



*Institut für
Neuroinformatik*

*Ruhr-Universität
Bochum*

Internal Report 99-03

**Untersuchungen zur Rekonstruierbarkeit lokaler
Bildmerkmale aus der Gaborwavellettransformierten**

by

Ingo Wundrich

Ruhr-Universität Bochum
Institut für Neuroinformatik
44780 Bochum



IR-INI 99-03
March 1999
ISSN 0943-2752

RUHR-UNIVERSITÄT BOCHUM



Fakultät für Elektrotechnik und Informationstechnik
Lehrstuhl für Signaltheorie
Prof. Dr.-Ing. J. F. Böhme

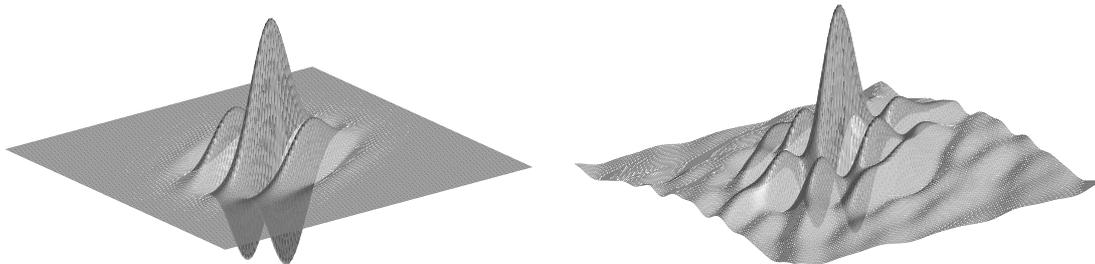
Institut für Neuroinformatik
Lehrstuhl für Systembiophysik
Prof. Dr. C. von der Malsburg

Diplomarbeit

Untersuchungen zur Rekonstruierbarkeit lokaler Bildmerkmale aus der
Gaborwavelettransformierten

von

Ingo Wundrich¹



Betreuer: Dr. Rolf Würtz

¹Institut für Neuroinformatik, Email: wundrich@neuroinformatik.ruhr-uni-bochum.de

Zusammenfassung

In den ersten beiden Teilen werden Grundlagen über zweidimensionale (Gabor)Wavelets und Frames aufgearbeitet, die für die Rekonstruierbarkeitsuntersuchungen unentbehrlich sind, sowie auf den Begriff der Rekonstruierbarkeit selbst eingegangen. Wichtiges Kriterium für die numerische Stabilität der inversen diskreten Gaborwavelettransformation sind die Frameschranken, deren Bedeutung detailliert dargestellt wird. Der dritte Teil handelt von der Rekonstruierbarkeit eines Bildes aus den Amplituden der Gaborwavelettransformierten (komplexe Zellen), wozu zunächst einmal Aussagen über die Eindeutigkeit von Fourieramplituden herangezogen werden müssen. Entsprechende Theoreme werden erläutert und erweitert. Eine weitere Interpretation der diskreten Gaborwavelettransformation ermöglicht die Anwendung dieser Erkenntnisse auf die Antworten komplexer Zellen, die in Theorem 8 formalisiert werden. Zu guter Letzt werden die numerischen Probleme eines darauf basierenden Rekonstruktionsalgorithmus diskutiert.

Inhaltsverzeichnis

I	Grundlagen	11
1	Das visuelle System: Biologie und Modell	12
1.1	Neuronen und rezeptive Felder	12
1.2	Frühe visuelle Verarbeitung	12
1.3	Bildverarbeitung im primären visuellen Kortex	13
1.3.1	Einfache Zellen	14
1.3.2	Komplexe Zellen	16
1.4	Technische Anwendung	16
2	Bildverarbeitung	17
2.1	Kontinuierliche Transformationen	17
2.1.1	Die Fouriertransformation	17
2.1.2	Die Gabortransformation	17
2.2	Diskrete Transformationen	18
2.2.1	Signalabtastung	18
2.2.2	Die zweidimensionale ortsdiskrete Fouriertransformation	19
2.2.3	Die zweidimensionale Z-Transformation	19
2.2.4	Die zweidimensionale diskrete Fouriertransformation	20
2.2.5	Angepaßte Abtastung von Gaborkernen	21
II	Einfache Zellen	23
3	Die kontinuierliche Gaborwavelettransformation	24
3.1	Zweidimensionale kontinuierliche Gaborwavelets	24
3.2	Signalrekonstruktion aus zweidimensionalen Gaborwavelets	25
3.2.1	Die inverse Gaborwavelettransformation	25
3.2.2	Alternative Methoden der Signalrekonstruktion	28
4	Frames	29
4.1	Kontinuierliche Wavelet-Frames	29
4.2	Allgemeine Frames	30
4.2.1	Die Frametransformation als Operator	31
4.2.2	Die Frameschranken	33
4.2.3	Rekonstruktionsmöglichkeiten	35

