

DEUTSCHE GEODÄTISCHE KOMMISSION
bei der Bayerischen Akademie der Wissenschaften

Reihe C

Dissertationen

Heft Nr. 570

Gert Beyer

Wavelettransformationen hybrider Geländemodelle

München 2005

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

ISSN 0065-5325

ISBN 3 7696 5009 3

Wavelettransformationen hybrider Geländemodelle

Von der Fakultät für Forst-, Geo- und Hydrowissenschaften
der Technischen Universität Dresden
zur Erlangung der Lehrbefähigung (venia legendi)
für das Lehrgebiet "Geoinformatik"
genehmigte Habilitationsschrift

vorgelegt von

Dr.rer.nat. Gert Beyer

aus Dresden

München 2005

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

Adresse der Deutschen Geodätischen Kommission:

Deutsche Geodätische Kommission

Marstallplatz 8 • D – 80 539 München

Telefon (089) 23 031 -1113 • Telefax (089) 23 031 -1283/ -1100

E-mail hornik@dgfi.badw.de • <http://dgk.badw.de>

Prüfungskommission

Vorsitzender: Prof. Dr.habil. H.-G. Maas

Gutachter: Prof. Dr.-Ing.habil. S. Meier
Prof. Dr.-Ing.habil. K. Kraus
Prof. Dr.sc.techn. W. Keller

Tag der Einreichung: 05.06.2002

Tag des Habilitationskolloquiums: 20.12.2002

Tag der Probevorlesung: 20.12.2002

© 2005 Deutsche Geodätische Kommission, München

Alle Rechte vorbehalten. Ohne Genehmigung der Herausgeber ist es auch nicht gestattet,
die Veröffentlichung oder Teile daraus auf photomechanischem Wege (Photokopie, Mikrokopie) zu vervielfältigen

Prof. Dr.-Ing. habil. Siegfried Meier
zum 65. Geburtstag

Inhaltsverzeichnis

Verzeichnis der Tabellen	7
Verzeichnis der Abbildungen	7
1 Einleitung	9
2 Wavelet-Transformation zur Analyse, Approximation und Kompression	12
2.1 Integraltransformationen	12
2.2 Kontinuierliche Wavelettransformation	13
3 Diskrete Wavelettransformation	15
3.1 Matrizenbeziehungen diskreter Funktionen	15
3.2 Konzept der Multiskalenanalyse	16
3.3 Konstruktion der diskreten Wavelets	20
3.3.1 Konstruktion der eindimensionalen Wavelets	20
3.3.2 Konstruktion der zweidimensionalen Wavelets	22
3.4 Verfahren der diskreten Wavelettransformation	23
3.4.1 Verfahren der 1D-Wavelettransformation	23
3.4.2 Verfahren der 2D-Wavelettransformation	25
3.5 Wavelettransformation einer zweidimensionalen Polynomfunktion	27
3.5.1 Wavelettransformation einer Funktion $z = xy$	27
3.5.2 Wavelettransformation einer Polynomfunktion $z = P(x, y)$	28
3.6 Momente	29
4 Digitale Geländemodelle	32
4.1 Modelle für Kurven und Flächen	32
4.2 Neigung und Krümmung	36
4.3 Raumkurven in Geländemodellen	39
4.3.1 Explizite Beschreibung	39
4.3.2 Implizite Beschreibung	44
5 Wavelettransformation digitaler Geländemodelle	49
5.1 Wavelettransformation der Gelände­fläche	49
5.2 Wavelettransformation von Kurven in expliziter Darstellung	51
5.3 Wavelettransformation von Kurven in impliziter Darstellung	53
6 Kompression von Geländedaten	55
6.1 Vorbemerkungen	55
6.2 Kompressionsverfahren	55
6.2.1 Sperrfilter	57
6.2.2 Schwellwertverfahren	59
6.3 Kompression von Gelände­flächen	60
6.4 Kompression von Gelände­kurven	62
6.5 Probleme bei hybriden Daten	67
6.5.1 Gemeinsame Kompression von Vektor- und Rasterdaten	67
6.5.2 Glättungsmaße	67
6.5.3 Kantenanpassung	72

7	Analyse von Geländemodellen mittels Wavelettransformation (Folgeprodukte)	74
7.1	Approximationseigenschaften	74
7.1.1	Theoretischer Hintergrund	74
7.1.2	Approximationseigenschaften bei der diskreten Wavelettransformation . .	76
7.1.3	Skalierung der Waveletkoeffizienten	77
7.2	Lokalisierungseigenschaften	80
7.2.1	Eindimensionale Wavelettransformation	80
7.2.2	Zweidimensionale Wavelettransformation	81
7.3	Höhen- und Gefällelinien	85
8	Ausblick	90
8.1	Wavelettransformation raumbezogener Geoinformationen	90
8.2	Weitere Fragestellungen	91
	Literatur	93
A	Probleme bei der Signalanalyse mit der schnellen 2D-Wavelettransformation	96
B	Filtereigenschaften des Daub4-Wavelets	100
B.1	Filtercharakteristik	100
B.1.1	Filtercharakteristik in Originallage	100
B.1.2	Filtercharakteristik in Symmetrielage	102
B.2	Testrechnung zur Verifizierung der Ergebnisse	105
C	Kantenextrapolation bei gekrümmten Kanten	107
C.1	Gleichung des hyperbolischen Paraboloids	107
C.2	Schnitt zweier hyperbolischer Paraboloiden	108
D	Algorithmus zur Kantenrekonstruktion	109
E	Testrechnung zur Qualität der Lagekorrektur	113
F	Das Testgebiet Schneealpe	116

Verzeichnis der Tabellen

3.1-1	Analogien zwischen stetigen und diskreten Funktionen	16
6.3-1	Geländeausschnitte unterschiedlicher Neigungsbereiche	60
6.3-2	Wavelets für den Test der Kompressionsraten	61
6.4-1	Eigenschaften der Beispielkurven	65
6.5-1	Dominierende Frequenzen in den einzelnen Skalen	69
B-1	Schätzung der Amplitude aus den Stützwerten	106
B-2	Vergleich der Amplituden	106
E-1	Vergleich der Standardabweichungen von Signal und Differenz.	115

Verzeichnis der Abbildungen

3.2-1	Schematische Darstellung des Baumalgorithmus	18
3.3-1	Rekursive Berechnung der Skalierungsfunktionen	21
3.3-2	Rekursive Berechnung der Wavelets	22
3.4-1	Schematische Darstellung der schnellen WT	24
3.4-2	Schematische Darstellung des Baumalgorithmus	24
3.4-3	Schematische Darstellung der stationären WT	25
3.4-4	Datenstruktur der WK bei der zweidimensionalen schnellen WT	26
3.4-5	Teilgitterquadrate, die von den Haarwavelets der ersten Skale überdeckt werden.	26
3.4-6	Zyklische Fortsetzung der Signalmatrix	27
3.5-1	Schematische Darstellung des Skalarprodukts $\langle HH_{k,l}, Z_{k,l} \rangle$	28
4.1-1	Geländefläche als reguläres Netz	33
4.1-2	Begriffe in einem DGM.	34
4.1-3	Hybrides digitales Geländemodell.	34
4.1-4	Darstellung einer Geländelinie über einem regulären Gitter	35
4.1-5	Gitterschnitte des Kantenpolygons mit den Gitterebenen	35
4.2-1	Differentialgeometrische Eigenschaften der Geländefläche	37
4.2-2	Transformation des Koordinatensystems	38
4.3-1	Richtungswinkel der Tangente	41
4.3-2	Nachbarpunkte	42
4.3-3	Diskrete Krümmung und Torsion	43
4.3-4	Gerastertes Profil mit zusätzlichen Knickinformationen	44
4.3-5	Profil	45
4.3-6	Berechnung der Gitterschnitte	46
4.3-7	Extrapolation	46
4.3-8	Bruchlinie im Testdatensatz Schneetalpe.	48
4.3-9	Transformation in die extrapolierten Höhen	48
5.1-1	Datenstruktur der WK bei der zweidimensionalen schnellen WT	49
5.1-2	Veranschaulichung der Approximation der partiellen Ableitungen	50
5.2-1	Geländekurve aus dem Datensatz Schneetalpe.	51
5.2-2	Koordinatenfunktionen und Tangentenwinkelfunktionen	51
5.2-3	Die Waveletkoeffizienten der ersten Skale	52
5.2-4	Die rücktransformierten Funktionen	52
5.2-5	Die orthogonalen Abweichungen	53
5.3-1	Wavelettransformation der extrapolierten Höhen	53

5.3-2	Wavelettransformierte der extrapolierten Höhen	54
6.2-1	Unterschiedliche Abszissenabtastung	56
6.2-2	Unterschiedliche Ordinatenabtastung	57
6.2-3	Skalenabhängige Amplituden der kontinuierlichen WT	58
6.2-4	Skalenabhängige Amplituden der diskreten WT	59
6.2-5	Fehler der rekonstruierten Funktion	60
6.3-1	Testgebiet Schneeanpe. Geländeausschnitte unterschiedlicher mittlerer Neigung.	61
6.3-2	Fehler in Abhängigkeit von der Kompressionsrate	62
6.4-1	Ablaufschema der Datenkompression bei Raumkurven.	63
6.4-2	Beispielkurven	64
6.4-3	Fehler der Beispielkurven	66
6.5-1	Modellapproximation der skalenabhängigen Amplituden	68
6.5-2	Amplituden A_L in den einzelnen Skalen	68
6.5-3	Fouriertransformierte vor und nach der Kompression (Coiflet3)	69
6.5-4	Schematische Darstellung der Fouriertransformierten	70
6.5-5	Fouriertransformierte vor und nach der Kompression (Daubechies4)	70
6.5-6	Fouriertransformierte vor und nach der Kompression (Haar-Wavelet)	70
6.5-7	Krümmungskreis und Schmiegeellipse	71
6.5-8	Änderung des Krümmungsradius	71
6.5-9	Kantenanpassung	72
7.1-1	Amplitudenkorrektur	78
7.2-1	Ortsverschiebung	80
7.2-2	Schematische Darstellung der zweidimensionalen stationären WT	82
7.3-1	Halbkugel als Testdatensatz.	86
7.3-2	Verbesserung der Iteration	87
7.3-3	Höhenlinien in einem realen Höhenmodell	87
7.3-4	Talmulde als Testdatensatz.	88
7.3-5	Gefällelinien ohne und mit Modifikation.	88
8.1-1	Projektiertes Mehrschichtenmodell	90
A-1	Schnelle 2D-WT des Datensatzes 'BoxWithCross'	96
A-2	Datensatz um ein Pixel nach rechts verschoben	97
A-3	Ausschnitt aus den beiden Signalmatrizen	97
A-4	Stationäre 2D-WT des Datensatzes 'BoxWithCross'	98
A-5	Stationäre 2D-WT in der 3. Skale	99
B-1	Ortskurve der Filtercharakteristik (Originallage)	102
B-2	Ortsverschiebung des Daub4-Wavelets (Originallage)	102
B-3	Ortskurve der Filtercharakteristik (Symmetrielage)	103
B-4	Ortsverschiebung des Daub4-Wavelets (Symmetrielage)	104
B-5	Frequenzabhängigkeit der Skalierungskorrektur	105
B-6	Vergleich der Skalierungen.	106
D-1	Zellendurchläufe und Matrixtopologien	109
D-2	Belegung der Gitterpunkte	110
D-3	Auswahl einer Startzelle	110
D-4	Mögliche Eckenbelegungen einer (Start-)Zelle	111
D-5	Mögliche Gitterschnitt-Topologien	111
E-1	Das synthetische Signal $z = \sin(x + \sin xy)$	113
E-2	Matrix der skalierten WK	114
F-1	Testgebiet Schneeanpe.	116

1 Einleitung

Durch die rasche technologische Entwicklung bei der Erfassung von Geoinformationen wächst gleichzeitig die Notwendigkeit, den dadurch immer größer werdenden Umfang der Daten, die u. a. in digitalen Geländemodellen (DGM) gespeichert werden, rationell zu verarbeiten und zu verwalten. Als wirkungsvolles Werkzeug für einige dieser Aufgaben hat sich die Wavelettransformation herausgestellt. Die Wavelettransformation (WT) ist eine Integraltransformation, die als eine Verallgemeinerung der Fouriertransformation angesehen werden kann. Sie erweitert einerseits die bekannten Signalanalyseigenschaften der Fouriertransformation vor allem um die bisher fehlende Lokalisierungseigenschaft und ermöglicht andererseits eine sehr effektive Datenkompression, wie entsprechende Ergebnisse in der Bildverarbeitung eindrucksvoll belegen (vgl. z.B. KIEFNER und HAHN (2000)). Für die Beherrschung der Datenflut ist insbesondere die Datenkompression wichtig.

Diese Eigenschaften der Wavelettransformation machen sie auch für die Geowissenschaften, insbesondere für die Geoinformatik interessant, wie die Zunahme an Publikationen zu diesem Thema zeigt. Dabei sind die konkret bearbeiteten Fragestellungen sehr verschieden. BARTHELMES et al. (1994) stellen Anwendungen der Wavelettransformation in der Geodäsie vor, wie z.B. zur Lösung großer linearer Gleichungssysteme, die bei geodätischen Fragestellungen oft vorkommen. Mit Hilfe der Wavelettransformation gelingt GROSS et al. (1996) eine adaptive Flächentriangulation auf der Basis lokaler Spektralschätzungen. Von BAYER et al. (1998) wird die Wavelettransformation z.B. im Bereich der Physikalischen Geodäsie angewendet.

Die Wavelettransformation ermöglicht unter anderem die Geländeanalyse auf der Basis der (eventuell komprimiert) in Form von Waveletkoeffizienten gespeicherten Daten, da die Waveletkoeffizienten wesentliche Signaleigenschaften des Geländes repräsentieren. Um diese Möglichkeiten effektiv zu nutzen, ist es erforderlich, die bestehenden Zusammenhänge zwischen Approximationseigenschaft, Lokalisierungseigenschaft und Kompressionseigenschaft herauszuarbeiten.

Wesentliche Teile der vorliegenden Arbeit entstanden im Rahmen des DFG-Forschungsprojekts „Analyse, Approximation und Kompression von Reliefdaten, speziell digitaler Geländemodelle, mit Hilfe der Wavelettransformation“. Dabei wurde der Schwerpunkt der Untersuchungen auf hybride Geländemodelle gelegt, die auf einem digitalen Höhenmodell (DHM) aufsetzen. Das hybride Modell wird hier als Mehrschichtenmodell verstanden, bei dem in den grundlegenden ersten beiden Schichten die reinen Höheninformationen (DHM, Rasterdaten, erste Schicht) und die reliefbezogenen Raumkurven (Vektordaten, zweite Schicht) gespeichert sind. Weitere Schichten sind zur Aufnahme weiterer raumbezogener Geo-Informationen gedacht.

Von der Vielzahl der daraus unmittelbar resultierenden Teilprobleme wurden die folgenden, die sich insbesondere auf die Wechselwirkung zwischen Geländefläche und Geländekurven beziehen, schwerpunktmäßig untersucht:

- Beschreibungsmöglichkeiten explizit gegebener Raumkurven für eine effektive Wavelettransformation.
- WT-gerechte Beschreibung von Raumkurven in hybriden DGM.
- Probleme der Kompressionsverfahren in Mehrschichtenmodellen.

Zusätzlich ergaben sich mehrere Probleme zum theoretischen Hintergrund der diskreten Wavelettransformation. Es erschien lohnenswert, einige Effekte der allgemein verbreiteten Programme, für die bisher in der Literatur keine deutlichen Erklärungen zu finden waren, genauer zu untersuchen. Dazu gehörte vor allem die Quantifizierung der Lokalisierungseigenschaft bei der

diskreten Wavelettransformation, die sonst nur für den Grenzübergang $a \rightarrow 0$ des Frequenzparameters bei der kontinuierlichen Wavelettransformation erklärt ist. Es zeigte sich, daß vor allem die sachlich richtige Interpretation der Waveletkoeffizienten als Output der Verfahren die auftretenden Effekte zu erklären hilft. Diese Interpretation ist besonders auch für die Fragestellungen der Signalanalyse von Bedeutung. Weitere wichtige Problemkreise waren daher:

- Die Aufbereitung der algorithmischen Zusammenhänge der üblichen Verfahren aus der Sicht der Indexbeziehungen, die sich aus der iterativen Konstruktion der Wavelets ergeben.
- Die Übertragung der Lokalisierungs- und Approximationseigenschaft auf die diskrete Wavelettransformation, ohne den Grenzübergang $a \rightarrow 0$ voraussetzen zu müssen.
- Untersuchung der Anwendbarkeit der diskreten Wavelettransformation für die Analyse von DGM im Waveletbereich, dem Bildbereich der WT.

Die Arbeit verfolgt damit die Zielstellung, die oben genannten ursprünglichen Aufgaben des DFG-Projektes in einem um diese Fragestellungen erweiterten Gesamtrahmen darzustellen.

Bei den Untersuchungen zu den theoretischen Grundlagen konnten einige interessante Ergebnisse erzielt werden, die in dieser konkreten Form in der Wavelet-Literatur nicht zu finden sind. Die Darstellung dieser Ergebnisse, bei der auf eine didaktische Aufbereitung besonderer Wert gelegt wurde, ist einer der Schwerpunkte der vorliegenden Arbeit. Gleichzeitig werden bereits veröffentlichte Einzelergebnisse in diesen Zusammenhang eingeordnet.

Im ersten Teil der Arbeit werden nach einer kurzen Einführung in die kontinuierliche Wavelettransformation (Kapitel 2) die wesentlichen Grundlagen der diskreten Wavelettransformation zusammengestellt und so aufbereitet, daß sie eine zweckmäßige formale Basis für die durchzuführenden Untersuchungen darstellen (Kap. 3). Die Verfahren der diskreten Wavelettransformation werden im unmittelbaren Zusammenhang mit der iterativen Konstruktion der diskreten Wavelets gesehen. Als mathematische Basis fungiert im wesentlichen der Matrizenkalkül.

Danach werden zunächst Raumkurven als ein Hauptbestandteil hybrider Geländemodelle unter dem Gesichtspunkt der Wavelettransformation betrachtet (Kapitel 4). Raumkurven als eindimensionale Mannigfaltigkeiten im dreidimensionalen Raum lassen sich nicht durch nur eine Gleichung und damit auch nicht als eine Funktion beschreiben. Folglich können die Ansätze, die es für ebene Kurven gibt, z.B. WERSCHLEIN (1996), nicht einfach übernommen werden, sondern die Ansätze müssen entsprechend verallgemeinert werden. In Abschnitt 4.3.1 werden derartige Beschreibungsmöglichkeiten diskutiert, vgl. auch BEYER (1999).

In hybriden Modellen mit im allgemeinen grundsätzlich unterschiedlicher Datenstruktur von Flächen- und Kurvendaten (Rasterdaten, Vektordaten), erscheint es im Interesse konsistenter Datensätze im Waveletbereich (WB) sinnvoll, die Datenstrukturen bereits vor der Wavelettransformation im Ortsbereich anzupassen. Dazu dient eine implizite Beschreibung der reliefbezogenen Kurven, die in Abschnitt 4.3.2 vorgestellt wird, vgl. auch BEYER (2000a), BEYER (2000b). Probleme der Rekonstruierbarkeit der so transformierten Daten werden im Anhang D behandelt.

Für die Analyse der Geländemodelle im Waveletbereich ist die Approximationseigenschaft im unmittelbaren Zusammenhang mit der Lokalisierungseigenschaft von Bedeutung. Die Waveletkoeffizienten sind bei einem Wavelet n -ter Ordnung proportional zur n -ten Ableitung der transformierten Funktion. Neben der Bestimmung des Proportionalitätsfaktors ist auch die genaue Lokalisierung für die Analyse wichtig. Es zeigt sich, daß durch den endlichen Support bei den meist asymmetrischen Wavelets die Indizierung der Waveletkoeffizienten nicht unmittelbar als Maß für die Lokalisierung geeignet ist. In den Abschnitten 7.1 und 7.2 werden diese Probleme ausführlich behandelt, vgl. auch BEYER und MEIER (2001), BEYER (2002).

Durch die Approximationseigenschaft ergeben sich spezielle Analyse­möglichkeiten und die Ableitung von Folgeprodukten im Waveletbereich. Zum einen können Geländeneigung und -krümmung bei entsprechender Ordnung der benutzten Wavelets unmittelbar aus den Waveletkoeffizienten bestimmt werden (Abschnitt 4.2). Zum anderen ist es darauf aufbauend möglich, Höhen- und Gefällelinien aus den Waveletkoeffizienten zu generieren (Abschnitt 7.3). Von RICHTER (2001) wurden diesbezügliche Algorithmen entwickelt.

Bei der Anwendung der Kompressionsverfahren, die in Abschnitt 6.2 beschrieben werden, treten bei der oben genannten konsistenten Beschreibung hybrider Modelle Anpassungsprobleme im Zuge der Rekonstruktion des Geländemodells durch die inverse Wavelettransformation auf. Diese Anpassungsprobleme ergeben sich durch die unterschiedlichen Glättungseffekte bei Flächen und Kurven. Diese Probleme wurden von WEICHSEL (2001) untersucht. Die Ergebnisse werden in Abschnitt 6.5 kurz zusammengefaßt.

2 Wavelet-Transformation zur Analyse, Approximation und Kompression

2.1 Integraltransformationen

Transformationen spielen bei der Untersuchung und Bearbeitung mathematisch beschreibbarer Sachverhalte schon lange eine wichtige Rolle. In der Analysis hat man das Problem der Darstellung beliebiger Funktionen durch spezielle Funktionen mit bestimmten Eigenschaften, d.h. der Approximation durch Funktionen aus einem geeigneten Unterraum. In der Signalverarbeitung sind für die Signalanalyse vor allem Integraltransformationen von Bedeutung. In der Geometrie dienen Koordinatentransformationen u. a. zur Klassifizierung von geometrischen Objekten, z. B. die Klassifizierung der Kurven zweiter Ordnung mittels Hauptachsentransformation.

Allgemein kann man sagen, daß ein Element eines Vektorraumes, das bezüglich einer Basis dieses Vektorraumes dargestellt ist, in die Darstellung bezüglich einer anderen Basis transformiert wird. Es handelt sich also um eine Transformation der Darstellung. Von der neuen Basis hängt ab, ob es sich nur um eine Approximation oder um eine Transformation ohne Informationsverlust handelt.

Eine Integraltransformation einer Funktion $f(t)$ ist gegeben durch

$$g(s) = \int_a^b K(s, t) f(t) dt . \quad (2.1-1)$$

Der sogenannte Kern $K(s, t)$ der Integraltransformation stellt gewissermaßen die neue Basis für die Beschreibung der Funktion dar.

Die Fouriertransformation hat als Kern die Funktion $K(s, t) = e^{ist}$ und ist mit einem Skalierungsfaktor gegeben durch

$$g(s) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ist} f(t) dt . \quad (2.1-2)$$

Der Kern $e^{ist} = \cos(st) + i\sin(st)$ besteht aus den harmonischen Schwingungen aller Frequenzen s . Damit liegt eine Zerlegung des Signals in Schwingungsanteile vor.

Die Fouriertransformation bildet eine Funktion im Orts- bzw. Zeitbereich in eine Funktion im Frequenzbereich¹ ab. Die transformierte Funktion enthält ausschließlich Frequenzinformationen. Es erfolgt eine vollständige Trennung zwischen Orts- und Frequenzbereich.

Bei der Wavelettransformation ist der Kern zweidimensional, er enthält einen Orts- und einen Frequenzparameter. Damit ist eine Lokalisierung im Bildbereich (Waveletbereich) möglich. Die Wavelettransformation ist gegeben durch

$$L_\psi f(a, b) := \sqrt{\frac{1}{c_\psi |a|}} \int_{-\infty}^{\infty} f(t) \psi \left(\frac{t-b}{a} \right) dt . \quad (2.1-3)$$

Dabei ist a der Frequenzparameter und b ist der Ortsparameter. Die Analysefunktionen ψ werden durch den Parameter a in Richtung der t -Achse skaliert und durch den Parameter b auf der t -Achse verschoben. Die transformierte Funktion enthält damit sowohl Frequenzinformationen als auch Ortsinformationen.

¹ Bei einer Transformation aus dem Ortsbereich nennt man den Bildbereich auch Wellenzahlbereich.

2.2 Kontinuierliche Wavelettransformation

Durch eine fast beliebig wählbare Funktion ψ (Wavelet) als Analysefunktion wird mit der Wavelettransformation eine gegenüber der Fouriertransformation wesentlich höhere Flexibilität erreicht. Den Namen erhält das Wavelet („Wellchen“) durch seine Form, eine weitgehend nur auf einen begrenzten Definitionsbereich ausgedehnte Schwingung. Durch Stauchung (Dilatation, Parameter a) und Verschiebung (Translation, Parameter b) des sogenannten Prototyps ψ entsteht die Funktionsbasis $\psi_{a,b}$ des Bildraumes, die damit als Analysefunktionen sowohl Frequenzen als auch Orte analysieren. Die Wavelettransformierte $L_\psi f$ einer Funktion f ergibt sich über das L^2 -Skalarprodukt $\langle f, \psi \rangle$ zu

$$L_\psi f(a, b) := \sqrt{\frac{1}{c_\psi |a|}} \int_{-\infty}^{\infty} f(t) \psi\left(\frac{t-b}{a}\right) dt \quad (2.2-1)$$

als Funktion der beiden Parameter a und b . Der Faktor $\sqrt{1/c_\psi |a|}$ dient der Erzielung gewisser zweckmäßiger Eigenschaften. c_ψ ist definiert durch

$$c_\psi := 2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|}{|\omega|} d\omega \quad (2.2-2)$$

und muß die Zulässigkeitsbedingung

$$0 < c_\psi < \infty \quad (2.2-3)$$

erfüllen.

Bei richtiger Wahl von c_ψ vermittelt die Wavelettransformation eine Isometrie zwischen gewichteten L^2 -Räumen. Daraus ergibt sich die Inversionsformel

$$f(t) := \sqrt{\frac{1}{c_\psi}} \int_{\mathbb{R}} \int_{\mathbb{R}} L_\psi f(a, b) \sqrt{\frac{1}{|a|}} \psi\left(\frac{t-b}{a}\right) \frac{da db}{a^2}. \quad (2.2-4)$$

Nach LOUIS et al. (1994) läßt sich die Wavelettransformierte $L_\psi f$ interpretieren als

- Phasenraumdarstellung von f ,
- Approximation einer Ableitung von f ,
- Aufspalten von f in Anteile zu verschiedenen Frequenzbändern.

Die Lokalisierungseigenschaft, die mit der ersten Interpretation korrespondiert, wird im Abschnitt 7.2 näher untersucht, da sich zeigt, daß die Lokalisierung nur für $a \rightarrow 0$ exakt ist und einer momentabhängigen Korrektur bedarf. Eine Verbesserung der Lokalisierung ist dennoch nur soweit möglich, wie es die Heisenbergsche Unschärferelation zuläßt. Die „exakte“ Lokalisierung ist u. a. dann erforderlich, wenn die Wavelettransformation zur Datenanalyse genutzt wird, z. B. zur Bestimmung der Ableitungen mittels ihrer Approximationseigenschaft. Diese Problematik wird insgesamt im Kap. 7 behandelt.

Für weitergehende Aussagen zur kontinuierlichen Wavelettransformation wird auf die Literatur (DAUBECHIES (1992), LOUIS et al. (1994), BLATTER (1998)) verwiesen, da das Gewicht dieser Arbeit auf Verfahren der diskreten Wavelettransformation liegt.

Während die kontinuierliche Theorie für das Verständnis der inneren Zusammenhänge der Transformation von Bedeutung ist, ist für die praktische Nutzung der Wavelettransformation mit Hilfe

der Computertechnik eine effiziente diskrete Version erforderlich. Die Multiskalenanalyse (MSA, engl. multi resolution analysis: MRA) leistet das hervorragend. Im Gegensatz zur Fouriertransformation, bei der die Fourierreihe nur die Teilmenge der 2π -periodischen Funktionen repräsentieren kann, lassen sich mit der MSA alle $f \in L^2(\mathbb{R})$ darstellen, nur begrenzt durch die Auflösung des diskret gegebenen Signals. Im folgenden Kapitel werden die Grundlagen der Multiskalenanalyse im Hinblick auf die zu lösenden Fragestellungen der vorliegenden Arbeit dargestellt. Ein wesentlicher Schwerpunkt liegt dabei auf der Interpretation der Indizierung der durch Vektoren bzw. Matrizen repräsentierten 1D- bzw. 2D-Signale. Die Indizes der Vektoren bzw. Matrizen sind die Repräsentation der Lokalisierung des Ortes bzw. der Zeit der zu transformierenden Signale.

3 Diskrete Wavelettransformation

3.1 Matrizenbeziehungen diskreter Funktionen

Die praktische Realisierung der diskreten Wavelettransformation durch die in 3.2 beschriebene Multiskalenanalyse nach MALLAT (1989) basiert in den Programmtools zur Wavelettransformation i. allg. auf dem Matrizenkalkül, da er eine sehr zweckmäßige Beschreibung der mathematischen Zusammenhänge in Hinblick auf die realisierbaren Datenstrukturen zuläßt.

Eine Funktion $f : \mathbb{Z} \rightarrow \mathbb{R}$ über endlich vielen diskreten Argumenten kann als endlichdimensionaler Vektor

$$\underline{f} = \begin{bmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{bmatrix} \quad (3.1-1)$$

aufgefaßt werden. Dieser läßt sich bezüglich einer endlichen Basis $\{\underline{b}_1, \underline{b}_2, \dots, \underline{b}_n\}$ mit den Koeffizienten b_1, b_2, \dots, b_n darstellen:

$$\underline{f} = \sum_{i=1}^n b_i \cdot \underline{b}_i. \quad (3.1-2)$$

In Matrixschreibweise hat man

$$\underline{f} = B \cdot \underline{f}_B \quad \text{mit} \quad B = \begin{bmatrix} | & | & \cdots & | \\ \underline{b}_1 & \underline{b}_2 & \cdots & \underline{b}_n \\ | & | & & | \end{bmatrix}, \quad \underline{f}_B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}. \quad (3.1-3)$$

Sind die Basisfunktionen orthogonal, so ist B eine orthogonale Matrix ($B^{-1} = B^\top$). Die Koeffizienten b_i ergeben sich demnach zu

$$\underline{f}_B = B^{-1} \cdot \underline{f} = B^\top \cdot \underline{f}, \quad (3.1-4)$$

d.h., die Koeffizienten b_i sind die Skalarprodukte der Funktion \underline{f} mit den Basisfunktionen \underline{b}_i :

$$b_i = \langle \underline{b}_i, \underline{f} \rangle. \quad (3.1-5)$$

Ist die Matrix B der Basisvektoren \underline{b}_i die Einheitsmatrix E mit

$$E = \begin{bmatrix} | & | & \cdots & | \\ \underline{e}_1 & \underline{e}_2 & \cdots & \underline{e}_n \\ | & | & & | \end{bmatrix}, \quad (3.1-6)$$

so sind die Koeffizienten b_i der Linearkombination (3.1-2) gleich den Elementen f_i des Funktionswertvektors \underline{f} . Man nennt deshalb die Basis $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ auch kanonische Basis. Die kanonische Basis der diskreten Funktionen entspricht der δ -Basis der kontinuierlichen Funktionen. Diese Analogie ist in Tab. 3.1-1 veranschaulicht.

Damit liegt, wenn die \underline{b}_i Waveletvektoren in der Terminologie von Abschnitt 3.2, (3.2-15) sind, die Matrizenform der Wavelettransformation mit den Waveletkoeffizienten b_i vor.

Tabelle 3.1-1: Analogien zwischen stetigen und diskreten Funktionen

$y = f(x)$	$\underline{y} = [y_1 \quad y_2 \quad \cdots \quad y_8]^T$
<p>Funktionsbasis: δ-Funktion</p>	<p>kanonische Basis $\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$</p>
$f(x) = \int_{-\infty}^{\infty} f(u)\delta(x - u)du$	$\underline{y} = y_1 \cdot \underline{e}_1 + y_2 \cdot \underline{e}_2 + \cdots + y_8 \cdot \underline{e}_8 = E\underline{y}$

3.2 Konzept der Multiskalenanalyse

Die diskrete Wavelettransformation verdankt ihren Erfolg dem Konzept der Multiskalenanalyse. Nachfolgend wird die Multiskalenanalyse in Anlehnung an LOUIS et al. (1994) eingeführt.

Definition 1 (Multiskalenanalyse) Die Multiskalenanalyse (MSA) des $L^2(\mathbb{R})$ ist eine aufsteigende Folge abgeschlossener Unterräume $V_m \subset L^2(\mathbb{R})$,

$$\{0\} \subset \dots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \dots \subset L^2(\mathbb{R}),$$

so daß gilt

$$\overline{\bigcup_{m \in \mathbb{Z}} V_m} = L^2(\mathbb{R}), \tag{3.2-1}$$

$$\bigcap_{m \in \mathbb{Z}} V_m = \{0\}, \tag{3.2-2}$$

$$f(\cdot) \in V_m \iff f(2^m \cdot) \in V_0, \quad m \in \mathbb{Z}. \tag{3.2-3}$$

Es gibt eine Funktion $\varphi \in L^2(\mathbb{R})$, deren ganzzahlige Translate eine Riesz-Basis von V_0 erzeugen, d.h.

$$V_0 = \overline{\text{span}\{\varphi(\cdot - k) \mid \}}$$

und

$$A \sum_{k \in \mathbb{Z}} c_k^2 \leq \left\| \sum_{k \in \mathbb{Z}} c_k \varphi(\cdot - k) \right\|_{L^2}^2 \leq B \sum_{k \in \mathbb{Z}} c_k^2 \tag{3.2-4}$$

für alle $\{c_k\}_{k \in \mathbb{Z}} \in l^2(\mathbb{Z})$. Es bezeichnen A und B positive Konstanten.

Die Räume V_m sind folglich skalierte Versionen des Grundraumes V_0 , der durch Translation einer Funktion $\varphi(x) \in V_0$, der Skalierungsfunktion, aufgespannt wird. Die Skalierungsfunktion wird auch „Prototyp“ der Basisfunktionen genannt. Es ist dann $\{\varphi_{0n}\}$ eine Basis von V_0 mit $\varphi_{0n}(x) = \varphi(x - n)$, $n \in \mathbb{Z}$. Eine Funktion $f \in V_0$ hat demzufolge eine Darstellung

$$f = \sum_n a_n \cdot \varphi_{0n}. \quad (3.2-5)$$

Das gilt auch für die Basisfunktionen, insbesondere für den Prototyp $\varphi(x)$. Wegen

$$\begin{aligned} \varphi(x) &\in V_0 \subset V_1 \\ \varphi(x) &\in V_1 \stackrel{(3.2-3)}{\implies} \varphi\left(\frac{x}{2}\right) \in V_0 \end{aligned}$$

folgt

$$\varphi\left(\frac{x}{2}\right) = \sum_n c_n \cdot \varphi(x - n). \quad (3.2-6)$$

Die Substitution $\frac{x}{2} \rightarrow x$ führt auf die Skalierungsgleichung

$$\varphi(x) = \sum_n c_n \cdot \varphi(2x - n). \quad (3.2-7)$$

Um eine eindeutige Lösung der Skalierungsgleichung zu erhalten, betrachtet man die Normierung $\int \varphi(x) dx = 1$. Wegen

$$2 = 2 \int_{\mathbb{R}} \varphi(x) dx = 2 \int_{\mathbb{R}} \sum_n c_n \varphi(2x - n) dx = 2 \sum_n c_n \int_{\mathbb{R}} \varphi(2x - n) dx \quad (3.2-8)$$

folgt mit der Substitution $y := 2x - n$

$$2 = 2 \sum_n c_n \int_{\mathbb{R}} \varphi(y) \frac{dy}{2} = \sum_n c_n \underbrace{\int_{\mathbb{R}} \varphi(y) dy}_{=1} \quad (3.2-9)$$

und damit als notwendige Bedingung für die c_n

$$\sum c_n = 2. \quad (3.2-10)$$

Die verbleibenden Freiheitsgrade für die Koeffizienten führen auf eine Vielzahl verschiedener Folgen von Unterräumen mit jeweils speziellen Eigenschaften für die Wavelettransformierten.

Die c_i werden auch Skalierungskoeffizienten genannt und in der Literatur teilweise mit h_i bezeichnet.

Die diskrete Wavelettransformation ist dann im Grunde eine fortlaufende Zerlegung des Signals f aus einem Unterraum V_0 des $L^2(\mathbb{R})$ in einen „glatten“ und einen „rauen“ Anteil. Es handelt sich dabei um eine Projektion $P_1 f$ des Signals $f \in V_0$ in einen Unterraum $V_1 \subset V_0$ und eine Projektion $Q_1 f$ in das mit W_1 bezeichnete orthogonale Komplement von V_1 in V_0 . Wird jeweils die Projektion $P_i f$ in den Unterraum $V_i \subset V_{i-1}$ weiter in die Unterräume V_{i+1} und W_{i+1} projiziert, entsteht eine Folge von Projektionen $\{Q_1 f, Q_2 f, \dots, Q_n f, P_n f\}$, aus denen sich das Signal f rekonstruieren läßt:

$$f = Q_1 f + Q_2 f + \dots + Q_n f + P_n f. \quad (3.2-11)$$

Diese Zerlegung des Signals durch einen Baumalgorithmus illustriert Abb. 3.2-1. Die Projektionen $Q_i f$ enthalten die Anteile von f zu bestimmten Frequenzbändern, die auch Skalen genannt werden (LOUIS et al. (1994), S. 104).

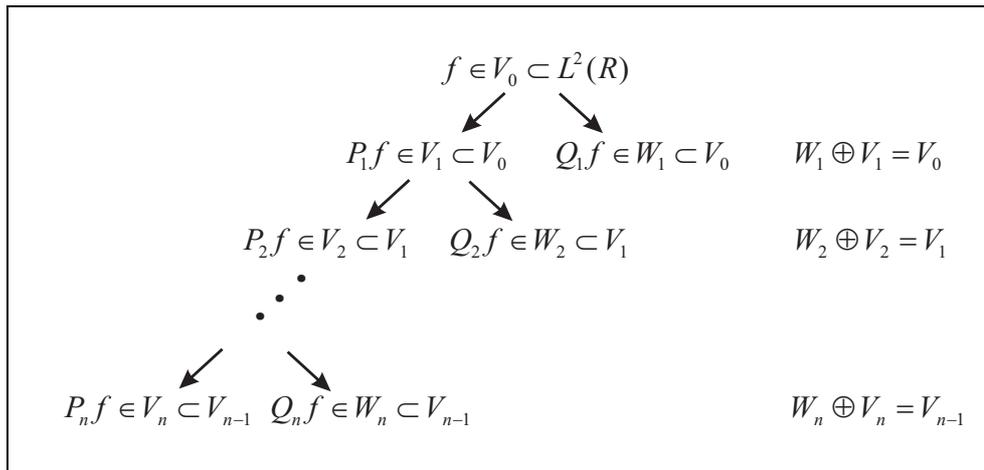


Abbildung 3.2-1: Schematische Darstellung der schrittweisen Zerlegung des Signals in „glatte“ und „rauhe“ Anteile.

Praktisch werden diese Projektionen durch die Waveletkoeffizienten ausgedrückt, die als Skalarprodukte des Signalvektors mit den Basisvektoren der Unterräume V_i ($i = 1, \dots, n$) und W_n berechnet werden.

Die Basisfunktionen der orthogonalen Komplemente W_i sind die Wavelets (sog. Mutterfunktionen). Aus gegebenen Koeffizienten c_i , die die Skalierungsgleichung befriedigen, lassen sich die Wavelets für die verschiedenen Skalen iterativ bestimmen; siehe STRANG and NGUYEN (1997).

Man startet im kontinuierlichen Fall mit der Basis $\{\tilde{\varphi}_{0n}\}$ des Funktionenraumes V_0 mit

$$\tilde{\varphi}_{0n}(x) = \begin{cases} 1 & \text{für } n \Delta x \leq x < (n+1) \Delta x \\ 0 & \text{sonst} \end{cases} .$$

Dieser Funktionenraum entspricht einer Diskretisierung mit der Tastweite Δx . In den untersuchten Algorithmen der diskreten Wavelettransformation übernimmt die kanonische Basis¹ $\{\underline{\varphi}_{0n}\}$ mit

$$\underline{\varphi}_{0n} = \left[\cdots \underbrace{0 \ 0 \ 1 \ 0 \ 0}_{n\text{-te Stelle}} \cdots \right]^T, \quad n \in \mathbb{Z}$$

mit dem sogenannten Prototyp $\underline{\varphi}_{00}$ die Rolle der oben erwähnten Basis $\{\tilde{\varphi}_{0n}\}$. Die Prototypen der Vaterfunktion $\underline{\varphi}_{k0}$ und der Mutterfunktion $\underline{\psi}_{k0}$ (Wavelet) der k -ten Skale ergeben sich dann iterativ zu

$$\begin{aligned}
 \underline{\varphi}_{k0} &= \frac{1}{2} \sqrt{2} \sum_n c_n \cdot \underline{\varphi}_{k-1,n}, \\
 \underline{\psi}_{k0} &= \frac{1}{2} \sqrt{2} \sum_n (-1)^n c_{1-n} \cdot \underline{\varphi}_{k-1,n}.
 \end{aligned} \tag{3.2-12}$$

Der Faktor $1/2\sqrt{2}$ sichert $\|\varphi_{kn}\| = \|\psi_{kn}\| = 1$. Die vollständige Basis der k -ten Skale ergibt sich durch fortlaufende Verschiebung des Prototyps um 2^k .

¹ Die kanonische Basis (Standardbasis) einer diskreten Funktion in der gegebenen Signalaufösung entspricht der δ -Basis einer kontinuierlichen Funktion.

Um nicht durch zu viele Indizes den Blick für das Wesentliche zu verlieren, werden an dieser Stelle einige vereinfachende Schreibweisen eingeführt, die verwendet werden, wenn durch den Kontext Verwechslungen ausgeschlossen sind.

Das diskret gegebene endliche Signal (die Signalfolge)

$$\{ f(x_1), f(x_2), \dots, f(x_n) \} \quad (3.2-13)$$

werde durch den Funktionswertvektor (Signalvektor)

$$\underline{f} = \begin{bmatrix} f_1 & f_2 & \cdots & f_n \end{bmatrix}^\top \quad (3.2-14)$$

und das verwendete Wavelet durch den Waveletvektor

$$\underline{\psi} = \begin{bmatrix} v_1 & v_2 & \cdots & v_s \end{bmatrix}^\top \quad (3.2-15)$$

beschrieben. Dabei sind n die Länge des Signals und s die Länge (Support) des Wavelets in der betrachteten Skale². Unter dem Support des Wavelets versteht man den Bereich vom ersten bis zum letzten von Null verschiedenen Element des diskreten Wavelets bzw. auch die Länge dieses Bereiches. Die oben konstruierten Wavelets sind eigentlich unendlich lange Folgen, die auf diese Weise auf ihren Support beschränkt als endliche Vektoren beschrieben werden³. Die Wavelettransformation, die durch das Skalarprodukt⁴ $w_i = \langle \underline{\psi}_i, \underline{f} \rangle$ realisiert wird (vgl. 3.1), liefert einen Vektor von Waveletkoeffizienten

$$\underline{w} = \begin{bmatrix} w_1 & w_2 & \cdots & w_n \end{bmatrix}^\top. \quad (3.2-16)$$

Dieser kann wiederum als diskretes Signal aufgefaßt werden und erlaubt unter Umständen eine lokale Zuordnung der Waveletkoeffizienten zu den Argumenten x_1, x_2, \dots, x_n des Signals. In 3.4.1 wird die Struktur des Vektors \underline{w} für die zwei in dieser Arbeit untersuchten Verfahren im Detail beschrieben.

Analog zum eindimensionalen Fall werde das diskret gegebene endliche Signal $z_{ij} = f(x_i, y_j)$ durch die Funktionswertematrix (Signalmatrix)

$$Z = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nn} \end{bmatrix} \quad (3.2-17)$$

beschrieben. Für die Konstruktion der zweidimensionalen Wavelets werden sowohl das eindimensionale Wavelet ψ (Hochpaß) als auch die eindimensionale Skalierungsfunktion φ (Tiefpaß) benötigt. Sie sollen hier kurz als Hochpaßvektor

$$\underline{h} := \psi = \begin{bmatrix} v_1^h & v_2^h & \cdots & v_s^h \end{bmatrix}^\top \quad (3.2-18)$$

und Tiefpaßvektor

$$\underline{t} := \varphi = \begin{bmatrix} v_1^t & v_2^t & \cdots & v_s^t \end{bmatrix}^\top \quad (3.2-19)$$

² Im Interesse einer einfachen Indizierung ist in (3.2-15) der Skalenindex weggelassen worden.

³ In der Praxis hat man es ohnehin mit endlichen Signalvektoren bzw. Signalmatrizen zu tun.

⁴ Dem Skalarprodukt $\langle \psi, f \rangle = \int \psi \cdot f \, dx$ im kontinuierlichen Fall entspricht im diskreten Fall das Skalarprodukt $\langle \underline{\psi}_i, \underline{f} \rangle = \underline{\psi}_i^\top \underline{f}$.

bezeichnet werden. Dabei sind (n, n) das Format der Signalmatrix und s analog zum eindimensionalen Fall der Support des Wavelets bzw. der Skalierungsfunktion in der betrachteten Skale.

Beispiel 1 Daubechies4-Wavelet

Die Koeffizienten des Daubechies4-Wavelet sind $c_0 = \frac{1}{4}(1 + \sqrt{3})$, $c_1 = \frac{1}{4}(3 + \sqrt{3})$, $c_2 = \frac{1}{4}(3 - \sqrt{3})$, $c_3 = \frac{1}{4}(1 - \sqrt{3})$. Man erhält für den Hochpaßvektor \underline{h} und den Tiefpaßvektor \underline{t} in der ersten Skale gemäß (7.1-2)

$$\underline{h} := \begin{bmatrix} -0.1294 & -0.2241 & 0.8365 & -0.4830 \end{bmatrix}^\top$$

$$\underline{t} := \begin{bmatrix} 0.4830 & 0.8365 & 0.2241 & -0.1294 \end{bmatrix}^\top,$$

vgl. WICKERHAUSER (1994).

In den üblichen Verfahren der diskreten Wavelettransformation, vgl. z.B. BUCKHEIT et al. (1995), wird die Signalmatrix Z bei der sogenannten schnellen Wavelettransformation in die Matrix

$$W = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & & \vdots \\ w_{n1} & w_{n2} & \cdots & w_{nn} \end{bmatrix} \quad (3.2-20)$$

der Waveletkoeffizienten überführt. Diese ist aus Teilmatrizen zusammengesetzt, die aus den Skalarprodukten $\langle HH, Z \rangle$, $\langle HT, Z \rangle$, $\langle TH, Z \rangle$ und $\langle TT, Z \rangle$ entstehen (vgl. Abb. 3.4-4).

Bei der sogenannten stationären Wavelettransformation werden die Wavelets zur Berechnung der Waveletkoeffizienten in jeder Richtung um jeweils Eins über die Signalmatrix geschoben. Damit entstehen in jeder Skale vier Matrizen W_{HH} , W_{HT} , W_{TH} , W_{TT} von Waveletkoeffizienten im Format der Signalmatrix Z , die die partiellen Ableitungen approximieren.

3.3 Konstruktion der diskreten Wavelets

3.3.1 Konstruktion der eindimensionalen Wavelets

Die Momente des Wavelets, die, wie in 7.2 gezeigt wird, die Lokalisierung der Approximation beeinflussen, sind von der Verschiebung des Wavelets, d.h. seiner Indizierung, abhängig. Es ist deshalb notwendig, die verfahrensbedingten Indexbeziehungen korrekt zu berücksichtigen. In den Abbildungen 3.3-1 und 3.3-2 wird zur Verdeutlichung der Indexbeziehungen bei der rekursiven Konstruktion der diskreten Wavelets die Konstruktion einer Wavelet-Basis mit vier Filterkoeffizienten aus der kanonischen Basis grafisch dargestellt⁵. Insbesondere wird der Einfluß auf die zyklische Verschiebung der Wavelets und der Skalierungsfunktionen sichtbar.

Der Support aller Vaterfunktionen ϕ_{k0} beginnt gemäß (3.2-12) immer an der 0-ten Stelle bezüglich der vereinbarten Zählung. Der Support der Wavelets ψ_{k0} beginnt jedoch versetzt um $d_1 = -2^{k-1}(n_c - 2)$, wobei k die Nummer der Skale und n_c die Anzahl der Filterkoeffizienten c_i des betrachteten Wavelets ist.

⁵ Da nur die Indexbeziehungen veranschaulicht werden sollen, ist bei der Beschriftung der Skizzen der Skalierungsfaktor $\frac{1}{2}\sqrt{2}$ weggelassen worden.

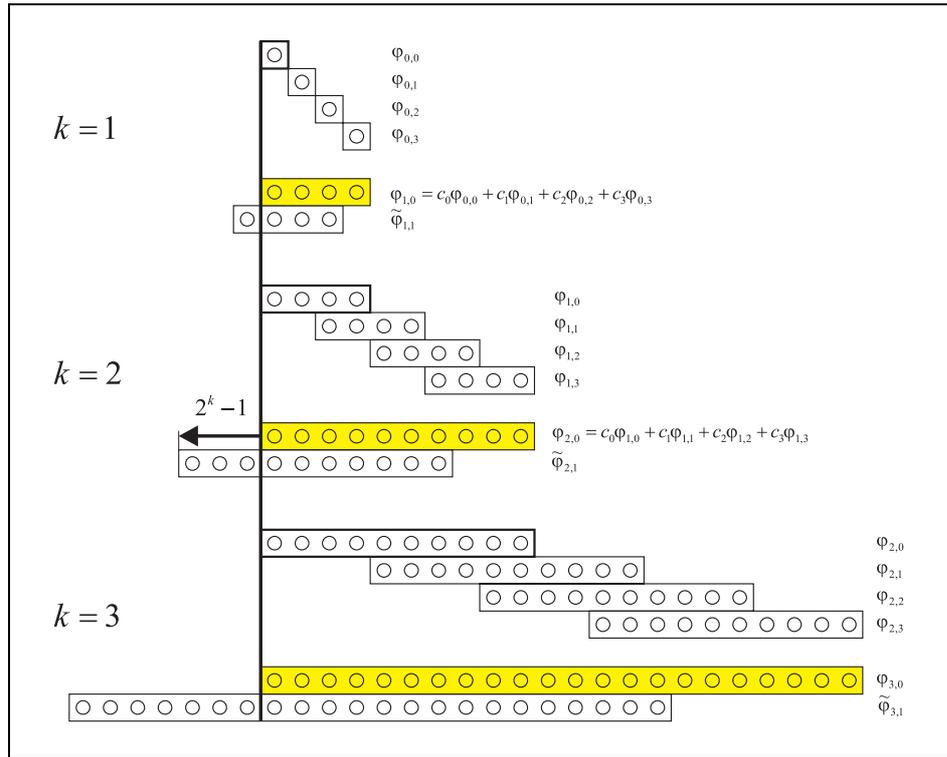


Abbildung 3.3-1: Schematische Darstellung der rekursiven Berechnung der Skalierungsfunktionen für ein Wavelet mit vier Filterkoeffizienten c_0, c_1, c_2, c_3 (z.B. Daubechies4-Wavelet) in den ersten drei Skalen $k = 1, 2, 3$.

