E. Trouvé, and J. M. Nicolas. “Adaptive Multitemporal SAR Image Filtering Based on the Change Detection Matrix”. In: *IEEE Geoscience and Remote Sensing Letters* 11.10 (Oct. 2014), pp. 1826–1830. ISSN: 1545-598X. DOI: 10.1109/LGRS.2014.2311663 (cit. on p. 96).

[Leb+12] M. Lebrun, M. Colom, A. Buades, and J. M. Morel. “Secrets of image denoising cuisine”. In: *Acta Numerica* 21 (2012), pp. 475–576. DOI: 10.1017/S0962492912000062 (cit. on p. 27).

[Lee+94] Jong-Sen Lee, K. W. Hoppel, S. A. Mango, and A. R. Miller. “Intensity and phase statistics of multilook polarimetric and interferometric SAR imagery”. In: *IEEE Transactions on Geoscience and Remote Sensing* 32.5 (Sept. 1994), pp. 1017–1028. ISSN: 0196-2892. DOI: 10.1109/36.312890 (cit. on p. 24).

[Lee+98] J. S. Lee, K. P. Papathanassiou, T. L. Ainsworth, M. R. Grunes, and A. Reigber. “A new technique for noise filtering of SAR interferometric phase images”. In: *IEEE Transactions on Geoscience and Remote Sensing* 36.5 (Sept. 1998), pp. 1456–1465. ISSN: 0196-2892. DOI: 10.1109/36.718849 (cit. on pp. 37, 49, 50).

[Lee81] Jong-Sen Lee. “Speckle analysis and smoothing of synthetic aperture radar images”. In: *Computer Graphics and Image Processing* 17.1 (1981), pp. 24–32. ISSN: 0146-664X. DOI: http://dx.doi.org/10.1016/S0146-664X(81)80005-6. URL: http://www.sciencedirect.com/science/article/pii/S0146664X81800056 (cit. on p. 64).

[LF02] C. Lopez-Martinez and X. Fabregas. “Modeling and reduction of SAR interferometric phase noise in the wavelet domain”. In: *IEEE Transactions on Geoscience and Remote Sensing* 40.12 (Dec. 2002), pp. 2553–2566. ISSN: 0196-2892. DOI: 10.1109/TGRS.2002.806997 (cit. on p. 38).

[LF16] M. Lachaise and T. Fritz. “Update of the Interferometric Processing Algorithms for the Tan-DEM-X High Resolution DEMs”. In: *EUSAR 2016; Proc. of.* June 2016, pp. 1–4 (cit. on p. 68).

[LFB14] Marie Lachaise, Thomas Fritz, and Helko Breit. "InSAR processing and dual-baseline phase unwrapping for global TanDEM-X DEM generation". In: *International Geoscience and Remote Sensing Symposium (IGARSS)*. July 2014, pp. 2229–2232. ISBN: 978-1-4799-5775-0 (cit. on p. 68).

[LFB18] M. Lachaise, T. Fritz, and R. Bamler. "The Dual-Baseline Phase Unwrapping Correction Framework for the TanDEM-X Mission Part 1: Theoretical Description and Algorithms". In: *IEEE Transactions on Geoscience and Remote Sensing* 56.2 (Feb. 2018), pp. 780–798. ISSN: 0196-2892. DOI: 10.1109/TGRS.2017.2754923 (cit. on p. 68).

[Lin+15] X. Lin, F. Li, D. Meng, D. Hu, and C. Ding. "Nonlocal SAR Interferometric Phase Filtering Through Higher Order Singular Value Decomposition". In: *IEEE Geoscience and Remote Sensing Letters* 12.4 (Apr. 2015), pp. 806–810. ISSN: 1545-598X. DOI: 10.1109/LGRS.2014.2362952 (cit. on p. 41).

[Liu+15] H. Liu, R. Xiong, J. Zhang, and W. Gao. "Image denoising via adaptive soft-thresholding based on non-local samples". In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2015, pp. 484–492. DOI: 10.1109/CVPR.2015.7298646 (cit. on p. 34).

[LN11] A. Levin and B. Nadler. "Natural image denoising: Optimality and inherent bounds". In: *CVPR 2011*. June 2011, pp. 2833–2840. DOI: 10.1109/CVPR.2011.5995309 (cit. on p. 27).

