

REMOTE SENSING SERIES 44

ADRIAN SCHUBERT

Stereo-Assisted Interferometric SAR



Remote Sensing Laboratories
Department of Geography
University of Zurich, 2004

ISBN: 3-03703-010-0

REMOTE SENSING SERIES 44

ADRIAN SCHUBERT

Stereo-Assisted
Interferometric SAR



Remote Sensing Laboratories
Department of Geography
University of Zurich, 2004

REMOTE SENSING SERIES 44

ADRIAN SCHUBERT

Stereo-Assisted
Interferometric SAR



Remote Sensing Laboratories
Department of Geography
University of Zurich, 2004

Editorial Board of the Remote Sensing Series:
Prof. Dr. K. I. Itten, Prof. Dr. D. Nüesch, Dr. T. Kellenberger, Dr. E. Meier,
Dr. M. Kneubühler, and Dr. U. Frei.

*Die vorliegende Arbeit wurde von der Mathematisch-naturwissenschaftlichen
Fakultät der Universität Zürich auf Antrag von Prof. Dr. Klaus Itten und
Prof. Dr. Robert Weibel als Dissertation angenommen.*

Remote Sensing Laboratories
Department of Geography
University of Zürich
Winterthurerstrasse 190
CH-8057 Zürich
Switzerland

<http://www.geo.unizh.ch>

Copyright 2004 Adrian Schubert, University of Zürich, Switzerland

Printed by the *Druckerei der Zentralstelle der Studentenschaft der Universität
Zürich*

ISBN 3-03703-010-0

Summary

The two most reliable methods for extracting surface topography from SAR image pairs are interferometry (InSAR) and stereogrammetry (stereo SAR). The high-resolution results obtained by interferometry have been the main focus for research into digital surface model (DSM) generation in recent years. While a more accurate technique, InSAR has several weaknesses as compared to stereo SAR:

- InSAR topography estimation requires a delicate phase-unwrapping step, which is especially difficult, or even impossible, in alpine terrain.
- Ground control points (GCPs) are required for interferometric phase calibration in the form of reflectors whose positions have been measured, or tiepoints visible in the SAR images and topographic maps. This hinders the automatability of the InSAR technique.
- Areas of low coherence in interferograms, due to temporal decorrelation in the multi-pass case, vegetation presence, and steep topography lead to erroneous or altogether missing height estimates (data holes).

Although a stereo DSM will not provide nearly the height resolution of an InSAR DSM, in this thesis it is demonstrated that first processing a SAR stereo pair for the same area can assist in the phase-unwrapping, phase-calibration, and hole-filling steps of the InSAR chain. Depending on the particular scene, acquisition parameters, and processing algorithms, these three weaknesses may be more or less relevant. The stereo-assisted technique described here can address each of these points individually, making it case-adaptable.

The core of the stereo processing chain is a wavelet-based multiresolution matching algorithm. The multiresolution framework proved to be successful in coping with certain particularities of SAR images, namely, the slant-range geometry and the presence of speckle noise. However, problems were encountered for image areas dominated by radar shadow or layover. A useful secondary product is also generated during matching: the match confidence, which provides an indication of the quality of the final derived heights. This information is then used during the automatic GCP selection and hole-filling steps of the stereo-assisted InSAR processing chain.

The proposed technique is intended for use in those cases where no reliable digital elevation model (DEM) is already available, or automation of the InSAR processing chain is desired. With the increasing availability of stereoscopic and interferometric data generated by air- and spaceborne sensors, the combination of stereo SAR with InSAR will become increasingly feasible.

Three test sites were studied in the context of this work, with stereo and interferometric acquisitions obtained for each. Interferometric data were obtained over two Swiss test sites near Berne and Lucerne by a pair of X-band sensors mounted in the cross-track orientation on an aircraft. Parallel flight tracks were flown, providing the larger baseline required for stereo SAR at the same time. A spaceborne case was also studied, for the general area between Lucerne and Zurich, Switzerland. Stereo data were obtained from the European Space Agency's (ESA) ENVISAT Advanced SAR (ASAR) C-band sensor, using its swaths IS3 and IS6. InSAR data came from ESA's ERS-1/2 satellites, the C-band pair acquired in tandem mode one day apart.

DSMs were created from both the air- and spaceborne datasets using stereo and InSAR techniques individually and combined. The results were compared to the best references (DSMs, DHMs, or GCPs) available for the particular site, and the optimal combination of stereo and InSAR was determined for each case.

The limits of the stereoscopic technique were determined through simulation of the radar backscatter using a reference DSM and nominal backscatter values mapped from the DSM into radar geometry. Stereo pairs for various combinations of incidence angles were generated for one of the airborne test sites. It was determined that same-side stereo SAR, with both sensors having incidence angles of about 45 degrees, is the preferred configuration for the matching method used. The real data from the X-band airborne sensors were, in fact, obtained under nearly these conditions.

For the airborne cases, it was discovered that for scenes containing flat-to-rolling terrain, for which the interferometric phase does not vary too quickly, phase-flattening using the stereo DSM was not necessary. The stereo-based GCP generation did, however, provide the means to automatically perform the subsequent combined phase-calibration and baseline estimation. Finally, if data holes remained in the InSAR DSM, such as may be due to volume scattering within forested areas, the proposed stereo-InSAR hole-filling method was helpful in increasing the total DSM coverage without severely damaging the overall accuracy of the DSM.

The use of the stereo DSM for phase unwrapping was helpful in the case of the spaceborne data, where the fringe rates caused by alpine terrain were beyond the capability of the available phase-unwrapping software. Stereo-based phase calibration and hole-filling also proved beneficial.

For all cases studied, the stereo-assisted InSAR technique generated results superior to those obtained with InSAR alone. Finally, the combined technique requires no reflectors or searches for tiepoints, which would normally be required by InSAR alone. The method described therefore provides the basis for a fully-automatic processing chain.

Zusammenfassung

Die zwei zuverlässigsten Methoden für topographische Höhenextraktion aus SAR Bildpaaren sind die Interferometrie (InSAR) und die Stereogrammetrie (Stereo-SAR). Die hochauflösten, interferometrisch hergestellten Resultate standen in den letzten Jahren im Mittelpunkt der Forschung auf dem Gebiet der Generierung digitaler Oberflächenmodelle, (*digital surface models* oder DSMs). Obwohl InSAR die genauere Technik ist, hat sie im Vergleich zu Stereo-SAR doch mehrere Schwächen:

- Die Berechnung eines Höhenmodells mittels InSAR erfordert eine schwierige Phasenabwicklung (*phase unwrapping*), welche in alpinen Regionen oft beinahe unmöglich ist.
- Die Kalibrierung der interferometrischen Phase benötigt Kontrollpunkte (*Ground Control Points*, oder GCPs) in Form von Reflektoren, deren Positionen vermessen wurden, oder Messpunkte, die sowohl auf topographischen Karten als auch in SAR-Bildern erkennbar sind. Dadurch wird die Automatisierbarkeit von InSAR stark eingeschränkt.
- Regionen tiefer Kohärenz in Interferogrammen, die im *multi-pass* Fall aufgrund temporaler Dekorrelation, Vegetation und steiler Topographie auftreten, führen zu falschen oder sogar fehlenden Höhenwerten (Datenlöchern).

Obwohl ein Stereo-DSM nicht annähernd die Höhenauflösung eines InSAR DSMs bietet, wird in dieser Arbeit gezeigt, dass durch eine vorangehende Prozessierung eines Stereo-Paares für die gleiche Szene die Phasenabwicklung, die Phasenkalibrierung und das Beseitigen fehlender Höhenwerte die InSAR Prozessierungskette unterstützen kann. Je nach Szene, Aufnahmeparameter und Prozessierungsalgorithmen stellen diese drei Schwächen sehr oft ein Problem dar. Mit dem in dieser Arbeit vorgestellten Stereo-SAR-Verfahren ist es möglich, jeden dieser Punkte separat anzusprechen, wodurch die Technik sehr flexibel und auf viele Fälle anwendbar wird.

Der Kern der Stereo-Prozessierungskette ist ein *wavelet*-basierter *Multiresolution-Matching* Algorithmus. Das Multiresolution Verfahren stellte sich bei den SAR-Bild-spezifischen Problemen als erfolgreich heraus, insbesondere unter der Annahme der *slant-range* Geometrie (Schrägentfernung) und des *Speckle*-Rauschens. Es traten jedoch Probleme auf, wenn SAR-Bilder mit starkem Radarschatten bzw. *Layover* prozessiert wurden. Während des *Matchings* wird ausserdem ein nützliches Nebenprodukt produziert: Die sogenannte *match confidence*, welche eine

Schätzung der Qualität der Höhenwerte liefert. Diese wird nachher während der automatischen Suche nach GCPs sowie für das Füllen von Lücken verwendet.

Die vorgestellte Technik ist besonders für diejenigen Fälle geeignet, in welchen entweder kein digitales Höhenmodell (DHM) bereits vorliegt oder ein vollautomatisiertes Verfahren erwünscht ist. Weil immer mehr Stereo- und InSAR-Daten vorliegen, die auf satelliten- sowie flugzeuggestützten Systemen basieren, werden die Möglichkeiten zunehmen, die beiden Techniken zu kombinieren.

Drei Testgebiete wurden im Rahmen dieser Arbeit untersucht; in allen drei Fällen lagen Stereo- sowie InSAR-Aufnahmen vor. Interferometrische Daten wurden über zwei Gebieten in der Nähe von Bern bzw. Luzern von zwei flugzeuggestützten X-Band-Sensoren in *cross-track* Konfiguration bearbeitet. Es wurden parallele Flugtracks geflogen, deren Abstand speziell für Stereo-SAR geeignet waren. Satellitengestützte SAR-Daten wurden ebenfalls verwendet. Ein grösseres Gebiet zwischen Luzern und Zürich in der Schweiz war das dritte Testgebiet. Von diesem Gebiet wurden Stereodaten der *swaths* IS3 und IS6 des *Advanced SAR* (ASAR) Sensors verwendet. ASAR ist ein C-Band Sensor auf dem von der Europäischen Weltraumorganisation (ESA) betriebenen Satelliten ENVISAT. Die interferometrischen Daten stammen von den Fernerkundungssatelliten ERS-1 und -2 der ESA. Es wurde ein C-Band Interferometrie-Paar verwendet, welches im *tandem*-Modus mit einem zeitlichen Abstand von einem Tag aufgenommen wurde.

DSMs wurden für die flugzeug- sowie satelliten-gestützten Daten berechnet, indem Stereo und InSAR sowohl separat als auch kombiniert eingesetzt wurden. Die Resultate wurden mit den besten Referenzdaten (DSMs, DHMs, und GCPs) verglichen, die für die Testgebiete erhältlich waren. Für jeden Fall wurde die optimale Kombination von Stereo- und InSAR definiert.

Die Grenzen der stereoskopischen Technik wurden untersucht, indem Simulationen der Radarrückstreuung durchgeführt wurden. Diese Simulationen basierten auf einem Referenz-DSM und den nominalen Rückstreucoeffizienten, die vom DSM aus auf die *slant-range* Ebene projiziert wurden. Zudem wurden Stereopaare aus unterschiedlichen Kombinationen von Einfallswinkeln für den flugzeuggetragenen Sensor berechnet. Aufnahmen, die beide bei gleichen Einfalls winkeln von 45 Grad und von der selben Blickrichtung aufgenommen wurden, haben sich am besten für das vorgestellte Stereoverfahren erwiesen. Die verwendeten realen Daten von den flugzeuggestützten X-Band-Sensoren wurden in sehr guter Näherung in dieser optimalen Konfiguration aufgenommen.

In den flugzeuggestützten Fällen wurde festgestellt, dass bei der Phasenabwicklung die Unterstützung von Stereo-Daten nicht nötig war, solange das Gebiet eine flache bis leicht hügelige Oberfläche aufweist. Dies sind Gebiete mit sich nur langsam räumlich ändernden interferometrischer

Phase. Jedoch bot auch in solchen Fällen das Stereo-Verfahren die Möglichkeit, die Kalibrierung der Phase und der Basislinie automatisch durchzuführen. Falls zuletzt noch Lücken im InSAR-DSM übrig blieben, wie es zum Beispiel über Waldregionen oft der Fall sein kann, war es mit der stereo-kombinierten Methode möglich, die Löcher teilweise mit gültigen Höhenwerten zu füllen, ohne die Qualität des DSMs wesentlich zu reduzieren.

Der Einsatz des Stereo-DSMs bei der Phasenabwicklung für den satelliten-gestützten Fall hat sich hingegen als sehr vorteilhaft erwiesen, weil sich die Phase in alpinen Regionen für die vorhandene *phase-unwrapping* Software zu schnell änderte. Stereo-basierte Phasenkalibrierung und das Füllen von Lücken waren hier ebenfalls von Vorteil.

Für alle Fälle, die studiert wurden, waren die InSAR Ergebnisse, die mit Hilfe von Stereo-SAR gerechnet wurden, besser als diejenigen, die ohne Stereo-SAR prozessiert wurden. Ausserdem erfordert die beschriebene Methode weder Reflektoren noch Kontrollpunkte, die bei alleiniger Verwendung von InSAR nötig gewesen wären. Somit wurde eine wichtige Grundlage für eine vollautomatische Prozessierungskette geschaffen.

Résumé

Les deux plus grandes méthodes actuellement utilisées pour l'extraction de la topographie à partir de paires d'images radar à ouverture synthétique (*Synthetic Aperture Radar*, ou SAR) sont l'interférométrie (InSAR) et la stéréogrammétrie (*stereo SAR*). Ces dernières années, les travaux de recherche concernant la génération de modèles numériques de surface (*digital surface models*, ou DSM) ont mis l'accent sur les résultats à haute résolution de l'interférométrie. Bien qu'elle soit une technique plus précise, l'interférométrie présente certaines faiblesses par rapport à la stéréogrammétrie :

- La restitution du relief avec l'InSAR nécessite une étape délicate de déroulement de phase (*phase unwrapping*), particulièrement difficile, voire impossible, pour un terrain montagneux.
- Des points de contrôle au sol (*ground control points*, ou GCP) sont nécessaires pour l'étalonnage de la phase interférométrique, soit sous la forme de réflecteurs dont les positions ont été mesurées ou encore celle de points de repère visibles à la fois sur les images SAR ainsi que sur des cartes topographiques.
- Les zones de basse cohérence dans les interférogrammes, qui sont dues, soit à la décorrélation temporelle dans les cas *multi-pass*, soit à la présence de végétation ou d'un terrain au relief abrupt, peuvent produire des valeurs topographiques fausses ou même manquantes (trous).

Alors qu'un DSM produit avec la technique stéréoscopique n'aura nullement la précision altimétrique d'un DSM produit par l'InSAR, cette thèse démontre qu'un premier traitement d'une paire stéréoscopique pour la même région peut rendre plus aisés le déroulement et l'étalonnage de la phase, ainsi que le remplissage des valeurs manquantes pendant le traitement interférométrique. Selon la scène à travailler, les paramètres d'acquisition et les algorithmes InSAR employés, ces trois étapes seront plus ou moins délicates. La technique proposée traite chacune de ces étapes individuellement et elle est donc adaptable au cas considéré.

Le cœur du processeur stéréo est un algorithme de mise en correspondance intégré dans une représentation multirésolution par ondelettes. Le cadre multirésolution s'est avéré justifié face aux particularités des images SAR, notamment la géométrie *slant range* et la présence de bruit *speckle*. Néanmoins, des difficultés ont été rencontrées pour les régions d'images dominées par l'ombre ou le *layover*. Par ailleurs, comme résultat secondaire utile, la procédure de mise en correspondance fournit une mesure de la confiance des correspondances, qui est une indication de la qualité des valeurs

topographiques dérivées. Cette information est utilisée pendant la sélection automatique de GCP et le remplissage de trous dans le cadre de la chaîne de traitement combinée stéréo-InSAR.

La technique proposée est prévue pour les cas où aucun modèle numérique de terrain (*digital elevation model*, ou DEM) fiable n'est encore disponible, ou bien pour ceux où on désire automatiser la chaîne de traitement InSAR. Grâce à la disponibilité croissante de capteurs à bord de satellites et d'avions et donc de scènes pour lesquelles il existe des prises stéréométriques aussi bien qu'interférométriques, la combinaison de stéréo SAR avec l'InSAR sera d'autant plus envisageable dans le futur.

Trois sites d'essai ont été étudiés pour ce travail, avec des acquisitions stéréo ainsi qu'interférométriques dans chaque cas. Des données InSAR ont été obtenues au-dessus de deux sites Suisses, près de Berne et Lucerne, par deux capteurs dans la bande X montés sur un avion et orientés selon la normale à la direction de vol (*cross-track*). Des lignes de vol en parallèle ont permis l'acquisition simultanée en mode stéréo, la stéréo exigeant une base beaucoup plus importante qu'en InSAR. Un cas satellitaire a aussi été analysé pour la région générale entre Lucerne et Zurich en Suisse. Des données stéréoscopiques ont été obtenues avec le capteur *Advanced SAR* (ASAR), embarqué sur le satellite ENVISAT de l'Agence Spatiale Européenne (ESA). Les données ont été acquises dans la bande C avec les fauchées IS3 et IS6. Les prises InSAR pour la même région sont issues des satellites ERS-1 et -2 de l'ESA. La paire, en bande C, a été acquise en mode *tandem*, soit avec un intervalle d'une journée entre les deux prises.

Des DSM ont été créés pour l'ensemble des sites et des capteurs en employant les techniques stéréoscopiques et InSAR individuellement et conjointement. Les résultats ont été comparés aux meilleures références (DSM, DEM, ou GCP) disponibles pour le site en question et, dans chaque cas, la combinaison optimale entre les deux techniques a été définie.

Les limites de la technique stéréoscopique ont été déterminées par le biais de simulations d'images radar créées à partir d'un DSM de référence, basées sur des calculs liant l'amplitude du signal reçu à la superficie illuminée. Des paires stéréoscopiques ont été simulées pour plusieurs combinaisons d'angles d'incidence pour l'un des capteurs aéroportés. Pour la méthode de mise en correspondance utilisée, il a été déterminé que la configuration géométrique optimale était celle où les deux images sont prises sous un angle d'incidence d'environ 45 degrés et dans la même direction (*same-side stereo*). Les données réelles provenant du capteur aéroporté ont été acquises pratiquement dans ces mêmes conditions.

Pour les sites aéroportés, on a constaté que pour les terrains relativement plats, pour lesquels la phase interférométrique ne varie pas trop rapidement, l'utilisation du DSM stéréogrammétrique n'était pas nécessaire pour le déroulement de la phase. Par contre, la génération de GCP basée sur le processeur stéréo a permis d'automatiser l'étalonnage du modèle combiné de

la phase et de la base interférométrique. Enfin, en cas de trous dans le DSM InSAR, dus par exemple à la basse cohérence au-dessus de la végétation, la méthode de remplissage de trous basée sur la technique stéréo-InSAR combinée a aidé à compléter, au moins partiellement, le DSM sans trop dégrader la qualité générale de ce dernier.

L'utilisation du DSM-stéréo pour le déroulement de la phase interférométrique s'est avérée intéressante dans le cas de l'acquisition satellitaire, où la densité des cycles de phase dans les zones montagneuses a été trop importante pour permettre un déroulement avec les logiciels disponibles. L'étalonnage de la phase et le remplissage des trous ont aussi bénéficié de la technique combinée stéréo-InSAR.

Dans tous les cas étudiés, la technique combinée a généré de résultats supérieurs à ceux qui avaient été produits par l'interférométrie seule. En outre, la technique stéréo-InSAR n'exige ni réflecteurs ni recherche de points de repère, normalement indispensables pour l'InSAR. La méthode décrite fournit donc la base d'une chaîne de traitement entièrement automatisée.

Table of Contents

Summary	1
Zusammenfassung	3
Résumé	6
Table of Contents	9
List of Figures	13
List of Tables	19
1. Introduction and Techniques	21
1.1 Preliminary Definitions	21
1.2 Motivation for this Work.....	21
1.3 Overview of Techniques.....	23
1.4 SAR Basics	25
1.4.1 Range Compression	28
1.4.2 Azimuth Compression	29
1.5 Overview of SAR-Based Height-Extraction Methods	30
1.5.1 Single-Image Object-Height Extraction	30
1.5.2 Radarclinometry	31
1.5.3 Stereogrammetry (Stereo SAR).....	33
1.5.4 Interferometric SAR (InSAR)	34
2. Stereo Image Matching Using Wavelets	37
2.1 Introduction	37
2.2 Wavelets and Multiresolution Decomposition	38
2.2.1 Fourier vs. Wavelet Decomposition	38
2.2.2 From the Short Term Fourier Transform (STFT) to the Wavelet Transform (WT)	39
2.2.3 Wavelets and Multirate Filter Banks	41
2.3 The Dyadic Discrete Wavelet Transform Using Magarey-Kingsbury Wavelets.....	44
2.4 Stereo Matching Algorithm	48
2.4.1 Measuring Similarity	48
2.4.2 Disparity Field Error Correction	50
2.4.3 Algorithm Structure	50
3. Height Extraction Using Stereo SAR	53
3.1 Introduction	53

3.2	Rough Coregistration	53
3.3	Topographic Disparity	56
3.4	Geocoding	57
3.5	Summary of the Stereo Processing Chain	60
3.6	General Weaknesses of the Stereoscopic Method	61
4.	Interferometric Height Extraction	63
4.1	Introduction	63
4.2	Generation of a SAR Interferogram	63
4.2.1	Registration	63
4.2.2	Interferogram Calculation	64
4.2.3	Ambiguity Height	65
4.2.4	Flattening	65
4.2.5	Coherence Calculation	66
4.3	Height Map Generation	67
4.3.1	Phase Unwrapping	67
4.3.2	Geocoding	68
4.4	Weaknesses of the Interferometric Method	69
5.	Performing Interferometry Using Stereo SAR Results	71
5.1	Introduction	71
5.2	Phase Flattening	73
5.3	Ground Control Point Generation for Phase- and Baseline Calibration	73
5.4	Data Merging and Interpolation (Hole-Filling)	76
5.4.1	Data Merging	76
5.4.2	Interpolation	77
6.	SAR Image Simulation: a Virtual Testbed	79
6.1	Introduction	79
6.2	Amplitude Image Simulation	80
6.2.1	Qualitative Comparison of Simulated and Real Images	80
6.2.2	Verification of Matching Algorithms and Geocoding	81
6.2.3	Simulation of Variable Acquisition Conditions	82
6.3	Qualitative Comparison of Simulated and Real Interferograms	83
7.	Characterization of Height References, Test Sites, and Data Sets ..	85
7.1	Introduction	85
7.2	Overview of the Test Sites	85
7.3	Airborne Case: Küttigkofen	86
7.3.1	Experimental Configuration	87

7.3.2	Data Acquired	89
7.3.3	Height References	89
7.3.3.1	The DoSAR DSM from Infoterra	89
7.3.3.2	The Laser DSM from TopoSys.....	92
7.3.3.3	Ground Control Points	93
7.4	Airborne Case: Emmen	94
7.4.1	Experiment Configuration	94
7.4.2	Data Acquired.....	95
7.4.3	Height References.....	95
7.4.3.1	The DHM25 from SwissTopo	95
7.4.3.2	The DSM25 from sarmap	95
7.4.3.3	Ground Control Points	98
7.5	Spaceborne Case: Zurich.....	99
7.5.1	Experiment Configuration	99
7.5.2	Data Acquired.....	101
7.5.3	Height References.....	103
7.5.3.1	DHM25	103
7.5.3.2	Ground Control Points	103
8.	Stereogrammetry Results	105
8.1	Introduction	105
8.2	Airborne Stereo: Küttigkofen and Emmen	105
8.2.1	Stereo Matching Results	105
8.2.1.1	Evaluating Match Accuracy	106
8.2.1.2	Matching Results for Küttigkofen	106
8.2.1.3	Matching Results for Emmen	108
8.2.2	DSM Results and Discussion	109
8.2.2.1	Küttigkofen	109
8.2.2.2	Emmen	112
8.2.3	Algorithm Robustness	114
8.2.3.1	Sensitivity to Acquisition Geometry	114
8.2.3.2	Simulated versus Real Results.....	115
8.2.4	Match Confidence for Küttigkofen.....	119
8.3	Spaceborne Stereo: Zurich	122
8.3.1	Matching Results for Zurich	122
8.3.2	Error Correction versus Accuracy Retention	124
8.3.3	DSM Result for Zurich	126
8.3.4	Match Confidence.....	128
9.	Stereo-Assisted Interferometry Results	131
9.1	Introduction	131

9.2	Airborne Case: Küttigkofen	132
9.2.1	Introduction	132
9.2.2	Phase Flattening	133
9.2.3	Coherence	134
9.2.4	Ground Control Point Collection	136
9.2.5	Data Merging and Interpolation	138
9.2.6	DSM Validation	138
9.3	Airborne Case: Emmen	142
9.3.1	Introduction	142
9.3.2	Phase Flattening	142
9.3.3	Coherence	143
9.3.4	Ground Control Point Collection	144
9.3.5	DSM Validation	144
9.4	Spaceborne Case: Zurich	147
9.4.1	Introduction	147
9.4.2	Phase Flattening	148
9.4.3	Coherence	150
9.4.4	Ground Control Point Collection	150
9.4.5	Data Merging and Interpolation	151
9.4.6	DSM Validation	151
10.	Conclusions and Recommendations	157
10.1	Stereo SAR	157
10.1.1	Rough Coregistration	157
10.1.2	Matching	158
10.1.3	Radar Shadow and Layover	160
10.2	Stereo-Assisted Interferometry	160
10.2.1	Choice of Approach	160
10.2.2	Choice of Baseline	161
10.2.3	Phase Unwrapping	161
10.2.4	Automatic Ground Control Point Selection	162
10.2.5	Interpolation of Large Holes.....	162
10.2.6	Robustness of Stereo-Assisted InSAR	163
10.2.7	Final Remarks and Outlook	163
Appendix A:	List of Symbols	165
Appendix B:	Acronyms and Abbreviations	168
References	170
Acknowledgements	178

List of Figures

Figure 1.1	Geometry of a SAR system	27
Figure 1.2	Stages of SAR data processing: (a) raw data (b) range-compressed (c) range- and azimuth-compressed data	29
Figure 1.3	Object-height calculation using radar shadow length, and associated problems.....	31
Figure 1.4	Typical stereo SAR configuration	34
Figure 1.5	InSAR principle	35
Figure 1.6	Formation of interferometric fringes: (a) for a flat surface (b) for a varying surface	36
Figure 2.1	Difference between: (a) short-time Fourier analysis (b) wavelet analysis	40
Figure 2.2	Single wavelet transform of a SAR image: (a) input image (b) wavelet transform of input	44
Figure 2.3	Approximate filter alignments for the M-K wavelet basis	46
Figure 2.4	Greyscale representation of M-K mother wavelets: (a) real part (b) imaginary part	46
Figure 2.5	Two iterations of a DDWT on an image	48
Figure 2.6	The similarity distance (SD) surface	49
Figure 2.7	Multi-resolution matching: algorithm structure	51
Figure 3.1	The stereo window in range geometry	54

Figure 3.2	Rough coregistration: remapping the slave's range geometry and radiometry to those of the master: (a) before rough coregistration (b) after rough coregistration	55
Figure 3.3	The origin of topographic range disparity	57
Figure 3.4	Solution of the range-Doppler equations for a point P1 viewed by slave (S) and master (M) sensors.	59
Figure 5.1	Stereo-assisted InSAR processing chain	74
Figure 6.1	Real image and simulation of the same scene: (a) real image (b) simulated image	81
Figure 6.2	Variable acquisition conditions: (a) variable incidence angles (b) variable terrain types	83
Figure 6.3	Real interferogram and simulation for the same scene: (a) real interferogram (b) simulated interferogram	84
Figure 7.1	Overview of the test sites	86
Figure 7.2	AeS-1 system and choice of baselines	87
Figure 7.3	Küttigkofen flight geometry	88
Figure 7.4	The Küttigkofen test site, with master and stereo slave scene perimeters	90
Figure 7.5	Küttigkofen SAR slant-range amplitude images: (a) stereo master, also used for InSAR (b) stereo slave	90
Figure 7.6	The DoSAR DSM for Küttigkofen	91
Figure 7.7	Difference between the DoSAR DSM and DHM25 for Küttigkofen	91
Figure 7.8	The laser DSM for Küttigkofen	93
Figure 7.9	Reference GCP positions used for Küttigkofen.....	93

Figure 7.10	Emmen flight geometry	94
Figure 7.11	The Emmen test site, with master and stereo slave scene perimeters.....	96
Figure 7.12	Emmen SAR slant-range amplitude images: (a) stereo master, also used for InSAR (b) stereo slave	96
Figure 7.13	(a) DHM25 height model reference for Emmen (b) DSM25 for Emmen.....	97
Figure 7.14	Difference between DSM25 and DHM25 for Emmen	98
Figure 7.15	Spot height locations for Emmen	99
Figure 7.16	Zurich flight geometry	100
Figure 7.17	The Zurich stereo-InSAR test site	101
Figure 7.18	Zurich ENVISAT stereo-pair slant-range amplitudes: (a) master (b) slave.....	102
Figure 7.19	Zurich ERS-1 SAR slant-range amplitude (ERS-1)	102
Figure 7.20	DHM25 for the approximate Zurich region of stereo-InSAR overlap	103
Figure 7.21	GCP locations for the Zurich ERS-1/2 interferometric pair... 104	104
Figure 8.1	Red-green overlay between stereo master and resampled slave for Küttigkofen	108
Figure 8.2	Red-green overlay between stereo master and resampled slave for Emmen, with an example of a feature not visible in the slave (the red cross)	109
Figure 8.3	Estimated stereo DSM for Küttigkofen	110
Figure 8.4	Difference between Küttigkofen stereo DSM and (a) DoSAR DSM (b) laser DSM	111

Figure 8.5 Estimated stereo DSM for Emmen	112
Figure 8.6 Difference between Emmen stereo DSM and (a) DHM25 (b) DSM25	113
Figure 8.7 Close-up of real and simulated images: (a) real master (b) real remapped-slave (c) simulated master (d) simulated remapped slave	118
Figure 8.8 Küttigkofen stereo results: (a) match confidence (bright areas indicate high confidence) versus (b) height error (relative to the DoSAR reference).	120
Figure 8.9 Absolute height error versus match confidence for the Küttigkofen stereo DSM	122
Figure 8.10 Red-green overlay of master and remapped slave with two zooms (Zurich)	123
Figure 8.11 Effect of the match-smoothness parameter Ω on the DSM error statistics using the ENVISAT-ASAR stereo pair	125
Figure 8.12 (a) Estimated stereo DSM for Zurich (b) DHM25 reference for same perimeter	126
Figure 8.13 Difference between Zurich stereo DSM and DHM25	127
Figure 8.14 Zurich stereo results: (a) match confidence (b) height error...	129
Figure 8.15 Absolute height error versus match confidence for the Zurich stereo DSM.	130.
Figure 9.1 Original and flattened phase for Küttigkofen: (a) unflattened phase (b) flattened using an ellipsoid-Earth model (c) stereo-DSM-flattened phase	135
Figure 9.2 Küttigkofen coherence	136
Figure 9.3 Locations of automatically-extracted GCPs for Küttigkofen .	137

Figure 9.4 Data merging and interpolation ("smart"-filling) for the ellFlat_sterCal_smartFill case in Köttigkofen (a) after ellipsoid flattening and stereo-based phase calibration (b) after merging (for area covered by stereo-DSM) (c) after cubic interpolation of the smaller holes	139
Figure 9.5 Köttigkofen stereo-assisted InSAR-DSM height error relative to DoSAR, for all seven experiments	140
Figure 9.6 Original and flattened phase for Emmen: (a) unflattened phase (b) flattened using an ellipsoid-Earth model (c) stereo-DSM-flattened phase.....	143
Figure 9.7 Emmen coherence	144
Figure 9.8 Locations of automatically-extracted GCPs for Emmen. Background is AeS-1 slant-range stereo master amplitude.....	145
Figure 9.9 Emmen stereo-InSAR DSM, generated by ellipsoid-flattening, stereo phase calibration, and cubic interpolation of the small holes.....	145
Figure 9.10 Emmen ellipsoid-flattened InSAR-DSM results relative to DHM25 and DSM25: (a) map-calibrated case relative to DHM25 (b) map-calibrated case relative to DSM25 (c) stereo-calibrated case relative to DHM25 (d) stereo-calibrated case relative to DSM25	146
Figure 9.11 Phase flattening of the Zurich ERS-1/2 pair: (a) unflattened (b) ellipsoid-flattened (c) DHM25-flattened (d) stereo DSM-flattened	149
Figure 9.12 Zurich ERS-1/2 coherence	150
Figure 9.13 Locations of 44 Zurich GCPs used for phase calibration	151
Figure 9.14 Data merging and interpolation steps ("smart"-filling) for Zurich (a) stereo-based flattening and phase calibration leaves large holes (b) stereo values are used to partially close the holes (c) interpolation closes up the smallest of the remaining holes	152

Figure 9.15 Comparison of stereo-assisted and unassisted InSAR DSMs for Zurich, relative to DHM25: (a) ellipsoid-flattened, map-calibrated (unassisted) (b) stereo-flattened and -calibrated, (c) stereo-flattened, -calibrated, and -filled..... 153

List of Tables

Table 7.1	AeS-1 dual-antenna airborne SAR system parameters	88
Table 8.1	Küttigkofen stereo-DSM error statistics	111
Table 8.2	Emmen stereo-DSM error statistics	114
Table 8.3	Simulation results for various incidence angles	116
Table 8.4	Zurich stereo-DSM error statistics	128
Table 9.1	Küttigkofen stereo-InSAR experiment configurations.....	133
Table 9.2	Error statistics for Küttigkofen InSAR experiments	141
Table 9.3	Error statistics for Zurich InSAR experiments	155

Chapter 1

Introduction and Techniques

1.1 Preliminary Definitions

Digital Height Models (DHMs), also called *Digital Elevation Models* (DEMs), have a number of uses in a wide range of fields. Several examples are land mapping, transportation and military applications, hydrological studies, recreation and education, to name but a few. Specific applications making use of DEMs include the generation of digital orthophotos and the correction of geometric and radiometric distortions in remotely-sensed datasets.

There is some debate as to precise meaning of the terms *DHM* and *DEM*; the term *DEM* will be used in this work when referring to height models of unspecified (or at best, ambiguous) characteristics [63]. The results produced during this work are better classified as *Digital Surface Models*. A Digital Surface Model (DSM) is a DEM that describes the topography of the *true* surface, which includes the upper surfaces of buildings and forest stands, or anything else elevated above the bare earth [63]. This is in contrast to a *Digital Terrain Model* (DTM), which describes the bare-earth variation without the consideration of vegetation or man-made objects. A DEM of any sort, as defined within the context of this work, takes the form of a matrix formed of square pixels, whose values are terrain elevations above some reference height model, such as a geoid or ellipsoid model of the Earth's surface. The DSMs produced here are *geocoded*, that is, the pixel raster is laid out in map geometry; the pixel positions are referenced using eastings and northings (meters), and a constant pixel spacing is assumed.

1.2 Motivation for this Work

Of all existing techniques for DEM estimation using SAR technology, interferometry (InSAR) produces the most accurate DEMs. It relies on the phase difference across two antennas of an incoming signal reflected from the topographic surface to deduce the distance, or range, from the sensor to the

surface with an accuracy on the order of the sensor wavelength. While its strength is rooted in the phase, its reliance on the phase *coherence* is also its weak point. When situations arise that cause the two arriving signals to decorrelate significantly with respect to one another, no information can be extracted from the phase difference, effectively rendering the measurement useless. In practice, DEMs based on InSAR are often incomplete due to phase coherence loss over particular areas. The final result is visible as missing data, or holes, in the end product.

Holes can also be the result of difficulties resolving certain phase-related ambiguities in areas of rapidly-changing topography such as is typical of mountainous regions.

In addition to the holes that are possible with interferometric DEMs, some effort is required to attain the final product. In particular, ground control points (GCPs) are required as height references for the conversion of the phase differences to topographic heights. These can be obtained by placing reflectors with carefully-measured positions (using *differential GPS*, or D-GPS) for example; see [47]) into open areas to be imaged by the radar, or they can be collected from precise topographic maps for points clearly visible in both the SAR images and on the maps. Both versions require manual intervention, with the time and cost rising with increasing accuracy requirements.

The problems described above can, in many cases, be resolved or at least reduced by combining InSAR with an older, less accurate height-estimation method: stereogrammetry, or stereo SAR.

Stereogrammetry is oblivious to the phase of the returned signal: only the amplitude is required for DEM estimation. This means that phase ambiguities and coherence loss are of no significance to a stereo processor. It produces height maps without holes, since it relies entirely on the matching of the amplitude images acquired from two points of view.

A further advantage of stereo is that it is automatable; DSMs can be produced without resorting to maps or placing reflectors in the field.

The weakness of stereo SAR in comparison to InSAR lies almost entirely in the accuracy of the resulting DEMs: a stereo-estimated height model will typically have a height resolution on the order of one tenth that of an InSAR model. The results are hole-free but inaccurate.

With the ever-increasing availability of SAR sensors in orbit and mounted on airborne platforms, the potential to combine the two techniques is also increasing.

The goal of the technique presented here is to reinforce InSAR with stereo

SAR in such a way as to address the weaknesses of the former. Using the proposed combined method it is possible to:

- take advantage of the height accuracy obtainable by InSAR
- reduce the number of holes in the final DSM
- make InSAR automatable

This is accomplished using the output of the stereo processor described in the following two chapters. The lower-quality stereo DSM generated with such a processor can be used to:

- 1) **initialize the interferometric height estimation process,**
- 2) **calibrate the interferometric phase automatically, and**
- 3) **provide height data for areas where none were obtainable by InSAR.**

These three points form the basis for the stereo-InSAR combined technique described in this thesis.

1.3 Overview of Techniques

A number of common techniques exist for the generation of height models using optical and radar imaging systems:

- **Stereoscopic photogrammetry:** This is the oldest of the techniques used to generate DHMs. Currently, it involves the acquisition of digital photogrammetric stereo pairs, combined with manual as well as automatic stereo matching techniques for the subsequent topographic reconstruction. Sensors may be air- or spaceborne, and the wavelengths used range from visible to near-infrared. A prime example from the spaceborne side is the French SPOT 5 satellite, whose High Resolution Stereoscopic (HRS) instrument is capable of acquiring near-instantaneous stereo views providing subsequent DSM height accuracies over open ground of 5 m [45]. The major difficulty with photogrammetric techniques is the requirement that the ground be visible, which excludes acquisition during the night and through cloud cover.
- **Laser scanning or LIDAR:** LIDAR stands for *Light Detection And Ranging*. Topographic LIDAR involves use of an airborne, pulsed laser beam, which is carried over the surface to be measured. Its return signals provide direct distance measurements based on the speed of light. Over forests, first- and last-pulse returns provide information about the canopy and ground heights, respectively. Height accuracies as good as 10 cm can

be attained, depending on the flight height and other acquisition characteristics [3][53]. As with photogrammetry, the major constraint is ground visibility. However, because it is an active system, it can be operated at night. It typically employs a near-infrared laser.

- **Interferometric Synthetic Aperture Radar (InSAR):** InSAR is a technique whereby the phase difference between two separate SAR acquisitions are obtained and combined using one or more air- or spaceborne sensors. It was first applied to topographical mapping by Graham [38]. The two acquisitions may be from two antennas mounted on a single platform (single-pass InSAR) or two separate acquisitions of a single-antenna system. Interference bands caused by the overlapping of the radar signals returning to the sensor(s) are transformed into topographic variations at the surface. Because of the operating frequencies (typically ~ 1 to 10 GHz) and the nature of active systems, this technique can be used at night as well as in the presence of clouds. Typical height accuracies achievable are ~ 10 m for spaceborne InSAR [78][87], and ~ 1 m using an airborne system [64].
- **Stereoscopic SAR, or stereo SAR:** This technique is similar to stereoscopic photogrammetry, but uses SAR images as input. It relies on radiometric matching of the radar amplitude values for the subsequent geometric reconstruction of the surface. Baselines are much larger than those used for InSAR, and matching strategies are similar to those employed for photogrammetric surface reconstruction. It has been studied in detail by Leberl [50]. Height accuracies are strongly dependent on the ability of the processor to accurately coregister the images; typical accuracies of ~ 30 - 100 m are reported for spaceborne stereo SAR [87], and ~ 5 - 50 m for airborne sensors, depending on the height variability of the terrain [85][50]. Indeed, the typical height errors over flat terrain are on the order of the pixel dimension in the sensor look direction (range resolution), whether space- or airborne.
- **Radargrammetry:** This term actually describes several techniques for topography extraction, usually employing single SAR amplitude images. Good examples are: *radarclinometry*, or *shape-from-shading* [92][50], which relates radiometric to topographic variation, and shadow-based methods that make use of radar shadow to calculate object heights, such as described in [48].

Radar-based methods are active, generally weather-independent, high-resolution techniques. Airborne InSAR is capable of providing DSM height accuracies between 10 cm and several meters [64], depending on the system configuration, terrain variability, and processing methods. Stereo SAR

provides a much lower height resolution, but because the types of errors present in InSAR DSMs are generally not present in those obtained by stereo SAR, the techniques can be combined to improve the InSAR result.

Although terrestrial topographic mapping using InSAR was described by Graham in 1974 [38], it was not until 1986 that NASA's Jet Propulsion Laboratory (JPL) conducted the first experiments in Pasadena, California. The studies, published by Zebker and Goldstein [93], involved two SAR antennas mounted on an aircraft. The results demonstrated the usefulness as well as the limitations of the technique for topographic mapping, and paved the way for further studies in the following years. These included crossed-orbit [29], differential [30], and multibaseline [51] InSAR.

With the launch of the ERS-1 spaceborne SAR system in 1991, it became possible for the international research community to study the applications and limitations of spaceborne InSAR. The addition of a second ERS sensor in nearly the same orbit in 1995, called ERS-2, made datasets taken one day apart available for more detailed studies including temporal and atmospheric effects. SAR interferometry had become a widespread topic of research.

Canada's RADARSAT satellites and the Advanced SAR (ASAR) instrument on the recently-launched European ENVISAT satellite are further examples of spaceborne radar sensors that are currently generating more interest than ever in SAR applications. Additionally, high-resolution data is available from a number of airborne sensors, such as the data from Intermap Technologies' AeS-1 system (see [43]) used in this work. Airborne sensors have the additional advantage over spaceborne sensors of being more immediately available and configurable, as well as less costly to manage.

1.4

SAR Basics

As an introduction to both the theory and the terminology used in this work, the acquisition and focusing of a SAR image will be briefly described in this section. It is assumed that the basic theory and terminology related to Side-Looking Airborne Radar (SLAR) has already been understood (see [21] or [88] for example). In particular, the range resolution of a SLAR depends only on the pulse length. Resolution in azimuth depends on the physical length of the antenna in the azimuth direction for a given wavelength, sensor orientation, and sensor height.

While the range resolution of a SAR system is fundamentally the same as for a SLAR, the use of pulse compression, or chirping, permits the range resolution to be improved without severely affecting the signal to noise ratio

(SNR); this is described in Section 1.4.1 on page 28.

The major advantage of using a SAR over a SLAR system is the drastically-improved azimuth resolution obtainable using an antenna that is synthesized along the flight direction. This antenna-aperture synthesis is performed for a given point P on the ground by integrating its returns to the sensor as the beam footprint sweeps over it. The goal of range- and azimuth compression, described in the following sections, is to perform this integration first in range, then in azimuth. The azimuth (along-track) resolution on the ground of a real antenna ρ_{real} increases (i.e. worsens) with its height (H) and wavelength (λ), and decreases with its length (L) and the cosine of the beam incidence angle (θ_i), as shown in Equation 1.1. The azimuth resolution for a synthesized antenna ρ_{SAR} , however, depends solely on the antenna length (Equation 1.2) [17]. In other words, its azimuth resolution is theoretically independent of the sensor height. Also, the smaller the antenna, the better one would expect the resolution to be. These counter-intuitive conclusions can be understood as follows [21]:

- A smaller antenna creates a larger the beam opening, hence a longer synthetic azimuth array.
- A higher sensor produces a wider beam azimuth footprint on the ground, also resulting in a longer synthetic array.

$$\rho_{\text{real}} = \frac{H \cdot \lambda}{L \cdot \cos \theta_i} \quad (1.1)$$

$$\rho_{\text{SAR}} = \frac{L}{2} \quad (1.2)$$

The azimuth resolution of Equation 1.2, while theoretically correct, can never be fully achieved in reality. This is due in part to a limit on how well the Doppler frequency estimation can be performed ([17], section 1.2.2), and in part because of a lower limit on the real antenna area, thus constraining its dimensions ([17], section 1.2.1). A SAR system must also obey the far-field approximation to achieve proper focusing [82]; this places a lower limit on the antenna azimuth length. Nonetheless, the $L/2$ result serves to demonstrate the fundamental difference between a SLAR and a SAR system.

Figure 1.1 illustrates the geometry of such a system. Several pulses, or bursts, are shown on their way to the ground at the same time. An airborne system, being much nearer to the ground than a spaceborne one, waits for all returns from a pulse to be received before sending out the next pulse. Spaceborne sensors send several pulses in the time it takes for one pulse to be returned from the ground to achieve the desired resolution, due to their greater

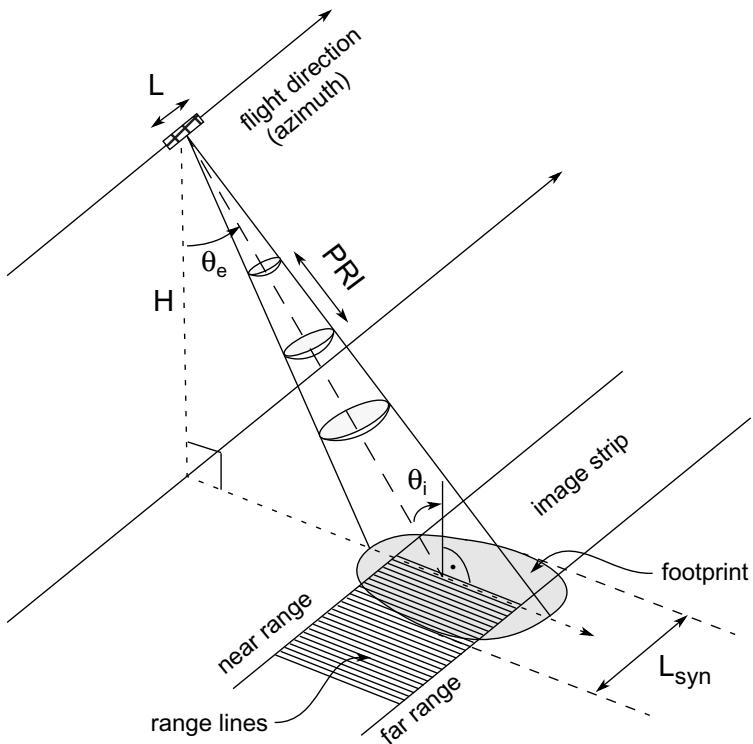


Figure 1.1 Geometry of a SAR system

altitude (for example, ENVISAT's ASAR system has about ten pulses underway at a given time, in image mode).