Für die Untersuchungen zur Approximationseigenschaft werden nicht die oben hergeleiteten Waveletbasisfunktionen ϕ und ψ der schnellen Wavelettransformation, sondern die Waveletbasisfunktionen $\tilde{\phi}$ und $\tilde{\psi}$ der stationären Wavelettransformation (vgl. 3.4.1) betrachtet. Diese entstehen durch die Zuordnung $\tilde{\phi}_{k,2^k} = \phi_{k,0}$ bzw. $\tilde{\psi}_{k,2^k} = \psi_{k,0}$.

Durch fortlaufende Verschiebung um Eins in beide Richtungen werden diese Basisfunktionen zur Waveletmenge $\{\tilde{\psi}_{k,1}, \tilde{\psi}_{k,2}, \dots, \tilde{\psi}_{k,n}\}$ (Hochpaß) und zur Menge der Skalierungsfunktionen $\{\tilde{\phi}_{k,1}, \tilde{\phi}_{k,2}, \dots, \tilde{\phi}_{k,n}\}$ (Tiefpaß) vervollständigt. Die Beschränkung auf je n Basisfunktionen gilt für Signale mit der endlichen Signallänge n .

Die jeweils erste Funktion $\tilde{\phi}_{k,1}$ bzw. $\tilde{\psi}_{k,1}$ ist dann um $d_2 = -(2^k - 1)$ gegenüber $\phi_{k,0}$ bzw. $\psi_{k,0}$ verschoben. Insgesamt ergibt sich damit eine Verschiebung des i -ten Wavelets $\tilde{\psi}_{k,i}$ um $d^h + i$ mit

$$d^h = d_1 + d_2 - 1 = -2^{k-1}(n_c - 2) - (2^k - 1) - 1 ,$$

also

$$\boxed{d^h = -2^{k-1}n_c ,} \tag{3.3-1}$$

und der i -ten Skalierungsfunktion $\tilde{\phi}_{k,i}$ um $d^t + i$ mit

$$d^t = d_2 - 1 = -(2^k - 1) - 1 ,$$

also

$$\boxed{d^t = -2^k .} \tag{3.3-2}$$

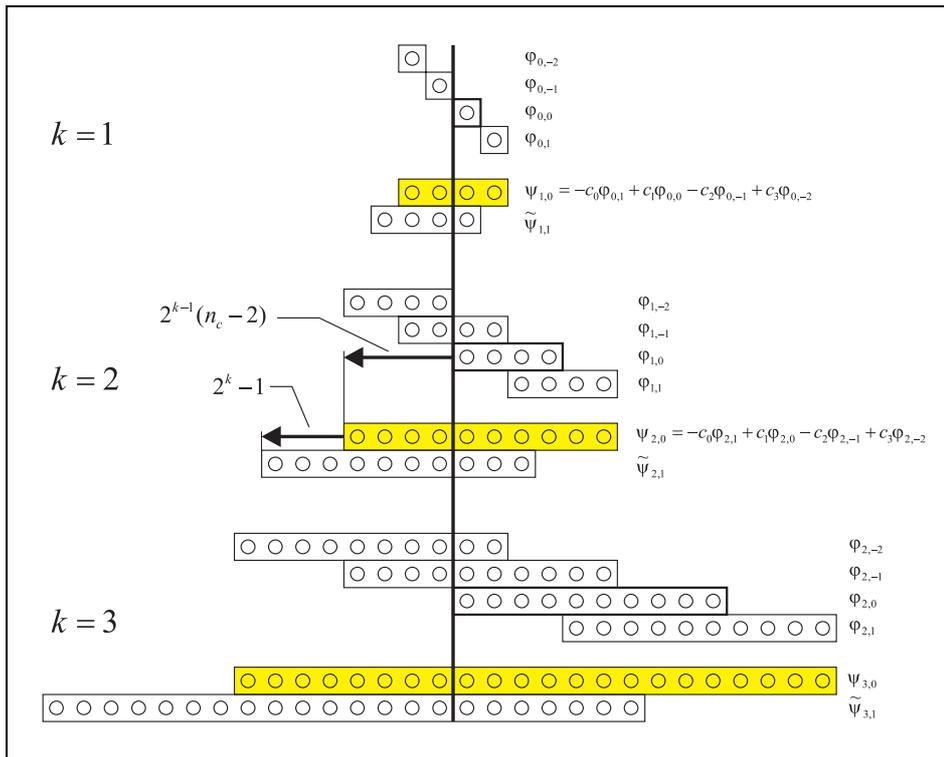


Abbildung 3.3-2: Schematische Darstellung der rekursiven Berechnung der Wavelets mit vier Filterkoeffizienten c_0, c_1, c_2, c_3 (z.B. Daubechies4-Wavelet) in den ersten drei Skalen $k = 1, 2, 3$.

3.3.2 Konstruktion der zweidimensionalen Wavelets

Der Tensorproduktansatz ist eine Möglichkeit, aus einer eindimensionalen Waveletbasis eine zweidimensionale Waveletbasis zu konstruieren. Diese Konstruktion wird in den üblichen Verfahren der diskreten Wavelettransformation, vgl. z.B. BUCKHEIT et al. (1995) angewendet. Ist die eindimensionale Basis durch das Paar Hochpaß (Wavelet, Mutterfunktion)

$$\underline{h} := \psi = \left[v_1^h \quad v_2^h \quad \dots \quad v_s^h \right]^\top \tag{3.3-3}$$

und Tiefpaß (Skalierungs-, Vaterfunktion)

$$\underline{t} := \phi = \left[v_1^t \quad v_2^t \quad \dots \quad v_s^t \right]^\top, \tag{3.3-4}$$

gegeben, ergibt sich eine zweidimensionale Waveletbasis als Quadrupel HH, HT, TH, TT durch die Matrizenprodukte

$$HH = \underline{h}\underline{h}^\top, \quad HT = \underline{h}\underline{t}^\top, \quad TH = \underline{t}\underline{h}^\top, \quad TT = \underline{t}\underline{t}^\top. \tag{3.3-5}$$

Diese Konstruktion der zweidimensionalen Waveletbasis führt damit gewissermaßen auf eine Hintereinanderausführung zweier eindimensionaler Wavelettransformationen (siehe 3.5).

Beispiel 2 Haar-Wavelet

Für das Haar-Wavelet ergeben sich aus den eindimensionalen Hoch- und Tiefpaßvektoren

$$\underline{h} = \frac{1}{2}\sqrt{2} \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \underline{t} = \frac{1}{2}\sqrt{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{3.3-6}$$

die Tensorproduktwavelets

$$\begin{aligned} HH = \underline{h}\underline{h}^\top &= \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, & HT = \underline{h}\underline{t}^\top &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, \\ TH = \underline{t}\underline{h}^\top &= \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}, & TT = \underline{t}\underline{t}^\top &= \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}. \end{aligned} \quad (3.3-7)$$

3.4 Verfahren der diskreten Wavelettransformation

3.4.1 Verfahren der 1D-Wavelettransformation

Im Rahmen dieser Arbeit werden zwei wesentliche Verfahren der diskreten Wavelettransformation behandelt, die hier als die übliche Schnelle Wavelettransformation (engl. fast wavelet transformation, FWT) und in Anlehnung an die Prozedurbezeichnung FWT_Stat des Wavelet-Tools WaveLab der Universität Stanford als Stationäre Wavelettransformation bezeichnet werden (siehe BUCKHEIT et al. (1995)).

Die Extraktion lokaler Eigenschaften des Signals aus den Waveletkoeffizienten mittels der Approximationseigenschaften setzt eine Lokalisierung der Waveletkoeffizienten bezüglich des Signals voraus. Es muß also zunächst geklärt werden, für welchen Signalbereich ein Waveletkoeffizient eine Aussage liefert. Da für den Support der Wavelets $s > 1$ gilt, ist eine scharfe Lokalisierung nicht möglich.

Die in 3.2 eingeführten Wavelets ergeben 2^m Waveletkoeffizienten für ein Signal der endlichen Länge $n = 2^m$. Bei der sogenannten Schnellen Wavelettransformation wird der Signalvektor $\underline{f} = [f_1 \ f_2 \ \cdots \ f_n]^\top$ in den Vektor $\underline{w} = [w_1 \ w_2 \ \cdots \ w_n]^\top$ der Waveletkoeffizienten mit

$$\begin{aligned} w_1 &= \langle \underline{f}, \underline{\phi}_{k,0} \rangle \\ w_{2^{j-1}+i} &= \langle \underline{f}, \underline{\psi}_{k+1-j,i-1} \rangle, \quad j = 1, 2, \dots, k, \quad i = 1, 2, \dots, 2^{j-1} \end{aligned} \quad (3.4-1)$$

überführt. In Abb. 3.4-1 ist das Prinzip des Verfahrens für ein Wavelet mit zwei Koeffizienten schematisch dargestellt. Dieses Verfahren entspricht dem Baumalgorithmus gemäß (3.2-11). Nach dem Baumalgorithmus wird das Signal schrittweise mit dem Wavelet der ersten Skale

$$\underline{\psi} = [v_1 \ v_2 \ \cdots \ v_{n_c}]^\top \quad \text{mit} \quad v_i = (-1)^{n_c+1-i} c_{n_c-i} \quad (i = 1, \dots, n_c)$$

und der Skalierungsfunktion

$$\underline{\phi} = [c_0 \ c_1 \ \cdots \ c_{n_c-2} \ c_{n_c-1}]^\top$$

zerlegt, wobei die c_i die n_c Koeffizienten der Skalierungsgleichung (3.2-7) sind. In der ersten Skale gilt für den Support s der Wavelets $s = n_c$. Für vier Koeffizienten c_i hat man gerade

$$\underline{\psi} = [c_3 \ -c_2 \ c_1 \ -c_0]^\top \quad \text{und} \quad \underline{\phi} = [c_0 \ c_1 \ c_2 \ c_3]^\top.$$

In Abb. 3.4-2 ist diese schrittweise Berechnung schematisch dargestellt. Die temporär mit den Skalierungsfunktionen berechneten Waveletkoeffizienten w'_i und w''_i sind dunkel hervorgehoben.

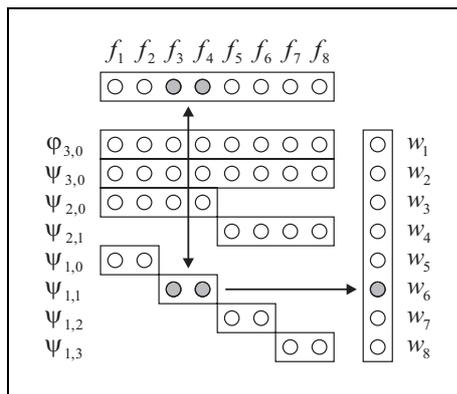


Abbildung 3.4-1: Schematische Darstellung der schnellen WT für ein Wavelet mit zwei Koeffizienten (z.B. Haar-Wavelet). Die Waveletkoeffizienten w_i werden als Skalarprodukt eines Wavelets $\psi_{k,l}$ mit dem Signalvektor f berechnet. Hervorgehoben ist die Berechnung von w_6 .

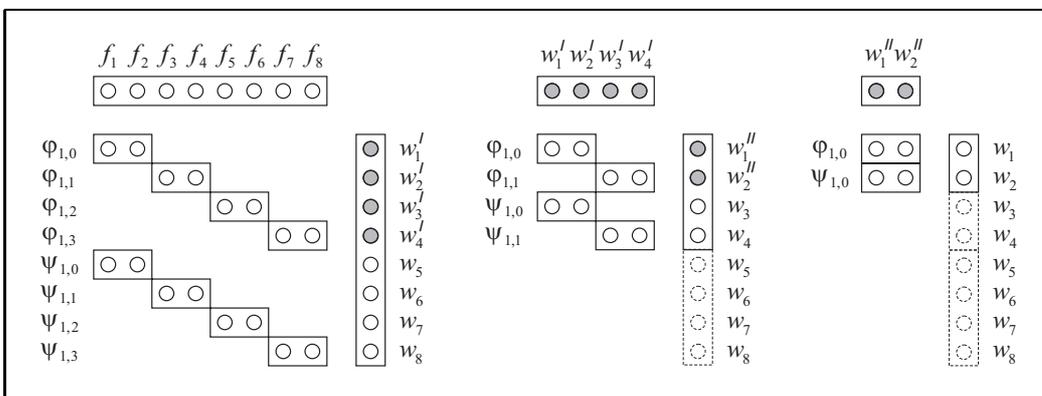


Abbildung 3.4-2: Schematische Darstellung der Berechnung der Waveletkoeffizienten mit dem Baumalgorithmus für ein Wavelet mit zwei Koeffizienten (z.B. Haar-Wavelet). Die Waveletkoeffizienten w'_i und w''_i (dunkel hervorgehoben), die als Skalarprodukt der Skalierungsfunktion $\phi_{k,l}$ mit dem Signalvektor f berechnet wurden, werden jeweils weiter zerlegt.

Die Berechnung der Waveletkoeffizienten durch die schrittweise Zerlegung in Hochpaß und Tiefpaß nach dem Baumalgorithmus und die direkte Berechnung über die Wavelets der einzelnen Skalen sind äquivalent.

Die schnelle Wavelettransformation eignet sich vorwiegend zur Datenkompression, da nur die minimale Anzahl an Waveletkoeffizienten berechnet wird, die zur Rekonstruktion erforderlich sind. Es ist ersichtlich, daß eine Signalanalyse im Sinne der in der Einleitung genannten Zielstellung nicht zufriedenstellend möglich ist. Man erhält in der ersten Skale nur in der Hälfte der Positionen (vier im Beispiel Abb. 3.4-1), d.h. nur in der doppelten Tastweite, eine Information über die Ableitung des Signals, in der zweiten Skale sogar nur noch in einem Viertel der Positionen usw.

Für die Extraktion von lokalen Eigenschaften eignet sich ein Verfahren, bei dem die Wavelets zur Berechnung der Koeffizienten in alle Skalen mit der Schrittweite $\Delta x = \Delta y = 1$ über das Signal geschoben werden. Diese Wavelets $\tilde{\psi}$ entstehen aus der Basis $\{\psi\}$ durch die Zuordnung $\tilde{\psi}_{k,2^k} = \psi_{k,0}$ (vgl. Abb. 3.3-2). Durch fortlaufende Verschiebung um Eins werden diese Wavelets zur Waveletmenge $\{\tilde{\psi}_{k,1}, \tilde{\psi}_{k,2}, \dots, \tilde{\psi}_{k,2^m}\}$ vervollständigt. Bei diesem Verfahren werden für jede

Skale k die $n = 2^m$ Waveletkoeffizienten

$$w_{k,i} = \langle f, \psi_{k,i} \rangle \quad i = 1, \dots, n \quad (3.4-2)$$

berechnet. Es wird auch als stationäre Wavelettransformation bezeichnet (BUCKHEIT et al. (1995)). Die Zuordnung der Waveletkoeffizienten zu den Signalbereichen wird für ein Wavelet mit vier Skalierungskoeffizienten c_i in Abb. 3.4-3 beispielsweise illustriert⁶. Der Waveletkoeffi-

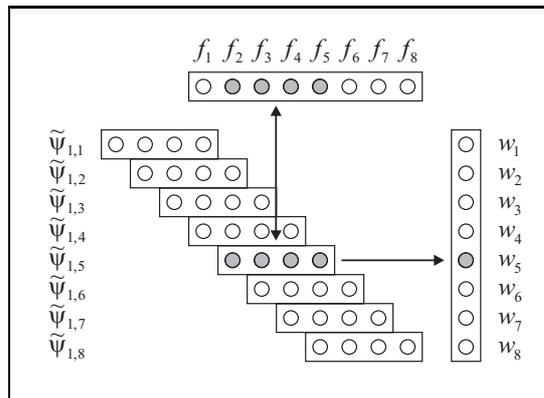


Abbildung 3.4-3: Schematische Darstellung der stationären WT für ein Wavelet mit vier Koeffizienten (z.B. Daubechies4-Wavelet) in der 1. Skale. Die Berechnung von w_5 ist hervorgehoben.

zient w_5 z.B. liefert damit bei einem Wavelet mit vier Koeffizienten eine aus den Signalwerten $\{f_2, f_3, f_4, f_5\}$ „gemittelte“ lokale Information, die bei einer unmittelbaren Zuordnung nach dieser Indizierung auf den Signalwert f_5 lokalisiert wird. In der zweiten Skale (hier nicht illustriert) liefert z.B. der Waveletkoeffizient w_9 eine Information, die sich aus den Signalwerten $\{f_2, f_3, \dots, f_{11}\}$ ergibt und auf den Signalwert f_9 lokalisiert wird. Es ist anzunehmen, daß es dabei zu einer Verschiebung der lokalen Information kommt, die bei der Auswertung unbedingt berücksichtigt werden muß. Diese Verschiebung wird im Abschnitt 7.2 untersucht.

3.4.2 Verfahren der 2D-Wavelettransformation

Bei der zweidimensionalen Wavelettransformation mittels Tensorproduktwavelets ist meist nur die schnelle Wavelettransformation implementiert, die die Signalmatrix auf eine Matrix von Waveletkoeffizienten (WK) der gleichen Dimension abbildet. Analog zur Anordnung der Waveletkoeffizienten gemäß (3.4-1) im Vektor \underline{w} (vgl. Abb. 3.4-1) ist die WK-Matrix aus den Teilmatrizen W_i^{HH} , W_i^{HT} , W_i^{TH} ($i = 1, \dots, k$) und W_k^{TT} zusammengesetzt (Abb. 3.4-4).

Die zweidimensionale stationäre Wavelettransformation liefert bei m Skalen $m + 1$ Matrizen, was zu einer enormen Vergrößerung des Datenumfanges führt. Für die Signalanalyse bestehen aber bei der schnellen Wavelettransformation die gleichen Probleme, die bereits bei der eindimensionalen Wavelettransformation erwähnt wurden. Anhand des in der Literatur oft erwähnten Beispiels eines Quadrates mit seinen beiden Diagonalen (z. B. Datensatz 'BoxWithCross' im Tool Wavelab) kann dies leicht illustriert werden (siehe Anhang A).

Bei der schnellen Wavelettransformation wird die $(2^k, 2^k)$ -Matrix der Höhen in eine $(2^k, 2^k)$ -Matrix von Waveletkoeffizienten transformiert. Die schrittweise Zerlegung in Hochpaß und Tiefpaß erfolgt analog zum Baumalgorithmus (Abb. 3.2-1).

⁶ Im weiteren Text werden auch die Wavelets für die stationäre WT zur Vereinfachung ohne Tilde geschrieben.

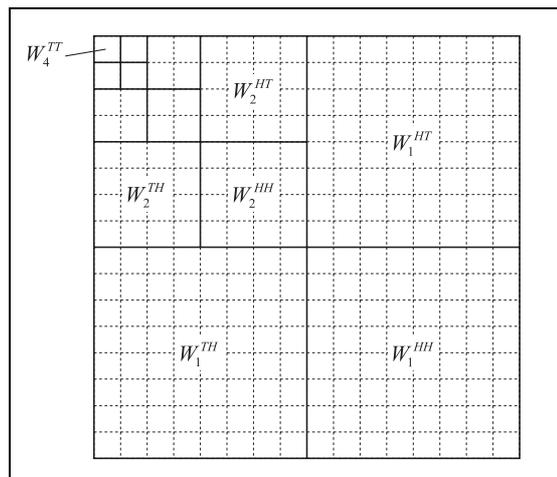


Abbildung 3.4-4: Datenstruktur der WK bei der zweidimensionalen schnellen WT

Zur Berechnung der Waveletkoeffizienten der 1. Skale wird das Wavelet der 1. Skale mit der Schrittweite $\Delta x = 2^1$ in beiden Koordinatenrichtungen über die Signalmatrix geschoben. Das wird in Abb. 3.4-5 am Beispiel des Haarwavelets illustriert. Da der Support s des Haarwavelets der 1. Skale gleich der Schrittweite Δx der Wavelet-Translation ist ($\Delta x = s = 2$), wird die Signalmatrix durch die verschobenen Wavelets exakt überdeckt. Ist der Support größer als die

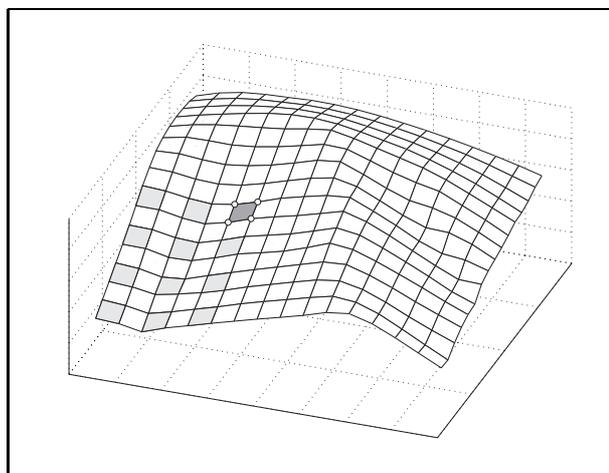


Abbildung 3.4-5: Teilgitterquadrate, die von den Haarwavelets der ersten Skale überdeckt werden.

Schrittweite (z. B. $s = 4$ beim Daubechies4-Wavelet), wird die Signalmatrix i. allg. durch zyklische Fortsetzung modifiziert. Die zyklische Fortsetzung wird realisiert, indem die Signalmatrix S in alle Richtungen, in denen die Wavelets über die Signalmatrix hinausragen, als S_Z zyklisch wiederholt wird (Abb. 3.4-6).

Jedes Skalarprodukt der vier Tensorproduktwavelets (vgl. Abschnitt 2) mit der Signalmatrix erfaßt damit s^2 Höhenwerte (beim Haar-Wavelet $2 \cdot 2 = 4$, Abb. 3.4-5). Man erhält für jede Waveletposition vier Waveletkoeffizienten: w^{HH} , w^{HT} , w^{TH} , w^{TT} , die in vier Matrizen der Dimension $(2^{k-1}, 2^{k-1})$ abgebildet werden. Im Interesse einer praktisch handhabbaren Datenstruktur werden die Matrizen W^{HH} , W^{HT} und W^{TH} als Teilmatrizen einer $(2^k, 2^k)$ -Matrix in der in Abb. 5.1-1 dargestellten Weise angeordnet.

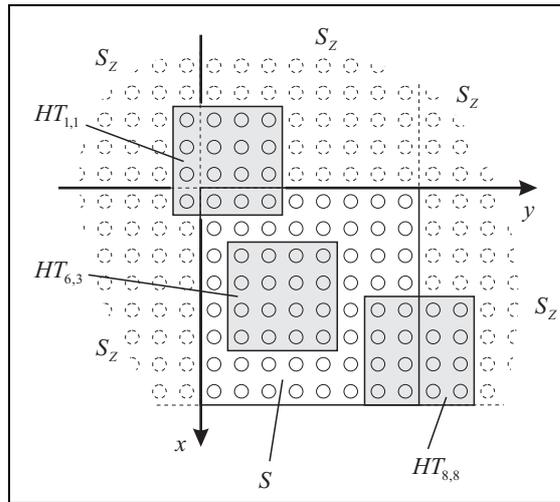


Abbildung 3.4-6: Notwendige zyklische Fortsetzung der Signalmatrix wegen des Überragens der Wavelets. Die Indizierung der beispielhaft eingezeichneten Daubechies4-Wavelets entspricht der stationären WT.

Die Matrix W^{TT} wird erneut der Wavelettransformation mit den Wavelets der 1. Skale unterworfen und liefert wiederum vier Matrizen, diesmal der Dimension $(2^{k-2}, 2^{k-2})$. Auf die Originalmatrix angewendet ist das gleichbedeutend mit einer Wavelettransformation mit den Wavelets der 2. Skale, die mit der Schrittweite $\Delta x = 2^2 = 4$ über die Signalmatrix geschoben werden (Die Schrittweite in der k -ten Skale ist $\Delta x = 2^k$). Im Beispiel des Haar-Wavelets (Support $s = 4$ in der 2. Skale) werden dann beim Skalarprodukt jeweils $4 \cdot 4 = 16$ Höhenwerte erfaßt, und es ergeben sich nur noch 16 Waveletkoeffizienten, die in je eine $(2^{k-2}, 2^{k-2})$ -Matrix abgebildet werden.

Diese sukzessive Wavelettransformation wird so lange fortgesetzt, bis die vier Matrizen die Dimension $(1, 1)$ haben. Letztendlich erhält man eine Matrix der Struktur, wie sie in Abb. 5.1-1 zu sehen ist.

3.5 Wavelettransformation einer zweidimensionalen Polynomfunktion

3.5.1 Wavelettransformation einer Funktion $z = xy$

Nach 3.3.2 erhält man aus der eindimensionalen Waveletbasis h (Hochpaß) und t (Tiefpaß) die zweidimensionalen Tensorprodukt-Wavelets HH , HT , TH und TT .

Mit den diskreten Argumentvektoren

$$\begin{aligned} \underline{x} &= [x_1 \ x_2 \ \cdots \ x_n]^\top, \\ \underline{y} &= [y_1 \ y_2 \ \cdots \ y_n]^\top \end{aligned}$$

ergibt sich aus der Funktion $z = xy$ die Signalmatrix

$$Z = \underline{xy}^\top. \tag{3.5-1}$$

Die Waveletkoeffizienten werden als Skalarprodukt⁷ der translatierten Tensorprodukt-Wavelets HH , HT , TH und TT mit der Signalmatrix berechnet. In den folgenden Betrachtungen (beispielsweise für das Wavelet HH) beziehen sich die Indizes der Argumentvektoren auf den vom Wavelet $HH_{k,l}$ überdeckten Ausschnitt $Z_{k,l}$ der Signalmatrix (vgl. Abb. 3.5-1). Man erhält für

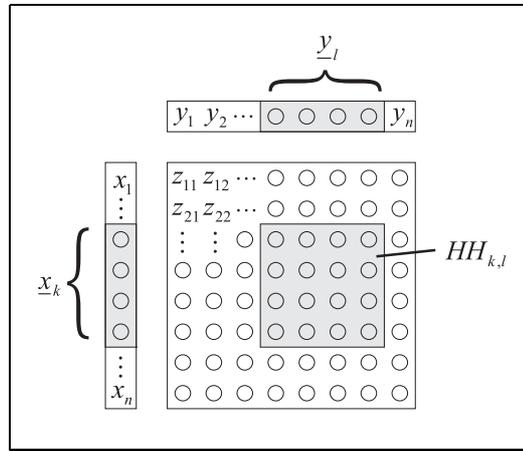


Abbildung 3.5-1: Schematische Darstellung des Skalarprodukts $\langle HH_{k,l}, Z_{k,l} \rangle$.

das Skalarprodukt von HH mit der Signalmatrix

$$\begin{aligned}
 w_{k,l} &= \langle HH_{k,l}, Z_{k,l} \rangle \\
 &= \sum_{i,j=1}^s h_i h_j \cdot x_i y_j = \sum_{i,j=1}^s h_i x_i \cdot h_j y_j = \sum_{i=1}^s h_i x_i \cdot \sum_{j=1}^s h_j y_j \\
 &= \langle \underline{h}, \underline{x}_k \rangle \cdot \langle \underline{h}, \underline{y}_l \rangle .
 \end{aligned} \tag{3.5-2}$$

Damit ist die zweidimensionale Wavelettransformation auf das Produkt der Wavelettransformationen der eindimensionalen Signale \underline{x} und \underline{y} zurückgeführt.

Analog gilt:

$$\begin{aligned}
 \langle HT_{k,l}, Z_{k,l} \rangle &= \langle \underline{h}, \underline{x}_k \rangle \cdot \langle \underline{t}, \underline{y}_l \rangle , \\
 \langle TH_{k,l}, Z_{k,l} \rangle &= \langle \underline{t}, \underline{x}_k \rangle \cdot \langle \underline{h}, \underline{y}_l \rangle , \\
 \langle TT_{k,l}, Z_{k,l} \rangle &= \langle \underline{t}, \underline{x}_k \rangle \cdot \langle \underline{t}, \underline{y}_l \rangle .
 \end{aligned} \tag{3.5-3}$$

3.5.2 Wavelettransformation einer Polynomfunktion $z = P(x, y)$

Die oben beschriebene Zerlegung einer zweidimensionalen Wavelettransformation in das Produkt zweier eindimensionaler Wavelettransformation kann auch für eine allgemeine Polynomfunktion formuliert werden.

Sei $z = f(x, y)$ eine zweidimensionale Polynomfunktion

$$z = \sum_{i,j} a_{ij} \cdot x^i y^j \tag{3.5-4}$$

⁷ Das hier verwendete Skalarprodukt zweier quadratischer Matrizen vom Format (n, n) ist definiert als $\langle A, B \rangle = \sum_{i,j=1}^n a_{i,j} b_{i,j}$

mit den diskreten Argumentvektoren

$$\begin{aligned}\underline{x} &= [x_1 \ x_2 \ \cdots \ x_n]^\top, \\ \underline{y} &= [y_1 \ y_2 \ \cdots \ y_n]^\top\end{aligned}$$

und den daraus abgeleiteten diskreten Potenzfunktionen⁸

$$\begin{aligned}\underline{x}^i &:= [x_1^i \ x_2^i \ \cdots \ x_n^i]^\top, \\ \underline{y}^i &:= [y_1^i \ y_2^i \ \cdots \ y_n^i]^\top.\end{aligned}$$

Dann erhält man die Matrix des zweidimensionalen diskreten Signals zu

$$Z = \sum_{i,j} a_{ij} \cdot \underline{x}^i \underline{y}^j{}^\top \quad (3.5-5)$$

und analog zu 3.5.1

$$\begin{aligned}\langle \mathbf{H}\mathbf{H}_{k,l}, Z \rangle &= \sum_{i,j} a_{ij} \langle \underline{h}, \underline{x}_k^i \rangle \cdot \langle \underline{h}, \underline{y}_l^j{}^\top \rangle, \\ \langle \mathbf{H}\mathbf{T}_{k,l}, Z \rangle &= \sum_{i,j} a_{ij} \langle \underline{h}, \underline{x}_k^i \rangle \cdot \langle \underline{t}, \underline{y}_l^j{}^\top \rangle, \\ \langle \mathbf{T}\mathbf{H}_{k,l}, Z \rangle &= \sum_{i,j} a_{ij} \langle \underline{t}, \underline{x}_k^i \rangle \cdot \langle \underline{h}, \underline{y}_l^j{}^\top \rangle, \\ \langle \mathbf{T}\mathbf{T}_{k,l}, Z \rangle &= \sum_{i,j} a_{ij} \langle \underline{t}, \underline{x}_k^i \rangle \cdot \langle \underline{t}, \underline{y}_l^j{}^\top \rangle.\end{aligned} \quad (3.5-6)$$

Durch diesen Zusammenhang zwischen ein- und zweidimensionaler Wavelettransformation werden für die in 7.1 und 7.2 durchgeführten Herleitungen nur die Momente der eindimensionalen Waveletbasis benötigt, die im Abschnitt 3.6 zusammengestellt werden.

3.6 Momente

Die Approximationseigenschaften der Wavelets stehen in einem unmittelbaren Zusammenhang mit ihren Momenten. Mit Hilfe der Momente der Wavelets und Skalierungsfunktionen (Hoch- und Tiefpaßfilter) lassen sich die in Kapitel 7 angegebenen Approximationsformeln übersichtlich und waveletunabhängig herleiten. Deshalb werden nachfolgend die benötigten Momente bereitgestellt.

Nach der allgemeinen Definition der Momente $m_k := \langle x^k, \psi \rangle$ berechnen sich die Momente der translatierten diskreten Wavelets der stationären Wavelettransformation mit dem Translationsparameter i ($i = 1, 2, \dots, 2^k$) zu

$$m_k(i) = \sum_{j=1}^s (j + i + d)^k v_j \quad (3.6-1)$$

⁸ Die Schreibweise der Vektoren \underline{x}^i bzw. \underline{y}^i ist symbolisch und nicht als Potenz im Sinne des Matrizenkalküls zu verstehen.

mit $\underline{\psi} = [v_1 \ v_2 \ \cdots \ v_s]^T$, wobei i durch $w(i) = \langle \psi_i, f \rangle$ definiert ist. Der Parameter d legt den Beginn der zyklischen Verschiebung fest und ist für Wavelet und Skalierungsfunktion verschieden. Zur Unterscheidung der Wavelets (Hochpaß) und der Skalierungsfunktionen (Tiefpaß), die im zweidimensionalen Fall erforderlich ist, werden die oberen Indizes h und t verwendet (siehe (3.3-1), (3.3-2)). Im einzelnen sind für die Untersuchung der Approximationseigenschaft der diskreten Wavelettransformation die folgenden Momente der Wavelets wichtig.

0. Moment (Mittelwert):

$$\boxed{m_0^h(i) = \sum_{j=1}^s v_j^h =: m_0^h \quad (\text{unabhängig von } i \text{ und } d^h)} \quad (3.6-2)$$

Für ein Wavelet ist eine notwendige Voraussetzung $m_0^h = 0$. Das ist bei den folgenden Momenten bereits berücksichtigt.

1. Moment:

$$\begin{aligned} m_1^h(i) &= \sum_{j=1}^s (j + i + d^h) v_j^h \\ &= (i + d^h) \cdot \underbrace{\sum_{j=1}^s v_j^h}_{=m_0^h=0} + \sum_{j=1}^s j \cdot v_j^h \end{aligned}$$

$$\boxed{m_1^h(i) = \sum_{j=1}^s j \cdot v_j^h =: m_1^h \quad (\text{unabhängig von } i \text{ und } d^h)} \quad (3.6-3)$$

Je nach Ordnung des Wavelets ergibt sich für das

2. Moment⁹:

$$\begin{aligned} m_2^h(i) &= \sum_{j=1}^s (j + i + d^h)^2 v_j^h \\ &= (i + d^h)^2 \cdot \underbrace{\sum_{j=1}^s v_j^h}_{=m_0^h=0} + 2(i + d^h) \underbrace{\sum_{j=1}^s j \cdot v_j^h}_{=m_1^h} \\ &\quad + \underbrace{\sum_{j=1}^s j^2 \cdot v_j^h}_{=: \tilde{m}_2^h} \end{aligned}$$

$$\boxed{m_2^h(i) = 2(i + d^h)m_1^h + \tilde{m}_2^h} \quad (3.6-4)$$

Ist die Ordnung des Wavelets > 1 , so ist auch $m_1^h = 0$ und es gilt für das 2. Moment speziell

$$\boxed{m_2^h(i) = \sum_{j=1}^s j^2 \cdot v_j^h =: m_2^h \quad (\text{unabhängig von } i \text{ und } d^h)} \quad (3.6-5)$$

⁹ Der „Basis“-Term des i -ten Moments wird hier im Interesse einer kurzen Schreibweise mit \tilde{m}_i bezeichnet. Er ist gleich dem Moment m_i , falls alle Momente m_j für $j < i$ verschwinden.

Für die nachfolgenden Betrachtungen benötigen wir für ein Wavelet 2. Ordnung (d.h. $m_1^h = 0$, $m_2^h \neq 0$) noch das

3. Moment:

$$\begin{aligned}
m_3^h(i) &= \sum_{j=1}^s (j+i+d^h)^3 v_j^h \\
&= (i+d^h)^3 \cdot \underbrace{\sum_{j=1}^s v_j^h}_{=m_0^h=0} + 3(i+d^h)^2 \underbrace{\sum_{j=1}^s j \cdot v_j^h}_{=m_1^h=0} \\
&\quad + 3(i+d^h) \underbrace{\sum_{j=1}^s j^2 \cdot v_j^h}_{=m_2^h} + \underbrace{\sum_{j=1}^s j^3 \cdot v_j^h}_{=: \tilde{m}_3^h} \\
\boxed{m_3^h(i) = 3(i+d^h)m_2^h + \tilde{m}_3^h} & \tag{3.6-6}
\end{aligned}$$

Für den zweidimensionalen Fall werden zusätzlich noch die Momente der Tiefpaßfilter benötigt. Diese werden analog bestimmt.

Für die in dieser Arbeit durchgeführten Untersuchungen sind das 0. und das 1. Moment der Tiefpaßfilter wichtig, wobei das 0. Moment nicht wie beim Wavelet nach Voraussetzung verschwindet.

0. Moment:

$$\boxed{m_0^t(i) = \sum_{j=1}^s v_j^t =: m_0^t \quad (\text{unabhängig von } i \text{ und } d^t)} \tag{3.6-7}$$

1. Moment:

$$\begin{aligned}
m_1^t(i) &= \sum_{j=1}^s (j+i+d^t)v_j^t \\
&= (i+d^t) \cdot \underbrace{\sum_{j=1}^s v_j^t}_{=m_0^t} + \underbrace{\sum_{j=1}^s j \cdot v_j^t}_{=: \tilde{m}_1^t} \\
\boxed{m_1^t(i) = (i+d^t) \cdot m_0^t + \tilde{m}_1^t} & \tag{3.6-8}
\end{aligned}$$

4 Digitale Geländemodelle

4.1 Modelle für Kurven und Flächen

Digitale Geländemodelle (DGM) stellen die geometrische Basis für raumbezogene Geoinformationssysteme (GIS) dar. Der Begriff kann weiter gefaßt werden. Analog zu Höhenmodellen können auch Modelle mit nichttopographischen Informationen, z. B. Temperaturverteilungen, Geoidhöhen, die Bebauungsdichte usw. betrachtet werden; BARTELME (2000). In dieser Arbeit geht es ausschließlich um die Geländeform. Die Ergebnisse lassen sich jedoch bei Vorliegen der getroffenen Voraussetzungen, die abstrakt-mathematischer Natur sind, ohne daß in dieser Arbeit an jeder Stelle explizit darauf hingewiesen wird, auf allgemeinere Modelle im Sinne von BARTELME (2000) übertragen.

Digitale Geländemodelle entstehen aus Primärdaten: 3D-Koordinaten von ausgewählten Geländepunkten, die durch terrestrische Geländeaufnahme, photogrammetrische Auswertung oder Laserscanning gewonnen werden. Das können Punkte von Höhenlinien, Bruchkanten und Formlinien, charakteristische Einzelpunkte oder, wie beim Laserscanning, mehr oder weniger zufällige Punkte der Geländeoberfläche sein. Zu Daten in einer vernünftigen, handhabbaren Datenstruktur (Sekundärdaten) gelangt man durch ein geeignetes Approximationsverfahren, wie es z. B. im Programmsystem SCOP, IPF (1991) enthalten ist.

Im engeren Sinne ist das Gelände eine Fläche, die sich umkehrbar eindeutig durch Projektion in die Ebene abbilden läßt. Felsüberhänge, Bauten mit senkrechten Wänden und weitere derartige Geländeformen, die sich einer eindeutigen Abbildung verschließen, sind dabei ausgeschlossen. Mit diesen, für die meisten praktischen Anwendungen zulässigen, Einschränkungen läßt sich die Lokalisierung der Geländepunkte in Lage (x, y - Koordinaten) und Höhe (z - Koordinate) trennen. Man spricht dann von einer sogenannten $2\frac{1}{2}$ D- Modellierung. Das hier untersuchte digitale Höhenmodell (DHM) ist ein solches $2\frac{1}{2}$ D- Modell.

Die Zielstellung der vorliegenden Arbeit ist, Anwendungsmöglichkeiten der Wavelettransformation in der Geoinformatik zu untersuchen. Der Schwerpunkt liegt dabei auf der Beschreibung hybrider Geländemodelle.

Ein reguläres Rechteckgitter erlaubt die Verwendung einer Matrixstruktur, die zu sehr effizienten Algorithmen für die Verwaltung und Bearbeitung des Geländemodells führt. In einem regulären Gitter lassen sich charakteristische Geländelinien zunächst nicht unterbringen. Sie können mittels einer zusätzlichen Datenstruktur für 3D-Einzelpunkte und 3D-Punktfolgen explizit in das Modell aufgenommen werden (vgl. z. B. Programmsystem SCOP, IPF (1991)). Darunter leidet jedoch im Hinblick auf die beabsichtigte Wavelettransformation die Effektivität des Modells, die durch die Matrixstruktur gegeben war.

Mit irregulären Gittern (engl. triangulated irregular network, TIN) läßt sich eine höhere Qualität der Geländebeschreibung erreichen, wenn charakteristische Geländelinien (insbesondere Bruchkanten) als Gitterlinien und charakteristische Einzelpunkte als Gitterpunkte erhalten bleiben. Diese führen jedoch auf komplexere Datenstrukturen mit wesentlich höherem Aufwand in der Datenorganisation.

Hier wird ein anderer Weg zur Beschreibung hybrider Geländemodelle, nämlich auf der Basis eines regulären Gitters beschritten. Die Geländelinien sollen ebenfalls in einer Matrixstruktur beschrieben werden. Die Lösung dieses Problems ist dem Hauptziel der Arbeit, hybride Geländemodelle einer diskreten Wavelettransformation zugänglich zu machen, untergeordnet.

Geoinformationssysteme befinden sich zur Zeit in einer Phase stürmischer Entwicklung. In einer solchen Zeit bleibt es nicht aus, daß die Begriffsbildung sehr uneinheitlich vonstatten geht. Das liegt einerseits daran, daß in verschiedensten Forschungszentren an ähnlichen Themen gearbeitet wird und nicht immer die Möglichkeit zu einer Abstimmung oder gar Normung besteht. Zum anderen werden Begriffe, oftmals entlehnt aus allgemeinverständlichen Begriffen, für neue Sachverhalte benutzt und bekommen ihre spezielle Bedeutung im Zusammenhang mit dem jeweiligen Forschungsgegenstand. Solange es keine Überschneidungen der Forschungsrichtungen gibt, ergeben sich keine Widersprüche. Diese Überschneidungen sind aber bei interdisziplinärer Forschung kaum zu vermeiden. Es ist deshalb erforderlich, an dieser Stelle die in dieser Arbeit verwendeten Begriffe zu klären, um Mißverständnissen vorzubeugen. Insbesondere geht es dabei um Begriffe im Zusammenhang mit hybriden Modellen.

In der Literatur wird der Begriff hybrid im Zusammenhang mit DGM in zwei verschiedenen Bedeutungen verwendet. Einmal im Zusammenspiel zwischen Vektor- und Rasterdaten (BARTELME (2000), S.121ff) und zum anderen zwischen Flächen- und Liniendaten. In der vorliegenden Arbeit steht die zweite Version im Vordergrund (KRAUS (2000), S.181). Es soll deshalb für die hier betrachtete regelmäßige Datenstruktur nicht die auch oft verwendete Bezeichnung Raster sondern konsequent die Bezeichnung Gitter verwendet werden. Nachfolgend werden auf der Basis dieser grundsätzlichen Entscheidung die Detailbegriffe definiert.

Digitale Höhenmodelle basieren meist (im Interesse einer einfachen Datenstruktur) auf einem rechteckigen Gitter in einem ebenen Koordinatensystem. Werden die Gitterlinien gleichabständig gewählt, können die Höhenwerte in einer Matrix abgelegt werden, und es genügt für die Lage- definition die Angabe der x, y -Koordinaten eines Punktes und der beiden Gitterweiten in x - und y -Richtung. Im Falle eines quadratischen Gitters hat man sogar nur eine Gitterweite zu speichern (Abb. 4.1-1).

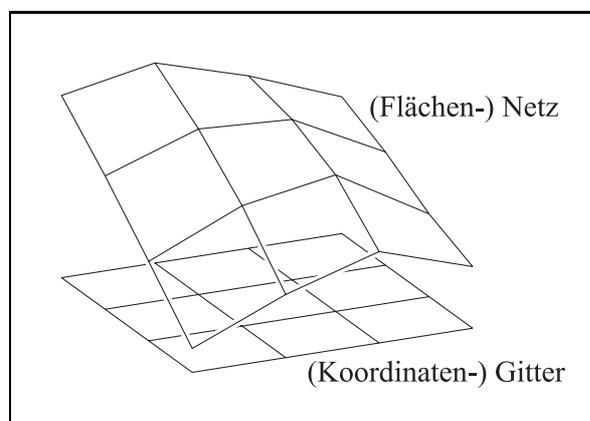


Abbildung 4.1-1: Darstellung einer Gelände- fläche als reguläres Netz über einem regulären Gitter.

Der Bereich, der durch zwei zueinander senkrechte Paare jeweils benachbarter *Gitterlinien* begrenzt wird, wird als (*Gitter-*) *Zelle* bezeichnet. Begrifflich soll zwischen den reinen Lage- beziehungen bezüglich eines (*Koordinaten-*) *Gitters* in der x, y -Ebene und der räumlichen Lage der Fläche durch Einbeziehung der Höhe als (*Flächen-*) *Netz* unterschieden werden. Die sich daraus ergebenden Bezeichnungen sind in Abb. 4.1-2 zu sehen.

Für die später zu untersuchende Transformation der Beschreibung von Geländelinien in eine Matrixstruktur sollen noch die folgenden Bezeichnungen eingeführt werden. Der der Gitterzelle entsprechende Bereich im Flächennetz heißt *Netzmasche*. Die Grenzlinie zwischen zwei benachbarten Zellen wird *Zellenseite* genannt. Analog heißen die Grenzlinie zwischen zwei benachbarten Maschen *Maschenseite* und die der Gitterlinie entsprechende Linie im Flächennetz *Netzlinie*. Die

Kreuzungspunkte der Gitterlinien, denen die Höheninformation der Fläche zugeordnet wird, heißen *Gitterpunkte* und die entsprechenden Punkte im Flächennetz *Netzpunkte*. Die Gitterpunkte sind in lokalen Betrachtungen gleichzeitig Eckpunkte der Zellen bzw. Endpunkte der Zellenseiten.

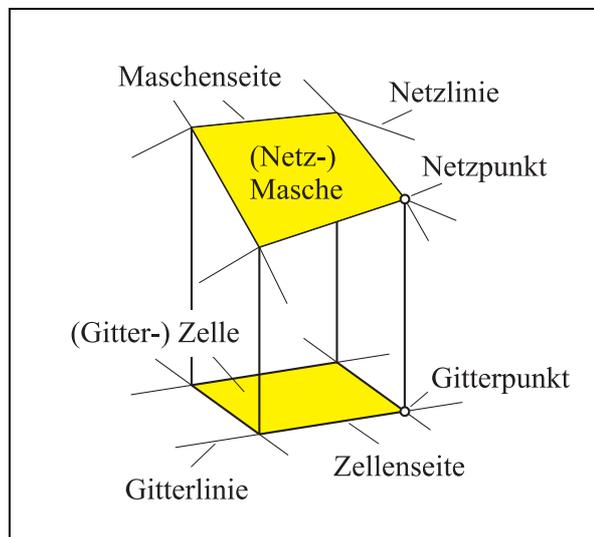


Abbildung 4.1-2: Begriffe in einem DGM.

Die Höhenwerte der Geländefläche werden den Gitterpunkten (den Schnittpunkten der Gitterlinien) zugeordnet. Die Höhen in den Gitterpunkten entstehen zumeist durch geeignete Interpolation aus gemessenen Höhen in Punkten, die i.allg. nicht mit den Gitterpunkten zusammenfallen. Dabei gehen jedoch die speziellen Informationen über Geländekanten und markante Höhenpunkte, falls sie in den ursprünglich erfassten Daten enthalten waren, verloren. Die Aufnahme dieser Zusatzinformationen führt auf eine Erweiterung des Modells um eine Datenebene mit einer anderen Struktur und damit auf ein hybrides Modell (Abb. 4.1-3).

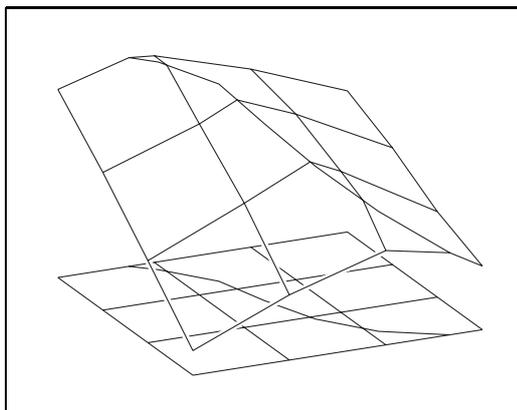


Abbildung 4.1-3: Hybrides digitales Geländemodell.

Dieses Modell besteht aus der oben genannten Matrixstruktur für die Beschreibung der Fläche und eine Listenstruktur (Folge von 3D-Koordinaten) für die Beschreibung der Linien als Polygonzüge (Abb. 4.1-4).

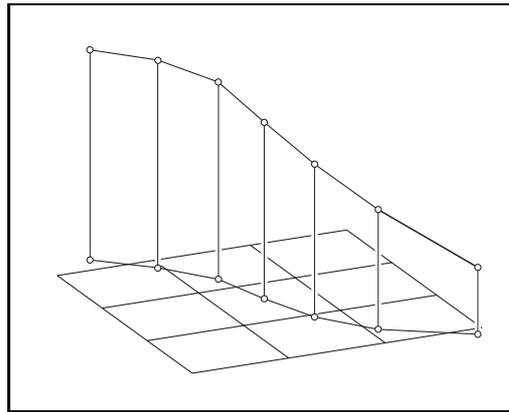


Abbildung 4.1-4: Darstellung einer Geländelinie über einem regulären Gitter. Die x, y -Koordinaten der Punkte der Polygonlinie liegen (i. allg.) nicht auf den Gitterpunkten. Die Matrixstruktur der Geländefläche ist deshalb zur Speicherung nicht geeignet.

Um ein solches hybrides Modell der Wavelettransformation zugänglich zu machen, kann ein Modell benutzt werden, das qualitativ eine Zwischenstufe zwischen dem reinen Gittermodell mit einfacher Matrixstruktur und einem hybriden Modell mit Matrixstruktur der Flächenbeschreibung und zusätzlicher Listenstruktur der Kantenbeschreibung darstellt.

Die Linien des hybriden Modells, die als Folge von 3D-Koordinaten in einer Listenstruktur vorliegen, können auch durch eine Folge von Schnittpunkten mit den *Gitterebenen* (*Gitterschnitte*, Abb. 4.1-5) beschrieben werden. In der Projektion sind das die Schnittpunkte mit den Gitterlinien. Für die spätere Visualisierung des hybriden Modells (Abb. 4.1-3) werden oft ohnehin diese Gitterschnitte benutzt. Damit hat man nur noch Punkte, die auf Gitterlinien liegen, aber noch zwischen den Gitterpunkten. Werden diese Gitterschnitte auf die benachbarten Gitterpunkte extrapoliert, liegen auch die Linien des hybriden Modells in einer Matrixstruktur vor.