[Lv+14] X. Lv, B. Yazıcı, M. Zeghal, V. Bennett, and T. Abdoun. "Joint-Scatterer Processing for Time-Series InSAR". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.11 (Nov. 2014), pp. 7205–7221. ISSN: 0196-2892. DOI: 10.1109/TGRS.2014.2309346 (cit. on p. 42).

[Mag+12] Matteo Maggioni, Giacomo Boracchi, Alessandro Foi, and Karen Egiazarian. "Video Denoising, Deblocking, and Enhancement Through Separable 4-D Nonlocal Spatiotemporal Transforms". In: *Trans. Img. Proc.* 21.9 (Sept. 2012), pp. 3952–3966. ISSN: 1057-7149. DOI: 10.1109/TIP.2012.2199324. URL: http://dx.doi.org/10.1109/TIP.2012.2199324 (cit. on p. 83).

[Mai+09a] Julien Mairal, Francis R. Bach, Jean Ponce, Guillermo Sapiro, and Andrew Zisserman. "Non-local sparse models for image restoration". In: *IEEE 12th International Conference on Computer Vision, ICCV 2009, Kyoto, Japan, September 27 - October 4, 2009*. 2009, pp. 2272–2279. DOI: 10.1109/ICCV.2009.5459452. URL: https://doi.org/10.1109/ICCV.2009.5459452 (cit. on p. 34).

[Mai+09b] Julien Mairal, Francis Bach, Jean Ponce, and Guillermo Sapiro. "Online Dictionary Learning for Sparse Coding". In: *Proceedings of the 26th Annual International Conference on Machine Learning*. ICML '09. Montreal, Quebec, Canada: ACM, 2009, pp. 689–696. ISBN: 978-1-60558-516-1. DOI: 10.1145/1553374.1553463. URL: http://doi.acm.org/10.1145/1553374.1553463 (cit. on p. 34).

[Mal03] R. K. Mallik. “On multivariate Rayleigh and exponential distributions”. In: *IEEE Transactions on Information Theory* 49.6 (June 2003), pp. 1499–1515. ISSN: 0018-9448. DOI: 10.1109/TIT.2003.811910 (cit. on pp. 86, 90).

[Mal92] H. S. Malvar. “Extended lapped transforms: properties, applications, and fast algorithms”. In: *IEEE Transactions on Signal Processing* 40.11 (Nov. 1992), pp. 2703–2714. ISSN: 1053-587X. DOI: 10.1109/78.165657 (cit. on p. 45).

[McC80] J. McClellan. “Artifacts in alpha-rooting of images”. In: *ICASSP ’80. IEEE International Conference on Acoustics, Speech, and Signal Processing*. Vol. 5. Apr. 1980, pp. 449–452. DOI: 10.1109/ICASSP.1980.1170896 (cit. on p. 37).

[Mey+06] F. Meyer, R. Bamler, N. Jakowski, and T. Fritz. “The Potential of Low-Frequency SAR Systems for Mapping Ionospheric TEC Distributions”. In: *IEEE Geoscience and Remote Sensing Letters* 3.4 (Oct. 2006), pp. 560–564. ISSN: 1545-598X. DOI: 10.1109/LGRS.2006.882148 (cit. on p. 22).

[Mil86] Gavin S P Miller. “The Definition and Rendering of Terrain Maps”. In: *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH ’86. New York, NY, USA: ACM, 1986, pp. 39–48. ISBN: 0-89791-196-2. DOI: 10.1145/15922.15890. URL: http://doi.acm.org/10.1145/15922.15890 (cit. on p. 60).

[Mit+14] J. Mittermayer, S. Wollstadt, P. Prats-Iraola, and R. Scheiber. “The TerraSAR-X Staring Spotlight Mode Concept”. In: *IEEE Transactions on Geoscience and Remote Sensing* 52.6 (June 2014), pp. 3695–3706. ISSN: 0196-2892. DOI: 10.1109/TGRS.2013.2274821 (cit. on p. 113).

[Mor+13] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. P. Papathanassiou. “A tutorial on synthetic aperture radar”. In: *IEEE Geoscience and Remote Sensing Magazine* 1.1 (Mar. 2013), pp. 6–43. ISSN: 2473-2397. DOI: 10.1109/MGRS.2013.2248301 (cit. on pp. 19, 20).