The meaning of the labels are as follows:

L	Azimuth length of real antenna
L_{syn}	Azimuth length of synthetic aperture
θ_e	Elevation angle
θ_i	Incidence angle
H	Sensor flight height
PRI	Pulse Repetition Interval (time between pulses)

Note that the range sample acquisition for a single burst begins and ends well within the range bounds of the footprint. This serves two purposes: (1) it guarantees a relatively constant signal-to-noise ratio (SNR) over the whole range line (i.e. no range bins containing zero signal), and (2) the azimuth

characteristics will be more consistent in range, since the somewhat deformed near- and far-range beam areas are being ignored.

Coherent electromagnetic waves have two characteristics: amplitude and phase [83]. The latter is what gives them their wave-like nature, and results in the interference pattern observed when the radar images are combined interferometrically. The raw data values, organized in the focused image as a raster grid of range lines, are the received power and phase for the given resolution cell - one power-phase pair per pixel - for all ranges and azimuth positions in the raster. Figure 1.2(a) shows an example of a raw-data raster for the city of Zurich, Switzerland and surrounding region. The data were recorded on the ERS-1 spaceborne sensor at C-band, and have a ground resolution of 25 m covering 50 x 50 km. The greyscale images ignore the phase information; they represent the amplitude of the received signal only.

1.4.1 Range Compression

In order to increase the range resolution without compromising the SNR, pulse compression is used [17]. This involves the use of *chirped*, or linearly frequency-modulated radar pulses, described by the function $s(t)$ (CHIRP also means Coherent Integration of Radar Pulses). The use of chirped pulses allow overlapping returns from the ground to be separated in frequency space, thus improving the ability to separate objects in range. In order to recover the terrain reflectivity function, given the received pulse echoes, the echoes for a given range line are convolved with the time-reversed, complex-conjugate chirp signal [17], called the *reference function*, or $s^*(-t)$; this procedure is performed for all range lines and is called *range compression*:

$$R(t) = E(t) \otimes s^*(-t) \quad (1.3)$$

with

- | | |
|-----------|--|
| $R(t)$ | range-compressed signal for a single burst (range line) |
| $E(t)$ | unprocessed received signal |
| $s^*(-t)$ | time-reversed, complex-conjugated chirp signal (reference) |

and \otimes is the convolution operator. This operation, performed on all range lines, produces a range-compressed image as the output. An example of this can be seen in Figure 1.2(b).

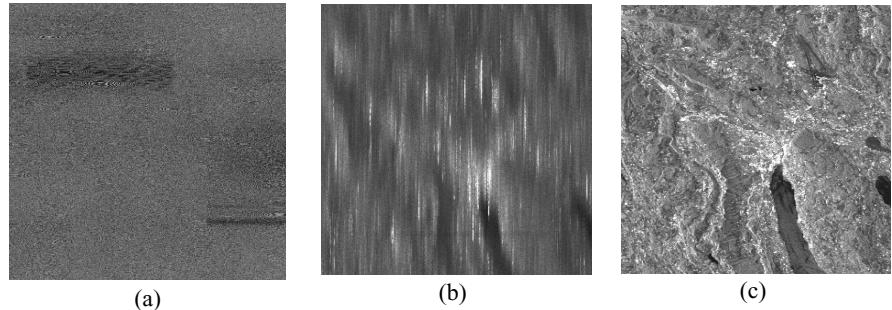


Figure 1.2 Stages of SAR data processing: (a) raw data (b) range-compressed (c) range- and azimuth-compressed data

1.4.2 Azimuth compression

While both a SLAR and a SAR system consist of a real antenna of similar dimensions, a SAR system integrates the real signal returned to the antenna along the azimuth direction, effectively simulating a much longer antenna. That is, all echoes received from a given reflector on the ground need to be coherently summed along the entire length of the beam footprint length along the flight direction. Although not obvious in Figure 1.2(b), all reflecting locations on the ground generate signals that not only vary in azimuth over the synthetic antenna length L_{syn} , but also in range. The sensor-to-point distance varies hyperbolically over L_{syn} , as the sensor approaches and then recedes from a given ground location. This creates a range dependency that must be taken into account during azimuth focusing.

Azimuth compression is performed by multiplying the range-compressed 2-D input by a reference function, itself also two-dimensional. This reference function in azimuth is the inverse of the Doppler-shift history for a single reflecting point over the length L_{syn} for all ranges [17]. The range-compressed image is convolved with the reference function as follows [82]:

$$I(t, a) = R(t, a) \otimes A^*(-t, -a) \quad (1.4)$$

with

$I(t, a)$ azimuth-compressed data (2-D)

$R(t, a)$ range-compressed input (2-D)

$A^*(-t, -a)$ reference signal (2-D)

The result of azimuth- and range-compression is called a *slant-range image*; an example of the result can be seen in Figure 1.2(c). The horizontal

direction represents range, the vertical edge the sensor flight direction. In this case the first acquired pixel is at the lower left. If no post-processing has been performed beyond simple range- and azimuth compression, this type of product is called a single-look complex (SLC) image. At this point, the image has been focused, and is ready for higher-level interpretation. For further information on azimuth compression algorithms, please refer to [4], [10],[12], and [16] for example.

The images in Figure 1.2 show the amplitude of the complex signal which is recorded. While the amplitude is all that is required for radiometric image matching of a stereo pair, InSAR requires the phase of the returned signals (section 1.5.4 and chapter 4 describe InSAR and interferometric height model estimation in greater detail).

1.5 Overview of SAR-Based Height-Extraction Methods

An overview of the most common methods currently used for surface topography estimation using air- or spaceborne radar sensors is given here.

1.5.1 Single-Image Object-Height Extraction

The oldest and simplest form of height extraction requires a single amplitude image of a scene, and involves the direct calculation of an object's height based on the length of the radar shadow produced by it. Various authors have attempted to use this method for the extraction of building heights [8][31][70][77], and even forest stand heights [48]. Figure 1.3(a) illustrates the basic concept. This method is subject to a number of severe restrictions, namely

- for a given object on the ground, its full shadow length must be visible (i.e. no interference by nearby buildings, trees, etc.),
- the tip/edge of the shadow must be cast by the "top" of the object being measured; for non-flat rooftops for example, the true maximum height may not be responsible for the tip of the shadow (Figure 1.3(b)),
- the length of the shadow also depends on the local ground slope. Without knowledge of the latter, one cannot be certain that the calculated height will be correct, since local slope alters the shadow length (Figure 1.3(c)).
- as shown in Figure 1.3(d) and (e), a more general restriction related to the use of single SAR images for object-height extraction exists. An object's shadow length increases at higher altitudes, even while remaining at the same absolute range position r . Similarly, the projection of h onto the range plane will decrease. Hence, without *a priori* knowledge of the

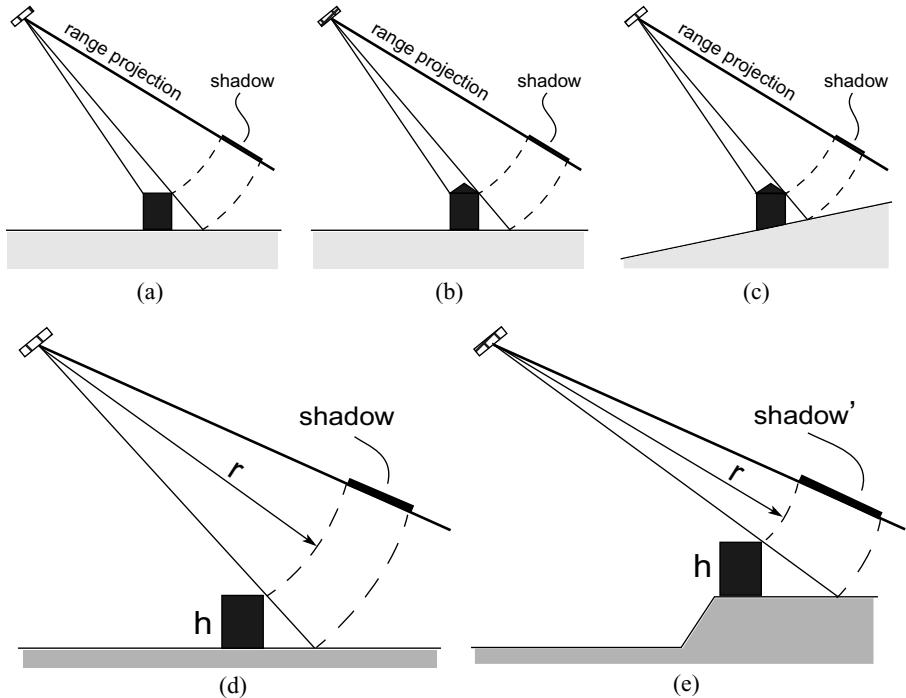


Figure 1.3 Object-height calculation using radar shadow length, and associated problems

elevation of the object's base, or equivalently, the view angle, one cannot correctly calculate the object's height.

In other words, unless a flat surface with isolated, simple geometric objects on it is being measured, shadow- or object-length measurements alone cannot be used to accurately estimate object heights.

A possible way to partially overcome the clutter problem (i.e. shadow not entirely visible), as well as improve object-shape recognition, is to fuse the results from multiple views, such as described in [7]. This is not, however, a general solution to the various problems listed above.

1.5.2 Radarcinometry

Radarclinometry, or shape-from-shading (SFS), is a technique first applied to radar by Wildey for topography estimation of Venus in 1984 [92]. Its major advantage over more widely used techniques lies in its dependence on a single SAR amplitude image for the extraction of 3-D information.

The main goal of SFS is the recovery of surface topography through analysis of variations in scattering intensity: the orientation of a given

reflecting element, relative to the sensor, is directly correlated to its brightness. In its most basic form, SFS assumes a simple reflectance model of the surface, which permits interpretation of pixel brightness in terms of relative orientation alone. For example, Wildey's reflectance model [92] assumes a constant albedo (or normalized radar cross-section, σ^0) over all local incidence angles θ_i , i.e. $\sigma^0(\theta_i) = \text{constant}$. In this way, the reflectance varies as the cosine of the incidence angle (the angle between the view direction and the surface normal).

Considering the general case of an angle-dependent cross-section, the pixel intensity can be written as in [50][37]:

$$I_{\text{pixel}} = f(\theta_i, \sigma^0(\theta_i)) \quad (1.5)$$

That is, a pixel's intensity (brightness) depends on the local incidence angle and the backscatter function, the latter itself a function of incidence angle.

Frankot introduced a Lambertian reflectance model into his version of SFS [27]:

$$\sigma^0(\theta_i) = \frac{\cos^m(\theta_i)}{\sin^n(\theta_i)} \quad (1.6)$$

where m and n are integers. The Lambert model describes perfectly diffuse reflection from a surface, which Frankot showed provides better results than earlier constant- σ^0 models [27]. An in-depth study of SFS using a Lambertian reflectance model can be found in [68].

By integrating the slopes of the surface elements along a given direction, it is theoretically possible to obtain a 3-D profile. In order for this to work, several simplifying assumptions are necessary, such as a constant backscatter function $\sigma^0(\theta_i)$ over the entire image and no multiple-bounce reflections.

These assumptions are still not sufficient to enable one to solve for the heights; the implementation of mathematical constraints is required. This is to overcome radical cumulative errors caused by speckle (a type of noise characteristic of radar images, due to their phase-coherent nature; see [21]), as well as ambiguities such as the fact that for a given reflectance model, the surface normal orientation is constrained to a conical surface and not to a single direction. In particular: (a) the height values must be obtainable by integration along any path, and (b) the surface must be regularized (smoothed) to enforce a stable solution along the chosen path of integration.

The simplifications and constraints required by SFS limit its use to certain types of geographical areas, and certainly the absolute accuracy obtainable is

significantly lower than that obtainable using InSAR or even stereo SAR, described briefly below.

1.5.3 Stereogrammetry (Stereo SAR)

When two SAR amplitude images are available for a given region, the area of overlapping ground coverage, called the *stereo window*, contains points that have been imaged from two different viewpoints (Figure 1.4). Note that the sensor positions are given for a defined reference surface such as the local geoid or ellipsoid.

One of the sensors will need to serve as a reference for reasons which will become clear in chapter 3; the reference is called the *master*, the other the *slave*. In order to maintain consistency, for same-side stereo the master is defined as the sensor aimed away from the other one. In the context of the images formed by the two sensors, the terms master and slave will also refer to the images from the respective sensors.

The stereo intersection angle is sometimes defined simply as the difference between the beam-center incidence angles for two sensors. This definition is somewhat limited, because of the different ways the stereo window can be formed, depending on the sensor positions and beam footprints. A definition stressing the importance of the stereo window itself is used in the context of this work: **the stereo intersection angle is defined as the angle subtended by the two sensors at the point halfway along the reference surface between the near-range and far-range extents of the stereo window.**

Just as stereo vision permits our eye-brain system to localize points in three dimensions relative to our eyes, all points within the stereo window of the radar pair can be localized relative to the sensors. In both cases, an estimation of the disparity field between the two images is performed, which describes the transformation that brings one image into alignment with the other. Our brain does this automatically, and calculates the resulting depth field. Digital matching systems, however, either (a) require human users to define a disparity field manually by specifying *homologous*, or corresponding, point pairs, or (b) the matching can be based on an automatic or semi-automatic correlation algorithm. Because the main focus of this thesis is SAR depth-retrieval without manual intervention or *a priori* knowledge of ground topography, the subsequent discussions will be restricted to fully-automatic image matching.

The disparity field is a raster grid of the 2-D vectors relating all points in one image with their homologues in the other. Given accurate positional information for both sensors, the system can be solved for the ground-point positions, within the same coordinate system. In other words, the disparity

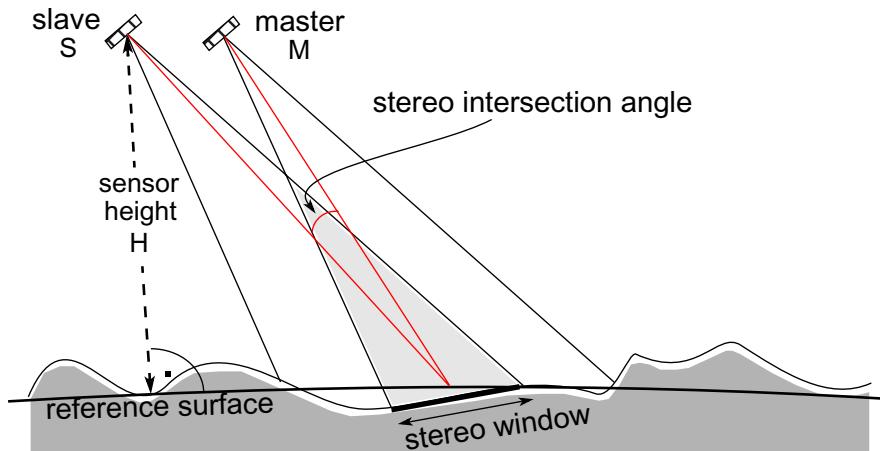


Figure 1.4 Typical stereo SAR configuration

field measured across the scene can be converted to a height map, as described in detail in chapter 3.

This method of height extraction, historically referred to under various names such as stereogrammetry [65] or radar stereoscopy [50], will be referred to as *stereo SAR* in this work.

1.5.4 Interferometric SAR (InSAR)

The fact that SAR sensors send and receive coherent radiation makes it possible for two sensors, properly synchronized, to measure topography with an accuracy on the order of the radiated wavelength. This accuracy may be reduced by factors discussed in chapter 4, such as faulty phase calibration, system noise, and phase unwrapping difficulties.

Figure 1.5 illustrates the basic concept. Two sensors are shown here mounted onto an airplane; however, the concept is identical for air- and spaceborne platforms. Two sensors may acquire data simultaneously (*single-pass* case), as is shown here, or one sensor may acquire two datasets from two different points of view (*multi-pass* case). Sensors 1 and 2 (S₁ and S₂) send pulses with a known initial phase. Point P on the ground returns the radar pulse, which arrives along two paths at S₁ and S₂. The path-length difference Δr relative to S₂ results in a phase difference between the two received echoes. The image spans a certain distance in the ground-range direction (the processed range footprint), and the path difference will increase with

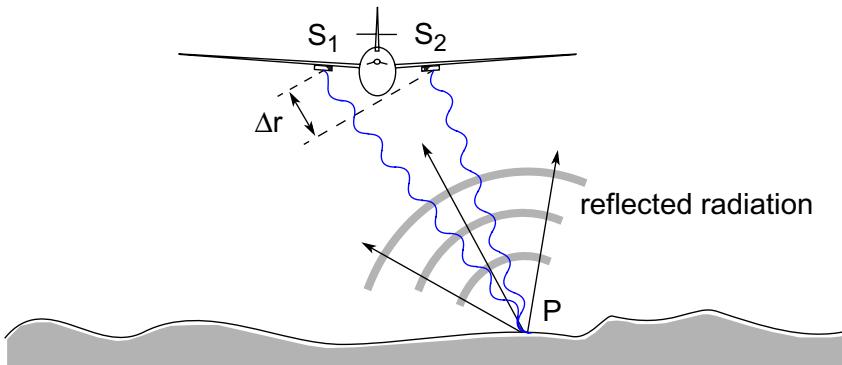


Figure 1.5 InSAR principle

increasing range. By definition, the phase difference, or interferometric phase, can only have values between 0 and 2π , which implies that direct measurement of the *absolute* phase difference is impossible. If one calculates the phase difference for each pixel location in the scene, assigning brightness values to its magnitude (i.e. $[0, 2\pi]$ is mapped to $[0, 255]$ for an 8-bit greyscale visualization), a smoothly-varying pattern of dark and light bands aligned perpendicularly to the range direction will be obtained. This is due to the destructive and constructive interference of the wave troughs and crests as Δr increases in the range direction. Such an interference-band image product is called an *interferogram*. The generation of an interferogram for a flat surface (a so-called *flat-Earth* interferogram) is illustrated in Figure 1.6(a). The dark and light bands are often referred to as *fringes*.

Naturally, the existence of a changing topographical surface complicates the interferogram. Specifically, the phase differences due to the increasing path difference in range are augmented by additional phase differences caused by the vertical variation of the topographic surface; this is called the *topographic phase*. The effect of topography on the flat-Earth interferogram is shown in Figure 1.6(b). If the flat-Earth interferogram is subtracted from the measured one the resulting image will contain the fringes that are due only to the topographic phase, relative to the reference surface (the "flat Earth" surface). Clearly, such a product contains the information required to produce a topographical map of the scene. Indeed, both InSAR and stereo SAR rely on the determination of path difference, which can subsequently be converted to topographical height.

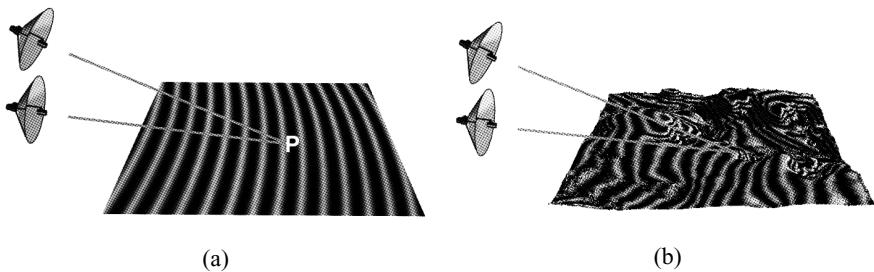


Figure 1.6 Formation of interferometric fringes: (a) for a flat surface (b) for a varying surface [from [34]]

What if the topographical surface is already known to a certain degree, due to an existing approximate or low-resolution DEM for the imaged area? One can calculate the expected interferogram for this surface instead of the flat-Earth surface, and subtract this from the measured interferogram (described in [78] for example). The result will represent the difference between the approximate and true surfaces. If they are very close, few - if any - complete fringe cycles will remain in the result. As will be seen later in chapter 4, it is often easier to estimate a height map with such an interferogram than with one that has been merely flat-Earth corrected, and certainly much easier than one containing all measured fringes.

Chapter 2

Stereo Image Matching Using Wavelets

2.1 Introduction

The extraction of 3-D information from a pair of images - whether they be SAR or optical - requires a disparity field to be found for the pair. For radar images in their native slant-range geometry, this disparity field is a direct measure of the path difference for all imaged points, thus making surface reconstruction possible in a manner similar to InSAR.

The disparity field describes the pixel-by-pixel translational vector field required to bring the slave image into optimal correspondence with the master (i.e. the images from the master and slave sensors, Figure 1.4 on page 34). In other words, the best homologous slave image position needs to be determined for each master pixel position. The reverse is also possible of course, but in this work all calculations are performed within the master geometry for consistency purposes. This disparity map then permits the calculation of 3-D positions, based on the acquisition geometry assumed already available (sensor position and velocity, called the *sensor state vectors*).

Stereo matching is the automatic determination of the disparity field for an image pair, and it is an area of research on its own. The matching problem has been studied actively for the last 20 years by researchers in fields such as computer vision, medicine, biology, and remote sensing because of the need to correlate images that vary slightly from one to the next, either spatially or temporally (such as for video sequences, see [54]). The matching algorithms described in the literature can be roughly classified as area-based, feature-based, or hybrids [87]. Their implementation within *multiresolution*, or coarse-to-fine frameworks has led to hybrid algorithms that now make stereo matching all the more feasible.

The matching approach used during this work was first outlined by H-P. Pan in 1996, which he called *uniform full-information image matching* [66] [67]. It was designed for stereo matching of optical images, and later specifically adapted to human facial reconstructions [55]. The technique involves *multiresolution wavelet decomposition* of the images, which provides

the basis for a coarse-to-fine matching implementation.

The wavelet-based technique, although designed for optical images, can be just as easily adapted for SAR amplitude images. After all, the matching problem is a purely radiometric one; it is entirely independent of the subsequent conversion of the disparities to heights. The existence of speckle noise in SAR images [21] will be shown to have an exceedingly minor effect on the match quality; this is due to the coarse-to-fine nature of the algorithm.

Because wavelets and wavelet decomposition are relatively new, this chapter will briefly introduce the topic, before describing the operation of the matching algorithm in greater detail.

2.2 Wavelets and Multiresolution Decomposition

2.2.1 Fourier vs. Wavelet Decomposition

Fourier decomposition of a 1-D signal results in its representation as a sum of sinusoidal waves, each with a constant amplitude. These amplitudes are a measure of the frequency content of the signal: the larger the amplitude, the higher the similarity between the signal and the analyzing wave. Such a decomposition into frequency space is called a Fourier transform (FT).

In an image, which is a 2-D signal, frequency refers to how quickly the radiometry is changing in a given direction. Applied to a digital (i.e. discrete-valued) image, the Discrete Fourier Transform (DFT) yields a representation of the image in frequency-space. As with the 1-D case, the result is a measure of the overall frequency content of an image. For a detailed discussion on Fourier analysis, the reader is referred to [49].

In this section and the next one the focus will be on the 1-D case of a time-amplitude signal for simplicity's sake. It is, however, understood that the time dimension is equivalent to the spatial dimensions of a 2-D image; the theory remains unchanged when moving from the 1-D to the 2-D case. Indeed, each line or column of a digital image can be considered a single 1-D signal.

Fourier analysis is well-adapted to *continuous* (or near-continuous) signals: signals whose frequency content changes little or not at all over time (or over space, in the case of an image). That is, if one is not interested in *locating* the frequency content within a signal, Fourier analysis is the ideal tool. To take an example from digital audio, one might be interested in suppressing or boosting a narrow band of frequencies in the signal, irrespective of their location in time; this is the job of a multi-band equalizer.

But what if information on the location in time of the various frequencies

is required? *Where* in the signal are low or high frequencies prevalent? This sort of information may be required for *non-stationary* signals - ones whose frequency content varies in time. One solution from the Fourier camp is the *short-time Fourier transform* (STFT). This is a version of the FT in which it is assumed that sufficiently-narrow regions of a signal are stationary. An STFT analyzes the signal for a finite-length window, which is then moved over the signal, yielding a time-frequency decomposition [58].

The problem with the STFT lies in its resolution characteristics. A large window allows for an accurate analysis of the contained frequencies; however, it is impossible to say at what point within the window the various frequencies occur. Conversely, using a narrow window to analyze the signal increases the location accuracy while reducing the certainty in the measured frequencies and their amplitudes. After all, to firmly establish the existence of a given frequency, it needs to be allowed to fluctuate - and this requires a window of sufficient size. Because the STFT employs a fixed window size, its frequency characteristics are fixed as well. The optimal choice of its width is, naturally, application-dependent - but fixed in any case. The greater the desired resolution in frequency, the lower the resolution in time/space, and *vice versa*. This resolution problem is in fact fundamental, and related to Heisenberg's uncertainty principle [42].

2.2.2 From the Short Term Fourier Transform (STFT) to the Wavelet Transform (WT)

The orthogonal analyzing functions of Fourier decomposition, namely sine and cosine functions, are defined for the range $[-\infty, +\infty]$. This corresponds to an STFT with an infinitely-wide window. Such functions are said to have infinite *support*. It is this property that allows the Fourier transform to *either* represent a signal with perfect frequency resolution (frequency-amplitude representation), *or* with perfect time resolution (amplitude-time representation).

In the mid-1970s Jean Morlet, a graduate of France's *Ecole Polytechnique*, was trying to improve crude-oil detection using the STFT [42]. Because of the STFT's inherent resolution inflexibility, he decided to modify the concept as follows: instead of fixing the window width and varying the frequency of the analyzing functions within the window (i.e. the STFT), he decided to fix the number of oscillations and vary the window width. The difference between the two concepts is shown in Figure 2.1. Part (a) illustrates STFT analysis using a fixed window size while varying the number of oscillations; (b) fixes the number of oscillations and varies the window size, thereby creating analyzing functions with varying sensitivity to the different signal frequencies.

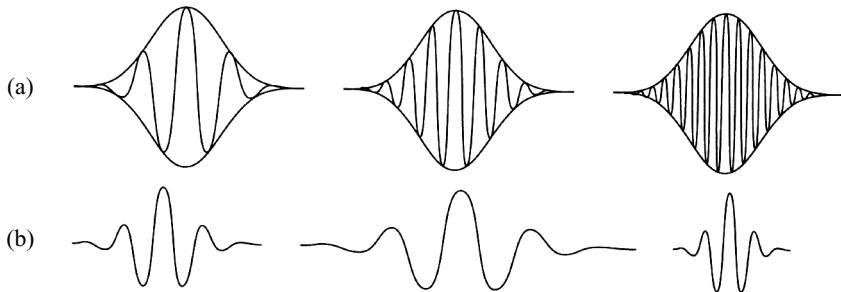


Figure 2.1 Difference between: (a) short-time Fourier analysis (b) wavelet analysis (from [42])

Upon experimenting with these functions for awhile, he realized it was possible to decompose signals into wavelets, and reconstruct the signal afterwards. He enlisted the help of Alex Grossman, a physicist from the University of Marseille. Together, they developed the concept of lossless decomposition and reconstruction of signals into and back from the wavelet domain [39]. The process of representing a signal by a linear combination of orthogonal wavelet functions is called the wavelet transform (WT).

As can be seen in Figure 2.1(b), wavelets are functions that have *compact support*: they are nonzero only on a finite interval. It is just this property that allows the WT to efficiently represent signals with localized features.

Because wavelet decomposition analyzes a signal at various resolution levels, it lends itself easily to *multiresolution analysis* (MRA) of signals, a technique developed by Mallat in 1989 [59][57]. As implied by its name, MRA is the analysis of a signal at different frequencies with different resolutions. Every spectral component is not resolved equally as is the case in the STFT: MRA gives good time resolution but poor frequency resolution at high frequencies, and good frequency resolution but poor time resolution at low frequencies. Fortunately, this situation is often adapted to real signals. A digital image will typically contain low frequencies over large areas, such as radiometric changes in the ground cover; high-frequency components such as edges and objects will tend to be localized.

In the following, the case of a 1-D function $f(x)$ and its representation using wavelets will be studied.

The continuous wavelet transform (CWT) of a 1-D function of an independent variable x , $f(x)$, is given in [42] as:

$$\text{CWT}_f^\psi(s, \tau) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} f(x) \psi^*\left(\frac{x-\tau}{s}\right) dx \quad (2.1)$$

with

τ translation

s scale

$\psi(x)$ A wavelet, which is a function of x

$\psi(x - \tau)$ A wavelet, translated by τ

$\psi\left(\frac{x}{s}\right)$ A wavelet, dilated by s

Equation 2.1 states that any (periodic) function $f(x)$ can be represented as an infinite sum of wavelet functions $\psi(x)$ that have been translated and dilated. It generates a representation of the signal's large- and small-scale features, regardless of where in the signal they might be found.

The CWT is, however, highly-redundant. Because an infinite number of dilations and translations of the mother wavelet are multiplied with the signal, a great deal of information overlap will occur during the decomposition. For purposes of efficient representation and computing time, a *dyadic discrete wavelet transform* (DDWT) is used. In this version, the transform is only calculated for scales which are integer multiples of two, and at translations which are integer multiples of the current scale factor, preserving just the information necessary for signal reconstruction:

$$\text{DDWT}_f^\psi(j, k) = \frac{1}{2^{j/2}} \int_{-\infty}^{\infty} f(x) \psi^*\left(\frac{x}{2^j} - k\right) dx \quad (2.2)$$

which corresponds to Equation 2.1 with $s = 2^j$ and $\tau = k \cdot s = k \cdot 2^j$, for $j, k \in \mathbb{Z}$.

The signals analyzed in this work are digital images. Therefore, a DDWT-equivalent that can operate on discrete, two-dimensional input was needed.

2.2.3 Wavelets and Multirate Filter Banks

Consider a pair of length-two filters operating on a discrete, 1-D signal:

they are given the labels AVG and DIFF. AVG has kernel (weighting) coefficients of $(\frac{1}{2}, \frac{1}{2})$, whereas DIFF is defined as $(\frac{1}{2}, -\frac{1}{2})$. The application of AVG to a pair of numbers will produce their mean. Thus, application of the filter at each array location generates a signal which approximates the input; in other words, AVG is a low-pass filter. On the other hand, the application of DIFF to the same signal will only generate a strong output at those locations where the difference between the two inputs is large. DIFF responds mainly to rapid changes in the input, whereas AVG simply smooths the signal.

The case of AVG and DIFF demonstrates another important characteristic of some filter pairs: it is impossible to reconstruct a signal based on the output of a single filter, but a pair of filters can be chosen such that the entire signal information is retained between them. This is indeed the case for AVG and DIFF: for a given average and difference, there is a unique pair of corresponding inputs (verification of this fact, which requires solving a trivial pair of simultaneous equations, is left to the reader!).

The information-retention property of a well-chosen filter pair, or *filter bank* containing two or more filters, permits one to break down the signal iteratively. *Multirate filtering* is the process of applying the filter pair (or group) repeatedly, first on the input signal, and subsequently on the low-frequency output only, until the signal average has been reduced to a single value. Each application of the filter pair occurs at a given level of resolution: filtering begins on the full-resolution signal defined to be at *level 0*, and continue on through levels 1, 2, etc. until the signal length has decreased to below that of the filters. The result of the multiscale filtering is an array of coefficients from the high-pass filter for each resolution level, in addition to the final low-pass output.

In the case of the filter pair AVG and DIFF, note that the input resolution is being halved at each new level. It is also clear that in order to generate the coefficients allowing signal reconstruction, filtering is only required at every second array position, since the two filter outputs generate the two subsequent inputs for a given location. In this way, one generates exactly as many coefficients as there are array inputs. This dyadic, non-overlapping multiscale process generates a new, lossless representation of the input signal in such a way as to provide information about rapid changes in the signal (*via* DIFF). More importantly, these rapid-change measurements are sensitive to regions of the signal ranging from two elements to half of the signal. Because of the way it analyzes the signal over such a wide range of scales, this sort of filtering process is called *multiresolution decomposition*. Also note that because of the finite length of the filters, a finite portion of the input signal is analyzed.

The wavelet transform also analyzes a signal according to scale and

position. Indeed, a multirate filter bank can be designed in such a way as to comply with the definition of a wavelet transform - but in this case, one that works on discrete input.

A valid scaling function must satisfy certain requirements[60], ensuring that all signals in L^2 (i.e. square-integrable functions) can be expressed as a weighted sum of dilations and translations of the wavelet function ψ .

For the wavelet decomposition of a discrete, 2-D signal such as a digital image, it is possible to build a wavelet basis by combining 1-D scaling and wavelet functions [20]:

$$\Phi(y, x) = \phi(y)\phi(x) \quad (2.3)$$

$$\Psi_h(y, x) = \psi(y)\phi(x) \quad (2.4)$$

$$\Psi_v(y, x) = \phi(y)\psi(x) \quad (2.5)$$

$$\Psi_d(y, x) = \psi(y)\psi(x) \quad (2.6)$$

with

y, x vertical, horizontal index

h, v, d vertical, horizontal, diagonal edges detected

An example of performing the above four operations on an image is shown in Figure 2.2, shown in the standard "pyramid" representation. Here, a test image is shown before and after the transformation, which was based on a wavelet basis first defined by Ingrid Daubechies [18]. A number of wavelets are commonly used; the Daubechies wavelet was arbitrarily chosen for this example. The output of Equation 2.3 - the low-pass function - is at the upper left of Figure 2.2(b). The upper right, lower left, and lower right images are the output from Equation 2.4 through Equation 2.6, respectively. When linear features exist, they are mainly visible in the image created with the filter most nearly perpendicular to the feature orientation. Thus, the filter sensitive to horizontal features produces the upper-right output, the one sensitive to vertical features the one on the lower-left, and the combined horizontal-vertical filter (i.e. diagonal) output is at the lower-right.

This decomposition of the input into four outputs can be repeated for the new low-resolution output successively, until the low-resolution image is reduced to a single pixel. At each new level, the wavelet filters are applied to increasingly-coarse input, providing the multi-scale sensitivity.

When the wavelet and scaling functions have been properly chosen (as described earlier), the original image can be perfectly reconstructed given the

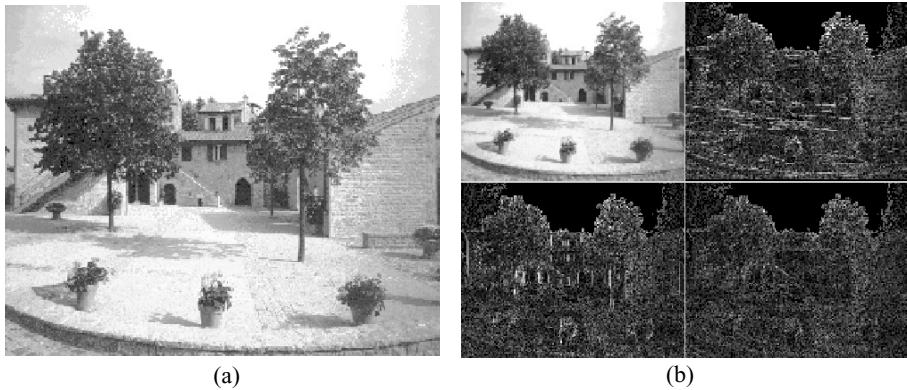


Figure 2.2 Single wavelet transform of an image: (a) input image (b) wavelet transform of input

output of all four filters in the final step (naturally, discounting the effect of limited computational floating-point precision).

For more detailed reviews of wavelet analysis and filter banks, see for example [58] or [9].

2.3 The Dyadic Discrete Wavelet Transform Using Magarey-Kingsbury Wavelets

The wavelet transform illustrated in the example from the last section is a real-valued transform: four real coefficients are generated for each pixel per transform iteration. The wavelet transform was used in this work to generate multiresolution decompositions for two images constituting a SAR stereo pair. Wavelet functions can be chosen according to the desired application; the matching technique used for this work is *phase-based*, which gives it a high sensitivity to feature orientation in images rather than radiometry. For the phase to be available, the image pair needs to be transformed using a complex-valued transform, or wavelet basis. The more standard amplitude-based matching methods, such as the cross-correlation, are not as well suited to matching oriented edges for example. Therefore, a real-valued transform such as the Daubechies transform of Figure 2.2 is not suitable for phase-based matching. However, it was presented as an example of the nature of wavelet transforms in general, which is similar for all wavelets.

The use of phase as a feature to be matched is rooted in the Fourier shift theorem [73], which relates translations in space to phase rotations in frequency. A localized version of the Fourier shift theorem is required in order to be able to use it locally within an image. The *Gabor transform* is a form of

Short Term Fourier Transform (STFT, see Section 2.2.1 on page 38) with Gaussian windows on the Fourier basis functions (sines and cosines). The Gabor phase is particularly well-adapted to image matching, in which pixel brightnesses and affine distortions can superficially change the appearance of a given feature from image to image.

The wavelet basis, which has a Gabor form, was designed by Magarey and Kingsbury [55], and will henceforth be referred to as the *M-K basis*. The basis is composed of a pair of complex-valued Finite Impulse Response (FIR) filters of length 4, one a high-pass, the other a low-pass filter. Their values at integral pixel positions are given by functions of the form:

$$g(n) = a_0 v_0(n + 0.5) e^{j\omega_0(n + 0.5)} \quad (\text{low-pass}) \quad (2.7)$$

$$h(n) = a_1 v_1(n + 0.5) e^{j\omega_1(n + 0.5)} \quad (\text{high-pass}) \quad (2.8)$$

with

$$n = -2, \dots, 1$$

ω_0, ω_1 angular frequency (low-pass, high-pass)

a_0, a_1 complex constants (low-pass, high-pass)

v_0, v_1 real-valued window functions (low-pass, high-pass) of the form:

$$v_{0,1} = e^{-\frac{(n+0.5)^2}{2\sigma_{0,1}^2}} \quad (2.9)$$

The window functions are symmetric about $n = 0$, and decay to zero as $n \rightarrow \pm\infty$. Furthermore, ω_0 is equal to $\pi/6$, while $\omega_1=0.76\pi$. These values are a consequence of the equal-spaced angular distribution of the six filter orientations between $-\pi/2$ and $\pi/2$, among other desirable properties [54]. Its smaller of the two angular frequencies makes $g(n)$ the low-pass filter, with $h(n)$ the high-pass filter [20].

Similar as was indicated by Equations 2.3 through 2.6, each filter is applied one-dimensionally either vertically or horizontally to the image, that is, to the rows or the columns. A 2-D sub-image is produced by pairs of such operations. For example, application of the high-pass filter $h(n)$, or H , to the columns followed by application of the low-pass $g(n)$, or L , to the rows produces a result such as upper-right sub-image of Figure 2.2(b); this operation is labelled HL for convenience. Note that vertical filtering is always performed first. The result of the operations LL , LH , HL , and HH produces the three detail sub-images and single smoothed sub-image. When real-valued filters are used, the above four operations will enhance features aligned

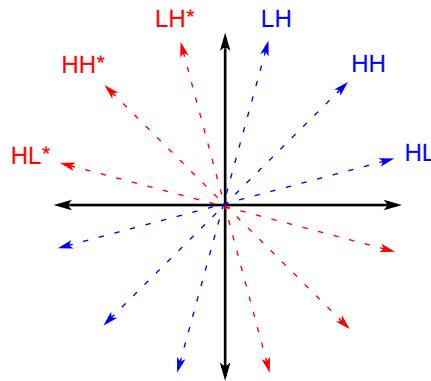


Figure 2.3 Approximate filter alignments for the M-K wavelet basis

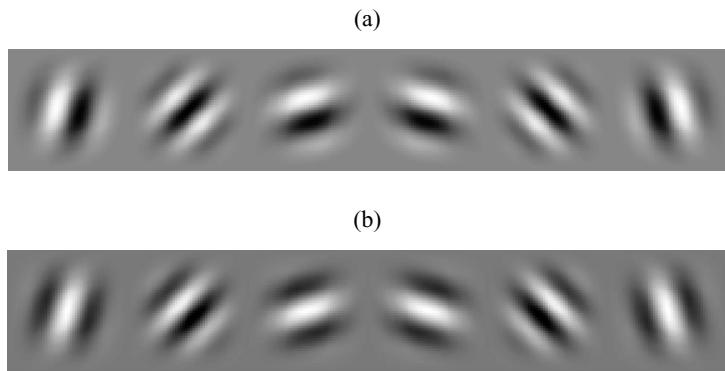


Figure 2.4 Greyscale representation of M-K mother wavelets: (a) real part (b) imaginary part (from [56])

vertically, horizontally, or diagonally, according to the order of filter application. It should be mentioned that any one of the above-mentioned pairwise operations can be seen as a single filtering operation, and the combined "filter" is aligned along a direction perpendicular to the edges it is the most sensitive to. Thus, the HL filter is aligned along the 90-degree axis, and is most sensitive to horizontal edges.

On the other hand, performing the same four operations on an image using a complex-valued wavelet basis will not, in general, detect edges that are aligned along the 0-degree and 90-degree axes. In our case, the M-K basis is

aligned approximately along directions illustrated in Figure 2.3. Here, the three filters resulting from the combinations of $g(n)$ and $h(n)$ (L and H) are seen to align along directions in the first and third quadrants of the complex plane (in blue); as stated above, the edges detected by a given filter will be perpendicular to the filter's alignment.

Taking the complex conjugate of the same filter triplet results in alignments which are symmetric about the 0-degree and 90-degree axes, also shown in Figure 2.3 (in red). The M-K complex wavelet basis is therefore sensitive to twice the number of edge alignments compared to a real-valued basis, which has only three alignments (horizontal, vertical, and diagonal). Greyscale representations of the real and imaginary parts of the M-K mother wavelets are shown in Figure 2.4 [56]; the filter orientations are clearly visible.

When the wavelet transform is applied to the image, six coefficients are produced for each pixel the transform is applied to, resulting from the LH , HL , HH , LH^* , HL^* , and HH^* outputs. The transform is applied to both images separately and iteratively, and it is on the resulting output "coefficient pyramids" that the actual matching algorithm is performed.

The result of the wavelet decomposition on our image pair is two coefficient pyramids. The term "pyramid" is used because of the two-fold reduction of the output resolution at each subsequent level; in practice, the quad-tree-style representation such as was shown as an example in Figure 2.2, is the most common way of illustrating the decomposition. A layered representation such as is shown in Figure 2.5 highlights the concept of resolution levels, each level being the output of the six combined filters described above. The coefficients on a given level are generated based on the LL , or smoothed, output from the previous level, and this iterative decomposition can be continued until only a single pixel remains as the LL output. In Figure 2.5 only two iterations are represented, whereas for an image of size 1024×1024 pixels, ten iterations will be required for decomposition down to the pixel level. However, in practice multi-resolution image matching will typically begin working on the 8×8 or 16×16 level.

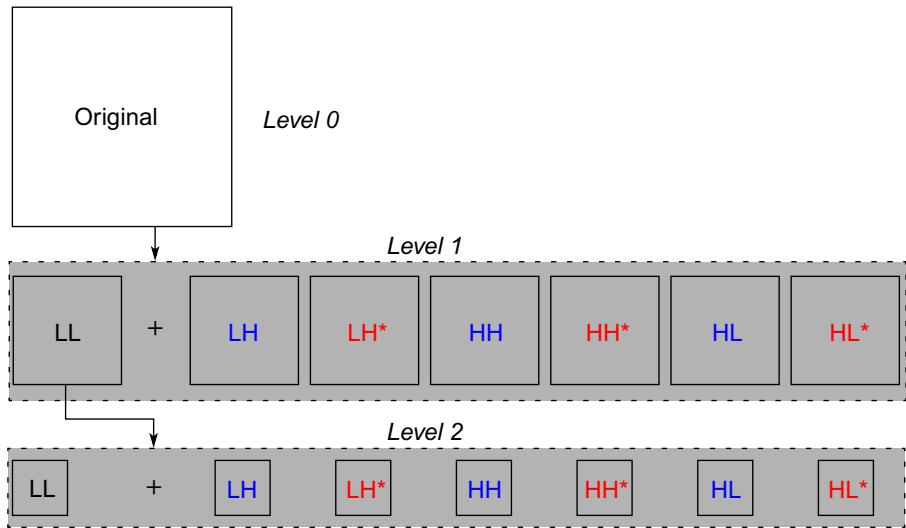


Figure 2.5 Two iterations of a DDWT on an image

2.4 Stereo Matching Algorithm

2.4.1 Measuring Similarity

The heart of the matching process is the calculation of the similarity between two pixels, one from each image at a given level of resolution. The goal of the matching process in the implementation used in this work can be stated as follows:

For each fixed, integral pixel position in the master image, find the best corresponding position in the slave image.

Note that sub-pixel accuracy is maintained during matching: the slave positions are not restricted to pixel centers.

In order to measure the similarity between a master pixel p and a slave pixel p' at a given level of resolution i , the function first described in [66] is used, called the similarity distance, or $SD^i(p, p')$:

$$SD^i(p, p') = \sum_{k=1}^{6} |D_{p'}^{(k, i)} - D_p^{(k, i)}| \quad (2.10)$$

with

p pixel position in the master image
 p' pixel position in the slave image
 i level of resolution

$D_p^{(1, i)}, \dots, D_p^{(6, i)}$ the six complex coefficients generated by the DDWT on level i for pixel p

$D_{p'}^{(1, i)}, \dots, D_{p'}^{(6, i)}$ the six complex coefficients generated by the DDWT on level i for pixel p'

The inherent interpolability of the Gabor-like wavelet filters allows the SD to be calculated between slave pixel locations, for a given integral master position. This way, the SD can be represented as a surface in three dimensions, whose shape is indicated in Figure 2.6. Its values represent the strength of the match between master position p and the current slave position. If p' is the most recent matched position for master pixel p , the green area represents the local search area around p' . The SD surface, drawn as a series of isovalue lines in cross-section [56], has its minimum value at p_0 (i.e. the greatest similarity), indicating that this is the best match for master pixel p . This floating-point pixel location is then rounded to the nearest pixel location at the next-finest resolution during interpolation to the next level.

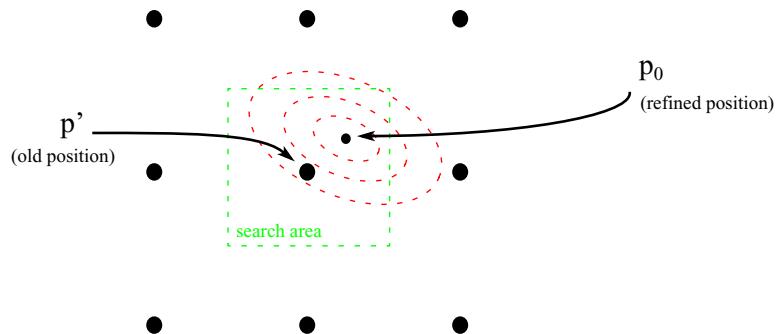


Figure 2.6 The similarity distance (SD) surface. The current disparity estimate is centered on p' , with the next-finest iteration at p_0 . The red lines are isovalue lines of the ellipsoidal SD in cross-section; the green-enclosed perimeter is the search area.

2.4.2 Disparity Field Error Correction

A consequence of the coarse-to-fine (multiresolution) design is that any errors that occur at a coarse level risk being propagated to the finer levels. In practice, the disparity field will become jagged and irregular if a correction operation is not carried out after the field estimate at a given resolution level. The regularization step proposed by Anandan [1] and described in the wavelet-matching context in [55], aims to remove sudden perturbations in the disparity field without overly-damaging its correctness. The procedure attempts to find the best compromise between a perfectly smooth field and the estimate at a given level. Mathematically, if $\{u\}$ is the smoothed disparity field to be found, it attempts to minimize the energy functional

$$E(\{u\}) = E_{sm}(\{u\}) + \Omega E_{ap}(\{u\}) \quad (2.11)$$

$E_{sm}(\{u\})$ is the smoothness term, a normalized measure of the difference between $\{u\}$ and a perfectly-smooth field; the greater the difference, the greater its influence on $E(\{u\})$. Balancing this is $E_{ap}(\{u\})$, which is the global "approximation error energy"; it increases with increasing difference between $\{u\}$ and the input field. The balance between these two forces is governed by the constant Ω ; if it is set to 1.0, the influence of each of the two opposing forces on the total energy $E(\{u\})$ is equal. To take an example: setting it to 0.1 would give the smoothness term $E_{sm}(\{u\})$ a tenfold-greater influence $E(\{u\})$ than the approximation error energy $E_{ap}(\{u\})$, resulting in a greatly-smoothed field. On the other hand, setting it very high will prefer jagged, noisy disparity field, which may contain more detail - but at the cost of a greater number of strong errors. The choice of Ω depends on the type of images being matched: images of topographic surfaces that change rapidly will prefer high values of Ω so as to retain high-frequency details, while smooth surfaces will fare better with an Ω well below 1.0.