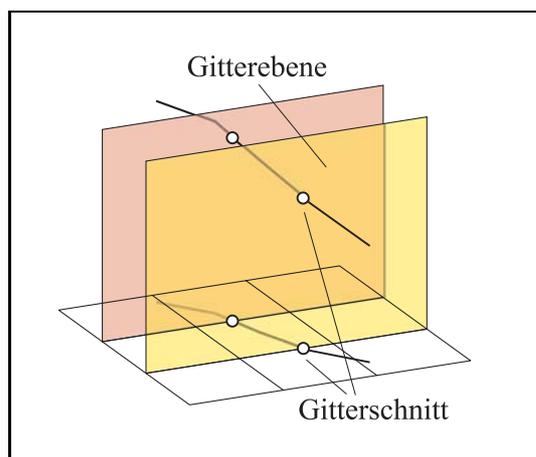


Abbildung 4.1-5: Die Gitterschnitte entstehen als Schnittpunkte des Kantenpolygons mit den Gitterebenen.

Die Beschreibung dieser Transformation der Listenstruktur einer Geländelinie in eine Matrixstruktur ist Gegenstand des Kapitels 4.3.2.

4.2 Neigung und Krümmung

Für viele Anwendungen im Geo-Bereich ist die Geländeneigung eine wichtige Größe (z.B. Abflußgebiete, Hochwasservorhersage, erosions- oder lawinengefährdete Hänge usw.); mitunter wird auch die Krümmung benötigt (BETHGE (1994)). Ist nun ein digitales Geländemodell (DGM) in Form von Waveletkoeffizienten abgespeichert, so könnten unter Ausnutzung der Approximationseigenschaft der Wavelettransformation der Gradient, bei Bedarf auch Wölbungs- und Krümmungsgrößen direkt aus den Waveletkoeffizienten bestimmt werden, ohne in den Originalbereich zurücktransformieren und dort numerisch differenzieren zu müssen.

Ist die Geländefläche F (Relief) durch eine Funktion $z = f(x, y)$ gegeben, so lassen sich aus den partiellen Ableitungen nach x und y differentialgeometrische Eigenschaften des Geländes, wie Neigung und Krümmung bestimmen. Aus den ersten partiellen Ableitungen z_x und z_y ergeben sich unmittelbar (vgl. Abb. 4.2-1) die Tangenteneinheitsvektoren in x - und y -Richtung zu

$$\underline{t}_x = \frac{1}{\sqrt{1 + z_x^2}} \begin{bmatrix} 1 \\ 0 \\ z_x \end{bmatrix}, \quad \underline{t}_y = \frac{1}{\sqrt{1 + z_y^2}} \begin{bmatrix} 0 \\ 1 \\ z_y \end{bmatrix}. \quad (4.2-1)$$

Der Flächennormaleneinheitsvektor ergibt sich zu

$$\underline{n}_F = \frac{1}{\sqrt{1 + z_x^2 + z_y^2}} \begin{bmatrix} -z_x \\ -z_y \\ 1 \end{bmatrix}. \quad (4.2-2)$$

Daraus läßt sich schließlich der Tangenteneinheitsvektor an die Gefällelinie g bestimmen:

$$\underline{v}_g = \frac{1}{\sqrt{(1 + z_x^2 + z_y^2)(z_x^2 + z_y^2)}} \begin{bmatrix} -z_x \\ -z_y \\ -z_x^2 - z_y^2 \end{bmatrix}. \quad (4.2-3)$$

Im weiteren sollen die wesentlichen Flächenkrümmungen als Krümmungen ausgezeichneter Flächenkurven bestimmt werden. Die Krümmung einer ebenen Kurve, die als Funktion $y = f(x)$ gegeben ist, wird bekanntlich nach

$$\kappa = \frac{y''}{(1 + y'^2)^{\frac{3}{2}}} \quad (4.2-4)$$

berechnet. Damit sind die Krümmungen der senkrechten Schnitte in x - und y - Richtung gegeben durch

$$\kappa_x = \frac{z_{xx}}{(1 + z_x^2)^{\frac{3}{2}}}, \quad \kappa_y = \frac{z_{yy}}{(1 + z_y^2)^{\frac{3}{2}}}. \quad (4.2-5)$$

BETHGE (1994) analysiert die Krümmungseigenschaften der ebenen Kurven in der Kartenprojektion. In Anlehnung an die dort dargestellten Zusammenhänge sollen hier einige weitere Beziehungen im Raum angegeben werden. Der Bezug auf den Neigungswinkel ψ des Gradienten im ebenen Koordinatensystem der Projektion ist im Gegensatz zur Zielstellung von BETHGE an dieser Stelle im dreidimensionalen Raum nicht zweckmäßig. BETHGE gibt u. a. eine Beziehung für die Krümmung κ_h der Höhenlinie h in einem Punkt $P(x, y, z)$ des Geländes in Abhängigkeit

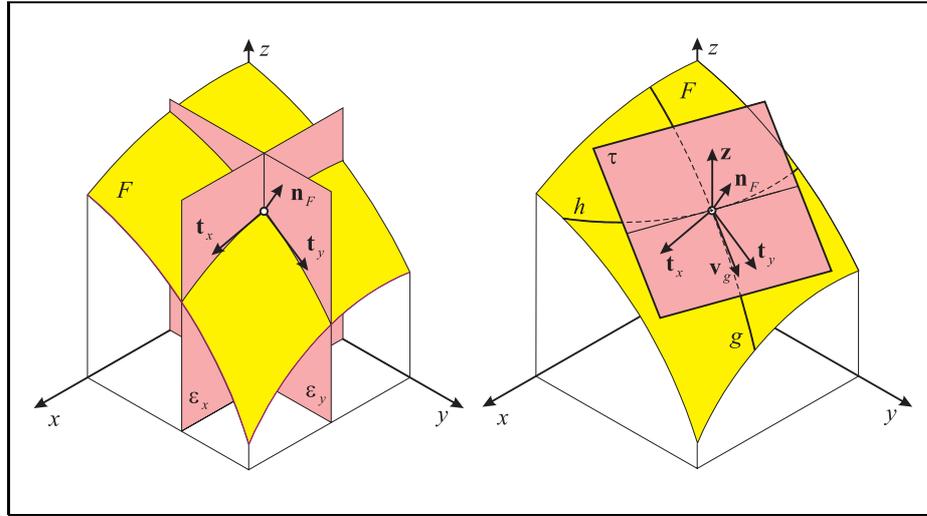


Abbildung 4.2-1: Differentialgeometrische Eigenschaften der Geländefläche F . Die Tangentenvektoren \mathbf{t}_x und \mathbf{t}_y an die senkrechten Schnitte ergeben sich unmittelbar aus den partiellen Ableitungen z_x und z_y . Über den Tangentenvektor \mathbf{v}_g erhält man die Neigung des Geländes.

von den partiellen Ableitungen der zweidimensionalen Flächenfunktion an. Mit den dortigen Bezeichnungen und der Vorgabe $\Delta\psi = \pi/2$ ergibt sich

$$\kappa_h = \frac{\dot{x}\ddot{y} - \dot{y}\ddot{x}}{(\dot{x}^2 + \dot{y}^2)^{3/2}} = -\frac{z_x^2 z_{yy} - 2z_x z_y z_{xy} + z_y^2 z_{xx}}{(z_x^2 + z_y^2)^{3/2}} \quad (4.2-6)$$

wegen

$$\begin{aligned} \dot{x} &= z_y & \dot{y} &= -z_x \\ \ddot{x} &= z_{xy}\dot{x} + z_{yy}\dot{y} & \ddot{y} &= -(z_{xx}\dot{x} + z_{xy}\dot{y}) \end{aligned} \quad (4.2-7)$$

Nach dem Satz von MEUSNIER gilt für die Krümmung der Schnittkurven der Geländefläche F mit Ebenen ε_i , die alle die gleiche Tangente an F enthalten,

$$\kappa_i = \frac{\kappa_n}{\cos \gamma_i}, \quad (4.2-8)$$

wobei ε_n die Normalschnittebene, κ_n die Normalschnittkrümmung und γ_i der Winkel zwischen ε_n und ε_i sind. Speziell für die Krümmung κ_s des senkrechten Schnitts und die Krümmung κ_h der Höhenlinie (waagerechter Schnitt) gilt

$$\begin{aligned} \kappa_h &= \frac{\kappa_n}{\cos \gamma_h}, \\ \kappa_s &= \frac{\kappa_n}{\cos \gamma_s}. \end{aligned} \quad (4.2-9)$$

Mit dem Einheitsvektor \mathbf{z} in z -Richtung gilt für den Winkel γ_s

$$\cos \gamma_s = \langle \mathbf{z}, \mathbf{n}_F \rangle = \frac{1}{\sqrt{1 + z_x^2 + z_y^2}} \quad (4.2-10)$$

und mit $\gamma_h + \gamma_s = \pi/2$ für den Zusammenhang zwischen κ_s und κ_h

$$\kappa_s = \kappa_h \sqrt{z_x^2 + z_y^2} \quad (4.2-11)$$

wegen

$$\cos \gamma_h = \sin \gamma_s = \sqrt{1 - \cos^2 \gamma_s} = \sqrt{\frac{z_x^2 + z_y^2}{1 + z_x^2 + z_y^2}}. \quad (4.2-12)$$

Zur Bestimmung der Normalschnittkrümmung in Richtung der Gefällelinie wird zweckmäßigerweise die Flächengleichung $z = f(x, y)$ im Koordinatensystem $(O; \mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3)$ in das in den Punkt $P = \hat{O}$ verschobene und um die z -Achse gedrehte Koordinatensystem $(\hat{O}; \hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3)$ mit der Richtung des Gradienten in P mit $\underline{p}^\top = [p_1 \ p_2 \ p_3]$ als neue x -Richtung transformiert (Abb. 4.2-2).

Mit der Basistransformation $\hat{\underline{e}} = A \cdot \underline{e}^{-1}$ erhält man die Koordinatentransformation

$$\begin{aligned} \underline{x} &= A^\top \hat{\underline{x}} + \underline{p}, \\ \hat{\underline{x}} &= A(\underline{x} - \underline{p}), \end{aligned} \quad (4.2-13)$$

die innerhalb der Normalschnittebene in Richtung der Gefällelinie auf die Kurve

$$\hat{z}(t) = f(x(t), y(t)) - p_3, \quad (4.2-14)$$

mit

$$\begin{aligned} x(t) &= \frac{z_x}{w} t + p_1, \\ y(t) &= \frac{z_y}{w} t + p_2. \end{aligned} \quad (4.2-15)$$

führt. Die Gefällelinienkrümmung, genauer die Krümmung des Normalschnitts in Richtung der

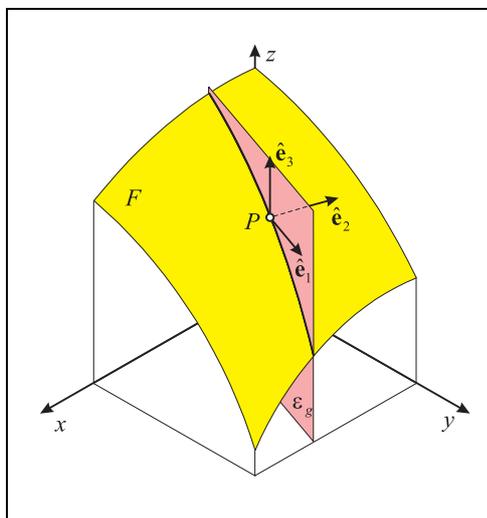


Abbildung 4.2-2: Transformation des Koordinatensystems in den senkrechten Schnitt in Richtung der Tangente an die Gefällelinie.

Gefällelinie, läßt sich dann bekanntlich nach (4.2-4) berechnen und man erhält

$$\kappa_g = \frac{z_{xx}z_x^2 + 2z_{xy}z_xz_y + z_{yy}z_y^2}{(z_x^2 + z_y^2)(1 + z_x^2 + z_y^2)^{\frac{3}{2}}}. \quad (4.2-16)$$

¹ Es sind $\hat{\underline{e}} := [\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2, \hat{\mathbf{e}}_3]^\top$ und $\underline{e} := [\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3]^\top$ die als Matrix zusammengefaßten Basisvektoren der beiden Koordinatensysteme

4.3 Raumkurven in Geländemodellen

Kurven kommen in digitalen Geländemodellen in verschiedensten Ausprägungen vor. Einerseits können Kurven unmittelbar flächenbezogen sein, das soll heißen, sie lassen sich aus der Geometrie der Gelände­fläche ableiten. Das sind z. B. Höhenlinien, Gerippelinien, Bruchkanten usw. — Im allgemeinen sind flächenbezogene Kurven lediglich eindimensionale Teilmengen der Gelände­fläche, z. B. Grenzen, Straßen usw. — Andererseits gibt es auch Kurven, die mehr oder weniger unabhängig von der Gelände­fläche existieren: Hochstraßen, Brücken, Seilbahnen usw. Es sind damit sowohl flächenbezogene als auch eigenständige Kurvenrepräsentationen von Interesse. Für die erforderliche Qualität der Kurvenbeschreibung ist der Zweck entscheidend. Es können ausschließlich topologische Informationen interessant sein, wie bei Liniennetzplänen, bei denen die Vernetzung und die Umsteigemöglichkeiten wesentlich sind. In diesem Fall wird keine geometrische Genauigkeit benötigt. Werden jedoch geometrische Informationen wie die Länge einer Straße oder der Inhalt einer durch eine Kurve begrenzten Fläche benötigt, ist die erreichbare geometrische Genauigkeit durch eine gewählte Beschreibung zu berücksichtigen. Das führt zu der Frage, ob bereits die topologische Struktur mehrerer in Wechselbeziehung stehender Kurven in Form eines Graphen, ein Polygonzug in Form einer Folge von Punkten mit geometrisch exakten Koordinaten oder erst eine funktionale Beschreibung (u. U. stückweise) die Genauigkeitsanforderungen erfüllt. Diese Fragen werden z. B. in BARTELME (2000) diskutiert.

In der vorliegenden Arbeit werden ausschließlich flächenbezogene Raumkurven als Polygonzüge betrachtet. Diese Einschränkung ergibt sich aus dem gewählten Anwendungsfall. Im folgenden Unterabschnitt, der sich zunächst mit der expliziten Kurvenbeschreibung befaßt, wäre die Einschränkung auf den Flächenbezug nicht notwendig.

Die Beschreibung von Raumkurven als Polygonzug ist die diskrete Version der Parameterdarstellung. Damit lassen sich Beziehungen zur funktionalen Beschreibung herstellen, die u. a. die Nutzung differentialgeometrischer Aussagen unter gewissen Voraussetzungen ermöglichen.

4.3.1 Explizite Beschreibung

Die geometrischen Gebilde Kurve und Fläche lassen sich unter gewissen Voraussetzungen als Funktion $y = f(x)$ bzw. $z = f(x, y)$ beschreiben. Das gilt insbesondere für Gelände­flächen als Objekte des dreidimensionalen Raumes, die auf Grund ihrer Natur normalerweise eine $2\frac{1}{2}$ -D-Beschreibung zulassen. Ausnahmen stellen z.B. Geländeüberhänge oder Höhlen dar.

Die diskrete Realisierung einer Funktion $z = f(x, y)$ sind in einer Matrix gespeicherte Höhenwerte, die zu Punkten eines gegebenen regulären Gitters gehören (DHM). Bei einer Ausdehnung von $n \times m$ Gitterpunkten sind das $n \cdot m$ skalare Daten. Die Topologie ist durch das Gitter und die Geometrie durch die Funktionswerte festgelegt. Das DHM kann damit unmittelbar einer zweidimensionalen Wavelettransformation unterworfen werden.

Ebene Kurven, die auf Grund ihrer Form (z.B. Höhenprofile) eine Funktionsbeschreibung $y = f(x)$ zulassen, können gitterbasiert mit Hilfe eines Gitters auf der x -Achse diskretisiert und die Funktionswerte in einem Vektor gespeichert werden. Bei einer Ausdehnung von n Gitterpunkten sind das n skalare Daten. Die Topologie ist wie bei den $2\frac{1}{2}$ -D-Flächen durch das Gitter und die Geometrie durch die Funktionswerte festgelegt.

Bei ebenen Kurven, die sich nicht in der Form $y = f(x)$ beschreiben lassen (z.B. Höhenlinien), ist die Topologie nicht durch ein (achsenparalleles) Gitter festgelegt, und die Geometrie ist damit auch nicht durch einen Funktionswertvektor über einem Gitter beschreibbar. Die Topologie muß über die Punktreihenfolge und die Geometrie durch die x - und y -Koordinaten festgelegt werden

(stetiges Analogon: Parameterdarstellung $x(t), y(t)$). Es handelt sich dabei um eine zweidimensionale vektorwertige Funktion von einer Veränderlichen. Für die eindimensionale Kurve werden damit zwei skalare Funktionen benötigt.

Um dieses Dilemma zu umgehen, werden in der Literatur für ebene Kurven verschiedene spezielle Beschreibungen (Repräsentationen) mit einer skalaren Funktion vorgeschlagen. Im wesentlichen handelt es sich dabei um Modifikationen einer Anstiegswinkelfunktion (Tangentenwinkelfunktion STOYAN und STOYAN (1992), Psi-s-Plot O'NEILL und MARK (1996), t-alpha-Repräsentation WERSCHLEIN (1996)). Um die Originalkoordinaten zu erhalten, ist eine Rückrechnung aus einer solchen speziellen Repräsentation erforderlich. Voraussetzung für eine solche Beschreibung ist eine äquidistante Punktfolge (stetiges Analogon: Parameterdarstellung bez. der Bogenlänge s , also $x(s), y(s)$). Weiterhin werden für eine eindeutige Rückrechnung zusätzlich Informationen über den Anfangspunkt benötigt.

Eine Raumkurve ist wie eine ebene Kurve eine eindimensionale Mannigfaltigkeit. Es gibt jedoch keine analoge Beschreibung durch nur eine skalare Funktion.

Bei den oben genannten Repräsentationen handelt es sich genau genommen um die Beschreibung der Kurve als Sehnenzug, bei der neben dem Anfangspunkt von jeder Sehne die Richtung und ihre Länge bekannt sein muß, um die Kurve rekonstruieren zu können. Damit hat man ebenso wie bei den $2n$ Informationen der Koordinatenfunktionen 2 Koordinaten des Anfangspunktes, $n - 1$ Sehnenrichtungen und $n - 1$ Sehnenlängen. Die Beschränkung auf nur eine Funktion wird durch die einheitliche Sehnenlänge bei äquidistanten Punkten erreicht. Bei ebenen Kurven läßt sich die Sehnenrichtung durch *einen* Winkel beschreiben (die Richtungen in der Ebene sind isomorph zum Einheitskreis, also eindimensional). Soll dieses Prinzip auf Raumkurven übertragen werden, so ist klar, daß mindestens zwei skalare Funktionen benötigt werden, da die Sehnenrichtung im Raum zwei Freiheitsgrade besitzt (die Richtungen im Raum sind isomorph zur Einheitskugel, also zweidimensional).

Da bei Raumkurven eine Reduktion der Beschreibung auf eine skalare Funktion ohnehin nicht möglich ist, werden die drei Koordinatenfunktionen als eine mögliche Repräsentation in die folgende vergleichende Untersuchung einbezogen. Als unmittelbare Verallgemeinerung der Tangentenwinkelfunktion bei ebenen Kurven ergibt sich die Beschreibung der Sehnenrichtung durch zwei Winkel. Eine weitere Beschreibung durch zwei Funktionen sind die aus der Differentialgeometrie als natürliche Gleichungen bekannten Funktionen Krümmung und Torsion, die die innere Geometrie der Kurve vollständig festlegen, KREYSZIG (1957). Diese Funktionen werden im folgenden auch allgemein als Signalfunktionen bezeichnet.

Koordinatenfunktionen

Gespeicherte Raumkurven liegen meist als Punktfolge (geordnete 3D-Daten, z.B. SCOP-Datensatz, IPF (1991)) vor. Für eine Analyse der Geometrie der Raumkurve durch eine Wavelettransformation fehlt noch eine geeignete Parametrisierung, wobei sich wegen der weiter oben dargelegten Gründe der natürliche Parameter Bogenlänge anbietet. Als einfachste Approximation der Bogenlänge der diskret gegebenen Kurve kann die Bogenlänge des Sehnenzuges genommen werden. Da die Punkte i.allg. nicht gleichabständig vorgegeben sind, ist mittels einer geeigneten Kurvenapproximation (z.B. Walking Divider Algorithmus, GOODCHILD und MARK (1987)) eine gleichabständige Punktfolge auf der Kurve zu ermitteln. Diese dient als diskrete Parameterdarstellung der Kurve mit äquidistanten Parameterwerten. Damit liegen drei Koordinatenfunktionen $x(s), y(s)$ und $z(s)$ mit äquidistanten Stützstellen vor. Jede der Koordinatenfunktionen kann einer Wavelettransformation unterworfen werden.

Tangentenwinkelfunktionen

In Analogie zur Tangentenwinkelfunktion bei ebenen Kurven kann für jeden Punkt der Kurve die Tangentenrichtung durch zwei Winkel angegeben werden. Naheliegend sind die Winkel λ und ϕ der Kugelkoordinaten, d.h. die Neigung der Projektion der Tangente in die x, y -Ebene gegen die x -Achse und die Neigung der Tangente gegen die x, y -Ebene (Abb. 4.3-1).

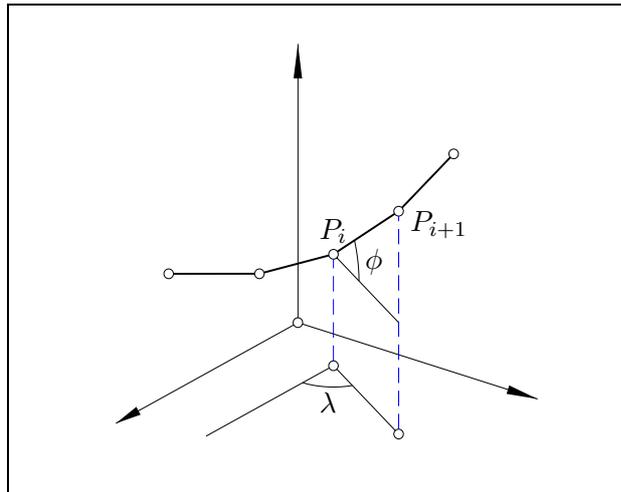


Abbildung 4.3-1: Richtungswinkel der Tangente

Eine einfache Approximation des Tangentenvektors im Punkt P_i ist der Vektor $\overrightarrow{P_i P_{i+1}}$. Werden die Punkte P_i durch die Ortsvektoren \underline{p}_i beschrieben und die Richtungseinheitsvektoren der Koordinatenachsen mit $\underline{e}_x, \underline{e}_y, \underline{e}_z$ bezeichnet, gilt für die Tangentenwinkel

$$\begin{aligned}\lambda_i &= \arctan \frac{\Delta y}{\Delta x} \quad , \quad -\frac{\pi}{2} < \lambda_i < \frac{\pi}{2} \\ \phi_i &= \arctan \frac{\Delta z}{\Delta d} \quad , \quad -\frac{\pi}{2} < \phi_i < \frac{\pi}{2}\end{aligned}\tag{4.3-1}$$

mit

$$\begin{aligned}\Delta x &= (\underline{p}_{i+1} - \underline{p}_i, \underline{e}_x) \quad , \quad \Delta y = (\underline{p}_{i+1} - \underline{p}_i, \underline{e}_y) \\ \Delta z &= (\underline{p}_{i+1} - \underline{p}_i, \underline{e}_z) \quad , \quad \Delta d = \sqrt{(\Delta x)^2 + (\Delta y)^2}\end{aligned}\tag{4.3-2}$$

Dabei bedeutet $(\underline{a}, \underline{b})$ das Skalarprodukt von \underline{a} und \underline{b} .

Im stetigen Fall kann die Kurve durch Integration nach der Bogenlänge zurückgewonnen werden. Im hier vorliegenden diskreten Fall erhält man die Kurve durch Summieren der Tangentenvektoren. Voraussetzung ist die bekannte konstante Tastweite, die die Länge der Tangentenvektoren festlegt, denn in den Tangentenwinkelfunktionen ist nur die Richtung gespeichert. Die Stetigkeit der Tangentenwinkelfunktionen (4.3-1) ist wegen der Periodizität bei beliebigen Kurven nicht unmittelbar gewährleistet. Auch die Auslegung in speziellen Programmierbefehlen, die den Wertebereich der arctan-Funktion auf den Wertebereich $[-\pi, \pi]$ ausdehnen, ändert nichts daran. Das Problem kann jedoch unkompliziert umgangen werden, wenn die Tangentenwinkel relativ zum vorhergehenden berechnet und sukzessive aufsummiert werden. Bei den Funktionen zur Rückrechnung darf nur der Definitionsbereich nicht auf eine Periode eingeschränkt sein.

Krümmungs- und Torsionsfunktion

Zur Ermittlung differentialgeometrischer Eigenschaften in einem Punkt werden Nachbarpunkte benötigt, deren gegenseitiger Abstand gegen Null geht. Werden bei einer punktwise gegebenen Kurve die nachfolgenden Punkte benutzt, ohne den Grenzübergang zu vollziehen, was im vorliegenden Fall einer punktwise gegebenen Kurve unmittelbar auch nicht möglich ist, erhält man jedoch nur eine sehr grobe und zusätzlich asymmetrische Näherung. Für die Beschreibung der Kurve zum Zwecke der Wavelettransformation mit dem Ziel einer Datenkompression ist das zunächst kein Hindernis. Erst wenn mit dieser Transformation auch eine Analyse der Kurvenform möglich sein soll, stehen Fragen der Approximationsgüte an.

Für das begleitende Dreibein, die Krümmungs- und die Torsionsfunktion im Punkt P_i werden als nachfolgende Punkte benötigt (Abb. 4.3-2):

Tangente:	P_{i+1}
Hauptnormale:	P_{i+1}, P_{i+2}
Binormale:	P_{i+1}, P_{i+2}
Krümmung:	P_{i+1}, P_{i+2}
Torsion:	$P_{i+1}, P_{i+2}, P_{i+3}$

Eine konsistente Beschreibung ist folglich bei n gegebenen Punkten nur für die Punkte P_1, \dots, P_{n-3} möglich. Anderenfalls sind die n Punkte um drei weitere zu ergänzen.

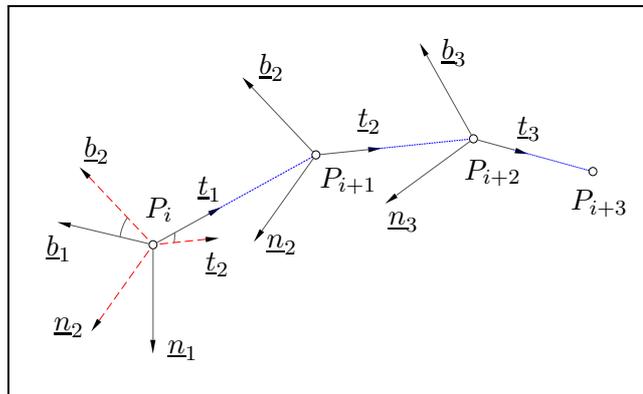


Abbildung 4.3-2: Nachbarpunkte

Bei einer konstanten Tastweite s mit

$$s = \|\underline{p}_{i+1} - \underline{p}_i\| = \text{const}, \quad i = 1, \dots, n-1 \quad (4.3-3)$$

lassen sich schrittweise die begleitenden Dreibeine, die Krümmung und die Torsion in den Punkten P_1, \dots, P_{n-3} bestimmen (Abb. 4.3-3).

Begleitendes Dreibein $(\underline{t}, \underline{n}, \underline{b})$:

$$\begin{aligned} \underline{t}_i &= \frac{1}{s}(\underline{p}_{i+1} - \underline{p}_i) & i = 1, \dots, n-1 \\ \underline{b}_i &= \frac{[\underline{t}_i, \underline{t}_{i+1}]}{\|[\underline{t}_i, \underline{t}_{i+1}]\|} & i = 1, \dots, n-2 \\ \underline{n}_i &= [\underline{b}_i, \underline{t}_i] & i = 1, \dots, n-2 \end{aligned} \quad (4.3-4)$$

(Diskrete) Krümmung κ :

$$\kappa_i = \frac{1}{s} \arctan \frac{(\underline{t}_{i+1}, \underline{n}_i)}{(\underline{t}_{i+1}, \underline{t}_i)} \quad i = 1, \dots, n-2 \quad (4.3-5)$$

(Diskrete) Torsion τ :

$$\tau_i = \frac{1}{s} \arctan \frac{(\underline{b}_{i+1}, \underline{t}_{i+1}^\perp)}{(\underline{b}_{i+1}, \underline{b}_i)} \quad i = 1, \dots, n-3 \quad (4.3-6)$$

mit $\underline{t}_{i+1}^\perp = [\underline{b}_i, \underline{t}_{i+1}]$

Dabei bedeutet $[\underline{a}, \underline{b}]$ das Kreuzprodukt von \underline{a} und \underline{b} .

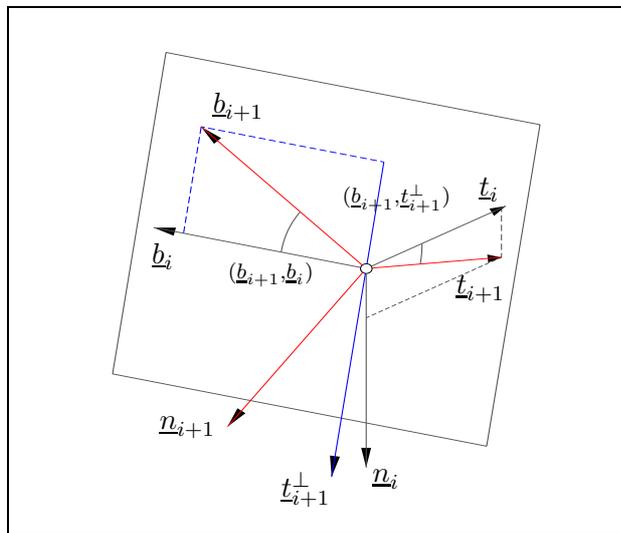


Abbildung 4.3-3: Diskrete Krümmung und Torsion

Die Rückrechnung der Koordinaten aus den Funktionen für Krümmung und Torsion bei gegebenen Koordinaten und begleitendem Dreibein im Anfangspunkt P_0 der Kurve ist ebenfalls schrittweise (im Sinne einer Integration nach der Bogenlänge) möglich.

Koordinaten des Folgepunktes:

$$\underline{p}_{i+1} = \underline{p}_i + s \cdot \underline{t}_i \quad (4.3-7)$$

Begleitendes Dreibein im Folgepunkt:

$$\begin{aligned} \underline{t}_{i+1} &= \cos(\kappa_i \cdot s) \cdot \underline{t}_i + \sin(\kappa_i \cdot s) \cdot \underline{n}_i \\ \underline{b}_{i+1} &= \cos(\tau_i \cdot s) \cdot \underline{b}_i - \sin(\tau_i \cdot s) \cdot \underline{t}_{i+1}^\perp \\ &\quad \text{mit } \underline{t}_{i+1}^\perp = [\underline{b}_i, \underline{t}_{i+1}] \\ \underline{n}_{i+1} &= [\underline{b}_{i+1}, \underline{t}_{i+1}] \end{aligned} \quad (4.3-8)$$

4.3.2 Implizite Beschreibung

In Abschnitt 4.1 wurde u. a. die Zielstellung genannt, flächenbezogene Geländelinien in einer Matrixstruktur zu beschreiben. Während bei der expliziten Beschreibung die Raumkurven losgelöst von der Gelände­fläche beschrieben werden und damit der Raumbezug zwischen Fläche und Kurve verlorengeht, wird in diesem Abschnitt eine flächenbezogene, implizite Beschreibung der Kurven vorgenommen. Als Motivation für diese Beschreibung dient der geometrische Zusammenhang zwischen Fläche und Kurve bei Bruchkanten. Bruchkanten sind insofern flächenbezogene Kurven, indem sie als Schnittkurven von Flächenstücken entstehen. Sie lassen sich also aus den Flächen ableiten, wenn es entsprechende Beschreibungen für die einzelnen Flächenstücke gibt.

Hier soll im Rahmen dieses geometrischen Zusammenhangs der umgekehrte Weg gegangen werden. Bei einer Beschreibung der Gelände­fläche in einer Matrixstruktur, die nur die Höhen in den Punkten eines regulären Gitters enthält, ist eine Ableitung von Schnittkurven (Bruchkanten) nicht sicher möglich. Es gibt zwar Algorithmen, die mögliche Schnittkanten extrahieren, sie sind aber nicht relevant, wenn bei der Gelände­aufnahme die Bruchlinien direkt erfaßt wurden und in einer weiteren Datenstruktur zusätzlich zur Flächenbeschreibung vorliegen.

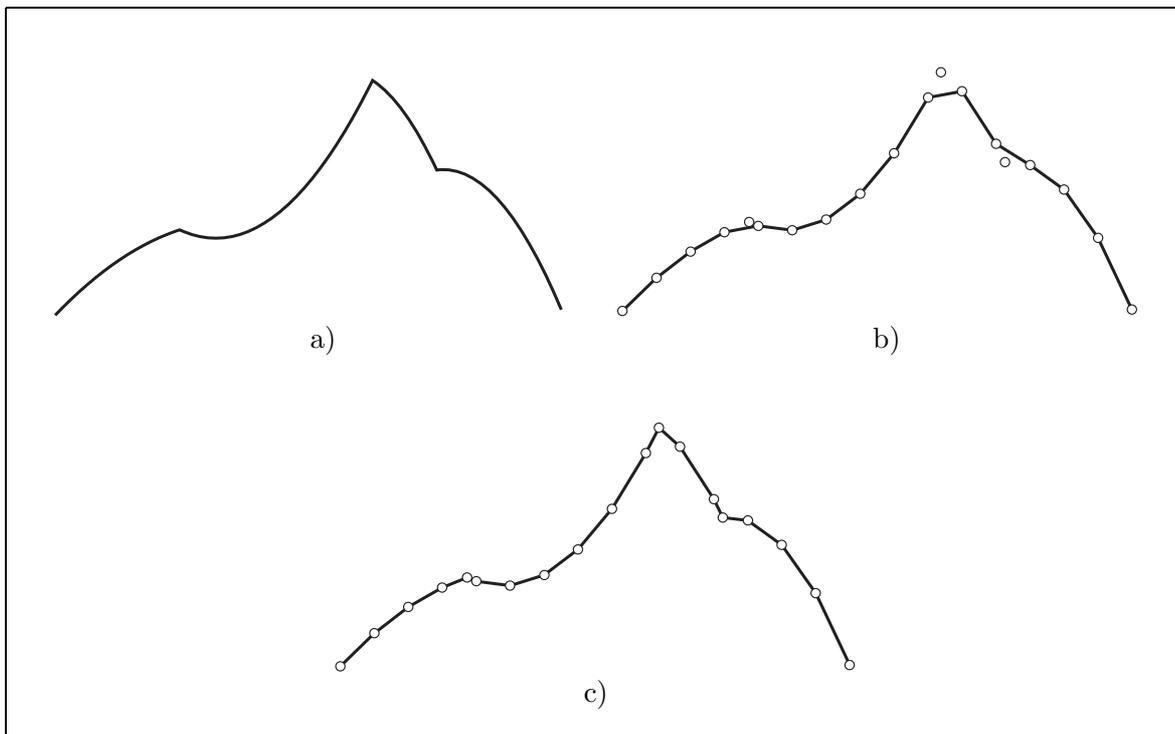


Abbildung 4.3-4: Gerastertes Profil mit zusätzlichen Knickinformationen

Im eindimensionalen Fall, d. h. bei Knicken in Profillinien sollen diese Überlegungen illustriert werden. Man betrachte im Hinblick auf die Zielstellung Wavelettransformation das Profil als eine nicht glatte Funktion (Abb. 4.3-4, a), die diskret durch äquidistante Abtastung gegeben ist (Gitterstruktur). Die Knickpunkte der Kurve, die Nichtdifferenzierbarkeitsstellen, liegen i. allg. nicht auf den Gitterpunkten und sind in dem diskreten Höhenvektor (dieser entspricht der Höhenmatrix des DHM im eigentlich zu untersuchenden zweidimensionalen Fall) nicht enthalten. Es soll aber davon ausgegangen werden, daß die Knickpunkte (Bruchkanten) explizit erfaßt wurden und als zusätzliche Information vorliegen (Abb. 4.3-4, b). Bei der Weiterverarbeitung des DGM, z.B. der Darstellung des Profils bzw. der Fläche, können die hybriden Informationen wieder aufgabenspezifisch zusammengeführt werden (Abb. 4.3-4, c).

Um konsistente Datensätze mit Raumbezug bei der Wavelettransformation eines hybriden DGM zu erhalten, sollen die Punkte der zusätzlich zu den Gitterdaten der Geländehöhen (DHM) aufgenommenen Geländekanten, die i.a. nicht auf dem Gitter liegen, in ein deckungsgleiches Gitter ohne Informationsverlust abgebildet werden. Damit entsteht ein zweites „DHM“ als Ersatzmodell für die Kanten. Im eindimensionalen Fall eines Geländeprofiles sind jeweils die beiden Koordinaten des Knicks (zwei skalare Informationen: Lage und Höhe) in zwei entsprechende Informationen des zweiten „DHM“ abzubilden.

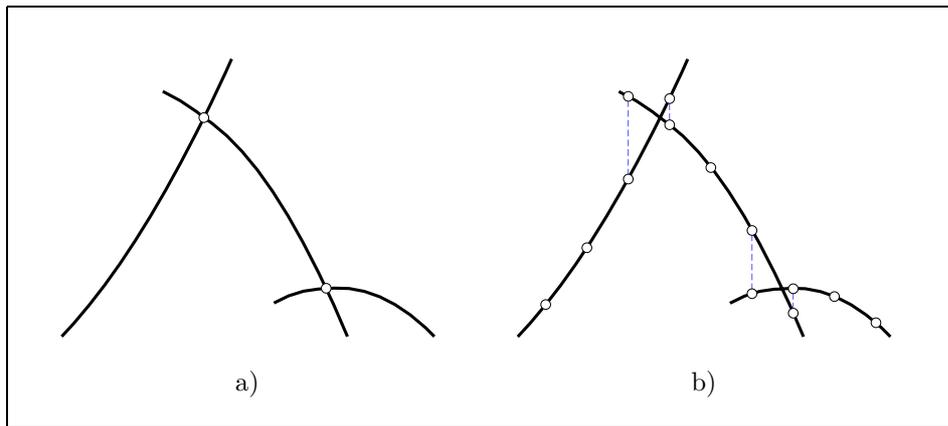


Abbildung 4.3-5: Profil

Einen Knick in einem Profil kann man als Schnittpunkt zweier glatter Kurvenstücke ansehen (Abb. 4.3-5, a).

Wird jedes der beiden Kurvenstücke über den Knick fortgesetzt, so erhält man für den jeweils nachfolgenden Gitterpunkt nach dem Knick Höhenwerte, die von den tatsächlichen Höhen des Profils abweichen (Abb. 4.3-5, b). Werden diese beiden Differenzen durch lineare Extrapolation ermittelt und als Gitterwerte (Kantensignal) eines zweiten Gittermodells (Ersatzmodell) gespeichert, so lassen sich aus diesen Werten gemeinsam mit den Höhenwerten des Originalprofils (Profilsignal) die Koordinaten des Knicks rekonstruieren.

Außerdem kann im Ersatzmodell der Knick lokalisiert werden, und es ist damit im Waveletbereich ein Raumbezug zum Profilsignal möglich.

Diese Überlegungen lassen sich unmittelbar auf den zweidimensionalen Fall übertragen, wenn die allgemeine Kurvenbeschreibung als Polygonzug geeignet modifiziert wird. Man kann den Polygonzug mit 3D-Punkten in allgemeiner Lage in vertretbaren Genauigkeitsgrenzen durch einen Polygonzug ersetzen, der aus den Schnittpunkten des Ausgangspolygons mit den Gitterebenen besteht (Abb. 4.3-6).

Aus dem Polygonzug $\{Q_k\}$ der Original-Kantenpunkte entsteht ein Polygonzug $\{R_j\}$ aus Gitterschnittpunkten. Im Programmsystem SCOP, IPF (1991) werden die $\{Q_k\}$ *Kantenstützpunkte* und die $\{R_j\}$ *Kantenknoten* genannt (vgl. auch KRAUS (2000), S. 193). Es handelt sich dabei bereits um eine weitere Approximation der Bruchlinie. Die Abbildung $\{Q_k\} \rightarrow \{R_j\}$ ist deshalb in die Fehlerbetrachtungen einzuschließen.

Aus der Zugehörigkeit zweier aufeinanderfolgender Kantenpunkte Q_k, Q_{k+1} zu den Gitterzellen des DHM ergeben sich die horizontalen und vertikalen Gitterschnitte R_j als Schnitt des Kantensegments (Q_k, Q_{k+1}) mit den Gitterebenen zwischen diesen beiden Zellen. Neben den x, y, z -Koordinaten ist für jeden Gitterschnittpunkt die Zugehörigkeit zur jeweiligen Zelle zu speichern, um die Zuordnung der extrapolierten Differenzen zu den Gitterpunkten zu ermöglichen. Diese Datenstruktur, die nur *integer*-Daten enthält, kann gleichzeitig die notwendigen

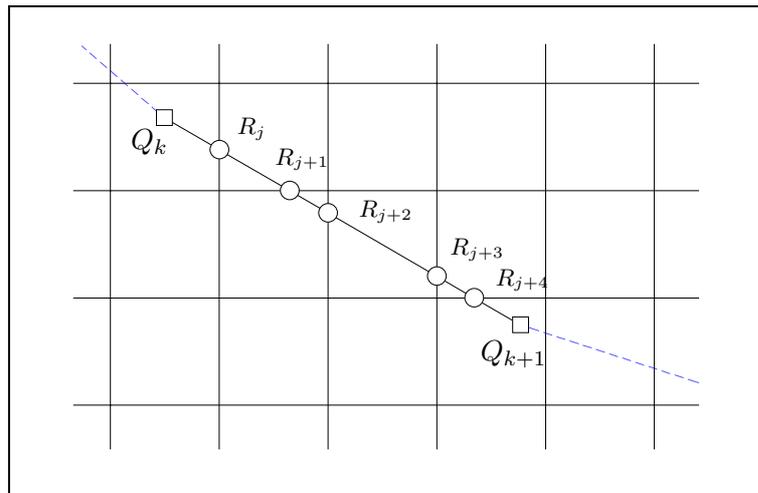


Abbildung 4.3-6: Berechnung der Gitterschnitte

Topologie-Informationen der Geländekante aufnehmen, die in der zu erzeugenden Matrixstruktur der extrapolierten Höhen nicht mehr unmittelbar vorhanden ist. Im Anhang D wird ein Algorithmus vorgestellt, der die Kantentopologie aus der Matrixstruktur rekonstruiert.

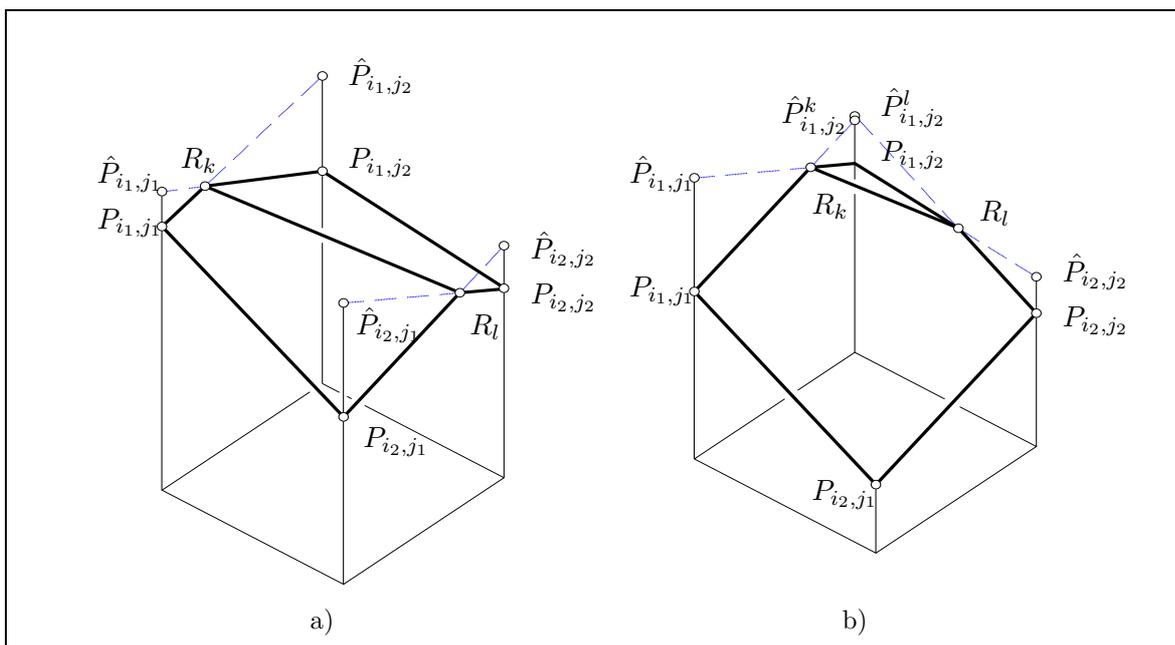


Abbildung 4.3-7: Extrapolation

Durch die Verbindungsstrecke der Gitterschnitte wird die Geländefläche in dieser Gitterzelle (bzw. Netzmasche) in zwei Teilflächen zerlegt. Dabei können zwei topologisch unterschiedliche Fälle eintreten, die Einfluß auf die Eindeutigkeit der berechneten Differenzen haben (Abb. 4.3-7). Im ersten Fall liegen die Gitterschnitte auf gegenüberliegenden und im zweiten Fall auf benachbarten Maschenseiten.

Entsprechend erfolgt die Berechnung der extrapolierten Höhen nach BEYER (2000b) im ersten Fall gemäß

$$\begin{aligned}
\tilde{y}_{i_1, j_1} &= \frac{y_k - y_{i_1, j_2}}{x_k - x_{i_1, j_2}} \cdot (x_{i_1, j_1} - x_{i_1, j_2}) + (y_{i_1, j_2} - y_{i_1, j_1}) \\
\tilde{y}_{i_1, j_2} &= \frac{y_k - y_{i_1, j_1}}{x_k - x_{i_1, j_1}} \cdot (x_{i_1, j_2} - x_{i_1, j_1}) + (y_{i_1, j_1} - y_{i_1, j_2})
\end{aligned} \tag{4.3-9}$$

und

$$\begin{aligned}
\tilde{y}_{i_2, j_1} &= \frac{y_k - y_{i_2, j_2}}{x_k - x_{i_2, j_2}} \cdot (x_{i_2, j_1} - x_{i_2, j_2}) + (y_{i_2, j_2} - y_{i_2, j_1}) \\
\tilde{y}_{i_2, j_2} &= \frac{y_k - y_{i_2, j_1}}{x_k - x_{i_2, j_1}} \cdot (x_{i_2, j_2} - x_{i_2, j_1}) + (y_{i_2, j_1} - y_{i_2, j_2})
\end{aligned} \tag{4.3-10}$$

und im zweiten Fall gemäß

$$\begin{aligned}
\tilde{y}_{i_1, j_1} &= \frac{y_k - y_{i_1, j_2}}{x_k - x_{i_1, j_2}} \cdot (x_{i_1, j_1} - x_{i_1, j_2}) + (y_{i_1, j_2} - y_{i_1, j_1}) \\
\tilde{y}_{i_1, j_2}^k &= \frac{y_k - y_{i_1, j_1}}{x_k - x_{i_1, j_1}} \cdot (x_{i_1, j_2} - x_{i_1, j_1}) + (y_{i_1, j_1} - y_{i_1, j_2})
\end{aligned} \tag{4.3-11}$$

und

$$\begin{aligned}
\tilde{y}_{i_1, j_2}^l &= \frac{y_k - y_{i_2, j_2}}{x_k - x_{i_2, j_2}} \cdot (x_{i_1, j_2} - x_{i_2, j_2}) + (y_{i_2, j_2} - y_{i_1, j_2}) \\
\tilde{y}_{i_2, j_2} &= \frac{y_k - y_{i_1, j_2}}{x_k - x_{i_1, j_2}} \cdot (x_{i_2, j_2} - x_{i_1, j_2}) + (y_{i_1, j_2} - y_{i_2, j_2}) \cdot
\end{aligned} \tag{4.3-12}$$

Auf diese Weise entsteht aus der Folge von 3D-Koordinaten eine i. allg. schwach besetzte Matrix extrapolierter relativer Höhen. Die Bruchlinie in Abb. 4.3-8 ist durch eine Folge von 3D-Koordinaten gegeben. Sie stellt damit ein Polygon dar, dessen Eckpunkte i. allg. nicht auf den Gitterpunkten liegen (Abb. 4.3-9, links). Nachdem die ursprüngliche Geländelinie in ein Polygon aus ihren Gitterschnittpunkten transformiert wurde (Abb. 4.3-9, Mitte oben), können die Höhen der Gitterschnitte nach den oben angegebenen Beziehungen auf die Gitterpunkte extrapoliert werden (Abb. 4.3-9, Mitte unten). Es entsteht eine schwach besetzte Matrix aus extrapolierten Höhen, die die Kantengeometrie „unscharf“ wiedergibt (Abb. 4.3-9, rechts). Der Raumbezug ist in dieser Matrixstruktur nur explizit „unscharf“ erhalten. Wegen der Rekonstruierbarkeit, die durch Interpolation (zu (4.3-9), (4.3-10), (4.3-11), (4.3-12) inverse Formeln) lokal für jeden Gitterschnitt möglich ist, ist der Raumbezug implizit sogar exakt erhalten. Diese zur Höhenmatrix konsistente Matrix kann in gleicher Weise der Wavelettransformation unterworfen werden.

Die Rekonstruktion der Kante setzt die Kenntnis über die Kantentopologie voraus, die aus der Matrix der extrapolierten Höhen mittels eines Verfolgungsalgorithmus gewonnen werden kann (vgl. Anhang D). Die Rücktransformation aus den in der Matrix gespeicherten extrapolierten

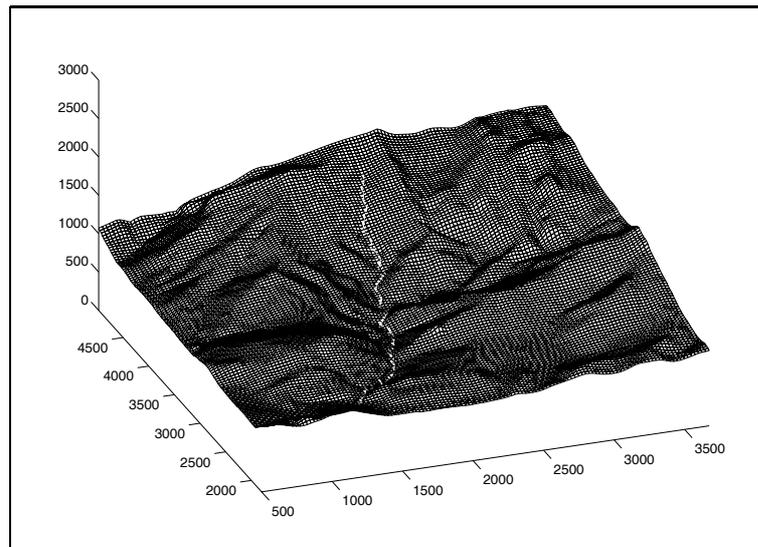


Abbildung 4.3-8: Bruchlinie im Testdatensatz Schneetalpe.