[Mor+15] A. Moreira, G. Krieger, I. Hajnsek, K. Papathanassiou, M. Younis, P. Lopez-Dekker, S. Huber, M. Villano, M. Pardini, M. Eineder, F. De Zan, and A. Parizzi. “Tandem-L: A Highly Innovative Bistatic SAR Mission for Global Observation of Dynamic Processes on the Earth’s Surface”. In: *IEEE Geoscience and Remote Sensing Magazine* 3.2 (June 2015), pp. 8–23. ISSN: 2473-2397. DOI: 10.1109/MGRS.2015.2437353 (cit. on p. 15).

[Mor00] Alberto Moreira. “Radar mit Synthetischer Apertur - Grundlagen und Signalverarbeitung”. habilitation. University of Karlsruhe, 2000 (cit. on p. 19).

[Mot+17] T. Motohka, Y. Kankaku, S. Suzuki, and M. Shimada. “Status of the advanced land observing satellite-2 (ALOS-2) and its follow-on L-band SAR mission”. In: *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. July 2017, pp. 2427–2429. DOI: 10.1109/IGARSS.2017.8127482 (cit. on p. 15).

[MS05] M. Mahmoudi and G. Sapiro. “Fast image and video denoising via nonlocal means of similar neighborhoods”. In: *IEEE Signal Processing Letters* 12.12 (Dec. 2005), pp. 839–842. ISSN: 1070-9908. DOI: 10.1109/LSP.2005.859509 (cit. on p. 103).

[MSF14] M. Maggioni, E. Sanchez-Monge, and A. Foi. “Joint Removal of Random and Fixed-Pattern Noise Through Spatiotemporal Video Filtering”. In: *IEEE Transactions on Image Processing* 23 (Oct. 2014), pp. 4282–4296. DOI: 10.1109/TIP.2014.2345261 (cit. on p. 83).

[Mul16] Jean-Michel Muller. *Elementary Functions: Algorithms and Implementation*. 3rd. Birkhäuser Basel, 2016. ISBN: 1489979816, 9781489979810 (cit. on p. 110).

[Nak60] M. Nakagami. “The m-Distribution—A General Formula of Intensity Distribution of Rapid Fading”. In: *Statistical Methods in Radio Wave Propagation*. Ed. by W.C. Hoffman. Pergamon, 1960, pp. 3–36. ISBN: 978-0-08-009306-2. DOI: https://doi.org/10.1016/B978-0-08-009306-2.50005-4 (cit. on p. 21).

[NF03] M. Nolan and D. R. Fatland. “Penetration depth as a DInSAR observable and proxy for soil moisture”. In: *IEEE Transactions on Geoscience and Remote Sensing* 41.3 (Mar. 2003), pp. 532–537. ISSN: 0196-2892. DOI: 10.1109/TGRS.2003.809931 (cit. on pp. 70, 77).

[NS92] M. Nitzberg and T. Shiota. “Nonlinear image filtering with edge and corner enhancement”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 14.8 (Aug. 1992), pp. 826–833. ISSN: 0162-8828. DOI: 10.1109/34.149593 (cit. on p. 45).

[Oll+12] E. Ollila, D. E. Tyler, V. Koivunen, and H. V. Poor. “Complex Elliptically Symmetric Distributions: Survey, New Results and Applications”. In: *IEEE Transactions on Signal Processing* 60.11 (Nov. 2012), pp. 5597–5625. ISSN: 1053-587X. DOI: 10.1109/TSP.2012.2212433 (cit. on p. 93).

[Par+09] Sylvain Paris, Pierre Kornprobst, Jack Tumblin, and Fŕedo Durand. “Bilateral filtering: Theory and applications”. English (US). In: *Foundations and Trends in Computer Graphics and Vision* 4.1 (Oct. 2009), pp. 1–73. ISSN: 1572-2740. DOI: 10.1561/0600000020 (cit. on p. 28).

[Par+12] S. Parrilli, M. Poderico, C.V. Angelino, and L. Verdoliva. “A Nonlocal SAR Image Denoising Algorithm Based on LLMMSE Wavelet Shrinkage”. In: *IEEE Transactions on Geoscience and Remote Sensing* 50.2 (Feb. 2012), pp. 606–616. ISSN: 0196-2892. DOI: 10.1109/TGRS.2011.2161586 (cit. on pp. 34, 99, 101).

[Pat07] Yale Patt. “The Transformation Hierarchy in the Era of Multi-core”. In: *High Performance Computing – HiPC 2007: 14th International Conference, Goa, India, December 18-21, 2007. Proceedings*. Ed. by Srinivas Aluru, Manish Parashar, Ramamurthy Badrinath, and Viktor K. Prasanna. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 5–5. ISBN: 978-3-540-77220-0. DOI: 10.1007/978-3-540-77220-0_4. URL: https://doi.org/10.1007/978-3-540-77220-0_4 (cit. on pp. 103, 104).