5	Die phasenraumdiskrete Gaborwavellettransformation	36
5.1	Abtastung des Phasenraums	36
5.2	Abschätzung der Frameschranken	39
6	Die diskrete Gaborwavellettransformation	41
6.1	Abtastung des Ortsraums bei diskretem Phasenraum	41
6.2	Skalenunabhängige Abtastung des Ortsraums und der Translationen	43
6.3	DGWT und IDGWT in Vektornotation	44
6.3.1	Die Inversion des Metrischen Operators	46
6.3.2	Die Singulärwertzerlegung von Ψ	46
6.4	Skalenabhängige Abtastung des Ortsraums und der Translationen	53
6.4.1	Die pyramidale Gaborwavellettransformation	53
6.4.2	Die inverse pyramidale Gaborwavellettransformation	55
6.4.3	DGWT mit Bandpaßabtastung	55
6.5	Lokale Rekonstruktion	55
6.6	Wavelet-Frames und orthogonale Wavelet-Basen	57
6.6.1	Analyse und Synthese: Filterbänke und Multiskalenanalyse	57
6.6.2	Die Wavellettransformierte: übervollständige und kritische Abtastung	58
6.6.3	Das Problem der „Verschiebbarkeit“	58
III	Komplexe Zellen	61
7	Rekonstruierbarkeit lokaler Merkmale aus Gaboramplituden	62
7.1	Fourieramplituden	62
7.1.1	Polynome in einer und mehreren komplexen Variablen	62
7.1.2	Die Hayesschen Theoreme	64
7.2	Die Eindeutigkeit der Jetrekonstruktion	67
7.2.1	Gabor- und Fourieramplituden	67
7.2.2	Anpassung der Abtastrate des Patches an die Bandpaßfrequenz	69
7.3	Bildrekonstruktion aus Gaboramplituden	69
7.3.1	Rekonstruktion des Patches $P_{\kappa_0, \vec{n}_0}(\vec{n})$ aus dem Jet $ \mathcal{J}_{\kappa_0 \vec{e}(\varphi)}^{BP}(\vec{n}_0) $	70
7.3.2	Rekonstruktion des Bildes $I(\vec{n})$ aus den Patches $P_{\kappa_0, \vec{n}_0}^{rek}(\vec{n})$	72
8	Bildrekonstruktion aus Antworten komplexer Zellen	74
8.1	Erweiterung der Hayesschen Theoreme auf komplexwertige Sequenzen	74
8.2	Die Eindeutigkeit der Bildrekonstruktion von Amplituden der DGWT	76
8.3	Ein Rekonstruktionsalgorithmus für die Antworten komplexer Zellen	79
IV	Ausblick und Anhang	83
A	Formeln zu Fouriertransformation und Systemtheorie	84
A.1	FT	84
A.2	DSFT	84
A.3	DFT	84
A.4	Das Abtasttheorem	85
B	Die Wavelletfunktionen $\hat{\chi}_{\cdot, a, \vartheta}[\omega \vec{e}(\vartheta)]$	85
	Literatur	87

Formelzeichen und Abkürzungen

Formelzeichen (1)