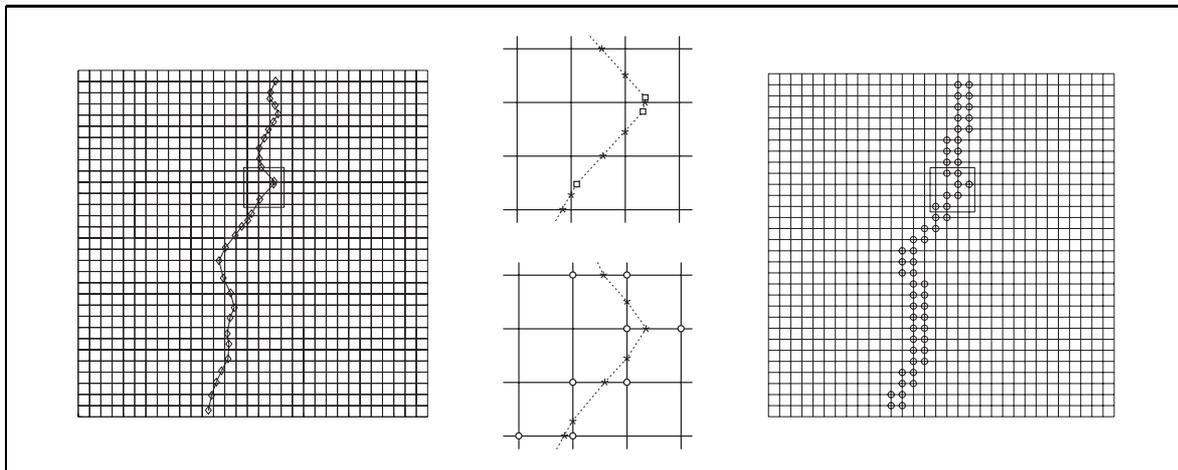


Abbildung 4.3-9: Polygon der Raumkurve (links). Transformation in das Gitterschnittpolygon (Ausschnitt Mitte, oben). Transformation in die extrapolierten Höhen (Ausschnitt Mitte, unten). Matrix der extrapolierten Höhen (rechts).

Höhen liefert nur die Gitterschnittpunkte. Die ursprünglichen Polygonpunkte sind verloren. Diesen „Verlust“ kann man jedoch durch kleine Gitterweiten in tolerierbaren Grenzen halten.

Das hier vorgestellte Verfahren zur Abbildung der Kanteninformationen leistet das Erforderliche für die genannte Zielstellung bei „gutartiger“ Verteilung der Kanten im Gelände, die im natürlichen Gelände in der Regel angenommen werden kann. Probleme bestehen bei im Verhältnis zur Rasterweite des DHM zu dicht liegenden Kanten (vorwiegend im verbauten Gelände, z.B. an terrassierten Hängen) und Knoten von Kantennetzen (vorwiegend im Hochgebirge). Diese lassen sich jedoch z.B. durch eine geeignete Aufteilung in mehrere Schichten beheben (vgl. 8.1).

5 Wavelettransformation digitaler Geländemodelle

5.1 Wavelettransformation der Geländefläche

Besteht das digitale Geländemodell ausschließlich aus den Höhen über einem regulären Gitter, so bereitet die Wavelettransformation keine Probleme. Es können unmittelbar die bekannten Verfahren der zweidimensionalen Wavelettransformation angewendet werden. Die Daten sind üblicherweise in einer Matrix gespeichert und stellen damit eine diskrete Funktion von zwei Veränderlichen dar. Unabhängig davon, ob die schnelle Wavelettransformation (FWT) oder die stationäre Wavelettransformation angewendet wird, ist ohne Datenkompression im Waveletbereich eine exakte Rekonstruktion der Geländefläche möglich.

Nach Abschnitt 3.4.2 wird die $(2^k, 2^k)$ -Matrix der Höhen bei der schnellen Wavelettransformation in eine $(2^k, 2^k)$ -Matrix von Waveletkoeffizienten (WK) der Struktur in Abb. 5.1-1 transformiert. Diese Struktur ist bei der Interpretation der WK zu berücksichtigen. Wegen der Approximationseigenschaft kann man die WK als i -te Ableitung interpretieren, wenn i die Ordnung des Wavelets ist. Bei dieser Approximation ist jedoch noch ein Skalierungsfaktor zu berücksichtigen, der für die stationäre Wavelettransformation in 7.1 für Wavelets verschiedener Ordnungen in Abhängigkeit von den Momenten der Wavelets bestimmt wird. Für die Wavelets der 2D-Wavelettransformation approximieren die Tensorproduktwavelets HT , TH die partiellen Ableitungen $\frac{\partial^i f}{\partial x^i}$ und $\frac{\partial^i f}{\partial y^i}$. Das Tensorproduktwavelet HH approximiert die partielle Ableitung $\frac{\partial^{2i} f}{\partial x^i \partial y^i}$, und das Tensorproduktwavelet TT liefert den reinen Tiefpaß, d.h. einen „Mittelwert“. Jede Teilmatrix ist bei der Abbildung des Signals in den Waveletkoeffizienten ein Bild des gesamten Signals. In den einzelnen Skalen stehen die WK damit wegen des (bedingt durch den

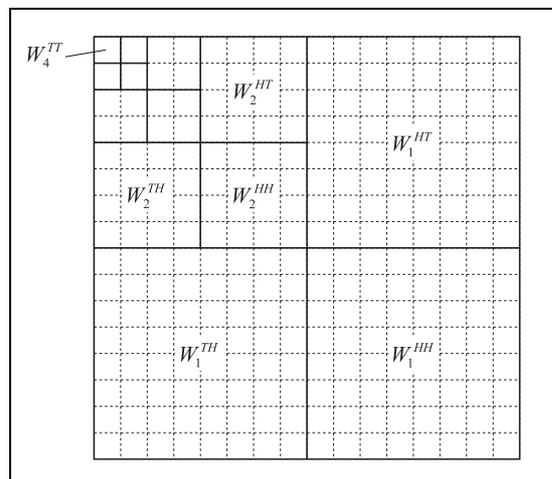


Abbildung 5.1-1: Datenstruktur der WK bei der zweidimensionalen schnellen WT

wachsenden Support) immer größer werdenden Einflußbereiches für mehr oder weniger „mittlere“ Ableitungen in einem Bereich der Geländefläche. In Abb. 5.1-2 sind die Teilmatrizen der Waveletkoeffizienten dreidimensional dargestellt, um die Approximation der partiellen Ableitungen zu veranschaulichen. Links ist jeweils die Geländefläche zu sehen. Rechts sind die Waveletkoeffizienten für drei Skalen als zweidimensionale diskrete Funktion abgebildet. Man kann erkennen, daß in allen Skalen die erste partielle Ableitung approximiert wird, jedoch von Skale zu Skale immer grober.

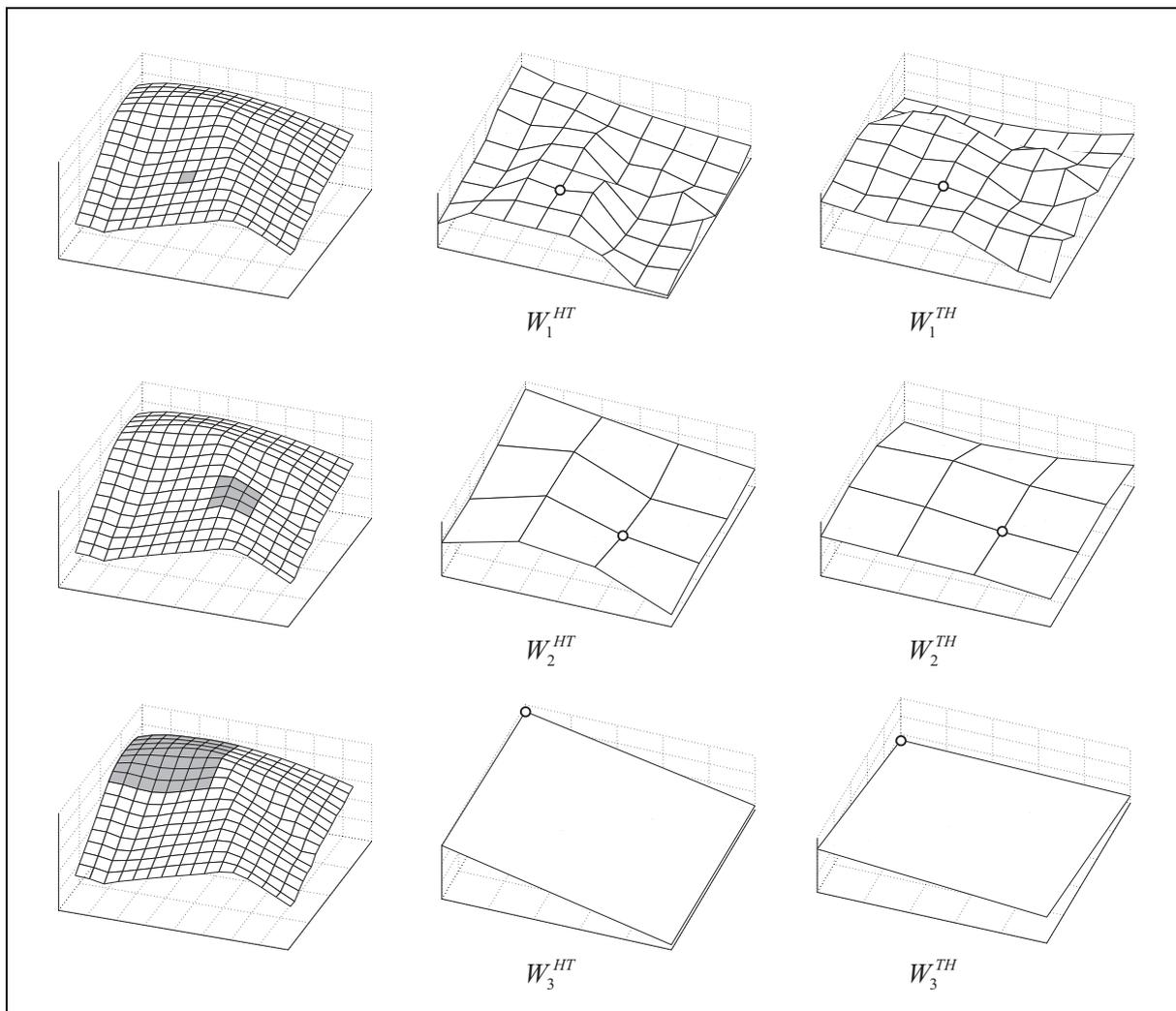


Abbildung 5.1-2: Veranschaulichung der Approximation der partiellen Ableitungen nach x und y . Jeder Netzpunkt entspricht einem Matrixelement.

Für jede Skale sind jeweils im linken Bild diejenigen Höhenwerte hervorgehoben, die zur Berechnung des rechts markierten Waveletkoeffizienten herangezogen werden. Die Gitterpunkte entsprechen jeweils den Matrixelementen. In dieser Weise können die WK bestimmten Signalbereichen als „mittlere“ Ableitungen durch einfache Indexumrechnung zugeordnet werden. Diese bei der schnellen Wavelettransformation offensichtlich notwendige Indexumrechnung stellt jedoch nur eine erste Näherung für die Lokalisierung der in den WK gespeicherten Ableitungswerte auf bestimmte Signalwerte dar. Bei der stationären Wavelettransformation, bei der diese Indexumrechnung nicht erforderlich ist, da jede Bildmatrix die gleiche Dimension wie die Signalmatrix hat, zeigt sich jedoch, daß durch den Support der Wavelets die indexbezogene Lokalisierung nicht ausreichend ist. Es ist eine Korrektur erforderlich, die ebenso wie der Skalierungsfaktor wesentlich von den Momenten der Wavelets abhängt. Die Bestimmung dieser notwendigen Lagekorrektur ist Gegenstand des Abschnittes 7.2.

5.2 Wavelettransformation von Kurven in expliziter Darstellung

In Abschnitt 4.3.1 wurde diskutiert, wie Kurven, die in einem digitalen Geländemodell explizit als 3D-Punktfolge vorliegen, so beschrieben werden können, daß sie einer Wavelettransformation zugänglich werden. Dazu wurden die 3D-Punktfolgen, die eine diskrete 3D-Vektorfunktion darstellen, durch zwei oder drei skalare Funktionen beschrieben. Diese Funktionen können mit den üblichen Verfahren der diskreten Wavelettransformation in drei oder zwei Sätze von Waveletkoeffizienten transformiert werden (siehe BEYER (1999)).

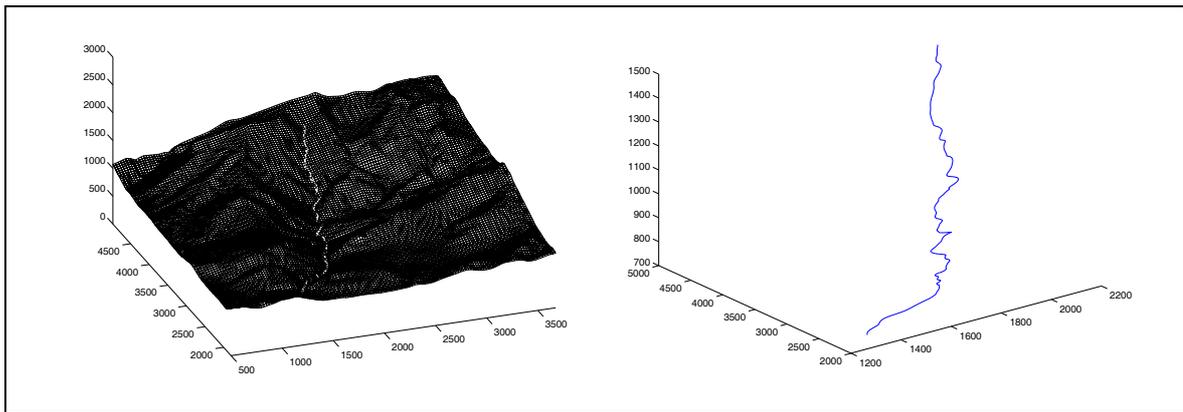


Abbildung 5.2-1: Geländekurve aus dem Datensatz Schneetalpe.

Es zeigte sich, daß die Beschreibung mittels Krümmungs- und Torsionsfunktion zu empfindlich gegenüber Fehlern, die durch die verlustbehaftete Datenkompression entstehen, ist. Es werden deshalb in diesem Abschnitt nur die Repräsentationen Koordinatenfunktionen und Tangentwinkelfunktionen vorgestellt. Eine Geländekurve (Abb. 5.2-1), hier als Beispiel eine Tallinie aus dem DGM Schneetalpe, läßt sich explizit durch ihre Koordinatenfunktionen (Abb. 5.2-2, links)

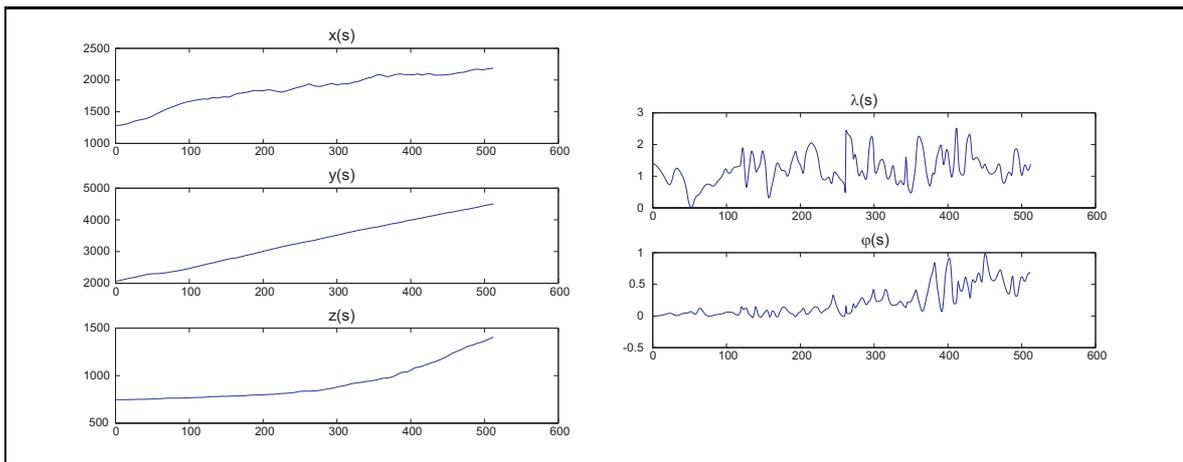


Abbildung 5.2-2: Die drei Koordinatenfunktionen (links) und die beiden Tangentwinkelfunktionen (rechts) der Geländekurve.

oder ihre Tangentwinkelfunktionen (Abb. 5.2-2, rechts) beschreiben. Jede dieser Funktionen kann als Signalvektor der Wavelettransformation unterworfen werden. Bei der schnellen Wavelettransformation entsteht für jeden Signalvektor ein gemäß (3.4-1) gegliederter Vektor von Waveletkoeffizienten. Um den lokalen Zusammenhang zwischen Signalvektor und Waveletvektor

(Signal und approximierter Ableitung) zu veranschaulichen, wird in Abb 5.2-3 jeweils nur die zweite Hälfte der WK-Vektoren, die die Waveletkoeffizienten der ersten Skale enthalten, grafisch dargestellt. Bei einer Kompression nach dem Schwellwertverfahren werden je nach Schwellwert betragsmäßig kleine Waveletkoeffizienten zu Null gesetzt. Die dadurch bei der Synthese der Funktionen fehlenden Basisfunktionen, die zwar nur gering in die Synthese eingehen, führen zu Abweichungen der rekonstruierten Koordinaten- und Tangentenwinkelfunktionen (Abb. 5.2-4) gegenüber den Originalfunktionen.

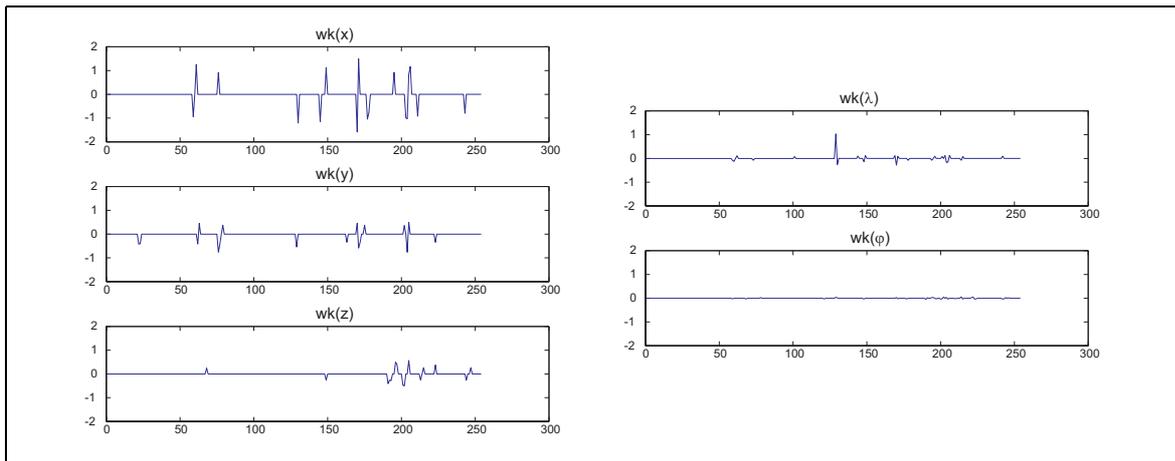


Abbildung 5.2-3: Die Waveletkoeffizienten der drei Koordinatenfunktionen und der zwei Tangentenwinkelfunktionen in der ersten Skale.

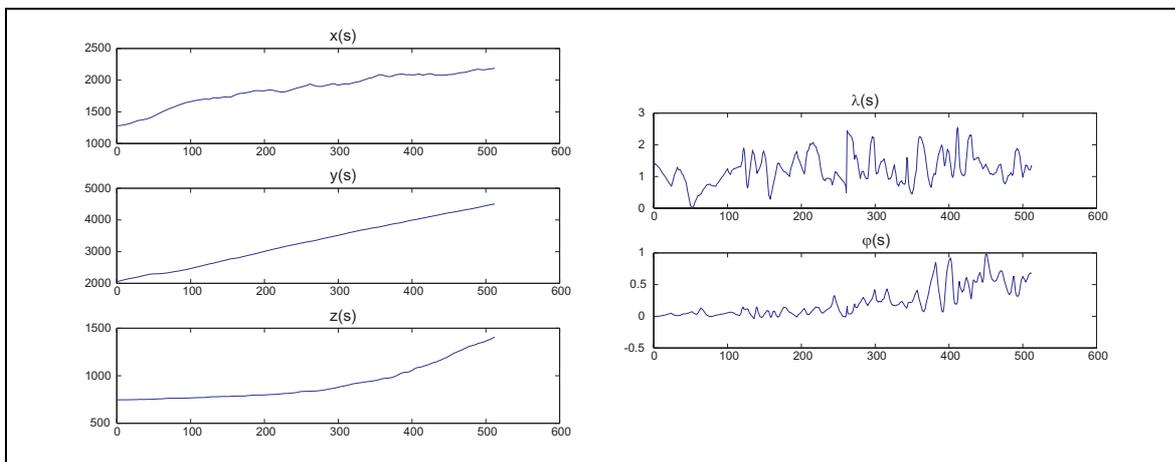


Abbildung 5.2-4: Die rücktransformierten Koordinaten- und Tangentenwinkelfunktionen.

Folglich sind die aus den fehlerhaften Beschreibungsfunktionen rekonstruierten Raumkurven ebenfalls fehlerbehaftet. Während jedoch die Fehler der Kurve bei der Beschreibung durch die Koordinatenfunktionen über die Bogenlänge stationäres Verhalten zeigen, sind die Fehler bei der Beschreibung durch die Tangentenwinkelfunktionen wegen der rekursiven Berechnung bei der Rekonstruktion abhängig vom vorhergehenden Fehler. Es ist ein Ansteigen des Fehlers über die Bogenlänge zu verzeichnen. Dieser Effekt ist in Abb. 5.2-5, in der die Abweichungen orthogonal zum Kurvenverlauf dargestellt sind, zu sehen.

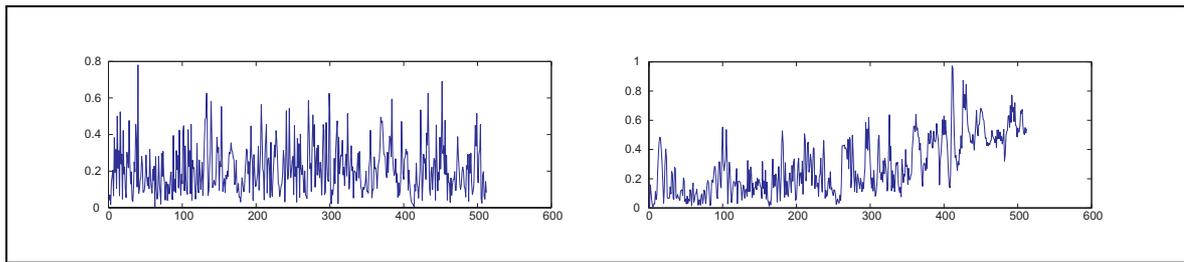


Abbildung 5.2-5: Die orthogonalen Abweichungen der rekonstruierten von der Original-Kurve.

5.3 Wavelettransformation von Kurven in impliziter Darstellung

Die in Abschnitt 4.3.2 diskutierte Möglichkeit, eine Raumkurve eines DGM in eine Matrixstruktur abzubilden, die der Matrixstruktur des Höhenmodells auf einem regulären Gitter entspricht, führt auf eine schwach besetzte Matrix. Diese läßt sich in der gleichen Weise wie die Höhenmatrix der Geländefläche der Wavelettransformation unterwerfen und liefert eine zur WK-Matrix der Geländefläche konsistente Matrix von Waveletkoeffizienten. Durch die Lokalisierungseigenschaft der Wavelettransformation besteht im Waveletbereich ein Raumbezug zwischen der Geländefläche und den Geländekanten.

Die schnelle zweidimensionale Wavelettransformation bildet die Matrix der extrapolierten Höhen (Abb. 4.3-9, rechts, Abb. 5.3-1, links) in die Matrix der Waveletkoeffizienten (Abb. 5.3-1, rechts) ab. Es ist zu sehen, daß der Raumbezug durch die Lokalisierungseigenschaft der Wavelettransformation erhalten bleibt.

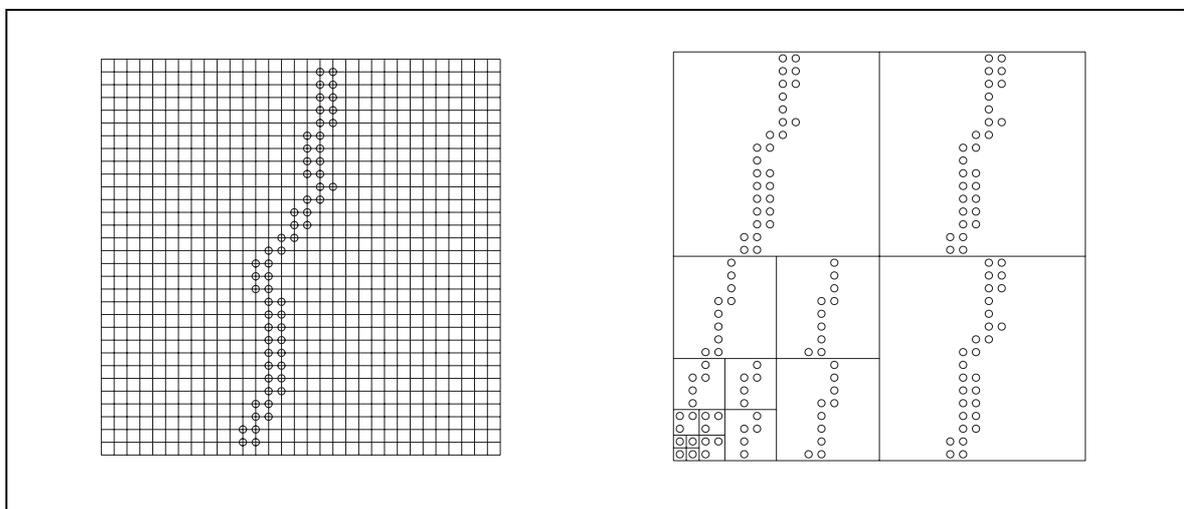


Abbildung 5.3-1: Matrix der extrapolierten Höhen (links). Wavelettransformierte der extrapolierten Höhen (rechts).

Die die Kurven beschreibenden Matrizen sind auf Grund ihrer Entstehung schwachbesetzte Matrizen¹ und in Folge davon auch die Matrizen der Waveletkoeffizienten. Die Notwendigkeit einer Datenkompression ist deshalb nicht unbedingt gegeben. Es treten hingegen sogar durch eine verlustbehaftete Datenkompression erhebliche Probleme bei der Rücktransformation auf. Bei der

¹ Schwachbesetzte Matrizen lassen sich mit geeigneten Methoden der Informatik effektiv speichern. Diese Methoden sind nicht Gegenstand der vorliegenden Arbeit.

Datenkompression im Waveletbereich werden i. allg. typische Eigenschaften glatter Funktionen durch die Wahl geeigneter Wavelets ausgenutzt. Diese typischen Eigenschaften liegen aber bei schwach besetzten Matrizen nicht vor. Bei der verlustbehafteten Datenkompression, um die es in dieser Arbeit gehen soll, entstehen bei der inversen Wavelettransformation Matrizen, die je nach Kompressionsgrad nicht mehr schwach besetzt sein müssen. Damit geht die zur Rekonstruktion der Kanten notwendige Topologieinformation, die in der ursprünglichen Matrix noch indirekt enthalten ist (vgl. Anhang D), weitgehend verloren. Um das zu veranschaulichen, betrachten wir die Wavelettransformierte nach Datenkompression mit der Kompressionsrate 2. Es bleiben die Waveletkoeffizienten in Abb. 5.3-2, links.

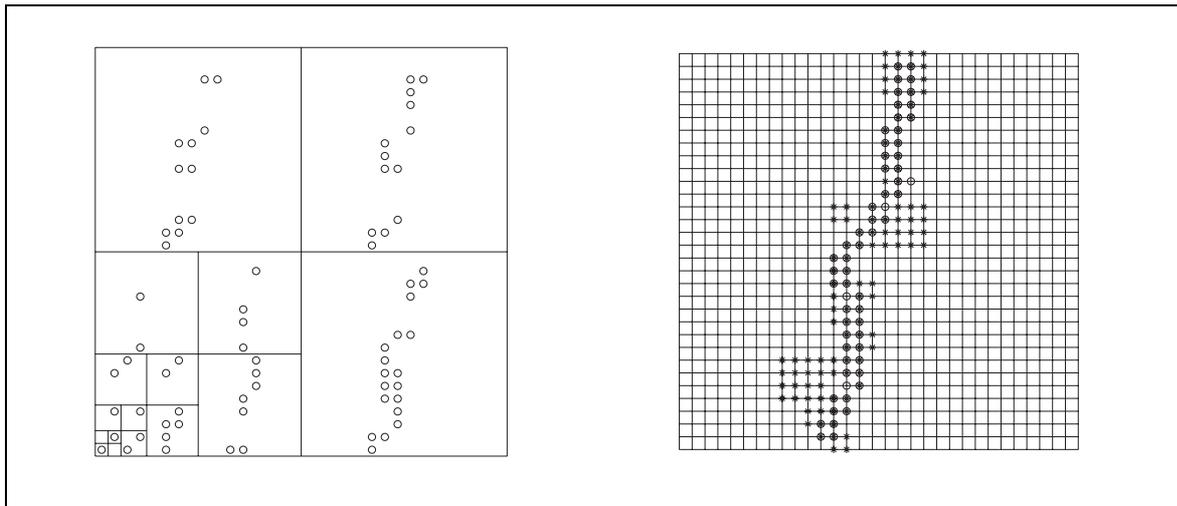


Abbildung 5.3-2: Wavelettransformierte der extrapolierten Höhen nach Datenkompression mit Kompressionsrate 2 (links). Matrix der extrapolierten Höhen nach Anwendung eines Schwellwertfilters (rechts).

Die inverse Wavelettransformation, angewendet auf die ausgedünnte Matrix der Waveletkoeffizienten, durch die die extrapolierten Höhen zurückgewonnen werden, liefert in diesem Falle eine vollbesetzte Matrix. Die Begründung liegt auf der Hand. Die Originalmatrix entsteht bei der inversen Wavelettransformation als Linearkombination der Wavelets mit den WK als Koeffizienten. Die Wavelets der hohen Skalen belegen wegen ihres breiten Supports die Matrix vollständig. Ohne Kompression kompensieren die Wavelets der niederen Skalen diese Belegung an den Stellen, an denen ursprünglich Nullen standen. Diese Kompensation kann nicht vollständig erfolgen, wenn wegen der gestrichenen Waveletkoeffizienten Wavelets in der Linearkombination der inversen Wavelettransformation fehlen. Die fehlbelegten Matrixelemente sind jedoch i. allg. klein. Filtert man die Matrixelemente heraus, die betragsmäßig unterhalb eines Schwellwertes liegen, bleiben z.B. die relativen Höhen übrig, die in Abb. 5.3-2, rechts, markiert durch „*“ zu sehen sind. Die ursprünglichen Matrixelemente sind zum Vergleich durch „o“ markiert. Es ist zwar noch grob die Lokalisierung erkennbar, aber eine Rekonstruktion der Kante aus diesen durch die Kompression verfälschten relativen Höhen nach den zu (4.3-9) bis (4.3-12) inversen Formeln ist nicht möglich. Die Wavelettransformation der implizit in einer Matrixstruktur gemäß 4.3.2 gespeicherten Kante ist möglich, und der Raumbezug im Waveletbereich ist gegeben, eine verlustbehaftete Datenkompression erweist sich jedoch nicht als zweckmäßig.

6 Kompression von Geländedaten

6.1 Vorbemerkungen

Durch die rasche Entwicklung der Erfassungsmethoden für Geländedaten, insbesondere durch die Möglichkeiten der Fernerkundung, werden Geo-Informationssysteme immer leistungsfähiger, da Daten in immer vielfältigerer Art und immer höherer Auflösung zur Verfügung stehen. Auf der anderen Seite müssen die GIS aber auch in der Lage sein, diese Datenflut zu bewältigen. Eine geeignete Datenkompression wird zur dringenden Notwendigkeit. Die steigenden Hardwarerekapazitäten werden dieses Problem allein nicht lösen können, wie die bisherige Entwicklung gezeigt hat. Eine gewisse Vorreiterrolle in der Datenkompression hat die digitale Bildverarbeitung übernommen.

Grundsätzlich sind zwei Kompressionsverfahren zu unterscheiden:

- Verlustfreie Verfahren (lossless compression)
Sie erlauben eine exakte Rekonstruktion der Daten nach der Dekodierung. Sie wurden z.B. seit den 60er Jahren in Raumfahrtmissionen eingesetzt. Damit lassen sich jedoch nur geringe Kompressionsraten (1,1:1 - 2,5:1, maximal 3,5:1, TRIEBFÜRST (1999)) erzielen.
- Verlustbehaftete Verfahren (lossy compression)
Diese bringen nach der Dekodierung einen gewissen Datenverlust mit sich. Es müssen deshalb für eine praktisch sinnvolle Kompression gewisse Redundanzen in den Daten vorhanden sein. Je nach Redundanz lassen sich damit wesentlich höhere Kompressionsraten erzielen.

Die ursprüngliche Anwendung der Kompression zur Einsparung von Speicherplatz bei der Archivierung von Bildern wurde bald um die temporäre Kompression zur Bildübertragung im Internet erweitert. Während die Datenstrukturen von digitalen Rasterbildern und von digitalen Höhenmodellen ähnlich sind, ist der Informationsgehalt durchaus unterschiedlicher Natur. Das wird dann von Interesse sein, wenn bei der Anwendung verlustbehafteter Kompressionsverfahren die Frage nach den verzichtbaren redundanten Informationen steht.

Von TRIEBFÜRST (1999) werden Kompressionsverfahren für Fernerkundungsdaten umfangreich diskutiert. Die vorliegende Arbeit konzentriert sich auf Probleme der Wavelettransformation hybrider Geländemodelle (vgl. Kapitel 4), wobei die verlustbehaftete Kompression im Waveletbereich ein Teilproblem darstellt. In diesem Sinne wird zur allgemeinen Fragestellung der Datenkompression auf die o.g. Arbeit verwiesen.

6.2 Kompressionsverfahren

Bei den verlustbehafteten Kompressionsverfahren mittels Schwellwertbildung im Waveletbereich kann man prinzipiell nach BETHGE et al. (1997) zwei Verfahren unterscheiden:

- Schwellwertbildung auf der Abszisse:
Es werden die Koeffizienten der hochfrequenten Skalen zu Null gesetzt. Damit wird gewissermaßen ein Sperrfilter realisiert.

- Schwellwertbildung auf der Ordinate:
In allen Skalen werden diejenigen Koeffizienten zu Null gesetzt, die einen gegebenen Schwellwert, der skalenabhängig sein kann, nicht übersteigen.

Dabei ist zu berücksichtigen, welche Effekte durch den Datenverlust akzeptiert werden bzw. sogar erwünscht sind. Steht die Filterung zum Zwecke der Generalisierung im Vordergrund, sollen hochfrequente Informationen herausgefiltert werden. Handelt es sich um eine Datenkompression im eigentlichen Sinne zur Verringerung des Speicherplatzes, sollen die wesentlichen Informationen erhalten bleiben.

Die hier angegebene Einteilung kann man weiter präzisieren. Geht man allgemein von einer Funktion aus (in unserem Falle sind das z. B. die Geländehöhen als Funktion ihrer Lage: $h = f(x, y)$), so erfolgt bereits bei der Datenerfassung eine gewisse Datenkompression. Die praktisch in allen drei Dimensionen kontinuierlich vorliegenden Daten des zu erfassenden Bereichs werden nämlich durch die technisch notwendige Abtastung in eine endliche Datenmenge abgebildet. Es handelt sich dabei nicht nur um eine Abtastung im Argumentbereich (Lagekoordinaten, „Abszisse“¹) sondern wegen der Speicherung der Funktionswerte mit einer endlichen Stellenzahl auch um eine Abtastung im Funktionswertebereich (Höhe, „Ordinate“).

Unter Kompression im engeren Sinne versteht man jedoch die weitere Kompression der nach der Datenerfassung vorliegenden Daten. Kompression heißt Verringerung des benötigten Speicherplatzes. Diese kann einerseits durch effektivere Kodierung, d.h. Verringerung der Coding-Redundanz (TRIEBFÜRST (1999)) und andererseits durch Verzicht auf Informationen geringer Relevanz erreicht werden. Hier soll der zweite Aspekt behandelt werden.

Der Datenumfang einer diskreten Funktion kann durch Verringerung der Informationsdichte (Verringerung der Auflösung) sowohl im Argumentbereich als auch im Funktionswertebereich reduziert werden. Wie bereits im Zusammenhang mit der Datenerfassung erwähnt, ist das durch eine Abtastung mit größerer Tastweite möglich. Der dabei entstehende Informationsverlust ist über das Abtasttheorem beschreibbar.

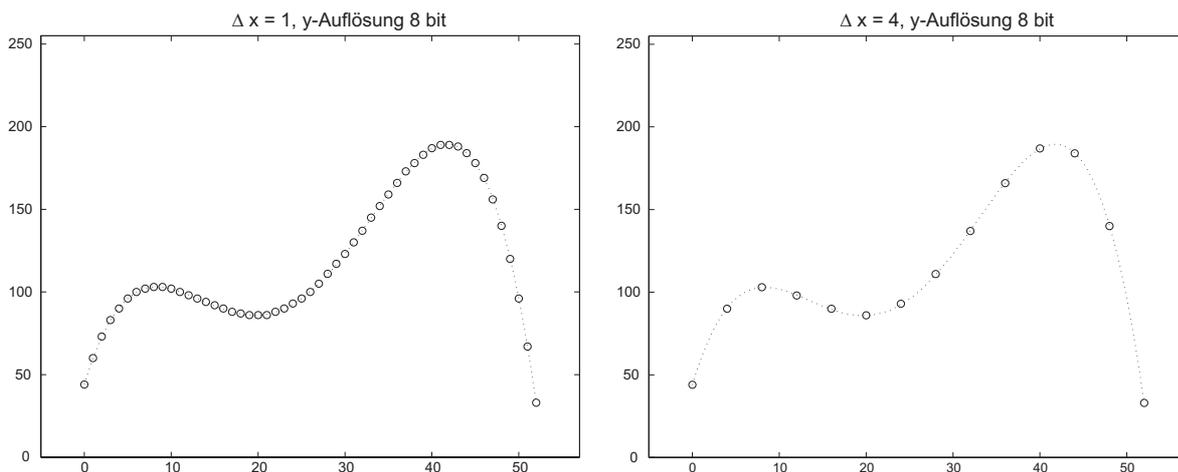


Abbildung 6.2-1: Unterschiedliche Abszissenabtastung durch Variation der Tastweite bei konstanter Speichertiefe von 8 bit.

¹ Die Bezeichnungen Abszisse und Ordinate stehen für die x - und die y -Achse im zweidimensionalen Koordinatensystem. Im dreidimensionalen Koordinatensystem heißen die Achsen Abszissen-, Ordinaten- und Applikatenachse. Hier werden die Bezeichnungen Abszisse und Ordinate, aus dem zweidimensionalen Koordinatensystem entlehnt, mit Argument und Funktionswert gleichgesetzt.

In Abb. 6.2-1 sind die Variation der Tastweite auf der Argumentachse und der damit verbundene Informationsverlust veranschaulicht. In Abb. 6.2-2 ist der zusätzliche Informationsverlust gegenüber der Abtastung in Abb. 6.2-1 durch Verringerung der Speichertiefe von 8 bit auf 5 bit dargestellt.

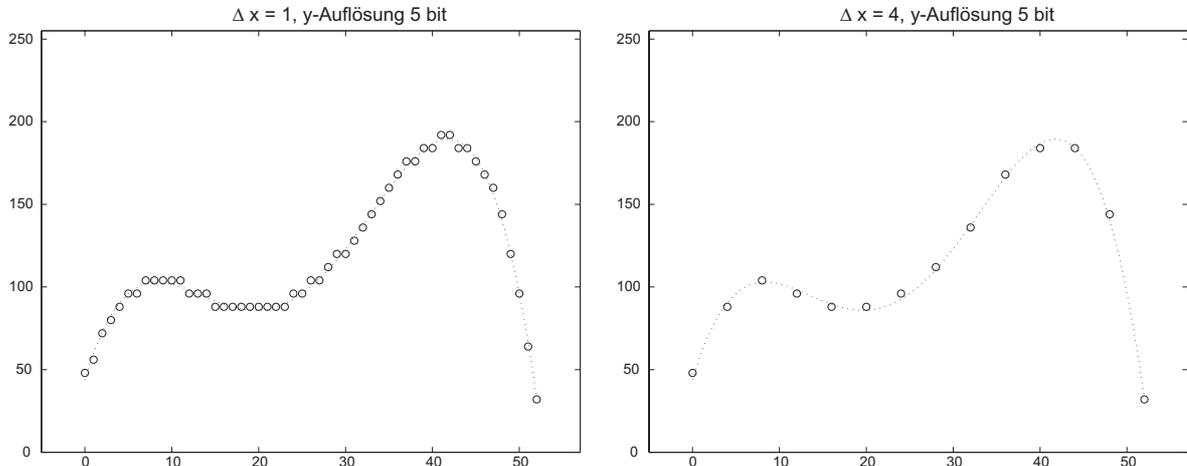


Abbildung 6.2-2: Unterschiedliche Ordinatenaufnahme durch Variation der Speichertiefe (5 bit im Vergleich zu 8 bit in Abb. 6.2-1).

Bei diesen Varianten der Datenreduktion handelt es sich gewissermaßen um eine Kompression im Ortsbereich. Die gleichen Betrachtungen kann man im Waveletbereich anstellen, wobei jedoch Argumente und Funktionswerte der Wavelettransformierten anders zu interpretieren sind. Die gegebene Funktion wird nach der Transformation bezüglich einer anderen Funktionsbasis (vgl. Kap. 2) beschrieben. Die Argumente der Wavelettransformierten sind die in der Funktionsbasis enthaltenen Wavelets und die Funktionswerte die Waveletkoeffizienten, d. h. die Gewichte, mit denen die Wavelets in die Linearkombination, die die Ausgangsfunktion beschreibt, eingehen. Diese andere Interpretation von Argument und Funktionswert ermöglicht andere Kriterien für das Weglassen von Informationen. Ein wesentliches Grundprinzip ist das Nullsetzen eines Teils der Waveletkoeffizienten und in Folge eine Datenreduktion durch eine effektivere Speicherung schwach besetzter Matrizen. Die Kriterien, nach denen Waveletkoeffizienten zu Null gesetzt werden, führen zu obiger Unterscheidung der Kompressionsverfahren nach BETHGE et al. (1997).

6.2.1 Sperrfilter

Die Wavelettransformation bildet Signalanteile unterschiedlicher Frequenz auf unterschiedliche Skalen ab. Jedoch muß diese Aussage durch den Zusatz „vorwiegend“ präzisiert werden, denn es findet im Gegensatz zur Fouriertransformation keine reine Frequenzanalyse statt. Die Wavelets haben keine Frequenz im Sinne harmonischer Schwingungen. Durch die Dilatation liegt aber eine der Frequenzvariation vergleichbare Variation der Analysefunktionen vor.

Von BETHGE et al. (1997) wurden diese Effekte bei der stetigen Wavelettransformation untersucht und an einem Beispiel illustriert. Für das Beispiel dienten das Signal $f(t) = A \cdot \sin(\omega_0 t)$ und das Wavelet $\psi(x) = -\sqrt{8/\pi} x e^{-x^2}$, proportional zur ersten Ableitung der Gaußfunktion. Für die Wavelettransformierte erhielt man

$$L_\psi f(a, b) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \widehat{L_\psi f}(a, \omega) e^{j\omega b} d\omega =: A_L f(\omega_0 b + (\pi/2)) \quad (6.2-1)$$

mit der skalenabhängigen Amplitude

$$A_L(a; A, \omega_0) = \frac{1}{2} \sqrt{\pi|a|} \psi(a\omega_0/2) A. \quad (6.2-2)$$

Diese skalenabhängige Amplitude ist für zwei Signale unterschiedlicher Frequenz und Amplitude in Abb. 6.2-3 zu sehen.

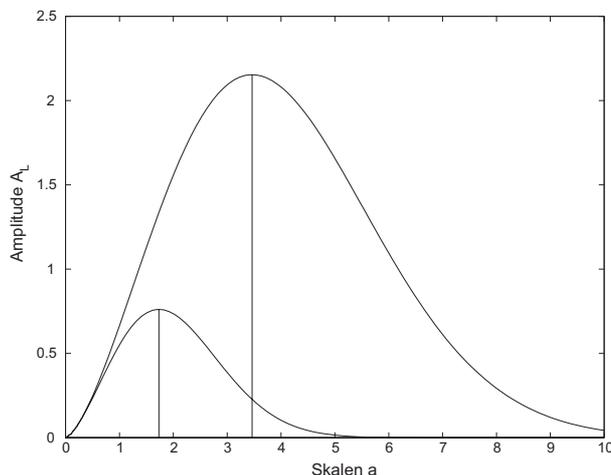


Abbildung 6.2-3: Skalenabhängige Amplituden A_L zweier Sinusschwingungen im WB: $A_1 = 1$, $\omega_{0,1} = 1$, $a_{0,1} = \sqrt{3}$ auf niederen Skalen a konzentriert; $A_2 = 2$, $\omega_{0,2} = 1/2$, $a_{0,2} = 2\sqrt{3}$ auf Höheren Skalen a konzentriert (nach BETHGE et al. (1997)).

Es zeigt sich, daß sin-Schwingungen je nach Frequenz auf unterschiedlichen Skalen dominant wiedergegeben werden. Da in den Koeffizienten der niederen Skalen die hohen Frequenzen dominieren, wird durch Nullsetzen der Koeffizienten der ersten n Skalen gewissermaßen ein Sperrfilter realisiert.

Bei der diskreten Wavelettransformation lassen sich diese Beziehungen nicht unmittelbar analytisch herleiten, es ergeben sich nur diskrete Amplituden auf den einzelnen Skalen. In Abb. 6.2-4 sind die diskreten Amplituden, die sich auf den einzelnen Skalen der diskreten Wavelettransformation ergeben, abgebildet. Verwendet wurde für dieses Beispiel das Haar-Wavelet.

Es ist zu sehen, daß sich in gleicher Weise die Amplituden der niedrigeren Frequenz auf höheren Skalen² konzentrieren.

Für die Datenkompression mittels Sperrfilter, das durch Nullsetzen der Waveletkoeffizienten auf den ersten n Skalen realisiert werden kann, wären die den Skalen zuzuordnenden Frequenzen interessant. Wie schon gesagt wurde, ist das nur in gewisser Analogie zur Fouriertransformation möglich. Man kann z.B. den Skalen diejenigen Frequenzen zuordnen, die in der jeweiligen Skale die maximale Amplitude aufweisen.

Für die diskrete Wavelettransformation wurden empirische Untersuchungen zu einer Frequenzzuordnung zu den Skalen von WEICHSEL (2001) durchgeführt. Die Ergebnisse werden in Abschnitt 6.5.2 im Zusammenhang mit der Ableitung von Glättungsmaßen diskutiert.

Da die Wavelets im Gegensatz zu den sin-Funktionen als Analysefunktionen bei der Fouriertransformation keine harmonischen Schwingungen beschreiben, werden durch das Nullsetzen von Waveletkoeffizienten ganzer Skalen nicht reine Frequenzen abgeschnitten, sondern Formbestandteile der Funktion, die durch die weggelassenen Wavelets bei der Synthese des Signals

² Die Numerierung der Skalen erfolgt hier aufsteigend beginnend mit Eins auf der feinsten Skale.

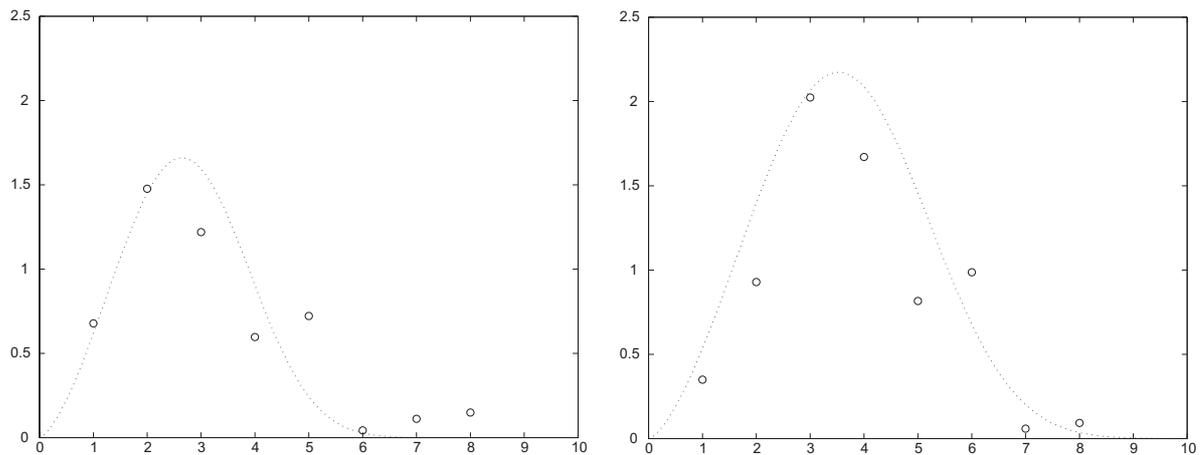


Abbildung 6.2-4: Skalenabhängige Amplituden A_L zweier Sinusschwingungen mit den Frequenzen $\omega_{0,1} = 1$, $\omega_{0,2} = 1/2$ und den hier gleichen Amplituden $A_1 = A_2 = 1$.

beschrieben werden. Die Dilatation der Wavelets von Skale zu Skale ist mit einer Frequenzhalbierung vergleichbar. Dadurch erzielt man bei der Datenkompression nach diesem Verfahren zu einem Sperrfilter analoge Effekte.

6.2.2 Schwellwertverfahren

Schwellwertverfahren im engeren Sinne sind eine Schwellwertbildung auf der Ordinate (BETHGE et al. (1997)). Werden die Waveletkoeffizienten, die eine vorgegebene Schranke nicht übersteigen, zu Null gesetzt und die anderen beibehalten, spricht man von *hard thresholding*. Werden zusätzlich die beibehaltenen Koeffizienten betragsmäßig um den Schwellwert verringert, spricht man von *soft thresholding*. Für diese Filtervorschrift läßt sich leider keine signalunabhängige Filtercharakteristik angeben.