[PB11] A. Parizzi and R. Brcic. “Adaptive InSAR Stack Multilooking Exploiting Amplitude Statistics: A Comparison Between Different Techniques and Practical Results”. In: *IEEE Geoscience and Remote Sensing Letters* 8.3 (May 2011), pp. 441–445. ISSN: 1545-598X. DOI: 10.1109/LGRS.2010.2083631 (cit. on pp. 42, 85).

[PBC08] Gabriel Peyré, Sébastien Bougleux, and Laurent Cohen. “Computer Vision – ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12-18, 2008, Proceedings, Part III”. In: ed. by David Forsyth, Philip Torr, and Andrew Zisserman. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008. Chap. Non-local Regularization of Inverse Problems, pp. 57–68. ISBN: 978-3-540-88690-7. DOI: 10.1007/978-3-540-88690-7_5. URL: http://dx.doi.org/10.1007/978-3-540-88690-7_5 (cit. on p. 34).

[PCE09] A. Parizzi, X. Y. Cong, and M. Eineder. “First results from multifrequency interferometry. A comparison of different decorrelation time constants at L, C and X band”. In: *Fringe 2009 Workshop: Advances in the Science and Applications of SAR Interferometry*. 2009 (cit. on pp. 84, 86).

[PE09] M. Protter and M. Elad. “Image Sequence Denoising via Sparse and Redundant Representations”. In: *IEEE Transactions on Image Processing* 18.1 (Jan. 2009), pp. 27–35. ISSN: 1057-7149. DOI: 10.1109/TIP.2008.2008065 (cit. on p. 83).

[Per73] M. Peritsky. “Statistical Estimation of Mean Signal Strength in a Rayleigh-Fading Environment”. In: *IEEE Transactions on Communications* 21.11 (Nov. 1973), pp. 1207–1213. ISSN: 0090-6778. DOI: 10.1109/TCOM.1973.1091577 (cit. on p. 88).

[Piž17] Aleksandra Pižurica. “Image denoising algorithms : from wavelet shrinkage to nonlocal collaborative filtering”. eng. In: *Wiley encyclopedia of electrical and electronics engineering*. Ed. by Mihai Peterca. Wiley, 2017, pp. 1–17. ISBN: 9780471346081. URL: http://dx.doi.org/10.1002/047134608X.W8344 (cit. on p. 27).

[Pla+11] A. Plaza, Q. Du, Y. L. Chang, and R. L. King. “High Performance Computing for Hyperspectral Remote Sensing”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 4.3 (Sept. 2011), pp. 528–544. ISSN: 1939-1404. DOI: 10.1109/JSTARS.2010.2095495 (cit. on p. 104).

[Rab+03] Bernhard Rabus, Michael Eineder, Achim Roth, and Richard Bamler. “The shuttle radar topography mission—a new class of digital elevation models acquired by spaceborne radar”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 57.4 (2003), pp. 241–262. ISSN: 0924-2716. DOI: https://doi.org/10.1016/S0924-2716(02)00124-7. URL: http://www.sciencedirect.com/science/article/pii/S0924271602001247 (cit. on p. 67).

[RE15] C. Rossi and E. Erten. “Paddy-Rice Monitoring Using TanDEM-X”. In: *IEEE Transactions on Geoscience and Remote Sensing* 53.2 (Feb. 2015), pp. 900–910. ISSN: 0196-2892. DOI: 10.1109/TGRS.2014.2330377 (cit. on p. 73).

[Ric48] S. O. Rice. “Statistical properties of a sine wave plus random noise”. In: *The Bell System Technical Journal* 27.1 (Jan. 1948), pp. 109–157. ISSN: 0005-8580. DOI: 10.1002/j.1538-7305.1948.tb01334.x (cit. on p. 21).

[Riz+17] Paola Rizzoli, Michele Martone, Carolina Gonzalez, Christopher Wecklich, Daniela Borla Tridon, Benjamin Bräutigam, Markus Bachmann, Daniel Schulze, Thomas Fritz, Martin Huber, Birgit Wessel, Gerhard Krieger, Manfred Zink, and Alberto Moreira. “Generation and performance assessment of the global TanDEM-X digital elevation model”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 132 (2017), pp. 119–139. ISSN: 0924-2716. DOI: https://doi.org/10.1016/j.isprsjprs.2017.08.008. URL: http://www.sciencedirect.com/science/article/pii/S092427161730093X (cit. on pp. 67, 69, 103).