2.4.3 Algorithm Structure

The matching algorithm operates on the output of the DDWT for both input images, beginning at the coarsest level. The steps followed during matching, as illustrated in Figure 2.7, are described below:

- 1) **Coarse registration:** by having the user define at least one pair of homologous points, an initial rotation, translation, and re-scaling are performed for the slave, bringing it into coarse alignment with the master.

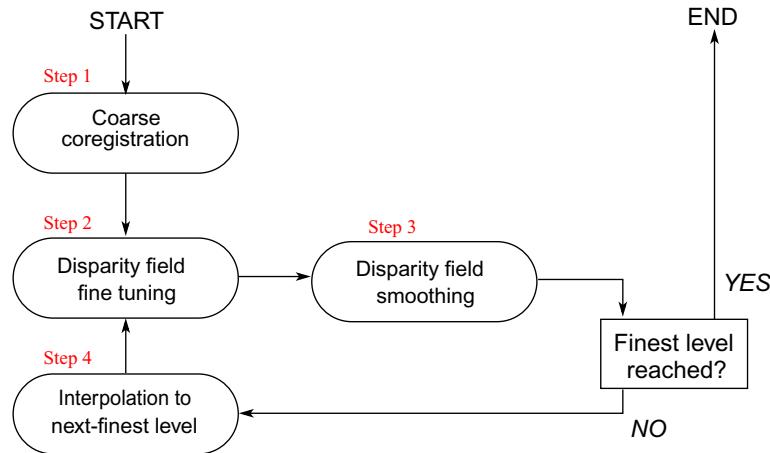


Figure 2.7 Multi-resolution matching: algorithm structure

- 2) **Disparity field fine tuning:** given the disparity field from the last (coarser) level as a starting point, re-calculate it at the current resolution. That is, for each master pixel position and its corresponding estimated position in the slave, the new local minimum of the SD function is determined for the slave, as shown in Figure 2.6. In other words, a fine adjustment is made to the previous slave position presumed to correspond to a given master position. It is noted again that sub-pixel accuracy is maintained in the slave: for each integral master position, one estimates a new homologous slave position within the sub-pixel domain of the new resolution level.
- 3) **Disparity field smoothing:** a field-smoothing operation is performed as described in section 2.4.2 on page 50, governed by two opposing forces: (a) on the one hand, it is desirable to retain the precision gained by step (2) at the finer resolution, while (b) on the other hand, one would like to reduce the errors introduced by the fine-tuning caused by poor matches. Poor matches can occur in regions with no distinct objects or edges, for example. The smoothing takes both (a) and (b) into account by smoothing more or less, according to the *match confidence* for a given master position [55]. This is a measure of the curvature of the SD function at its minimum. Because the field is ellipsoidal, the match confidence has components along the major and minor axes. Both components are considered during the smoothing: if the confidence (curvature) is high along one axis, the disparity will be retained in this direction; if the confidence

is low, more smoothing will occur, thus reducing the effect of the fine-tuning at this point and in this direction. The balance between smoothing and precision can be altered by the user, thus making it case-adaptable. The match confidence can also be used during automatic GCP generation, as will be described later (section 5.3 on page 73).

- 4) **Interpolation to next level:** at all intermediate resolution levels, the current disparity field is interpolated to the next-finest level, thus providing the initial estimate for the next iteration of step (2). All current disparity estimates are rounded to the nearest pixel-center location on the next-finest level; the sub-pixel precision is therefore lost at the start of each new iteration. However, it is once again corrected during the fine-tuning step on the new level. Sub-pixel precision is, of course, retained in the interpolation to the finest resolution in the very last step. The net effect of the process is that sub-pixel accuracy is maintained through to the end, through a series of roundings and subsequent sub-pixel-level corrections.

Chapter 3

Height Extraction Using Stereo SAR

3.1 Introduction

A brief overview of stereo SAR, as well as the geometric conventions used in the context of this work, was given in section 1.5.3 on page 33. The theory behind matching the two images produced by such a configuration was introduced in chapter 2. This chapter will examine the relationship of the disparity field to the surface topography in greater detail, as well as the theory behind the estimation of a DSM based on the disparities.

At this point it should be clarified that there is a fundamental difference between the shadows in an optical-stereo pair and those in a SAR pair, complicating the matching problem for SAR. In a passive system, where the sun illuminates the scene, the shadows are objects fixed on the ground. They change their aspects according to view position, as do all other objects in the scene. In a SAR pair, however, the shadows are not simply *viewed* from different angles. Rather, they are *a direct result* of the different angles. This highlights an important difference between optical and SAR stereo:

SAR stereo involves a different illumination source for each image, optical stereo only one (the sun). This makes SAR stereo fundamentally more challenging than optical stereo.

3.2 Rough Coregistration

The core of stereo SAR is the matching (coregistration) step between the two slant-range images. The output of the matching process is to serve as the input to the generation of a DSM. This said, the nature of SAR geometry poses a first major difficulty. Consider the typical case of same-side stereo, as depicted in Figure 3.1; the width of the stereo window will be significantly larger for the slave sensor S, than for the master M. An example of this effect

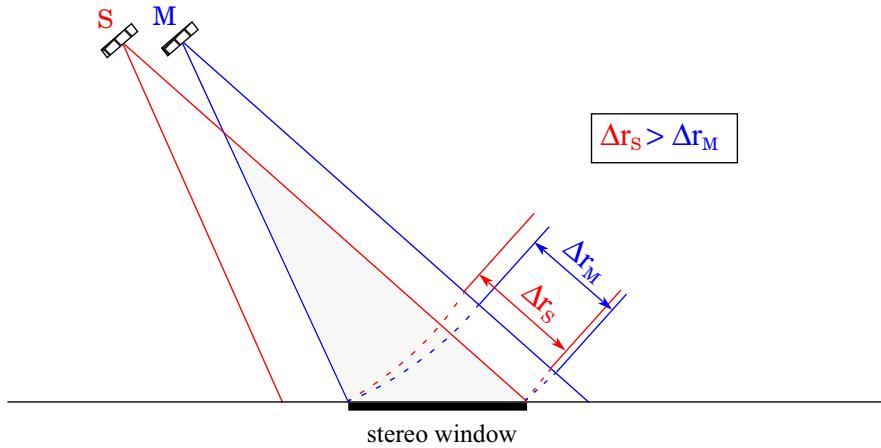


Figure 3.1 The stereo window in range geometry

is shown in Figure 3.2(a), where it can be seen that the stereo window covers a greater range in the slave than in the master. Since it is just this region that the matching estimation is performed on, the algorithm will either have to be sufficiently insensitive to gross scaling differences, or this difference will first need to be corrected for. Unfortunately, the wavelet-based approach used in this work fails when the images are too different, either in overall size or aspect ratio, or if too many internal dissimilarities exist. In fact, images stemming from opposite-side stereo were discovered to be unmatchable using the wavelet-based algorithm, precisely because of the radically-different viewing geometries. While relatively insensitive to radiometric differences, large-scale compression of one image with respect to the other, such as occurs in stereo SAR, cannot be properly handled.

Therefore, a first step is required to increase the similarity between the images. This is done here by resampling one image such as to bring it into the geometry of the other, based on the simple geometric range model depicted in Figure 3.1; this is called *rough coregistration*. The slave sensor will have the wider stereo window, for same-side stereo. For this work the slave was consistently resampled into master geometry rather than the reverse; although either is possible using bilinear resampling, data reduction rather than augmentation was chosen so as not to introduce new information that could influence the sub-pixel accuracy of the matching algorithm.

The output of rough coregistration is a new slave image, S' , one that is much more similar-looking to the master, as seen in Figure 3.2(b)

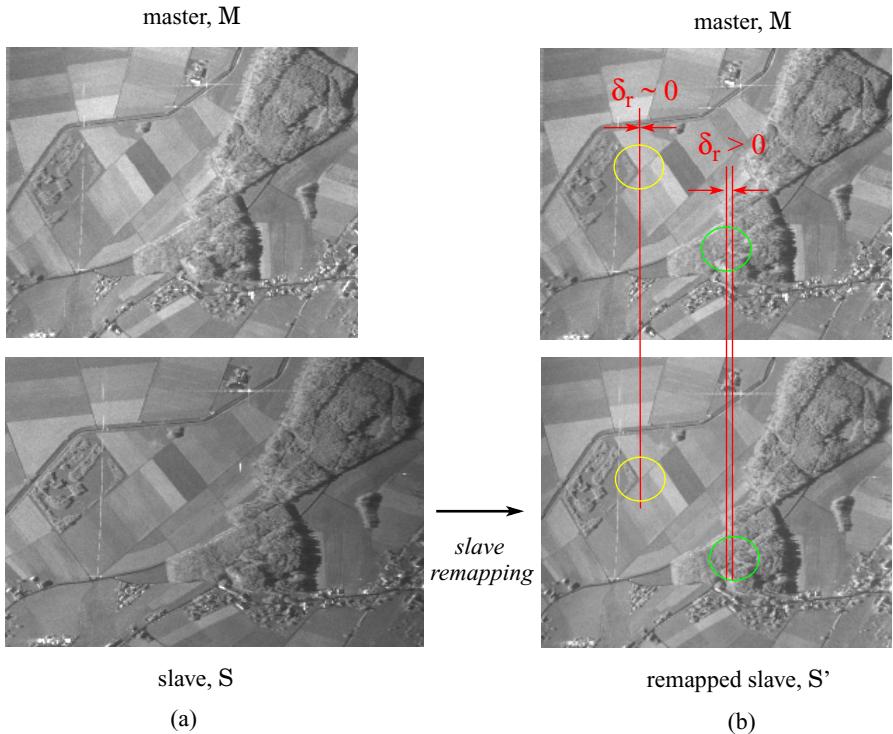


Figure 3.2 Rough coregistration: remapping the slave’s range geometry and radiometry to those of the master: (a) before rough coregistration (b) after rough coregistration; only range disparities due to the topography relative to the reference plane remain

(disregarding the colored annotations for now). Note that in addition to the geometric coregistration, the signal loss with increasing range has also been adjusted to better match that of the master. This is performed on a column-by-column basis after the geometric resampling: the average brightness of a given slave column is brought to the average master level for the same column. Because matching is a purely-geometric problem, it can do no harm to modify the slave image radiometry in order to make it appear more like the master. This is because the geometric transformation used to remap the slave into master geometry is later added to the measured disparity field so that the absolute range differences between the original master and slave images can be calculated. The only purpose of the coarse geometric and radiometric coregistration is to facilitate the calculation of a disparity field between the two images.

Naturally, were the topography as flat as depicted in Figure 3.1, the new

slave would lie perfectly over top of the master. The images would be perfectly matched; there would be no reason to follow up with an estimation of the disparities between the images.

In reality, the flat ground depicted in Figure 3.1 will serve as a reference height for the true topography, which can vary greatly and rapidly about this reference level. A consequence of the existence of the topography is indicated with the red lines in Figure 3.2(b). On the left, the road intersection at the center of the yellow circle has not changed its horizontal (range) position noticeably from M to S', as shown by the red line. This area lies at approximately the reference height used during remapping, and so its position does not change. Towards the right, a particular feature in the forest-stand crown (illuminated from the left by the radar) lies at a different range position in S' as compared to M. This is due to its nonzero height relative to the reference plane, hence its incorrect repositioning during coarse remapping. The difference between their range positions is indicated as δ_r , called the *topographic range disparity*. **These residual disparities in range are measured over the whole stereo window, using the wavelet-based matching technique.** The reason for the existence of these residual disparities is examined in the next section.

3.3 Topographic Disparity

If one assumes that there is no topography above or below the reference height, as in Figure 3.1, then the remapped slave S' will not exactly lie over top of the master. In fact, the greater an object's height above the reference, the greater its range disparity δ_r will be, as compared to the estimated position. The reason for this can be seen in Figure 3.3. During rough coregistration, topography is ignored; the treetop will appear at range $(r_M)_{\text{true}}$ in M, but because this range is assumed to lie at the reference height (i.e. flat-Earth assumption), the new slave range will be calculated as $(r_S)_{\text{flat}}$, instead of $(r_S)_{\text{true}}$. The difference in range between the estimated and true slave ranges is the topographic range disparity δ_r , which is the small range shift for the point in the forest crown in Figure 3.2(b). That is:

$$r = (r_s)_{\text{flat}} - (r_s)_{\text{true}} \quad (3.1)$$

All points lying above the reference plane will be incorrectly assigned over-estimated ranges during rough coregistration. The resulting disparity map is therefore directly related to the topographic variation about the reference plane. Note that $\delta_r < 0$ for all points lying above the reference plane,

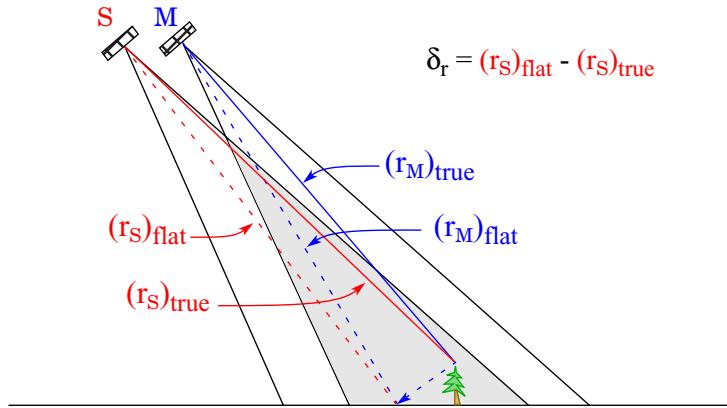


Figure 3.3 The origin of topographic range disparity

and $\delta_r > 0$ for points below the plane.

3.4 Geocoding

So far, only the role of the range disparity has been discussed in determining the height of an imaged point on the ground. That is, it has been assumed that the vertical (azimuth) disparity field is null over the entire stereo window. This is a valid approximation in many cases, notably for the case of a sensor that was flown along parallel flight tracks while imaging, and always with the same speed, therefore keeping the azimuth line spacing constant and the image orientation constant for both acquisitions.

Of course, a real sensor cannot be flown along perfectly parallel tracks or at exactly the same, constant, speed. Although this can be nearly achieved with spaceborne platforms, it is much more difficult for airborne platforms, as they are subject to variable wind conditions. However, this work will consider the general case of a nonzero azimuth disparity field during calculation of the DSM.

For a given pixel in the slave image, the following occurs:

- 1) its approximate position in the master is calculated via the rough coregistration described above, and then
- 2) this rough estimate is adjusted for the topographical disparities by the subsequent matching process.

The result is that for each slave pixel, a sub-pixel-accurate estimate of its position in the master is generated. Because the images are in their native slant

range geometries, the ranges to all points are known for both the master and the slave sensors. Given these ranges for all point pairs, and given accurate flight position and velocity information, the following four equations must be satisfied for each point \vec{P} imaged by both sensors, within an Earth-centered cartesian reference system:

$$|\vec{P} - \vec{S}_1| - |\vec{r}_1| = 0 \quad (3.2)$$

$$\frac{-2}{\lambda} \cdot \frac{(\vec{P} - \vec{S}_1)(\hat{v}_{S_1})}{|\vec{P} - \vec{S}_1|} + f_{d_1} = 0 \quad (3.3)$$

$$|\vec{P} - \vec{S}_2| - |\vec{r}_1| - \Delta r = 0 \quad (3.4)$$

$$\frac{-2}{\lambda} \cdot \frac{(\vec{P} - \vec{S}_2)(\hat{v}_{S_2})}{|\vec{P} - \vec{S}_2|} + f_{d_2} = 0 \quad (3.5)$$

with

\vec{P}	ground point position
1, 2	subscripts for master (1) and slave (2) antenna
\vec{S}	antenna positions
\vec{r}	range from \vec{P} to \vec{S}
\hat{v}	sensor velocities relative to \vec{P}
f_d	Doppler frequency shift
Δr	range differences (given by the range pixel disparities)
λ	radar wavelength

This system of equations is called a *range-Doppler* system [35]; it is over-determined since \vec{P} (which is three-dimensional) is the only unknown, but may be solved numerically. The solution implemented in this work is based on iterative Newton-Raphson approximation [72]. It yields a 3-D position in the same cartesian reference system used to describe the sensor positions [41]. This georeferenced height field is then transformed into a map geometry to allow comparison with other available height models.

The range-Doppler equations are 3-D functions in space, and the position of the point being imaged is given by the intersection of these functions at a point in space. For a given imaged point with 3-D position \vec{P} on the ground at a given instant, each sensor records a different Doppler shift of the reflected signal because of the different radial velocities. In the same instant each sensor measures a different range to the same point. The four equations 3.2 through 3.5 represent these range and Doppler-shift measurements. Given the range to a point, it is known to lie on the surface of a sphere with a radius equal to the

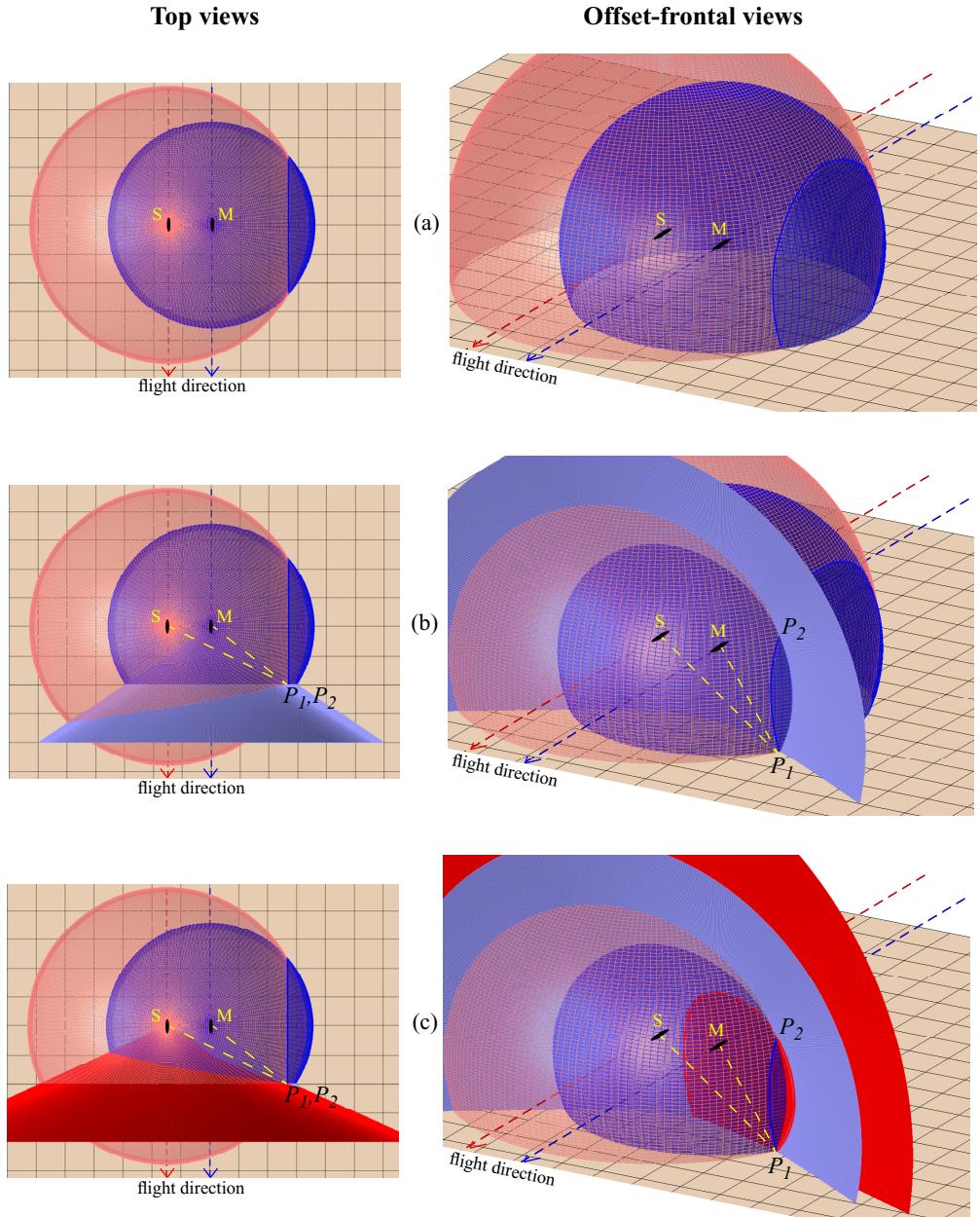


Figure 3.4 Solution of the range-Doppler equations for a point P_1 viewed by slave (S) and master (M) sensors. The gridded plane represents the surface containing P_1 ; it is shown as a viewing aid only. Top and offset-frontal views: (a) The intersection of the two range spheres is a circle. (b) Adding a Doppler cone permits the localization of P_1 . The second solution (P_2) is ignored, it being above the sensors. (c) Given the second Doppler cone, the system is over-determined; the location of P_1 is determined numerically.

measured range and the sensor at its center. For a fixed Doppler shift or radial velocity, the same point is known to lie on the surface of a cone that opens up in the flight direction and having the sensor at the cone's tip. Thus, for two sensors, the range-Doppler measurements for a given point describe two range spheres and two Doppler cones. Any three of these four equations suffice to analytically solve for the point position, but all four are solved simultaneously and numerically because only imperfect measurements of the state vectors and ranges are available. A numerical solution ensures that all measurements are optimally taken into account.

A graphic representation of the range-Doppler equations is shown in Figure 3.4 for two sensors acquiring data for stereo. The flight paths are indicated by the red- and blue-dotted arrows. The figure shows the two Doppler cones and two range spheres measured for some point P_1 on the surface being imaged, but split up into three steps for clarity. The left- and right-hand columns represent two different viewpoints (top view and offset-frontal view), shown for comparison purposes.

Figure 3.4(a) represents the two measured range spheres. Two ranges are not sufficient to pinpoint the location of the scatterer; this is illustrated by the fact two spheres create a 2-D circle where they intersect (rather than a point).

Figure 3.4(b) superposes the Doppler measurement from the master antenna. At this point two possible solutions emerge, given by the locations P_1 and P_2 . They are symmetrically below and above the sensor, respectively. P_2 can therefore be eliminated, since it is clear that a point below the sensor being measured.

With the addition of the second Doppler measurement in Figure 3.4(c), the system becomes over-determined: more information is available than is required to obtain a solution. However, because in reality the measurements are imperfect, a redundancy of information is not undesirable, and the iterative numerical solution arrives at the best possible solution for the position of P_1 with the given information.

3.5 Summary of the Stereo Processing Chain

As an overview of the entire stereo SAR processing chain, the steps are summarized here:

- **Acquisition of a SAR stereo pair.** Image pair can be acquired using airborne or spaceborne sensors, either flown along parallel tracks, crossing tracks, or in tandem. Simulation of stereo pairs can also be performed, given a predefined flight path and a DSM as input. Images

are required to be in their native slant range geometry.

- **Coarse mapping of slave into master geometry (rough coregistration).** Also referred to a "flat-Earth remapping" of the slave into the master geometry. Increases the geometric (and, optionally) radiometric similarity between the master and slave images, increasing the chances of successful disparity field estimation.
- **Disparity field estimation.** Generates a 2-D vector field describing the translation required to bring the roughly-coregistered slave into perfect alignment with the master. This disparity field, in combination with the information used to perform the coarse registration, provides the sensor-to-point ranges for all pixels, for both sensors.
- **Generation of a geocoded height map (DSM).** Given the sensor-to-point ranges from the last step, and given the position and velocity of both sensors during image acquisition (state vectors), it is possible to calculate the absolute position in space of all imaged points. This can be directly calculated within the coordinate system describing the sensor positions (for example the global cartesian *World Geodetic System 1984*, or WGS84) and later transformed into map geometry if required.

3.6 General Weaknesses of the Stereoscopic Method

The heart of any stereo processing chain is the automated matching, and the quality of the resulting disparity field determines, more than any other step, the quality of the estimated DSM. The matching scheme employed in this work is quite flexible in that several parameters can be modified by the user, making it case-adaptable to some extent. However, it was found that there is still room for improvement in the details of the algorithm itself; these will be discussed in chapter 10.

Errors in the height map can also be due to errors in the knowledge of the state vectors, but this is not unique to the stereo technique, and will typically not be the dominant source of the positional errors in the result. Multi-pass stereo and InSAR are both affected by the positional measurement errors of both sensors. This is a small but not insignificant disadvantage as compared to single-pass InSAR.

Finally, there is the issue of resolution. Matching measures shifts in amplitude pixels, whose sides are given by the range and azimuth resolutions, typically of the order of half a meter (airborne) to several tens of meters (spaceborne). In contrast, InSAR utilizes high-precision phase measurements of the returning signals, whose wavelengths are typically several centimeters

long. With this fundamental difference in the size of the entities being measured, it is not surprising that InSAR delivers higher-resolution height estimates.

Chapter 4

Interferometric Height Extraction

4.1 Introduction

While stereo SAR compares the signal amplitudes from two acquisitions for the estimation of a height map, interferometry relies on the phase differences of the reflected electromagnetic waves between the two, which is directly proportional to the path difference Δr of the waves reflected from a point in the terrain (Figure 1.5 on page 35). Because the topographical height depends ultimately on phase measurements, rather than visual matching of amplitude pixels, the accuracy achievable using this technique is much greater than that achieved by stereo processing. Because the departure point is the phase, and not the amplitude, the error sources associated with the interferometric chain are of a different nature than those encountered in the stereo technique. Combining the two techniques should therefore be aimed at utilising the advantages of both to achieve a better result than either can produce alone.

Interferometric height extraction will be discussed in this chapter, and its combination with stereo SAR in Chapter 5.

4.2 Generation of a SAR Interferogram

4.2.1 Registration

In order to calculate the measured interferometric phase difference for points on the ground, both images need to be coregistered. It is convenient to choose the geometry of one of the two as the reference; this image (or equivalently, this sensor) is called the master. By calculating the disparity field that brings the other image (the slave) into alignment with the master, the phase captured by both sensors is known for all pixel locations.

It should be noted that interferometric baselines are necessarily much smaller than stereo baselines because of coherence limitations, meaning that

the two amplitude images will appear to be nearly the same. The consequence of this is that the transformation required to register the slave to the master will be of a much subtler nature than one arising from a typical stereo pair. Special multiresolution techniques, such as the one used in this work for the stereo matching, are therefore not necessary. Nonetheless, because the phase varies from pixel to pixel, accurate phase measurement also requires accurate sub-pixel correlation measurements.

Given precise knowledge of the positions of both sensors, the baseline between them can be calculated for the master scene center. By assuming that this baseline applies to the entire scene, one can calculate the expected slave pixel positions for each master position. This process results in a first approximation of the disparity field.

In a further step, the first approximation of the disparity field is refined by performing a cross-correlation-based automatic matching, but this time using the measured phases as inputs [81]. Where little to no similarity is found between corresponding image areas, the *coherence* is said to be low (this will be defined in section 4.2.5). Given the new sub-pixel-accurate disparity field, the precise phase measured by each sensor is known for all image locations.

Details on the registration methods used here are given in [78].

4.2.2 Interferogram Calculation

Given two complex signals for a given scene position (master and slave returns), c_1 and c_2 , the path difference Δr (from Figure 1.5 on page 35) results in a phase difference equal to $c_1 c_2^*$. By averaging this quantity over N azimuth looks to reduce data quantity as well as increase the SNR, the interferogram is calculated as:

$$I = \sum_k^N c_1(k)c_2(k)^* \quad (4.1)$$

with the interferometric phase being expressed as

$$\phi = \text{atan}\left(\frac{\Im(I)}{\Re(I)}\right) \quad (4.2)$$

If I is calculated for each master pixel position, the result is the interferogram for the scene [78].

As was seen in Figure 1.6 on page 36, the phase difference caused by Δr increases continuously on the ground as one moves away from the sensors. Our knowledge of this phase difference is constrained to $[0, 2\pi]$, resulting in

an apparent repetition of the phase-difference cycles with increasing distance from the sensors. These phase differences can be intensity-coded and represented as the fringes seen in Figure 1.6, with the range $[0, 2\pi]$ being mapped to $[0, 255]$ for an 8-bit greyscale representation, for example.

4.2.3 Ambiguity Height

The topographic height variation producing one full fringe, called the *ambiguity height*, is given in [78] as:

$$\Delta h_{2\pi} = \left(\frac{\lambda}{2}\right) \frac{R \sin \theta_i}{B_{\perp}} \quad (4.3)$$

with

$\Delta h_{2\pi}$	ambiguity height
λ	wavelength
R	slant range
θ_i	incidence angle
B_{\perp}	baseline component normal to look direction

The ambiguity height is an indicator of the sensitivity of the interferometer to height change. Conversely, it determines the amount of height error given a certain error in the phase measurement.

When the topographic height change from one pixel to the next exceeds the ambiguity height, the 2π -modulo-constrained interferometric phase "wraps" back on itself, and determination of the absolute phase difference becomes impossible without first "undoing" this folding-back phenomenon.

While this formula is correct, attention must be paid to the normal baseline: it is often assumed to be a function of range only. In fact, the normal baseline also varies with topographical height, making it technically incorrect to deduce topographical height from phase change without taking this effect into account. The problem is described in [80]. However, the normal baseline clearly varies more from near- to far-range for a given data take than over the vertical extent of a relatively small topographical feature. Thus, the constant- B_{\perp} assumption is considered valid to the first order, with the goal of demonstrating the baseline-dependent sensitivity of an interferometer to height variation.

4.2.4 Flattening

In order to simplify the interferogram for subsequent processing, as well as

eliminate or reduce areas of phase ambiguity due to fast height variation, the phase is *flattened* before continuing.

A fringe pattern can be calculated in advance for an existing terrain model, and subtracted from the actual measured interferogram [78]. The result is a *flattened* interferogram that is greatly simplified, with a residual structure depending only on the difference between the measured terrain and the terrain model used to synthesize the fringe pattern. If no prior information about the terrain topography is available, an ellipsoid-Earth (or even a flat-Earth) model can be used to remove the equivalent of the regular fringe pattern seen in Figure 1.6(a) from the interferogram. The remaining fringe pattern will be due to the topography relative to the model. When a better approximation of the topography exists, such as the result obtained for the same scene using stereo SAR, a fringe pattern can be calculated that will more nearly approximate the measured interferogram. In this case, the difference between the two will result in an interferogram with greatly diminished phase variation across the scene; the residual phase will depend entirely on the accuracy of the terrain model.

Subsequent processing of the flattened interferogram will be greatly simplified, which is especially advantageous in regions of rapidly-varying topography. In fact, phase unwrapping in mountainous terrain is virtually impossible using current methods unless a good height model already exists for that area.

4.2.5 Coherence Calculation

The coherence describes the statistical similarity between the two SLC images used to generate the interferogram, and hence the reliability of the interferometric phase as a measure of the topography. It is simply the normalised cross-correlation between the images [82][78], and it ranges from 0 (no correlation) to 1 (perfect correlation):

$$\gamma = \frac{\langle c_1 c_2^* \rangle}{\sqrt{\langle |c_1|^2 \rangle \langle |c_2|^2 \rangle}} \quad (4.4)$$

Each radar echo is the vectorial sum of multiple returns from within the ground resolution cell. Over a 3-D volume containing many point scatterers, small variations in the sensor position can rapidly alter the net returned signal. Such *volume scattering* processes, typical of vegetation, result in a decorrelation between the two scenes, which are taken from slightly different positions. If the decorrelation is too strong, the interferometric phase for those

areas will be equally meaningless, since it is required to be proportional to the path difference Δr from Figure 1.5 on page 35.

Coherence loss due to volume scattering is the major source of low coherence in the case of single-pass interferometry.

In the multi-pass case, scatterers that change significantly during the *time* between acquisitions can also cause low coherence. Over vegetation, the length of time leading to total coherence loss depends on the time of year, the type of vegetation, and the time between acquisitions.

Other sources of low coherence are listed in section 4.4.

4.3 Height Map Generation

For a given sample at position (x,y) in the master image, the range is known to a high precision, since this is the fundamental measurement of a radar sensor. The main goal of InSAR topographical mapping - as well as of stereo SAR - is the accurate determination of the slave sensor's range to the same point. In fact, the registration process itself (section 4.2.1) provides an estimate of the slave range for each master pixel. At this point it is already mathematically possible to solve the range-Doppler equations to obtain a height model, as was described in section 3.4 on page 57. The reason this is not done is because the baseline is too small to allow stereo SAR to yield a useful height model; the SNR would be much too low. Put another way, the master and slave images are too similar to allow a stereo-based height mapping to be accurate. The range shifts between master and slave will typically not exceed 1-2 pixels, and subpixel accuracy is limited.

Because the phase is presumed not to vary too quickly from pixel to pixel, the accuracy achieved by the registration process is sufficient for interferogram generation. The information contained in the interferometric phase provides, in turn, a very elegant way of determining the path difference more directly than is possible during registration. The conversion of the phase to topographical height is described in the following sections.

4.3.1 Phase Unwrapping

For a given point on the Earth's surface, the path difference Δr is typically many wavelengths (phase cycles) long, making it impossible to determine how many phase cycles it includes; the interferometric phase is constrained to vary between $[0, 2\pi]$ (see also Equation 4.2 on page 64). If one knows the position - and therefore interferometric phase - of a single point on the ground, then it is theoretically possible to integrate the phase gradient from that point

outwards, in all directions, to obtain the absolute phase for all image points. Each time a fringe boundary is crossed, the absolute phase has changed by 2π since the last boundary. This process is called *phase unwrapping* - and it is made much simpler when the interferogram has been flattened already. The result of the phase unwrapping process is knowledge of the absolute interferometric phase for every point in the scene, that is, the total number of phase cycles completed by the reflected signal over the path difference Δr (Figure 1.5 on page 35). The absolute phase is given in [82] by

$$\phi_{abs} = \phi + \phi_c \quad (4.5)$$

which is simply Δr expressed as an integral number of phase cycles (ϕ_c) plus the final partial phase cycle given by Equation 4.2 on page 64 (ϕ). The absolute phase is related to the path difference $\Delta r = (\lvert \vec{r}_M \rvert - \lvert \vec{r}_S \rvert)$ as follows [82]:

$$\phi_{abs} = \left[\frac{4\pi}{\lambda} (\Delta r) \right]_{mod2\pi} + 2k\pi \quad (4.6)$$

where k full phase cycles have been completed along Δr .

This 2-D integration process can be hindered by areas of low coherence, where the phase noise is so great that integrating through these areas would cause the absolute phase estimate to drastically diverge from reality. Indeed, it is necessary that the phase integration process proceeds only through those areas where the phase is changing slowly and continuously, that is, along a path that provides the certainty that our absolute phase value is always valid. Many methods of performing the phase integration while avoiding cumulative errors have been devised, for example in [34], [13], [19], [90] and [91]; the one used in the context of this work is the well-known branch-cut method, described in [36].

4.3.2 Geocoding

At this point the fringe pattern that was removed from the interferogram during flattening is added back to the (absolute) unwrapped phase again; this is called *unflattening*. The result is knowledge of the absolute phase over the whole image; the 2π -ambiguity problem has been resolved at this point. One step remains before it is possible to convert the phase variation to height information. Our ultimate goal is to determine the total number of phase cycles completed along the entire path between the ground and the sensors for all points in the terrain. This yields a measure of the ranges to those points with an accuracy on the order of the wavelength.

One problem remains, however: the constant k from Equation 4.6 needs to

be known for a single, arbitrary image point. Given this value, the absolute phase can be calculated for all image points, since the *relative* phase fluctuations have already been determined (phase unwrapping). In order to estimate its value, while at the same time performing a final fine estimation of the interferometric baseline, several carefully-chosen GCPs for the image are required. They serve as the calibration points for a combined model of the baseline- and phase-variation for the scene [78]. The points should optimally be well-distributed across the scene, and be chosen in areas without rapid phase variations, since the local mean is used for the phase value at a GCP location. Given the absolute phase for each GCP and the relative phase variation over the whole image, it is possible to estimate the absolute phase for all pixels as well as fine-tune the estimate of the baseline cross-track variation along the flight path. The process of calibrating the unwrapped, unflattened phase is called *phase calibration*.

The result of phase calibration is that one has a measure of the absolute range, as measured by both sensors, to all points in the scene. It now suffices to solve the system of equations described in section 3.4 on page 57 in order to obtain a geocoded DSM in a system of cartographic coordinates chosen by the user.

4.4 Weaknesses of the Interferometric Method

Phase ambiguity resolution, or unwrapping, is decidedly the most difficult step in DSM estimation using InSAR. Pixel-to-pixel phase gradients exceeding π must be avoided (see [36] and [78]) during unwrapping. Such situations have several possible causes [91] [34]:

- phase noise due to volume scattering
- phase noise due to low SNR (e.g. slope back surfaces, shadow, water)
- layover and sudden discontinuities in the terrain shape
- temporal decorrelation (multi-pass only)
- system (thermal) noise

The phase-unwrapping problem is a difficult one, and has no unique solution. While the branch-cut technique used here is a straightforward one, it was observed to require a fair amount of manual guidance and case-by-case adjustment in order to achieve satisfactory results (chapter 9 describes the results in greater detail).

Assuming the phase unwrapping was completed successfully, the subsequent phase calibration requires several GCPs to be available for the

image. This requires either setting up reflectors in the terrain to be imaged and accurately measuring their locations using, for example, differential GPS (D-GPS, see [47]), or the measurement of tiepoints in the images, given that accurate topographical maps are available.

The combined stereo-InSAR method tested during this work makes it possible to fully-automate the GCP generation process, making it possible to generate high-precision DSMs without the aid of maps or reflectors.

Chapter 5

Performing Interferometry Using Stereo SAR Results

5.1 Introduction

The weaknesses of the interferometric method for height extraction were outlined in section 4.4 on page 69. For single-pass InSAR, one of the greatest difficulties arises over rapidly-varying topography, creating densely-packed fringes and pixel-to-pixel phase discontinuities often greater than π on slopes facing the sensors. This, combined with the noise due to speckle and the system itself, makes phase unwrapping a difficult process in these regions. Over forested or otherwise vegetated areas, volume scattering sometimes leads to strong decorrelation even for a single-pass system, depending largely on the baseline and emitted frequency. For multi-pass InSAR, the addition of the time delay between acquisitions worsens the noise over vegetation, often making phase unwrapping impossible in these areas.

Indeed, it is the phase-dependent nature of interferometry itself from which arise both its strengths and its weaknesses. Problem areas in which the phase ambiguities cannot be resolved automatically are usually marked as off-limits for the phase unwrapping, so as not to falsify the phase integration; this creates holes in the final height model.

Because stereo SAR is entirely dependent on the measurement of disparities between the amplitude images, it is subject to errors completely unrelated to those produced by InSAR. It is therefore natural to assume that the height map produced by one method will contain errors not present in the other, and a fusion should be possible.

InSAR provides results of superior resolution - assuming the results can be obtained. Its weaknesses can be partially compensated for by an existing stereo DEM. Because of the rapid increase in the number of operational sensors and especially their ability to acquire data at variable incidence angles using electronic beam steering (ScanSAR), it is not far-fetched to assume that the data required for stereo processing will often also exist for scenes being

analyzed interferometrically, and vice versa.

Indeed, various authors have proposed methods for combining stereo- and InSAR-based DEMs. Use of the stereo DEM for InSAR phase flattening was suggested by [22], and merging of the stereo and InSAR DEMs by coherence-weighted averaging is described in [33] and [22]. A two-level multiresolution fusion is attempted in [15], in which the low-frequency DEM component of the stereo case is merged with the high frequencies from the interferometric DEM.

In 2003, a commercial feasibility study was completed by the European Space Agency (ESA), investigating the potential of combined stereo-InSAR for DEM estimation using ENVISAT's ASAR instrument [24]. Hindered by difficulties obtaining the required datasets, many projected results were unable to be completed. Successful experiments were carried out for two test sites (Pyrénées and Philippines), where the stereo DEM was used to calibrate the interferometric phase over low-coherence areas, effectively increasing the area that could be successfully unwrapped, as compared to InSAR alone. In the same instances, large parts of the combined processing chain had been successfully automated. Other test sites proved disappointing, partly due to unavailable data, partly due to problems generating the stereo DEM, InSAR DEM, or the combined DEM. However, the results were at least in part promising.

All of the experiments listed above were performed using spaceborne sensors. No fully-automated stereo-InSAR scheme has yet been devised as of this writing; certainly, no such reported studies using airborne data had been encountered at the time of this writing.

The phase noise due to volume scattering or temporal decorrelation does not present a critical problem to a stereo processor. Steep terrain, while also a difficult stereo-matching problem, does not cause local errors to propagate as far as they can with InSAR. This is mainly due to the benign nature of matching errors: errors in one region have little or no effect on more distant regions, especially within a multiresolution framework. InSAR phase unwrapping errors, on the other hand, propagate over large distances during phase unwrapping. Outside of areas where no data is available at all - radar shadow, which presents a problem to both height extraction methods - the success of stereo SAR is virtually guaranteed *as long as the acquisition geometry is suitable* for matching. This is a key point: under the right conditions, which incorporate the terrain characteristics, the DEM produced will be nearly free of masked-off areas except where shadow makes this impossible. However, choosing the wrong combination of incidence angles for the terrain may make the stereo DEM highly unreliable or even impossible

to obtain. Everything depends on the type of matching method chosen, as well as the acquisition geometry [86].

If a stereo DEM and some measure of the confidence of the height values already exist for the area covered by an InSAR scene, they can be used in three independent ways to assist the InSAR processing without compromising the comparatively high resolution of the resulting InSAR DEM. These methods are described in the next three sections: phase flattening, phase and baseline calibration using automatically-generated GCPs, and data merging. These points of influence on the InSAR processing chain are indicated by the arrows in Figure 5.1, with the input requirements being the stereo DEM, match confidence (introduced in section 2.4.3 on page 50), or both, depending on the step being performed.

5.2 Phase Flattening

Any DEM representing the real topography more closely than an ellipsoidal flat-Earth approximation can only aid - and not hinder - the phase-unwrapping process. With the height model one can calculate the expected interferometric phase and subtract it from the real interferogram, thus lowering the fringe frequency and reducing or eliminating the absolute-phase ambiguity caused by the $[0, 2\pi]$ constraint on the interferometric phase. This is the most important use of a low-resolution DEM during InSAR processing of a scene with steep terrain causing high fringe rates, since it has the greatest effect on the success of the phase unwrapping. Its position in the InSAR chain is shown at point 1 in Figure 5.1.

5.3 Ground Control Point Generation for Phase- and Baseline Calibration

Phase integration produces an unwrapped-phase image, which is directly related to the topographical variation. Two problems remain at this point: the phase and baseline both need to be calibrated before absolute heights can be calculated. Several GCPs with well-known phase values are therefore required [78]; the problem was first introduced in section 4.3.2 on page 68.

The unwrapped, unflattened phase requires calibration because at no point during InSAR processing was the absolute number of phase cycles between the ground and the sensors known. Phase calibration provides an absolute reference for the phase model. Given a number of GCPs at locations with low local phase variation (i.e. flat areas), and distributed over the scene, the phase

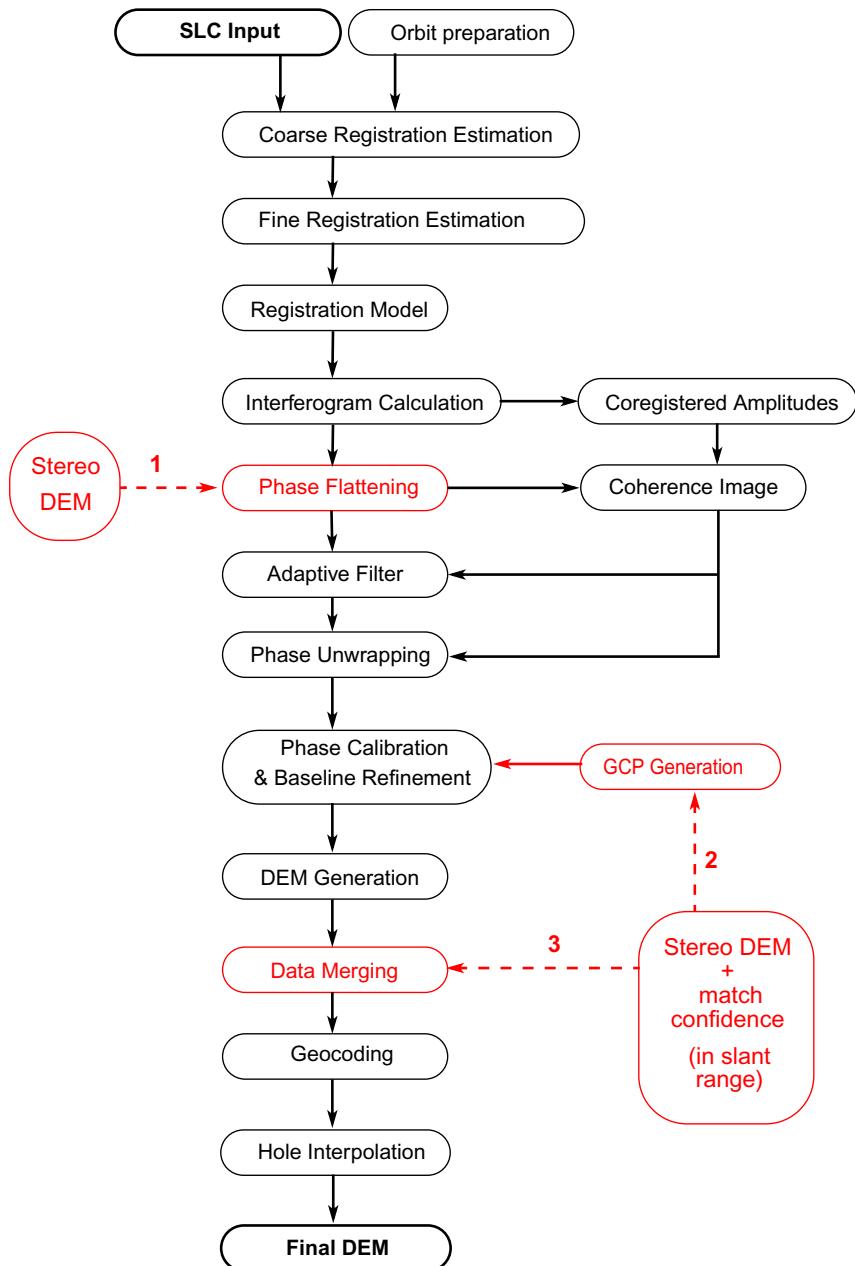


Figure 5.1 Stereo-assisted InSAR processing chain. Points of influence from the stereo data on the InSAR chain are shown in red and numbered 1 through 3.

can be calibrated.