Da bei einer Kompression Informationen verlorengehen, findet im weiteren Sinne ebenfalls eine Generalisierung statt. Die quantitative Analyse im Fourierbereich, wie sie beim Sperrfilter durchgeführt wurde, ist bei dieser Generalisierung jedoch nicht möglich. Hier fließt die Geometrie der Basisfunktionen, der Wavelets, ein.

Das Nullsetzen eines Waveletkoeffizienten w_i bedeutet das Streichen einer Synthesefunktion (des zu dem Waveletkoeffizienten gehörenden Wavelets ψ_i). Damit gilt für die rekonstruierte Funktion f_{rek}

$$f_{rek} = f - w_i \cdot \psi_i . \quad (6.2-3)$$

Dieser Effekt ist in Abb. 6.2-5 für eine mit dem Haarwavelet transformierte sin-Funktion dargestellt.

Der Absolutfehler Δf der rekonstruierten Funktion ist demnach

$$\Delta f = \sum_{i \in I} w_i \cdot \psi_i , \quad (6.2-4)$$

wobei I die Indexmenge der zu Null gesetzten Waveletkoeffizienten ist.

Die Synthesefunktionen ψ_i , deren Waveletkoeffizienten betragsmäßig eine vorgegebene Schranke nicht übersteigen, tragen folglich wenig zur Synthese bei und können deshalb ohne großen Informationsverlust weggelassen, d.h. ihre Waveletkoeffizienten zu Null gesetzt werden.

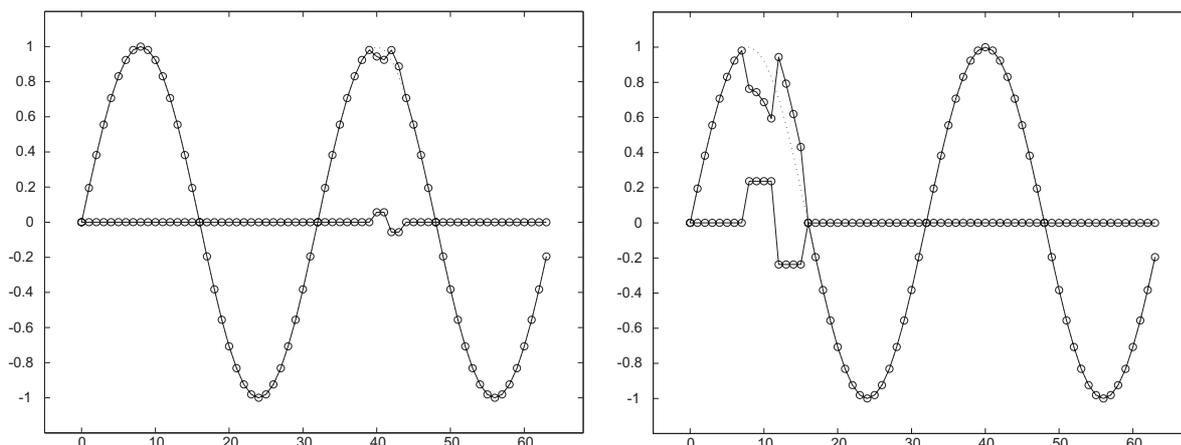


Abbildung 6.2-5: Fehler der rekonstruierten Funktion durch Nullsetzen eines WK der zweiten Skale (links) und eines WK der dritten Skale (rechts).

6.3 Kompression von Geländeflächen

Bei der Kompression von Geländeflächen kommen bei den hier betrachteten digitalen Höhenmodellen über einem regulären Gitter wegen der gleichartigen Datenstruktur die Verfahren der Bild-datenkompression in Betracht. Während bei der Kompression von Bilddaten wegen der hohen Pixelauflösung im Verhältnis zu den verarbeitbaren Daten eine ausreichende Redundanz vorliegt, ist das bei digitalen Höhenmodellen, deren Rasterweite i. allg. bereits dem späteren Zweck mit vorgegebenen Genauigkeitsanforderungen angepaßt ist, nicht der Fall. Eine Abschätzung der möglichen, eher niedrigen Kompressionsraten in Abhängigkeit von den vertretbaren Fehlern im Rahmen der ohnehin vorliegenden Erfassungsfehler wird von MEIER (2002) aus statistischer Sicht vorgenommen.

Tests an Ausschnitten eines realen Geländemodells zeigen die Größenordnung der möglichen Kompressionsrate bei auf Grund der ohnehin begrenzten Höhengenaugkeit des Modells tolerierbaren Höhenfehlern. Als Testdatensatz dient ein digitales Höhenmodell der Schneealpe (ca. 75 km südwestlich von Wien gelegen), das vom Institut für Photogrammetrie der TU Wien zur Verfügung gestellt wurde. Es handelt sich um ein DGM mit Quadratgitter der Gitterweite 25 m. In dem Modell sind auf einer Fläche von ca. 100 km² ca. 171 000 Höhenwerte, 290 Bruchkanten und 538 Formlinien erfaßt.

Die Erfassungsfehler sind von der Neigungszone abhängig und für dieses Modell in Tab. 6.3-1 angegeben.

Tabelle 6.3-1: Testgebiete innerhalb der Neigungsbereiche mit unterschiedlichen Erfassungsfehlern

Neigungs-Zone (in %)	0 - 20	20 - 40	40 - 60	60 - 80	80 - 100	über 100
Erfassungsfehler (in m)	1.30	2.18	2.85	3.12	3.98	6.56
Testgebiete	Gebiet 3			Gebiet 2		Gebiet 1
mittlere Neigung (in %)	27			52		67
Standardabweichung (in %)	21			21		19
Anteil des Testgebiets (in %)	40			52		61

Für die vergleichenden Tests wurden drei Gebiete gewählt, die deutlich unterschiedliche mittlere Neigungen mit relativ kleiner Standardabweichung aufweisen (Abb. 6.3-1). Um den Zusammenhang zu den Neigungszonen sichtbar zu machen, sind die Parameter der gewählten Gebiete in die Tabelle der Neigungszonen integriert worden. Die Daten ergeben sich aus den berechneten Neigungen in den Gitterpunkten. Anteil des Testgebietes bedeutet dabei die relative Häufigkeit der berechneten Neigungen in den Gitterpunkten, die in den Grenzen der Neigungszone liegen.

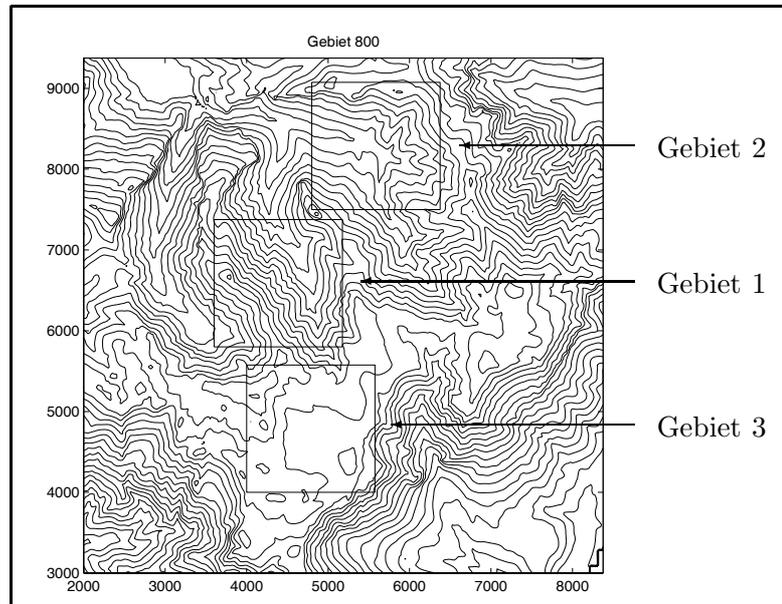


Abbildung 6.3-1: Testgebiet Schneealpe. Geländeausschnitte unterschiedlicher mittlerer Neigung.

Für die Wavelettransformation wurden die sechs in Tab 6.3-2 mit ihrer Ordnung aufgeführten Wavelets des Tools WaveLab (BUCKHEIT et al. (1995)) herangezogen.

Tabelle 6.3-2: Wavelets für den Test der Kompressionsraten

Wavelet	Ordnung
Haar	1
Daubechies4	2
Daubechies6	3
Daubechies8	4
Coiflet3	6
Symmlet4	4

Als Fehlermaß dient die Standardabweichung der Höhendifferenzen zwischen dem Originalhöhenmodell und dem nach Datenkompression im Waveletbereich rekonstruierten Höhenmodell. Für die sechs Wavelets in Abb. 6.3-2 ist die Standardabweichung der Höhendifferenzen jeweils für die drei Teilgebiete des Testdatensatzes über der variierten Kompressionsrate aufgetragen. Die obere Grenze der tolerierbaren Fehler in Abhängigkeit von der Neigungszone ist auf der jeweiligen Kurve markiert. Dadurch ist maximale Kompressionsrate für die drei betrachteten Neigungszonen ablesbar. Mit Ausnahme des Haar-Wavelets ist der Zusammenhang zwischen der Kompressionsrate und dem mittleren Abweichung vom Ausgangsmodell für alle Wavelets etwa

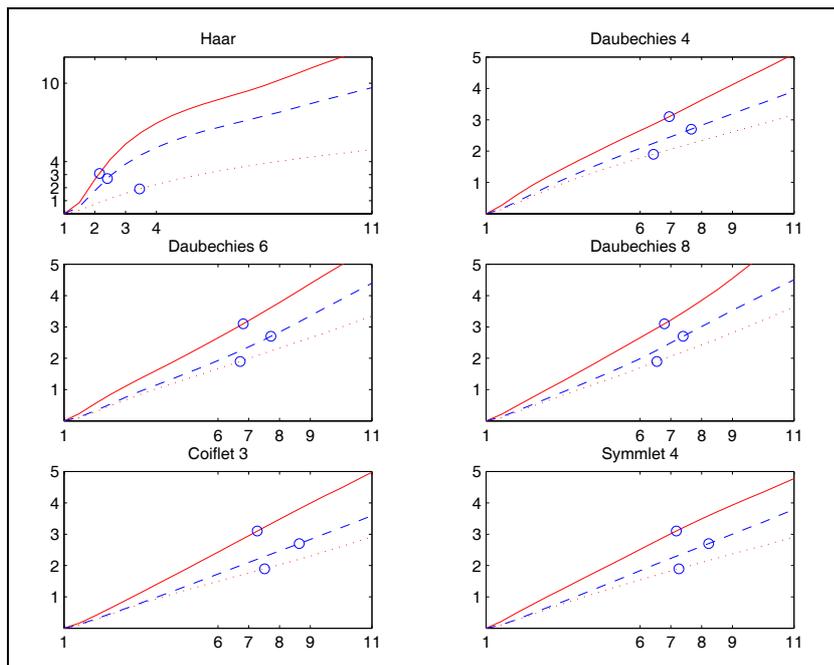


Abbildung 6.3-2: Mittlere Abweichungen der Höhen vom Originalmodell für die drei Geländeausschnitte in Abb. 6.3-1 in Abhängigkeit von der Kompressionsrate. Gebiet 1 (—), Gebiet 2 (- -), Gebiet 3 (···)

gleich. Läßt man eine mittlere Abweichung in der Größenordnung der Erfassungsfehler zu, so liegt die mögliche Kompressionsrate bei den Daubechies-Wavelets etwa bei 7, bei Coiflet3 und Symmlet4 bei etwa 8 und beim Haar-Wavelet bei nur etwa 2 bis 3. Bei allen Wavelets ist eine deutliche Abhängigkeit von der mittleren Neigung sichtbar: Die Fehler durch die Datenkompression nehmen etwa in der gleichen Weise wie die Erfassungsfehler mit wachsender Geländeneigung zu. Eine signifikante Abhängigkeit von der Ordnung des Wavelets ist in diesem Beispiel nicht erkennbar.

Es zeigt sich, daß eine Datenkompression für derartig effektive Geländemodelle nicht sinnvoll ist, es sei denn, die Datenkompression erfolgt im Zusammenhang mit einer gewünschten Generalisierung.

6.4 Kompression von Geländekurven

Im Gegensatz zur Wavelettransformation und Kompression von Geländeflächen ist Kompression von Raumkurven mit Hilfe der Wavelettransformation offensichtlich noch nicht von Interesse gewesen. Das ist möglicherweise darauf zurückzuführen, daß sich Raumkurven nicht durch eine skalare Funktion beschreiben lassen. Hier werden deshalb auf der Basis der Raumkurvenbeschreibungen in 4.3.1 Untersuchungen zur Datenkompression vorgestellt (vgl. auch BEYER (1999)). In 4.3.1 wurde bereits gesagt, daß für die Transformation der Raumkurven eine Parameterdarstellung nach der Bogenlänge bzw. im vorliegenden diskreten Fall eine äquidistante Punktfolge zweckmäßig ist. In einem ersten Schritt ist deshalb die punktweise gegebene Raumkurve neu zu sampeln. Dazu kann z.B. der Walking-Divider-Algorithmus (GOODCHILD und MARK (1987)) benutzt werden. Die Folge $\{(x_i, y_i, z_i), i = 1, \dots, n\}$ aus 3D-Koordinaten wird überführt in Folgen von Werten der in Abschnitt 4.3.1 beschriebenen Signalfunktionen, also $\{x_i, i = 1, \dots, n\}$, $\{y_i, i = 1, \dots, n\}$, $\{z_i, i = 1, \dots, n\}$ für die Koordinatenfunktionen,

$\{\lambda_i, i = 1, \dots, n\}$, $\{\phi_i, i = 1, \dots, n\}$ für die Tangentenwinkelfunktionen und $\{\kappa_i, i = 1, \dots, n\}$, $\{\tau_i, i = 1, \dots, n\}$ für die Krümmungs- und Torsionsfunktion.

Diese diskreten Signalfunktionen werden wavelettransformiert. Für die nachfolgend beschriebenen Tests wurden vier im Tool WaveLab verfügbare Wavelets unterschiedlicher Eigenschaften benutzt.

Mit einem Schwellwertfilter werden die Waveletkoeffizienten komprimiert. Um den Informationsverlust zu quantifizieren, wird nach inverser Wavelettransformation die Originalkurve aus den Signalfunktionen rekonstruiert und der Transformationsfehler ermittelt. Eine Übersicht gibt das Schema in Abb. 6.4-1.

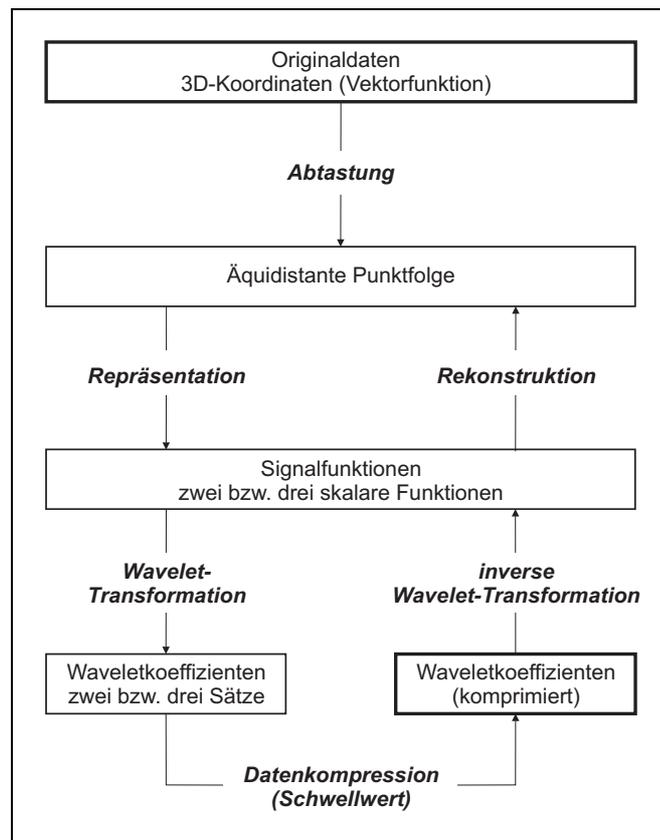


Abbildung 6.4-1: Ablaufschema der Datenkompression bei Raumkurven.

Aus diesen Transformationsschritten ergeben sich unterschiedliche Fehlerquellen, die in ihrer Gesamtheit die Qualität der Rücktransformation beeinflussen.

Abtastfehler entstehen bei der Berechnung der äquidistanten Punktfolge. Die tatsächlichen Abstände weichen von der vorgegebenen konstanten Tastweite ab. Bei der Rückrechnung (Rekonstruktion der Kurve aus den Signalfunktionen) wird jedoch die konstante Tastweite benutzt. Der daraus resultierende Fehler der rücktransformierten Kurve existiert unabhängig von der Datenkompression und kann als Grundfehler des Parametrisierungsverfahrens angesehen werden.

Weiterhin treten *Fehler durch Datenkompression* auf. Der Fehler, der sich aus der Kompressionsrate ergibt, wirkt sich zunächst auf die rücktransformierte Signalfunktion aus.

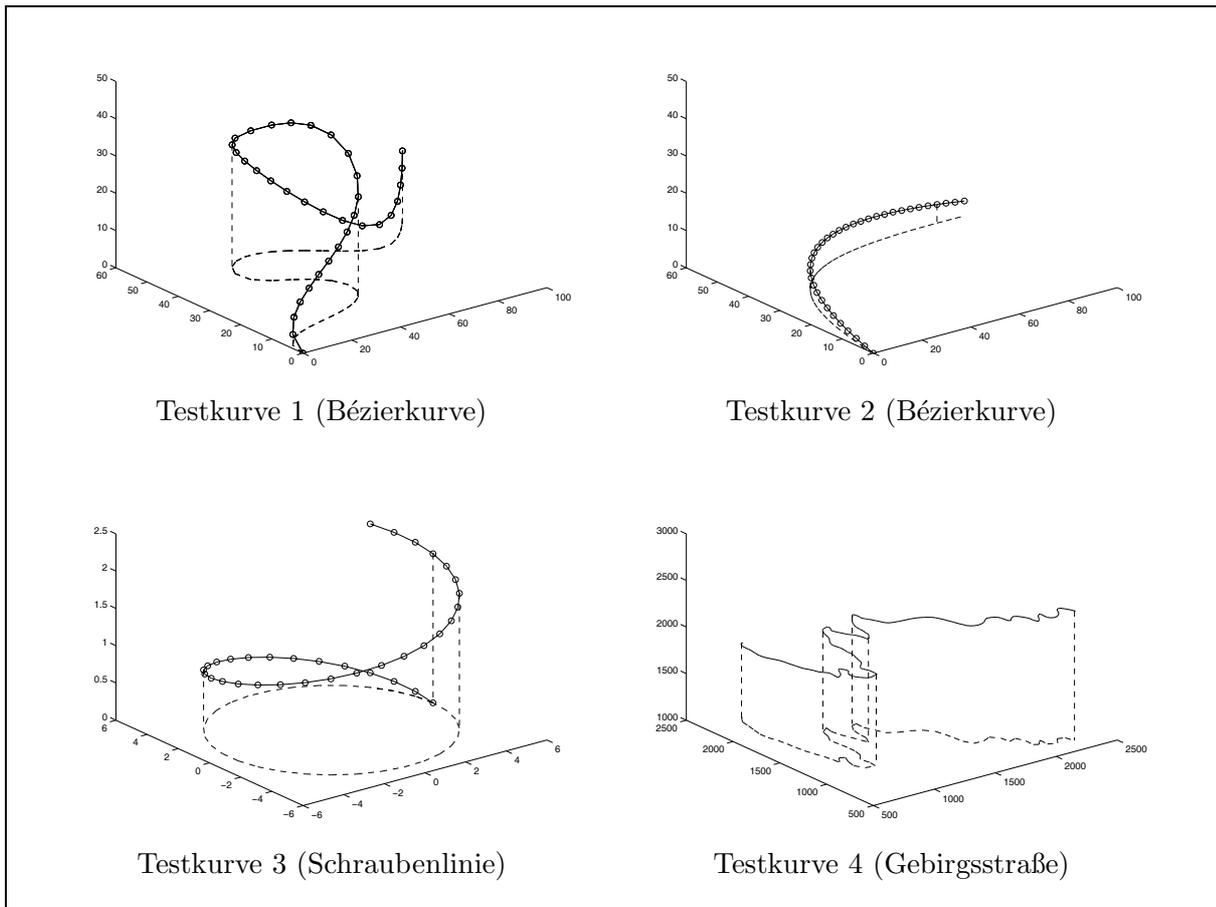


Abbildung 6.4-2: Beispielkurven

Diese Fehler überlagern sich bei der Rekonstruktion der Kurve mit den Abtastfehlern. Die Eignung des verwendeten Wavelets hängt von der Gestalt der Signalfunktion ab. Die Gestalt der Signalfunktion folgt aus der Gestalt der Originalkurve und aus dem Parametrisierungsverfahren; z.B. besitzt eine Gerade als Raumkurve konstante Tangentenwinkelfunktionen und eine Schraubenlinie eine konstante Krümmungs- und eine konstante Torsionsfunktion. Das Fehlerverhalten bezüglich der Datenkompression ist damit bei einer Geraden, parametrisiert mittels Tangentenwinkelfunktionen, und bei einer Schraubenlinie, parametrisiert mittels Krümmungs- und Torsionsfunktion, gleich. Wegen der gegenseitigen Abhängigkeiten der Einflußgrößen ist eine theoretische Analyse der Fehlereffekte sehr schwierig.

Um dennoch Aussagen über die Fehlerentwicklung zu bekommen, wurden Tests an unterschiedlichen Raumkurven vorgenommen. Als Testkurven wurden eine rationale Bézierkurve höherer Ordnung als Beispiel einer beliebig gekrümmten und gewundenen Kurve in zwei Varianten, eine Schraubenlinie mit konstanter Krümmung und Torsion und eine reale Gebirgsstraße (Großglocknerstraße von Hochmaiß bis Fuscher Törl, Hohe Tauern, Österreich) verwendet.

Aus den Parameterdarstellungen bzw. aus den realen diskreten Koordinaten des Straßenzuges wurden bezüglich der Bogenlänge äquidistante Punktfolgen konstruiert. In den hier dargestellten Beispielen sind es $2^5 + 3 = 35$, $2^7 + 3 = 131$ bzw. $2^{10} + 3 = 1027$ Punkte. Die drei zusätzlichen Punkte werden für die diskrete Krümmung und Torsion benötigt (vgl. 4.3.1). Die wesentlichen Charakteristiken der gewählten Beispielkurven sind in Tabelle 6.4-1 gegenübergestellt. Im einzelnen wurden die Varianten Koordinatenfunktionen (KOF), Tangentenwinkelfunktionen (TWF), Krümmungs- und Torsionsfunktion (KTF) untersucht.

Tabelle 6.4-1: Eigenschaften der Beispielkurven

	Bézierkurve	Schraubenlinie	Gebirgsstraße
Original- beschreibung	Parameter- darstellung $\underline{x} = \underline{x}(t)$	Parameter- darstellung $\underline{x} = \underline{x}(t),$ $t \sim s$	digitalisierte Punktfolge
Parametrisierung nach Bogenlänge	iterativ durch Parametervariation	Umrechnung über Proportionalitäts- faktor	Walking Divider Algorithmus
geometrische Eigenschaften	variierbare Kurven- form (durch Änderung der Steuerpunkte)	konstante Krümmung, konstante Torsion	reale Geländedaten (zufälliger Verlauf)
Abtastfehler	groß	klein	klein
beste Repräsentation	TWF	KTF	TWF

Die dargestellten Testreihen in Abb. 6.4-3 zeigen den durchschnittlichen Fehler der Kurvenpunkte nach der Rekonstruktion in Abhängigkeit von der Kompressionsrate. Der Fehler wurde zusätzlich nach der Tastweite normiert. Dieser Fehler ist nur bei den Koordinatenfunktionen stationär bezüglich der Bogenlänge s . In den Abbildungen sind für die drei Verfahren (KOF: ausgezogen, TWF: gerissen, KTF: punktiert) die Fehler in Abhängigkeit von der Kompressionsrate dargestellt. Nach LOUIS et al. (1994) gibt die Kompressionsrate das Verhältnis des Speicherplatzbedarfs der Originaldaten (S_{orig}) zum Speicherplatzbedarf der komprimierten Daten (S_{komp}) an: $k = S_{orig}/S_{komp}$. In den Grafiken ist im Unterschied zu dieser Definition die Kompression als Verhältnis $r = S_{komp}/S_{orig}$ in Prozent angegeben. Es wurden jeweils die Wavelets Haar, Daubechies 8, Coiflet 3 und Symmlet 4 untersucht.

Die Tests zeigen, daß die jeweils erkennbare Rangordnung zwischen den drei Verfahren nicht zu verallgemeinern ist. Die Tangentenwinkelfunktionen liefert jedoch in allen Beispielen die kleinsten Fehler nach der Datenkompression. Mit Ausnahme des Haar-Wavelets sind die Unterschiede zwischen den Test-Wavelets gering. Die Kurvenform ist die Hauptursache für die Fehler. Die sehr unterschiedlich gekrümmte und gewundene Test-Kurve 1 läßt ohne nennenswerte Fehler Kompressionen auf 40% des Datenumfangs zu. Bei der schwach gekrümmten Test-Kurve 2 sind es 15%. Bei der sehr speziellen Schraubenlinie mit konstanter Krümmung und Windung (Test-Kurve 3), die in der Praxis näherungsweise z.B. bei Gebirgsbahnen vorkommt, ist die mögliche Kompression bei der Repräsentation KTF noch wesentlich höher.

Um reliefbezogene Raumkurven in den Waveletbereich zu transformieren, eignen sich prinzipiell alle drei untersuchten Parametrisierungen. Die untersuchten Kurventypen sind mit Ausnahme der Testkurve 2 Extrembeispiele. Derartig stark strukturierte Kurven kommen an natürlichen und verbauten Reliefs eher selten vor. Sie stellen selbst in den Alpenländern nicht den häufigsten Fall dar. Sie sind jedoch für „robuste“ Tests geeignet. Insgesamt kann festgestellt werden, daß die Tangentenwinkelfunktion in ihrer räumlichen Verallgemeinerung meist die größte Kompressionsrate bei vorgegebener Fehlerschranke ermöglicht. Bei speziellen Kurven, z. B. der Schraubenlinie, ist die Parametrisierung nach Krümmung und Torsion die beste Kurvenrepräsentation für eine Datenkompression mittels Wavelettransformation. Die Abhängigkeit der untersuchten Verfahren von den vielfältigen speziellen Kurvenformen und den daraus abgeleiteten Signalfunktionen

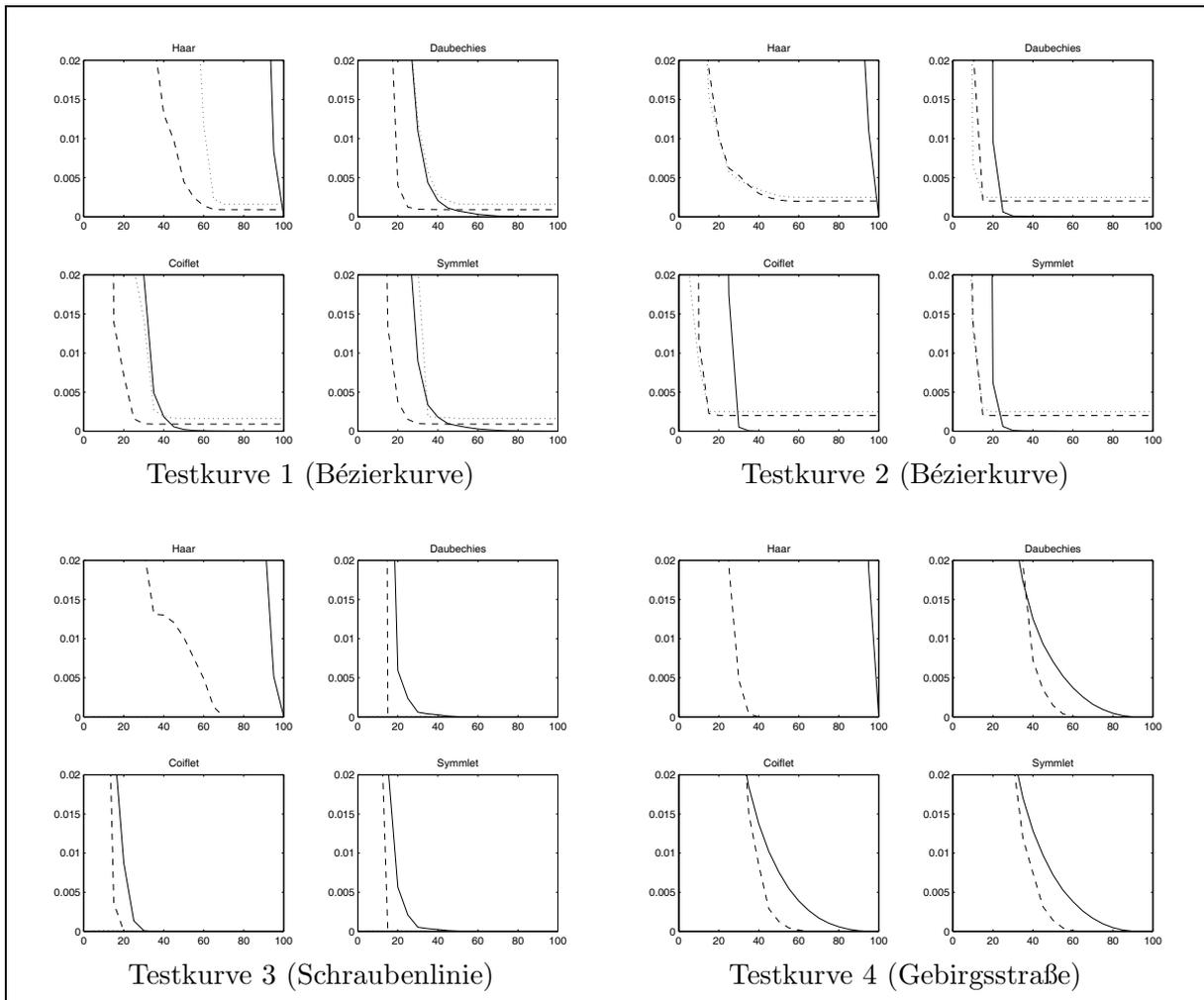


Abbildung 6.4-3: Fehler der Beispielkurven

erschwert eine übersichtliche Quantifizierung der auftretenden Effekte. Besonders deutlich ist das an den Testkurven 1 und 2 zu sehen. Detaillierte quantitative Aussagen zu den untereinander abhängigen Merkmalen Kurvenform, Kurvenrepräsentation, Kompressionsrate, Wavelet und Fehler sind nur über eine Klassifizierung der Originalkurven nach ihren Signalfunktionen möglich und könnten Gegenstand weiterer Untersuchungen sein.

6.5 Probleme bei hybriden Daten

6.5.1 Gemeinsame Kompression von Vektor- und Rasterdaten

Wie bereits in Kapitel 4 gesagt, werden hier hybride Geländemodelle betrachtet, die auf einer Flächenbeschreibung durch ein DHM auf regulärem Gitter (Rasterdaten) und einer Beschreibung von Geländelinien (insbesondere Bruchkanten) in Form von Polygonzügen (Vektordaten) beruhen. Zur Datenkompression im Waveletbereich ist das Geländemodell zunächst in den WB zu transformieren. Die Standardalgorithmen der Wavelettransformation setzen eine Matrixstruktur der zu transformierenden Funktion voraus. Diese ist für ein DHM über einem regulären Gitter in Form der Höhenmatrix unmittelbar gegeben. Die im hybriden Modell i.allg. als Polygone vorliegenden Kanten erfüllen diese Voraussetzung nicht unmittelbar. In Abschnitt 4.3 wurden verschiedene Möglichkeiten der Kurvenbeschreibung diskutiert. Aus der Sicht der diskreten Wavelettransformation ist die Beschreibung in einer Matrixstruktur (vgl. 4.3.2) die zweckmäßigste. Die Wavelettransformation kann dann vollkommen homogen auf Fläche und Kanten angewendet werden. Die inhaltliche Interpretation der Waveletkoeffizienten bleibt indes wegen des unterschiedlichen geometrischen Gehalts verschieden. Das hat natürlich Auswirkungen auf die Datenkompression im Waveletbereich. In 6.1 wurde darauf hingewiesen, daß das Vorhandensein redundanter Daten Voraussetzung für eine erfolgreiche verlustbehaftete Kompression ist. Bei einer Kompression mittels Schwellwertbildung auf der Abszisse findet eine Tiefpaßfilterung statt, deren Filtercharakteristik angegeben werden kann. Für die Kompression der Flächenfunktion bedeutet das, daß im Sinne einer eventuell sogar gewünschten Generalisierung die herausgefilterten Feinstrukturen ohnehin als redundant anzusehen sind. Bei den in eine Matrixstruktur transformierten Kantendaten kann man jedoch nicht von einer Generalisierung sprechen, da die Geometrie der Kantenfunktion keinen unmittelbaren Bezug zur Kantengeometrie hat. Es ist damit zwar einerseits eine getrennte Kompression mit unabhängig wählbaren Kompressionsraten möglich, andererseits aber kein meßbarer Zusammenhang ausnutzbar.

6.5.2 Glättungsmaße

Frequenzzuordnung zu den Skalen

Bei der Wavelettransformation eines Signals erfolgt analog zur Fouriertransformation eine Zerlegung in Frequenzanteile, die wegen der Lokalisierungseigenschaft der Wavelettransformation als transformiertes Signal auf verschiedene Skalen abgebildet werden. Um die vom gewählten Wavelet abhängige Frequenzzuordnung zu den einzelnen Skalen zu bestimmen, kann man die Amplituden wavelettransformierter sin-Funktionen mit variierender Frequenz untersuchen. In Abb. 6.5-1 sind die Amplituden über den Skalen für fünf verschiedene Frequenzen dargestellt. Der Skale mit maximaler Amplitude ist dann die Frequenz der zur Wavelettransformation verwendeten sin-Funktion zuzuordnen.

Das Maximum einer nur diskret gegebenen Funktion kann man mit Hilfe einer approximierenden stetigen Funktion bestimmen. Auf Grund ihrer Form könnte sich die Funktion

$$f(a, b, c; x) = a \cdot x^b \cdot e^{cx} \quad \text{mit} \quad c < 0, \quad b > 1 \quad (6.5-1)$$

eignen.

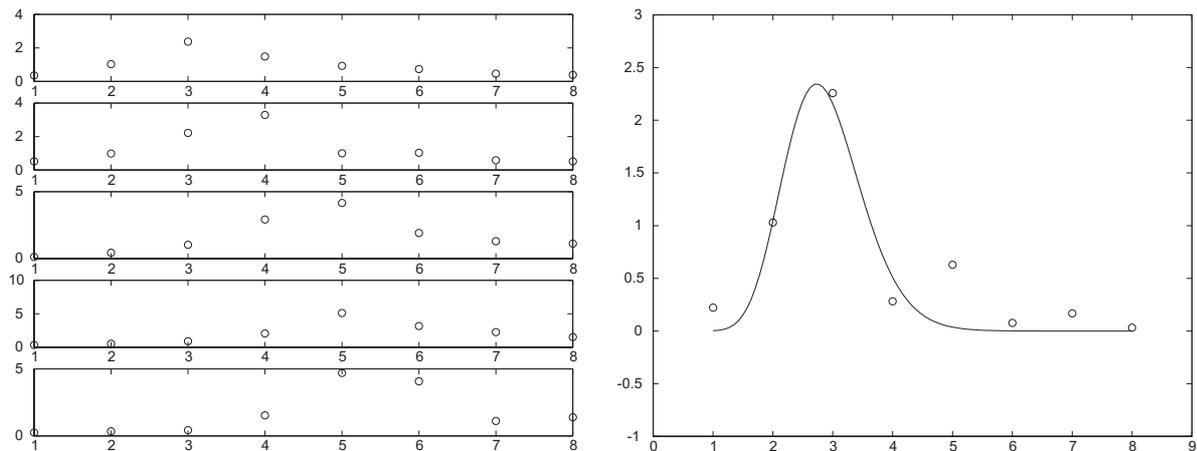


Abbildung 6.5-1: Skalensabhängige Amplituden A_L einer Sinusschwingung mit den Frequenzen $\omega = 2\pi/10$ bis $\omega = 2\pi/50$ (links) und Modellapproximation für $\omega = 2\pi/10$ (rechts).

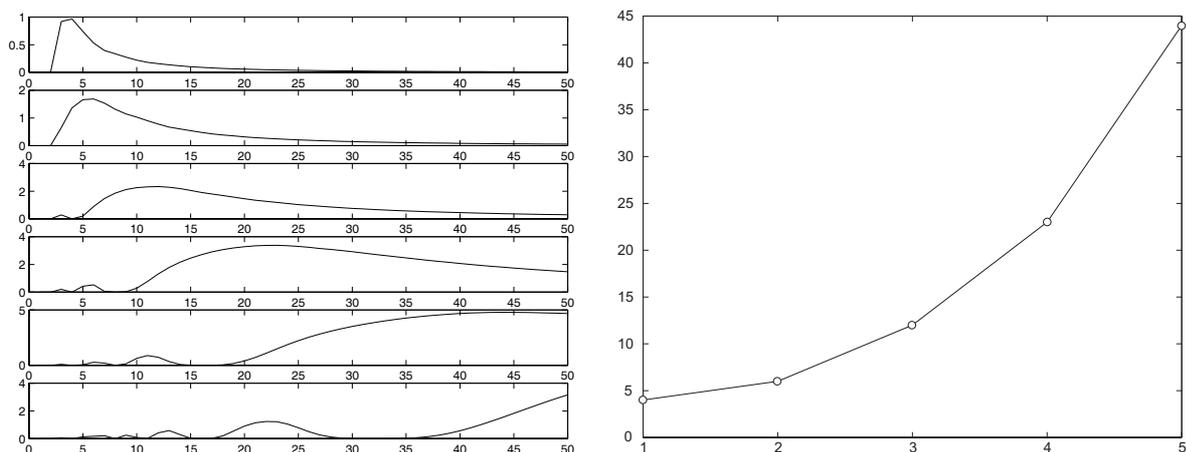


Abbildung 6.5-2: Amplituden A_L in den einzelnen Skalen in Abhängigkeit von der Frequenz (links). Die Maximumstellen entsprechen der auf der jeweiligen Skale dominierenden Frequenz. Diese Frequenzen sind im rechten Bild über den Skalen aufgetragen.

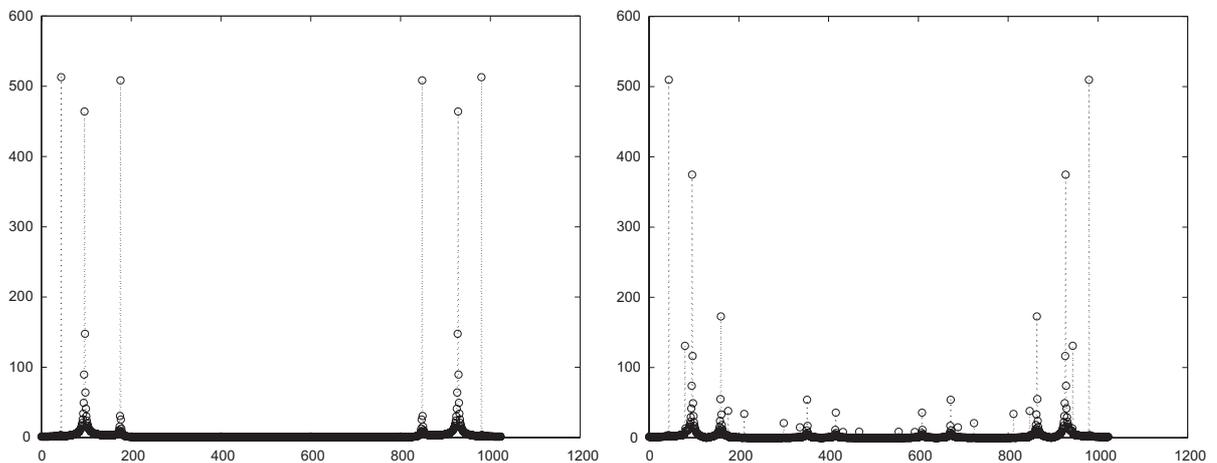
In der ersten Skale liefert die Modellapproximation z. B. für die Frequenz $\omega = \frac{2\pi}{10}$ mit den Startvorgaben $x_A = 3$, $y_A = 2$; $x_D = 1$ nach zweifacher Iteration die Maximumstelle $x_A = 2.7315$ (Abb. 6.5-1).

Die gewählte approximierende Funktion für die Modellapproximation führt in den höheren Skalen, d.h. für niedrigere Frequenzen zu Instabilitäten, die möglicherweise auf die unzulässig starke Vereinfachung des Modells durch die notwendige Linearisierung zurückzuführen sind. Durch einen anderen Zugang wird pro Skale durch Variation der Frequenz der Maximalwert der Amplitude gesucht, der damit die Frequenz liefert, die dem Maximum der Modellapproximation auf der jeweiligen Skale entspricht (Abb. 6.5-2).

Von WEICHSEL (2001) wurden diese Analysen detailliert durchgeführt. Für die dort untersuchten Wavelets ergaben sich die in Tab. 6.5-1 angegebenen dominierenden Frequenzen für die ersten neun Skalen.

Tabelle 6.5-1: Dominierende Kreisfrequenzen in den einzelnen Skalen (aus WEICHSEL (2001))

Skale	Haar	Daubechies4	Daubechies8	Coiflet3	Symmlet4
1	3,14	3,14	3,25	3,16	3,09
2	1,23	1,14	1,097	1,079	1,097
3	0,590	0,566	0,548	0,540	0,548
4	0,292	0,282	0,274	0,270	0,274
5	0,146	0,141	0,137	0,135	0,137
6	0,0729	0,0707	0,0685		0,0685
7	0,0364	0,0354			
8	0,0182				
9	0,0091				

Abbildung 6.5-3: Fouriertransformierte der Funktion $\sin(\omega_2 t) + \sin(\omega_3 t) + \sin(\omega_4 t)$ (links) und der rekonstruierten Funktion mit dem Wavelet Coiflet3 (rechts).

Mit Hilfe einer Funktion, die aus \sin -Funktionen der auf den Skalen dominierenden Frequenzen additiv zusammengesetzt ist, soll nachfolgend die Wirkung des Streichens der hochfrequenten Skalen demonstriert werden. Benutzt wird die Funktion

$$f(t) = \sin(\omega_2 t) + \sin(\omega_3 t) + \sin(\omega_4 t), \quad (6.5-2)$$

wobei $\omega_2, \omega_3, \omega_4$ die in Tab. 6.5-1 angegebenen Kreisfrequenzen der 2., 3. und 4. Skale des jeweils betrachteten Wavelets sind. In Abb. 6.5-3 ist die diskrete Fouriertransformierte bei Verwendung des Wavelets Coiflet3 zu sehen. Um die Änderung der Frequenzanteile besser sichtbar zu machen, werden in den nachfolgenden Beispielen nur noch die Peaks dargestellt. Zunächst ist diese „schematische“ Darstellung der Fouriertransformierten in Abb. 6.5-4 zum Vergleich noch einmal für das Beispiel in Abb. 6.5-3 zu sehen.

Es wurden im Waveletbereich die 1. und die 2. Skale gestrichen. Es ist zu erkennen, daß die Frequenz ω_2 deutlich gegenüber den Frequenzen ω_3 und ω_4 abgeschwächt wurde. Ein reines Sperrfilter kann die Wavelettransformation jedoch nicht realisieren.

Für das Daubechies4- und das Haar-Wavelet ergeben sich ganz analoge Ergebnisse, die in Abb. 6.5-5 und 6.5-6 zu sehen sind.

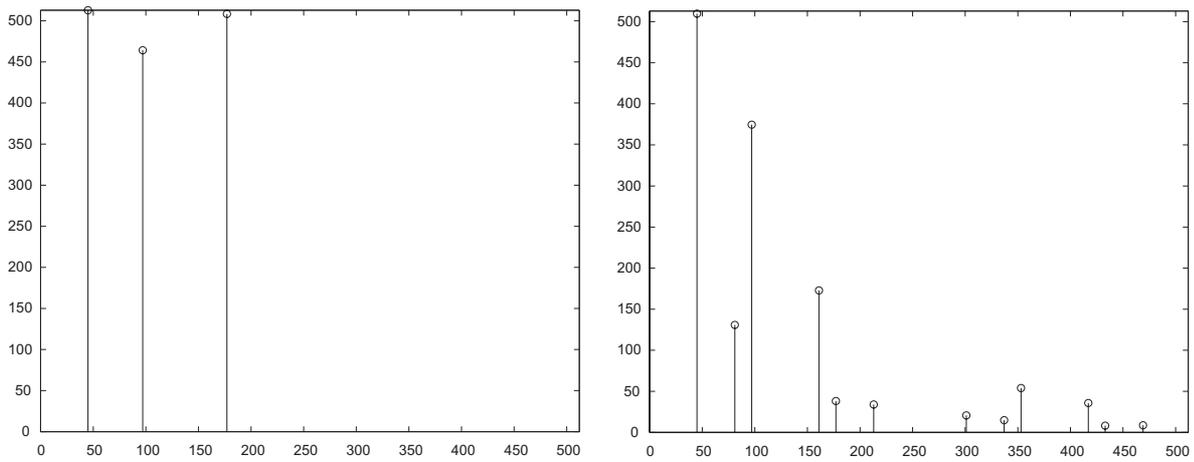
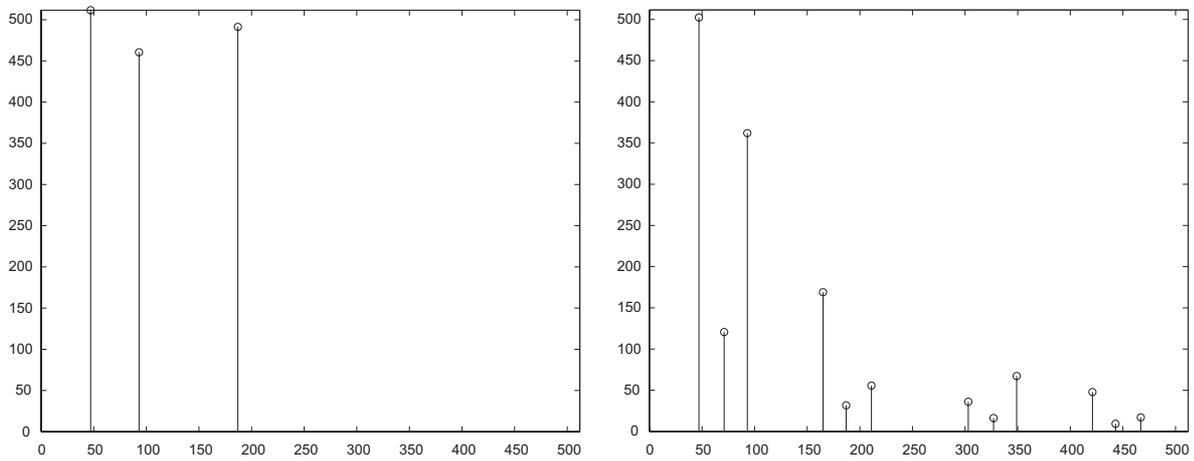
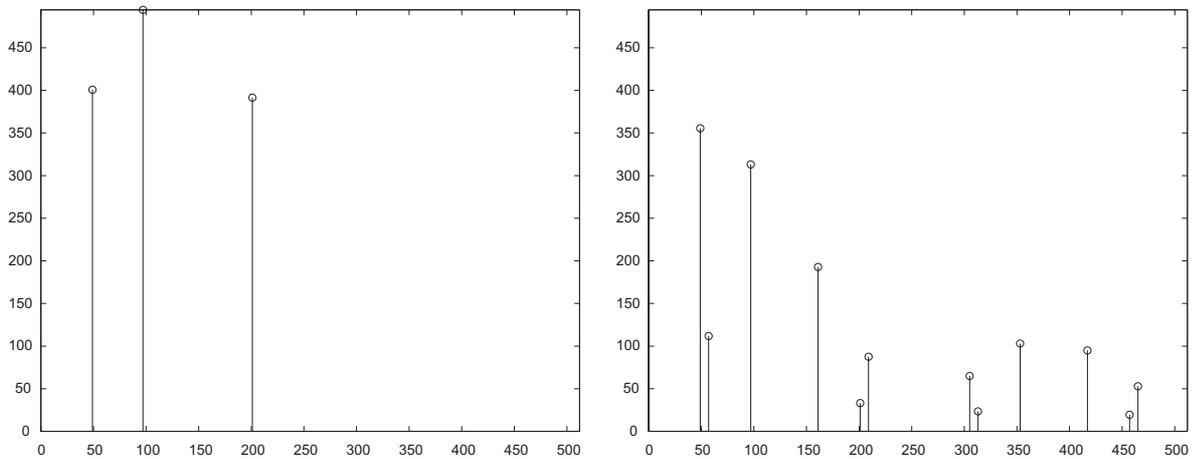


Abbildung 6.5-4: Schematische Darstellung der Fouriertransformierten aus Abb. 6.5-3.

Abbildung 6.5-5: Fouriertransformierte der Funktion $\sin(\omega_2 t) + \sin(\omega_3 t) + \sin(\omega_4 t)$ (links) und der rekonstruierten Funktion nach Kompression mit dem Daubechies4-Wavelet (rechts).Abbildung 6.5-6: Fouriertransformierte der Funktion $\sin(\omega_2 t) + \sin(\omega_3 t) + \sin(\omega_4 t)$ (links) und der rekonstruierten Funktion nach Kompression mit dem Haar-Wavelet (rechts).

Frequenzen und Krümmungen

In diesem Abschnitt wird der Frage nachgegangen, ob man einen Zusammenhang zwischen den vorkommenden Gelände­krümmungen und den auf den einzelnen Skalen dominierenden Frequenzen finden kann. Ein Gelände­profil in Form einer sin-Funktion bestimmter Frequenz hat in den einzelnen Punkten eine Krümmung, die von der Lage innerhalb der Periode abhängig ist. Die Frage ist nun, in welcher Weise sich die Krümmung bei Frequenzänderung in jedem dieser Punkte ändert. Betrachtet man den Krümmungskreis in einem Punkt der sin-Funktion der Frequenz f_0 ,

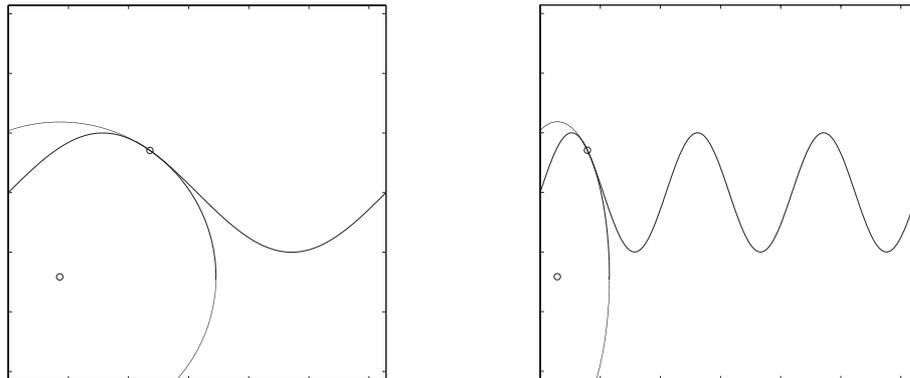


Abbildung 6.5-7: Der Krümmungskreis (links) in einem Punkt der sin-Kurve wird durch Stauchung der sin-Funktion in x -Richtung durch die gleiche Stauchung in die Schmiegeellipse (rechts) transformiert.

so wird sich dieser bei einer Frequenzänderung in eine Ellipse verformen, wobei nur die Achse parallel zur x -Achse gestreckt bzw. gestaucht wird, da die Frequenzänderung einer Streckung bzw. Stauchung der sin-Funktion in x -Richtung entspricht. Gilt für den Krümmungskreis (die Ellipse in der Ausgangssituation) $a_0 = b = r$, so ändern sich die Parameter der Ellipse bei einer Frequenzänderung mit $f = \alpha \cdot f_0$ in $a = \frac{1}{\alpha} \cdot a_0$, wobei b konstant bleibt.