[ROF92] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. “Nonlinear total variation based noise removal algorithms”. In: *Physica D: Nonlinear Phenomena* 60.1 (1992), pp. 259–268. ISSN: 0167-2789. DOI: https://doi.org/10.1016/0167-2789(92)90242-F. URL: http://www.sciencedirect.com/science/article/pii/016727899290242F (cit. on pp. 34, 38).

[Ros+00] P. A. Rosen, S. Hensley, I. R. Joughin, F. K. Li, S. N. Madsen, E. Rodriguez, and R. M. Goldstein. “Synthetic aperture radar interferometry”. In: *Proceedings of the IEEE* 88.3 (Mar. 2000), pp. 333–382. ISSN: 0018-9219. DOI: 10.1109/5.838084 (cit. on p. 19).

[Ros+12] Cristian Rossi, Fernando Rodriguez Gonzalez, Thomas Fritz, Nestor Yague-Martinez, and Michael Eineder. “TanDEM-X calibrated raw DEM generation”. In: *ISPRS Journal of Photogrammetry and Remote Sensing* 73 (2012), pp. 12–20. DOI: http://dx.doi.org/10.1016/j.isprsjprs.2012.05.014. URL: http://www.sciencedirect.com/science/article/pii/S0924271612001062 (cit. on p. 68).

[Ros+17] P. Rosen, S. Hensley, S. Shaffer, W. Edelstein, Y. Kim, R. Kumar, T. Misra, R. Bhan, and R. Sagi. “The NASA-ISRO SAR (NISAR) mission dual-band radar instrument preliminary design”. In: *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*. July 2017, pp. 3832–3835. DOI: 10.1109/IGARSS.2017.8127836 (cit. on p. 15).

[Sam+16] S. Samiei-Esfahany, J. E. Martins, F. van Leijen, and R. F. Hanssen. “Phase Estimation for Distributed Scatterers in InSAR Stacks Using Integer Least Squares Estimation”. In: *IEEE Transactions on Geoscience and Remote Sensing* 54.10 (Oct. 2016), pp. 5671–5687. ISSN: 0196-2892. DOI: 10.1109/TGRS.2016.2566604 (cit. on p. 86).

[SC11] E. Strekalovskiy and D. Cremers. “Total variation for cyclic structures: Convex relaxation and efficient minimization”. In: *CVPR 2011*. June 2011, pp. 1905–1911. DOI: 10.1109/CVPR.2011.5995573 (cit. on p. 38).

[Sch+12] R. Schmogrow, B. Nebendahl, M. Winter, A. Josten, D. Hillerkuss, S. Koenig, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, J. Becker, W. Freude, and J. Leuthold. "Error Vector Magnitude as a Performance Measure for Advanced Modulation Formats". In: *IEEE Photonics Technology Letters* 24.1 (Jan. 2012), pp. 61–63. ISSN: 1041-1135. DOI: 10.1109/LPT.2011.2172405 (cit. on p. 84).

[Sch07] Satu Elisa Schaeffer. "Survey: Graph Clustering". In: *Computer Science Review* 1.1 (Aug. 2007), pp. 27–64. ISSN: 1574-0137. DOI: https://doi.org/10.1016/j.cosrev.2007.05.001. URL: http://www.sciencedirect.com/science/article/pii/S1574013707000020 (cit. on p. 96).

[SDA14] C. Sutour, C. A. Deledalle, and J. F. Aujol. "Adaptive Regularization of the NL-Means: Application to Image and Video Denoising". In: *IEEE Transactions on Image Processing* 23.8 (Aug. 2014), pp. 3506–3521. ISSN: 1057-7149. DOI: 10.1109/TIP.2014.2329448 (cit. on pp. 35, 80).

[SDT11] A. Shabou, J. Darbon, and F. Tupin. "A Markovian Approach for InSAR Phase Reconstruction With Mixed Discrete and Continuous Optimization". In: *IEEE Geoscience and Remote Sensing Letters* 8.3 (May 2011), pp. 527–531. ISSN: 1545-598X. DOI: 10.1109/LGRS.2010.2090336 (cit. on p. 38).