Additionally, the state vectors of even the best current spaceborne sensor are not known accurately enough to guarantee a wavelength-accurate transformation from differential phase to topographic heights without calibration. For the baseline model used here, the cross-track component of the baseline is assumed to vary linearly along azimuth, and this baseline-variation model is calibrated using the phase values at the GCP locations. At least three GCPs are needed, but six or more are preferred. A least-squares fit to the phase-baseline model is calculated, as described in [78].

An algorithm has been implemented for the automatic generation of the required GCPs, using the available DEM and match confidence products calculated during stereo processing (see section 2.4.3 on page 50). The use of these two outputs for the GCP generation is indicated by the number 2 in Figure 5.1.

Traditional GCPs, such as reflectors placed on the ground in view of the sensors or tiepoints collected from maps, have positions that are well known, but such points are generally difficult or time-consuming to collect. It is preferable to sidestep this process, since a height model is already available, namely the stereo DEM. Naturally, GCPs chosen randomly from the stereo DEM would vary greatly in quality, because of the imprecise nature of the stereo result. The wavelet-based matching method used here fortunately includes a measure of the stereo match confidence, which provides an indication of how accurate the derived heights may be. It will be seen that **low match confidence is strongly correlated with high DEM error**; this confidence measure is **used to filter out points considered the most unreliable**, and in this way produce dozens of GCPs of moderate-to-good quality distributed across the scene. Naturally, "good" is a relative term: the selected points will still exhibit a high standard deviation due to limitations of the stereoscopic method, as well as the somewhat weak correlation between high match confidence and the highest-quality stereo heights (discussed in chapter 8). In addition, if larger areas of the stereo DEM are over- or underestimated, it may nevertheless be necessary to use points from these regions to ensure a good distribution across the scene.

A compensating factor for the high variability of the stereo heights is that a large number of them are available, making calibration of the phase-baseline model feasible. The least-squares nature of the model fitting tends to result in an accurate calibration as long as the mean height error of the selected GCPs is approximately zero. If this is the case, one can expect the calibration based on their positions to be successful, and the subsequent phase-to-height conversion will produce correct heights.

To prevent stereo-DEM points with enormous errors from upsetting the phase calibration, only those values fulfilling certain criteria are selected:

- Stereo DEM values associated with low match confidences are ignored, with the confidence threshold being empirically determined.
- DEM values stemming from areas of low SAR amplitude, such as is typical of water and shadow, are also not used.
- Local phase stability is required, to within an empirically-determined threshold. This requirement is imposed such that the local phase average, used during calibration, is representative of the true local phase.

The existence of the stereo DEM and the associated match confidence **allows the calibration to occur automatically, making InSAR topographic mapping over "difficult" terrain**, for which GCPs or image tiepoints cannot be easily obtained, a **simpler and less costly** procedure.

The minimum number of GCPs produced by the search algorithm can be chosen by the user, making it possible to produce a high number of GCPs. However, because points that satisfy the selection criteria are often found within small areas scattered across the scene, forcing the algorithm to produce too many points may result in point clustering. Too many points in the same small area are not only redundant, but could bias the phase calibration if the DSM heights there are in fact not of the best quality. Therefore, the ideal solution to the problem of how many GCPs to produce can be stated as follows: enough points are needed to bring the mean height error of the selected locations to nearly zero, while avoiding point clustering.

5.4 Data Merging and Interpolation (Hole-Filling)

5.4.1 Data Merging

Because of residues due to phase noise [34], a DEM resulting from InSAR will often contain areas that had to be excluded from the phase integration process. These data holes can reach large dimensions over vegetation even for single-pass InSAR, if the interferometric baseline was long enough to create severe phase decorrelation. Some authors have experimented with stereo-InSAR DEM fusion based on coherence-weighted averaging [22][33]. The idea is that low InSAR coherence is correlated to lower-quality InSAR heights, and consequently more weight is given to the stereo value for those areas. It was found that for the datasets and test sites used in this work, the low-coherence areas were blocked off altogether during the phase unwrapping, leaving data holes. The medium-to-high coherence areas showed no particular

correlation to the areas of low height error. InSAR heights of low-to-medium coherence usually turned out to be more accurate than stereo heights of medium-to-high confidence. Therefore, InSAR heights were preferred over stereo heights wherever they are available.

With respect to the data holes, the InSAR DEM can be *partially* filled with values considered reliable from an existing stereo DEM; this is referred to as data merging (point 3 in Figure 5.1). Certainly, even in areas of high match confidence, stereo heights are sometimes unreliable, thereby worsening the overall DEM statistics. Nonetheless, **it is preferable to have a DEM of lower mean quality but very good coverage to one that is very accurate but containing extended data holes.** This is especially true when one has the possibility of marking the regions containing the stereo heights, therefore providing the end-user the means to decide whether to use the height values or not.

Provided that the stereo DEM contains height values for an InSAR-DEM data hole, a technique needs to be developed to allow for the inclusion of the stereo information within the hole. The criteria used for the GCP selection (above) are retained in this step, but without the constraint on the phase stability. That is:

- Stereo DEM values with low match confidences are disregarded.
- Stereo DEM values stemming from areas of low SAR amplitude are disregarded.

The result of the selective merging is an InSAR DSM whose data holes have been *complemented* with values coming from the stereo DEM, although in most cases not filled entirely. This step is performed before geocoding, since the match confidence was generated in the slant-range geometry.

5.4.2 Interpolation

After the data merging, the slant-range DEM is geocoded according to the range-Doppler system described in section 3.4 on page 57.

Next, in order to smooth over the "spotted" appearance of the DEM holes arising from the stereo-merging step, it is now desirable to fill in the small holes using interpolation. The larger holes are left unfilled, since the likelihood of introducing errors increases with increasing hole size. In the center of a large hole, no reliable information to aid the interpolation exists. The maximum hole size for which interpolated filling is allowed is user-adjustable. In general, the flatter the topography, the larger the hole-size threshold can be reasonably set.

For the calculation of the locations of the large holes, defined as having a diameter greater than an empirically-determined threshold, a *closure* operator

(see, for example [76], for a description of this technique) was used on a binary hole mask for the DEM. The resulting hole mask contained only the large holes, thus indicating the locations of the pixels to ignore during the subsequent hole-filling. The closure operation was performed in *The Mathworks' MATLAB* [61] software package.

The hole-filling algorithm must take into account all available nearby information, which includes InSAR as well as stereo height values. The problem of hole-filling is a topic of research itself; it has been well-studied by researchers such as Caselles et al. [11], Bertalmio et al. [6], and Pérez et al. [69]. Two methods were tested:

- The solution to Laplace's partial differential equation was applied to the hole-interpolation, with the data points within and around a hole as fixed boundary conditions. This technique is described in [11] for example. MATLAB's *roifill* function is an implementation of this technique.
- Cubic interpolation on a Delaunay-triangulated network was also tested, the network nodes being defined at the data locations. The MATLAB function *griddata* was used to perform the interpolation. The triangulation algorithm at the heart of *griddata* is described in [2]. This technique was used for our work.

Both methods produce smooth interpolations, that is, their first derivatives do not contain discontinuities. However, the Delaunay-interpolated results systematically contained fewer artifacts, and appeared more "natural". Therefore, this technique was used in the context of this work.

It is, of course, also possible to fill in the InSAR-DSM holes directly with the stereo height values, where the latter are available. This is undesirable, since the weaknesses of the stereo method (such as low resolution, false heights assigned to shadows) are at least partially avoidable using the above-mentioned interpolation. This technique was used as a comparison for the more balanced approach to the data fusion used here.

The term hole-filling, as it is used in the context of this work, refers to the combination of data merging and subsequent interpolation of the smaller holes. Filling InSAR holes with stereo values wherever they are available without regard to their quality will be referred to as "blind" hole-filling; the combined partial data merging/cubic interpolation technique will be called "smart" hole-filling.

Chapter 6

SAR Image Simulation: a Virtual Testbed

6.1 Introduction

Although the ultimate goal of our work is to use real data acquired by airborne and spaceborne sensors to estimate surface height models, two practical reasons to work with simulated stereo datasets exist:

- Software testing and validation: the output of the stereo processor is a DEM. Verification of the accuracy of such height models based on real datasets is difficult, since a perfect height reference for real terrain cannot exist. Simulated data are created synthetically from the outset, making reference information unnecessary. If the stereo processor is functioning correctly, it will produce the correct DEMs as output.
- Airborne flight campaigns, when possible, open up the possibility to test new flight geometries, software, sensors, and so on; this is not the case with spaceborne sensors, which are relatively expensive and new missions rare. Simulating data for a number of possible flight configurations can help mission planners decide on a flight plan meeting the requirements of the particular processor to be used, for example. The choice of baseline, tracks to be flown, sensor orientation, and optimal stereo intersection angle are especially important when planning a flight campaign.

For these reasons, a SAR amplitude-image simulator called *radsim* (radiometric simulator) was developed in-house; it is described in [78] and will be summarized here.

Simulations of interferograms can be useful as well. During stereo-assisted InSAR processing it is also necessary to simulate the interferogram as a means to flatten the phase. The stereo DEM is used as the terrain model, and the flight path (state vectors) is given as well. Using the in-house interferogram simulator *syngram* [79], a synthetic interferogram is generated as a by-product of the InSAR chain. Section 6.3 provides an example of the qualitative

difference between a synthetic and a real interferogram.

6.2 Amplitude Image Simulation

The radar image simulator (*radsim*) employed in the context of this work requires two sources of input information:

- an existing surface height model for a given terrain, in the form of a digital raster grid (i.e. a DEM), and
- a formatted text file containing the state vector information for the sensor illuminating the surface model, as well as other sensor-related parameters.

The simulator cycles through the pixel positions of the DEM, creates facets defined by four adjacent pixels, and for each facet calculates the local illuminated area and local incidence angle. Upon confirmation that the pixel does not lie within radar shadow, its illuminated area is added to the running sum at the corresponding range/azimuth bin, the position being given by the transformation relating the map and radar geometries.

The simulator therefore takes into account that many map-geometry pixels may contribute to a single radar-geometry pixel (e.g. front slope of a mountain), or alternatively, that one pixel in map geometry may span several pixels in slant range (e.g. back slope of a mountain). The total illuminated area A for a position (rg, az) is calculated, that is, the total number of map-geometry pixels illuminated for the current range/azimuth position.

Because radar backscatter is proportional to the illuminated area contributing to a pixel, the simulations generated by *radsim* look very much like real SAR-amplitude images.

6.2.1 Qualitative Comparison of Real and Simulated Images

If one desires to simulate an image from a sensor for which real data exist, an input DEM for that area is required, as well as the state vectors of the sensor to simulate. In this way, stereo processing of a simulated dataset based on an existing real pair can serve as a control experiment. As long as the sensor characteristics and state vectors are available, sensor images can be simulated, and these compared to the real images and resulting DSMs, if this is desired. An example of a simulation of a real image is given in Figure 6.1(b) for an X-band airborne sensor from AeroSensing [43]. The input DSM used to generate the actual simulation for this example is of proven high accuracy, and is considered a surface model, although the tree stand heights are more likely

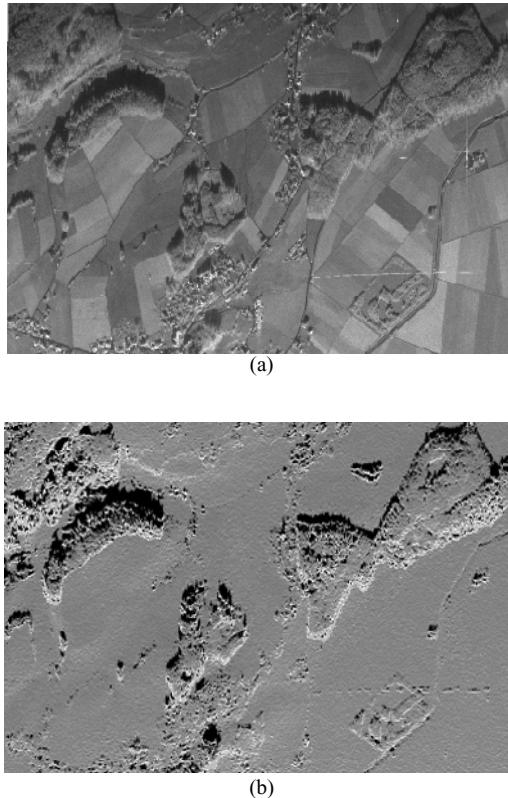


Figure 6.1 Real image and simulation of the same scene: (a) real image (b) simulated image

somewhat underestimated [26]. The flight path simulated was the same as the one flown during acquisition of the real data shown in Figure 6.1(a).

The main difference in the appearance between the real and simulated images is the lack of information in the simulation over areas where the backscatter depends on the surface characteristics. This is to be expected, since no information about the landcover types is introduced into the simulation. In spite of the comparative bareness of the simulations, the matching algorithm was observed to work well for the simulated images, even in image areas with few prominent features or textural variations.

6.2.2 Verification of Matching Algorithms and Geocoding

The stereo processing chain essentially consists of two steps: image-pair matching, and geocoding.

The matching step estimates the 2-D vector transformation that brings one

image into alignment with the other. In the case of a simulated radar amplitude pair, the disparity field calculated relative to the slave image can be used to resample it into the geometry of the master. In the case of a perfectly-functioning matching algorithm, the resampled slave should overlap perfectly with the master. This is simple enough to verify, either (1) by flickering back and forth between the two images and looking for coregistration errors, or (2) overlaying the master and remapped slave images in two colors; misregistration will manifest itself as color shifts at object edges and boundaries.

The true test of the stereo processor under simulated conditions, however, comes after geocoding. When the disparities are converted to a height map as described in section 3.4 on page 57, one should theoretically re-obtain the input DEM. If this ideal is sufficiently approached - given the impossibility of reconstructing shadow zones, and allowing for a certain amount of noise introduced by the stereo matching - then the internal consistency of the processing chain has been demonstrated.

6.2.3 Simulation of Variable Acquisition Conditions

Given a DEM, it is possible to simulate the data required for the stereo processor. One can perform stereo processing on these data, thereby providing the means to evaluate the performance of the processor without having real data available. The DEMs resulting from stereo processing of such simulated data are indicators of the results obtainable for real data acquired over the same terrain, under the same viewing conditions.

When testing various simulation configurations, it is useful to separate the sensor configuration from the terrain characteristics. From the sensor side, the two incidence angles can be varied, thereby altering the range resolution and stereo intersection angle. On the other hand, the sensor configuration can be held constant and the terrain height variability altered.

Figure 6.2 shows several examples of different acquisition geometries and terrain types. Figure 6.2(a) illustrates variable sensor incidence angles for a given terrain type; Figure 6.2(b) assumes a given viewing geometry, and allows for terrain variability. The robustness of the stereo matching algorithm can be tested by estimating disparity maps for all of these cases, geocoding them, and comparing the resulting DEMs with the input DEM used to generate the simulations in the first place [75].

The use of simulated pairs can also make advance planning for an airborne flight campaign simpler: flight paths and sensor orientations can be chosen such as to be optimally adapted to the stereo processor at hand.

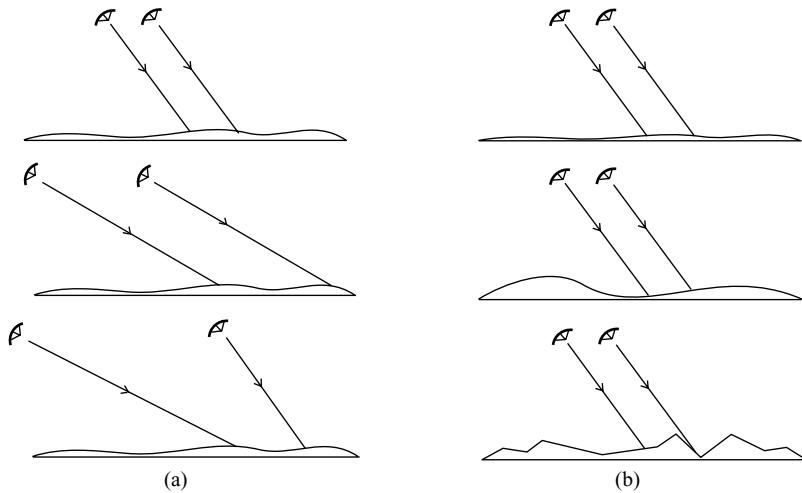


Figure 6.2 Variable acquisition conditions: (a) variable incidence angles (b) variable terrain types

6.3 Qualitative Comparison of Real and Simulated Interferograms

Synthetic interferograms are produced for the phase-flattening step during InSAR processing. Comparing a real (unwrapped and unflattened) interferogram with the synthetic interferogram calculated for the same area using *syngram* (Figure 6.3), one major difference is evident: the presence of more noise in the real interferogram for certain areas. The phase coherence is low over the treetops in this example because the baseline was large enough (about 1.9 m) for the volume scattering effect - which is different for each sensor - to have become visible as noise. The "volume scatterers" are mainly leaves and small branches for X-band sensors. Decreasing the baseline increases the chance that the scatterers contributing to the signal received by one sensor will also be the ones contributing to the signal of the other, hence lessening the phase noise. Completely decorrelated phases look like random noise in the interferogram. As one decreases the baseline, the phase correlation increases, and at some point a fringe pattern will emerge from the noise, indicating increased correlation between the two received signals.

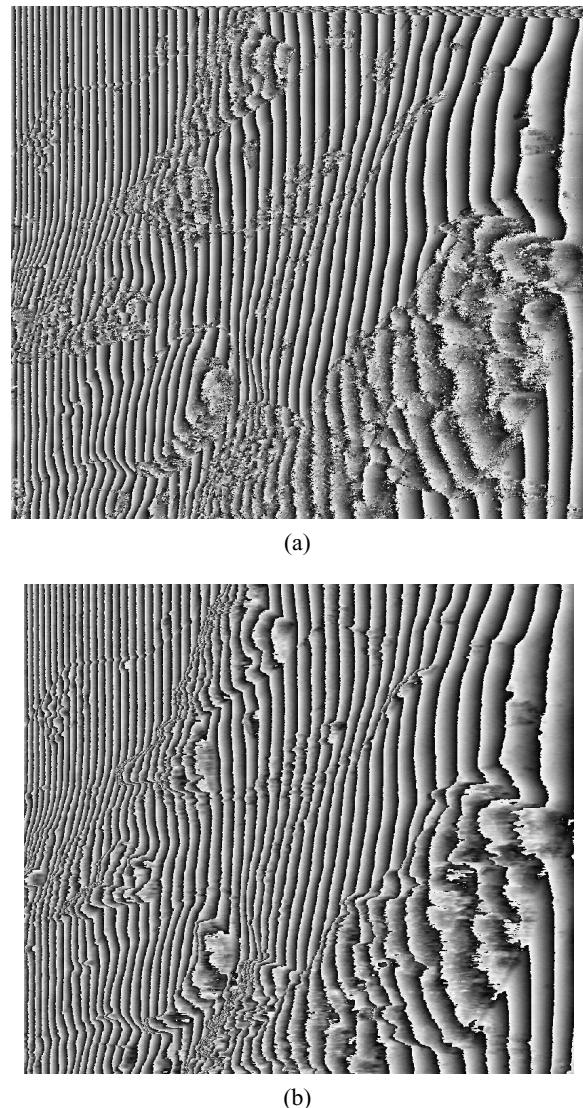


Figure 6.3 Real interferogram and simulation for the same scene: (a) real interferogram (b) simulated interferogram

Chapter 7

Characterization of Test Sites, Datasets, and Height References

7.1 Introduction

During the course of this work, data from one airborne and two spaceborne SAR sensors were analyzed for three Swiss test sites, each site with its own characteristics. In order to validate the height maps obtained by stereo and InSAR, the best available reference height information had to be used. The DEM references consisted of a high-resolution DSM generated using an airborne C-Band interferometer, a laser DSM, a national-scale DTM based on 1:25000 topographic maps, and a national-scale DSM generated by spaceborne C-Band interferometry. The references were adapted to the scale and resolution of the DSMs experimentally produced, except in one of the airborne cases, for which only the national-scale DTM was available. To partially compensate for this, individual GCPs were used as additional height references.

The goal of this chapter is to describe the test sites, the SAR data acquired, and the height reference information available for the three test sites, thus providing a context for the results in the next chapters. All reference DEMs and GCPs are projected in the Swiss national cartographic reference system, which is an oblique Mercator projection using the CH1903 datum shift parameters with the Bessel ellipsoid from 1841 [84]. Heights are specified above the Swiss geoid CHGEO98 for all references, described in [62].

7.2 Overview of the Test Sites

The approximate perimeters of the three test sites are shown in Figure 7.1. The two areas marked in white were imaged using the airborne AeS-1 sensor described in section 7.3.1. The green perimeter corresponds to the ERS-1/2 InSAR scene, while the red and blue areas are the ENVISAT stereo pair.



Figure 7.1 Overview of the test sites. Airborne scene perimeters are indicated in white, spaceborne sites in green, red, and blue. [Background ©The General Libraries, University of Texas at Austin]

7.3 Airborne Case: Küttigkofen

Of the three test sites, the Küttigkofen test site was studied in the most detail for a number of reasons:

- Three independent sources of high-quality reference data are available for validation purposes: a laser (LIDAR) DSM, a interferometrically-derived C-band DSM, and ground control points from 1:25000 topographic maps as well as national survey points.
- Both InSAR and stereo acquisitions are available from the same sensor (AeS-1) and flight campaign. The airborne platform flew its cross-track-InSAR system along several parallel tracks, making it possible to perform interferometry on the same images that were stereo-processed.
- A stereo data set was found to have been acquired with nearly the optimal geometric configuration (described later).
- The area consists of rolling hills, forests, and some housing, making it a balanced mixture of simple and more difficult terrain for stereo and InSAR processing.

- The literature concentrates on mainly spaceborne acquisitions for stereo-InSAR fusion; airborne-SAR has the ability to provide much higher-resolution DSMs and can often be more easily acquired when and where it is needed.

Details pertaining to the sensor, flight configuration, and datasets are given in the following sections.

7.3.1 Experiment Configuration

The acquisition system used in Küttigkofen was an X-band triple-antenna system mounted on an aircraft operated by AeroSensing GmbH, or AeS, now Intermap Technologies GmbH [43]. Figure 7.2 shows the antennas mounted on the aircraft, as well as the selectable interferometric baselines. The normal baselines are shown, which are defined as the projection of the antenna-to-antenna baseline onto a line normal to the beam direction. The AeS-1 SAR system is summarized in Table 7.1. Only data from the longer 1.89 m baseline

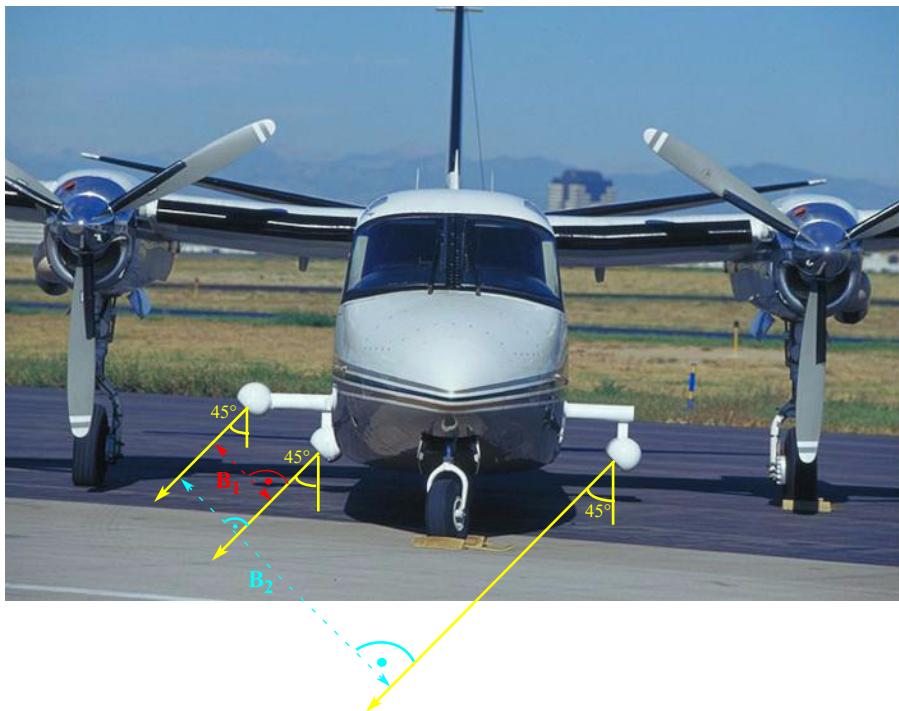


Figure 7.2 AeS-1 system and choice of baselines. $B_1=57\text{ cm}$, $B_2=189\text{ cm}$ [© Intermap Technologies GmbH]

was acquired by AeroSensing for the Küttigkofen campaign.

The flight geometry providing the stereo pair is illustrated in Figure 7.3. The InSAR and stereo datasets were acquired during the same flight campaign on April 22, 1999. Stereo SAR was possible because several tracks were flown, two of which provided a useful same-side-stereo baseline of 1 km and

Table 7.1 AeS-1 dual-antenna airborne SAR system parameters

Center frequency [GHz]	9.6
Polarization	HH
Bandwidth [MHz]	400
Incidence Angle [$^{\circ}$]	45
Pulse repetition frequency (PRF) [Hz]	10320
Radiometric resolution [dB]	1.8 - 1.0
Positioning	D-GPS and IMU
Selectable normal baselines (at 45°) [m]	0.57 or 1.89
Processed pixel spacing in range and azimuth [m]	0.37 and 0.58

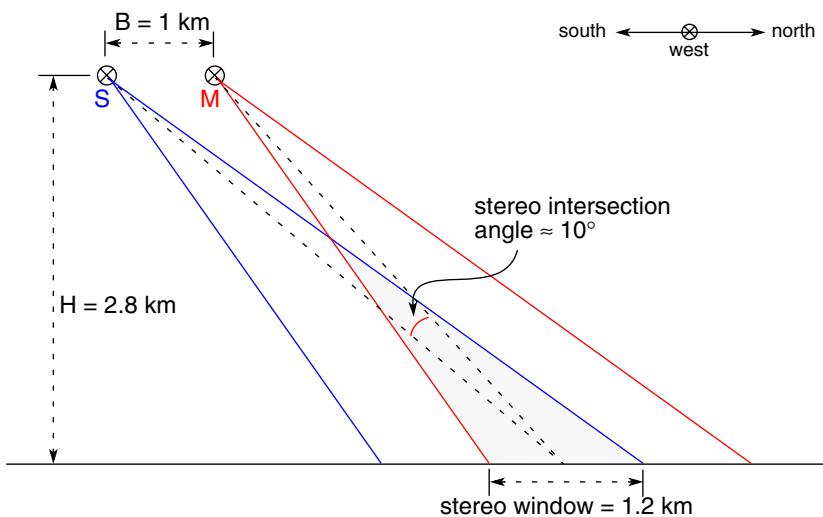


Figure 7.3 Küttigkofen flight geometry

a stereo window (overlap) width of about 1.2 km in the ground range direction. A stereo intersection angle of about 10.0° at the center of the stereo window was achieved in this way.

7.3.2 Data Acquired

The Küttigkofen test site and the perimeters of the datasets acquired are shown in Figure 7.4. The site is situated in the Swiss midlands to the north of Berne. The $\sim 2 \times 2$ km scenes consist mainly of rolling hills, forest stands, housing, and a man-made canal. The topographical variation of about 145 m across the scenes used for InSAR and stereo processing would normally not prove particularly challenging. However, the forest stands provide two challenges for SAR stereo and InSAR: the sudden height discontinuities at the borders, and the volume scattering occurring within the treetops. Radar shadow behind, and signal saturation in front of the forest stands can be seen in the amplitude images (Figure 7.5). This presents difficulties for both stereo and InSAR processing.

The stereo master of Figure 7.5(a) was also used as the InSAR master; the InSAR slave was from the other antenna mounted on the airplane's fuselage, forming the 1.89 m interferometric baseline. Thus, the area of overlap between the two acquisitions corresponded to the region where stereo-assisted InSAR is possible.

7.3.3 Height References

7.3.3.1 The DoSAR DSM from Infoterra

In May 1994, the German company Dornier (now Infoterra GmbH) and the Remote Sensing Laboratories (RSL) of the University of Zurich, Switzerland, produced an InSAR-derived height model of an area in Switzerland north of Berne that included the Küttigkofen site. Two 5.3 GHz C-band radar antennas were attached to an airborne platform in the cross-track orientation, with an interferometric baseline of 1.0 m and a 65° incidence angle. The resulting height model, shown in Figure 7.6, is considered a DSM mainly because the large off-nadir viewing angle guaranteed that reflections were received from within the upper forest canopies. The DoSAR system and the DSM derivation are described in [26].

The forest-covered areas stand out in Figure 7.7, which is the difference between the DoSAR DSM and the DHM25, Switzerland's national terrain model (described in section 7.4.3.1 on page 95) for Küttigkofen. The red zones indicate differences of about 25–40 m, marking the presence of tree canopies.

Heights for the DoSAR DSM are given above the Swiss geoid, whose

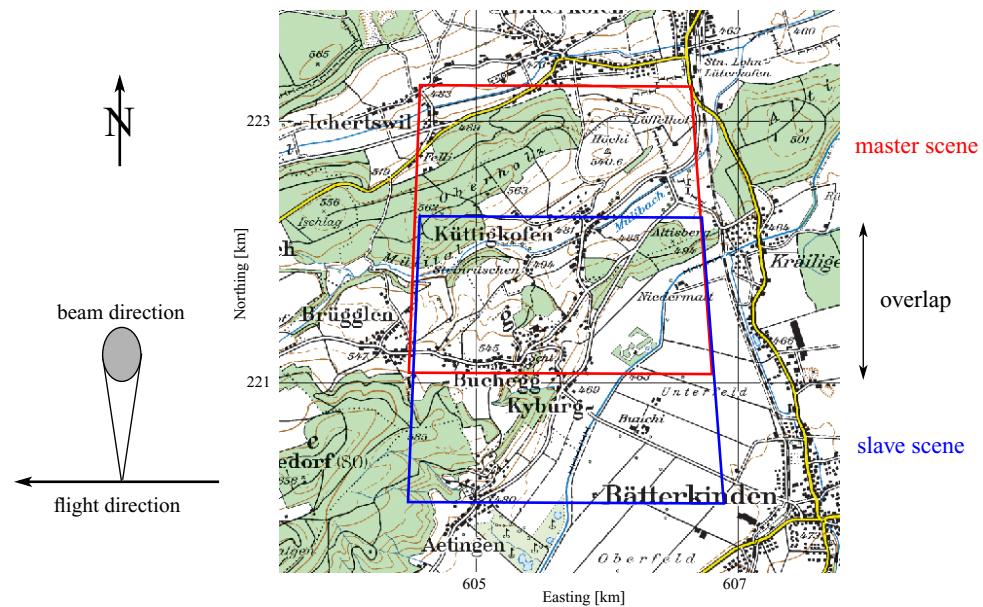


Figure 7.4 The Küttigkofen test site, with master and stereo slave scene perimeters [© SwissTopo]

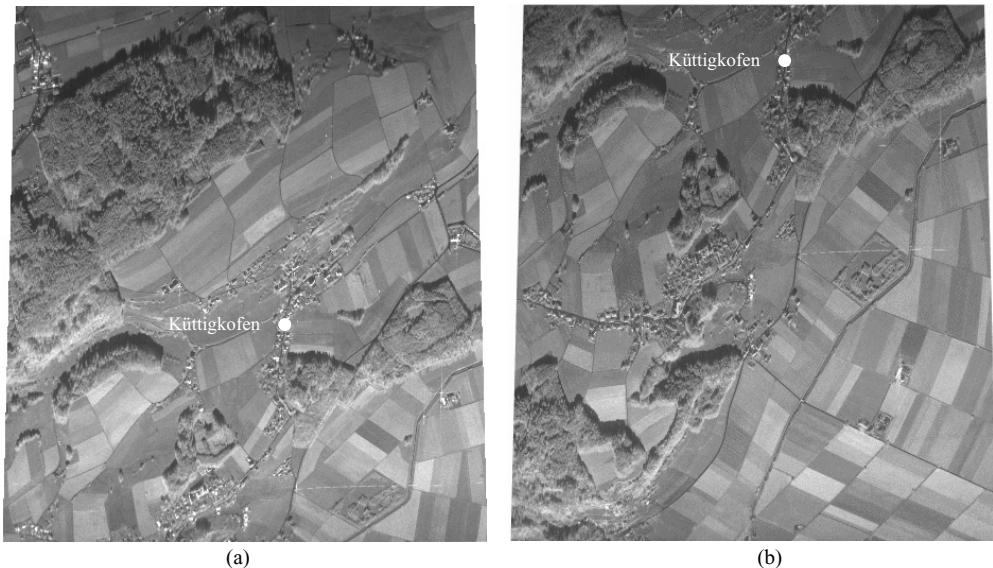


Figure 7.5 Küttigkofen SAR slant-range amplitude images: (a) stereo master, also used for InSAR (b) stereo slave. Azimuth/range looks: 4/1. Pixel spacing: range: 37 cm, azimuth: 60 cm [© Intermap Technologies GmbH]

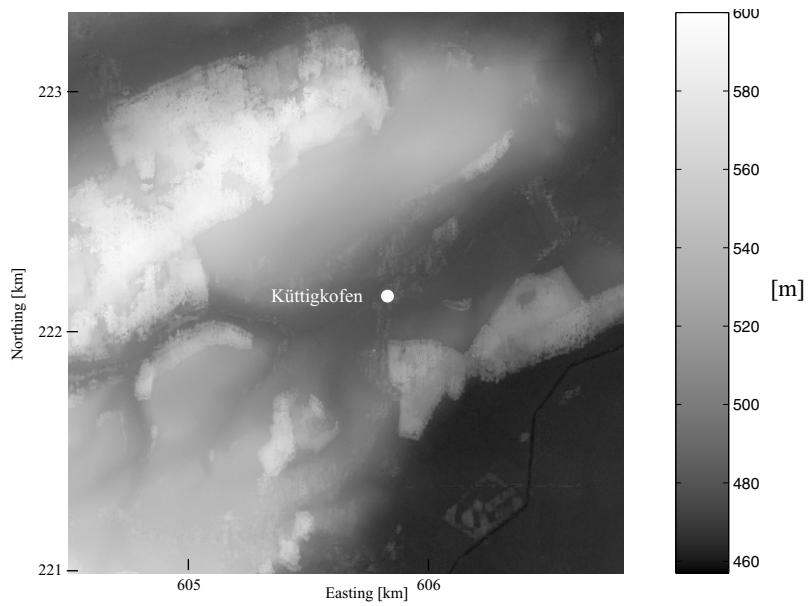


Figure 7.6 The DoSAR DSM for Küttigkofen. Pixel spacing: 1.0 m [© Infoterra GmbH]

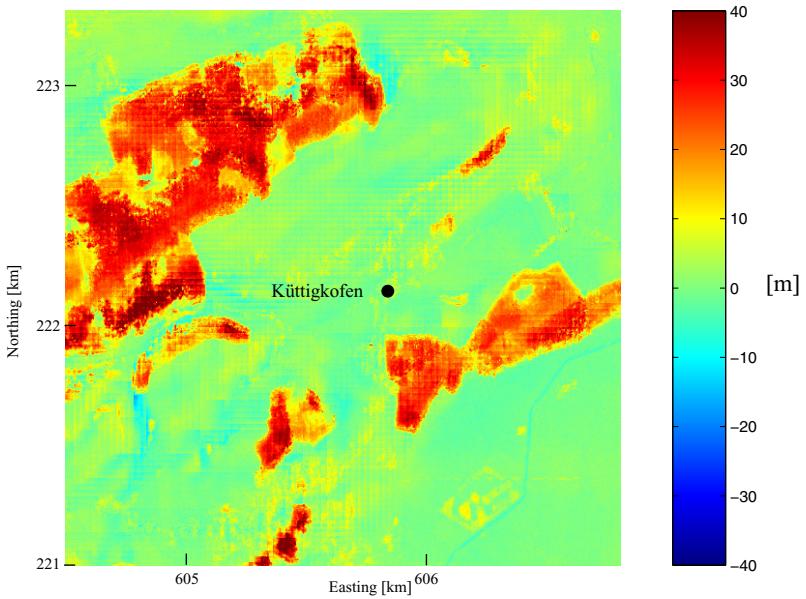


Figure 7.7 Difference between the DoSAR DSM and DHM25 for Küttigkofen

undulations vary between about 20 and 30 cm above the local Swiss ellipsoid for the Küttigkofen site. This is negligibly small compared to the stereo-DSM errors obtained later. Indeed, even the InSAR DSMs obtained will be seen to have error fluctuations roughly ten times larger than the geoid-ellipsoid differences. The DoSAR DSM is a high-resolution surface model of well-established quality. Furthermore, its representation of vegetation surfaces is hole-free, similar to the DSMs later obtained using the AeS-1 airborne system. Therefore, it served as an important reference model for our experiments in Küttigkofen.

7.3.3.2 The Laser DSM from TopoSys

Using an airborne laser scanner, or LIDAR, such as was described briefly in section 1.3 on page 23, TopoSys [44] created a height model for an area including Küttigkofen. It is a first-pulse height model, meaning that the canopy heights are represented wherever the narrow laser beam (with a footprint on the order of 1 m) is reflected from a part of the canopy surface. The ability of the laser scanner to capture greater detail is evident by comparing the DoSAR DSM (Figure 7.6) to the laser DSM (Figure 7.8). Besides improved planimetric resolution as compared to InSAR, the laser DSM is also virtually unaffected by shadow, the laser having a angular scan range of only ± 7 degrees from the vertical. Because of this, the laser DSM reveals holes in the canopy surfaces, which are missed with the InSAR technique.

While the laser DSM is a more accurate height reference than the DoSAR DSM over solid ground, the DoSAR DSM is used as the primary reference in the work presented later. This is justified because its representation of the vegetation canopies more closely resembles that produced by the stereo and InSAR systems used here, and because the difference in accuracies between the laser and DoSAR references fall well within the error fluctuations of the DSMs generated.

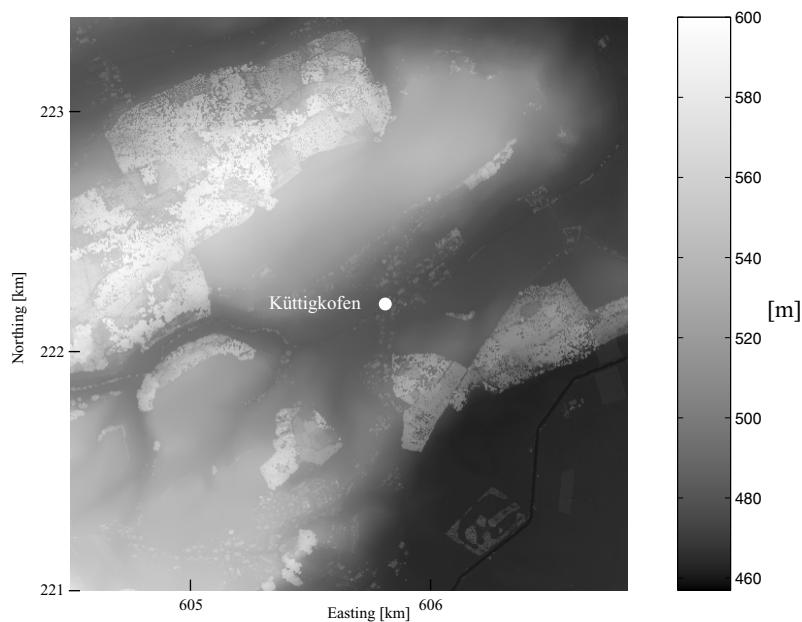


Figure 7.8 The laser DSM for Küttigkofen. Pixel spacing: 1.0 m [© TopoSys]

7.3.3.3 Ground Control Points

Eighteen tiepoints were collected for the Küttigkofen site from both SwissTopo's 1:25000 topographic map and from a 1:10000 digital cadaster map, also from SwissTopo. Their positions are shown against the 1:25000 map in Figure 7.9. While none lie within forested areas, their positions are

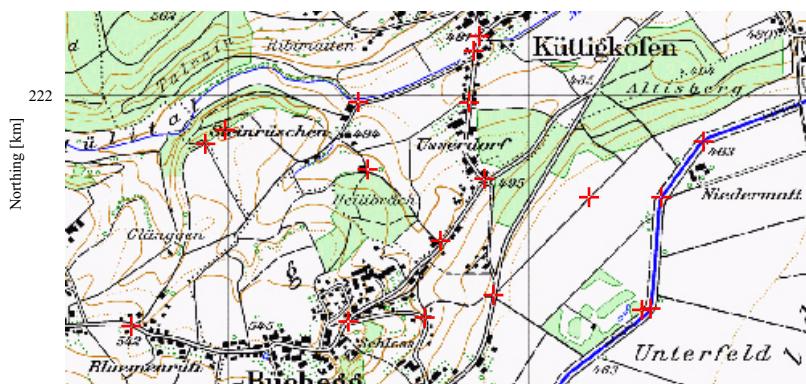


Figure 7.9 Reference GCP positions used for Küttigkofen [Background © SwissTopo]

provided with a vertical accuracy of 10 cm, making the points a reliable set of reference GCPs.

7.4 Airborne Case: Emmen

7.4.1 Experiment Configuration

The Emmen airborne data were acquired on November 14, 2000 using the same AeS-1 system also used in Küttigkofen, but from a mean altitude of approximately 4.3 km. Also, the shorter 57 cm baseline (shown in Figure 7.2 on page 87) was flown this time.

Stereo data were obtained during the same campaign, using a parallel-track configuration also similar to that flown in Küttigkofen. However, due to a 1.5 km increase in altitude in Emmen and a slight decrease in the stereo baseline as compared to Küttigkofen, the stereo intersection angle at the center of the stereo window was reduced to 6.9° , down from 10.0° .

The Emmen flight geometry is shown in Figure 7.10. Three parallel tracks were flown northwards, theoretically allowing for three different stereo coverages. However, only the outer two tracks were stereo-processed in order to maximize the stereo intersection angle, already considered low at 6.9° . The chosen master and slave pair is indicated in red and blue in the figure, with

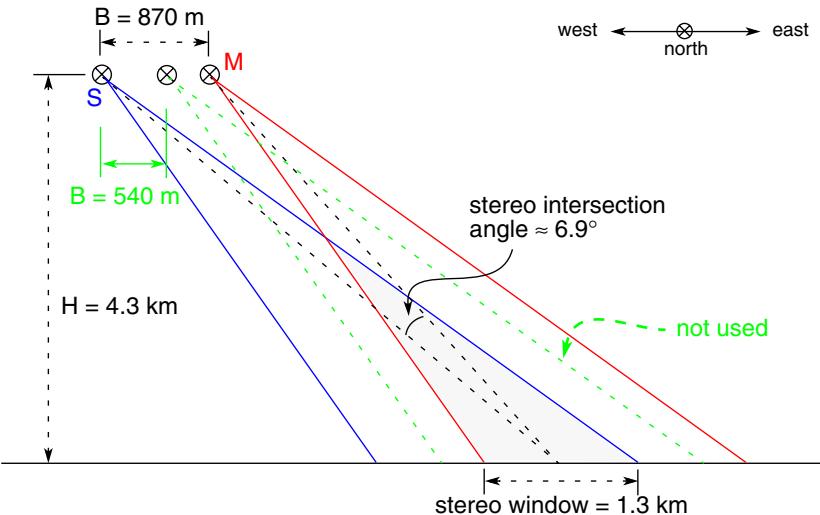


Figure 7.10 Emmen flight geometry

green representing the beam from the unused track.

7.4.2 Data Acquired

The $\sim 2 \times 2.2$ km Emmen scenes acquired in this study (Figure 7.11) have similar terrain characteristics as those of the Küttigkofen site: about 110 m height variation, with some forest and housing. However, because the 0.57 m interferometric baseline was used, a higher phase coherence was expected, but at the cost of a lower height resolution. As in Küttigkofen, the master dataset was common to the InSAR and stereo pairs; an additional slave from a parallel track made stereo possible for the overlap region.

Figure 7.12 shows the SAR amplitudes for the chosen stereo pair. Because the scenes were processed in such a way as to exhibit some azimuth misalignment, the stereo coverage was somewhat reduced as compared to the potential azimuth-aligned coverage.

7.4.3 Height References

7.4.3.1 The DHM25 from SwissTopo

The DHM25 is a DTM (bare ground elevation, without vegetation and buildings) with 25 m pixel spacing produced by the Swiss Federal Office of Topography, or SwissTopo. It covers all of Switzerland, being based on the complete set of national 1:25000 maps. An extract of the DHM25 corresponding to the 5×5 km area depicted Figure 7.11 is shown in Figure 7.13(a). The dataset perimeters are shown as well in red and blue, with the same color-coding as in Figure 7.11 to serve as a scale reference.

The height accuracy, as determined by ground control points, is similar to that of the national map set: $\sim 2\text{-}3$ m for the Swiss midlands and foothills, 5–10 m for the pre-Alps, and about 15–20 m for most of the Alps. A detailed description of the production and characteristics of the DHM25 are given in [84].

7.4.3.2 The DSM25 from sarmap

Using tandem pairs from the European Remote-Sensing Satellites ERS-1 and ERS-2, sarmap, a radar remote sensing company based in southern Switzerland, produced a DSM covering all of Switzerland. Based on data produced by the 5.3 GHz C-band SAR sensors carried by the ERS platforms, sarmap used InSAR to estimate the height model. It was gridded to a 25 m raster, which is the expected horizontal accuracy according to sarmap. Details on the techniques used are given in [71].