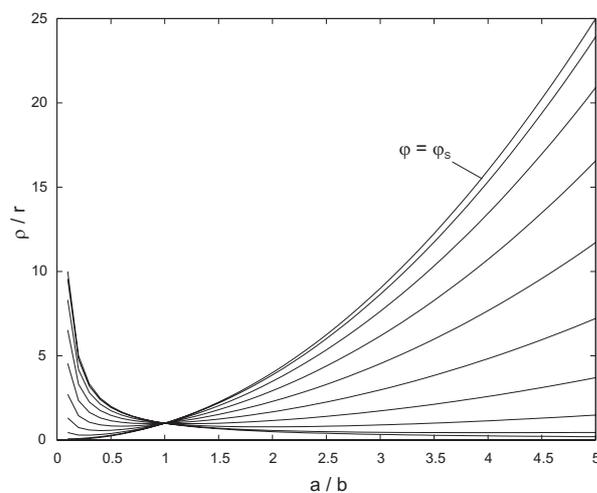


Abbildung 6.5-8: Änderung des Krümmungsradius in Abhängigkeit von der Lage des Punktes zwischen Maximum und Wendepunkt. Die Kurve, die rechts von (1,1) am steilsten ist, beschreibt die quadratische Abhängigkeit im Maximum der sin-Kurve.

Eine Ellipse, die durch Stauchung oder Streckung in x -Richtung aus einem Kreis mit dem Radius r entsteht, hat die Halbachsen $a = \frac{1}{\alpha} \cdot r$ und $b = r$. Polarkoordinaten (r, ϕ) hat, berechnet sich zu

$$\varrho = \frac{(a^2 \sin^2 \phi + b^2 \cos^2 \phi)^{\frac{3}{2}}}{ab} . \quad (6.5-3)$$

Speziell im Scheitel der Ellipse ($\phi_S = \frac{\pi}{2}$) gilt mit $\sin \phi_S = 1$ und $\cos \phi_S = 0$

$$\varrho = \frac{(a^2)^{\frac{3}{2}}}{ab} = \frac{a^2}{b} . \quad (6.5-4)$$

D.h., die Krümmung im Maximum der sin-Kurve (Scheitelpunkt der Schmiegeellipse) ist umgekehrt proportional zum Quadrat der Frequenz. In einem allgemeinen Punkt der Ellipse liegt keine Proportionalität zur Frequenz vor. Der allgemeine Zusammenhang ist in Abb. 6.5-8 dargestellt.

6.5.3 Kantenanpassung

Während bei einer Kompression der Geländefläche im Waveletbereich, insbesondere mittels Sperrfilter (6.2.1), mit einer Glättung des Reliefs zu rechnen ist, kann man bei einer Kompression der Geländekanten nach Transformation der Kantenmatrix in den Waveletbereich keine gleichartigen Glättungseffekte erwarten. In 5.3 wurde außerdem bereits eingeschätzt, daß eine verlustbehaftete Datenkompression der Kantenmatrix nicht zweckmäßig ist. Man kann deshalb davon ausgehen, daß nach Rücktransformation und Zusammenführung der Flächen- und Kanteninformationen die ungeglätteten Kanten auf die geglättete Fläche aufgesetzt werden. In BEYER (2000b) wurde bereits auf dieses Problem hingewiesen. Da der Kantenpunkt (Gitterschnittpunkt) nach der inversen Transformation durch Rückwärtsschnitt aus den extrapolierten Höhen rekonstruiert wird, bleibt die in den extrapolierten Höhen gespeicherte Überhöhung erhalten (Abb. 6.5-9).

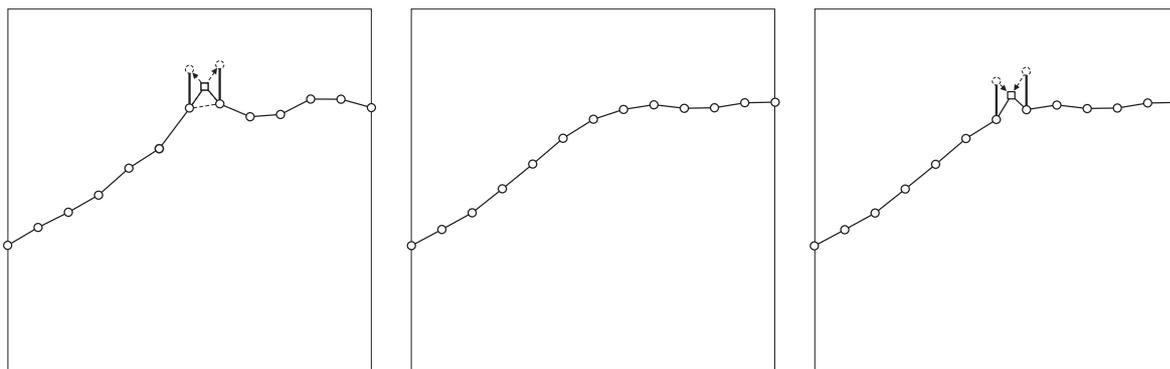


Abbildung 6.5-9: Notwendigkeit einer Kantenanpassung nach Datenkompression im WB. Originalprofil mit Gitterschnittpunkt und schematischer Darstellung der Extrapolation (links). Geglättetes Profil (Mitte). Profil mit aufgesetztem Gitterschnittpunkt durch Rückwärtsschnitt aus den extrapolierten Höhen (rechts).

Im linken Bild ist ein Originalprofil in einer Gitterebene mit einem Gitterschnittpunkt abgebildet. Die Operation „Extrapolation“ ist dabei graphisch veranschaulicht. Der Glättungseffekt bei der Geländefläche durch Datenkompression im Waveletbereich ist nach der Rücktransformation im mittleren Bild zu sehen. Wird nun der Gitterschnittpunkt mittels der extrapolierten Höhen

auf das Profil aufgesetzt, bleibt die Überhöhung erhalten, und im Gitterschnittpunkt hat man einen zu kleinen, nicht zum Geländeverlauf passenden Schnittwinkel.

In der Diplomarbeit WEICHSEL (2001) wurden deshalb Untersuchungen zur Kantenanpassung durchgeführt, die an dieser Stelle kurz vorgestellt werden sollen. Es wurden die folgenden Strategien unterschieden:

1. Globale Anpassung

Die bei einer Datenkompression durch Sperrfilter zu erwartende allgemeine Glättung des Reliefs könnte durch eine globale Glättung der Kanten zu einer geeigneten Anpassung der Kanten an das geglättete Relief führen. Es wäre dann eine allgemeine (globale) Vorschrift zur Korrektur der extrapolierten Höhen möglich.

(a) Krümmungsanpassung aus dem Verhältnis der Frequenzen

Aus dem Zusammenhang zwischen Frequenz und Krümmung und der in Tab. 6.5-1 aufgeführten dominierenden Frequenzen kann man ein Krümmungsverhältnis der Maximalkrümmungen vor und nach der Kompression je nach Anzahl der gestrichenen Skalen bei Kompression mittels Sperrfilter bestimmen. Dieses Krümmungsverhältnis soll zur Kantenanpassung genutzt werden.

Es wird dabei der Kreis durch die Gitterpunkte des Reliefs und den Gitterschnittpunkt gemäß dem oben genannten Krümmungsverhältnis im Radius geändert und der Gitterschnittpunkt auf den neuen Kreis projiziert.

(b) Verhältnis der Dreieckshöhen aus Tangentenschnittpunkt der maximalen Frequenz

Zwei sin-Schwingungen mit der dominierenden Frequenz der feinsten Skale und der dominierenden Frequenz der ersten nicht gestrichenen Skale werden als jeweilige Profilapproximation vor und nach der Datenkompression betrachtet. Das Höhenverhältnis des Tangendendreiecks wird dann zur Reduzierung der Überhöhung des rekonstruierten Gitterschnitts benutzt.

2. Lokale Anpassung

Ohne Ausnutzung der in 6.5.2 untersuchten Glättungseffekte beim Sperrfilter ist nach der Rücktransformation eine lokale Anpassung über eine Analyse des „geglätteten“ Reliefs denkbar. Diese Strategien nutzen zusätzlich das Originalrelief, das in der praktischen Anwendung nach der Datenkompression nicht mehr vorhanden ist. Sie haben deshalb bei den Untersuchungen nur vergleichenden Charakter.

(a) Verhältnis aus Dreieckshöhe zu Kantenhöhe

Der Generalisierungsfehler des Reliefs wird auf den Kantenpunkt übertragen, indem die Höhe des rücktransformierten Gitterschnitts im Höhenverhältnis des Reliefs korrigiert wird.

(b) Verhältnis der Dreieckshöhe aus Geradenschnitt

Die „Krümmungs“änderung des Geländes, die über die Schnittwinkel der angrenzenden Maschenseiten in der betrachteten Gitterebene ermittelt wird, dient zur Korrektur der Höhe des rücktransformierten Gitterschnitts.

Im Resümee der Diplomarbeit wird zusammenfassend festgestellt, daß die globalen Anpassungsverfahren bei Kompression mittels Sperrfilter die notwendige Abschwächung liefern, die jedoch zu stark ausgeprägt ist. Bei der Kompression mittels Schwellwertbildung zeigte sich in den Tests, daß eine Kantenanpassung wegen einer relativ geringen Glättung als nicht notwendig erscheint. Die umfangreichen Tests wurden am Testdatensatz Schneetalpe des IPF der TU Wien durchgeführt.

7 Analyse von Geländemodellen mittels Wavelettransformation (Folgeprodukte)

7.1 Approximationseigenschaften

7.1.1 Theoretischer Hintergrund

Die Wavelettransformierte eines Signals, multipliziert mit einem skalenabhängigen Faktor, approximiert die n -te Ableitung des Signals durch Grenzübergang des Skalenparameters $a \rightarrow 0$, sofern das benutzte Wavelet von n -ter Ordnung ist. Darunter versteht man solche, deren Mittelwert und die ersten $n - 1$ Momente verschwinden, während das n -te Moment endlich und ungleich Null ist. Die Ordnung des Wavelets bestimmt das Verhalten der Wavelettransformierten für kleine a . In der Monographie LOUIS et al. (1994) ist dieses sog. Hochfrequenzverhalten mit funktionalanalytischen Methoden, einschließlich distributiver Ableitungen, behandelt. Letztere können von Interesse sein, wenn man an hybride DGM mit Kanteninformation denkt.

Sei

$$L_\psi f(a, b) := \sqrt{\frac{1}{c_\psi |a|}} \int_{-\infty}^{\infty} f(x) \psi\left(\frac{x-b}{a}\right) dx \quad (7.1-1)$$

die Wavelettransformierte des Signals $f(x)$ (hier ein Höhenprofil) zum Wavelet ψ , mit dem Skalenparameter a , dem Ortsparameter b und dem Normierungsfaktor

$$c_\psi := 2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\omega)|}{|\omega|} d\omega, \quad (7.1-2)$$

wobei $\hat{\psi} := \mathcal{F}\{\psi\}$ die Fouriertransformierte (FT) von ψ ist. Dann wird die 1. Ableitung $f'(b)$ im Punkt b durch den Grenzwert

$$f'(b) = \lim_{a \rightarrow 0} \frac{\sqrt{c_\psi}}{a^{3/2} m_1} L_\psi f(a, b) \quad (a > 0) \quad (7.1-3)$$

mit dem nicht verschwindenden 1. Moment des Wavelets

$$m_1 := \int_{-\infty}^{\infty} x \psi(x) dx \neq 0 \quad (7.1-4)$$

approximiert. Die Voraussetzung $m_1 \neq 0$ schränkt die zur Approximation zugelassenen Wavelets auf wenige Kandidaten ein. Das bekannteste ist das Haar-Wavelet

$$\psi(x) = \begin{cases} 1 & \text{für } 0 \leq x < \frac{1}{2} \\ -1 & \text{für } \frac{1}{2} \leq x < 1 \\ 0 & \text{sonst} \end{cases} \quad (7.1-5)$$

mit $m_1 = -\frac{1}{4}$. Setzt man (7.1-5) in die Wavelettransformierte (7.1-1) ein, so ergibt sich aus (7.1-3)

$$f'(b) = \lim_{a \rightarrow 0} \left\{ -\frac{4}{a^2} \left[\int_b^{b+a/2} f(x) dx - \int_{b+a/2}^{b+a} f(x) dx \right] \right\}$$

und für kleine a mit Hilfe der Trapezregel

$$\begin{aligned} f'(b) &= \lim_{a \rightarrow 0} \left\{ -\frac{4}{a^2} [f(b) - f(b+a)] \frac{a}{4} \right\} \\ &= \lim_{a \rightarrow 0} \frac{f(b+a) - f(b)}{a}. \end{aligned}$$

Allgemein läßt sich (7.1-3) im Fourier-Bereich begründen. Die Wavelettransformation (7.1-1) ist vom Faltungstyp, entspricht daher einem linearen transversalen Filter mit dem Kern ψ als Gewichtsfunktion und seiner FT $\hat{\psi}$ als Filtercharakteristik; vgl. dazu auch BETHGE et al. (1997). Wegen $\hat{\psi}(0) = 0$, $\lim_{\omega \rightarrow \infty} \hat{\psi}(\omega) = 0$ ist $\hat{\psi}$ vom Typ Bandpaß: Für jedes feste $a > 0$ liefert die Wavelettransformation (7.1-1) eine bandgefilterte Version des Signals $f(x)$. Darauf beruht die Zerlegung in Signalanteile unterschiedlicher „Breite“. Im Grenzfall $a \rightarrow 0$ geht der Bandpaß in einen Hochpaß über, der genau dem Differentiationsfilter entspricht (vgl. BEYER und MEIER (2001), Anhang A.1).

Die 2. Ableitung $f''(b)$ im Punkt b wird durch den Grenzwert

$$f''(b) = \lim_{a \rightarrow 0} \frac{2\sqrt{c_\psi}}{a^{5/2}m_2} L_\psi f(a, b) \quad (a > 0) \quad (7.1-6)$$

mit dem nicht verschwindenden 2. Moment

$$m_2 := \int_{-\infty}^{\infty} x^2 \psi(x) dx \neq 0 \quad (7.1-7)$$

approximiert. Setzt man beispielsweise das Wavelet¹

$$\psi(x) = \begin{cases} 4 & \text{für } \frac{1}{2} \leq |x| \leq 1 \\ -4 & \text{für } |x| < \frac{1}{2} \\ 0 & \text{sonst} \end{cases} \quad (7.1-8)$$

mit $m_2 = 2$ in (7.1-1) ein, so folgt aus (7.1-6)

$$f''(b) = \lim_{a \rightarrow 0} \left\{ -\frac{4}{a^3} \left[\int_{b-a}^{b-\frac{a}{2}} f(x) dx - \int_{b-\frac{a}{2}}^b f(x) dx - \int_b^{b+\frac{a}{2}} f(x) dx + \int_{b+\frac{a}{2}}^{b+a} f(x) dx \right] \right\}$$

und daraus wieder mit der Trapezregel

$$\begin{aligned} f''(b) &= \lim_{a \rightarrow 0} \left\{ \frac{1}{a} \left[\frac{f(b+a) - f(b)}{a} - \frac{f(b) - f(b-a)}{a} \right] \right\} \\ &= \lim_{a \rightarrow 0} \frac{f(b-a) - 2f(b) + f(b+a)}{a^2} . \end{aligned}$$

Für jedes Wavelet mit $m_2 \neq 0$ realisiert (7.1-6) ein zweimaliges Differentiationsfilter (vgl. BEYER und MEIER (2001), Anhang A.2). Schließlich gilt für die n -te Ableitung

$$f^{(n)}(b) = \lim_{a \rightarrow 0} \frac{n! \sqrt{c_\psi}}{a^{n+1/2} m_n} L_\psi f(a, b) \quad (a > 0) \quad (7.1-9)$$

mit dem n -ten Moment

$$m_n := \int_{-\infty}^{\infty} x^n \psi(x) dx \neq 0 , \quad (7.1-10)$$

woraus speziell (7.1-3), (7.1-6) für $n = 1, 2$ folgen.

Bei der zweidimensionalen Wavelettransformation gilt entsprechendes für die partiellen Ableitungen. Die Wavelets HT bzw. TH n -ter Ordnung approximieren die n -te partielle Ableitung nach x bzw. y . Das Wavelet HH approximiert die $2n$ -te gemischte partielle Ableitung nach x und y . Das Wavelet TT entspricht der Skalierungsfunktion im 1D-Fall. Es liefert einen „Mittelwert“.

¹ Das Wavelet (7.1-8) entspricht dem Vierfachen der Differenz zweier Haar-Wavelets, wobei das erste auf der x -Achse um Eins nach links verschoben ist.

7.1.2 Approximationseigenschaften bei der diskreten Wavelettransformation

Die für die kontinuierliche Wavelettransformation nachgewiesene Approximationseigenschaft läßt sich auch für die diskrete Wavelettransformation zeigen.

Es sei \underline{v} ein diskretes Wavelet (Waveletvektor nach den Vereinbarungen in 3.2) mit Support $n + 1$:

$$\underline{v} = [v_0 \quad v_1 \quad \cdots \quad v_n] .$$

Weiterhin werde ein (ganzzahliger) Argumentvektor mit der Schrittweite $\Delta x = 1$, eingeschränkt auf den Support des Wavelets betrachtet:

$$\begin{aligned} \underline{x} &= [x_0 \quad x_0 + 1 \quad x_0 + 2 \quad \cdots \quad x_0 + n] \\ &= x_0 \cdot \underline{n}_0 + \underline{n}_1 \end{aligned}$$

mit

$$\begin{aligned} \underline{n}_0 &= [1 \quad 1 \quad 1 \quad \cdots \quad 1] , \\ \underline{n}_1 &= [0 \quad 1 \quad 2 \quad \cdots \quad n] , \\ \underline{n}_2 &= [0 \quad 1 \quad 4 \quad \cdots \quad n^2] . \end{aligned}$$

Dabei ist x_0 das erste Argument, das vom Wavelet \underline{v} mit dem Translationsparameter b überdeckt wird. Für die Berechnung der Momente $m_i = \langle \underline{x}^i, \underline{v} \rangle$ des Wavelets werden die Potenzen \underline{x}^i benötigt, z.B. gilt speziell für \underline{x}^2

$$\begin{aligned} \underline{x}^2 &= [x_0^2 \quad (x_0 + 1)^2 \quad (x_0 + 2)^2 \quad \cdots \quad (x_0 + n)^2] \\ &= x_0^2 \cdot \underline{n}_0 + 2x_0 \cdot \underline{n}_1 + \underline{n}_2 . \end{aligned}$$

Die Momente der Wavelets bis m_2 ergeben sich dann unter Berücksichtigung des endlichen Supports wie folgt (vgl. 3.6):

0. Moment

$$m_0 = \langle \underline{n}_0, \underline{v} \rangle = 0 ,$$

1. Moment

$$m_1 = \langle \underline{x}, \underline{v} \rangle = x_0 \underbrace{\langle \underline{n}_0, \underline{v} \rangle}_{=0} + \langle \underline{n}_1, \underline{v} \rangle = \sum_{i=0}^n i \cdot v_i ,$$

2. Moment

$$m_2 = \langle \underline{x}^2, \underline{v} \rangle = x_0^2 \underbrace{\langle \underline{n}_0, \underline{v} \rangle}_{=0} + 2x_0 \underbrace{\langle \underline{n}_1, \underline{v} \rangle}_{=m_1} + \langle \underline{n}_2, \underline{v} \rangle = 2x_0 m_1 + \sum_{i=0}^n i^2 \cdot v_i .$$

Da im diskreten Fall der Grenzübergang $a \rightarrow 0$ beim Skalenparameter a nicht möglich ist, kann umgekehrt eine Signalfunktion zur Untersuchung der Approximationseigenschaft herangezogen werden, die im Bereich des Supports des Wavelets eine konstante Ableitung besitzt. Es kann damit sogar die Approximationseigenschaft skalenunabhängig gezeigt werden.

Eine Funktion mit konstanter 1. Ableitung d ist z.B.

$$\underline{f}_1 = d \cdot \underline{x} = dx_0 \cdot \underline{n}_0 + d \cdot \underline{n}_1 .$$

Die Wavelettransformation mit einem Wavelet 1. Ordnung ($m_1 \neq 0$) ergibt für den Waveletkoeffizienten in der betrachteten Position b (Translationsparameter)

$$\begin{aligned} w &= \langle \underline{f}_1, \underline{v} \rangle = dx_0 \cdot \underbrace{\langle \underline{n}_0, \underline{v} \rangle}_{=0} + d \cdot \underbrace{\langle \underline{n}_1, \underline{v} \rangle}_{=m_1} \\ &= d \cdot m_1 \simeq d. \end{aligned}$$

Das 1. Moment des Wavelets ist ersichtlich unabhängig vom Translationsparameter b . Der Waveletkoeffizient w ist folglich proportional zur 1. Ableitung d der Signalfunktion.

Eine Funktion mit konstanter 2. Ableitung d ist z.B.

$$\underline{f}_2 = \frac{d}{2} \cdot \underline{x}^2 = \frac{d}{2} x_0^2 \cdot \underline{n}_0 + dx_0 \cdot \underline{n}_1 + \frac{d}{2} \cdot \underline{n}_2.$$

Die Wavelettransformation mit einem Wavelet 2. Ordnung ($m_1 = 0, m_2 \neq 0$) ergibt für den Waveletkoeffizienten in der betrachteten Position b (Translationsparameter)

$$\begin{aligned} w &= \langle \underline{f}_2, \underline{v} \rangle = \frac{d}{2} x_0^2 \cdot \underbrace{\langle \underline{n}_0, \underline{v} \rangle}_{=0} + dx_0 \cdot \underbrace{\langle \underline{n}_1, \underline{v} \rangle}_{=m_1=0} + \frac{d}{2} \underbrace{\langle \underline{n}_2, \underline{v} \rangle}_{=m_2} \\ &= \frac{d}{2} \cdot m_2 \simeq d. \end{aligned}$$

Das 2. Moment eines Wavelets 2. Ordnung ist wegen $m_1 = 0$ ersichtlich unabhängig vom Translationsparameter b . Der Waveletkoeffizient w ist folglich proportional zur 2. Ableitung d der Signalfunktion.

7.1.3 Skalierung der Waveletkoeffizienten

Eindimensional

Die erforderliche Amplitudenkorrektur läßt sich mittels einer Testfunktion mit bekannter konstanter Ableitung bestimmen. Für die Funktion $f_1(x) = x$ gilt für die 1. Ableitung $f_1'(x) = 1$ und für die Funktion $f_2(x) = \frac{1}{2}x^2$ gilt für die 2. Ableitung $f_2''(x) = 1$. Nach den bekannten Approximationseigenschaften gilt für ein Wavelet 1. Ordnung und eine Funktion f mit einem angenommenen Korrekturfaktor α_1 für die Amplitude und eine angenommene Verschiebung x_ϕ der Wavelettransformierten gegenüber der 1. Ableitung (Abb. 7.1-1)

$$f'(x_i - x_\phi) = \alpha_1 \cdot w(f; x_i), \quad (7.1-11)$$

wobei $w(f; x_i)$ der i -te Waveletkoeffizient $w_i = \langle f, \psi_i \rangle$ der stationären Wavelettransformation gemäß (3.4-2) ist², der bei einer Lokalisierung über den Index dem Argument x_i zugeordnet wird. Die gewählte Symbolik soll diese Zuordnung unterstreichen. Für die spezielle Funktion $f_1(x) = x$ mit $f_1'(x) \equiv 1$ gilt dann mit der Tastweite Δx

$$\begin{aligned} 1 &\equiv f_1'(x) = f_1'(x_i - x_\phi) = \alpha_1 \cdot w(f_1; x_i) \\ &= \alpha_1 \cdot \langle f_1, \psi_i \rangle = \alpha_1 \cdot \langle x, \psi_i \rangle = \alpha_1 \cdot m_1 \cdot \Delta x \end{aligned} \quad (7.1-12)$$

und damit wegen $\alpha_1 = \frac{1}{m_1 \Delta x}$ analog zu (7.1-3)

² Die nachfolgenden Herleitungen sind skalenunabhängig. Der Skalenindex k ist deshalb im Interesse der Übersichtlichkeit weggelassen worden.

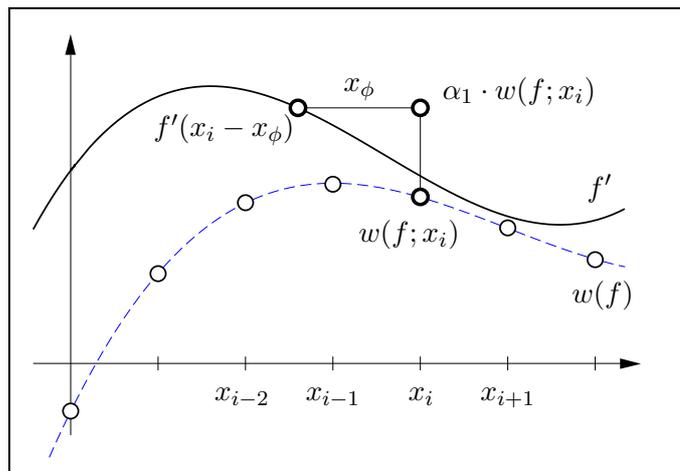


Abbildung 7.1-1: Zusammenhang zwischen den tatsächlichen Ableitungswerten (z.B. der 1. Ableitung $f'(x)$) einer Funktion $f(x)$ und den Waveletkoeffizienten $w(f)$ der diskretisierten Funktion $f(x_i)$. Der auf x_i lokalisierte WK $w(f; x_i)$ approximiert bis auf einen Faktor α_1 eine Ableitung der Funktion an der Stelle $x_i - x_\phi$.

$$\boxed{f'(x_i - x_\phi) = \frac{1}{m_1 \Delta x} \cdot w(f; x_i).} \quad (7.1-13)$$

Für ein Wavelet 2. Ordnung und eine Funktion f gilt in gleicher Weise

$$f''(x_i - x_\phi) = \alpha_2 \cdot w(f; x_i) \quad (7.1-14)$$

und dementsprechend für die spezielle Funktion $f_2(x) = \frac{1}{2}x^2$ mit $f_2''(x) \equiv 1$

$$\begin{aligned} 1 &\equiv f_2''(x) = f_2''(x_i - x_\phi) = \alpha_2 \cdot w(f_2; x_i) \\ &= \alpha_2 \cdot \langle f_2, \psi_i \rangle = \alpha_2 \cdot \langle \frac{1}{2}x^2, \psi_i \rangle \\ &= \frac{1}{2} \alpha_2 \cdot m_2 \cdot (\Delta x)^2. \end{aligned} \quad (7.1-15)$$

Schließlich folgt wegen $\alpha_2 = \frac{2}{m_2(\Delta x)^2}$ analog zu (7.1-6)

$$\boxed{f''(x_i - x_\phi) = \frac{2}{m_2(\Delta x)^2} \cdot w(f; x_i).} \quad (7.1-16)$$

Weiterhin läßt sich analog zu (7.1-9) für ein Wavelet n -ter Ordnung

$$\boxed{f^{(n)}(x_i - x_\phi) = \frac{n!}{m_n(\Delta x)^n} \cdot w(f; x_i).} \quad (7.1-17)$$

zeigen.

Zweidimensional

Ein Wavelet HT erster Ordnung approximiert die erste partielle Ableitung nach x . Zur Ermittlung der erforderlichen Amplitudenkorrektur eignet sich eine Testfunktion $z = f(x, y)$ mit konstanter Ableitung, speziell mit $z_x \equiv 1$:

$$f(x, y) = x + g_0(y) . \quad (7.1-18)$$

Über den Ansatz $1 \equiv z_x = f_x(x, y)$ mit dem angenommenen Skalierungsfaktor α ergibt sich

$$\begin{aligned} 1 \equiv f_x(x, y) &= \alpha \cdot w_{ht}(f) = \alpha \langle \langle f, h \rangle, t \rangle \\ &= \alpha \langle \langle x + g_0(y), h \rangle, t \rangle \\ &= \alpha \left\langle \underbrace{\langle x, h \rangle}_{=m_1^h \cdot \Delta x} + g_0(y) \underbrace{\langle 1, h \rangle}_{=m_0^h=0}, t \right\rangle \\ &= \alpha m_1^h \Delta x \underbrace{\langle 1, t \rangle}_{=m_0^t} \end{aligned}$$

und damit für den Skalierungsfaktor α der Amplitudenkorrektur $f_x(x, y) = \alpha \cdot w_{ht}(f)$

$$\boxed{\alpha = \frac{1}{m_1^h \Delta x \cdot m_0^t}} . \quad (7.1-19)$$

Ein Wavelet HT zweiter Ordnung approximiert die zweite partielle Ableitung nach x . Zur Ermittlung der erforderlichen Amplitudenkorrektur eignet sich folglich analog eine Testfunktion mit $z_{xx} = f_{xx}(x, y) \equiv 1$:

$$f(x, y) = \frac{1}{2}x^2 + x \cdot g_1(y) + g_0(y) . \quad (7.1-20)$$

Mit dem angenommenen Skalierungsfaktor α hat man

$$\begin{aligned} 1 \equiv f_{xx}(x, y) &= \alpha \cdot w_{ht}(f) = \alpha \langle \langle f, h \rangle, t \rangle \\ &= \alpha \left\langle \left\langle \frac{1}{2}x^2 + x \cdot g_1(y) + g_0(y), h \right\rangle, t \right\rangle \\ &= \alpha \left\langle \frac{1}{2} \underbrace{\langle x^2, h \rangle}_{=m_2^h \cdot (\Delta x)^2} + g_1(y) \underbrace{\langle x, h \rangle}_{=m_1^h \cdot \Delta x=0} + g_0(y) \underbrace{\langle 1, h \rangle}_{=m_0^h=0}, t \right\rangle \\ &= \frac{\alpha}{2} m_2^h (\Delta x)^2 \underbrace{\langle 1, t \rangle}_{=m_0^t} \end{aligned}$$

und der Skalierungsfaktor α ergibt sich zu

$$\boxed{\alpha = \frac{2}{m_2^h (\Delta x)^2 \cdot m_0^t}} . \quad (7.1-21)$$

Die Skalierungsfaktoren berechnen sich also analog zum 1D-Fall, wobei lediglich zusätzlich das nullte Moment m_0^t des Tiefpasses in die Formel eingeht. Für die Ableitungen nach y erhält man die gleichen Ergebnisse, wobei der Term Δx durch Δy zu ersetzen ist.

7.2 Lokalisierungseigenschaften

Bei der diskreten Approximation ist der Grenzübergang $a \rightarrow 0$ nicht möglich. Die Wavelets besitzen auf der feinsten Skale noch einen Support > 1 . Die durch einen Waveletkoeffizienten repräsentierte Approximation der Ableitung ist damit ein „Mittelwert“ aus mehreren benachbarten diskreten Signalwerten. Es stellt sich daher die Frage, ob bei der diskreten Wavelettransformation ebenfalls von *der* Lokalisierungseigenschaft gesprochen werden kann.

7.2.1 Eindimensionale Wavelettransformation

Die Verschiebung der Wavelettransformierten gegenüber der n-ten Ableitung des Signals läßt sich mit Hilfe einer Testfunktion ermitteln, deren n-te Ableitung linear ist, insbesondere $f^{(n)}(x) = x$. Die Ortsverschiebung x_ϕ ist dann gleich der Funktionswertdifferenz in jedem Argument.

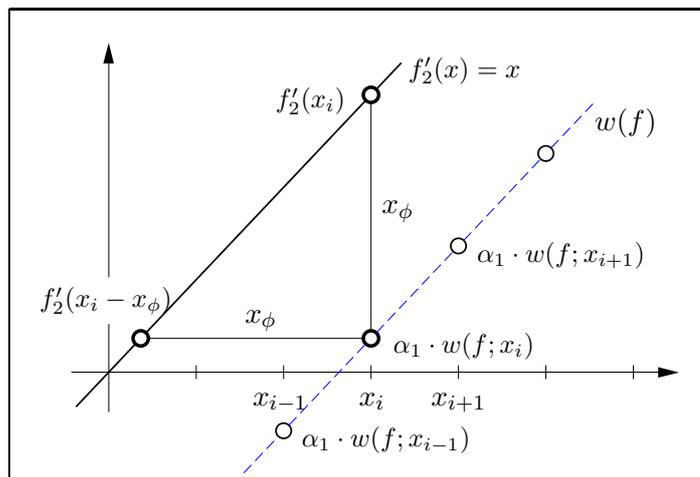


Abbildung 7.2-1: Zusammenhang zwischen den tatsächlichen Ableitungswerten und den Waveletkoeffizienten nach Amplitudenkorrektur bei einer Funktion mit $f'_2(x) = x$. Der korrigierte WK $\alpha_1 \cdot w(x_i)$ approximiert die Ableitung der Funktion f_2 an der Stelle $x - x_\phi$. Die Ortsverschiebung x_ϕ ist gleich der Differenz zwischen $f'_2(x_i) = x_i$ und $\alpha_1 \cdot w(x_i)$.

Für den Zusammenhang zwischen der 1. Ableitung und der Wavelettransformierten gilt

$$f'(x_i - x_\phi) = \alpha_1 \cdot w(f; x_i). \quad (7.2-1)$$

Mit der Testfunktion $f_2(x) = \frac{1}{2}x^2$ mit $f'_2(x) = x$ gilt mit der Tastweite Δx

$$\begin{aligned} x_i - x_\phi &= f'_2(x_i - x_\phi) = \alpha_1 \cdot w(f_2; x_i) \\ &= \alpha_1 \cdot \langle f_2, \psi_i \rangle = \frac{\alpha_1}{2} \cdot \langle x^2, \psi_i \rangle \\ &= \frac{\alpha_1}{2} \cdot m_2(i) \cdot (\Delta x)^2. \end{aligned} \quad (7.2-2)$$

Wenn $m_1 \neq 0$, dann ist das zweite Moment nicht mehr unabhängig von der Verschiebung des Wavelets. Es gilt jedoch $m_2(i) = 2(i+d) \cdot m_1 + \tilde{m}_2$. Daraus folgt mit $\alpha_1 = \frac{1}{m_1 \Delta x}$

$$\begin{aligned} x_\phi &= i \cdot \Delta x - \frac{2(i+d) \cdot m_1 + \tilde{m}_2}{2 \cdot m_1 \cdot \Delta x} \cdot (\Delta x)^2 \\ &= \left(i - \frac{2(i+d) \cdot m_1}{2 \cdot m_1} - \frac{\tilde{m}_2}{2 \cdot m_1} \right) \Delta x \\ &= \left(-d - \frac{\tilde{m}_2}{2 \cdot m_1} \right) \Delta x \quad (\text{unabhängig von } i). \end{aligned} \quad (7.2-3)$$

Damit ergibt sich mit $d = -n_c \cdot 2^{k-1}$ für die Verschiebung

$$\boxed{x_\phi = \left(n_c \cdot 2^{k-1} - \frac{\tilde{m}_2}{2 m_1} \right) \Delta x .} \quad (7.2-4)$$

Für die 2. Ableitung gilt mit der Testfunktion $f_3(x) = \frac{1}{6}x^3$ mit $f_3''(x) = x$

$$\begin{aligned} x_i - x_\phi &= f_3''(x_i - x_\phi) = \alpha_2 \cdot w(f_3; x_i) \\ &= \alpha_2 \cdot \langle f_3, \psi_i \rangle = \frac{\alpha_2}{6} \cdot \langle x^3, \psi_i \rangle \\ &= \frac{\alpha_2}{6} \cdot m_3(i) \cdot (\Delta x)^3. \end{aligned} \quad (7.2-5)$$

Damit ergibt sich für die Verschiebung mit $\alpha_2 = \frac{2}{m_2(\Delta x)^2}$, $m_3(i) = 3(i+d) \cdot m_2 + \tilde{m}_3$ und $d = -n_c \cdot 2^{k-1}$

$$\boxed{x_\phi = \left(n_c \cdot 2^{k-1} - \frac{\tilde{m}_3}{3 m_2} \right) \Delta x .} \quad (7.2-6)$$

7.2.2 Zweidimensionale Wavelettransformation

Bei der zweidimensionalen Wavelettransformation mit den in 3.3.2 konstruierten Tensorprodukt-wavelets werden die Wavelets als zweidimensionale Matrizen gemäß der zyklischen Translation von Hochpaß \underline{h} und Tiefpaß \underline{t} (vgl. 3.2) bei der Berechnung der Skalarprodukte in zwei Richtungen zyklisch über die Signalmatrix Z geschoben³.

Die Zuordnung der Waveletkoeffizienten $w_{i,j}$ zu den Indizes der Signalmatrix bei der zweidimensionalen Wavelettransformation ist für ein Wavelet mit vier Filterkoeffizienten in Abb. 7.2-2 dargestellt.

Bei einer Lokalisierung der Waveletkoeffizienten nach den Indizes wird die approximierte Ableitung z_{xx} (das zweidimensionale Wavelet HT approximiert die zweite Ableitung der Funktion $z = f(x, y)$ nach x) als z.B. $w_{i,j}$ dem Argument (x_i, y_j) zugeordnet⁴.

³ Um die Skalarprodukte auch bilden zu können, wenn die Waveletmatrix über die Signalmatrix hinausragt, wird bei den üblichen Verfahren die Signalmatrix periodisch fortgesetzt.

⁴ In der Matrixschreibweise bezeichnet der erste Index die Zeile und der zweite Index die Spalte. In der Funktionsschreibweise bezeichnet der erste Index das Argument x und der zweite Index das Argument y . Deshalb zeigt in der Matrizendarstellung die x -Achse nach unten und die y -Achse nach rechts.

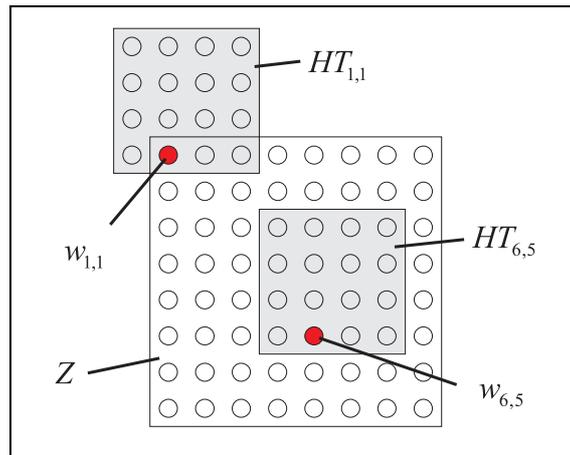


Abbildung 7.2-2: Schematische Darstellung der zweidimensionalen stationären WT für ein Wavelet mit vier Koeffizienten (z.B. Daubechies4-Wavelet) in der 1. Skale. Die Wavelets $HT_{1,1}$ und $HT_{6,5}$ und ihre nach Index lokalisierten WK $w_{1,1}$ und $w_{6,5}$ sind hervorgehoben.

Es ist zu sehen, daß diese Zuordnung im Ortsbereich offensichtlich nicht repräsentativ für die approximierte Ableitung ist. Es ist eher zu erwarten, daß der Waveletkoeffizient die Ableitung der Funktion im mittleren Bereich des Supports des Wavelets repräsentiert. D.h., die Lokalisierung ist gegenüber der Indexzuordnung zu verschieben. Die Bestimmung dieser erforderlichen Verschiebung wird nachfolgend untersucht.

Die durch die speziellen Filterkoeffizienten bedingte Translation in x - und y -Richtung wird wiederum exemplarisch für die partiellen Ableitungen nach x untersucht, da die Herleitungen für die Ableitungen nach y analog verlaufen. Dabei kann die Translation der approximierten Ableitungen in x - bzw. y - Richtung zweckmäßig mit einer Testfunktion bestimmt werden, deren Ableitung linear von x bzw. y abhängt. Wählt man speziell die Funktion $z = f(x, y)$ mit $f^{(n)}(x, y) = x$, dann ist die Ortsverschiebung x_ϕ gleich der Funktionswertdifferenz in jedem Argument (vgl. BEYER und MEIER (2001)).

Wavelet 1. Ordnung

Translation in x -Richtung Eine geeignete Testfunktion für die erste partielle Ableitung ist eine Funktion $z = f(x, y)$ mit $z_x = x$:

$$f(x, y) = \frac{1}{2}x^2 + g_0(y) . \quad (7.2-7)$$

Mit der angenommenen Verschiebung x_ϕ und der Skalierung α kann man $x_i - x_\phi = f_x(x_i - x_\phi, y) = \alpha \cdot w_{ht}(f)$ ansetzen und man erhält

$$\begin{aligned} x_i - x_\phi = f_x(x_i - x_\phi, y) &= \alpha \cdot w_{ht}(f) = \alpha \langle \langle f, h \rangle, t \rangle \\ &= \alpha \left\langle \left\langle \frac{1}{2}x^2 + g_0(y), h \right\rangle, t \right\rangle \\ &= \alpha \left\langle \frac{1}{2} \underbrace{\langle x^2, h \rangle}_{=m_2^h(i) \cdot (\Delta x)^2} + g_0(y) \underbrace{\langle 1, h \rangle, t}_{=0} \right\rangle \\ &= \frac{\alpha}{2} m_2^h(i) (\Delta x)^2 \underbrace{\langle 1, t \rangle}_{=m_0^t} . \end{aligned}$$

Daraus folgt mit $\alpha = \frac{1}{m_1^h \Delta x \cdot m_0^t}$, $m_2^h(i) = 2(i + d^h) \cdot m_1^h + \tilde{m}_2^h$

$$\begin{aligned}
 x_\phi &= x_i - \frac{\alpha}{2} m_2^h(i) (\Delta x)^2 \cdot m_0^t \\
 &= i \cdot \Delta x - \frac{2(i + d^h) \cdot m_1^h + \tilde{m}_2^h}{2 \cdot m_1^h \cdot \Delta x \cdot m_0^t} \cdot (\Delta x)^2 \cdot m_0^t \\
 &= \left(i - (i + d^h) - \frac{\tilde{m}_2^h}{2 \cdot m_1^h} \right) \Delta x \\
 &= \left(-d^h - \frac{\tilde{m}_2^h}{2 \cdot m_1^h} \right) \Delta x \quad (\text{unabhängig von } i)
 \end{aligned}$$

und insgesamt mit $d^h = -n_c 2^{k-1}$

$$\boxed{x_\phi = \left(n_c \cdot 2^{k-1} - \frac{\tilde{m}_2^h}{2 \cdot m_1^h} \right) \Delta x .} \quad (7.2-8)$$

Das Ergebnis ist identisch mit der Formel im 1D-Fall.

Translation in y -Richtung Die Translation in y -Richtung kann mit einer Testfunktion bestimmt werden, deren Ableitung linear von y abhängt.

Eine geeignete Testfunktion $z = f(x, y)$ mit $z_x = y$ ist

$$f(x, y) = x \cdot y + g_0(y) . \quad (7.2-9)$$

Mit einem zu oben analogen Ansatz erhält man

$$\begin{aligned}
 y_i - y_\phi &= f_x(x, y_i - y_\phi) = \alpha \cdot w_{ht}(f) = \alpha \langle \langle f, h \rangle, t \rangle \\
 &= \alpha \langle \langle x \cdot y + g_0(y), h \rangle, t \rangle \\
 &= \alpha \left\langle y \underbrace{\langle x, h \rangle}_{=m_1^h \cdot \Delta x} + g_0(y) \underbrace{\langle 1, h \rangle, t}_{=0} \right\rangle \\
 &= \alpha m_1^h \Delta x \underbrace{\langle y, t \rangle}_{=m_1^t(i) \Delta y} .
 \end{aligned}$$

Daraus folgt mit $\alpha = \frac{1}{m_1^h \Delta x \cdot m_0^t}$, $m_1^t(i) = (i + d^t) \cdot m_0^t + \tilde{m}_1^t$

$$\begin{aligned}
 y_\phi &= y_i - \alpha m_1^h \Delta x \cdot m_1^t(i) \Delta y \\
 &= i \cdot \Delta y - \frac{m_1^h \Delta x ((i + d^t) m_0^t + \tilde{m}_1^t) \Delta y}{m_1^h \Delta x \cdot m_0^t} \\
 &= \left(i - (i + d^t) - \frac{\tilde{m}_1^t}{m_0^t} \right) \Delta y \\
 &= \left(-d^t - \frac{\tilde{m}_1^t}{m_0^t} \right) \Delta y
 \end{aligned}$$

und insgesamt mit $d^t = -2^k$

$$y_\phi = \left(2^k - \frac{\tilde{m}_1^t}{m_0^t} \right) \Delta y. \quad (7.2-10)$$

Die Verschiebung senkrecht zur Ableitungsrichtung läßt sich nicht mit dem 1D-Fall vergleichen, da diese im 1D-Fall gegenstandslos ist. Die Formel unterscheidet sich jedoch deutlich von der Verschiebung in Ableitungsrichtung, was auf die unterschiedliche Form von Wavelet und Skalierungsfunktion zurückzuführen ist.

Wavelet 2. Ordnung

Translation in x -Richtung Mit der Testfunktion $z = f(x, y)$ mit $z_{xx} = x$

$$f(x, y) = \frac{1}{6}x^3 + x \cdot g_1(y) + g_0(y) \quad (7.2-11)$$

erhält man für die Translation in x -Richtung

$$\begin{aligned} x_i - x_\phi = f_{xx}(x_i - x_\phi, y) &= \alpha \cdot w_{ht}(f) = \alpha \langle \langle f, h \rangle, t \rangle \\ &= \alpha \left\langle \left\langle \frac{1}{6}x^3 + x \cdot g_1(y) + g_0(y), h \right\rangle, t \right\rangle \\ &= \alpha \left\langle \frac{1}{6} \underbrace{\langle x^3, h \rangle}_{=m_3^h(i) \cdot (\Delta x)^3} + g_1(y) \underbrace{\langle x, h \rangle}_{=0} + g_0(y) \underbrace{\langle 1, h \rangle}_{=0}, t \right\rangle \\ &= \frac{\alpha}{6} m_3^h(i) (\Delta x)^3 \underbrace{\langle 1, t \rangle}_{=m_0^t}. \end{aligned}$$

Daraus folgt mit $\alpha = \frac{2}{m_2^h(\Delta x)^2 \cdot m_0^t}$, $m_3^h(i) = 3(i + d^h) \cdot m_2^h + \tilde{m}_3^h$

$$\begin{aligned} x_\phi &= x_i - \frac{\alpha}{6} m_3^h(i) (\Delta x)^3 \cdot m_0^t \\ &= i \cdot \Delta x - \frac{3(i + d^h) \cdot m_2^h + \tilde{m}_3^h}{3 \cdot m_2^h \cdot (\Delta x)^2 \cdot m_0^t} \cdot (\Delta x)^3 \cdot m_0^t \\ &= \left(i - (i + d^h) - \frac{\tilde{m}_3^h}{3 \cdot m_2^h} \right) \Delta x \\ &= \left(-d^h - \frac{\tilde{m}_3^h}{3 \cdot m_2^h} \right) \Delta x \quad (\text{unabhängig von } i) \end{aligned}$$

und insgesamt mit $d^h = -n_c \cdot 2^{k-1}$

$$x_\phi = \left(n_c \cdot 2^{k-1} - \frac{\tilde{m}_3^h}{3 \cdot m_2^h} \right) \Delta x. \quad (7.2-12)$$

Translation in y -Richtung Die Translation in y -Richtung kann wiederum mit einer Testfunktion bestimmt werden, deren Ableitung linear von y abhängt.

Eine geeignete Testfunktion ist die Funktion $z = f(x, y)$ mit $z_{xx} = y$

$$f(x, y) = \frac{y}{2}x^2 + x \cdot g_1(y) + g_0(y). \quad (7.2-13)$$

Mit dem analogen Ansatz mit der angenommenen Verschiebung y_ϕ ergibt sich

$$\begin{aligned}
 y_i - y_\phi = f_{xx}(x, y_i - y_\phi) &= \alpha \cdot w_{ht}(f) = \alpha \langle \langle f, h \rangle, t \rangle \\
 &= \alpha \left\langle \left\langle \frac{y}{2} x^2 + x \cdot g_1(y) + g_0(y), h \right\rangle, t \right\rangle \\
 &= \alpha \left\langle \frac{y}{2} \underbrace{\langle x^2, h \rangle}_{=m_2^h(\Delta x)^2} + g_1(y) \underbrace{\langle x, h \rangle}_{=0} + g_0(y) \underbrace{\langle 1, h \rangle}_{=0}, t \right\rangle \\
 &= \frac{\alpha}{2} m_2^h(\Delta x)^2 \underbrace{\langle y, t \rangle}_{=m_1^t(i)\Delta y} .
 \end{aligned}$$

Daraus folgt mit $\alpha = \frac{2}{m_2^h(\Delta x)^2 \cdot m_0^t}$, $m_1^t(i) = (i + d^t) \cdot m_0^t + \tilde{m}_1^t$

$$\begin{aligned}
 y_\phi &= y_i - \frac{\alpha}{2} m_2^h(\Delta x)^2 \cdot m_1^t(i) \Delta y \\
 &= i \cdot \Delta y - \frac{m_2^h(\Delta x)^2 ((i + d^t)m_0^t + \tilde{m}_1^t) \Delta y}{m_2^h(\Delta x)^2 \cdot m_0^t} \\
 &= \left(i - (i + d^t) - \frac{\tilde{m}_1^t}{m_0^t} \right) \Delta y \\
 &= \left(-d^t - \frac{\tilde{m}_1^t}{m_0^t} \right) \Delta y
 \end{aligned}$$

und insgesamt mit $d^t = -2^k$

$$\boxed{y_\phi = \left(2^k - \frac{\tilde{m}_1^t}{m_0^t} \right) \Delta y .} \quad (7.2-14)$$

Im Vergleich der Ergebnisse für Wavelets erster und zweiter Ordnung sind wie im 1D-Fall einfache Verallgemeinerungen für Wavelets n -ter Ordnung zu erwarten, die für die gestellte praktische Aufgabenstellung jedoch nicht von Belang sind. Für die gemischten Ableitungen lassen sich die Untersuchungen analog durchführen.