[Sic+15] F. Sica, D. Reale, G. Poggi, L. Verdoliva, and G. Fornaro. "Nonlocal Adaptive Multilooking in SAR Multipass Differential Interferometry". In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8.4 (Apr. 2015), pp. 1727–1742. ISSN: 1939-1404. DOI: 10.1109/JSTARS.2015.2421554 (cit. on p. 42).

[Sic+18] F. Sica, D. Cozzolino, X. X. Zhu, L. Verdoliva, and G. Poggi. "InSAR-BM3D: A Nonlocal Filter for SAR Interferometric Phase Restoration". In: *IEEE Transactions on Geoscience and Remote Sensing* PP.99 (2018), pp. 1–12. ISSN: 0196-2892. DOI: 10.1109/TGRS.2018.2800087 (cit. on p. 41).

[SLB10] Z. Suo, Z. Li, and Z. Bao. "A New Strategy to Estimate Local Fringe Frequencies for InSAR Phase Noise Reduction". In: *IEEE Geoscience and Remote Sensing Letters* 7.4 (Oct. 2010), pp. 771–775. ISSN: 1545-598X. DOI: 10.1109/LGRS.2010.2047935 (cit. on p. 53).

[SM90] M. Sekine and Y. Mao. *Weibull Radar Clutter*. Electromagnetics and Radar Series. P. Peregrinus Limited, 1990. ISBN: 9780863411915 (cit. on p. 21).

[Ste81] Charles M. Stein. "Estimation of the Mean of a Multivariate Normal Distribution". In: *Ann. Statist.* 9.6 (Nov. 1981), pp. 1135–1151. DOI: 10.1214/aos/1176345632. URL: https://doi.org/10.1214/aos/1176345632 (cit. on p. 33).

[Suo+16] Z. Suo, J. Zhang, M. Li, Q. Zhang, and C. Fang. "Improved InSAR Phase Noise Filter in Frequency Domain". In: *IEEE Transactions on Geoscience and Remote Sensing* 54.2 (Feb. 2016), pp. 1185–1195. ISSN: 0196-2892. DOI: 10.1109/TGRS.2015.2476355 (cit. on p. 38).

[SW18] M. Storath and A. Weinmann. “Fast Median Filtering for Phase or Orientation Data”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 40.3 (Mar. 2018), pp. 639–652. ISSN: 0162-8828. DOI: 10.1109/TPAMI.2017.2692779 (cit. on p. 38).

[SWA12] J. Salmon, R. Willett, and E. Arias-Castro. “A two-stage denoising filter: The preprocessed Yaroslavsky filter”. In: *2012 IEEE Statistical Signal Processing Workshop (SSP)*. Aug. 2012, pp. 464–467. DOI: 10.1109/SSP.2012.6319733 (cit. on p. 47).

[Tis+04] C. Tison, J. M. Nicolas, F. Tupin, and H. Maitre. “A new statistical model for Markovian classification of urban areas in high-resolution SAR images”. In: *IEEE Transactions on Geoscience and Remote Sensing* 42.10 (Oct. 2004), pp. 2046–2057. ISSN: 0196-2892. DOI: 10.1109/TGRS.2004.834630 (cit. on p. 21).

[TM98] C. Tomasi and R. Manduchi. “Bilateral filtering for gray and color images”. In: *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*. Jan. 1998, pp. 839–846. DOI: 10.1109/ICCV.1998.710815 (cit. on pp. 28, 30, 31).

[Tou+99] R. Touzi, A. Lopes, J. Bruniquel, and P. W. Vachon. “Coherence estimation for SAR imagery”. In: *IEEE Transactions on Geoscience and Remote Sensing* 37.1 (Jan. 1999), pp. 135–149. ISSN: 0196-2892. DOI: 10.1109/36.739146 (cit. on p. 78).

[TTT14] Junichi Takaku, Takeo Tadono, and K Tsutsui. “Generation of high resolution global DSM from ALOS PRISM”. In: XL-4 (Mar. 2014) (cit. on p. 67).

[Vas+06] G. Vasile, E. Trouve, Jong-Sen Lee, and V. Buzuloiu. “Intensity-driven adaptive-neighborhood technique for polarimetric and interferometric SAR parameters estimation”. In: *IEEE Transactions on Geoscience and Remote Sensing* 44.6 (June 2006), pp. 1609–1621. ISSN: 0196-2892. DOI: 10.1109/TGRS.2005.864142 (cit. on p. 37).