While the surfaces of solid structures are represented in any case, C-band

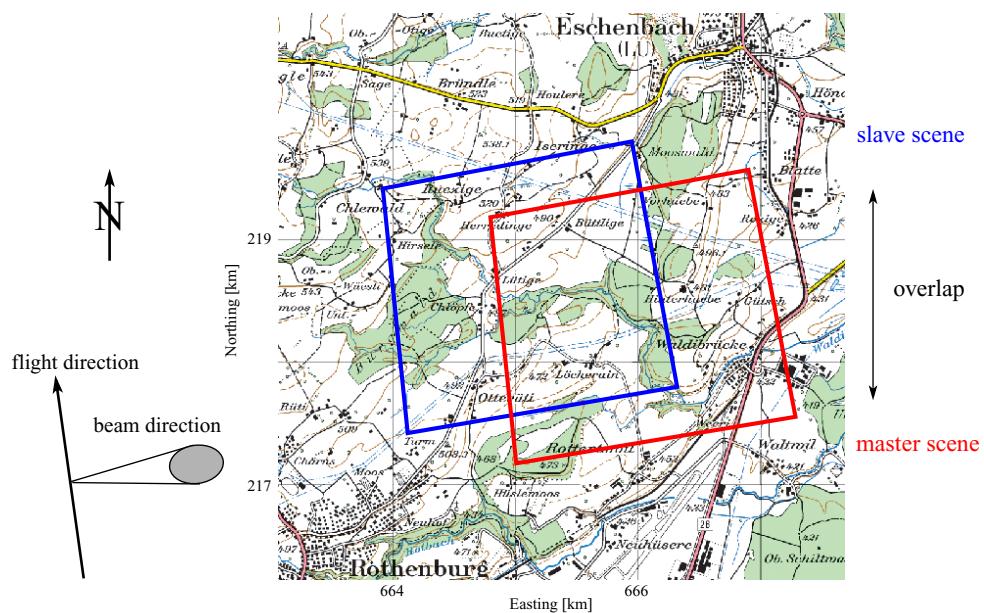


Figure 7.11 The Emmen test site, with master and stereo slave scene perimeters [Background © SwissTopo]

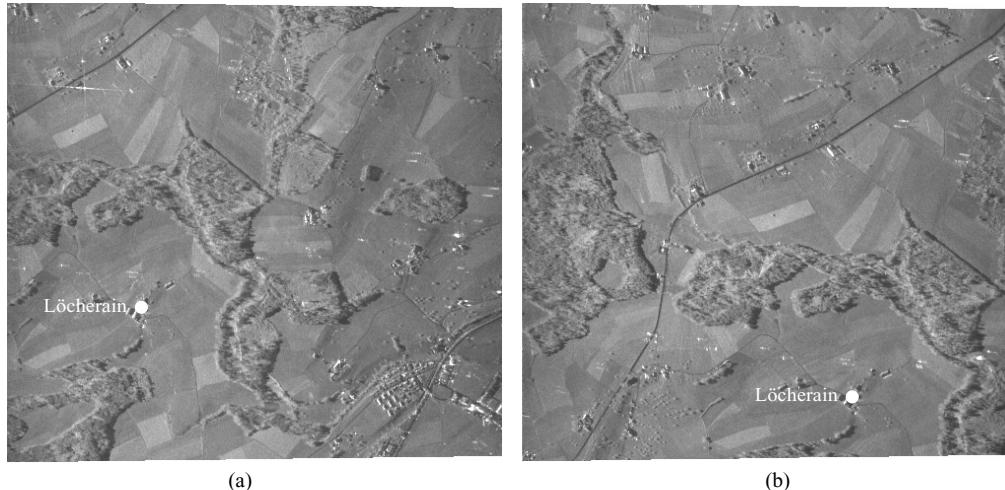


Figure 7.12 Emmen SAR slant-range amplitude images: (a) stereo master, also used for InSAR (b) stereo slave. Azimuth/range looks: 2/1. Pixel spacing: range: 37 cm, azimuth: 56 cm [© Intermap Technologies GmbH]

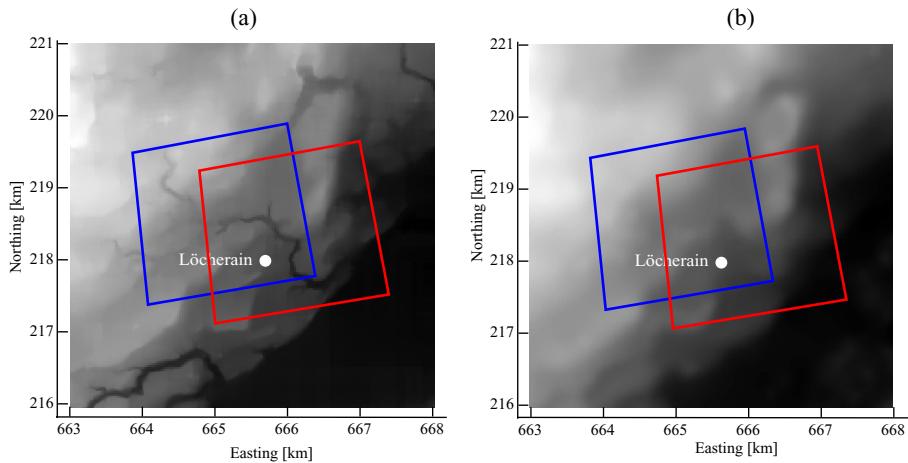


Figure 7.13 (a) DHM25 height model reference for Emmen. Pixel spacing: 25 m [© SwissTopo] (b) DSM25 for Emmen. Pixel spacing: 25 m [© sarmap] Black=415 m, White=560 m Blue perimeter = **slave**, red perimeter = **master**.

radiation, with a wavelength of nearly 6 cm, partially penetrates vegetation before being reflected. Therefore it is expected that the sarmap model will slightly underestimate forest stand heights and vegetation canopies. Although the off-nadir angle of about 23° decreases the average penetration depth with respect to nadir-aligned viewing, the angle is not steep enough to guarantee that the upper canopy will be the main reflecting structure over forested areas.

Microwave penetration into vegetation canopies depends on a number of factors, such as the microwave frequency, canopy density and water content, and the size distribution of the scattering elements. Ulaby measured typical penetration depths for mature crops of several meters at 1 GHz, down to 1 m or less at 10 GHz or higher [89]. One can expect penetration into forest canopies to be somewhat higher, since the density of scatterers such as leaves and branches is typically lower than for agricultural crops. These facts, combined with the incidence-angle effect, make it uncertain to what degree C-band or X-band DSMs will underestimate canopy heights; the only certainty is that X-band will provide a more accurate estimate of the true canopy surface.

Sarmap lists the achievable height accuracy of its DSM25 as ranging from about 5-8 m in flat-to-rolling terrain, to 10-15 m in areas of steep topography. However, in extreme cases, the accuracy drops to ± 50 m. The DSM25 is shown in Figure 7.13(b) for Emmen. The area covered is roughly the same as in Figure 7.11. The DSM25, based on InSAR measurements, shall be

considered the less accurate of the two in general; its accuracy is especially variable in the Swiss Alps. Nonetheless, it serves as a useful comparison in forested areas outside of the mountains, where it should be expected to more closely represent the canopy heights.

However, the canopy heights are in fact not clearly represented in the DSM25, as Figure 7.14 demonstrates. Referring back to Figure 7.11 on page 96 for the forest locations, the heights of the tree-covered areas do not stand out as compared to the DHM25. Within the central forest stand, the narrow deforested passage centered on the meandering stream is severely over-estimated in the DSM25, but this is an error in this case. With an overall standard deviation of 3.9 m and errors of up to 10 m, the DSM25 is considered a less reliable reference than the DHM25 for airborne SAR. However, because it is nonetheless an independent height reference, and because the reference information for this area is limited, it will be used along with the other sources.

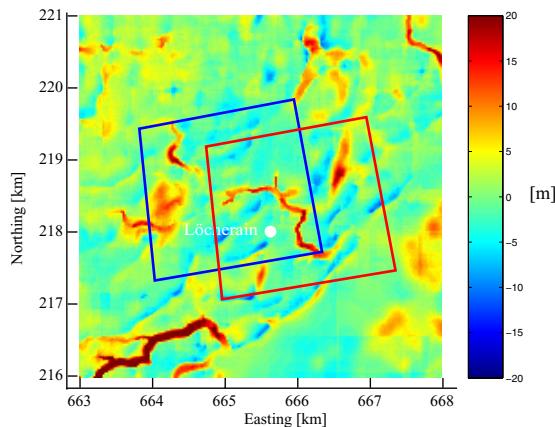


Figure 7.14 Difference between DSM25 and DHM25 for Emmen. Blue perimeter = **slave**, red perimeter = **master**.

7.4.3.3 Ground Control Points

Since no high-resolution DSM reference is available, GCPs in the form of spot heights were extracted manually from the 1:25000 digital topographical maps from SwissTopo, whose positions were given with a vertical accuracy of about 10 cm. The spot height locations are shown in Figure 7.15.

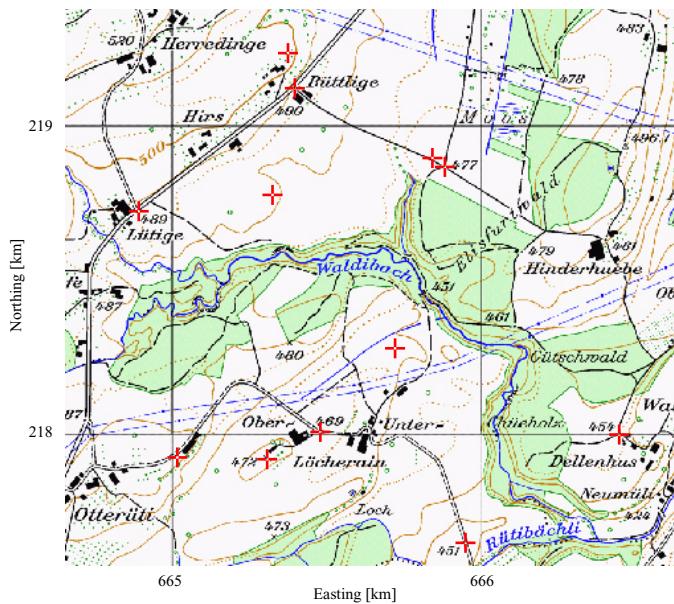


Figure 7.15 Spot height locations for Emmen [Background © SwissTopo]

7.5 Spaceborne Case: Zurich

7.5.1 Experiment Configuration

The coverage of ESA's ENVISAT and ERS-1/2 stereo and InSAR pairs was seen in the overview (Figure 7.1 on page 86). Both satellites carried C-band sensors at an altitude of about 790 km. Because a limited number of scenes were available, only one combination existed that permitted stereo and InSAR with some overlap between the two regions. For ENVISAT amplitude pair to be matched, the processed pixel spacing was about 16 m in range and 31 m in azimuth after multi-looking. More information on the ENVISAT platform and its SAR antenna can be found in [23] for example.

The flight geometry for ENVISAT stereo and ERS-1/2 tandem InSAR is illustrated in Figure 7.16. While AeroSensing arranged their sensors with fixed incidence angles of 45° and flew closely-spaced tracks for stereo coverage, ENVISAT is capable of electronically steering its SAR antenna; this is called *ScanSAR*. It permits stereo SAR without tightly-spaced orbits, since data from different incidence angles, or *beams*, can be used. The stereo pair used here is from two descending orbits east of the city of Zurich, first on

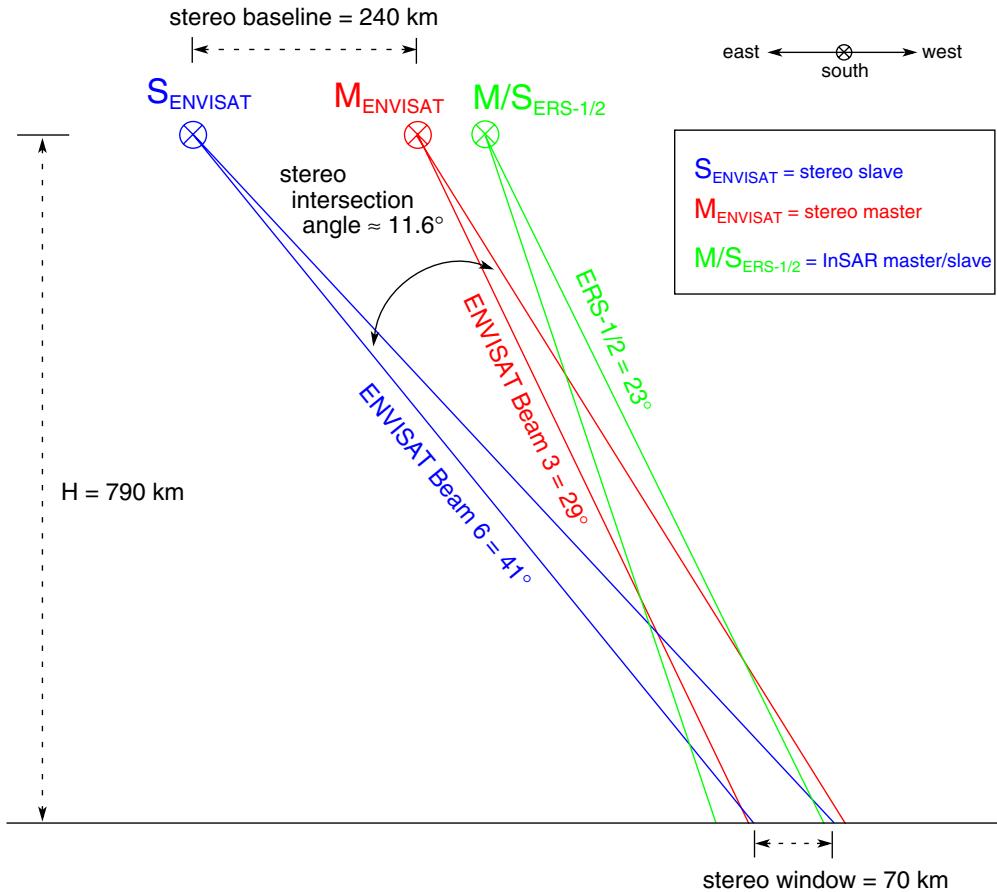


Figure 7.16 Zurich flight geometry.

September 8, 2003, using the IS6 beam ($\sim 41^\circ$ incidence angle), then on September 14 using the IS3 beam ($\sim 29^\circ$ incidence angle). The ground area covered in each case is nearly the same, and the stereo incidence angle near mid-scene is about 11.6° .

The ERS-1/2 interferometric pair was acquired on October 19-20, 1995 in tandem mode from 790 km altitude, using a fixed incidence angle of about 23° . The baseline component normal to the incidence angle was $\sim 140 \text{ m}$. Because of its short length compared to the stereo baseline, it is not explicitly indicated in Figure 7.16. Details on the ERS-1/2 platform can be found on ESA's web site or publications such as [25].

It is interesting to note that with ENVISAT's center frequency of 5.33 GHz and that of ERS at 5.30 GHz, performing interferometry between the two sensors is a complex matter. Although the theory was presented in 1994 (see [32]), practical application of this technique is still in its infancy.

7.5.2 Data Acquired

While given the name *Zurich*, the spaceborne acquisitions shown in Figure 7.1 on page 86 in fact comprise a good portion of Switzerland, including the cities of Zurich and Lucerne. The topography over this area spans a wide range, from moderate in the north to mountainous in the south, totalling about 2600 m of height difference between the low and high points (ERS-1/2 scenes). In fact, it is not height disparity itself that creates difficulties for SAR image interpretation and processing, but rather the quick height variation typical in the mountains. Indeed, this sort of terrain generates a great deal of terrain distortion in the form of layover in radar images, creating especially difficult stereo-matching and phase-unwrapping situations. The layover is clearly visible as bright areas towards the lower-right of the amplitude images in Figure 7.18 (ENVISAT) and Figure 7.19 (ERS-1).

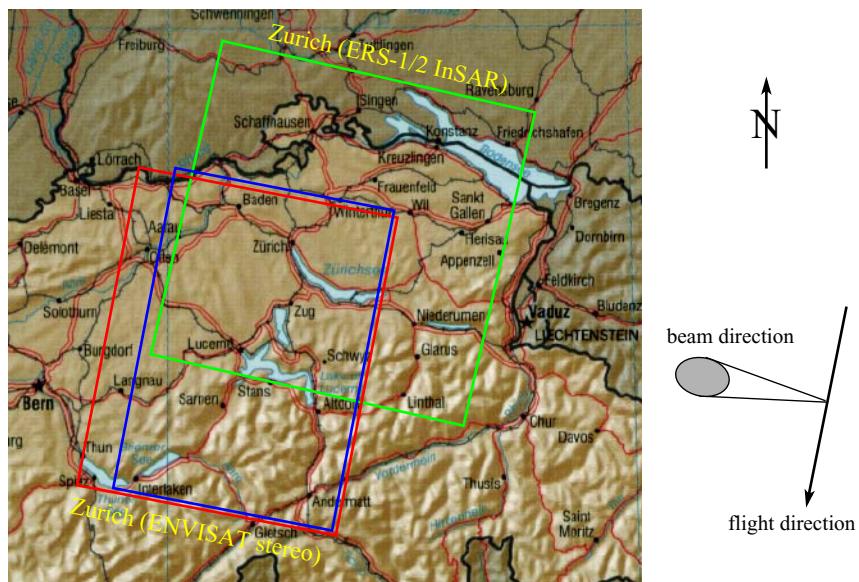


Figure 7.17 The Zurich stereo-InSAR test site. The ENVISAT stereo master (red) and slave (blue) overlap partly with the ERS-1/2 interferometric perimeter (green). [© The General Libraries, University of Texas at Austin]

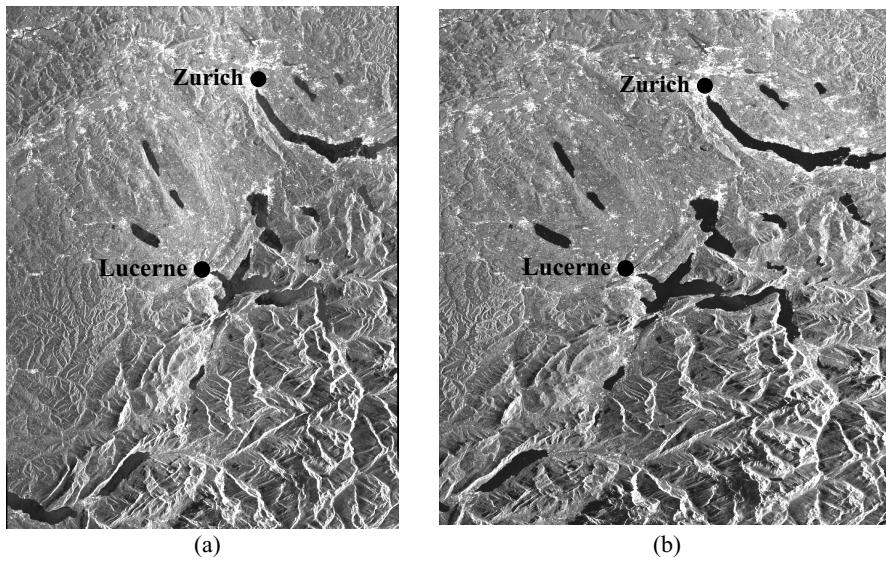


Figure 7.18 Zurich ENVISAT stereo-pair slant-range amplitudes (a) master (b) slave. Azimuth/range looks: 10/2 (master), 8/2 (slave). Pixel spacing: range: 16 cm, azimuth: 31 cm [© ESA]

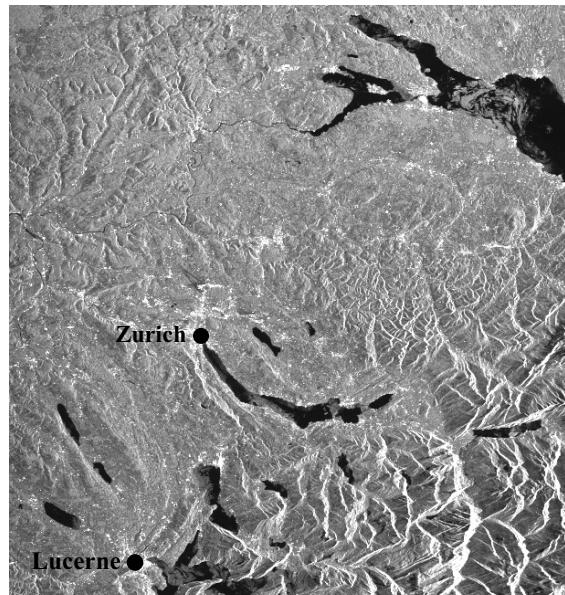


Figure 7.19 Zurich ERS-1 SAR slant-range amplitude. Azimuth/range looks: 5/1. Pixel spacing: range: 8 cm, azimuth: 20 cm [© ESA]

7.5.3 Height References

7.5.3.1 DHM25

The DHM25 terrain model covers all of Switzerland and is well-established as a high-quality height reference for spaceborne imagery. Furthermore, it is based on the 1:25000 topographical maps; therefore, map-collected GCPs were not strictly necessary for height model validation.

An extract of the DHM25 for the stereo-InSAR overlapping region is shown in Figure 7.20.

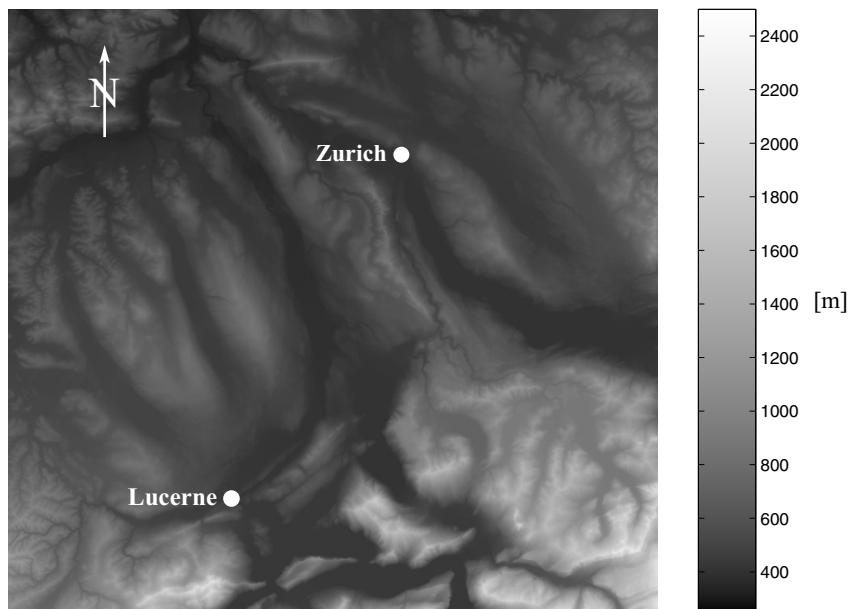


Figure 7.20 The DHM25 for the approximate Zurich region of stereo-InSAR overlap. Pixel spacing: 25 m [© SwissTopo]

7.5.3.2 Ground Control Points

While GCPs collected manually from maps are not strictly necessary for DSM validation, they were required in the context of the stereo-assisted InSAR experiments when comparing manual to automatic phase calibration using GCPs.

Ten points such as bridges and highway intersections were selected for their visibility in the ERS-1 image and on topographical maps. Precise image tiepoints were not simple to collect, with the speckle and altitude of the

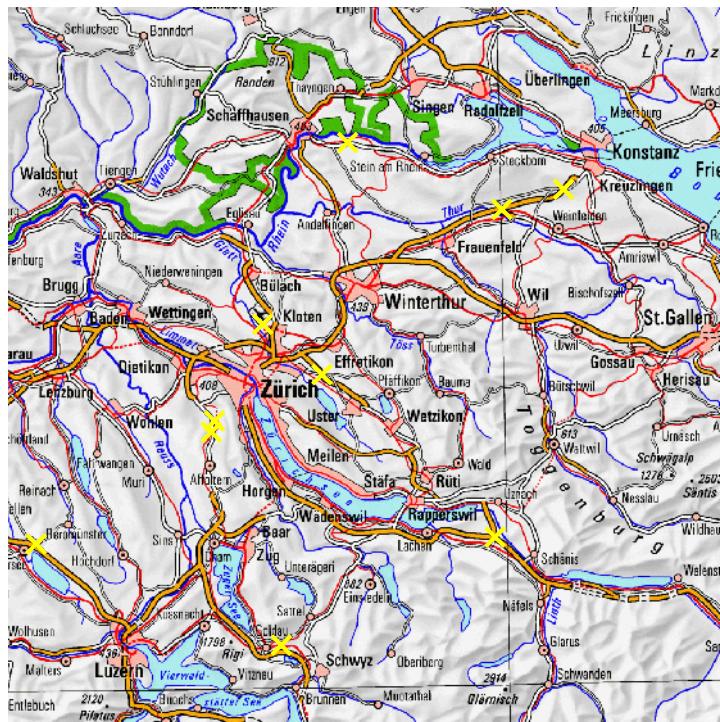


Figure 7.21 GCP locations for the Zurich ERS-1/2 interferometric pair [© SwissTopo]

spaceborne sensor making identification of distinct features difficult. Their positions are shown in Figure 7.21 as yellow crosses.

Chapter 8

Stereogrammetry Results

8.1 Introduction

The primary goal of this work is to show how, and under what conditions, DSMs estimated using InSAR can be improved by parallel processing of a SAR stereo pair. The output of the stereo processor, generated before any interferometry is done, is the focus of this chapter. The results obtained using the stereo processor outlined in section 2.4 on page 48 will be described for the three test sites described in chapter 7.

There are two outputs of the stereo processor that are used as InSAR inputs: the estimated heights and the match confidence (section 2.4.3 on page 50, step 3). The latter is of major importance to the proposed method, because it provides an indication of the reliability of the height estimates. This, in turn, serves as one of the criteria in determining which points to retain for the phase calibration (section 5.3 on page 73) and hole-filling (section 5.4 on page 76). A wavelet-based matching strategy was chosen here, but in practice it should be possible to perform the stereo matching with an algorithm of choice, provided that some measure of the height quality is available.

Some matching results are presented and discussed in this chapter, followed by the derived DSMs for the three test sites.

The robustness of the matching algorithm for various stereo configurations is tested using SAR images simulated for various airborne flight geometries.

Finally, DSM height errors are compared against the match confidence, and the utility of the stereo processing algorithm within an InSAR context is discussed.

8.2 Airborne Stereo: Küttigkofen and Emmen

8.2.1 Stereo Matching Results

Stereo processing begins in each case by estimating a disparity field for the

master and slave amplitude images, both in slant range, as described in section 2.4.3 on page 50. Based on the disparities and the sensor positions (state vectors), geocoding proceeds according to section 3.4 on page 57, generating the corresponding height field within the reference frame describing the sensor positions. Naturally, inaccuracies in the state vectors will directly affect the results as well, but small inaccuracies are of little concern for stereo processing, where the baseline is much greater than the magnitude of positional errors. It will be shown in this chapter that the major source of height error is disparity-field error.

8.2.1.1 Evaluating Match Accuracy

It is possible to evaluate the accuracy of the matching algorithm directly by resampling the slave image into a new geometry, as specified by the estimated disparity field. A perfect matching will map the slave exactly to the master, as far as this is possible, given that the image radiometries and noise patterns are not identical. There should be no discernable planimetric shifts between homologous points.

Verification of the point-to-point accuracy can be done in practice by flipping back and forth between the master and remapped-slave images; any features that appear to change position from one image to the other were not properly correlated. Further, the positions of these features can be measured in each image, making error calculations possible for those features.

Alternatively, the two images can be combined as a two-color image. One image is displayed as red, the other as green, and they are overlaid on top of one another. If the disparity field exactly maps the slave to the master, the resulting combined image should be uniformly yellow, its intensity modulated by the average combined SAR intensities. A feature that has not been properly matched will appear in double as a red and a green feature. Thus, any red or green areas indicate mismatching.

8.2.1.2 Matching Results for Küttigkofen

Radar shadows are the most difficult examples of the fundamental difference between SAR and optical stereo pairs, first mentioned in section 3.1 on page 53. Indeed, while correctly mapping a "slave" shadow to that of the master will result in the correct topographical height in optical stereo (this is done automatically by the human eye-brain system), doing the same for a SAR pair will result in a false height being assigned to the shadow.

Forest stands and other objects with vertical sides present this sort of extreme topographical difficulty for airborne SAR stereo. A forest stand is a sudden, large terrain discontinuity, topographically speaking. As can be seen

in Figure 7.5 on page 90 (Küttigkofen), the forests produce narrow zones of extreme layover on the sides facing the sensor (in this case, the sensor was positioned south of the scene) and significant shadows behind them.

As might be expected, the shadows perturb the accuracy of the disparity field near them. An example of the color-overlay technique is shown in Figure 8.1, which shows a red-green overlay of the master and remapped slave for the Küttigkofen pair. At first glance, the image is yellow-tinged everywhere. Looking more closely near the shadows caused by trees, one can observe that the road appears twice, once in red and once in green, indicating incorrect matching here. It is important to note that color separation is only visible where strong features or discontinuities exist; a featureless background will still look yellow even if the red and green channels are shifted with respect to each other. Therefore, in areas where many features tend to display color separation, it is likely that the problem extends beyond the features themselves.

These localized color shifts can be directly measured by magnifying small portions of the image in places where compact, recognizable features are seen separated into the red and green channels. Knowing the range pixel spacing, the range disparity error can be converted to meters. The corresponding expected height error can be estimated using Equation 8.1 [50][33]:

$$\Delta h \approx \frac{\Delta\delta_r}{(\cot\theta_i^M - \cot\theta_i^S)} \quad (8.1)$$

with

- | | |
|------------------|--|
| Δh | expected height error |
| $\Delta\delta_r$ | measured range disparity error |
| $\theta_i^{M,S}$ | incidence angle for the master (or slave); see Figure 1.1 on page 27 |

This describes how sensitive an estimated stereo height is to an error in the disparity field for a given image location. Under the assumption that the state vector reliability does not significantly vary over the scene, the measured dependence of height error on disparity error can be assumed to hold approximately for the entire scene, as the disparity and state vector measurements are the only possible error sources.

The mismatched road in Figure 8.1 were measured as having a vertical (range) error 13 pixels near the radar shadow. At a pixel spacing of 0.3747 m, this is 4.87 m. The feature location is at range pixel 1243 for the master, 3206 for the slave. Converting to meters and adding the near range constant of

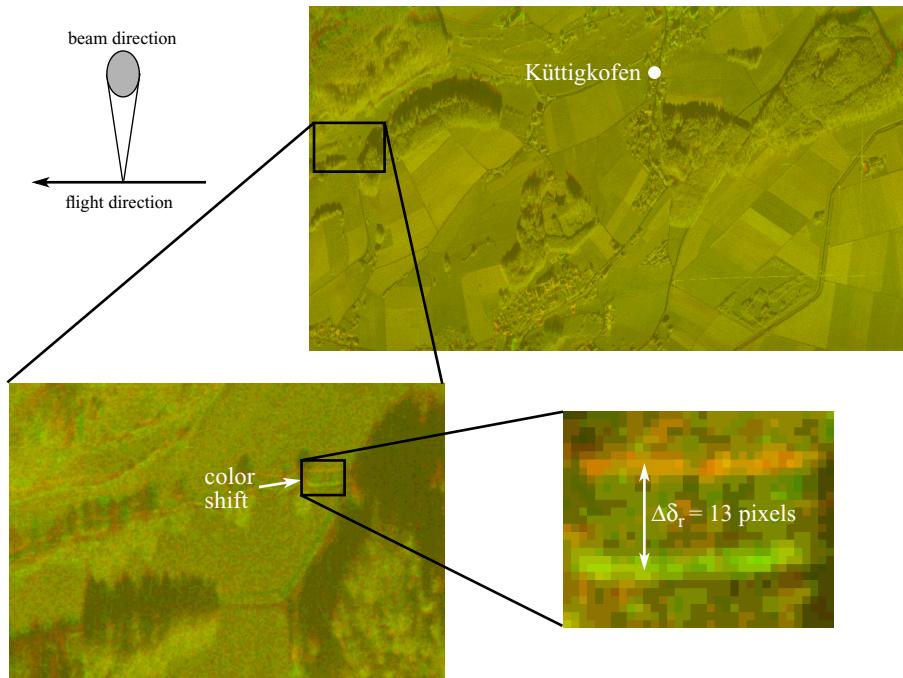


Figure 8.1 Red-green overlay between stereo master and resampled slave for Küttigkofen

3307 m to each, the range to the feature is determined to be 3772 m for the master, 4507 m for the slave. Thus, the incidence angle (for a flat-Earth approximation) for the flight height of 2.8 km is 42.1° for the master, 51.6° for the slave. Solving Equation 8.1, one arrives at the approximate height estimate error of 15.5 m; this nearly perfectly accounts for the measured height error at this location, as will be seen in section 8.2.2.1.

8.2.1.3 Matching Results for Emmen

The relatively small stereo intersection angle of 6.9° obtained over Emmen is expected to result in a stereo pair with a higher resemblance between the master and slave. Indeed, no shifts in any roads are visible in the overlay image shown in Figure 8.2, even near radar shadow. Some small red or green features are visible, but no obvious double images in both red and green. Occasional isolated red or green features are expected for scatterers that are significantly brighter in one image than the other, usually man-made objects with strongly direction-dependent reflectance characteristics. The best example of such an effect is shown in the zoom of Figure 8.2, where a strong scatterer appears as

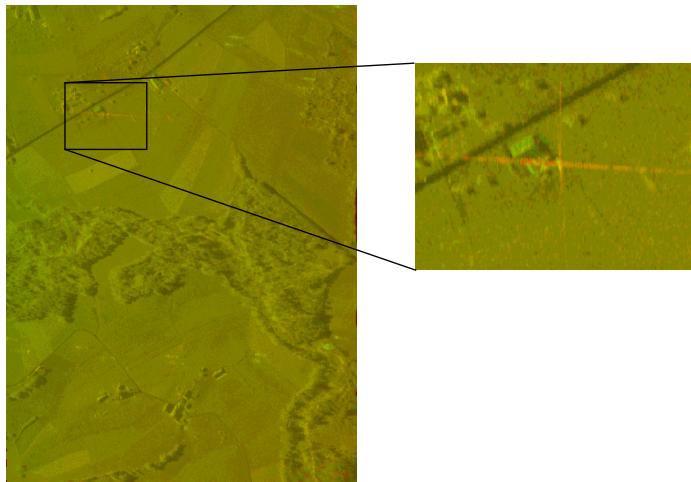


Figure 8.2 Red-green overlay between stereo master and resampled slave for Emmen, with an example of a feature not visible in the slave (the red cross)

red. This object is clearly visible at the upper left of Figure 7.12(a) on page 96, but not in (b). Such examples only highlight the differences between radar images taken from different viewpoints, but do not constitute match errors, since the feature being matched is effectively present in only one image.

8.2.2 DSM Results and Discussion

After the matching step, the disparity field is converted to a geocoded height field, given the absolute sensor positions at all times. This is done by solving the range-Doppler system described in section 3.4 on page 57. The results for the three test sites are shown in the following pages.

8.2.2.1 Küttigkofen

Figure 8.3 shows the DSM obtained for Küttigkofen. Subtracting the DoSAR and laser DSMs from it, the results in Figure 8.4 (a) and (b) are obtained, respectively. The major difference between the two is visible over the forest canopies; the laser beam, with its extremely steep incidence angle, occasionally reaches the ground through the trees. The 42°-angle SAR antennas rarely penetrate the upper canopy.

The regions of strong error within and near radar shadow, and to a lesser extent within layover zones, are immediately obvious. In particular, the shadows and their surroundings are assigned heights that have been overestimated by 20-30 m. Not coincidentally, the forest stand heights

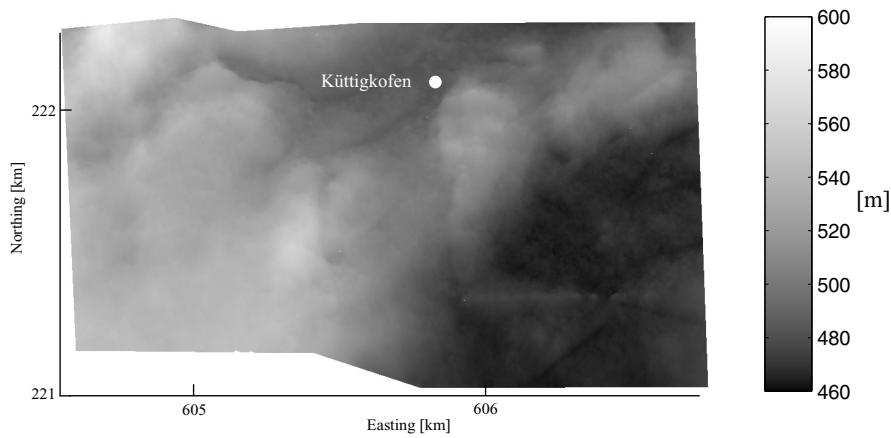


Figure 8.3 Estimated stereo DSM for Küttigkofen

towards the northern edges lie precisely in this range. Indeed, this is a consequence of matching two images in different slant-range geometries; **the height assigned to a radar shadow corresponds approximately to the height of the object responsible for the shadow** (the northern edge of the forest stands in this case).

On the opposite side of radar shadow, there seems to be a slight tendency to underestimate the heights for the zones along the front sides of raised objects, such as the southern rims of the forest stands. As with the shadows, the layover zones affect the immediate surroundings. This is especially visible in the north-western (upper-left) corner of Figure 8.4, where the side of the forest facing the sensor is severely underestimated, as are the surroundings. The layover here is stronger than in front of the other forest stands because the north-western trees are on a slope facing the sensor; this increases the extent of the layover in front of and behind the forest front edge.

The statistical results corresponding to Figure 8.4(a) are given in Table 8.1 for the height references described in section 7.3.3 on page 89. None of the GCPs lie within shadow or layover zones, explaining the improved statistics as compared to those for the two DSMs, with all points under 10 m absolute error.

It is clear that radar shadow and layover pose a problem for stereo SAR. Without further data acquired from different directions, even the best possible matching algorithm would contain strong errors where shadow and layover occur.

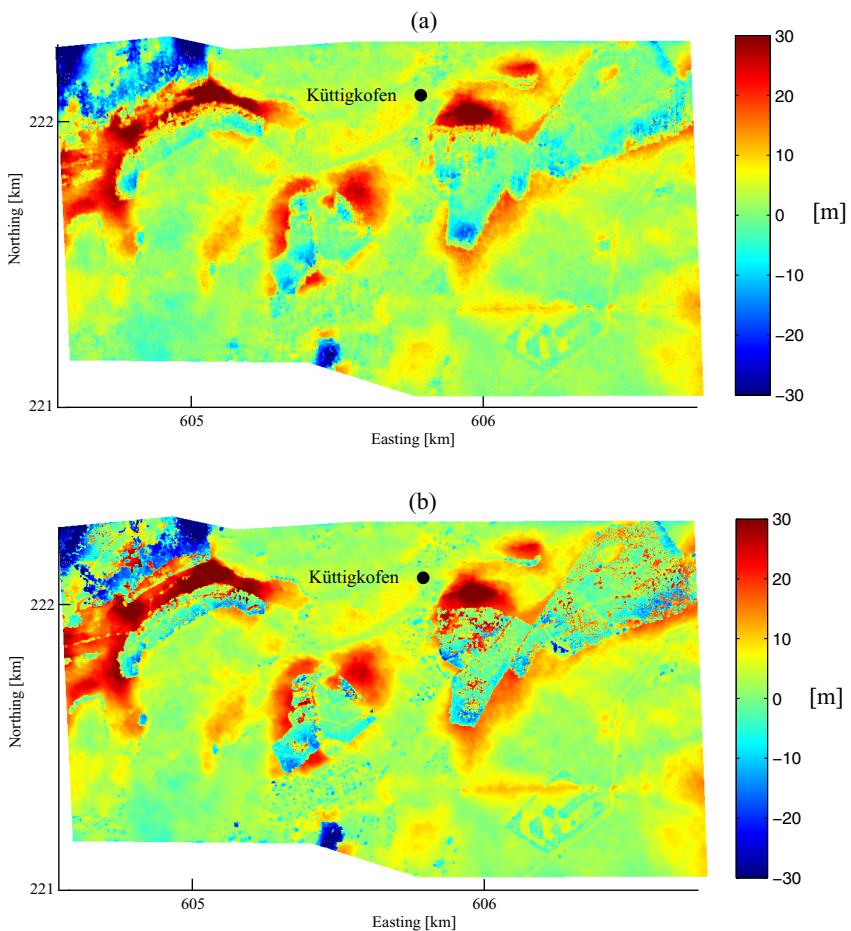


Figure 8.4 Difference between Küttigkofen stereo DSM and (a) DoSAR DSM (b) laser DSM

Table 8.1 Küttigkofen stereo-DSM error statistics

Reference	Mean Height Difference [m]	Standard Deviation [m]	Pixels with < 5 m difference [%]	Pixels with < 10 m difference [%]	Pixels with < 20 m difference [%]
DoSAR	3.86	7.59	59.5	84.3	95.5
Laser	4.52	8.24	54.2	80.4	94.2
18 Tiepoints	3.61	1.76	77.8	100.0	100.0

8.2.2.2 Emmen

Although parallel flight tracks were flown in Emmen as in Küttigkofen, the image azimuth limits did not line up between pairs, as they had in Küttigkofen. The result is that the stereo window for any given stereo pair only covers about 80% of the original azimuth length, as opposed to nearly 100% in Küttigkofen. The range overlap was slightly under 50%. In the end, despite Emmen's higher flight altitude, the Emmen DSM covers only 83% of the total area covered in Küttigkofen.

The DSM obtained for the stereo window is shown in Figure 8.5. The general outline of a part of the central forest is visible.

Since no high-resolution surface model reference exists for this site, it was compared to the three available sources: DHM25, DSM25, and the ten spot heights listed from the 1:25000 topographical map (shown in Figure 7.15 on page 99). Statistics for all three references are given in Table 8.2.

That the DSM25 is not really a surface model was seen in Figure 7.14 on page 98: an ideal surface model should show the forested regions clearly when compared to a terrain model (DTM), and this is not the case here. Figures 8.6 (a) and (b) both display the central forest as red, and this should not be the case when comparing one surface model to another. Ignoring the

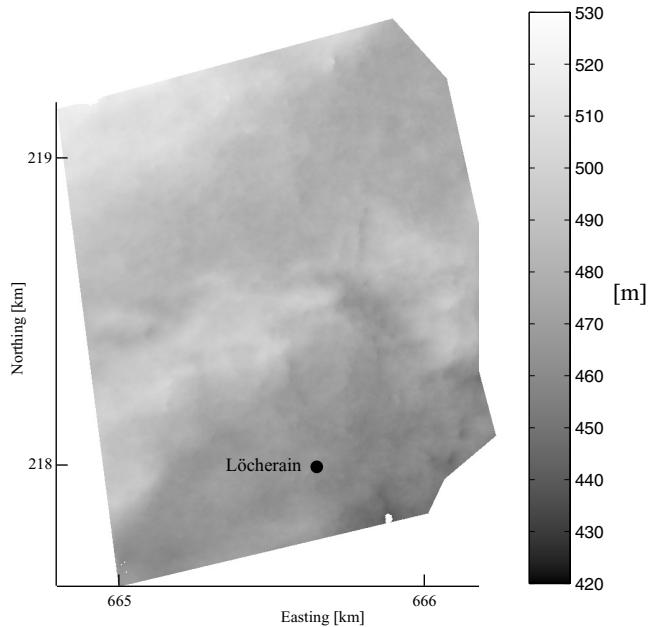


Figure 8.5 Estimated stereo DSM for Emmen

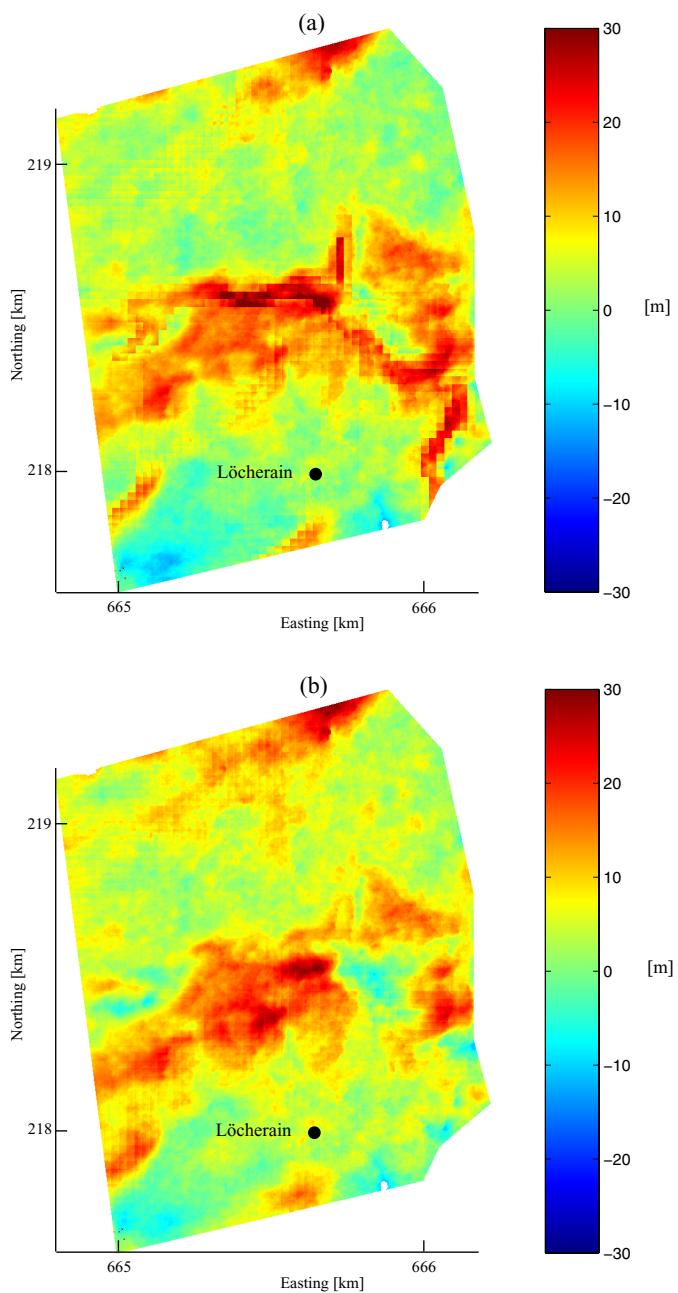


Figure 8.6 Difference between Emmen stereo DSM and (a) DHM25 (b) DSM25

Table 8.2 Emmen stereo-DSM error statistics

Reference	Mean Height Difference [m]	Standard Deviation [m]	Pixels with < 5 m difference [%]	Pixels with < 10 m difference [%]	Pixels with < 20 m difference [%]
DHM25	6.18	6.54	49.8	74.7	96.4
DSM25	6.75	5.57	40.9	77.1	97.6
10 GCPs	1.83	2.88	80.0	100.0	100.0

forest, there seems to be a slight trend of increasing overestimation as one progresses towards the north.

The tiepoints are outside of the forest, and the statistics show that 80% of them have errors under 10 m, similar to Küttigkofen. The standard deviation is higher than for the Küttigkofen tiepoints, while the mean is lower. However, the relatively low number of GCPs and the inadequate reference DEMs in Emmen make direct comparison difficult. Qualitatively comparing the stereo DSMs themselves, it seems safe to conclude that they are quite similar.

8.2.3 Algorithm Robustness

8.2.3.1 Sensitivity to Acquisition Geometry

In order to determine the optimal acquisition geometry and test the stereo matching robustness at the same time, six experiments using simulated airborne data were carried out. Similar experiments could have been carried out for simulated spaceborne pairs. However, if the main goal is to test the robustness of the matching algorithm, it suffices to use a single dataset, as long as the difficulty of the matching problem - which is purely a radiometric one - is also typical of other datasets, whether air- or spaceborne. That is, the matching does not behave differently for images based on spaceborne as compared to airborne data; its performance depends solely on how *different* the two images appear when viewed from the two viewpoints. For this reason we consider the use of an airborne-simulated pair a sufficient test of the matching performance under different viewing conditions.

Each of the six experiments simulated a different stereo flight geometry for the AeS-1 X-band airborne sensor. Amplitude pairs for parallel flight tracks flown at an altitude of 2800 m were simulated as described in chapter 6, for six combinations of incidence angles. The sensor characteristics (central frequency, PRF, etc.), altitude, and test site were all chosen to correspond to the real data acquired in Küttigkofen. Only one aspect was simplified: the

simulations were performed for a perfect side-looking geometry, whereas the real AES-1 amplitude data were processed for a non-zero beam forward-offset (called *squint*) corresponding approximately to the mean for the entire scene. While our stereo processor takes squint into account during geocoding, it was not considered important to simulate a non-zero squint.

Two aspects of the view geometry were of interest to us: incidence angle, and stereo intersection angle (see section 1.5.3 on page 33 for our definition of the latter). The six experiment characteristics and the corresponding results are given in Table 8.3(a)-(f). In cases (a) - (c), the incidence angles are varied, but kept the same for each sensor. The stereo view is created by the intersection of the two beams as depicted in Figure 1.4 on page 34. For cases (d) - (f) one sensor is kept fixed, while the incidence angle of the other is varied. Note that for both cases (e) and (f), the stereo intersection angle is the same. This is because the sensor positions are the same for both; only the incidence angle of the slave sensor is changed.