7.3 Höhen- und Gefällelinien

In den Abschnitten 7.1 und 7.2 wurde gezeigt, wie aus der Wavelettransformierten der Höhenmatrix eines Geländemodells die partiellen Ableitungen gewonnen werden können, und im Abschnitt 4.2 wurden die Beziehungen angegeben, mit denen aus den partiellen Ableitungen einer Flächenfunktion Neigungen und Krümmungen eines Geländemodells bestimmt werden können. Es ist daher naheliegend, der folgenden Frage nachzugehen: Können aus dieser Wavelettransformierten, die unter bestimmten Voraussetzungen als Vektorfeld der Gradienten aufgefaßt werden kann, auch Höhen- und Gefällelinien abgeleitet werden?

Im Rahmen einer Diplomarbeit (RICHTER (2001)) wurde dieses Problem erfolgreich bearbeitet. In dieser Diplomarbeit werden Algorithmen entwickelt, die die Höhen- und Gefällelinien eines Geländemodells mit hoher Qualität generieren. An dieser Stelle sollen diese Ergebnisse kurz vorgestellt werden.

Aus 4.2-2 und 4.2-3 erhält man einen Tangenteneinheitsvektor an die Höhenlinie zu

$$\underline{t}_h = [\underline{v}_g, \underline{n}_F] = \frac{1}{\sqrt{z_x^2 + z_y^2}} \begin{bmatrix} -z_y \\ z_x \\ 0 \end{bmatrix}. \quad (7.3-1)$$

Ausgehend von einem Startpunkt x_n , dessen Lage und Höhe bekannt sein muß, kann die Höhenlinie mit Hilfe der Richtungsvektoren \underline{t}_h sukzessive jeweils in Tangentenrichtung verfolgt werden. Dabei sind notwendige Zwischenwerte, die zu Punkten gehören, die keine Gitterknoten sind, zu interpolieren. Die Punktfolge, die iterativ nach

$$\underline{x}_{n+1} = \underline{x}_n + s \cdot \underline{t}_h \quad (7.3-2)$$

berechnet werden kann, wird erwartungsgemäß, insbesondere bei gleichbleibender Orientierung der Krümmung, von der tatsächlichen Höhenlinie abdriften. Dieser Effekt ist in Abb. 7.3-1, in der eine Höhenlinie eines halbkugelförmigen Geländes wiedergegeben ist, deutlich zu sehen.

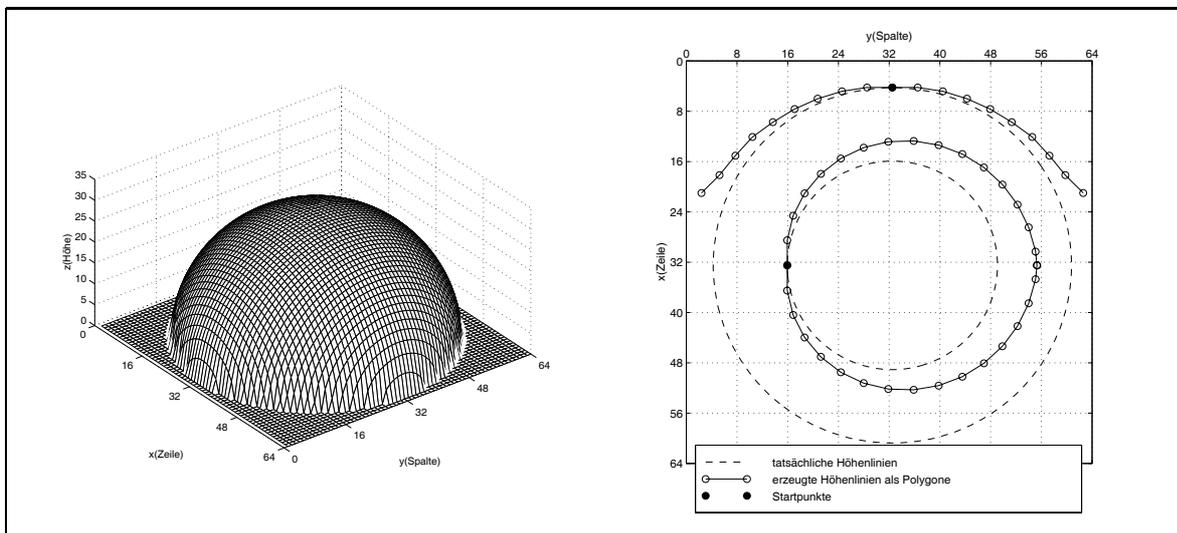


Abbildung 7.3-1: Halbkugel als Testdatensatz.

Um dieses Abdriften zu kompensieren, wurde der Algorithmus geeignet modifiziert. Dazu wird nach einem einfachen Prinzip indirekt die Krümmung der Höhenlinie berücksichtigt. Der Grundgedanke geht von den folgenden Voraussetzungen aus:

- Die Krümmung der Höhenlinien zwischen den Polygonpunkten ist konstant.
- Die Niveaulinien sind in der Umgebung der Polygonpunkte parallel.
- Die Tangentenrichtungen in P und Q' unterscheiden sich um weniger als $\frac{\pi}{2}$.

Und zwar kann der berechnete Höhenlinienpunkt Q' zur Korrektur radial in Richtung der Höhenlinie verschoben werden. Die Krümmung bzw. der Krümmungsradius ergeben sich mit den oben angegebenen Voraussetzungen indirekt aus der Tangentenrichtung an die Höhenlinie im Punkt Q' (Abb. 7.3-2).

Der Radius der kreisförmig angenommenen Höhenlinie ergibt sich zu

$$R = \frac{SW}{\tan \gamma} \quad \text{mit} \quad \tan \gamma = \frac{\| [\underline{t}_h^P, \underline{t}_h^{Q'}] \|}{\left(\underline{t}_h^P, \underline{t}_h^{Q'} \right)}. \quad (7.3-3)$$

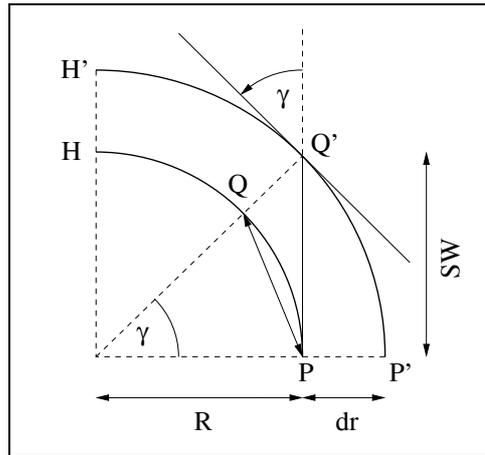


Abbildung 7.3-2: Geometrische Beziehungen, die zur Verbesserung der Iteration unter den getroffenen Voraussetzungen ausgenutzt werden.

Daraus folgt wegen

$$\overline{PQ} = 2 \cdot R \cdot \sin \frac{\gamma}{2} \quad \text{und} \quad \underline{r}_{PQ} = \frac{t_h^P}{\|L_h^P\|} + \frac{t_h^{Q'}}{\|L_h^{Q'}\|} \quad (7.3-4)$$

für den gegenüber (7.3-2) verbesserten Iterationsschritt

$$\underline{x}_Q = \underline{x}_P + \overline{PQ} \cdot \frac{\underline{r}_{PQ}}{\|\underline{r}_{PQ}\|}. \quad (7.3-5)$$

Eine weitere Verbesserung ist zu erwarten, wenn die Voraussetzung der Parallelität der Höhenlinien aufgegeben wird. Da sich diese Erwartungen in den Tests nicht signifikant erfüllen, soll diese Modifikation hier nicht weiter ausgeführt werden.

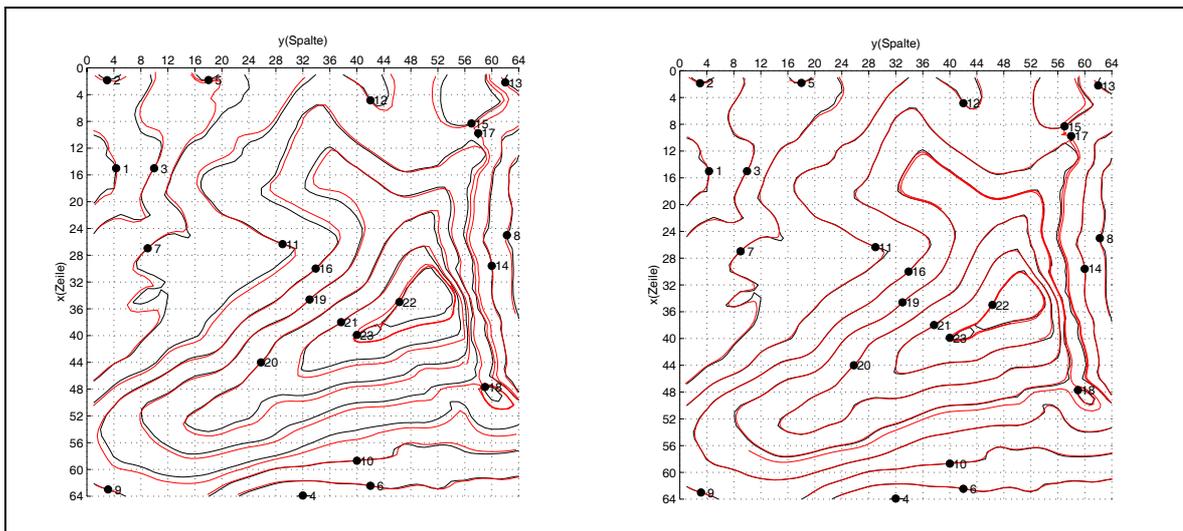


Abbildung 7.3-3: Höhenlinien ohne und mit Modifikation in einem realen Höhenmodell. Die berechneten Höhenlinien sind rot und die Referenzhöhenlinien der Matlab-Funktion contour schwarz dargestellt. Die numerierten Punkte (•) sind die Startpunkte für die jeweilige Höhenlinie.

Zum Test der Algorithmen an realen Geländemodellen wurde ein Referenzverfahren herangezogen: Die contour-Funktion von Matlab. Da keine „wahren“ Höhenlinien existieren, die zum

Qualitätsvergleich herangezogen werden können, ist der Vergleich mit einem solchen Verfahren ausreichend aussagekräftig. Es konnte so gezeigt werden, daß die von RICHTER (2001) entwickelten Algorithmen einem Vergleich mit den klassischen Algorithmen eines renommierten Programmsystems standhalten. Der Test an einem realen Höhenmodell zeigt die gute Qualität des entwickelten Algorithmus. In Abb. 7.3-3, links, sind die Höhenlinien, die mit dem unmodifizierten Algorithmus berechnet wurden, im Vergleich zu den Höhenlinien, die die Matlab-Funktion `contour` erzeugt, zu sehen. Im rechten Teil der Abbildung ist das Ergebnis des modifizierten Algorithmus gegenübergestellt.

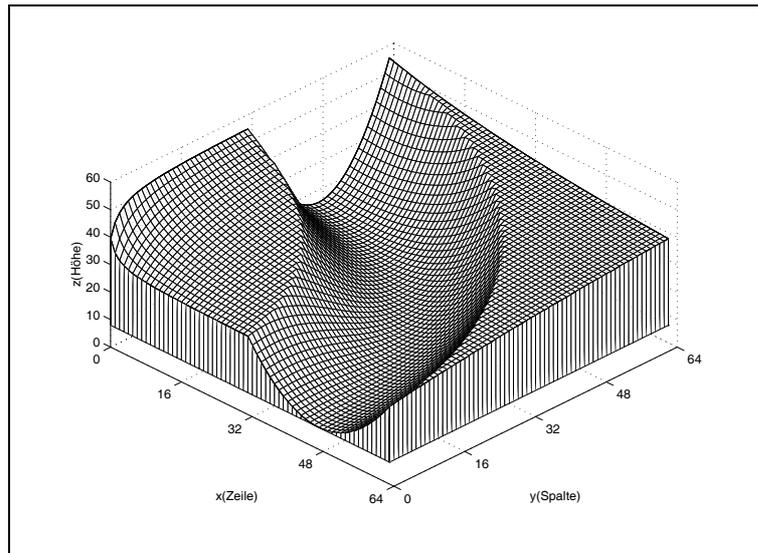


Abbildung 7.3-4: Talmulde als Testdatensatz.

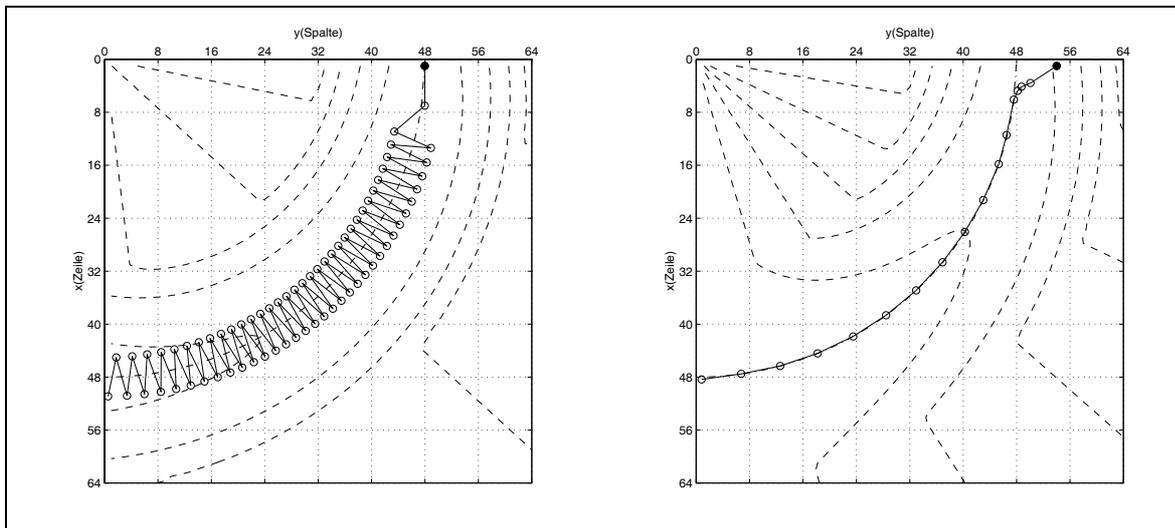


Abbildung 7.3-5: Gefälleslinien ohne und mit Modifikation.

Die Berechnung der Gefälleslinien erfolgt nach einem analogen Verfolgungsalgorithmus. Im Prinzip ist im o. g. Algorithmus nur die Tangentenrichtung an die Höhenlinie durch die Richtung der Gefälleslinie zu ersetzen:

$$x_{n+1} = x_n + SW \cdot \underline{u}_g. \tag{7.3-6}$$

Zum Test des Algorithmus wurde der Testdatensatz in Abb. 7.3-4 generiert. Das Abdriften der berechneten Punkte von der tatsächlichen Gefällelinie ist wesentlich geringer als bei den Höhenlinien. Jedoch ist hier ein Pendeln festzustellen (Abb. 7.3-5, links). Durch geeignete Modifikationen des Algorithmus läßt sich dieses Pendeln weitgehend einschränken. Das Ergebnis ist bei dem gleichen Testdatensatz deutlich sichtbar (Abb. 7.3-5, rechts).

8 Ausblick

8.1 Wavelettransformation raumbezogener Geoinformationen

Die bisherige Beschreibung hybrider Geländemodelle in einer einheitlichen Matrixstruktur stellt zunächst ein Zwei-Schicht-Modell aus Flächendaten und Kantendaten dar. Bei der Beschreibung der Kantenextrapolation (4.3.2) wurde bereits das Problem zu dicht liegender Kanten angesprochen. Dabei wurde auch die Hinzunahme einer weiteren Schicht als möglicher Ausweg angedeutet. Auch für markante Einzelpunkte, z.B. relative Extrema und Sattelpunkte des Reliefs, für die eine mögliche gitterbasierte Beschreibung in BEYER (2000b) vorgestellt wird, kann eine eigene Schicht vorgesehen werden. Ist man einmal dabei, das Zwei-Schicht-Modell aufzustocken, so kann man auch an weitere Schichten für raumbezogene (nicht topographische) Geoinformationen denken. Damit wäre der Weg zu einem allgemeinen Mehrschichtmodell auf der Basis des ursprünglichen hybriden Geländemodells frei. Raumbezogene Geoinformationen können schicht-

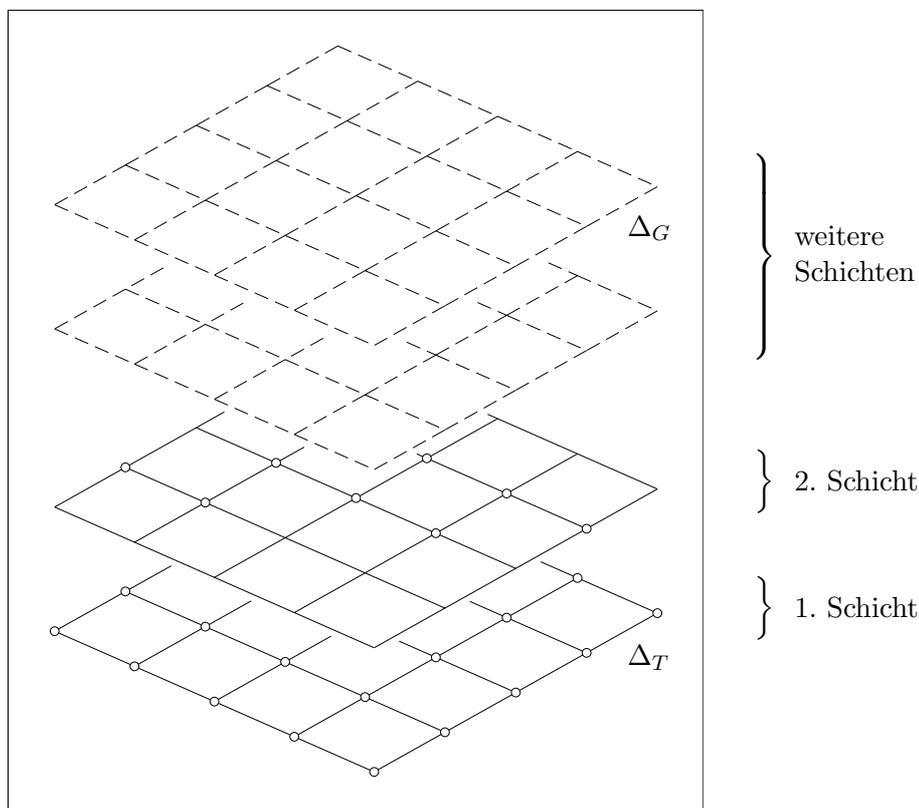


Abbildung 8.1-1: Projektiertes Mehrschichtenmodell für raumbezogene Geoinformationen. Auf der 1. Schicht liegen die diskreten Höhenwerte des DHM, auf der 2. Schicht die gitterbasierten Kanteninformationen. In den darüber liegenden Schichten (gestrichelt) finden die nicht-topographischen Informationen Platz.

weise so über das topographische Grundraaster gelegt werden, daß der Raumbezug auch nach Transformation in den Waveletbereich erhalten bleibt (Abb. 8.1-1). Dazu müssen die Informationen, gleich welcher Art und Struktur, auf deckungsgleiche Raster bezogen werden, wobei die Gleichheit der Rasterweiten $\Delta_G = \Delta_T$ nicht notwendig gelten muß. Es muß lediglich Δ_T ein ganzzahliges Vielfaches von Δ_G (ausgedünnte Information) oder Δ_G ein ganzzahliges Vielfaches von Δ_T (verdichtete Information) sein, damit jeweils eine Teilmenge von Rasterpunkten über-

einanderliegt und den Lagebezug sichert. Ein wesentliches Problem bei der Zusammenführung hybrider Geoinformationen besteht in ihrer Herkunft (Datenquellen): Man hat mit unterschiedlichen Datenformaten, mit Daten unterschiedlicher Auflösung (Raster- oder Vektordaten, Gebiets- und/oder Zeitmittelwerte usw.) und unterschiedlicher Erfassungsgenauigkeit zu tun.

Um geeignete Lösungswege zu konzipieren und zu testen, sind real verfügbare Daten erforderlich. In einer Studienarbeit (STEINIGER (2001)) wurde eine umfangreiche Recherche durchgeführt. Es handelt sich dabei um Vorarbeiten für die oben genannte Aufgabenstellung, die jedoch nicht Gegenstand der vorliegenden Arbeit ist.

8.2 Weitere Fragestellungen

Der große Erfolg der Wavelettransformation bei der Datenkompression in der digitalen Bildverarbeitung gab Anlaß, die Anwendbarkeit der Wavelettransformation auf digitale Geländemodelle zu untersuchen. Die im weiteren Sinne gleichartige Datenstruktur ließ auf vergleichbar gute Resultate hoffen. Die Ergebnisse der vorliegenden Arbeit zeigen jedoch, daß die Hauptanwendung der Wavelettransformation bei digitalen Geländemodellen sicherlich nicht die reine Datenkompression sein wird. Die eher geringen Kompressionsraten ergeben sich aus der Tatsache, daß digitale Geländemodelle bereits gegenüber den Originaldaten (Laser-Scanner-Daten, digitale Luftbilder), dem Anwendungszweck angepaßt, stark ausgedünnt sind.

Dagegen erscheint es sinnvoll, die Breite der Eigenschaften der Wavelettransformation zu nutzen, um in einem Zuge mehrere Probleme zu lösen, auch wenn das jeweilige Einzelproblem mit anderen Werkzeugen der Mathematik oder Signalverarbeitung evtl. besser zu lösen wäre. Die Kombination aus Kompressions-, Approximations- und Lokalisierungseigenschaft kann der Schlüssel für eine effektive Anwendung der Wavelettransformation in der Geoinformatik sein. Die vorhandenen, sehr effektiven Algorithmen der diskreten Wavelettransformation bieten dafür eine solide Grundlage.

Aufbauend auf den in dieser Arbeit gewonnenen vertieften Aussagen zur Approximation und Lokalisierung und der daraus abgeleiteten Analyse der Geländemodelle bezüglich ihrer differentialgeometrischen Eigenschaften ergeben sich eine Reihe weiterer interessanter Probleme. Es erscheint z.B. sinnvoll, die Wavelettransformation anstatt für eine globale Datenkompression besser für eine Lokalisierung und Klassifizierung kompressionwürdiger Teilgebiete zu nutzen.

Eine Aufgabenstellung im Rahmen der Geländeanalyse, die sich aus Diskussionen im Rahmen der vorliegenden Arbeit ergab, ist die Lokalisierung von Geländekanten im regulären Gittermodell. Es gibt viele Ansätze, Kanten aus Punktwolken (eine reguläre oder irreguläre Menge von Flächenpunkten) zu extrahieren. Die Wavelettransformation kann durch ihre Approximationseigenschaft im Zusammenhang mit der Lokalisierungseigenschaft Lösungen bieten, wobei die in (7.2) dargestellten Korrekturen unbedingt zu berücksichtigen sind. Wenn dann gleichzeitig noch eine, wenn auch geringe Kompression möglich ist, hat man, falls diese gewünscht ist, einen doppelten Effekt erzielt.

In 7.3 wurden unter Ausnutzung der Approximationseigenschaft aus den Waveletkoeffizienten Höhen- und Gefällelinien abgeleitet. Die allgemeinen Zusammenhänge zwischen der Approximation der Ableitungen und den differentialgeometrischen Beziehungen (4.2) ermöglichen die Lösung weiterer Aufgabenstellungen; z.B. könnte man durch die unterschiedliche Auflösung in den einzelnen Skalen, die mehr oder weniger grobe Mittelwerte der Ableitungen liefern, die Probleme des Rauschens bei der Berechnung der Gelände­krümmung und anderer differentialgeometrischer Größen in den Griff bekommen.

Die Extrapolation der Gitterschnittpunkte auf die Gitterknoten (Matrixelemente) nach 4.3.2 nutzt den Schnitt zweier die Geländeﬂäche approximierenden Ebenen. Es scheint überlegenswert, ob die beiden an die Geländekante angrenzenden Flächen durch hyperbolische Paraboloiden approximiert werden können, um auch gekrümmte Geländekanten zu beschreiben. Ein Ansatz dazu wird im Anhang C skizziert.

Literatur

- BARTELME, N. (2000). *Geoinformatik. Modelle, Strukturen, Funktionen*. Springer. Berlin, Heidelberg, New York.
- BARTHELMES, M. H., BALLANI, O. G. UND KLEES, R. (1994). *On the application of wavelets in geodesy*. In Hotine-Marussi Symposium on Mathematical Geodesy, L'Aquila, Italy.
- BAYER, M., BETH, S. UND MAIER, T. (1998). *Wavelet modelling of vectorial components in gravitational field determination*. In Freedon, W., Hrsg., Progress in Geodetic Science at GW 98.
- BETHGE, F. (1994). *Genauigkeit geometrischer Grössen aus Vektordaten*. Deutsche Geod. Komm., Reihe C, Nr.473. München.
- BETHGE, F., MEIER, S. UND SEEGRAEF, C. (1997). *Glättungsfilter im Wavelet- und Fourier-Bereich*. Allgemeine Vermessungsnachrichten, S. 341–348.
- BEYER, G. (1999). *Representation and wavelet transformation of relief-related space curves*. In International Archives of Photogrammetry and Remote Sensing, Band 32, S. 49–54.
- BEYER, G. (2000a). *A grid-based edge model for the wavelet transformation of hybrid dtm*. In International Archives of Photogrammetry and Remote Sensing, Band 33, S. 39–42.
- BEYER, G. (2000b). *Wavelet-Transformation hybrider Geländemodelle mit rasterbasierter Kanteninfor-mation*. Photogrammetrie, Fernerkundung, Geoinformation, S. 247–257.
- BEYER, G. (2002). *Terrain inclination and curvature from wavelet coefficients. approximation formulae for the relief*. Journal of Geodesy (reviewed).
- BEYER, G. UND MEIER, S. (2001). *Geländeneigung und -wölbung aus Waveletkoeffizienten. Approxima-tionsformeln für Profile*. Zeitschrift für Vermessungswesen, S. 23–33.
- BLATTER, C. (1998). *Wavelets - Eine Einführung*. Vieweg.
- BORKOWSKI, A. (1994). *Stochastisch-geometrische Beschreibung, Filterung und Präsentation des Reliefs*. Reihe C, Heft 431. Deutsche Geodätische Kommission.
- BÄR, G. (1996). *Geometrie: eine Einführung in die analytische und konstruktive Geometrie*. Mathematik für Ingenieure und Naturwissenschaftler. Teubner.
- BUCKHEIT, J. B., CHEN, S., DONOHO, D. L. UND SCARGLE, J. D. (1995). *WaveLab Reference Manual*. Stanford University.
- DAUBECHIES, I. (1992). *Ten Lectures on Wavelets*, Band 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia.
- EBNER, H. (1987). *Digital Terrain Models for High Mountains*. Mountain Research and Development, 7(4):353–356.
- GOODCHILD, M. F. UND MARK, D. M. (1987). *The fractal nature of geographic phenomena*. In Annals of the Association of American Geographers, 77(2), S. 265–278.
- GROSS, M. H., STAADT, O. G. UND GATTI, R. (1996). *Efficient triangular surface approximations using wavelets and quadtree data structures*. In IEEE Transactions on Visualisation and Computer Graphics, Band 2.
- HAKE, G. (1982). *Kartographie I*. Sammlung Göschen. Walter de Gruyter.
- HALMER, A., HEITZINGER, D. UND KAGER, H. (1996). *3D-Surface Modelling with Basic Topologic Elements*. In International Archives of Photogrammetry and Remote Sensing. Vol. XXXI, Part B4, S. 407–412.
- HODGSON, M. E. (1998). *Comparison of Angles from Surface Slope / Aspect Algorithms*. Cartographie and Geographic Information Systems, 25(3):173–185.
- IPF (1991). *Programmsystem SCOP zur Erstellung, Wartung und Anwendung digitaler Geländemodelle*. Institut für Photogrammetrie und Fernerkundung. TU Wien.

- KIEFNER, M. UND HAHN, M. (2000). *Image compression versus matching accuracy*. In International Archives of Photogrammetry and Remote Sensing, Band 33, S. 316–323.
- KRAUS, K. (2000). *Photogrammetrie. Band 3: Topographische Informationssysteme*. Duemmlerbuch. Duemmler.
- KREYSZIG, E. (1957). *Differentialgeometrie*. Akademische Verlagsgesellschaft Geest & Portig. Leipzig.
- LOUIS, A. K., MAAS, P. UND RIEDER, A. (1994). *Wavelets*. B.G.Teubner. Stuttgart.
- MALLAT, S. (1989). *A theory for multiresolution signal decomposition: The wavelet representation*. In IEEE Transactions on Pattern Analysis and Machine Intelligence, Band 11, S. 674–693.
- MALLAT, S. UND ZHONG, S. (1992). *Wavelet transform maxima and multiscale edges*. In et. al., M. B. R., Hrsg., *Wavelets and Their Applications*, S. 67–104. Jones and Bartlett Publishers.
- MEIER, S. (2002). *Zur K-Frage: Kompressionsraten der schnellen Wavelettransformation aus statistischer Sicht*. Zeitschrift für Vermessungswesen (eingereicht).
- MIETUSCH, L. (1999). *Interpolation von diskreten Wavelettransformierten*. Diplomarbeit. Institut für Planetare Geodäsie. Technische Universität Dresden.
- O'NEILL, M. P. UND MARK, D. M. (1996). *The psi-splot: A useful representation for digital cartographic lines*. In Proceedings AutoCarto 8. Baltimore, Maryland, S. 231–240.
- PFEIFER, N. UND POTTMANN, H. (1996). *Surface Models on the Basis of a Triangular Mesh - Surface Reconstruction*. In International Archives of Photogrammetry and Remote Sensing. Vol. XXXI, Part B3, S. 638–643.
- RICHTER, M. (2001). *Konstruktion von Gefälle- und Höhenlinien des Reliefs aus Waveletkoeffizienten*. Diplomarbeit. Institut für Planetare Geodäsie. Technische Universität Dresden.
- SCHÖNE, W. (1987). *Differentialgeometrie*. Mathematik für Ingenieure und Naturwissenschaftler. Teubner.
- SCHNEIDER, R. (1996). *Basiseigenschaften von Wavelets zur Multiskalenauswertung digitaler Signale*. Zeitschrift für Photogrammetrie und Fernerkundung, 64(3):75–83.
- SINGER, T. (1999). *Approximation der Geländeneigung mittels Wavelettransformation eines digitalen Höhenmodells (DHM)*. Diplomarbeit. Institut für Planetare Geodäsie. Technische Universität Dresden.
- STEINIGER, S. (2001). *Vorbereitung von raumbezogenen Geoinformationen zur Wavelettransformation*. Studienarbeit. Institut für Planetare Geodäsie. Technische Universität Dresden.
- STOYAN, D. UND STOYAN, H. (1992). *Fraktale - Formen - Punktfelder : Methoden der Geometrie-Statistik*. Akademie-Verlag. Berlin.
- STRANG, G. UND NGUYEN, T. (1997). *Wavelets and Filter Banks*. Wellesley-Cambridge Press.
- TRIEBFÜRST, B. (1999). *Entwicklung waveletbasierter Kompressionsverfahren für Fernerkundungsdaten und deren Einsatz in einem geographischen Informationssystem*. Dissertation. Institut für Informatik der Universität Freiburg.
- VOLK, M. UND KOLEM, R. (1998). *Möglichkeiten bei der Berechnung eines Gletscherinventars mittels GIS und digitaler Bildverarbeitung am Beispiel des Liefde- und Bochfjordgebietes, NW Spitzbergen(Svalbard)*. In PFG 1/1998, S. 27–40.
- VOLK, M. UND STEINHARDT, U. (1998). *Integration unterschiedlich erhobener Datenebenen in GIS für landschaftsökologische Bewertungen im mitteldeutschen Raum*. In PFG 6/1998, S. 349–362.
- VOSS, H. H. UND MORGENSTERN, D. (1997). *Interoperable Geowissenschaftliche Informationssysteme (IOGIS)*. GIS, (2):5–8.
- WALZ, U. UND SCHUMACHER, U. (1998). *Zum Einsatz des Geoinformationssystems SPANS für landschaftsökologische Fragestellungen*. PFG.

-
- WEICHSEL, M. (2001). *Reliefglättung mit Erhaltung von Bruchkanten*. Diplomarbeit. Institut für Planetare Geodäsie. Technische Universität Dresden.
- WERSCHLEIN, T. (1996). *Frequenzbasierte Linienrepräsentationen für die kartografische Generalisierung*. Diplomarbeit. Geographisches Institut der Universität Zürich.
- WICKERHAUSER, M. V. (1994). *Adapted Wavelet Analysis from Theory to Software*. A. K. Peters, Ltd. Wellesley.
- ZAHN, C. T. UND ROSKIES, R. Z. (1972). *Fourier descriptors for plane closed curves*. In IEEE Trans. Comp., vol. c-21 (1972) 3, S. 269 – 281.

Anhang A

Probleme bei der Signalanalyse mit der schnellen 2D-Wavelettransformation

Die zweidimensionale stationäre Wavelettransformation liefert bei n Skalen $n + 1$ Matrizen, was zu einer Vergrößerung des Datenumfanges durch eine Vielzahl redundanter Daten führt. Für die Signalanalyse bestehen aber bei der schnellen Wavelettransformation die gleichen Probleme, die bereits bei der eindimensionalen Wavelettransformation erwähnt wurden. Anhand des in der Literatur oft erwähnten Beispiels eines Quadrates mit seinen beiden Diagonalen (z. B. Datensatz 'BoxWithCross' im Tool Wavelab) kann dies leicht illustriert werden.

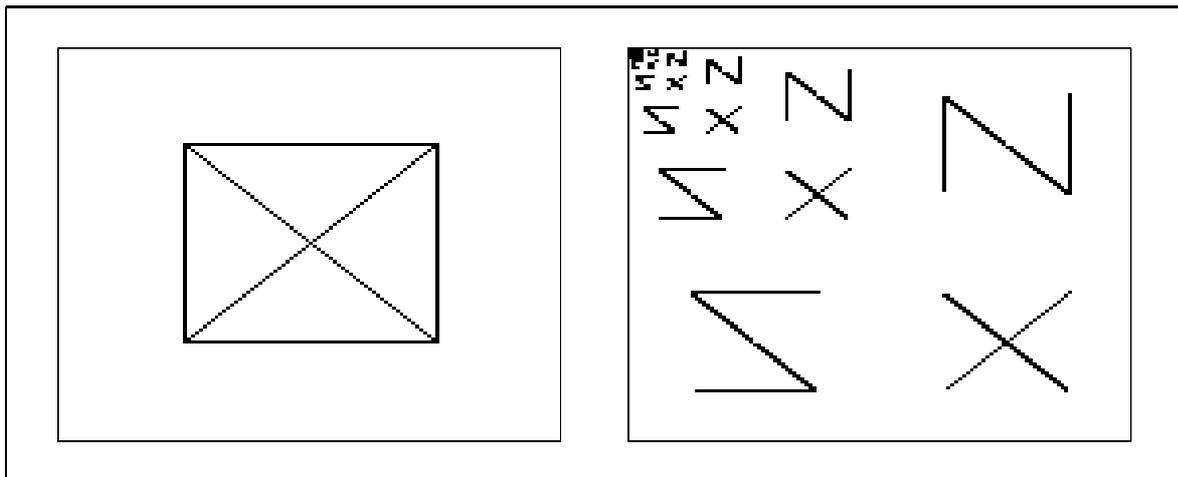


Abbildung A-1: Schnelle 2D-Wavelettransformation des Datensatzes 'BoxWithCross'.

In Abb. A-1, links, ist die Matrix des Originalsignals zu sehen. Die schwarzen Pixel entsprechen dem Wert Eins und die weißen Pixel dem Wert Null. Im rechten Teil des Bildes ist die visualisierte Matrix der Waveletkoeffizienten bei der schnellen Wavelettransformation mit dem Haar-Wavelet dargestellt. Die schwarzen Pixel entsprechen Waveletkoeffizienten $w_{ik} \neq 0$ und die weißen Pixel Waveletkoeffizienten $w_{ik} = 0$. Es ist zu sehen, daß dabei eine Diagonale nicht detektiert wird. Die Ursache ist in den zu großen Schrittweiten $\Delta x, \Delta y$ beim Verschieben des Wavelets zu suchen. Wird andererseits das zyklisch um ein Pixel nach rechts verschobene Signal wavelettransformiert, erhält man das Ergebnis in Abb. A-2. Diesmal wird die andere Diagonale nicht detektiert.

Betrachtet man einen Ausschnitt des Signals z. B. in der Umgebung des Kreuzungspunktes (Abb. A-3), wird der Sachverhalt sofort klar. Die Überdeckung des Signals mit den Wavelets der ersten Skale, die mit der Schrittweite $\Delta x = \Delta y = 2$ über das Signal geschoben werden, sind durch Trennlinien hervorgehoben. Links ist ein Ausschnitt des Originalsignals zu sehen. Die $(2,2)$ -Teilmatrizen Z_H , die die Elemente der „Hauptdiagonale“ enthalten, sind dunkel und die $(2,2)$ -Teilmatrizen Z_N , die die Elemente der „Nebendiagonale“ enthalten, hell hinterlegt.

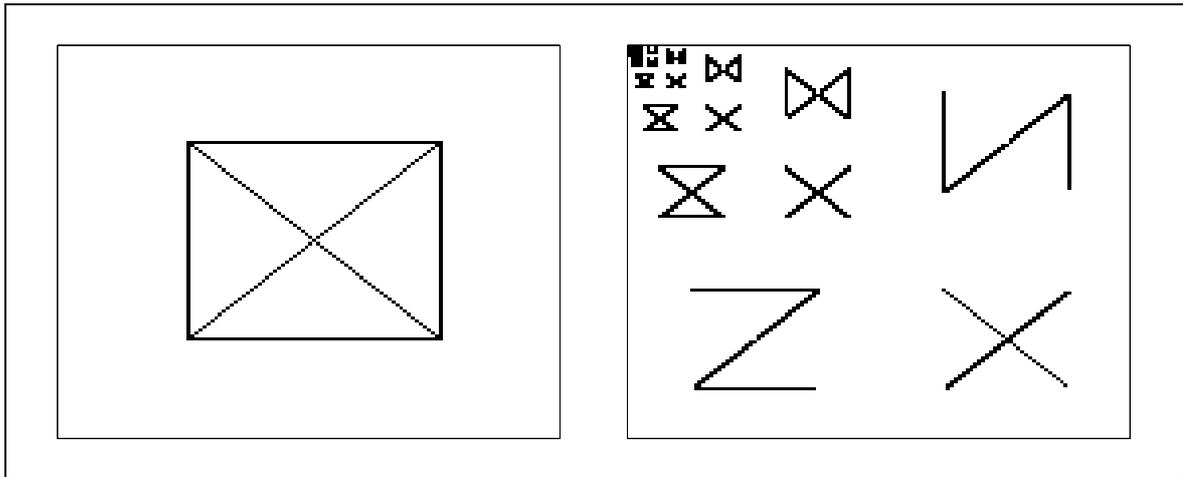


Abbildung A-2: Schnelle 2D-Wavelettransformation des um ein Pixel nach rechts verschobenen Datensatzes 'BoxWithCross'.

1	0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	1	
0	1	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	1
0	0	1	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0
0	0	0	1	0	0	0	0	0	0	0	0	0	1	0	0	0	0
0	0	1	0	1	0	0	0	0	0	0	0	1	0	1	0	0	0
0	1	0	0	0	1	0	0	0	0	0	1	0	0	0	1	0	0

Abbildung A-3: Ausschnitt aus den beiden Signalmatrizen mit markierten Bereichen der überdeckenden Wavelets.

Für die von den Wavelets überdeckten Signalbereiche treten bei den Diagonalen folgende Fälle auf:

$$Z_H = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad Z_{N_1} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad Z_{N_2} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}. \quad (1.0-1)$$

Das Skalarprodukt mit den drei Wavelets

$$w_{HH} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}, \quad w_{HT} = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}, \quad w_{TH} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix} \quad (1.0-2)$$

ergibt dann

$$\begin{aligned} \langle w_{HH}, Z_H \rangle &= 1, & \langle w_{HT}, Z_H \rangle &= 0, & \langle w_{TH}, Z_H \rangle &= 0, \\ \langle w_{HH}, Z_{N_1} \rangle &= \frac{1}{2}, & \langle w_{HT}, Z_{N_1} \rangle &= -\frac{1}{2}, & \langle w_{TH}, Z_{N_1} \rangle &= -\frac{1}{2}, \\ \langle w_{HH}, Z_{N_2} \rangle &= \frac{1}{2}, & \langle w_{HT}, Z_{N_2} \rangle &= \frac{1}{2}, & \langle w_{TH}, Z_{N_2} \rangle &= \frac{1}{2}. \end{aligned} \quad (1.0-3)$$

Die Hauptdiagonale wird von den Wavelets w_{HT} und w_{TH} nicht erkannt, da die entsprechende Waveletposition durch die Schrittweite $\Delta x = \Delta y = 2$ übersprungen wird.

In Abb. A-3, rechts ist ein Ausschnitt des verschobenen Signals zu sehen. Die (2,2)-Teilmatrizen Z_N , die die Elemente der „Nebendiagonale“ enthalten, sind diesmal dunkel und die (2,2)-Teilmatrizen Z_H , die die Elemente der „Hauptdiagonale“ enthalten, hell hinterlegt.

Für die von den Wavelets überdeckten Signalbereiche des verschobenen Signals treten jetzt folgende Fälle auf:

$$Z_{H_1} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad Z_{H_2} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad Z_{ND} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad (1.0-4)$$

Das Skalarprodukt mit den drei Wavelets w_{HH} , w_{HT} , w_{TH} ergibt dann

$$\begin{aligned} \langle w_{HH}, Z_{H_1} \rangle &= -\frac{1}{2}, & \langle w_{HT}, Z_{H_1} \rangle &= \frac{1}{2}, & \langle w_{TH}, Z_{H_1} \rangle &= -\frac{1}{2}, \\ \langle w_{HH}, Z_{H_2} \rangle &= -\frac{1}{2}, & \langle w_{HT}, Z_{H_2} \rangle &= -\frac{1}{2}, & \langle w_{TH}, Z_{H_2} \rangle &= \frac{1}{2}, \\ \langle w_{HH}, Z_N \rangle &= -1, & \langle w_{HT}, Z_N \rangle &= 0, & \langle w_{TH}, Z_N \rangle &= 0. \end{aligned} \quad (1.0-5)$$

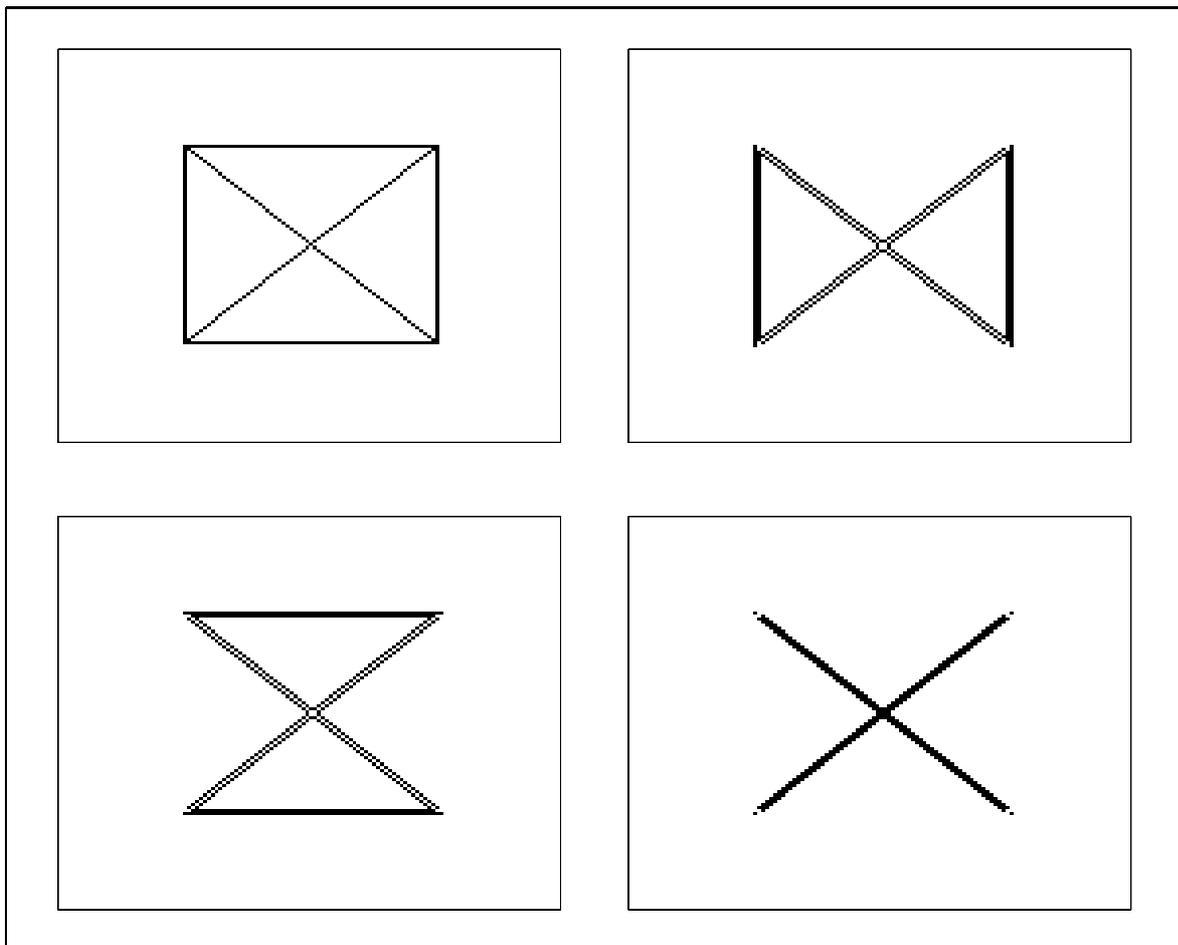


Abbildung A-4: Stationäre 2D-Wavelettransformation des Datensatzes 'BoxWithCross'.

Analog zu oben wird hier die Nebendiagonale nicht erkannt. Die stationäre Wavelettransformation liefert in der ersten Skale die WK-Matrizen, die in Abb. A-4 visualisiert sind. Entsprechend der Approximationseigenschaft des Haar-Wavelets werden mit w_{HT} die senkrechten, mit w_{TH} die waagerechten und mit w_{HH} die senkrechten und waagerechten Linien nicht detektiert. In den höheren Skalen werden die Linien in gleicher Weise detektiert, jedoch wegen des breiter werdenden Supports immer unschärfer. In Abb. A-5 sind die drei Matrizen der Waveletkoeffizienten in der 3. Skale dargestellt.

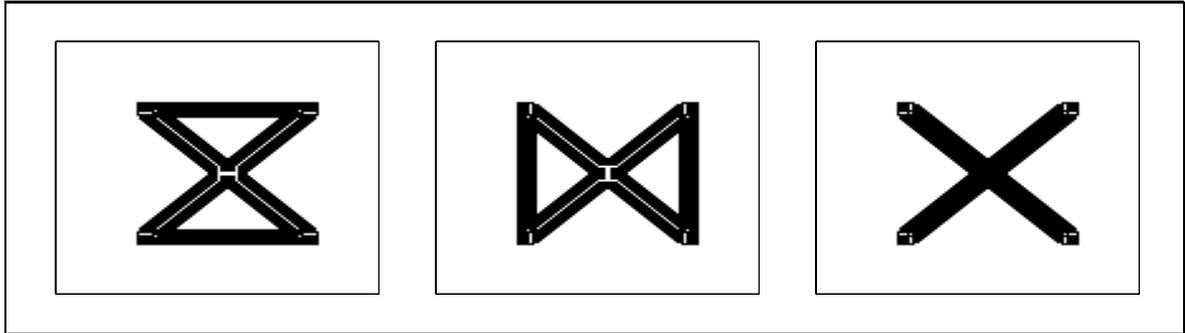


Abbildung A-5: Stationäre 2D-Wavelettransformation des Datensatzes 'BoxWithCross' in der 3. Skale.

Anhang B

Filtereigenschaften des Daub4-Wavelets

In den in 7.1 und 7.2 hergeleiteten Skalierungs- und Translationsformeln blieb die Frequenzabhängigkeit unberücksichtigt. Es handelt sich damit um Näherungsformeln, deren Güte von der jeweiligen Frequenz des Signals abhängt. Da sich eine geschlossene Darstellung der Fehleranalyse allgemein sehr schwierig gestaltet, wird eine Fehlerbetrachtung am Beispiel des Daub4-Wavelets durchgeführt.

B.1 Filtercharakteristik

Die Durchlaßcharakteristik eines digitalen Filters berechnet sich nach

$$\begin{aligned} G &= \sum_k g_k e^{-j\omega t_k} \\ &= \sum_k g_k (\cos \omega t_k - j \sin \omega t_k) \\ &= \sum_k g_k \cos k \omega \Delta t - j \sum_k g_k \sin k \omega \Delta t . \end{aligned}$$

Für eine gerade Filterfunktion ($g_k = g_{-k}$) gilt speziell

$$G = g_0 + 2 \sum_{k=1}^N g_k \cos k \omega \Delta t$$

und für eine ungerade Filterfunktion ($g_k = -g_{-k}$) gilt speziell

$$G = -2j \sum_{k=1}^N g_k \sin k \omega \Delta t .$$

B.1.1 Filtercharakteristik in Originallage

Die Filterkoeffizienten des Daub4-Wavelets sind

$$g = h = \frac{1}{8} \sqrt{2} \begin{bmatrix} 1 - \sqrt{3} & -3 + \sqrt{3} & 3 + \sqrt{3} & -1 - \sqrt{3} \end{bmatrix} .$$

Die Indizierung der Filterkoeffizienten entsprechend der Indizierung bei der stationären Wavelettransformation ist wie folgt vorzunehmen:

$$\begin{array}{cccc} -3 & -2 & -1 & 0 \\ | & | & | & | \\ [1 - \sqrt{3} & -3 + \sqrt{3} & 3 + \sqrt{3} & -1 - \sqrt{3}] . \end{array}$$

Für die Filtercharakteristik ergibt sich mit $\alpha := \omega \frac{\Delta t}{2}$, $\cos -\theta = \cos \theta$ und $\sin -\theta = -\sin \theta$

$$\begin{aligned} G_{Daub4} = & \frac{1}{8} \sqrt{2} \left\{ (1 - \sqrt{3}) \cos 6\alpha + (-3 + \sqrt{3}) \cos 4\alpha + (3 + \sqrt{3}) \cos 2\alpha + (-1 - \sqrt{3}) \right. \\ & \left. + j \left\{ (1 - \sqrt{3}) \sin 6\alpha + (-3 + \sqrt{3}) \sin 4\alpha + (3 + \sqrt{3}) \sin 2\alpha \right\} \right\} . \end{aligned}$$

Für die sin- und cos-Funktionen des Vielfachen eines Winkels gilt

$$\begin{aligned} \sin 2\alpha &= 2 \sin \alpha \cos \alpha \\ \sin 4\alpha &= 8 \sin \alpha \cos^3 \alpha - 4 \sin \alpha \cos \alpha \\ \sin 6\alpha &= 32 \sin \alpha \cos^5 \alpha - 32 \sin \alpha \cos^3 \alpha + 6 \sin \alpha \cos \alpha \\ \cos 2\alpha &= 2 \cos^2 \alpha - 1 \\ \cos 4\alpha &= 8 \cos^4 \alpha - 8 \cos^2 \alpha + 1 \\ \cos 6\alpha &= 32 \cos^6 \alpha - 48 \cos^4 \alpha + 18 \cos^2 \alpha - 1 . \end{aligned}$$

Damit ergibt sich für die Filtercharakteristik schließlich

$$\begin{aligned} \operatorname{Re}(G_{Daub4}) &= \frac{1}{8} \sqrt{2} \left\{ 32 \cos^6 \alpha - 24 \cos^4 \alpha + 48 \cos^2 \alpha - 8 \right. \\ & \quad \left. + \sqrt{3} \left\{ -32 \cos^6 \alpha + 56 \cos^4 \alpha - 24 \cos^2 \alpha \right\} \right\} , \\ \operatorname{Im}(G_{Daub4}) &= \frac{1}{8} \sqrt{2} \left\{ \sin \alpha \left\{ 32 \cos^5 \alpha - 56 \cos^3 \alpha + 24 \cos \alpha \right\} \right. \\ & \quad \left. + \sqrt{3} \sin \alpha \left\{ -32 \cos^5 \alpha + 40 \cos^3 \alpha - 9 \cos \alpha \right\} \right\} . \end{aligned}$$

Die Ortskurve der Filtercharakteristik ist in Abb. B-1 zu sehen.