[VK09] D. Van De Ville and M. Kocher. “SURE-Based Non-Local Means”. In: *IEEE Signal Processing Letters* 16.11 (Nov. 2009), pp. 973–976. ISSN: 1070-9908. DOI: 10.1109/LSP.2009.2027669 (cit. on p. 33).

[VOK10] R. Vignesh, B. T. Oh, and C. C. J. Kuo. “Fast Non-Local Means (NLM) Computation With Probabilistic Early Termination”. In: *IEEE Signal Processing Letters* 17.3 (Mar. 2010), pp. 277–280. ISSN: 1070-9908. DOI: 10.1109/LSP.2009.2038956 (cit. on p. 103).

[Wal+14] Stéfan van der Walt, Johannes L. Schönberger, Juan Nunez-Iglesias, François Boulogne, Joshua D. Warner, Neil Yager, Emmanuelle Gouillart, and Tony and Yu. “scikit-image: image processing in Python”. In: *PeerJ* 2 (June 2014), e453. ISSN: 2167-8359. DOI: 10.7717/peerj.453. URL: https://doi.org/10.7717/peerj.453 (cit. on p. 29).

[Wan+04] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. “Image quality assessment: from error visibility to structural similarity”. In: *IEEE Transactions on Image Processing* 13.4 (Apr. 2004), pp. 600–612. ISSN: 1057-7149. DOI: 10.1109/TIP.2003.819861 (cit. on p. 41).

[Wan+18] Y. Wang, Y. Deng, R. Wang, and J. Wang. “Adaptive Multilooking Based on Complex Patch for Multitemporal Interferometry”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11.3 (Mar. 2018), pp. 907–918. ISSN: 1939-1404. DOI: 10.1109/JSTARS.2018.2795012 (cit. on p. 42).

[WB10] R. Werninghaus and S. Buckreuss. “The TerraSAR-X Mission and System Design”. In: *IEEE Transactions on Geoscience and Remote Sensing* 48.2 (Feb. 2010), pp. 606–614. ISSN: 0196-2892. DOI: 10.1109/TGRS.2009.2031062 (cit. on p. 15).

[Wei97] Joachim Weickert. “A review of nonlinear diffusion filtering”. In: *Scale-Space Theory in Computer Vision: First International Conference, Scale-Space'97 Utrecht, The Netherlands, July 2–4, 1997 Proceedings*. Ed. by Bart ter Haar Romeny, Luc Florack, Jan Koenderink, and Max Viergever. Berlin, Heidelberg: Springer Berlin Heidelberg, 1997, pp. 1–28. ISBN: 978-3-540-69196-9. DOI: 10.1007/3-540-63167-4_37. URL: http://dx.doi.org/10.1007/3-540-63167-4_37 (cit. on p. 47).

[Wen+13] A. Wendleder, B. Wessel, A. Roth, M. Breunig, K. Martin, and S. Wagenbrenner. “TanDEM-X Water Indication Mask: Generation and First Evaluation Results”. In: *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6.1 (Feb. 2013), pp. 171–179. ISSN: 1939-1404. DOI: 10.1109/JSTARS.2012.2210999 (cit. on p. 79).

[WOG06] Holger Winnemöller, Sven C. Olsen, and Bruce Gooch. “Real-Time Video Abstraction”. In: *ACM Transactions on Graphics (Proc. of SIGGRAPH'06)* 25.3 (July 2006), pp. 1221–1226. DOI: http://doi.acm.org/10.1145/1141911.1142018 (cit. on p. 47).

[WZ16] Y. Wang and X. X. Zhu. “Robust Estimators for Multipass SAR Interferometry”. In: *IEEE Transactions on Geoscience and Remote Sensing* 54.2 (Feb. 2016), pp. 968–980. ISSN: 0196-2892. DOI: 10.1109/TGRS.2015.2471303 (cit. on p. 85).

[Yan93] M.-S. Yang. “A survey of fuzzy clustering”. In: *Mathematical and Computer Modelling* 18.11 (1993), pp. 1–16. ISSN: 0895-7177. DOI: https://doi.org/10.1016/0895-7177(93)90202-A. URL: http://www.sciencedirect.com/science/article/pii/089571779390202A (cit. on p. 95).

[Yar85] L.P. Yarovslasky. *Digital Picture Processing: An Introduction*. Springer-Verlag, 1985. ISBN: 3-054-11934-5 (cit. on p. 48).