The slightly different coverages are due to the different viewing geometries. This raised the question of whether to calculate the error statistics for a single region common to all experiments or for all available information in each case. The latter solution was chosen, since it gives the best overall picture of the statistical situation for each case; the cost of this choice is strict quantitative inter-comparability. Nevertheless, the consistent color-coding of the error images make inter-comparison possible to a certain extent.

Case (b) from Table 8.3 is clearly the best, with a scene-wide standard deviation of 4.4 m, the lowest of them all. This suggests that a fixed incidence angle of about 45 degrees is preferred by the matching algorithm.

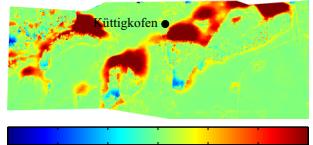
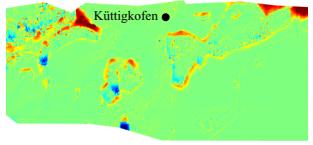
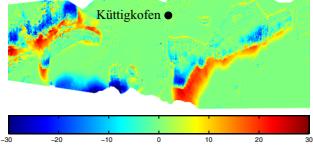
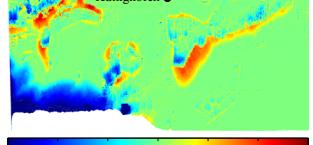
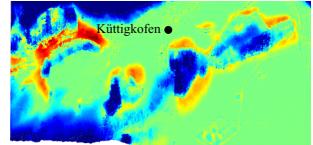
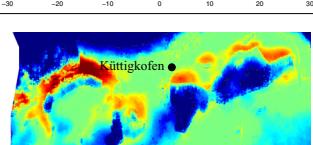
Opposite-side simulations were attempted for the same scene as well, but the radically-different appearance of the forest stands from one sensor to the other made it impossible to obtain the correct disparity field. Opposite-side stereo is no doubt the preferred technique for relatively flat terrain, for which case a large intersection angle is desired. In the context of tests using RADARSAT stereo data, it was recommended that same-side stereo be used for rolling or mountainous terrain [86].

In general, it is clear from the series of experiments that the matching algorithm is very sensitive to small differences in the view geometry, although this is probably in large part due to the forest stands. Such sudden increases in the terrain topography create a difficult matching situation, in spite of the smoothly rolling terrain they are placed in.

8.2.3.2 Simulated versus Real Results

It is interesting to compare Table 8.3(b) with Figure 8.4(a) on page 111.

Table 8.3 Simulation results for various incidence angles

Simulated geometry	Height error [m]	Mean height error [m]	Std Dev [m]	Pixels with < 5 m error [%]	Pixels with < 10 m error [%]	Pixels with < 20 m error [%]
(a) Off-nadir angle 1: 57° Off-nadir angle 2: 57° Intersection angle: 17.0°		5.1	10.3	74.2	83.4	91.3
(b) Off-nadir angle 1: 42° Off-nadir angle 2: 42° Intersection angle: 10.5°		1.1	4.4	90.4	95.6	98.8
(c) Off-nadir angle 1: 27° Off-nadir angle 2: 27° Intersection angle: 8.5°		0.2	6.3	79.2	87.9	98.0
(d) Off-nadir angle 1: 37° Off-nadir angle 2: 27° Intersection angle: 12.9°		-1.3	8.0	78.3	86.2	95.0
(e) Off-nadir angle 1: 47° Off-nadir angle 2: 27° Intersection angle: 25.7°		-4.7	12.4	62.2	73.5	86.6
(f) Off-nadir angle 1: 57° Off-nadir angle 2: 27° Intersection angle: 25.7°		-8.8	18.5	46.8	63.9	79.4

The former is based on a simulation of the latter, to a very good approximation. The only differences in the simulated versus the real geometry are the zero-squint assumption, perfectly parallel flight tracks instead of the nearly-parallel ones flown in reality, and an incidence angle of 42° instead of the true $\sim 45^\circ$. In fact, these differences are almost negligible, as a simulation using the true squint and incidence angle showed: the error map was almost indistinguishable from that of Table 8.3(b). This raises the question of why the simulated result does not contain the same strong errors surrounding the radar shadows and layover zones as are visible in the real result. Unfortunately, the definitive answer to this question had not been found at the end of this work's writing. However, several possible reasons are immediately clear:

- 1) The most striking difference between the real and simulated images (see Figure 6.1 on page 81) is the assumption of a constant backscatter function over the entire terrain during the amplitude simulations. Features of the real terrain such as agricultural plots have a strong angle-dependent effect on the brightness and texture of the real images; none of this occurs for the simulation. Therefore, one could hypothesize that images with simpler textural characteristics represent a simpler case for the wavelet-based matching.
- 2) When validating stereo DSMs based on simulated images, the DEM first used to generate the simulations are used again to validate the final results. This is natural: in a perfect stereo processor, the original DSM would be perfectly reconstructed, thus making it possible to evaluate the performance of the processor. By definition, the error analysis takes only the errors resulting from the stereo processor into account. On the other hand, the result based on the real amplitude pair is compared against a height model which is imperfect itself; the errors measured for the stereo DSM therefore include those inherent in the reference DEM, as compared to the true topography at the time of acquisition.
This hypothesis is weakened by a further observation, however: similar errors are measured when compared to the laser DSM, which was produced at a different time, using an unrelated technique. Therefore, the strong errors surrounding radar shadows should be considered real errors stemming from the stereo processor.
- 3) The real images contain speckle, which may present difficulties for the stereo matching (see [40] for an example of a matching algorithm that incorporates knowledge of speckle statistics directly). However, using the matching algorithm presented here, image preprocessing using a Frost speckle filter [28] was discovered to have a negligible effect on DSM estimation. In one experiment, a 5×5 Frost filter was applied to the Küettig-

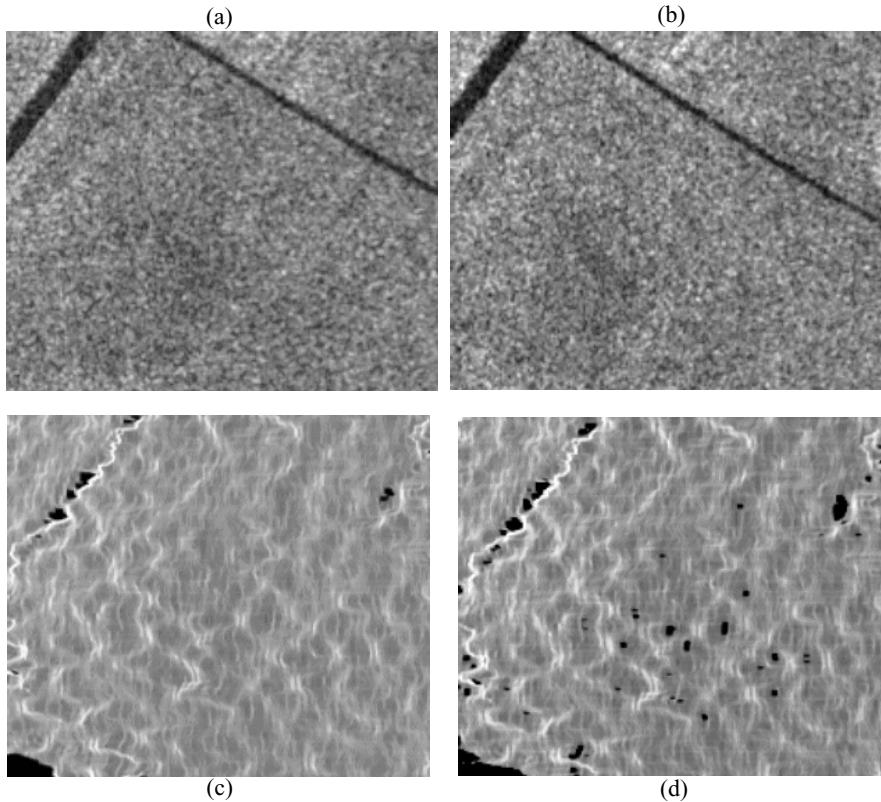


Figure 8.7 Close-up of real and simulated images: (a) real master (b) real remapped-slave (c) simulated master (d) simulated remapped slave

kofen pair from Figure 7.5 on page 90. Stereo matching was performed, and the resulting disparity field compared to the one generated using the unfiltered pair by using the color-overlay technique (see section 8.2.1 on page 105). No red-green offset was visible, even at the edges of roads and fields. It was therefore concluded that speckle filtering was not a necessary step in the stereo processing chain. Indeed, one reason a multiresolution matching framework was chosen was to avoid the speckle problem.

- 4) Confident matches are possible in relatively featureless expanses for simulated pairs, but not for real pairs. A close-up within one of the strongly-overestimated zones from the real and simulated pairs used during matching, shown in Figure 8.7, clarifies the situation somewhat. The images are $\sim 250 \times 300$ -pixel extracts, which reveals the finer structure of each. The simulated pair has wave-like structures stemming from fluctuations in the

DoSAR DSM used as the input to the simulations. These structures exist in both the master and the slave images. In the real pair, no such structures are visible; instead, speckle noise has become dominant at this scale.

The latter hypothesis seems the most likely: in real pairs the high match confidence associated with radar shadows and layover carries over to neighbouring featureless regions, where match confidence is low. For the simulated images, high match confidence is possible in the "featureless" areas because of the existence of wave-like structures even in the absence of real topographic variation.

As the airborne simulation results in Table 8.3 on page 116 already indicated, the wavelet-based matching algorithm seems to be prefer subtle cross-pair differences to large ones. Indeed, even for case (c), with a stereo intersection of merely 8.5° , the ground height was calculated as accurately as for all other cases. This is true in spite of the general lack of topographic features over much of the terrain, especially in the simulated images.

In summary, one may conclude that the stereo matching algorithm:

- is extremely robust for flat or smoothly-changing topography. The topography of the rolling terrain was accurately derivable even in the most extreme geometric cases.
- is negligibly unaffected by the presence of speckle
- produces strong errors when the input images become dominated by shadow and layover.
- may place too much importance on the continuity of the disparity field, resulting in errors in the vicinity of shadow and layover zones.
- prefers the acquisition geometry for the case shown in Table 8.3(b) on page 116, which corresponds approximately to that of the real data acquired with the AeS-1 sensor in Küttigkofen.

8.2.4 Match Confidence for Küttigkofen

The use of the match confidence for GCP selection and hole-filling in the stereo-assisted InSAR context was described in chapter 5. To what extent the match confidence serves as a predictor for the quality of the height estimates subsequently derived is described here for the Küttigkofen scene.

Figure 8.8 shows the confidence and error maps for an airborne case, that of Küttigkofen. Five areas of particularly strong or weak confidence marked A through E are shown in the figure. Areas A and B are examples of height overestimation. In favor of the proposed confidence-based methods, both correspond to low match confidence. Generally, the other zones of high match

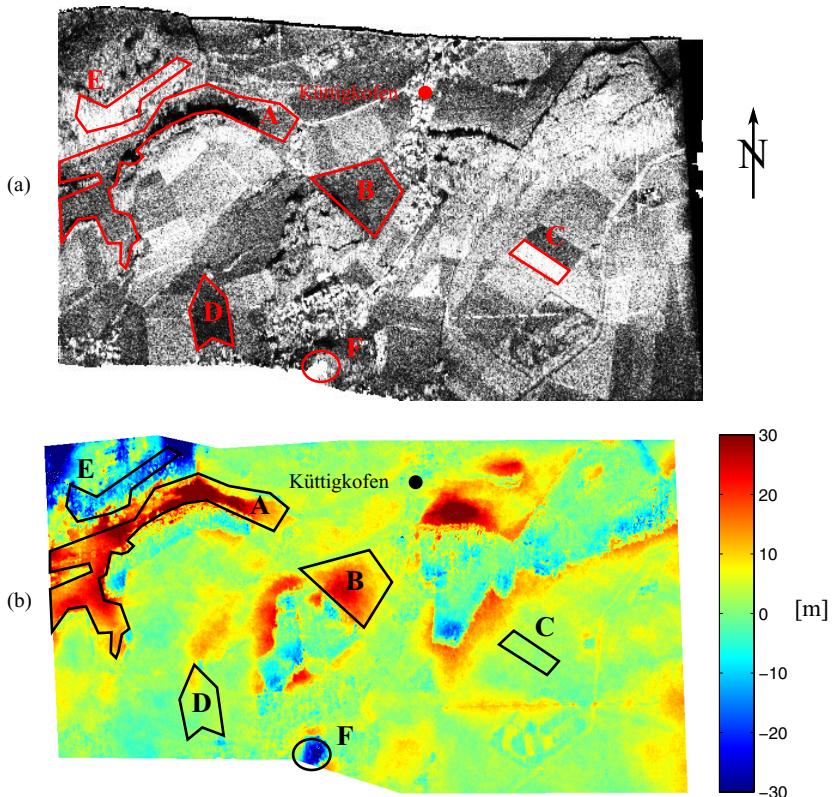


Figure 8.8 Küttigkofen stereo results: (a) match confidence (bright areas indicate high confidence) versus (b) height error (relative to the DoSAR reference). Several areas of especially low or high match confidence are outlined in both images for comparison.

confidence, such as area C, correspond to high height quality.

Area D demonstrates that low confidence does not necessarily imply low height quality.

Areas E and F are examples of exceptions to the general rule that high match confidence implies high height quality. They are the most severe counter-examples in the image. Comparing area E to the original amplitude image in Figure 7.5 on page 90, it can be seen that it corresponds to the bright layover pixels near the upper edge of the forest stand on the slope. As was seen earlier, the heights of layover regions tend to be underestimated. This is in conflict with the nature of the match confidence and matching in general: bright pixels are required for confident matches. Indeed, the lowest match

confidence is consistently seen in the radar shadows, where nearly no signal is present in the amplitude images. However, because layover is bright by definition, confident but false matches can occur; false because the layover is an effect that is unique to the particular viewing geometry. The same uniqueness problem is caused by shadows; however, their near-zero signal return creates low match confidence, thus making their elimination as quality height candidates possible.

The extended layover in the upper-left is the only clear example of an exception to the general rule relating high confidence to high height quality.

The correct of the relationship between match confidence and final height error is twofold: first, **high height error generally implies low match confidence**. The converse is also true: high match confidence generally implies low height error, albeit not without some exceptions. However, **low match confidence does not generally imply high height error**. This is not surprising; low confidence typically results from a lack of strong features, but this does not necessarily correspond to an incorrect match. It is a measure of *confidence*, not accuracy. Thus, low match confidence does not necessarily imply height errors, but the probability of errors is increased. If one would like to preferentially filter out points of high height error from a stereo DSM, one must filter out points with low match confidence. Indeed, the higher one sets a minimum value for confidence, the greater the proportion of good points to bad points - and the lower the total number of points as well.

A scatter plot for the height error versus match confidence is shown for the whole scene in Figure 8.9. It validates the above generalization: the higher the confidence, the fewer large height errors are to be found.

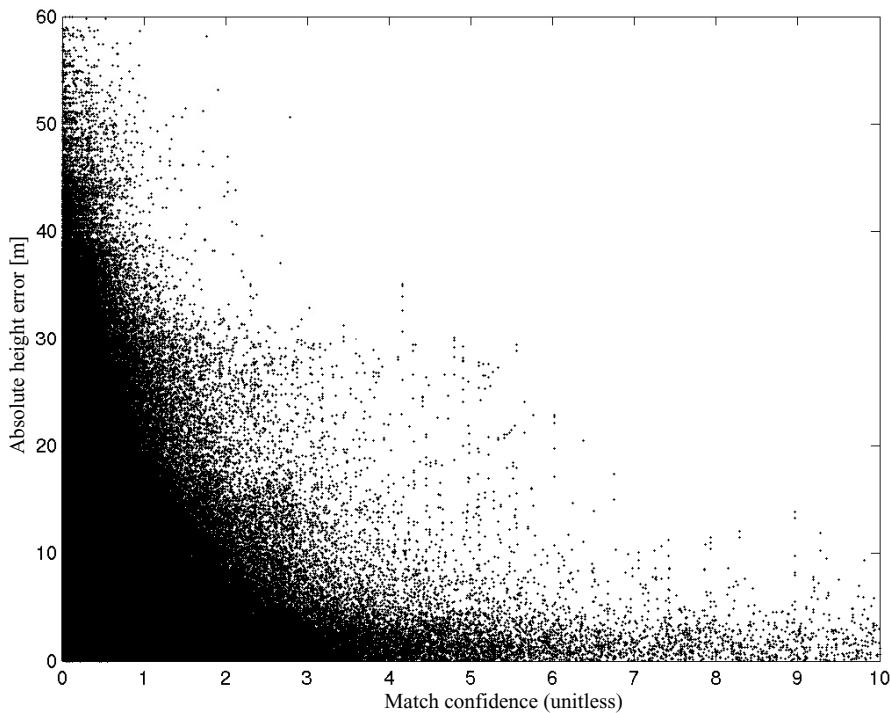


Figure 8.9 Absolute height error versus match confidence for the Küttigkofen stereo DSM.

8.3 Spaceborne Stereo: Zurich

8.3.1 Matching Results for Zurich

A red-green color layover for the stereo master and the resampled (remapped) slave is shown in Figure 8.10. Expanses of slight reddish or greenish tinge are not due to match errors, but rather differences in the intensities of the red and green input images resulting from the different acquisition geometries. The upper zoom, containing the city of Zurich and surroundings, shows no obvious color separation around the lake edges and hill ridges. Within the Swiss Alps, however, areas of color separation are visible, indicating false matches.

For this scene, the range disparities can be roughly measured within the mountains for a number of localized features; mountain ridge and valley features are the simplest to discern. A sample calculation was demonstrated

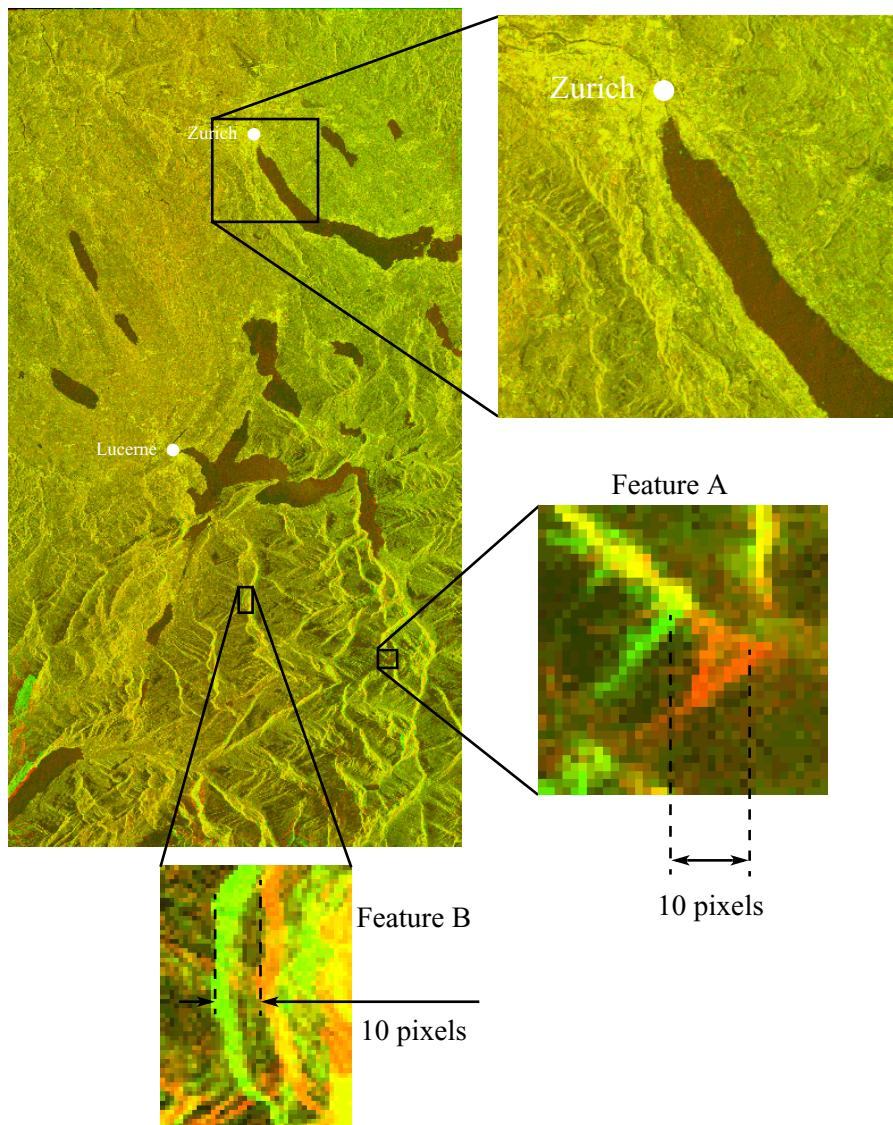


Figure 8.10 Red-green overlay of master and remapped slave with two zooms (Zurich)

for Küttigkofen in section 8.2.1.2 on page 106. While precise measurements are difficult, matching errors in the range direction of approximately 10 pixels were measured, as indicated by the two zoom windows for features A and B. The estimated measurement uncertainty was 2 pixels due to featural

indistinctness; the measurement is therefore written as 10 ± 2 pixels. This is equivalent to 156 ± 31 m in range. Feature A was found to lie at 900270 m from the master, and 1019779 m from the slave, which convert to 28.7° and 39.2° incidence angles, respectively. Solving Equation 8.1 on page 107, a corresponding height error range of $\Delta h \approx 260 \pm 52$ m is calculated.

Performing the same calculation for feature B, the expected height error at that location is 217 ± 43 m.

For both features the true height error was calculated by subtracting the DHM25 values at the same locations. In both cases, the approximate height error was seen to be near the predicted value based on the disparity error.

While two point-location measurements do not quantify an image-wide phenomenon; rather, they are indicative of the direct dependency of height error on disparity error.

Any scene containing mountainous terrain highlights presents a difficult matching problem: rapidly-changing, steep terrain leads to severe distortions and saturation in the slant-range geometry of a SAR image, with vast differences between the two input images.

8.3.2 Error Correction versus Accuracy Retention

The Zurich scene presents an additional complication: it includes not only difficult alpine terrain, but also smoothly-rolling terrain in the northern regions. As it turns out, these two types of terrain are best matched using different smoothing constants for optimal results during matching. This smoothing process was briefly described in section 2.4.2 on page 50 for a given resolution level as a compromise between retaining the estimated disparity field and requiring a smoothly-varying field, which reduces the chances of matching errors propagating from the coarse to the finer levels of resolution. This weighted smoothing process is governed by a single parameter, Ω , which controls the relative influence of the smoothness and accuracy requirements on the resulting disparity field [55]. With $\Omega = 1.0$, the influence of the smoothness and accuracy will be approximately equal. In practice, this value can be "tuned" to the type of terrain imaged: $\Omega > 1$ will favor accuracy, $\Omega < 1$ will favor smoothness.

Preferring too much smoothing for alpine regions causes the valleys and ridges to be "pulled towards" the local average height such as would occur with a regular low-pass filter, resulting in overestimation of the valley heights and underestimation of the ridge heights. On the other hand, not smoothing enough allows match errors to influence the height results, creating a noisier DSM. This is especially undesirable for flat-to-rolling terrain, where smoothing works more to reduce noise than decrease height accuracy.

An experiment was conducted to determine the effect of Ω for three sub-regions of the Zurich scene: one including only rolling terrain, one only mountains, and one with a mixture of the two terrain types. Two values of Ω were used during matching: 0.5 and 10, favoring smoothing and accuracy, respectively. After geocoding (DSM and error results are shown in section 8.3.3 on page 126), the DHM25 reference was subtracted from the two resulting DSMs, resulting in error maps. The percentages of pixels with error magnitudes under 5, 10, 20, 50, 100, 200, and 500 m were calculated for the three terrain types in each case, and the results plotted (Figure 8.11).

The results clearly indicate that the high value of Ω is preferred for alpine terrain, while a low value is preferred for rolling terrain. When there is a mixture of both, the lower value is preferred for heights up to about 50 m, beyond which the high value is preferred. It should be noted that the rising part of the red and black curves are quite compressed, making the curves appear nearer each other than is the case; the curves must be compared in the vertical (percentage) direction. Increasing Ω even more for the alpine region, it was discovered that favoring accuracy in the Alps is only advantageous up to a

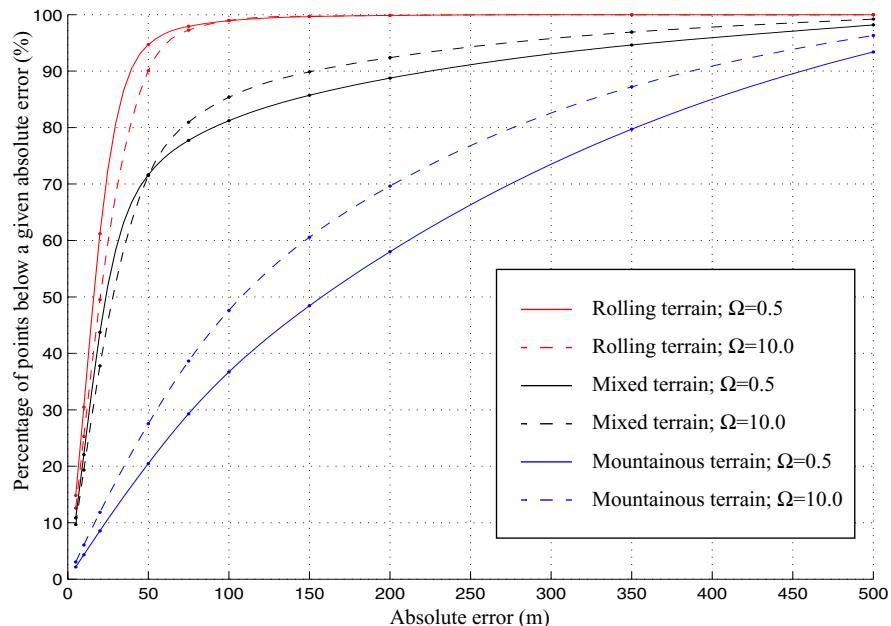


Figure 8.11 Effect of the match-smoothness parameter Ω on the DSM error statistics using the ENVISAT-ASAR stereo pair

point: at $\Omega=100$ (not plotted), which almost neglects smoothing altogether, a slightly *degraded* alpine accuracy was observed as compared to the $\Omega=10$ case. At the same time, even more errors were introduced into the rolling terrain. Consequently, $\Omega=10$ was found to be the optimal compromise for the entire scene, which is best represented by the mixed-terrain case in Figure 8.11.

8.3.3 DSM Result for Zurich

As was first seen in Figure 8.11, the balance between smoothness and accuracy can be fixed before matching begins, but the value representing this compromise has different optimal values for different terrain types. The terrain varies from flat to alpine in the Zurich ENVISAT pair, requiring a compromise between accurate mountains and accurate midlands.

As in the airborne cases, the disparity field was then geocoded along the satellite orbits, which were provided with the products. The 3-D positions resulting from the geocoding are in the same reference geometry as the state vectors themselves, namely, the global cartesian system WGS84. These are

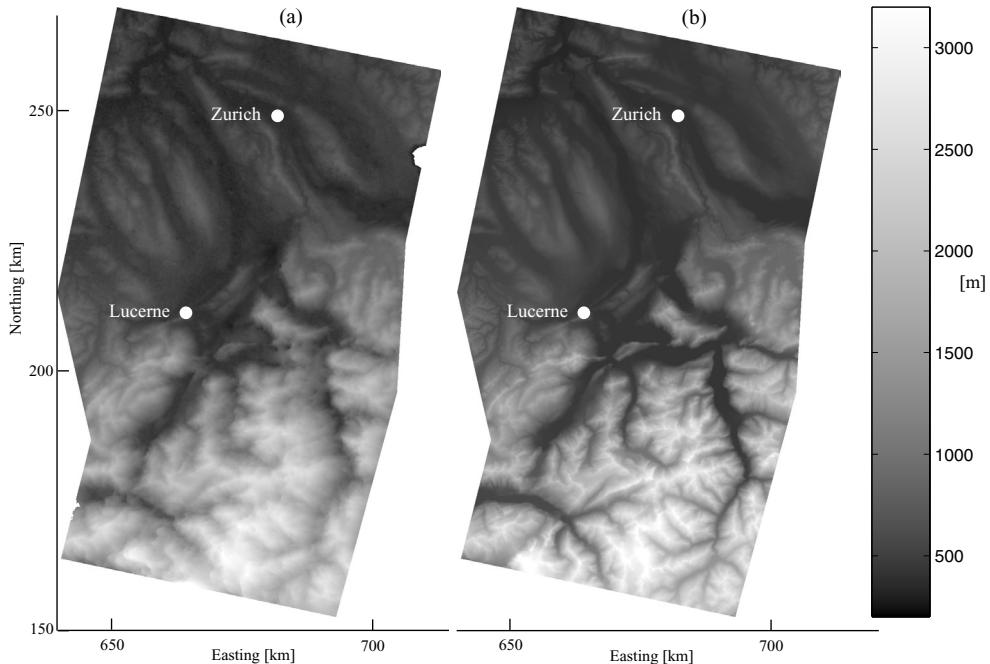


Figure 8.12 (a) Estimated stereo DSM for Zurich (b) DHM25 reference for same perimeter [© SwissTopo]

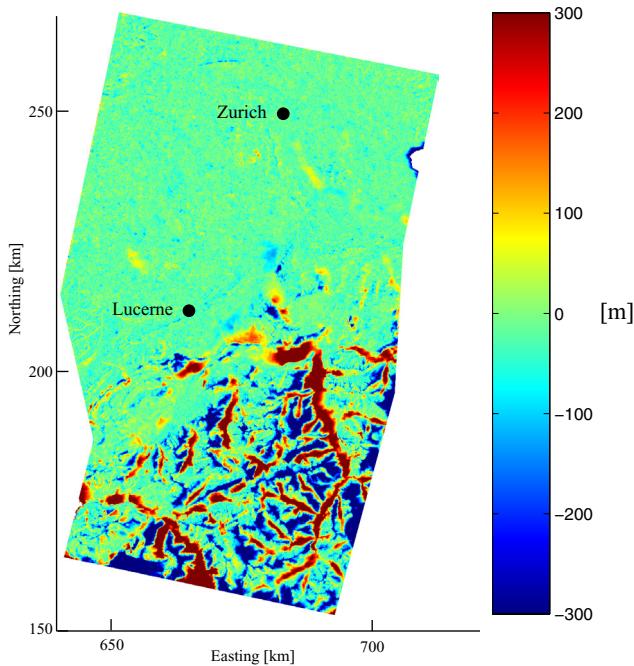


Figure 8.13 Difference between Zurich stereo DSM and DHM25

then transformed into the Swiss map geometry described in section 7.1 on page 85.

That the smoothing effect is harmful in the mountains is immediately clear by comparing the DSM obtained for Zurich (Figure 8.12(a)) with the reference DHM25 (Figure 8.12(b)). The mountains appear to have been low-pass filtered, and indeed this is essentially what occurred. However, even preferring accuracy over smoothness only improves this slightly, but at a cost to the rolling-terrain accuracy.

Although the general shape and contours of the Alps have been reconstructed, the errors relative to DHM25 are often quite high, as seen in Figure 8.13. Comparing the DSMs to the error map, one recognizes that the errors are not random: the valleys are overestimated, the ridges and peaks underestimated. This is precisely what is expected for a rapidly-changing terrain that is low-pass filtered.

Comparing the error statistics for the Alps, midlands, and the whole image in Table 8.4, and referring back to Figure 8.11, one is reminded that the use of a lower value of Ω would have dramatically improved the percentage of points

Table 8.4 Zurich stereo-DSM error statistics

Zone	Mean Height Difference [m]	Standard Deviation [m]	Pixels with < 50 m difference [%]	Pixels with < 100 m difference [%]	Pixels with < 200 m difference [%]	Pixels with < 500 m difference [%]
Whole image	-20.8	148.4	70.0	82.5	89.8	98.1
Rolling terrain (north)	-17.4	21.7	93.4	99.8	100.0	100.0
Alpine terrain (south)	-7.4	196.9	39.6	58.4	75.6	97.2

under, say, 25 m error, having as dramatic an effect on the standard deviation and probably also the mean. The rolling-terrain statistics obtained here represent a "worst case" scenario, but the smaller overall errors magnitudes for the entire scene are what led to the decision to compromise; the mountains have a much greater effect on the statistics for the image as a whole than do the midlands.

Because the stereo processor is required to be fully-automatic, it cannot *a priori* be assumed to have information on the topography of the images being given to it. If this were the case, then it may be possible to process regions of vastly different topographic characteristics (such as flat terrain versus alpine terrain) separately, accordingly assigning different values of Ω to the processing chain. However, because we prefer a fully-automatable solution, an initial choice of Ω is mandatory. Future versions of the matching algorithm may be able to determine the general magnitude of the local terrain variation and adjust the balance between accuracy and smoothness dynamically.

8.3.4 Match Confidence

Figure 8.14 shows the geocoded match confidence and height error for the ENVISAT stereo pair (Figure 8.14(b) is a reproduction of Figure 8.13). Area 'A' contains the highest height errors, and also corresponds to low match confidence. Based purely on this observation, one might be tempted to conclude that low match confidence consistently implies high height error. However, smaller areas of low match confidence can be found in Figure 8.14(a) for which no high height errors can be seen in Figure 8.14(b);

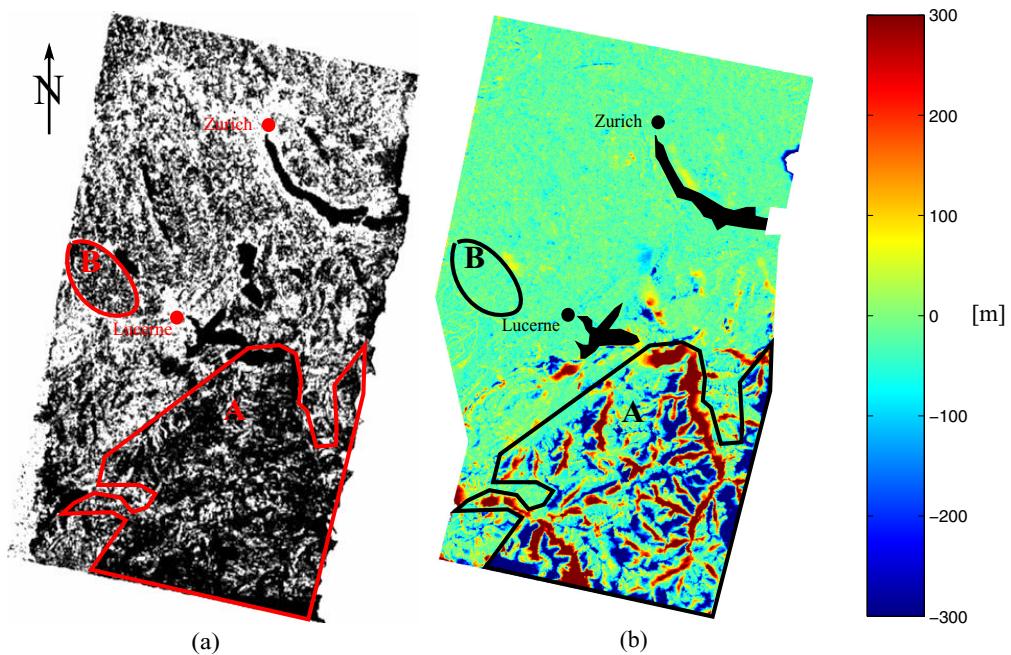


Figure 8.14 Zurich stereo results: (a) match confidence (b) height error. Area within marked perimeter corresponds to large height error in (b) and low match confidence in (a). The Lakes of Zurich and Lucerne are drawn in (b) for orientation purposes.

area 'B' is an example of this.

Figure 8.15 shows the scatter plot of the absolute height errors for Zurich versus the match confidence for all points in the DSM. Most of the points with errors above 150 m are confined to match confidences below about 2. Comparing its general shape to that of Küttigkofen from Figure 8.9 on page 122, a similar relationship can be seen, although the height error declines somewhat more rapidly in the low-confidence zone than for Küttigkofen. This is indicative of a greater number of higher-quality heights for a given match confidence. In both cases the trend is clear: higher match confidence values correspond to lower the average height errors.

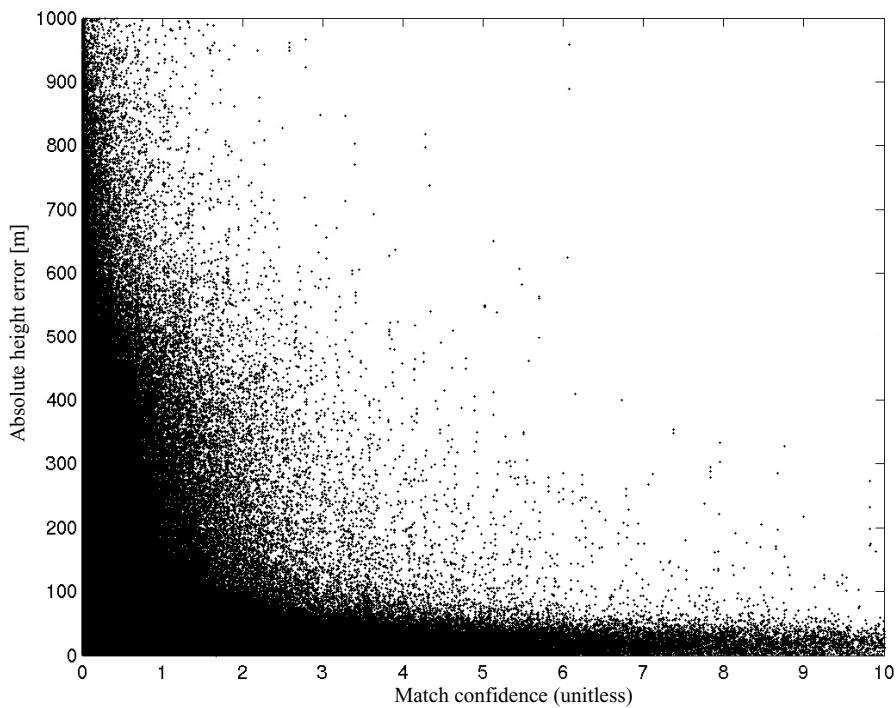


Figure 8.15 Absolute height error versus match confidence for the Zurich stereo DSM.

Chapter 9

Stereo-Assisted Interferometry Results

9.1 Introduction

While the stereo processing results presented in the last chapter are of interest in themselves, the goal of this work is to show how the combination of the match confidence and low-resolution DSM can be used to improve the results obtainable by InSAR. To clarify: it is not the primary goal of this work to obtain an InSAR DSM of unparalleled accuracy. Rather, it is shown that the use of parallel stereo processing can greatly improve the *achievable* InSAR result, while making the processor automatable at the same time.

If a low-resolution DSM of similar (or better) quality as is possible with stereo already exists for the scene to be InSAR-processed, the utility of the technique becomes doubtful. The role of the stereo DSM within an InSAR framework, as was shown in Figure 5.1 on page 74, can likely be performed equally well by the existing DSM.

The stereo-assisted technique becomes useful for those cases where no DSM is available at the outset. In this case the proposed technique provides a possible means to automatically generate an InSAR DSM of *higher quality than would have been obtainable otherwise*. In practice, "higher quality" translates to increased coverage (fewer holes) at an acceptable cost to the overall DSM accuracy.

This chapter will examine the usefulness of the proposed technique for the three test sites Küttigkofen, Emmen, and Zurich. The results were obtained in two ways: first, without the aid of an existing DSM (InSAR alone), and second, by applying the proposed stereo-assisted InSAR technique.

The goal of this work was to determine in what ways parallel stereo processing is the most useful, taking into account the significantly-lower quality of the stereo DSMs as compared to that of typical InSAR DSMs, as well as the unique nature of each site. While a stereo model generally provides near-100% coverage, one must take care not to let it overly-falsify the more reliable interferometric measurements. For each case, the stereo-assisted results are compared to those obtained without any external height models,

and the optimal fusion of the two techniques is discussed.

The results in this chapter were produced according to the diagram shown in Figure 5.1 on page 74. This includes an InSAR processor, whose individual steps are performed according to [78], with the addition of the stereo interventions shown in positions 1-3 in the same figure. The goal of this thesis is to show how such an InSAR processor can benefit from previous stereo processing. Therefore, we will concentrate on these aspects of the processing only.

9.2 Airborne Case: Küttigkofen

9.2.1 Introduction

The results from seven experiments, or processing chains, are described here. At the end of each chain is an interferometrically-derived DSM, which will be validated. These experiments are derived from seven different ways to combine the three steps shown in Figure 5.1 on page 74, namely:

- 1) Phase flattening: given an approximate DSM, the fringe density can be greatly reduced, easing phase unwrapping. In the absence of a DSM, ellipsoid flattening must be performed.
- 2) GCP generation: GCPs are required for the phase- and baseline calibration before a height model can be obtained. This auxiliary height information can come from topographical maps (image tiepoints), reflectors placed on the ground whose positions have been precisely measured, or a DSM such as is generated using stereo SAR. Automation is only possible in the latter case.
- 3) Data merging: if a stereo model is available, along with a measure of its quality (which is obtained from the match confidence), one can partially fill any holes remaining in the InSAR DSM. These values provide "stepping stones" (boundary conditions) for the subsequent cubic interpolation across the holes, performed in the final step. In the absence of a stereo DSM, the cubic interpolation is performed directly. In all cases, however, a maximum hole size is set, and only holes smaller than this threshold will be filled.

The seven experiment configurations are shown in Table 9.1. They have been assigned numbers for quick reference, as well as names that reflect what was in fact done. For each case the Küttigkofen data sets were interferometrically processed, and a DSM produced in each case. Depending on the configuration, one or more of the three stereo-based steps were

Table 9.1 Küttigkofen stereo-InSAR experiment configurations

Experiment number	Experiment name	Phase flattening	GCP Collection	Hole Filling
1	ellFlat_mapCal_noFill	Ellipsoid	1:25000 digital map, survey points	Small-hole interpolation
2	ellFlat_sterCal_noFill	Ellipsoid	Stereo DSM (automatic)	Small-hole interpolation
3	ellFlat_sterCal_smartFill	Ellipsoid	Stereo DSM (automatic)	"Smart" stereo-merging + Small-hole interpolation
4	sterFlat_sterCal_noFill	Stereo DSM	Stereo DSM (automatic)	Small-hole interpolation
5	sterFlat_sterCal_blindFill	Stereo DSM	Stereo DSM (automatic)	"Blind" stereo-merging + small-hole interpolation
6	sterFlat_sterCal_smartFill	Stereo DSM	Stereo DSM (automatic)	"Smart" stereo-merging + small-hole interpolation
7	sterFlat_mapCal_smartFill	Stereo DSM	1:25000 digital map, survey points	"Smart" stereo-merging + small-hole interpolation

performed. These steps are described in the sections below.

9.2.2 Phase Flattening

Before converting the interferometric phase to topographic heights, the topographic phase is removed using the best available terrain height model, whether a DEM or an ellipsoid approximation (this was described in section 4.2.4 on page 65). Using a stereo model to reduce the fringe density as opposed to an ellipsoid model limits the unwrappable area to the geographic overlap zone between the InSAR phase and the stereo DSM. Because of the

particular flight configuration flown over Küttigkofen, namely, that the interferometric master was also the stereo master, the overlap (shown in Figure 7.4 on page 90) covered approximately 50% of the InSAR scene. Thus, the interferometric phase could only be unwrapped over 50% of the area using the stereo DSM. The rest of the scene outside of the overlap zone could not be further processed; it was automatically masked off for the rest of the stereo-InSAR processing.

It should be noted that the stereo-InSAR processor does not require the InSAR and stereo masters to be the same; indeed, in general one can expect the stereo acquisitions to come from another pair of flight tracks or even sensors altogether. The Küttigkofen flight campaign was in a sense ideal: given a fixed antenna look direction, the maximum possible overlap between stereo and InSAR was attained. That is, the entire stereo window was also acquired with InSAR.

The utility of stereo-based flattening as compared to ellipsoid flattening is shown in Figure 9.1. As a comparison, Figure 9.1(a) illustrates the unflattened phase. The phase cycles (fringes) are color-coded, with the range from blue through yellow indicating the phase value in the range $[0, 2\pi]$. The color cycling illustrates the 2π -modulo phase ambiguity inherent in InSAR.

The ellipsoid-flattened phase in Figure 9.1(b) illustrates that a significant fringe pattern remains as a result of the topography over the ellipsoid model. Were the terrain extremely steep, such as in the mountains, phase unwrapping would be impossible here, since the phase ambiguity has not been resolved. In this case, the ellipsoid-flattened phase does not present an especially-difficult unwrapping case, however, stereo flattening was done here as well for comparison of the final results.

Figure 9.1(c) shows the stereo-flattened phase. Because the general shape of the real terrain is captured by the stereo model, no clear fringe development is visible. Partial fringes remain, in particular in the shadow zones on the northern edges of the tree stands, where the stereo DSM exhibits large errors. Nonetheless, the complexity of the phase-unwrapping problem has been simplified by flattening with a DSM.

9.2.3 Coherence

The coherence was now calculated for the flattened interferogram, not the original one, to avoid underestimation due to the range phase cycling (the "flat-Earth" fringes) present in the original [78]. The 1.89 m interferometric baseline flown in Küttigkofen from the low altitude of 2800 m was large enough to cause severe coherence loss of the ~ 3 cm-wavelength radiation over the forested areas. This is due to the significant change in the scattering-

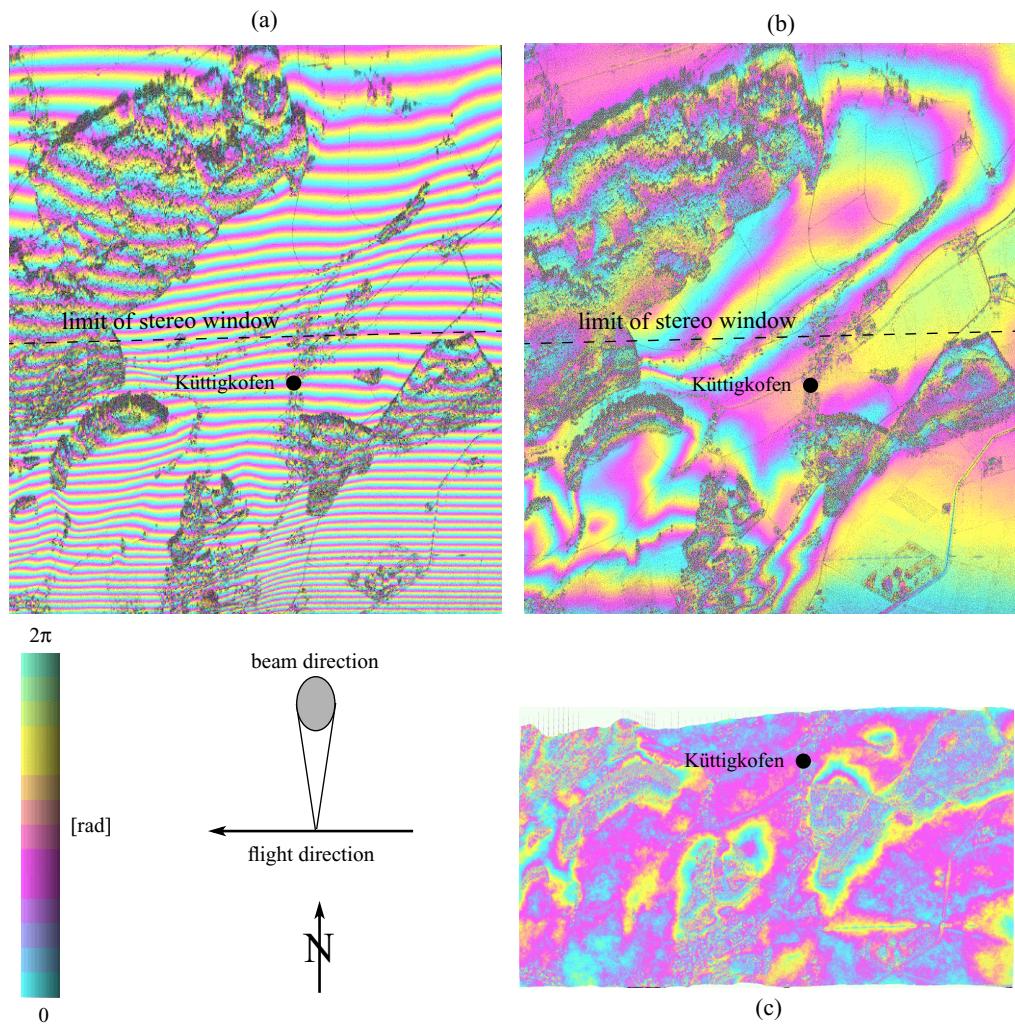


Figure 9.1 Original and flattened phase for Küttigkofen: (a) unflattened phase (b) flattened using an ellipsoid-Earth model (c) stereo-DSM-flattened phase. The coverage of the stereo DSM constrains the stereo-assisted processing chain to the same area, in this case the southern (lower) half of the source acquisition.

return properties of vegetation with relatively small changes in the arrival direction of the waves.