Das Verhältnis zwischen Phasenwinkel und Zeit ist die Kreisfrequenz ω . Das bedeutet für den Zusammenhang zwischen Phasenverschiebung ϕ und Ortsverschiebung x_ϕ

$$x_\phi = \frac{\phi}{\omega} = \frac{\phi}{2\alpha} \Delta t .$$

Aus der Filtercharakteristik kann damit die Ortsverschiebung in Abhängigkeit von α bzw. $\frac{\omega}{\omega_g}$ bestimmt werden.

Die frequenzabhängige Ortsverschiebung x_ϕ ist in Abb. B-2 dargestellt.

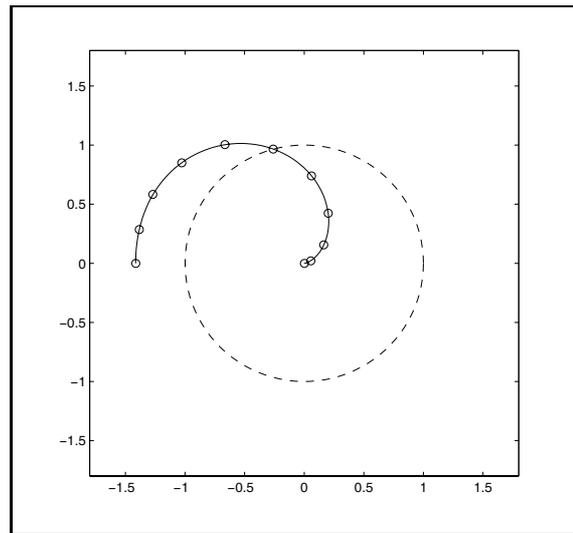


Abbildung B-1: Ortskurve der Filtercharakteristik des Daub4-Wavelets

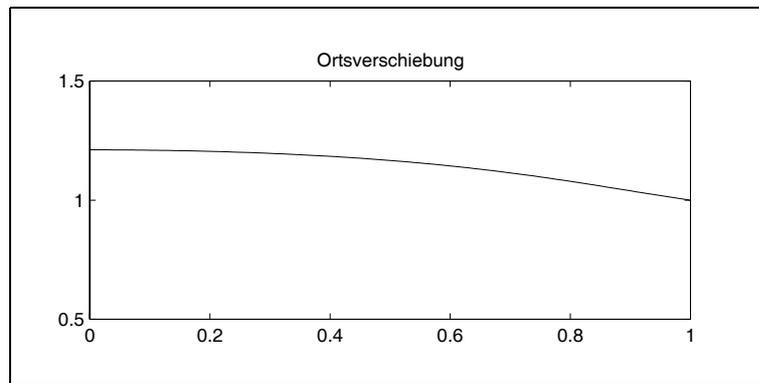


Abbildung B-2: Ortsverschiebung des Daub4-Wavelets bei der stationären WT als Vielfaches der Tastweite

B.1.2 Filtercharakteristik in Symmetrielage

Zur übersichtlichen Berechnung der Filtercharakteristik des Daub4-Wavelets werden die Filterkoeffizienten zweckmäßig in einen geraden und einen ungeraden Anteil zerlegt.

$$g = h = \frac{1}{8}\sqrt{2}\left\{ \begin{bmatrix} 1 & -3 & 3 & -1 \end{bmatrix} + \sqrt{3}\begin{bmatrix} -1 & 1 & 1 & -1 \end{bmatrix} \right\}.$$

Die Filterkoeffizienten sind dann wie folgt zu indizieren

$$\begin{array}{cccccc} -3 & -2 & -1 & 0 & 1 & 2 & 3 & & -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ | & | & | & | & | & | & | & & | & | & | & | & | & | & | \\ [& 1 & & -3 & & 3 & & -1 &] & [& -1 & & 1 & & 1 & & -1 &] . \end{array}$$

Mit der dadurch zu berücksichtigenden Tastweite $\frac{\Delta t}{2}$ und mit $\alpha := \omega \frac{\Delta t}{2} = \frac{\pi}{2} \frac{\omega}{\omega_g}$ ergibt sich für die Filtercharakteristik

$$\begin{aligned}
 G_{Daub4} &= \sum_k g_k \cos k \alpha - j \sum_k g_k \sin k \alpha \\
 &= \frac{1}{8} \sqrt{2} \left\{ -2j(3 \sin \alpha - \sin 3\alpha) + 2\sqrt{3}(\cos \alpha - \cos 3\alpha) \right\} \\
 &= \frac{1}{8} \sqrt{2} \left\{ -2j \cdot 4 \sin^3 \alpha + 2\sqrt{3} \cos \alpha \cdot \sin^2 \alpha \right\} \\
 &= \sqrt{2} \left\{ -j \sin^3 \alpha + \sqrt{3} \cos \alpha \sin^2 \alpha \right\} \\
 &= \sqrt{2} \sin^2 \alpha (-j \sin \alpha + \sqrt{3} \cos \alpha)
 \end{aligned}$$

und in Exponentialschreibweise

$$\begin{aligned}
 G_{Daub4} &= \sqrt{2} \sin^2 \alpha \sqrt{\sin^2 \alpha + 3 \cos^2 \alpha} \cdot e^{j\phi} \\
 \text{mit} \quad \tan \phi &= \frac{-\sin \alpha}{\sqrt{3} \cos \alpha} = -\frac{\sqrt{3}}{3} \tan \alpha .
 \end{aligned}$$

Damit erhält man die Ortskurve in Abb. B-3.

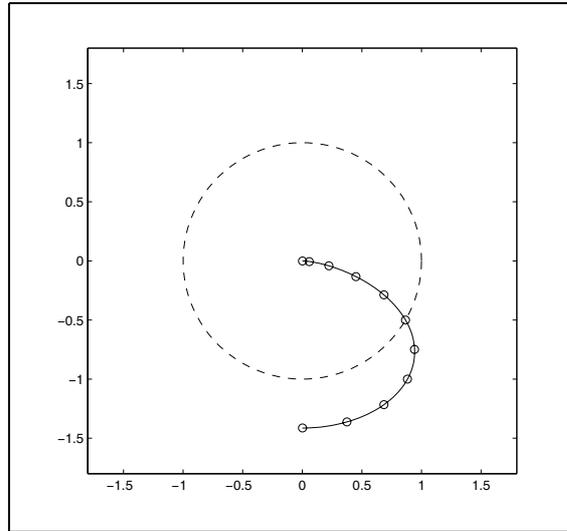


Abbildung B-3: Ortskurve der Filtercharakteristik des Daub4-Wavelets

Für das auf das Zentrum des Supports bezogene Filter mit den Koeffizienten des Daub4-Wavelets ist die frequenzabhängige Ortsverschiebung x_ϕ in Abb. B-4 zu sehen.

Betragsbildung und Ausklammern der Filtercharakteristik des zweifachen Differenzenfilters $|G_{Diff2}| = 4 \sin^2 \alpha$ führt auf

$$|G_{Daub4}| = 4 \sin^2 \alpha \cdot \frac{\sqrt{2}}{4} \sqrt{\sin^2 \alpha + 3 \cos^2 \alpha} .$$

Der in 7.1.3 hergeleitete Skalierungsfaktor zur Korrektur der Amplitude war für das Daubechies4-Wavelet

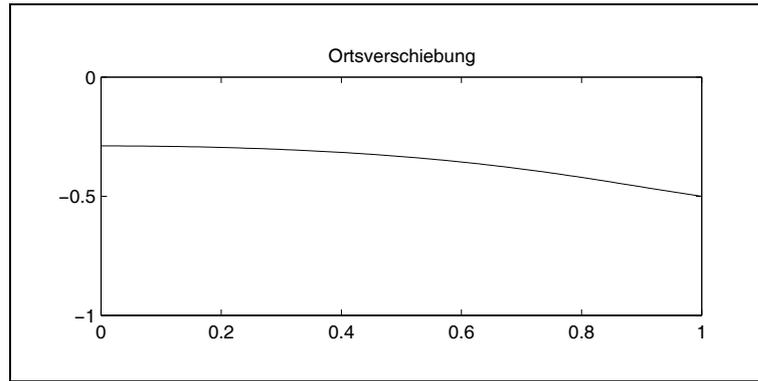


Abbildung B-4: Ortsverschiebung bei der stationären WT als Vielfaches der Tastweite

$$\alpha_{skal} = \frac{2}{m_2(\Delta t)^2}.$$

Bei den hier betrachteten harmonischen Schwingungen variabler Kreisfrequenz ω gilt für die Tastweite $\Delta\phi$ des Arguments ϕ bei einer festen Tastweite Δt im Zeitbereich

$$\Delta\phi = \omega\Delta t,$$

d.h. $\Delta\phi$ ist frequenzabhängig und damit auch der Skalierungsfaktor α_{skal} . Insgesamt läßt sich damit der Skalierungsfaktor α_{skal} in Abhängigkeit von $\alpha := \omega \frac{\Delta t}{2}$ darstellen:

$$\begin{aligned} \alpha_{skal} &= \frac{2}{m_2(\Delta\phi)^2} \stackrel{\Delta\phi=\omega\Delta t}{=} \frac{2}{m_2\omega^2(\Delta t)^2} \stackrel{\Delta t=\frac{\pi}{\omega_g}}{=} \frac{2\omega_g^2}{m_2\omega^2\pi^2} \stackrel{\alpha=\frac{\pi}{2}\frac{\omega}{\omega_g}}{=} \frac{2\pi^2}{m_2\pi^2 4\alpha^2} \\ &= \frac{1}{2m_2\alpha^2}. \end{aligned}$$

Unter Berücksichtigung des Skalierungsfaktors α_{skal} ergibt sich frequenzabhängig für die korrigierte Amplitude des Filterergebnisses

$$\begin{aligned} |G_{kor}| &= |G_{Daub4}| \cdot |\alpha_{skal}| = 4 \sin^2 \alpha \cdot \frac{\sqrt{2}}{4} \sqrt{\sin^2 \alpha + 3 \cos^2 \alpha} \cdot \frac{1}{2|m_2|\alpha^2} \\ &= \sqrt{\frac{1 + 2 \cos^2 \alpha}{3}} \cdot (\text{sinc } \alpha)^2. \end{aligned}$$

In Abb. B-5 sind die Durchlaßcharakteristiken für das zweifache Differenzenfilter und das Daub4-Wavelet sowie die korrigierte Amplitude der gefilterten harmonischen Schwingung zu sehen. Mit steigender Frequenz ist die approximierte Ableitung mit der frequenzunabhängig korrigierten Skalierung zu klein.

Im „gutartigen“ Gelände mit niedrigen Frequenzen relativ zur Grenzfrequenz bei der vorgegebenen Tastweite sind die Approximationsformeln mit ausreichender Genauigkeit zur Analyse der Geländeneigung und -wölbung geeignet.

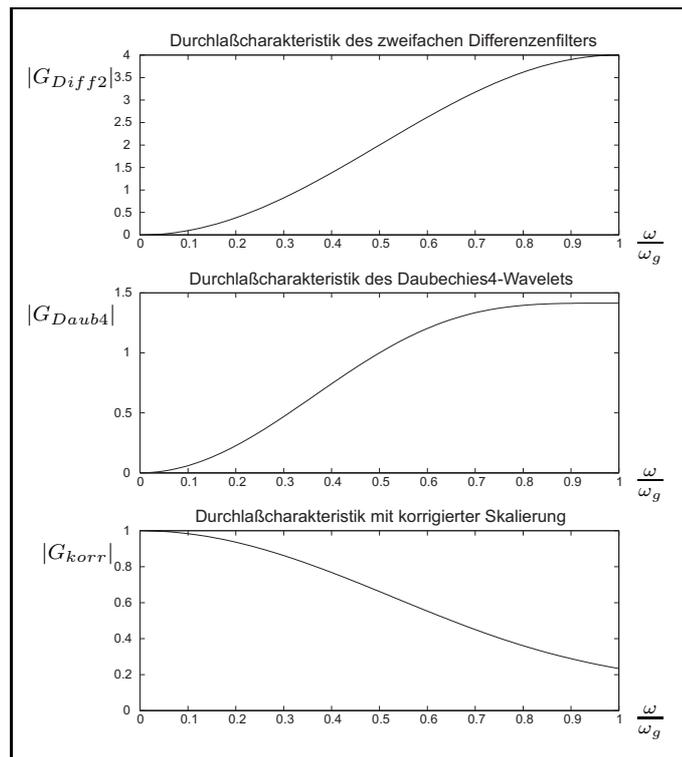


Abbildung B-5: Darstellung der Frequenzabhängigkeit der Skalierungskorrektur im Vergleich zur Durchlaßcharakteristik des zweifachen Differenzenfilters des Daub4-Wavelets

Es bleibt die Frage offen, ob in Abhängigkeit des konkreten Frequenzspektrums des Geländes eine jeweilige zusätzliche feste Korrektur an den Skalierungsfaktor α_{skal} angebracht werden sollte, um den mittleren Fehler zu minimieren.

B.2 Testrechnung zur Verifizierung der Ergebnisse

Um die Ergebnisse zu testen und zu veranschaulichen, wurde eine synthetische Funktion mit vier verschiedenen Auflösungen benutzt. Die verwendeten Tastweiten sind $\Delta t = \frac{\pi}{8}$, $\Delta t = \frac{\pi}{4}$ (Abb. B-6), $\Delta t = \frac{\pi}{2}$ und $\Delta t = \pi$. Das untersuchte Signal ist aus zwei \cos -Funktionen ($\cos t$, $\cos \frac{t}{2}$) zusammengesetzt. Die aus den unterschiedlichen Tastweiten resultierenden relativen Frequenzen $\frac{\omega}{\omega_g}$ für die beiden Funktionen $\cos t$, $\cos \frac{t}{2}$ sind in Tabelle B-1 zu sehen. Um einen Vergleich zwischen den zu erwartenden Amplituden lt. Filtercharakteristik (Abb. B-5, unten) und den Amplituden der Beispielrechnung zu erhalten, wurden die jeweils drei um das Maximum liegenden Waveletkoeffizienten (Abb. B-6, Tab. B-1). quadratisch interpoliert und das Maximum der Interpolation berechnet (Tab. B-1). In Tab. B-2 sind die Vergleichswerte für drei Frequenzen gegenübergestellt und zeigen die zu erwartende Übereinstimmung. In der unteren Zeile stehen die tatsächlichen Amplituden der Ableitungen, die lt. Filtercharakteristik für die Frequenz $\omega = 0$ angenommen werden.

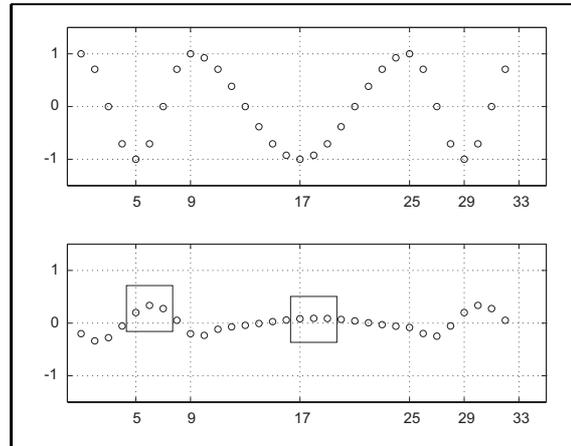


Abbildung B-6: Vergleich der Skalierungen.

Tabelle B-1: Schätzung der Amplitude aus den Stützwerten. Maximum der quadratischen Interpolation

Δt	$\cos t$					$\cos \frac{t}{2}$				
	$\frac{\omega}{\omega_g}$	w_1	w_2	w_3	max	$\frac{\omega}{\omega_g}$	w_1	w_2	w_3	max
π	1					$\frac{1}{2}$	-0.0642	0.1598	0.0428	0.1640
$\frac{\pi}{2}$	$\frac{1}{2}$	-0.1713	0.6393	0.1713	0.6508	$\frac{1}{4}$	0.1324	0.2227	0.1826	0.2251
$\frac{\pi}{4}$	$\frac{1}{4}$	0.5296	0.8909	0.7303	0.9005	$\frac{1}{8}$	0.2167	0.2428	0.2320	0.2436
$\frac{\pi}{8}$	$\frac{1}{8}$	0.8668	0.9713	0.9279	0.9745	$\frac{1}{16}$				

Tabelle B-2: Vergleich der Amplituden lt. Filtercharakteristik mit den erzielten Amplituden in der Beispielrechnung.

$\frac{\omega}{\omega_g}$	$\cos t$		$\cos \frac{t}{2}$	
	Amplitude lt. Filtercharakteristik	korrigierte Amplitude der WK	Amplitude lt. Filtercharakteristik	korrigierte Amplitude der WK
0.5	0.6618	0.6508	0.1655	0.1640
0.25	0.9021	0.9005	0.2255	0.2251
0.125	0.9746	0.9745	0.2437	0.2436
0	1		0.25	

Anhang C

Kantenextrapolation bei gekrümmten Kanten

Eine Möglichkeit, nicht Ebenen zur Extrapolation zu verwenden, wie in 4.3.2 getan, sondern hyperbolische Paraboloiden, die auch gekrümmte Geländelinien als Schnittlinie beschreiben können, soll hier kurz skizziert werden.

Im Gittermodell einer Fläche wird das Flächensegment (Netzmasche) einer Gitterzelle durch vier Gitterwerte repräsentiert. Lokal handelt es sich um ein i.allg. windschiefes Viereck, das als Rand eines hyperbolischen Paraboloids aufgefaßt werden kann. Verläuft durch die Masche eine Bruchlinie, so wird die Fläche in zwei „glatte“ Flächen zerlegt, die sich in der Bruchlinie schneiden. In einem regulären Gittermodell ist das hyperbolische Paraboloid eine naheliegende, einfache Flächenapproximation mit günstigen differentialgeometrischen Eigenschaften.

Aus den Höhen in den Gitterpunkten (Höhenmatrix) und den Gitterschnittpunkten können zwei hyperbolische Paraboloiden definiert werden. Man hat dann in einer Zelle zweimal vier Punkte, die vier Punkte des Höhenmodells und vier extrapolierte Punkte zu speichern. Nachfolgend werden die geometrischen Zusammenhänge bereitgestellt.

C.1 Gleichung des hyperbolischen Paraboloids

Die allgemeine Gleichung einer Fläche zweiter Ordnung ist

$$\begin{aligned} 0 &= a_{00} + 2\underline{a}^\top \underline{x} + \underline{x}^\top A \underline{x} \\ &= a_{00} + 2a_{01}x_1 + 2a_{02}x_2 + 2a_{03}x_3 \\ &\quad + 2a_{11}x_1^2 + 2a_{22}x_2^2 + 2a_{33}x_3^2 + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + 2a_{23}x_2x_3. \end{aligned} \quad (3.1-1)$$

Das hier interessierende hyperbolische Paraboloid ist eine Regelfläche mit zwei Geradenscharen, deren Geraden jeweils parallel zu den Gitterebenen sind. D.h., die Gleichung der Fläche ist sowohl für $x_1 = \text{const}$ als auch für $x_2 = \text{const}$ eine Geradengleichung, also linear in x_2 und x_3 bzw. in x_1 und x_3 . Daraus folgt, daß für die Koeffizienten gilt

$$a_{11} = a_{22} = a_{33} = a_{13} = a_{23} = 0.$$

Die Gleichung des hyperbolischen Paraboloids π ist damit

$$0 = a_{00} + 2a_{01}x_1 + 2a_{02}x_2 + 2a_{03}x_3 + 2a_{12}x_1x_2. \quad (3.1-2)$$

Wegen der Homogenität der Koeffizienten kann man $a_{00} = \text{const}$ setzen (i.allg. $a_{00} = 1$ für $\underline{x} = [0 \ 0 \ 0]^\top \notin \pi$). Die Koeffizienten können somit aus den Koordinaten von vier Punkten der Fläche (z.B. der vier Eckpunkte) ermittelt werden.

Nach x_3 aufgelöst ergibt sich die Darstellung der Fläche als Funktion

$$x_3 = x_3(x_1, x_2) = \frac{a_{00} + 2a_{01}x_1 + 2a_{02}x_2 + 2a_{12}x_1x_2}{-2a_{03}}. \quad (3.1-3)$$

C.2 Schnitt zweier hyperbolischer Paraboloiden

Die beiden hyperbolischen Paraboloiden seien durch ihre Koordinatenvektoren

$$\begin{aligned} h_a &= [a_{00} \ a_{01} \ a_{02} \ a_{03} \ a_{12}]^\top, \\ h_b &= [b_{00} \ b_{01} \ b_{02} \ b_{03} \ b_{12}]^\top \end{aligned}$$

gegeben. Dann erhält man in der durch $x_1 = x_1^0 = \text{const}$ definierten, zu den Gitterebenen parallelen Ebene die beiden Geradengleichungen

$$\begin{aligned} x_3 = x_3(x_2) &= \frac{2a_{02} + 2a_{12}x_1^0}{-2a_{03}}x_2 + \frac{a_{00} + 2a_{01}x_1^0}{-2a_{03}} \\ &= m_1x_2 + n_1, \\ x_3 = x_3(x_2) &= \frac{2b_{02} + 2b_{12}x_1^0}{-2b_{03}}x_2 + \frac{b_{00} + 2b_{01}x_1^0}{-2b_{03}} \\ &= m_2(x_1^0)x_2 + n_2(x_1^0). \end{aligned} \tag{3.2-1}$$

Die Koordinatenvektoren der beiden Geraden sind

$$\begin{aligned} g_a &= [n_1 \ m_1 \ -1]^\top, \\ g_b &= [n_2 \ m_2 \ -1]^\top \end{aligned}$$

und der homogene Koordinatenvektor des Schnittpunkts damit

$$s = g_a \times g_b.$$

Daraus ergeben sich die inhomogenen Schnittpunktkoordinaten zu

$$\begin{aligned} x_2^s(x_1^0) &= \frac{n_1 - n_2}{m_2 - m_1}, \\ x_3^s(x_1^0) &= \frac{n_1m_2 - m_1n_2}{m_2 - m_1}. \end{aligned} \tag{3.2-2}$$

Das ist, wenn x_1^0 als Parameter eine Zellenfläche durchläuft, die Parameterdarstellung der Schnittkurve innerhalb einer Netzmasche.

Anhang D

Algorithmus zur Kantenrekonstruktion

Bei der Transformation der Folge von 3D-Koordinaten der Kanten (Bruchlinien, ggf. auch Formlinien) in eine der Flächenbeschreibung analogen Matrixstruktur geht der topologische Zusammenhang der Punktfolge verloren. Unter gewissen Voraussetzungen läßt sich jedoch diese Topologie zurückgewinnen.

Als Vorbereitungsschritt für den Extrapolationsalgorithmus wird die originale 3D-Punktfolge der Kante (Primärdaten) in eine Folge von Gitterschnittpunkten (Sekundärdaten) transformiert. Die Sekundärdaten werden nunmehr als Repräsentation der Kante angesehen. Die anschließende Extrapolation der Gitterschnittpunkte auf die Gitterpunkte transformiert die geometrischen 3D-Informationen der Kante in die Elemente einer Matrix extrapolierter Höhen. Die durch die Listenstruktur der ursprünglichen 3D-Punktfolge gegebene Kantentopologie ist in der Matrixstruktur nicht mehr unmittelbar vorhanden. Sie liegt nur mittelbar und nicht hundertprozentig rekonstruierbar in der Topologie der mit extrapolierten Höhen belegten Matrix vor. Es ist damit ein inverses Problem zu lösen:

- Es ist diejenige Kantentopologie zu finden, die die vorliegende Matrixtopologie mit dem Extrapolationsalgorithmus erzeugt.

Der Extrapolationsalgorithmus bildet eine gegebene Rasterschnitt-Topologie als Folge von Rasterschnitten in eine Matrix-Topologie als Netz von belegten Matrixelementen ab. Die Umkehrabbildung ist nicht immer eindeutig, jedoch in der überwiegenden Anzahl praktisch vorkommender Situationen.

Die Kante als Folge von Gitterschnittpunkten durchläuft eine Folge von Gitterzellen, wobei sie über eine Zellenkante in die Zelle eindringt und über eine andere Zellenkante die Zelle verläßt. Dabei gibt es zwei topologisch verschiedene Fälle, die zu zwei verschiedenen lokalen Matrixtopologien führen (Abb. D-1).

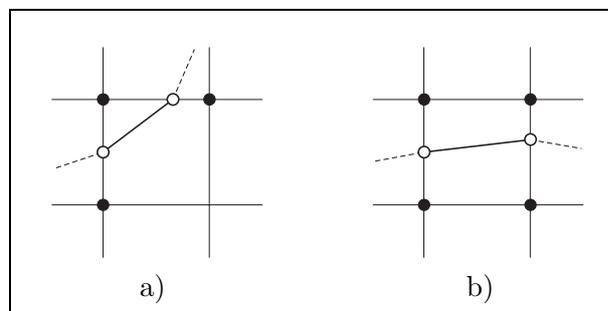


Abbildung D-1: Die beiden topologisch verschiedenen Zellendurchläufe und die resultierenden Matrixtopologien.

Es ist sofort zu sehen, daß aus der Teiltopologie einer Zelle keine eindeutige Inverse folgt. Im Fall a) ist die Inverse nur unter der Voraussetzung, daß die Zelle tatsächlich durchlaufen wird,

eindeutig. Diese lokale Matrixbelegung kann aber auch das Bild einer Gitterschnittfolge sein, die diese Zelle nicht durchläuft (Abb. D-2). Im Fall b) ist die Eindeutigkeit nicht einmal unter dieser Voraussetzung gegeben. Die Zelle könnte auch in senkrechter Richtung durchlaufen werden.

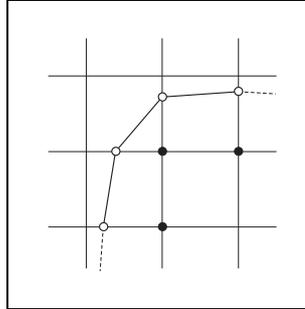


Abbildung D-2: Belegung der Gitterpunkte einer Zelle mit drei belegten Ecken ohne Kantendurchlauf.

Ein Algorithmus zur Invertierung läßt sich also nicht aus lokalen Entscheidungen in einer Zelle zusammensetzen. Es muß immer die nächste Nachbarschaft mit einbezogen werden.

Zunächst läßt sich eine notwendige Voraussetzung für eine durchlaufene Zelle formulieren:

- Wird eine Zelle von einer Gitterschnittfolge durchlaufen (Abb. D-3, links), so hat sie mindestens drei belegte Gitterpunkte.

Diese Voraussetzung legt den ersten Schritt der Analyse der vorliegenden Matrixtopologie fest:

- Suche alle Zellen, die mindestens drei belegte Ecken haben (Abb. D-3, rechts).

Wird diese Suche zeilenweise (z.B. o.B.d.A. von unten nach oben und von links nach rechts) durchgeführt, so ergibt sich in der letzten gefundenen Zelle R_0 , die diese Bedingung erfüllt, die topologische Situation in Abb. D-3. Allgemein sind für R_0 die Eckenbelegungen in Abb. D-4 möglich.

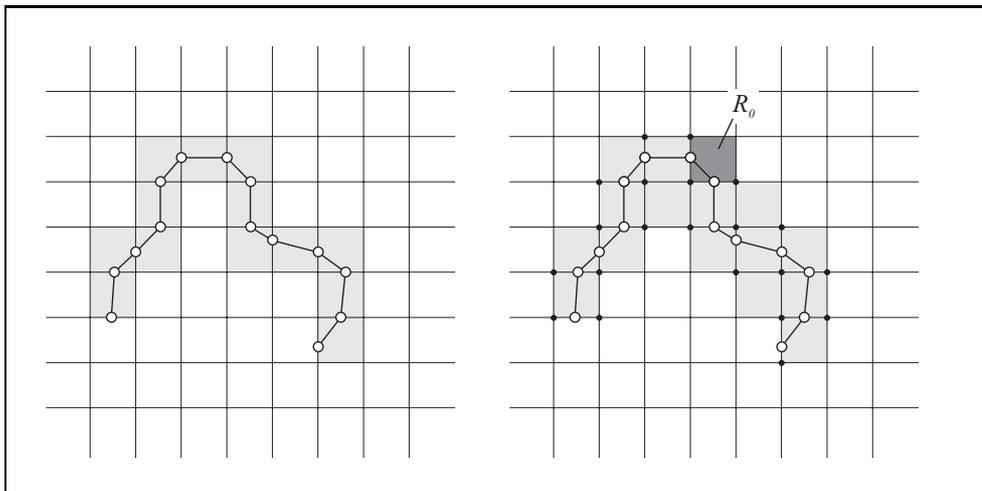


Abbildung D-3: Auswahl einer Startzelle nach Analyse der Eckenbelegung aller Zellen. Links: Tatsächlich durchlaufene Zellen. Rechts: Zellen, die die notwendige Bedingung für eine durchlaufene Zelle erfüllen.

Es gelten die folgenden Aussagen:

- Die beiden Zellen oberhalb und rechts der Zelle R_0 können keine durchlaufenen Zellen sein.
- Unter der Voraussetzung, daß die Zelle R_0 eine tatsächlich durchlaufene Zelle ist, bei der drei Ecken belegt sind, liegt die Kantentopologie eindeutig fest, da es nur zwei Zellenseiten gibt, die zwei belegte Endpunkte haben.
- Ist R_0 eine tatsächlich durchlaufene Zelle ist, bei der vier Ecken belegt sind, enthält die Zelle einen Endpunkt der Gitterschnittfolge auf der rechten oder auf der oberen Zellenseite, und die Kantentopologie ist i. allg. über die Analyse der unteren und linken Nachbarzellen entscheidbar.

Von speziellen Sonderfällen abgesehen ist R_0 eine geeignete Startzelle, wenn es eine tatsächlich durchlaufene Zelle ist. Um die Zelle R_0 als allgemeine Startzelle für einen Verfolgungsalgorithmus zu nutzen, muß geklärt werden, unter welchen Bedingungen R_0 keine Startzelle sein kann, weil es keine durchlaufene Zelle ist. Die Zelle R_0 erfüllt zunächst die notwendige Bedingung für eine durchlaufene Zelle, d.h. es sind die lokalen Topologien in Abb. D-4 möglich.

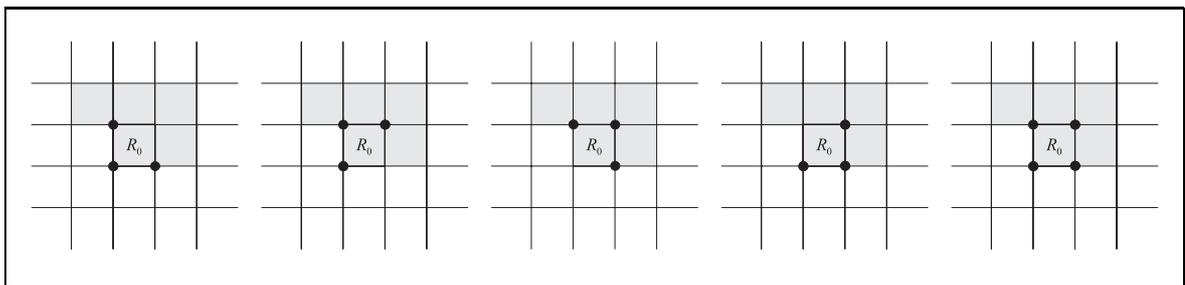


Abbildung D-4: Mögliche Eckenbelegungen bei einer gefundenen, nicht durchlaufenen (Start-) Zelle, die die notwendige Bedingung für eine durchlaufene Zelle erfüllt.

Unter der Voraussetzung, daß R_0 keine durchlaufene Zelle ist, ist nur die Topologie a) möglich, denn die rechte obere Ecke kann keine extrapolierte Höhe enthalten, da keine der angrenzenden Zellenseiten einen Gitterschnitt enthalten kann. Damit ist nur noch diese Topologie zu untersuchen (Abb. D-5). Die grau unterlegten Nachbarzellen sind unter den getroffenen Voraussetzungen keine durchlaufenen Zellen.

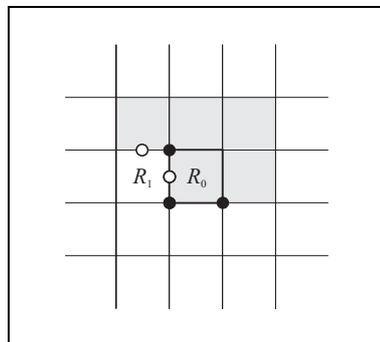


Abbildung D-5: Mögliche Gitterschnitt-Topologien bei einer nicht durchlaufenen Startzelle.

Die Belegung der linken oberen Ecke signalisiert, daß wenigstens auf einer der beiden links und unten angrenzenden Zellenseiten ein Gitterschnitt liegt, der nur ein Endpunkt der Gitterschnittfolge sein kann. Daraus folgt, daß die Zelle R_1 eine durchlaufene Zelle und gleichzeitig

eine Endzelle der Folge der durchlaufenen Zellen sein muß, wenn die Zelle R_0 keine durchlaufene Zelle ist. Umgekehrt gilt damit:

- Ist die linke Nachbarzelle R_1 keine Endzelle, so ist R_0 eine durchlaufene Zelle.

Dieser Fall kann damit als allgemeiner Fall angesehen und R_0 als Startzelle für eine Verfolgungsstrategie genommen werden. Der andere Fall, die linke Zelle R_1 ist eine Endzelle, muß als Spezialfall gesondert untersucht werden.

Für die Stabilität des Algorithmus muß noch eine weitere Voraussetzung getroffen werden, die bei den praktisch vorkommenden Kantenverläufen i. allg. erfüllt sein dürfte:

- Sind zwei benachbarte Raster durchlaufene Raster, so werden sie von einer Rasterschnittfolge auch unmittelbar nacheinander durchlaufen.

Das bedeutet praktisch, daß es keine zu dicht liegende parallele Kantenverläufe und keine zu großen Krümmungen gibt.

Der Verfolgungsalgorithmus, der hier nicht weiter im Detail beschrieben werden soll, analysiert ausgehend von einem der beiden Gitterschnitte auf einer Zellenseite der Startzelle die mögliche Fortschreiterichtung in der jeweiligen Nachbarzelle. Da die Startzelle i. allg. keine Endzelle der Gitterschnittfolge ist, läuft der Algorithmus in einem ersten Schritt zu einem Ende der Gitterschnittfolge, und von da in einem zweiten Schritt komplett in der Gegenrichtung zum anderen Ende. Im zweiten Schritt werden die ermittelten Topologieinformationen gespeichert.

Anhang E

Testrechnung zur Qualität der Lagekorrektur

Die Leistungsfähigkeit der in dieser Arbeit hergeleiteten Korrekturformeln für die Waveletkoeffizienten zur Approximation der Ableitungen einer Funktion von zwei Veränderlichen soll in diesem Abschnitt an einer möglichst allgemeinen Funktion demonstriert werden. Dabei stehen vor allem zwei Eigenschaften im Vordergrund. Zum einen sollten wegen der Frequenzabhängigkeit der Phasenverschiebung die Frequenzen über beide Argumente variieren. Zum anderen sollten, um die unterschiedlichen Verschiebungen in x - und y - Richtung im Zusammenhang zu testen, die 1. und 2. Ableitungen der zu wählenden Funktion noch von x und y abhängig sein. Als Funktion, die diese Eigenschaften besitzt, wurde

$$z = f(x, y) = \sin(x + \sin xy) \quad (5.0-1)$$

gewählt.

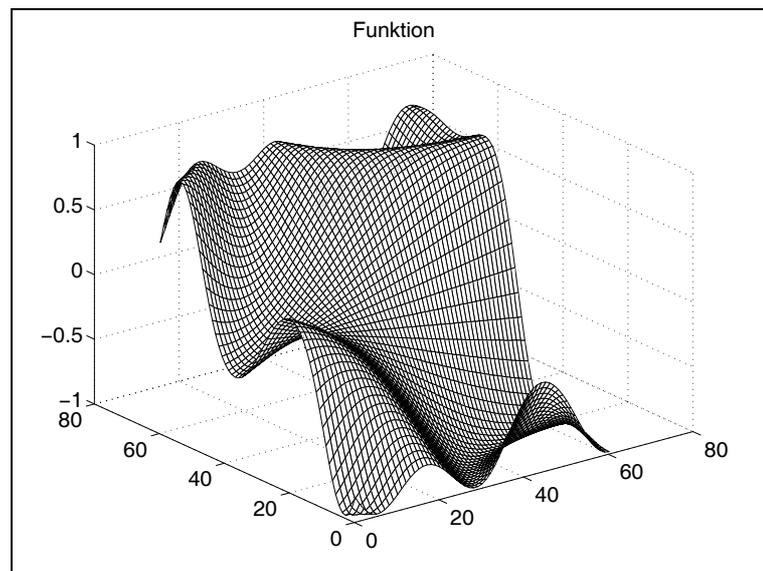


Abbildung E-1: Das synthetische Signal $z = \sin(x + \sin xy)$

Bei dem Test wird die Wavelettransformation mit dem Tensorproduktwavelet HT des Daubechies4-Wavelets betrachtet. Deren Waveletkoeffizienten approximieren die zweite Ableitung nach x . Die 1. und 2. Ableitungen der Funktion (5.0-1) nach x ergeben sich zu

$$z_x = \cos(x + \sin xy) + y \cdot \cos(x + \sin xy) \cdot \cos xy, \quad (5.0-2)$$

$$\begin{aligned} z_{xx} = & -\sin(x + \sin xy) \\ & -2y \cdot \sin(x + \sin xy) \cdot \cos xy \\ & -y^2 \left\{ \sin(x + \sin xy)(\cos xy)^2 + \cos(x + \sin xy) \cdot \sin xy \right\}. \end{aligned} \quad (5.0-3)$$

Diese diskreten Werte der zweiten Ableitung sind mit den Waveletkoeffizienten der Wavelets HT zu vergleichen.

In Abb. E-1 sind die Funktionswerte der Signalmatrix Z der diskretisierten Funktion $z = f(x, y)$ für den Definitionsbereich $-2 \leq x, y \leq 2$ dargestellt¹.

In Abb. E-2 sind die Werte der Waveletkoeffizienten und die Werte der Ableitungen in den zugehörigen Argumenten gegenübergestellt. Die Werte der Ableitungen wurden zum Vergleich mit den Waveletkoeffizienten in den Argumenten berechnet, für die die Waveletkoeffizienten gemäß Abschnitt 7.2 die Approximation der Ableitung repräsentieren. Die Ortsverschiebung der Waveletkoeffizienten wird damit für den Vergleich indirekt realisiert. Im unteren Teil der Abbildung ist der absolute Fehler als Differenz zu sehen.

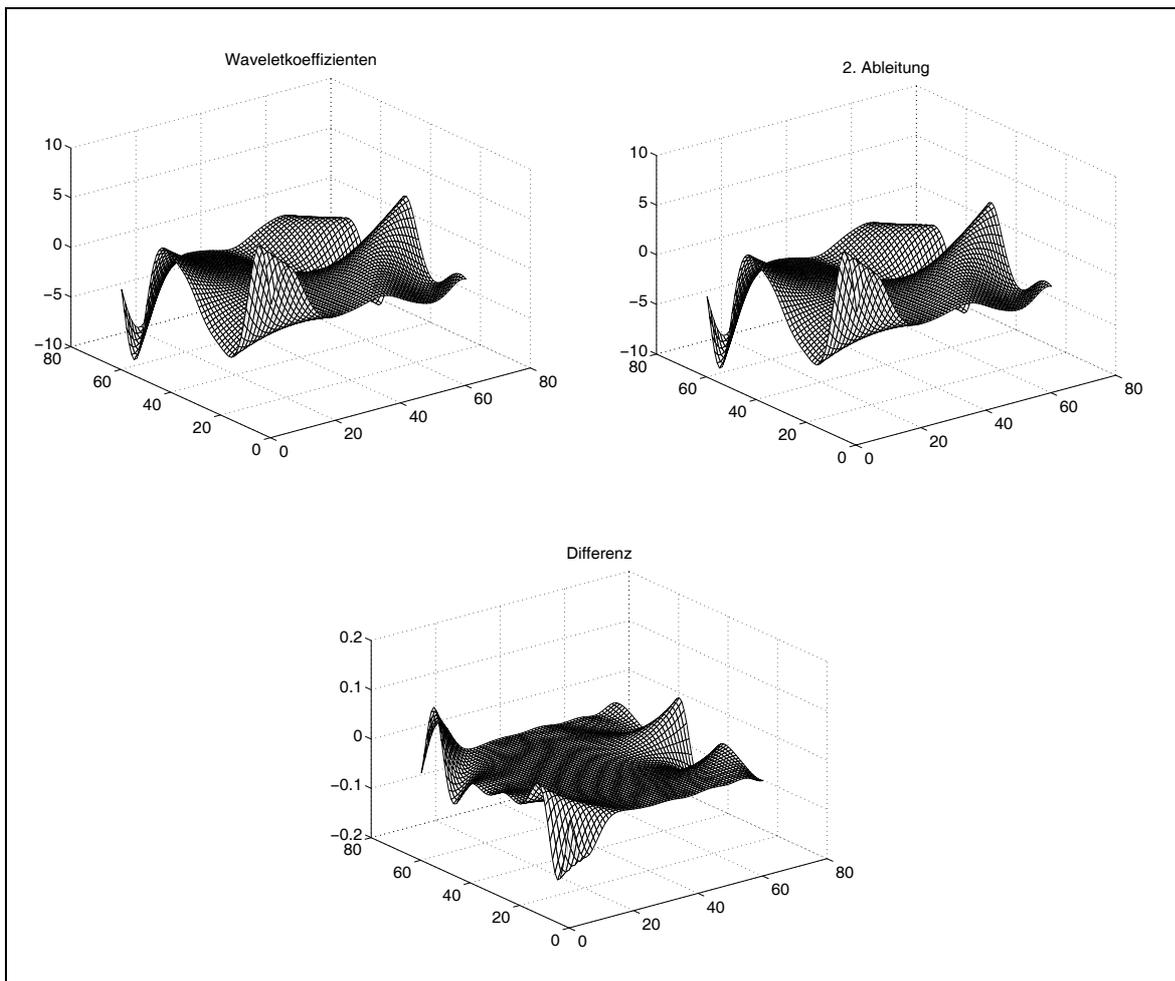


Abbildung E-2: Graphische Darstellung der Matrix der skalierten WK (links oben), der Matrix der zweiten partiellen Ableitung z_{xx} der diskretisierten Funktion $z(x, y)$ in den zugeordneten Argumenten (rechts oben) und der Differenzmatrix (unten).

In Tabelle E-1 sind die Standardabweichung des Signals und die Standardabweichung der Differenz (Fehler) gegenübergestellt. Es ist zu sehen, daß die Standardabweichung des Fehlers durch die Korrektur der Lokalisierung auf $\approx 3,8\%$ des ursprünglichen Wertes zurückgeht.

¹ Wegen der zyklischen Fortsetzung des Signals bei der Berechnung der Waveletkoeffizienten an den Rändern des Signalbereichs entstehen dort verfahrensbedingte Fehler, die in den Abbildungen durch Einschränkung des Indexbereiches der Matrizen ausgeblendet wurden.

Tabelle E-1: Vergleich der Standardabweichungen von Signal und Differenz.

	ohne Korrektur	mit Korrektur
Standardabweichung σ_S des Signals	1.96	1.98
Standardabweichung σ_F des Fehlers	0.45	0.017
Quotient (σ_S/σ_F)	4.35	117.4

Die verbleibenden Fehler sind nicht nur auf die üblichen Diskretisierungsfehler zurückzuführen. Die Frequenzabhängigkeit der notwendigen Korrektur der Lokalisierung, die in den Formeln aus 7.1 und 7.2 nicht berücksichtigt ist, trägt ebenfalls zu einem Restfehler bei. Dieser ist jedoch, wie diese Testrechnung belegt, in einem Gelände mit nicht zu hochfrequenten Anteilen tolerierbar. In Anhang B wurde die Frequenzabhängigkeit für das Daubechies4-Wavelet untersucht. Die in der Fußnote auf S. 114 erwähnten Randeffekte sind vor allem in den höheren Skalen zu berücksichtigen, da sie auf das Übertagen der Wavelets mit großem Support über das Signal zurückzuführen sind.

Anhang F

Das Testgebiet Schneetalpe

Vom Institut für Photogrammetrie und Fernerkundung der Technischen Universität Wien wurde freundlicherweise ein Datensatz eines DGM für Testzwecke zur Verfügung gestellt.

Es handelt sich um das Gebiet der Schneetalpe, ca. 75 km südwestlich von Wien. Das Gebiet hat eine Ausdehnung von rund $10\text{km} \times 10\text{km}$. Die Höhendaten liegen auf einem regulären Gitter mit der Gitterweite 25m. Zusätzlich zu den ca. 171000 Höhenwerten in einer (414,413)-Höhenmatrix sind 290 Geländekanten und 538 Formlinien erfasst. Die Geländekanten und Formlinien liegen in der SCOP-Datenstruktur als 3D-Punktfolgen vor. In Abb. F-1 ist das Testgebiet in einer verkleinerten Kopie (60 %) des SCOP-Originalausdrucks dargestellt. Der angegebene Maßstab gilt im Original.

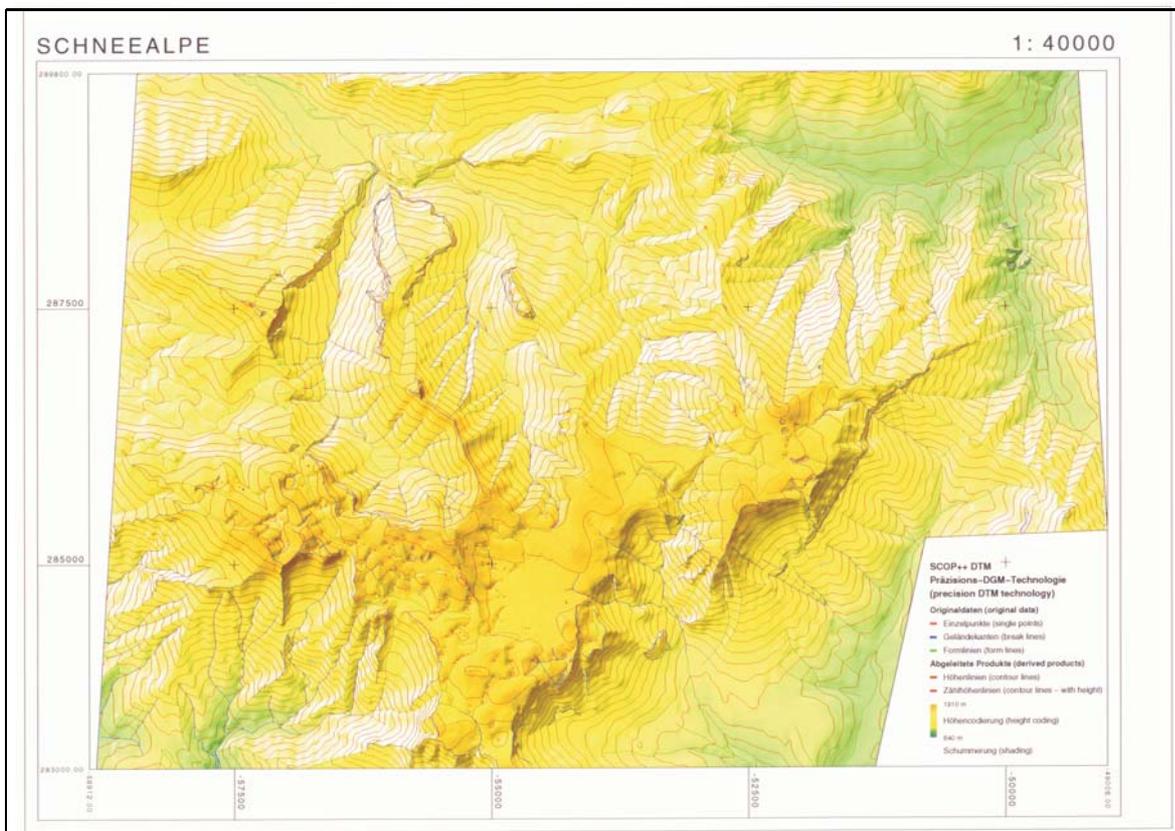


Abbildung F-1: Testgebiet Schneetalpe.

Dank

Die Hauptinhalte der Arbeit entstanden im Rahmen des von der Deutschen Forschungsgemeinschaft geförderten Projekts „Reliefbearbeitung im Waveletbereich“. Dem Projektleiter, Herrn Prof. Dr. S. Meier, danke ich für die wohlwollende Unterstützung und die hervorragende Arbeitsatmosphäre. In diesem Sinne möchte ich mich auch bei allen Kollegen des Instituts für Planetare Geodäsie der TU Dresden für die stete Hilfsbereitschaft bedanken.

Wertvolle Anregungen für das Gesamtprojekt sowie für weiterführende Untersuchungen, die zu der Breite der bearbeiteten Themen führten, bekam der Autor anlässlich mehrerer Forschungsaufenthalte am Institut für Photogrammetrie und Fernerkundung der TU Wien von Herrn Prof. K. Kraus und seinen Mitarbeitern, denen an dieser Stelle ebenfalls herzlich gedankt wird. Besonderer Dank gebührt auch Herrn Prof. W. Keller, der ein sehr kritischer und geduldiger Diskussionspartner für meine aus Zweckmäßigungsgründen etwas praktische Herangehensweise an die Behandlung der mathematischen Theorie war.