[Zhu+14] Xiao Xiang Zhu, Richard Bamler, Marie Lachaise, Fathalrahman Adam, Yilei Shi, and Michael Eineder. “Improving TanDEM-X DEMs by Non-local InSAR Filtering”. In: *EUSAR 2014; Proc. of*. June 2014 (cit. on p. 15).

[ZV92] H. A. Zebker and J. Villasenor. “Decorrelation in interferometric radar echoes”. In: *IEEE Transactions on Geoscience and Remote Sensing* 30.5 (Sept. 1992), pp. 950–959. ISSN: 0196-2892. DOI: 10.1109/36.175330 (cit. on p. 22).

List of figures

List of tables

Acknowledgments

First and foremost, I am deeply indebted to Dr. Marie Lachaise and Dr. Cristian Rossi for their help with processing TerraSAR-X and TanDEM-X CoSSCs. Without their knowledge, help, and babysitting this work would simply not have been possible. Of course, I would also like to thank my supervisor Prof. Zhu for pointing me to an interesting research topic and our scientific discussions. I also want to express my gratitude towards Prof. Bamler, who often participated in these meetings, contributing critical comments and insights. I am also grateful to Prof. Reigber, who on very short notice agreed to review my dissertation, and whose comments and questions have already sparked some new ideas.

I also want to thank in particular Dr. Schmitt, Dr. Yokoya, Dr. Körner, Dr. Ghamisi, and Dr. Werner, who helped me tremendously with advice and discussions to overcome some of the obstacle I encountered during the past years. Of course, I also need to mention Dr. Hoegner, who showed me to my surprise that social-democratic policy exists in Bavaria, and, obviously, for organizing visits to Oktoberfest.

I also had the pleasure to work alongside great people in the same predicament as me. All of them contributed to or eased, in some way or another, the work on my thesis. I would like to thank Dr. Wang, who together with Prof. Zhu introduced me to hotpot, Dr. Shazad, for sharing laughter and experience, Dr. Grohnfeldt, for always reminding me that there is stuff to do outside of a Ph.D., Nan Ge, for his German precision and accuracy, Homa Ansari, who wrote her dissertation at the same time as me, which allowed us to share our misery of this last stage, Sina Montazeri, with whom one can crack jokes about anything, Jingliang Hu, the most funny Chinese colleague one can imagine, Yao Sun, for her patience and humour, Lukas Krieger, one of the few people who can actually translate Bavarian to German, Erling Johnson, for making so many jokes during lunch, Danfeng Hong, the best guide one can have in Beijing, Jian Kang, a close run-up for being the most funny Chinese colleague, on occasion even surpassing Jingliang, Hossein Bagheri, who never seems to lose his calm, Lichao Mou, for his almost mother-like concern in Houhai, Beijing, Chunping Qiu, who never seems to be in a bad mood, but also for her strange Chinese-European fusion cuisine, Lloyd Hughes, for understanding my frustration about everything IT related, Rong Huang, who is unfazed by any obstacle or difficulty, Tobias Koch, for diversion during coffee breaks, but mostly for his IT support for others, Sandra Aigner, for her concern about my work-life balance and also company during coffee breaks, Lukas Liebel, for his contribution in finding the glorious name NL-SWAG, Yusheng Xu, for his insights into Chinese customs, politics and putting so many things into perspective, Richard Boerner, who always was a helping hand, Alexander Hanel, for explaining to me what an average geodesy student knows and is interested in and often providing a different point of view, and finally Matthias Häberle, who, due to our common cultural background, can correctly classify my degree of complaining about things. In case you have not realized it yet from reading the lines above, when doing a Ph.D. on can not wish

for a better company then these guys. I did receive an incredible amount of support and we cracked so many jokes on so many occasions.

From people who are not doing a Ph.D. but actually do some real work, I would like thank Dr. Fuh Chuo for the advice about and welcome distractions from my Ph.D., Prof. Eineder, Dr. Fritz, and Helko Breit for all things DEM and processing related, and finally Christian Minet for his insights about live in Japan.

I also need to thank Dr. Elsner and Dr. Braun, who supervised my Bachelors and Masters theses, and put me on the track of pursuing a Ph.D. in the first place.

At last, I need to thank Andre Richter, Michael Vonbun, Christian Herber, Stefan Wallentowitz, and Andreas Oeldemann for their Latex template, which made writing this thesis so much easier.