The low coherence over vegetation in Küttigkofen is indicated by the dark zones in Figure 9.2. It makes phase unwrapping a nearly impossible task for these areas irrespective of the success of phase flattening: it is ultimately responsible for the data holes seen later on in the interferometric DSM.



Figure 9.2 Küttigkofen coherence (black = 0, white = 1.0)

9.2.4 Ground Control Point Collection

For the two map-calibrated experiments (numbers 1 and 7 on page 133), GCPs were collected manually, making these experiments non-automatic. The other cases used a list of automatically-generated GCPs based on a subset of all points from the stereo DSM selected according to the criteria outlined in section 5.3 on page 73. The criteria were designed to select points with match confidences greater than a pre-defined threshold, outside of shadow or dark areas, over non-vegetated ground, and over relatively-flat ground (implying a near-constant local topographic phase). The distribution of the GCPs across the image is also very important; too much clustering may bias the phase calibration, especially when the GCPs are not extremely accurate. The current search algorithm produces a wide distribution in most cases by way of its design, although it does not specifically attempt to optimize it.

Setting the minimum number of points to about 30 tended to produce between 30 and 40 points across the scene. This was deemed sufficient, because the distribution was good and point clustering was avoided (section 5.3 on page 73 explained why this is desirable).

In order to validate the automatic choice of GCPs, the mean height error is compared to the standard deviation for this set of points to the statistics for the whole stereo DSM, relative to a given reference. For a set of randomly-chosen

points, the GCP error statistics should resemble those of the whole DSM, given a statistically-valid sample. If the selection criteria improve the error statistics of the GCP ensemble as compared to those of the stereo DSM, then the criteria are considered to be effective.

As was seen in Table 8.1 on page 111, the overall height error for the stereo DSM relative to DoSAR is 3.9 ± 7.6 m, with 59.5% of the points under 5 m. The automatic search algorithm produced a set of 34 GCPs based on the stereo DSM, whose positions are marked by crosses in Figure 9.3 against the slant-range amplitude image (since the point coordinates are returned in slant range). The points are spread across the area covered by the stereo DSM; in this case the lack of coverage in the northern half can be expected to result in a slight phase calibration bias. The 34 GCPs present a height error of -2.1 ± 2.4 m, with 91.2% under 5 m.

For the automatically-selected GCPs, the mean has moved closer to zero and the standard deviation has decreased by a factor of 3.2, with a greatly-increased percentage of points under the 5 m-error mark. This is a strong indication that the chosen set of selection criteria works quite well in filtering out the less accurate points from the stereo model.

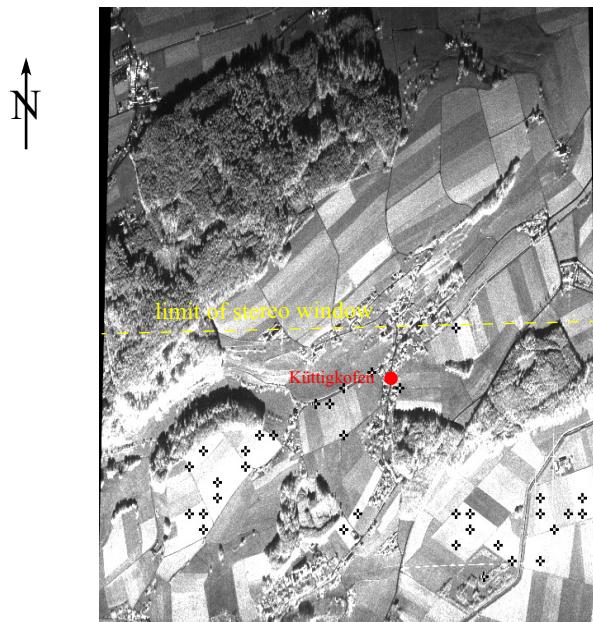


Figure 9.3 Locations of automatically-extracted GCPs for Küttigkofen. Background is AeS-1 stereo master amplitude

9.2.5 Data Merging and Interpolation

The stereo-based "smart" hole-filling is illustrated in Figure 9.4 for the ellFlat_sterCal_smartFill experiment. Figure 9.4(a) shows the initial result after ellipsoid phase flattening and stereo-based phase calibration. The "stereo-padded" holes can be seen in Figure 9.4(b), and the final result after interpolation of the small holes in Figure 9.4(c). The holes that remain are those whose initial diameter was over the pre-defined size threshold, 25 pixels in this case. Holes larger than this threshold are considered too large to justify interpolation.

9.2.6 DSM Validation

The final height errors relative to DoSAR for all seven experimental DSMs can be seen in Figure 9.5. The DoSAR model, rather than the laser DSM, is the major focus for reasons described in section 7.3.3.2 on page 92. The corresponding statistics are given in Table 9.2. Statistics relative to the laser DSM are given as well, although the images are not shown.

The error for the DSM created entirely without the use of the stereo data is shown in Figure 9.5(a) [case 1; refer to Table 9.1 on page 133 for experiment numbers]. Ellipsoid flattening is feasible for this case, since little steep terrain exists. Although it has the lowest standard deviation of all the cases (relative to a given reference), its calculation required a manual search for GCPs and tiepoints, which may in some cases not be available. Also, the forest stands are entirely missing, which helps explain the low error rate.

Replacing the manual phase calibration with the automatic stereo-based version, the result is Figure 9.5(b) [case 2]. The quality has only slightly been degraded, and this is due to the limited overlap between the stereo and interferometric acquisitions, resulting in a sub-optimal distribution of GCPs.

Taking the (b) result and performing the "smart" hole-filling described in section 5.4 on page 76, Figure 9.5(c) is obtained [case 3]. Result (c) is an overall improvement in (b), since more height information has been incorporated in the final result, with the strongest errors restricted to small regions.

Figure 9.5(d) [case 4] is of similar quality to (c), but due to the stereo-based phase flattening, the region north of the stereo window has been masked off.

Adding the stereo heights to the holes without considering their quality, Figure 9.5(e) is the result. The strong errors from the radar shadow are allowed in, bringing down the overall DSM quality significantly. Performing a "smart" hole-filling on the (d) result instead, such as was demonstrated in Figure 9.4, the strongest errors from the stereo DSM are not allowed in, and

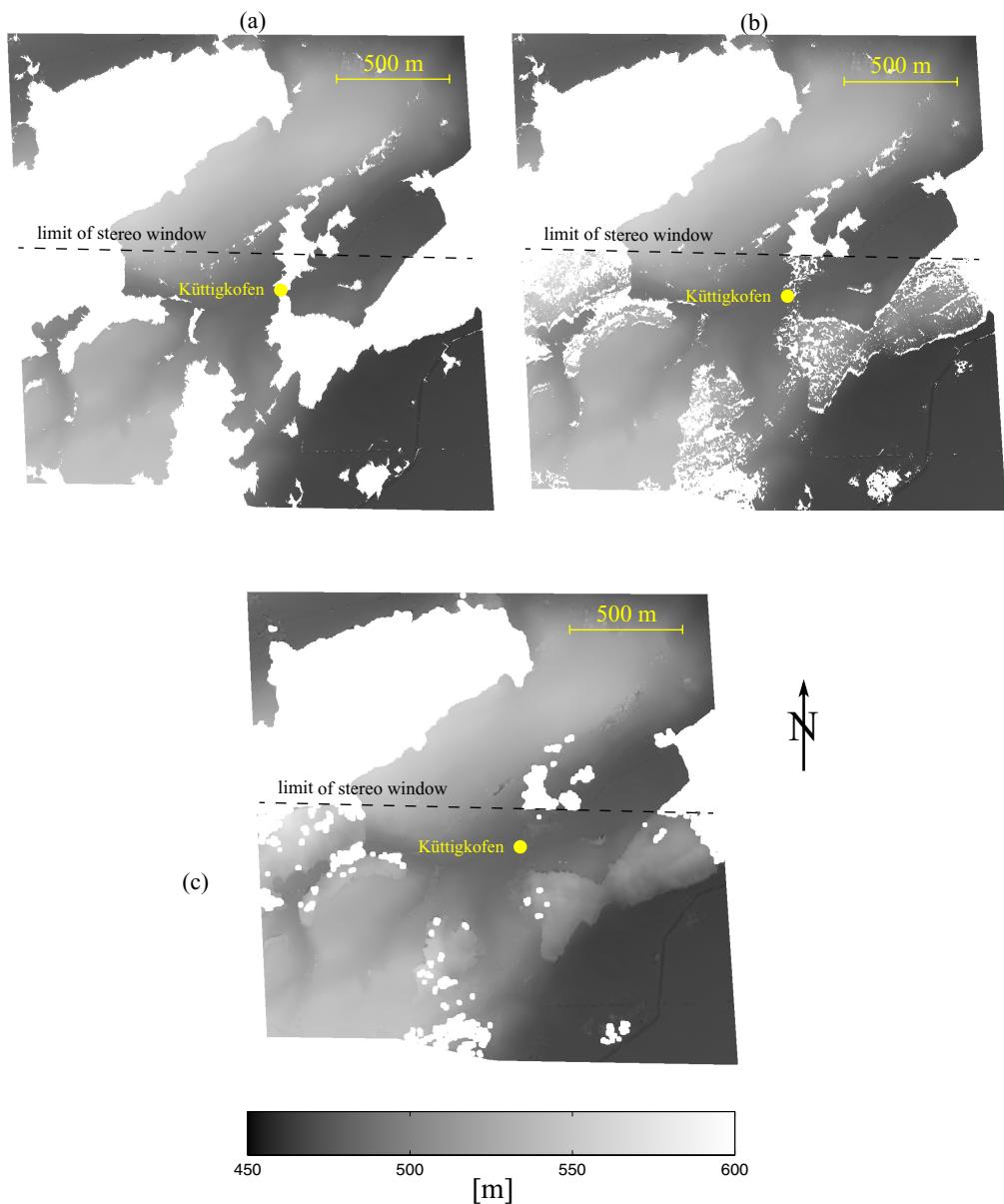


Figure 9.4 "Data merging and interpolation ("smart"-filling) for the ellFlat_sterCal_smartFill case in Küttigkofen (a) after ellipsoid flattening and stereo-based phase calibration (b) after data merging (for area covered by stereo-DSM) (c) after cubic interpolation of the smaller holes

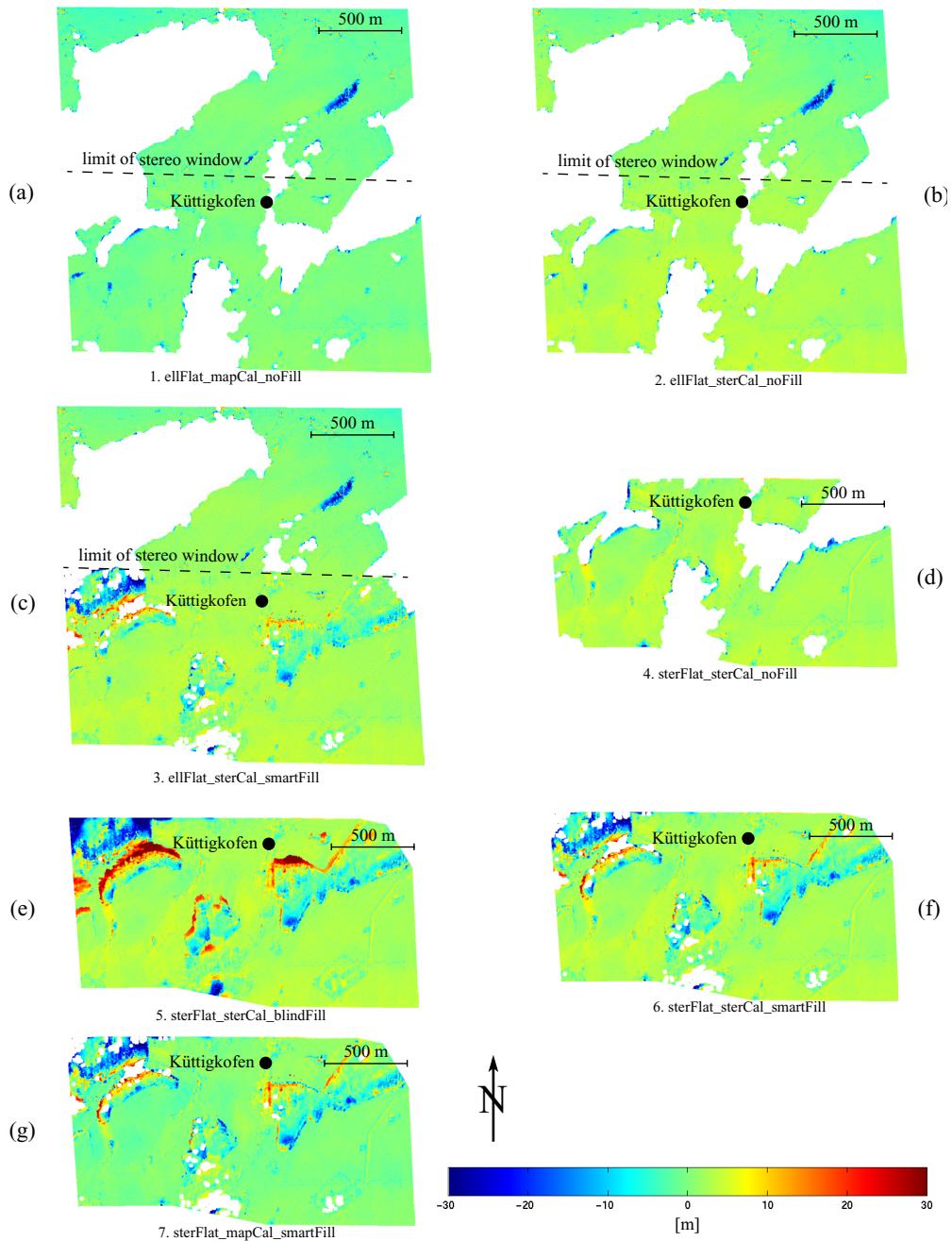


Figure 9.5 Küttigkofen stereo-assisted InSAR-DSM height error relative to DoSAR, for all seven experiments, labelled by number and name. The color bar at the bottom applies to all seven results.

Table 9.2 Error statistics for Küttigkofen InSAR experiments ("best" case highlighted)

Experiment	Reference	Mean height error [m]	Std Dev [m]	Pixels with < 5 m error [%]	Pixels with < 10 m error [%]	Pixels with < 20 m error [%]	Total valid area [km ²]
1. ellFlat_mapCal_noFill	DoSAR	-0.71	2.48	96.5	98.7	99.8	2.89
	Laser	-0.43	2.59	96.7	98.6	99.6	
2. ellFlat_sterCal_noFill	DoSAR	1.46	2.84	96.1	99.0	99.8	2.89
	Laser	1.75	2.91	96.5	99.7	100.0	
3. ellFlat_sterCal_smartFill	DoSAR	1.17	3.88	91.7	97.0	99.5	3.45
	Laser	1.66	4.63	89.6	95.0	98.8	
4. sterFlat_sterCal_noFill	DoSAR	2.23	2.79	94.0	98.0	99.8	1.60
	Laser	2.52	2.72	95.4	98.7	99.7	
5. sterFlat_sterCal_blindFill	DoSAR	1.80	6.75	82.3	91.2	96.9	2.22
	Laser	2.44	7.47	80.3	88.5	95.6	
6. sterFlat_sterCal_smartFill	DoSAR	1.50	4.73	86.4	94.9	99.1	2.11
	Laser	2.09	5.53	84.6	92.3	98.2	
7. sterFlat_mapCal_smartFill	DoSAR	-0.43	4.52	87.7	94.7	99.0	2.11
	Laser	0.17	5.44	85.3	92.2	98.1	

Figure 9.5(f) [case 5] is the result.

As a further comparison, re-running the experiment from Figure 9.5(d), but substituting the map-extracted GCPs for the automatic set during phase calibration, Figure 9.5(g) is obtained. The result is only slightly improved using the manually-selected GCPs.

Taking into account the above results, it can be concluded that the best possible DSM obtainable for the datasets available - while still remaining fully-automatic - is one which combines ellipsoid flattening, stereo-based phase calibration, and "smart" stereo-based hole-filling, seen in Figure 9.5(g). Not only is the large area covered by the InSAR pair retained, but some of the data holes have been filled, although not without occasional strong errors. With a total valid area of 3.45 km² of assigned heights - the highest of all experiments - and ~ 92% of points with errors under 5 m, it is believed that the "ellFlat_sterCal_smartFill" method (experiment 3) is a good compromise between accuracy and total coverage. Certainly, it is a more satisfying result than the one produced without any access to the stereo DSM (Figure 9.5(a)),

especially considering that it was obtained through an automatable process.

9.3 Airborne Case: Emmen

9.3.1 Introduction

The case of Emmen is similar to that of Küttigkofen, with one major difference: the normal interferometric baseline is much smaller, at 0.57 m. In airborne InSAR the normal baseline varies rapidly from near- to far-range, differently in each case. By calculating the inputs to Equation 4.3 on page 65 for both cases in the near and far ranges, it is possible to estimate that the Küttigkofen interferometer is roughly four times as sensitive as the Emmen configuration near mid-range, but this factor can range from three and five times, depending on the location. Indeed, this effect can be confirmed by comparing the fringe frequencies between Figures 9.1(a) and 9.6(a). Thus, a given phase error is expected to produce a height error in Emmen roughly four times greater than in Küttigkofen. In addition, the higher flight altitude in Emmen will reduce the SNR as compared to Küttigkofen, and thus one can expect this error factor to increase further as well.

For Emmen, as for Küttigkofen, InSAR and stereo processing had one data set in common. The stereo slave data set was obtained from the nearest parallel track, while the second InSAR sensor was fixed on the same fuselage. This resulted in the partial overlap between the InSAR and stereo datasets once again, reducing the area of the stereo-flattened phase and DSM based on it.

9.3.2 Phase Flattening

The Emmen interferometer has an ambiguity height at mid-range of slightly over 100 m. It is immediately clear that the scene's mere 110 m height variation will not affect the fringe pattern a great deal beyond the range-cycling trend. Indeed, the interferogram in Figure 9.6(a) seems nearly unaffected by the topography. Subtracting the ellipsoid phase confirms that approximately one topographic fringe remains, in rough agreement with the ambiguity height.

Naturally, flattening the phase using the stereo DSM is not desirable in this case, since the phase unwrapping is straightforward and it is preferable to continue with the full coverage of the InSAR scene. The stereo-flattened result is shown in Figure 9.6(c) nonetheless for comparison; it is not surprising that no full fringes remain.

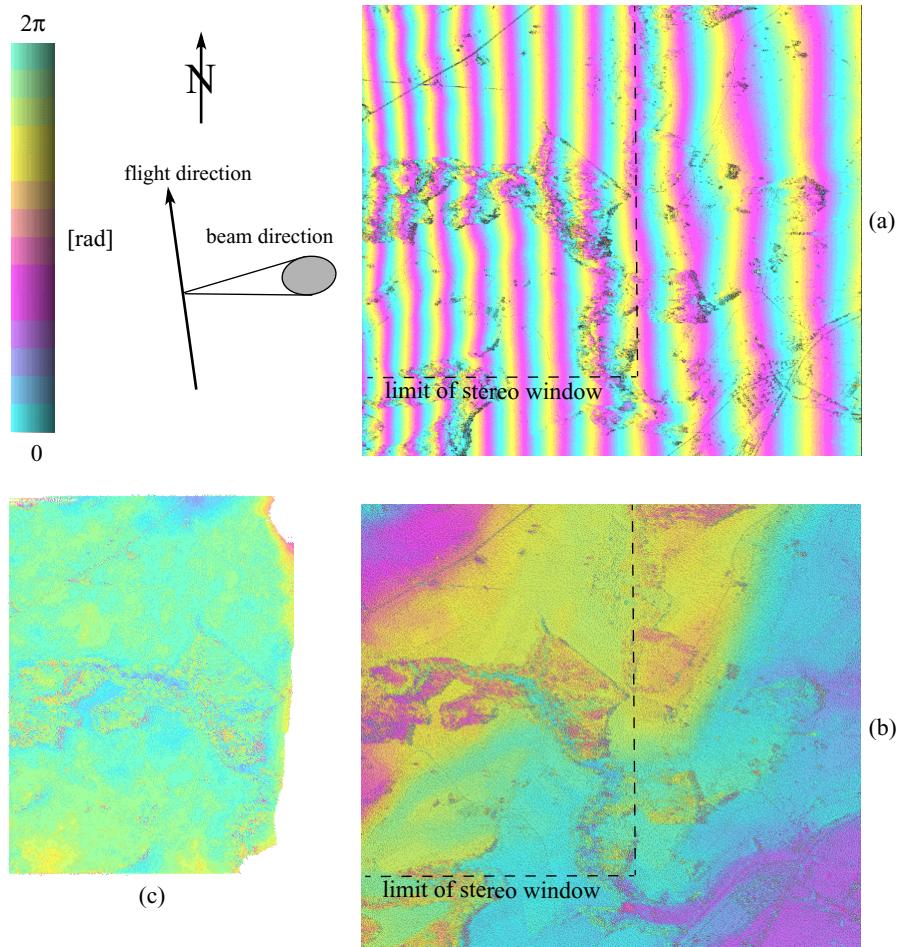


Figure 9.6 Original and flattened phase for Emmen: (a) unflattened phase (b) flattened using an ellipsoid-Earth model (c) stereo-DSM-flattened phase. The coverage of the stereo DSM sets the boundaries of the stereo-assisted processing chain to the same area, in this case the western (left-most) half of the scene indicated in (b).

9.3.3 Coherence

One advantage is gained with the short baseline: phase decorrelation over vegetation is reduced, making phase unwrapping feasible. As a result, nearly no data holes exist in the InSAR-only DSM, so hole-filling becomes unnecessary. As can be seen in Figure 9.7, the coherence over the forested areas, while darker than elsewhere, is not as low as it was in Küttigkofen.

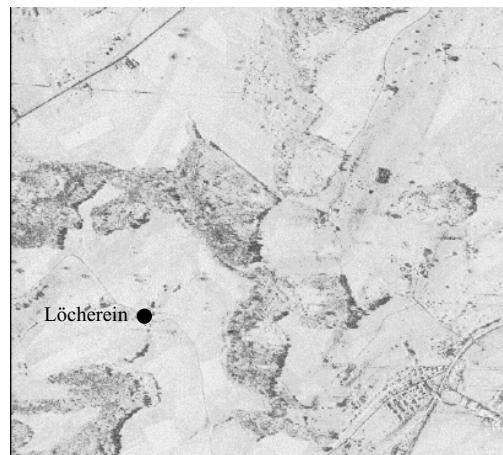


Figure 9.7 Emmen coherence (black = 0, white = 1.0)

9.3.4 Ground Control Point Collection

The automatically-extracted GCPs for Emmen are marked in Figure 9.8. Fifty points were found, a few of which are over the trees. The statistical difference between the GCP set and the DHM25 was -3.9 ± 3.9 m, although the points over the forest slightly worsen the result, since the DHM25 is a terrain-only model. The DSM-versus-DHM effect is more pronounced when subtracting the DHM25 from the stereo DSM, giving 6.2 ± 6.5 m. Based on the results alone, it would seem that the automatic GCP selection slightly improved on the statistics of the stereo model. However, for lack of a high resolution surface-model reference, it is not possible to compare these figures directly.

9.3.5 DSM Validation

The short baseline in Emmen made interferometric height estimation over the forested areas possible, with nearly no holes remaining in the DSMs. After phase calibration and geocoding for experiment configurations similar to those used for Küttigkofen (Table 9.1 on page 133), but without the hole-filling, the results were compared to the available references. As expected, the stereo-flattened cases were not more accurate than the ellipsoid-flattened ones, while at the same time limiting the coverage. Therefore, only the ellipsoid-flattened case results are shown here.

The DSM obtained by the ellipsoid-flattened, stereo phase-calibrated result

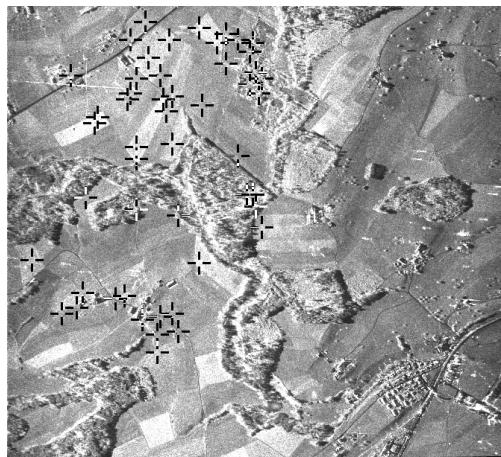


Figure 9.8 Locations of automatically-extracted GCPs for Emmen. Background is AeS-1 slant-range stereo master amplitude.

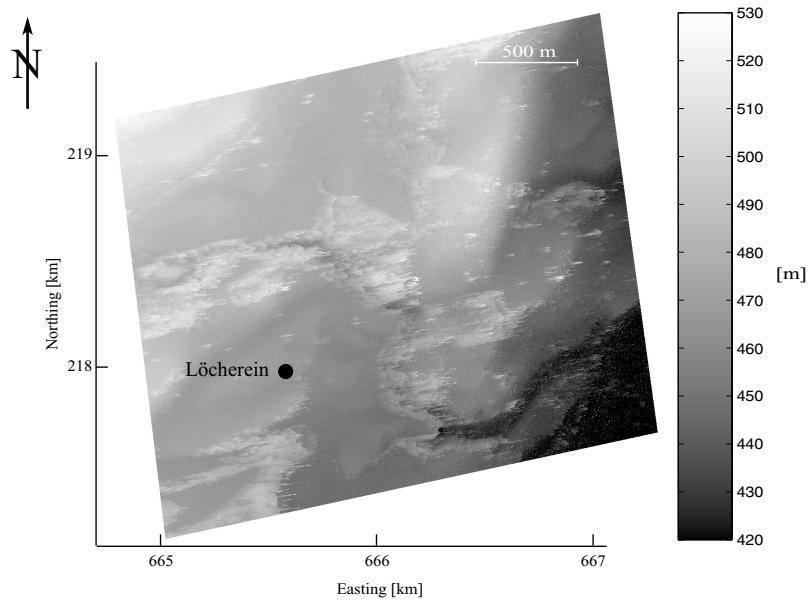


Figure 9.9 Emmen stereo-InSAR DSM, generated by ellipsoid-flattening, stereo phase calibration, and cubic interpolation of the small holes.

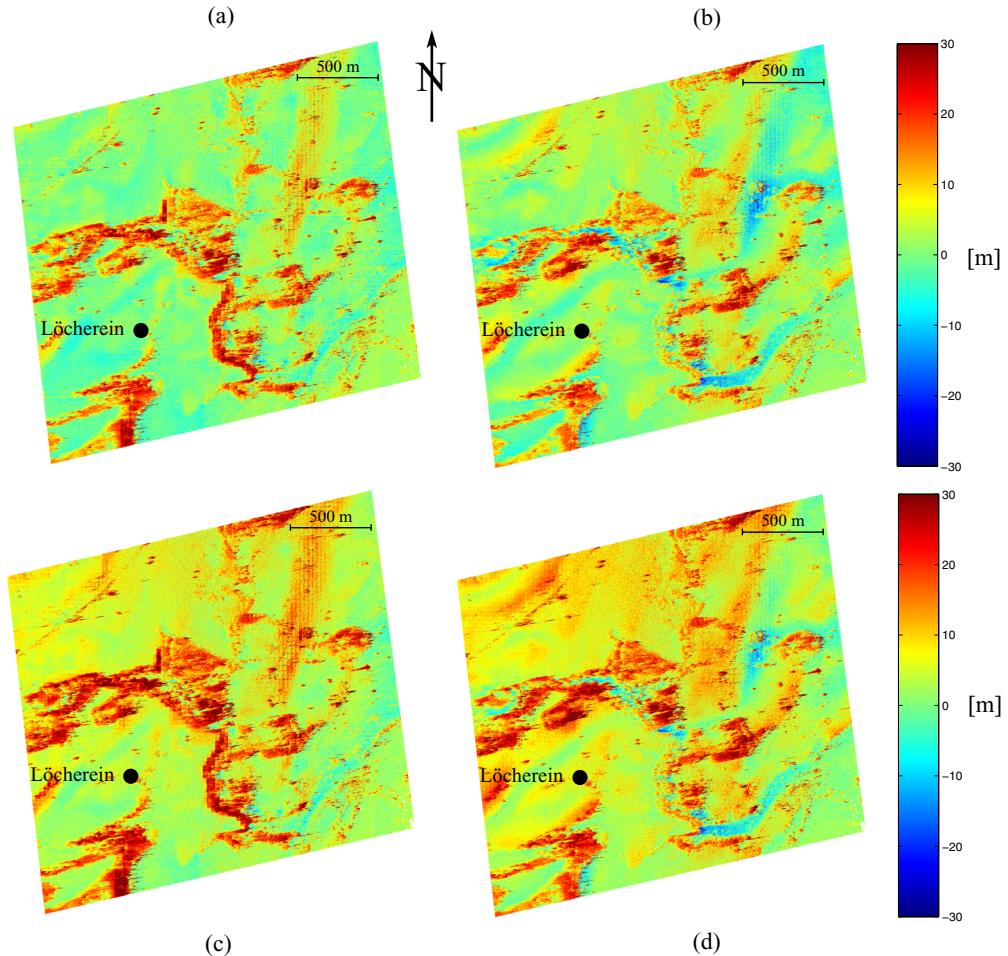


Figure 9.10 Emmen ellipsoid-flattened InSAR-DSM results relative to DHM25 and DSM25: (a) map-calibrated case relative to DHM25 (b) map-calibrated case relative to DSM25 (c) stereo-calibrated case relative to DHM25 (d) stereo-calibrated case relative to DSM25

is shown in Figure 9.9. Comparing it to the slant-range amplitude images or the map extract in Figure 7.11 on page 96 (the red area is the InSAR perimeter), it is clear that the forest stand canopies have been successfully reconstructed. The coverage and detail recovery are both greater than for the pure-stereo DSM (Figure 8.5 on page 112).

The most obvious feature of the differences between the InSAR DSMs and the two reference DSMs, shown in Figure 9.10, is the forest overestimation, which is expected. It was seen in Figure 7.14 on page 98 that while technically

a surface height model, the DSM25 does not show the forest canopy - at least not at the small scale of the airborne data set. The non-forested areas are approximately correct, with a tendency to overestimate by a few meters in the stereo-calibrated cases.

In order to compare the two DSMs against the map-based GCPs, shown in Figure 7.15 on page 99, a 7x7-pixel low-pass filtering of the DSM was first performed. This smooths out the high-frequency height variation typical of small-baseline DSMs, and makes the validation against so few points more reliable. The filtering is also justified by the relative flatness of the scene.

The ellipsoid-flattened, map-calibrated result (i.e. no stereo intervention) has a height error of 0.4 ± 1.8 m for the 13 GCPs. The slight overestimation seen in Figure 9.10(c) and (d) for the stereo-calibrated version is confirmed by the GCP validation, the latter displaying an error of 4.6 ± 2.1 m.

One can also compare the stereo-calibrated result to that of the ellipsoid-flattened, stereo-calibrated DSM obtained for Küttigkofen (the ellFlat_sterCal_noFill case; Figure 9.5(b) on page 140 and Table 9.2 on page 141). Emmen is compared to the case without the filled holes, since the set of Emmen GCPs are mainly over non-forested ground. The 2.1 m map-based standard deviation of the stereo-calibrated Emmen DSM is in fact less "noisy" than the 2.8 m value obtained for Küttigkofen, although the mean (4.6 m) is quite high. According to the calculations performed in section 9.3.1 on page 142 based on the ambiguity heights for Emmen and Küttigkofen, a standard deviation in Emmen of ~ 10 m or more had been expected.

With these figures in mind, and given the lack of reliable and statistically-significant reference information for Emmen, it can be tentatively concluded that the Emmen DSM produced automatically by ellipsoid flattening and stereo-based phase calibration is of good quality, considering the short interferometric baseline.

9.4 Spaceborne case: Zurich

9.4.1 Introduction

There are two major differences between the Zurich experiment and both others:

- 1) It is spaceborne, thus covers a vast expanse of terrain. The Alps present a unique challenge, while smaller-scale features such as forest canopies and housing do not play a significant role.
- 2) The stereo and InSAR datasets are from different sensors (ENVISAT and

ERS-1/2, respectively). This means that the geometries of the stereo and InSAR data are completely unrelated, whereas each AeS airborne campaign involved a shared data set (from the master sensor) between the stereo and InSAR pairs, simplifying the geometric transformation relating the stereo- and InSAR-processed outputs. The Zurich scenario, with different sensors and acquisition geometries providing the stereo and InSAR data-sets, is more representative of what is typically available - in particular when dealing with spaceborne scenes.

The usefulness of stereo-assisted InSAR processing for these data will be demonstrated, as compared to InSAR without auxiliary information (which is referred to as unassisted InSAR).

9.4.2 Phase Flattening

The foothills and mountains towards the south-east of the Zurich interferometric site pose a difficult phase-unwrapping problem. With a mid-range ambiguity height (Equation 4.3 on page 65) of ~ 67 m for the ERS-1/2 pair, rapid increases in terrain elevation over short distances result in high fringe rates. Put another way, an elevation change of 67 m corresponds to a full fringe, and this fringe is expressed in the slant-range geometry. When the sides of steep terrain are facing the sensors, layover will cause fringes to be compressed to within several range pixels or less.

The critical nature of the phase flattening is evident in Figure 9.11, which is a magnified extract, improving the fringe visibility. While ellipsoid flattening (Figure 9.11(b)) of the interferometric phase (Figure 9.11(a)) greatly reduces the fringe rates around the lake of Zurich, it is nearly useless in the mountainous terrain towards the south-east. As a best-case comparison, Figure 9.11(c) shows the phase after flattening using the DHM25, which is considered an accurate height model. No fringe pattern is clearly visible, however, fringe residuals (partial fringes) are common in the image, in particular in the Alps. These are mainly due to differences between the DHM25 and the true topography.

It should be noted that there is one other effect, other than topographical errors, that can result in such phase residuals: atmospheric interference. When viewing the surface from 800 km altitude it is possible that differences in the ionosphere and troposphere between the two acquisition times can influence the interferometric path difference by up to several meters [46]. This will be visible as systematic height- and time-dependent phase residuals even if the topographic phase has been perfectly subtracted from the interferogram. However, it is statistically unlikely that large differences will exist from one

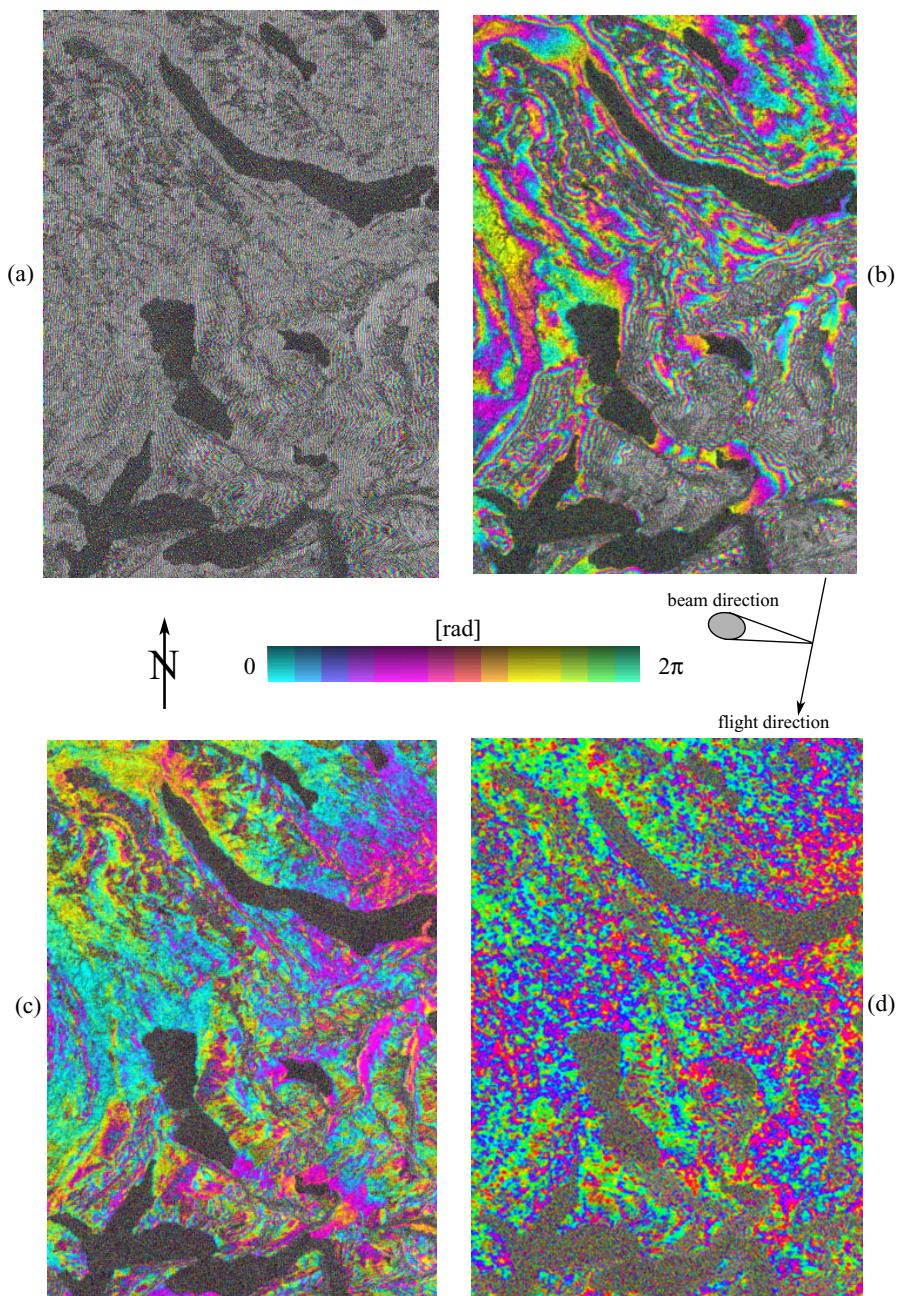


Figure 9.11 Phase flattening of the Zurich ERS-1/2 pair (scene extract only): (a) unflattened phase (b) ellipsoid-flattened phase (c) DHM25-flattened phase (d) stereo DSM-flattened phase

day to the next, especially for the same time of day.

Figure 9.11(d) is largely similar to (c), with the addition of a granular noise component from the ENVISAT stereo DSM. This granular structure, however esthetically displeasing, is not of a high enough frequency to seriously hinder phase unwrapping.

9.4.3 Coherence

The coherence for Zurich is shown in Figure 9.12. ERS-1 and -2 were in a repeat-pass configuration, the acquisitions separated by a day. The coherence for the ERS-1/2 pair over the Zurich site is therefore low over forested areas, which change their volumetric scattering properties over the course of time. Water bodies also have low coherence, due to the weak - if any - signal return.

9.4.4 Ground Control Point Collection

In order to contrast the stereo-assisted with the unassisted InSAR results, phase calibration was performed using both a set of 10 GCPs hand-picked from topographical maps and 44 points selected automatically.

The points chosen by the "smart" algorithm for the ENVISAT-based stereo

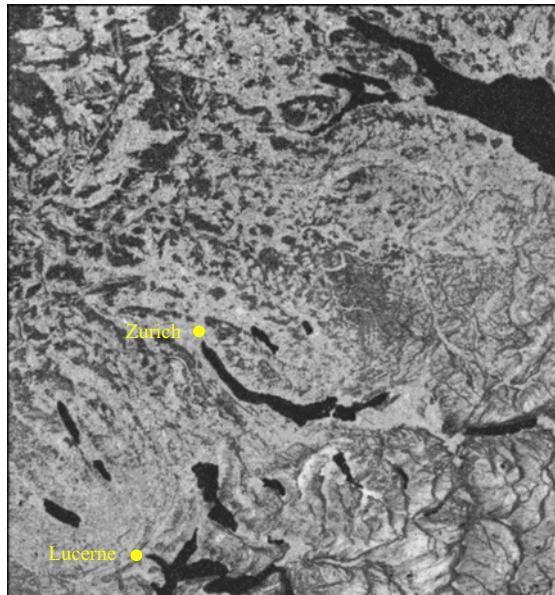


Figure 9.12 Zurich ERS-1/2 coherence (black = 0, white = 1.0)

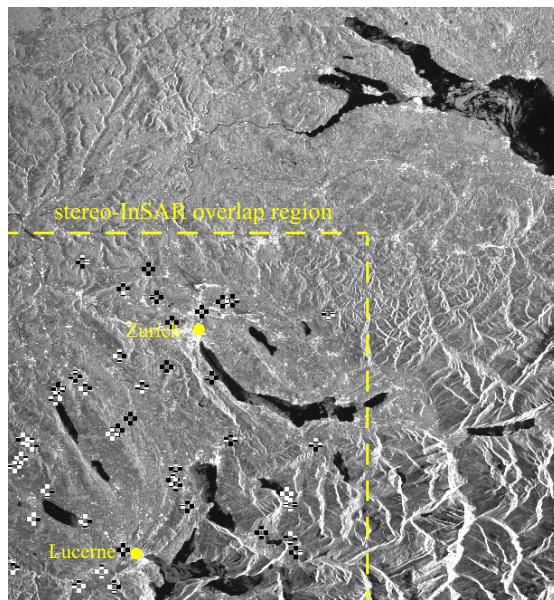


Figure 9.13 Locations of 44 Zurich GCPs used for phase calibration. The background is the ERS-1 slant-range amplitude. [© ESA]

DSM are visible in Figure 9.13. The 44 points are distributed quite evenly across the part of the scene covered by ENVISAT. Compared to the DHM25, they have a mean height error of 2.6 m with a standard deviation of 19.8 m. By comparison, the stereo DSM from which they were selected has a height error of -20.8 ± 148.4 m. This represents an over seven-fold improvement in the standard deviation, and a significantly-lower error bias.

9.4.5 Data Merging and Interpolation

The dramatic effect of the "smart" hole-filling algorithm, which uses similar selection criteria as for the GCP selection, can be seen in Figure 9.14. The partial filling of the InSAR-DSM holes with a judiciously-chosen set of points from the stereo DSM greatly improves the overall coverage (Figure 9.14(b) versus Figure 9.14(a)). The final small-hole-interpolation step closes the smallest holes, resulting in an esthetically-pleasing result (Figure 9.14(c)) with few large holes aside from the lakes.

9.4.6 DSM Validation

The validity of the stereo-assisted interferometry is best visible by

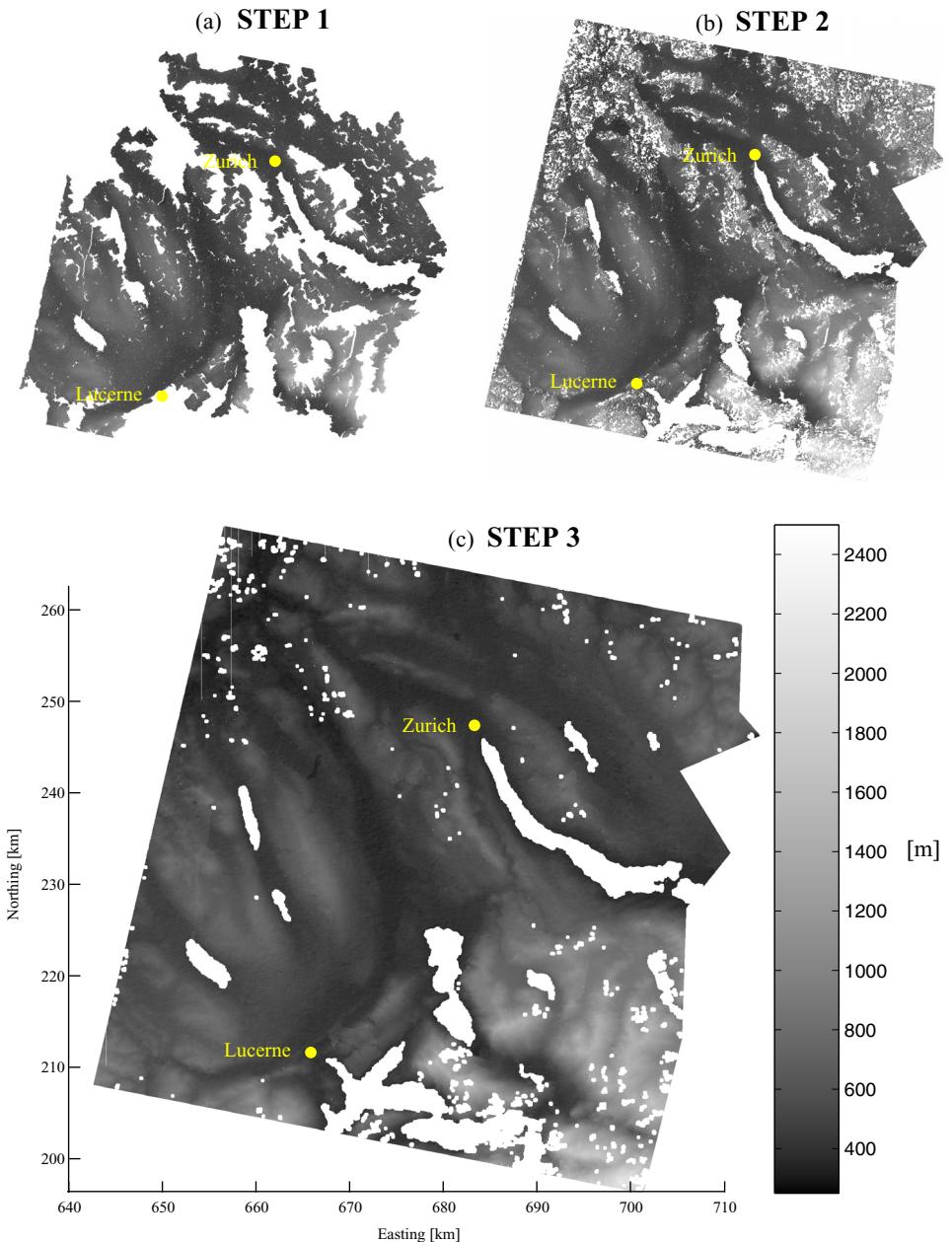


Figure 9.14 Data merging and interpolation steps ("smart"-filling) for Zurich (a) stereo-based flattening and phase calibration leaves large holes (b) stereo values are used to partially close the holes (c) cubic interpolation closes up the smallest of the remaining holes

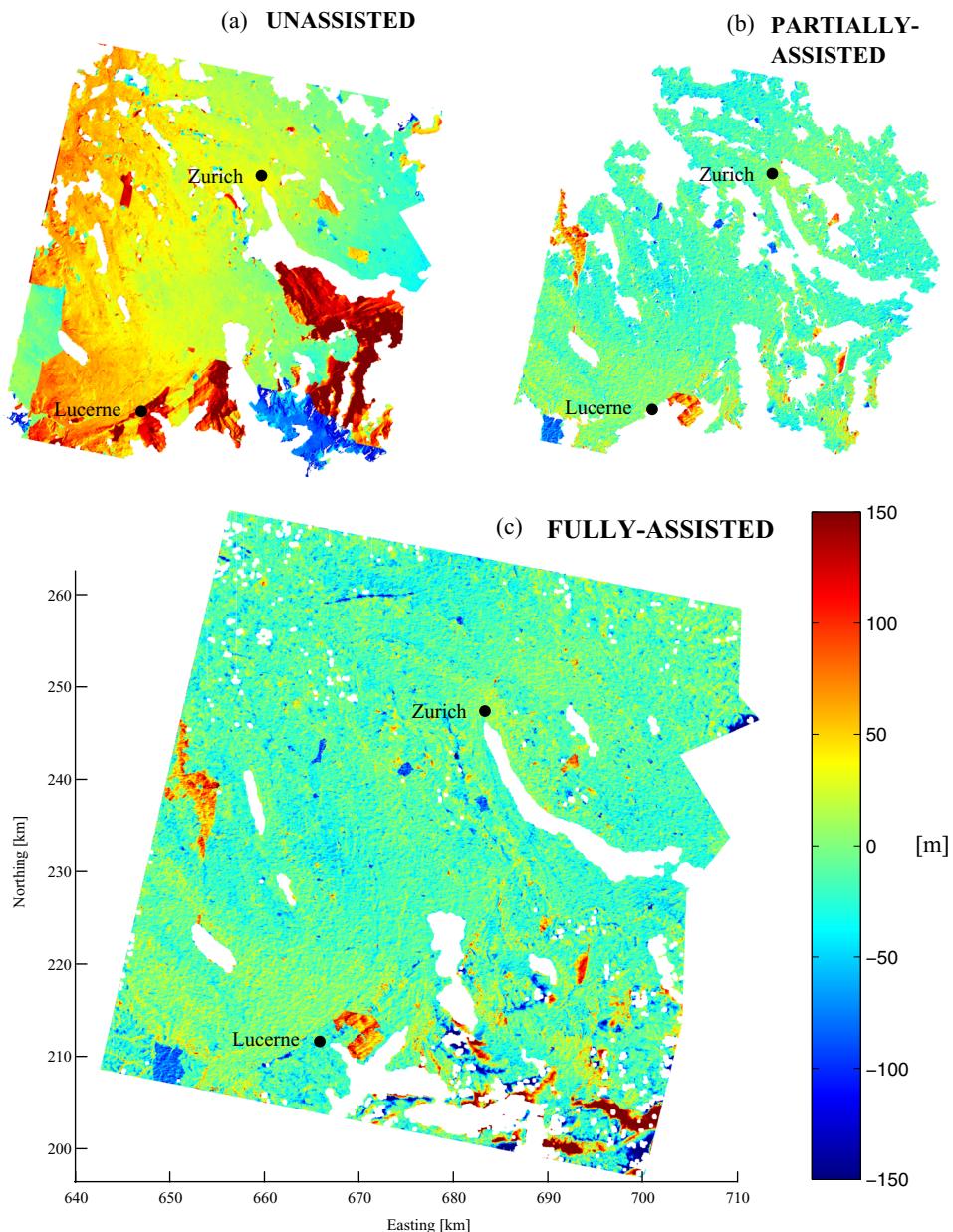


Figure 9.15 Comparison of stereo-assisted and unassisted InSAR DSMs for Zurich, relative to DHM25: (a) ellipsoid-flattened, map-calibrated (unassisted) (b) stereo-flattened and -calibrated (c) stereo-flattened, -calibrated, and "smart" hole-filled. Strong red and blue features common to (b) and (c) are unwrapping errors.

comparing the results obtained from the unassisted and fully stereo-assisted techniques.

Two major sources of error are immediately obvious in the error map for the DSM generated without the aid of the stereo DSM, illustrated in Figure 9.15(a). First, there are several areas where the phase-unwrapping algorithm, the branch-cut algorithm described in [36], incorrectly integrated the wrapped phase. The largest can be seen in the south-western and southern sectors of the image, as well as along the south shore of the lake of Zurich. The second gross error is the trend in the diagonal direction, from the south-west to the north-east: the height errors vary from consistent underestimation through consistent overestimation from corner to corner. This type of effect, to some extent also visible in the Küttigkofen experiments (Figure 9.5), is typical of a problem during phase calibration: it highlights the sensitivity of the final height model to the choice of the GCPs.

The comparison is restricted to the intersection of the ERS-1/2 and ENVISAT height models, although the unassisted result extends somewhat further northwards. There are two reasons for this: first, the DHM25 does not extend past the borders of Switzerland, already limiting the extent of the validation. Second, the large errors present in the unassisted DSM do not improve in the north extension; the extract shown here suffices as a comparison for the stereo-assisted DSM.

Figure 9.15(b) is the DSM produced after stereo-based flattening and phase calibration. It is to be compared to Figure 9.15(c), which is the same DSM after stereo-based "smart" hole-filling. Several phase-unwrapping errors can be seen in (b) as dark red and dark blue areas that stand out against the homogeneous light green-blue background. Stereo-based hole-filling introduces several new errors, mainly strong overestimations, but these are limited to the mountains in the south-east corner of the DSM.

The statistics corresponding to Figure 9.15 are given in Table 9.3. Here, the advantage of the filled versus non-filled version is quantified. While the error statistics are quite similar, the coverage is 64% greater for the stereo-filled version.

It can be concluded that for flat to rolling terrain, similar spaceborne situations have much to gain by stereo-assisted processing, including phase-flattening, phase calibration, and hole-filling. Alpine areas are perhaps best left unfilled; a future version of the hole-filling algorithm could automatically determine the local terrain-slope variation and ignore the stereo data completely in those areas, thus improving further on the result shown in Figure 9.15(c).

Table 9.3 Error statistics for Zurich InSAR experiments ("best" case highlighted)

Experiment	Mean height error [m]	Std Dev [m]	Pixels with < 20 m error [%]	Pixels with < 40 m error [%]	Pixels with < 100 m error [%]	Pixels with < 200 m error [%]	Total valid area [km ²]
Unassisted DSM	36.4	60.1	27.3	54.0	88.5	98.1	2731
Stereo-assisted DSM, holes not filled	-10.7	26.3	55.3	87.6	99.4	100.0	1954
Stereo-assisted DSM, holes filled	-10.8	35.8	54.6	85.6	98.3	99.5	3213

Chapter 10

Conclusions and Recommendations

It has been shown that initial stereo processing can improve the DSM obtainable by conventional interferometry in three ways: during phase flattening, during phase calibration, and as a means to DSM hole-filling. The usefulness of stereo depends on the sensor geometry and other acquisition parameters. In general, the more difficult the case for interferometry based on phase-unwrapping and -calibration, the greater the potential for an improved result using the stereo-assisted technique.

The conclusions resulting from the experiments using stereo processing alone as well as stereo-assisted InSAR are summarized here, and suggestions for improvements on the techniques are outlined.

10.1 Stereo SAR

10.1.1 Rough Coregistration

In order to reduce the gross differences between the two images comprising the stereo pair, a rough co-registration of the slave to the master is performed, based on the assumption of a topography-free geometry (Figure 3.1 on page 54). This involves compressing the slave in the range direction to roughly line up with the master. Possibly a better method would be to resample the master into slave geometry, or equivalently, define the master as the one whose stereo window is the widest. This would prevent information loss caused by the current approach.

More importantly, the current rough co-registration assumes roughly parallel flight tracks for the two sensors, such that only a range resampling is required to bring one image roughly into the other's native geometry. This excludes crossing-track stereo, with flight path intersections deviating from parallel by an amount dependent on the robustness of the matching algorithm. In order to extend this resampling step to such geometries, the move from a 1-D to a 2-D resampling algorithm must be made. Performing a forwards-

followed by a backward-geocoding from the slave to the master geometry would accomplish this task, taking an ellipsoid-Earth height model as a replacement for the simple 1-D flat-Earth range model.

The current algorithm permits the possibility of manually specifying an initial global rotation/scaling/translation operation, thus bringing the images into rough alignment before the automated matching takes over. However, the internal geometric differences due to viewpoints rotated with respect to one another will, for some critical path-crossing angle, make matching impossible. Opposite-side stereo was seen to be difficult to impossible with this method, precisely because of the high degree of internal dissimilarities such as differently-placed shadows and reversed layover. Radar images taken from different viewpoints become decorrelated more quickly than optical pairs obtained from the same viewpoints.

10.1.2 Matching

The core of any stereo processor is the matching algorithm. A wavelet-based algorithm was chosen for its multiresolution and phase-based properties, in the hope of overcoming the particular difficulties posed by the SAR matching problem. Specifically, the multiresolution framework was hoped to provide the desired speckle insensitivity, while the use of a combined shape- and area-based algorithm was chosen as a solution for the distortions caused by the range geometry of a SAR system.

It was confirmed that the errors that occur during stereo matching are the primary source of final height error, given state vector knowledge of nominal accuracy.

Matching errors occur when the differences between the two images are too great. With the right combination of incidence angles, sub-pixel accuracy is common, and errors are mainly confined to areas with strong radar shadow and layover. In spite of the multiresolution framework and integrated error-correction mechanism, strong errors were observed to have a significant detrimental effect on nearby surroundings. Radar shadows are a particularly difficult problem, because even a perfect match between shadows will result in erroneous derived heights; this is a consequence of the range geometry, and cannot as such be avoided.

It was discovered that the match confidence, a product of the stereo processor, was correlated to DSM height error. That is, the highest stereo-DSM errors were usually associated with low match confidences. However, the converse is not necessarily true: not all low-confidence pixels have heights with large errors. The strongest trend, exploited in this work, was the relatively strong correlation between high match confidence and accurate derived

heights.

The designs of the automatic GCP-selection and "smart" hole-filling algorithms depend on the elimination of heights associated with confidence values below a given threshold. Other selection criteria were pixel brightness in the case of the hole-filling, and additional local phase stability in the case of automated GCP selection. These algorithms were seen to be quite effective in eliminating the stereo heights with the greatest errors.

It was shown how the combined error-correction-and-smoothing step could be calibrated according to the type of terrain: flat to rolling through to mountainous. A current limitation arises when the terrain types vary greatly across the scene, which would be better served by a variable weighting between match accuracy and DEM smoothness. A future version of the matching algorithm could perform this calibration dynamically, adjusting the amount of smoothing according to the local variability of the pixel disparities, which is directly related to the terrain variability. This would eliminate the need to compromise when matching images containing both steep terrain features as well as flat terrain.

Another, more fundamental improvement to the matching algorithm is possible; it involves the transmission of information from one resolution level to the next-finest level. In the current implementation, the sub-pixel disparities are transmitted to the next level and used as starting guesses for the next round. However, the wavelet coefficients, which describe the match confidence at a given level, are not transmitted. At each new level, the similarity distance surface is calculated anew, resulting in near-zero match confidences for the interiors of bounded, featureless patches such as water bodies. A future implementation of the algorithm should keep a cumulative tally of the wavelet coefficients for all previous levels, while giving the highest weight to the confidences at the current level. This information retention would make it possible to assign high-confidence matches to featureless areas, as long as these areas are near strong features. In a sense, the match confidences obtained at a given level would be inherited at the next-finest level, lacking further strong matches. For example, a lake is typically assigned a low match confidence, since it is dark and featureless in a radar image. Once the algorithm begins matching the lake interior, it fails, oblivious to the lake perimeter. However, the shores of the lake are confidently matched. If this confidence were transmitted to the central regions of the lake, where no new strong features are to be found, then even the feature-free water bodies would be matched confidently; this resembles how the human visual system interprets such cases. Featureless terrain in SAR images often lacks significant topographical variation; this implies that its height will often correspond to the

height of the surrounding features.

10.1.3 Radar Shadow and Layover

As seen and described in section 8.2.2.1 on page 109, radar shadow will cause stereo-based height overestimation roughly equal to the height of the object casting the shadow. The shadows themselves generate extremely low match confidences, which permits the elimination of the corresponding heights from the automatic GCP generation and InSAR hole-filling. By performing a match-confidence-based fusion of the results from one or more additional stereo datasets of the same scene, taken from different view angles, it should be possible to greatly reduce the extent of such shadow-related errors.

Layover is a less severe problem in most cases, but in one way more dangerous: its high amplitude creates high match confidence, making it difficult for an automatic GCP-collection algorithm to ignore these points. A possible solution may be to implement a mechanism which would avoid the zones in front of raised objects.

With respect to viewing geometry, the ideal situation was determined to be one in which neither radar shadow nor layover were excessive. For cases with sharp changes in the terrain, such as forest stands, same-side stereo with equal incidence angles of $\sim 45^\circ$ is preferred by the matching algorithm. If only flat or rolling terrain is present, without large terrain discontinuities, the stereo processor is able to reproduce the terrain topography with very high accuracy even under more extreme viewing conditions.

10.2 Stereo-Assisted Interferometry

10.2.1 Choice of Approach

In [33], the authors perform a pixel-by-pixel fusion of stereo- and InSAR-derived DEMs using the interferometric coherence as a measure of the InSAR-derived height reliability. In the work performed here, the areas of low coherence are not even unwrapped, much less converted into height values. This is due to the algorithm's inability to deal with low-coherence phase values, since it uses a pixel-by-pixel approach rather than attempting to find a global phase model that fits with the data.

Also, the results obtained show no clear link between medium-to-high height accuracy and medium-to-high phase coherence; this can be seen by comparing Figures 9.12 (page 150) and 9.15(b) (page 153), for example. Thus,

the type of method proposed in [33] seems unsuitable for the work performed here. Rather, a consistently-higher reliability of the InSAR heights was observed - given that InSAR heights were derived at all. The current choice of approach is therefore justifiable: It attempts to combine the strengths of the stereo processor with the weaknesses of the InSAR processor without compromising the strengths of the latter method.

The Küttigkofen experiment demonstrated the full potential of the stereo-assisted InSAR technique. The result combining high accuracy with maximum coverage was the one that had been stereo phase-flattened, automatically calibrated based on GCPs extracted from the stereo DSM, and whose data holes had been filled through a combination of careful fusion with the stereo data and cubic interpolation along a Delaunay network. Not only was the final DSM a higher-quality product than the rest, but the whole processing chain is fully automatable, unlike traditional InSAR processors.

The fully stereo-assisted result for the spaceborne Zurich site was also extremely successful, when compared to the result obtained without any assistance. Some phase-unwrapping errors remained in the end product, while few strong errors were introduced directly by the stereo DSM itself. Given a more advanced phase-unwrapping algorithm than the branch-cut method used, the stereo-assisted result would have contained very few errors; it already covers nearly the entire stereo-InSAR overlap region with hardly any data holes.

10.2.2 Choice of Baseline

The use of a short a baseline for interferometric height extraction, such as was the case for Emmen, is not recommended as the primary means for high-quality DSM estimation. The 0.87 m baseline permits the reconstruction of the vegetation, when used in a single-pass configuration, but at a cost to overall DSM quality. When conditions permit, it would be desirable to use the phase arising from the small baseline to resolve the phase ambiguity while processing the larger baseline. This is possible because they record the same phase changes, but with different sensitivities; the less sensitive interferometer can be used to calibrate the phase of the other.

10.2.3 Phase Unwrapping

While offering a high precision, the 1.89 m normal baseline of the Küttigkofen configuration also made unwrapping of the forested areas impossible with the available software. Currently, an alternative phase-unwrapping algorithm based on a statistical cost-flow model [13] is being

tested, and the results look promising. Nonetheless, it will certainly also benefit from stereo-based assistance, both over difficult terrain (phase flattening; section 5.2 on page 73) and in the context of an automatic processor (for example, the GCP extraction; section 5.3 on page 73). Finally, low phase coherence will hinder accurate height extraction in some areas irrespective of the unwrapping technique, leaving open the possibility of fusion with a stereo DSM.

There is another way of increasing the chances of success of a given phase-unwrapping algorithm: improved phase filtering. Currently, a Fourier-based filter is used to adaptively smooth the phase before unwrapping [78]. Wavelets are used during stereo matching already; another possible application for wavelets is in the phase filtering. Wavelets are particularly adapted to the filtering of noise at a wide range of frequencies, without damaging the signal properties. A wavelet-based phase filter for interferometry is described in [52].

10.2.4 Automatic Ground Control Point Selection

The automated selection of GCPs to be used for phase calibration depends on three separate criteria in the current implementation:

- The stereo match confidence must be above a pre-determined threshold, thus eliminating stereo heights corresponding to low match confidences.
- The local mean SAR amplitude must exceed a pre-determined threshold, which eliminates stereo heights stemming from regions in shadow, water, or otherwise regions with low backscatter.
- The local phase variation must lie underneath a pre-determined threshold, such that a local average will represent the true local phase during phase calibration.

An additional requirement could be considered, not imposed on the values themselves, but on the point distribution. For a given distribution, a measure of the overall clustering of the GCPs can be calculated. The greater a candidate GCP increases clustering, the greater the pressure not to select it would be. The goal would be to obtain a list of GCP positions with as regular as possible a distribution, without overly-deviating from the three primary requirements. This may reduce the occurrences of error bias in the final height maps.

10.2.5 Interpolation of Large Holes

The current implementation of the final hole-filling step depends on a user-defined maximum hole diameter. Holes smaller than this threshold will be filled using cubic interpolation, while larger holes will be left unfilled. Since

the choice of threshold should reasonably depend on the terrain variability, a future implementation of the hole-filling algorithm should automatically adjust the hole-size threshold as a function of the terrain surrounding a given hole. In this way, a hole of a given size may be filled if it lies within smoothly-varying terrain, but not if the terrain varies quickly.

10.2.6 Robustness of Stereo-Assisted InSAR

In section 8.2.3 on page 114 it was seen how an existing DEM could be used to simulate SAR amplitude images, and how these could be used to test the robustness of the stereo processor for a variety of acquisition geometries. As a possible extension to this concept, one can simulate the interferometric phase for the same scene, obtained either in single- or repeat-pass mode. Then, combined stereo-InSAR processing could proceed according to the methods described in this work, using the stereo-generated DSM and match confidence as inputs to the InSAR chain.

Note, however, that the stereo configuration is completely independent of the InSAR configuration. Thus, one would ideally choose a stereo geometry corresponding to favorable conditions for the stereo processor, generate the DSM and match confidence, and then alter InSAR parameters such as baseline and incidence angle in a series of experiments similar to those performed in section 8.2.3 on page 114. In this way, the optimal *combined* configuration for a successful stereo-InSAR flight campaign could be planned in advance.

10.2.7 Final Remarks and Outlook

When no reliable DEM is already available for a given site to be InSAR-processed, initial stereo processing with the wavelet-based matching algorithm presented here has been shown to improve the InSAR processing in up to three different ways:

- **Phase flattening:** in extreme cases of alpine topography, InSAR is impossible without the stereo DSM to aid with the ambiguity resolution. If a flat-to-rolling scene is being unwrapped, ellipsoid-based flattening should suffice.
- The stereo DSM and match confidence generated by the stereo processor can be used to **automate the calibration of the combined phase-baseline model** before the phases are converted to heights. This avoids manual intervention such as placing reflectors into the scene to be imaged or searching for tiepoints in the images and in maps.
- Any **holes remaining in the InSAR DEM can often be filled** with stereo-DSM heights of medium-to-high quality, since a means to

filtering out the worst heights has been demonstrated, based on the stereo match confidence.

The technique has been shown to improve the results attainable using InSAR alone, for all three test sites, two airborne and one spaceborne. This is a strong indication of the extrapolability of the technique to other sites.

Furthermore, the general principles described should be applicable to a number of other processing configurations and algorithms. First, phase flattening is possible with any available height model, irrespective of the technique used to obtain it. Secondly, the automatic GCP selection and hole-filling depend only on the existence of some measure of the quality of the stereo disparity field. This should be available in coherence-based matching algorithms, as well as those that depend on the minimization of other functionals similar to the similarity distance that was described in this work.

Besides the near- to mid-term changes proposed in the above sections, a longer-term improvement to the presented method could involve the minimization of DSM errors due to radar shadow and layover. Such areas presented a problem during this thesis mainly because only the minimum number of acquisitions required to do stereo-assisted InSAR were available. Thus, image points within problems zones were either excluded from the final DSMs or led to strong localized errors, slightly damaging the final DSM quality. With acquisitions taken from several directions such as is described in [7] for example, all terrain features would be imaged from at least one direction, potentially eliminating the errors due to shadow. Moreover, the different range geometries would also aid in the resolution of layover-related errors. These improvements would require adapting both the stereo and InSAR processors to handle multiple input datasets. A more complete DSM would be the result of the stereo processor; this would then be integrated into the InSAR processing as before, with the multiple views being subsequently InSAR-processed. The actual fusion of the two methods would remain unchanged.

One other important factor may influence the future direction of the work presented here: low-resolution DEMs will become increasingly available and affordable, for example, the 30 m DEM produced by the Shuttle Radar Topography Mission (SRTM), described in [5]. Such DEMs may often have problems such as data holes or insufficient representation of surface features that will continue to inhibit their effective combination with InSAR. It is therefore not expected that the utility of the stereo-assisted method presented here will disappear in the near- to mid-term future.

Appendix A: List of Symbols

Symbol	Parameter	Units
a	Azimuth (slow) time	s
$A^*(-t, -a)$	Reference signal used during azimuth compression as a function of range time and azimuth	
$a_{0, 1}$	Complex constants	
B	Baseline	m
B_{\perp}	Normal baseline	
$c_{1, 2}$	Complex signal returns for the master and slave, respectively	
d	diagonal	
D_p	Complex coefficient generated by the DDWT for pixel p	
$E(t)$	Unprocessed radar-return as a function of range time	
f_d	Doppler frequency	Hz
$g(n)$	Low-pass FIR filter as a function of filter index	
h	horizontal	
$h(n)$	High-pass FIR filter as a function of filter index	
Δh	Expected height error	m
$\Delta h_{2\pi}$	Ambiguity height	m
H	Sensor flight height	m
i	Level of resolution	
I	Interferogram	
I_{pixel}	Pixel intensity	
$I(t, a)$	Azimuth-compressed image signal as a function of range time and azimuth	
k	Integer	
L	Azimuth length of real antenna	m
L_{syn}	Azimuth length of synthetic array	m
M	Master image	
N	Number of azimuth looks	
n	Integral index	
p, p'	Pixel location in the master and slave images, respectively	
\hat{p}	Ground point position in a Cartesian reference system	m
\hat{r}	Slant range between ground point and sensor in a Cartesian reference system	m

Appendix A

Symbol	Parameter	Units
$r_{1, 2}$	Slant range from a given image point to master (1) or slave (2) antenna	m
r_m	Slant range for master antenna	m
r_s	Slant range for slave antenna	m
Δr	Path length difference between two sensors	m
R	Slant range	m
$R(t)$	1-D range-compressed signal for a single burst as a function of range time	
$R(t, a)$	Range-compressed image signal as a function of range time and azimuth	
s	Wavelet scale	
S	Slave image	
\hat{s}	Antenna position in a Cartesian reference system	m
$SD^i(p, p')$	Similarity distance at resolution level i between master pixel p and slave pixel p'	
$s(t)$	Chirp signal as a function of time	
t	Slant range time	s
v	vertical	
$v_{0, 1}$	Real-valued, low-pass window functions for low- and high-pass FIR filters, respectively	
δ_r	Range disparity	pixels
$\Delta\delta$	Measured disparity error	m
γ	Backscatter coefficient	
\hat{v}	Sensor radial velocity relative to a ground point	m/s
θ_i	Beam incidence angle (relative to local normal)	degrees
$\theta_i^{M, S}$	Beam incidence angle for master (M) or slave (S)	degrees
θ_e	Beam elevation angle (relative to nadir)	degrees
λ	Antenna transmission wavelength	m
ρ_{real}	Azimuth resolution of a SLAR system	m
ρ_{SAR}	Azimuth resolution of a SAR system	m
σ^0	Normalized radar cross-section	dB
$\sigma_{0, 1}$	Standard deviation for low- and high-pass FIR filters, respectively	
τ	Wavelet translation	
γ	Coherence between two SLCs	
ϕ	Interferometric phase	radians
ϕ_{abs}	Absolute interferometric phase	radians
$\phi(i)$	Wavelet scaling function as a function of pixel index i	

Symbol	Parameter	Units
$\phi(i, j)$	Wavelet scaling function as a function of pixel indices i and j (vertical, horizontal)	
$\psi(x)$	Wavelet function as a function of independent variable x	
$\psi(i)$	Wavelet function as a function of pixel index i	
$\psi_h(i, j)$	Horizontal wavelet function as a function of pixel indices i and j	
$\psi_v(i, j)$	Vertical wavelet function as a function of pixel indices i and j	
$\psi_d(i, j)$	Diagonal wavelet function as a function of pixel indices i and j	
$\omega_{0, 1}$	Angular frequencies of low- and high-pass FIR filters, respectively	radians/s
Ω	Stereo matching scaling factor	

Appendix B: Acronyms and Abbreviations

1-D	One-dimensional
2-D	Two-dimensional
3-D	Three-dimensional
AeS	AeroSensing Radarsysteme GmbH
ASAR	Advanced Synthetic Aperture Radar
CHIRP	Coherent Integration of Radar Pulses
CWT	Continuous Wavelet Transform
D-GPS	Differential Global Positioning System
DDWT	Dyadic Discrete Wavelet Transform
DEM	Digital Elevation Model
DFT	Discrete Fourier Transform
DHM	Digital Height Model
DHM25	Digital Height Model (25 m grid spacing)
DoSAR	Dornier Synthetic Aperture Radar
DSM	Digital Surface Model
DSM25	Digital Surface Model (25 m grid spacing)
DTM	Digital Terrain Model
ENVISAT	Environmental Satellite
ERS	European Remote-Sensing Satellite
ESA	European Space Agency
FIR	Finite-Impulse Response
FT	Fourier Transform
GCP	Ground Control Point
GPS	Global Positioning System
HL, LH, LL, HH	output of <u>high</u> or <u>low</u> -pass filters applied in sequence
HRS	High-Resolution Stereoscopic instrument (SPOT 5 instrument)
IMU	Inertial Measurement Unit
InSAR	Interferometric Synthetic Aperture Radar
JPL	Jet Propulsion Laboratory
M-K	Magarey-Kingsbury
MATLAB	Matrix Laboratory
MRA	Multiresolution Analysis
NASA	National Aeronautics and Space Administration
PRF	Pulse Repetition Frequency
PRI	Pulse Repetition Interval
RADARSAT	Radar Satellite

RSL	Remote Sensing Laboratories
SAR	Synthetic Aperture Radar
ScanSAR	Scanning Synthetic Aperture Radar
SD	Similarity Distance
SFS	Shape-From-Shading
SLAR	Side-Looking Airborne Radar
SLC	Single-Look Complex
SNR	Signal-to-Noise Ratio
SPOT	Satellite Pour l'Observation de la Terre
SRTM	Shuttle Radar Topography Mission
STFT	Short-Time Fourier Transform
WGS84	World Geodetic System, 1984
WT	Wavelet Transform

References

- [1] Anandan P., *A computational framework and an algorithm for the measurement of visual motion*, Intern. Jour. Computer Vision, Vol. 2, pp. 283-310, 1989.
- [2] Aurenhammer F., *Voronoi diagrams - A survey of a fundamental geometric data structure*, ACM Computing Surveys, Vol. 23, pp. 345-405, 1991.
- [3] Baltsavias E.P., *Airborne Laser Scanning: Existing Systems and Firms and other Resources*, ISPRS Jour. Photogr. Rem. Sens., Vol. 54, Issues 2-3, 1999.
- [4] Bamler R., *A Comparison of Range-Doppler and Wavenumber Domain SAR Focusing Algorithms*, IEEE Trans. Geos. Rem. Sens., Vol. 30, No. 4, pp. 706-713, 1992.
- [5] Bamler, R., *The SRTM Mission: A World-Wide 30m Resolution DEM from SAR Interferometry in 11 Days*, Proc. 47th Photogram. Week, Stuttgart, 1999.
- [6] Bertalmio M., Sapiro G., Caselles V. and Ballester C., *Image Inpainting*, Proc. SIGGRAPH'00, pp. 417-424, 2000.
- [7] Bolter R. and Leberl F., *Detection and Reconstruction of Buildings from Multiple View Interferometric SAR Data*, Proc. IGARSS'00, pp. 749-751, 2000.
- [8] Burkhardt G., Bergen Z. and Carande R., *Elevation Correction and Building Extraction from Interferometric SAR Imagery*, Proc. IGARSS'96, pp. 659-661, 1996.
- [9] Burrus C.S., Gopinath R.A., Guo H., *Introduction to Wavelets and Wavelet Transforms*, Prentice Hall, U.S.A., 1997.
- [10] Cafforio C., Prati C. and Rocca F., *SAR Data Focusing Using Seismic Migration Techniques* IEEE Trans. Aero. Elec. Sys., Vol. 27, No. 2, pp. 194-207, 1991.

- [11] Caselles V., Morel J.-M. and Sbert C., *An axiomatic approach to image interpolation*, IEEE Trans. Im. Proc., Vol. 7, No. 3, pp. 376-386, 1998.
- [12] Chang C.Y., Jin M. and Curlander J.C., *SAR Processing Based on the Exact Two-Dimensional Transfer Function*, IGARSS'92, pp 355-359, 1992.
- [13] Chen C.W., *Statistical-Cost Network-Flow Approaches to Two-Dimensional Phase-Unwrapping for Radar Interferometry*, Doctoral Dissertation, Stanford University, U.S.A., 2001.
- [14] Chen C.W. and Zebker H.A., *Phase unwrapping for large SAR interferograms: Statistical segmentation and generalized network models*, IEEE Trans. Geosc. Rem. Sens., Vol. 40, pp. 1709-1719, 2002.
- [15] Crosetto M. and Aragues F.P., *Radargrammetry and SAR Interferometry for DEM Generation: Validation and Data Fusion*, Proc. CEOS SAR Workshop, France, 1999.
- [16] Cumming I.G., Neo Y.L. and Wong F.H., *Interpretations of the Omega-K Algorithm and Comparisons with other Algorithms*, Proc. IGARSS'03, pp. 1455-1458, 2003.
- [17] Curlander John C. and McDonough Robert N., *Synthetic Aperture Radar: Systems & Signal Processing*, John Wiley & Sons, Inc., U.S.A., 1991.
- [18] Daubechies, I., *Orthonormal Bases of Compactly Supported Wavelets*, Comm. Pure Appl. Math., Vol. 41, pp. 909-996, 1988.
- [19] Davidson G.W. and Bamler R., *Multiresolution phase unwrapping for SAR interferometry*, IEEE Trans. Geosci. Rem. Sens., Vol. 37, No. 1, pp. 163-174, 1999.
- [20] Dick A., *Multiresolution Stereo Image Matching using Wavelets*, Honours project thesis, University of Adelaide, Australia, 1997.
- [21] Elachi C., *Spaceborne Radar Remote Sensing: Applications and Techniques*, The Institute of Electrical and Electronics Engineers, Inc., U.S.A., 1988.
- [22] Elizavetin I.V., *Merging of Stereo and Interferometric Processing for Improved DEM Generation*, Proc. IGARSS'99, pp.479-481, 1999.

- [23] European Space Agency - ASAR Science Advisory Group, *ENVISAT ASAR: Science and Applications*, ESA Publications Division, The Netherlands, 1998.
- [24] European Space Agency - ESRIN, *TISAR: DEM / ORI Production and Service for the ERS & ENVISAT ASAR System, ESRIN Contract Number 15722, Final Report, Version 6.2*, Italy, 2004.
- [25] European Space Agency, *The European Remote Sensing Satellite: ERS-1 Data Book: A Summary of the Technical Elements of the ERS-1 Spacecraft and its Payload*, ESA Publications Division, The Netherlands, 1991.
- [26] Faller N.P. and Meier E.H., *First Results with the Airborne Single-Pass DO-SAR Interferometer*, IEEE Trans. Geosc. Rem. Sens., Vol 33, No. 5, pp. 1230-1237, 1995.
- [27] Frankot R.T. and Chellappa R., *Application of a Shape-from-Shading Technique to SAR Imagery*, Proc. IGARSS'87, pp. 1323-1329, 1987.
- [28] Frost V.S., Stiltes J.A., Shanmugan K.S. and Holtzman J.C, *A model for radar images and its application to adaptive digital filtering of multiplicative noise*, IEEE Trans. Patt. Anal. Mach. Intell., Vol. 4, No. 24, pp. 157-166, 1982.
- [29] Gabriel A.K. and Goldstein R.M., *Crossed orbit interferometry: Theory and experimental results from SIR-B*, Int. J. Rem. Sens., Vol. 9, No. 5, pp. 857-872, 1988.
- [30] Gabriel A.K., Goldstein R.M. and Zebker H.A., *Mapping small elevation changes over large areas: Differential radar interferometry*, J. Geophys. Research, Vol. 94, No. B7, pp. 9183-9191, 1989.
- [31] Gamba P., Houshmand B., and Saccani M., *Detection and Extraction of Buildings from Interferometric SAR Data*, IEEE Trans. Geosc. Rem. Sens., Vol. 38, No. 1, pp. 611-618, 2000.
- [32] Gatelli F., Guarneri A.M., Parizzi F., Pasquali P., Prati C. and Rocca F., *The Wavenumber Shift in SAR Interferometry*, IEEE Trans. Geosc. Rem. Sens., Vol. 32, No. 4, pp. 855-865, 1994.
- [33] Gelautz M., Paillou P., Chen C., and Zebker H., *Radar Stereo- and Interferometry-Derived Digital Elevation Models: Comparison and Combination Using Radarsat and ERS-2 Imagery*, Intl. Jour. Rem. Sens., Vol. 24, No. 24, pp. 5243-5264, 2003.

- [34] Ghiglia D.C. and Pritt M.D., *Two-Dimensional Phase Unwrapping: Theory, Algorithms, and Software*, John Wiley & Sons, Inc., U.S.A., 1998.
- [35] Goblirsch W., *Optimization of Geometric Parameters for Interferometric Surface Model Generation*, Ph.D. thesis, Remote Sensing Laboratories, Geography Department, University of Zurich, Switzerland, 1997.
- [36] Goldstein R.A., Zebker H.A., and Werner C.L., *Satellite Radar Interferometry: Two-Dimensional Phase Unwrapping*, Radar Science, Vol. 23, No. 4, pp. 713-720, 1988.
- [37] Horn B.K.P. and Brooks M.J., *Shape From Shading*, MIT Press, U.S.A., 1989.
- [38] Graham L.C., *Synthetic interferometer radar for topographic mapping*, Proc. of IEEE, Vol. 62, No. 6, pp. 763-768, 1974.
- [39] Grossman A. and Morlet J., *Decomposition of Hardy functions into square integrable wavelets of constant shape*, SIAM, Jour. Mathematical Analysis, Vol. 15, pp. 723-736, 1984.
- [40] Hervet E., *Appariement d'images Radar à Synthèse d'Ouverture en conditions radargrammétiques*, Ph.D. thesis, Institut National Polytechnique de Toulouse, France, 2000.
- [41] Holecz F., Pasquali P., Moreira J., Meier E. and Nüesch D., *Automatic generation and quality assessment of digital surface models generated from AeS-1 InSAR data*. Proc. EUSAR'98, pp. 57-60, 1998.
- [42] Hubbard B.B., *Ondes et Ondelettes: Le Saga d'un Outil Mathématique*, Pour La Science, France, 1995.
- [43] <http://www.intermaptechnologies.com>
- [44] <http://www.toposys.com>
- [45] Jacobsen K., *Analysis of SPOT HRS Stereo Data*, High Resolution Mapping from Space 2003 (Joint Workshop, Universität Hannover), Germany, 2003.
- [46] Jehle M., Frey O., Meier E., Small D. and Nüesch D., *Improved Knowledge of SAR Geometry through Atmospheric Modelling*, Proc. EUSAR'04, pp. 909-911, 2004.

- [47] Kaplan E.D., *Understanding GPS: Principles and Applications*, Artech House, U.S.A., 1996.
- [48] Kirsch M. and Rinke M., *3D Reconstruction of Buildings and Vegetation from Synthetic Aperture Radar (SAR) Images*, IAPR Workshop MVA'98, pp. 228-231, 1998.
- [49] Körner T.W., *Fourier Analysis*, Cambridge University Press, U.K., 1988.
- [50] Leberl F.W., *Radargrammetric image processing*, Artech House, Inc., 1990.
- [51] Li F.K. and Goldstein R.M., *Studies of multibaseline spaceborne interferometric synthetic aperture radars*, IEEE Trans. Geosc. Rem. Sens., Vol. 28, No. 1, pp. 86-96, 1990.
- [52] Lopez-Martinez C. and Fabregas X., *Modeling and reduction of SAR interferometric phase noise in the wavelet domain*, IEEE Trans. Geosc. Rem. Sens., Vol. 40, No. 12, pp. 2553-2566, 2002.
- [53] Maas H.-G., *Planimetric and height accuracy of airborne laserscanner data: User requirements and system performance*, Proc. 49. Photogrammetric Week, Wichmann Verlag, pp. 117-125, 2003.
- [54] Magarey J., *Motion estimation using complex wavelets*, Ph.D. thesis, University of Cambridge, U.K., 1997.
- [55] Magarey J. and Dick A., *Multiresolution stereo image matching using complex wavelets*, Proc. Intern. Conf. Patt. Recog., No. 14, pp. 4-7, 1998.
- [56] Magarey J. and Kingsbury N., *Motion Estimation Using a Complex-Valued Wavelet Transform*, IEEE Trans. Sig. Proc., Vol. 46 No.4, pp. 1069-1084, 1998.
- [57] Mallat S.G., *A Theory for Multiresolution Signal Decomposition: A Wavelet Representation*, IEEE Trans. Patt. Anal. Mach. Intell., Vol 11, No. 7, 1989.
- [58] Mallat S.G., *A Wavelet Tour of Signal Processing*, 2nd ed., Academic Press, U.S.A., 1999.
- [59] Mallat S.G., *Multiresolution approximation and wavelets*, Trans. Amer. Math. Soc., Vol 315, pp.69-88, 1989.

- [60] Martinoni D., *Models and experiments for quality handling in digital terrain modelling* (PhD. Thesis), GIS Division, Department of Geography, University of Zurich, Switzerland, *Appendix D: An Introduction to Wavelets*, pp. 272-273, 2002.
- [61] <http://www.mathworks.com>
- [62] Marti, U., *The New Geoid CHGEO98 of Switzerland*, Federal Office of Topography, Switzerland, 1998.
- [63] Maune D.F., *Digital Elevation Model Technologies and Applications: The DEM Users Manual*, American Society for Photogrammetry and Remote Sensing, U.S.A., 2001.
- [64] Moreira J., Schwäbisch M., Wimmer C., Rombach M. and Mura J., *Surface and Ground Topography Determination in Tropical Rainforest Areas Using Airborne Interferometric SAR*, Photogrammetrische Woche 2001, Fritsch & Spiller (Eds.), Wichmann Verlag, Germany, 2001.
- [65] Paillou P. and Gelautz M., *Relief Reconstruction from SAR Stereo Pairs: The "Optimal Gradient" Matching Method*, IEEE Trans. Geosc. Rem. Sens., Vol. 37, No. 4, pp. 2099-2107, 1999.
- [66] Pan H-P., *General stereo image matching using symmetric complex wavelets*, Proc. SPIE IV, Vol. 2825, pp. 697-721, 1996.
- [67] Pan H-P. and Magarey J., *Phase-Based Bidirectional Stereo in Coping with Discontinuity and Occlusion*, Proc. Intl. Workshop Im. Anal. Inf. Fusion, pp. 239-250, 1997.
- [68] Paquerault S., *Restitution du relief à partir d'images radar par radar-clinométrie*, PhD thesis, Ecole Nationale Supérieur des Télécommunications, Paris, 1998.
- [69] Pérez P., Gangnet M. and Blake A., *Poisson Image Editing*, Proc. SIGGRAPH'03, Vol. 22, No. 3, pp. 313-318, 2003.
- [70] Pink T.H. and Ramanathan, U., *Extraction of height information from target shadow for applications in ATC*, Proc. IGARSS '99, Vol. 1, pp. 351-353, 1999.
- [71] Poidomani C.G., Costantini D., Pasquali P. and Jaeger P., *National-Scale DEM Generation Using ERS Tandem Data in Alpine Regions*, Proc. ERS-ENVISAT Symposium, Göteborg, Sweden, 2000.

- [72] Press W.H., Teukolsky S.A., Vetterling W.T. and Flannery B.P., *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed., Cambridge University Press, 1992.
- [73] Rabiner L.R. and Gold B., *Theory and Application of Digital Signal Processing*, Prentice-Hall, New Jersey, U.S.A., 1975.
- [74] Schubert A., Small D., Holecz F., Meier E., and Nüesch D., *Extraction of Surface Topography from SAR Stereo Pairs Using an Airborne X-Band Sensor: Preliminary Results*, Proc. LANDSAR'01, pp. 151-156, 2001.
- [75] Schubert A., Small D., Meier E., and Nüesch D., *Robustness of Wavelet-Based Stereo Matching for Variable Acquisition Geometries Using Simulated SAR Images*, Proc. IGARSS'02, Vol. 5, pp. 2759-2761, 2002.
- [76] Serra J., *Image Analysis and Mathematical Morphology*, Academic Press, U.S.A., 1982.
- [77] Simonetto E., Oriot H., and Garello R., *Radargrammetric Processing for 3-D Building Extraction from High-Resolution Airborne SAR Data*, Proc. IGARSS'03, Vol. 3, pp. 2002-2004, 2003.
- [78] Small D., *Generation of Digital Elevation Models through Spaceborne SAR Interferometry*, Remote Sensing Series, Vol. 30, Dept. of Geography, University of Zurich, 1998.
- [79] Small D., Meier E. and Nüesch D., *Phase Noise Countermeasures for Synthetic Interferogram Generation*, Proc. IGARSS'02, Vol. 4, pp. 2546-2548, 2002.
- [80] Small D., Pasquali P. and Füglistaler S., *A Comparison of Phase to Height Conversion Methods for SAR Interferometry*, Proc. IGARSS '96, pp. 342-344, 1996.
- [81] Small D., Werner C., and Nüesch D., *Registration of ERS-1 SLC Products for Interferometry*, Proc. Fourth GEOSAR Workshop, Loipersdorf, Austria, May 26-28, 1993.
- [82] Stebler O. and Nüesch D., *Einführung in die Radar-Fernerkundung*, Geog. Inst. der Universität Zürich-Irchel, Vorlesung Fernerkundung III (Block 12) von Prof. Dr. D. Nüesch, 2002.
- [83] Stimson G.W., *Introduction to Airborne Radar*, 2nd ed., SciTech Publishing, Inc., U.S.A., 1998.

- [84] Swiss Federal Office of Topography (SwissTopo, see <http://www.swisstopo.com>), *Digital Height Model DHM25: Product Information*, product description booklet, November 1993.
- [85] Toutin T., *Accuracy assessment of stereo-extracted data from airborne SAR images*, Int. J. Rem. Sens., Vol. 18, No. 18, pp. 3693-3707, 1997.
- [86] Toutin T., *Radar Stereo Pairs for DEM Generation: RADARSAT for Stereoscopy*, Geomatics Info Magazine Intl., Vol. 13, No. 1, pp. 6-9, 1999.
- [87] Toutin T. and Gray L., *State-of-the-art of elevation extraction from satellite SAR data*, ISPRS Jour. Photogramm. Rem. Sens., No. 55, pp. 13-33, 2000.
- [88] Ulaby F.T., Moore R.K. and Fung, A., *Microwave Remote Sensing: Active and Passive*, Volume I, Artech House Inc., U.S.A., 1981.
- [89] Ulaby F.T., Moore R.K. and Fung, A., *Microwave Remote Sensing: Active and Passive*, Volume II, Artech House Inc., U.S.A., 1981.
- [90] Vidal-Pantaleoni A., Oviol R. and Ferrando M., *A Comparison of Phase Unwrapping Techniques in Synthetic Aperture Radar Interferometry*, Proc. IGARSS'99, 1999.
- [91] Werner C.L., Wegmüller U. and Strozzi T., *Processing Strategies for Phase Unwrapping for InSAR Applications*, Proc. EUSAR'02, pp. 353-356, Germany, 2002.
- [92] Wildey R. L., *Topography from Single Radar Images - Venus Radar Mapper*, Science, Vol. 224, No. 4645, pp. 153-156, 1984.
- [93] Zebker H.A. and Goldstein R.M, *Topographic mapping from interferometric synthetic aperture radar observations*, J. Geophys. Research, Vol. 91, No. B5, pp. 4993-4999, 1986.

Acknowledgements

Many thanks to the following companies for generously allowing me to use their height references:

- sarmap for the DSM25 extract.
- SwissTopo for the DHM25 and extracts from their topographical maps.
- TopoSys for the laser-DSM extract
- Infoterra, GmbH for the DoSAR-DSM extract

The following companies provided the SAR data used in this thesis:

- Intermap Technologies GmbH provided the AeS-1 data.
- The European Space Agency (ESA) provided the ERS-1/2 and ENVISAT data.

I would also like to thank the following people for their support:

- Prof. Dr. Klaus Itten for his ongoing dedication to the maintenance of a healthy and competitive research environment at RSL, which made this work possible as well as enjoyable.
- Daniel Nüesch for his unflagging patience and words of wisdom during my tortuous journey towards the completion of this work, as well as his detailed comments on the manuscript.
- Erich Meier for his strong personal investment in the intellectual and spiritual health of the SAR group, his never-ceasing drive to solve problems as cleanly and swiftly as possible, and his helpful suggestions.
- David Small for providing me with many of the programs used to process my SAR data, and for his eternal willingness to help me with even the most mundane of problems.
- Roland Brodbeck, Maurice Rüegg, Erich Meier and Cyril Humbert for their help translating the German and French summaries.
- My colleagues in the SAR group and RSL, who were always there to listen to my ruminations and bolster my sense of humor, whether in the evening after a long day or in the morning after a long night.
- My parents and brothers for their encouragement and proof-reading.
- Last but far from least: Iwona, for her on-site support in the evenings by way of amazing improvised meals and musical accompaniment, her open ears and heart, her loving patience, and her constant assurance that I would make it.