\mathbb{R}	Menge der reellen Zahlen
\mathbb{C}	Menge der komplexen Zahlen
\mathbb{Z}	Menge der ganzen Zahlen
i	imaginäre Einheit, $i^2 = -1$
$\vec{x} \in \mathbb{R}^2$	Vektor im kontinuierlichen Ortsraum
$\vec{\omega} \in \mathbb{R}^2$	Vektor im kontinuierlichen Frequenzraum (FT)
$\vec{\nu} \in \mathbb{R}^2$	Vektor im kontinuierlichen und periodischen Frequenzraum (DSFT)
\vec{z}	Frequenzvariable der ZT
$\vec{n} \in \mathbb{Z}^2$	Vektor im abgetasteten Ortsraum (DSFT, DFT)
$\vec{\rho} \in \mathbb{Z}^2$	Vektor im abgetasteten Frequenzraum (DFT)
$\vec{\rho}$	$= \begin{bmatrix} \frac{\rho_1}{N_1} \\ \frac{\rho_2}{N_2} \end{bmatrix}$
Δ, Δ_0	Abtastintervall im Ortsraum (für beiden Dimensionen gleich). Δ und Δ_0 werden in dieser Arbeit nicht als Präfix benutzt
ω_A	Abtastfrequenz (für beide Dimensionen)
$\hat{\Delta}$	Abtastintervall im Frequenzraum
f	Funktion (im Ortsraum)
\hat{f}	Funktion im Frequenzraum (FT)
\tilde{f}	Funktion im Frequenzraum (DSFT)
\check{f}	Funktion im Frequenzraum (DFT)
N_1, N_2	Anzahl der Pixel in der jeweiligen Richtung
$\psi_{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}$	Gabormutterwavelet (INI)
\vec{k}	Vektor mit Skalen und Orientierungen für Gaborwavelets (INI)
$\psi_{\vec{k}}$	Gaborwavelet ohne Translationen (INI)
χ	Gabormutterwavelet (allg.)
a	Skalierungsfaktor eines Gaborwavelets (allg.)
ϑ	Orientierungswinkel eines Gaborwavelets (allg.)
$\vec{x}_0 \in \mathbb{R}^2$	Vektor im kontinuierlichen Translationsraum (GWT)
$\chi_{\vec{x}_0, a, \vartheta}$	Familie von Gaborwavelets (allg.)
$\vec{n}_0 \in \mathbb{Z}^2$	Vektor im abgetasteten Translationsraum (DPSGWT, DGWT)
M	Anzahl der Skalen bei DPSGWT und DGWT
L	Anzahl der Orientierungen bei DPSGWT und DGWT
$\tilde{\chi}_l$	skalenunabhängige Gaborwavelets für die pyramidale Transformation
I	aufgenommenes Bildsignal
$\mathcal{J}_{\vec{k}}$	Gaborwavelettransformierte (INI)
\mathcal{I}	kontinuierliche oder pyramidale Gaborwavelettransformierte
\mathcal{I}	skalenunabhängig abgetastete diskrete Gaborwavelettransformierte
A, B	allg. Frameschranken
$A_{a_0, L}, B_{a_0, L}$	Frameschranken für DPSGWT und DGWT
Y	Metrische (Diagonal-) Operator für IGWT
$Y_{a_0, L}$	Metrische (Diagonal-) Operator für IDPSGWT und IDGWT
\mathcal{F}	Frame-transformierte von f
T	Frameoperator
T^+	adjungierter Operator zu T
T^\dagger	verallgemeinerte Inverse (Moore-Penrose-Inverse) zu T
G	Metrischer Operator
$h_{\vec{m}}$	Frame-Funktion
$h^{\vec{m}}$	zu $h_{\vec{m}}$ duale Frame-Funktion

Formelzeichen (2)

Ψ, \mathbf{X}	Wavelet-Frame-Matrix
\underline{I}	Vektor der Bilddaten
$\underline{\mathcal{I}}, \underline{\mathcal{I}}_0$	Vektor der gabortransformierten Bilddaten
$\underline{J}(\vec{n}_0)$	Jet an der Stelle \vec{n}_0
\mathcal{H}	Hilbertraum
$L^2(\mathbb{R}^2)$	Menge der energiebeschränkten Signale in \mathbb{R}^2
$\vec{e}(\varphi)$	$= \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix}$
\vec{e}_1, \vec{e}_2	Einheitsvektoren für die entsprechenden Dimensionen
Q	Drehmatrix
\mathcal{N}	eindimensionales Nyquistintervall
\mathcal{B}	Träger eines Bildsignals $[0; N_1 - 1] \times [0; N_2 - 1]$
$\vec{\mathcal{B}}$	Träger eines Bildsignals im diskreten Frequenzraum, so daß $\vec{\rho} = \vec{0}$ möglichst in der Mitte liegt
\mathcal{R}_j	Fläche des rezeptiven Feldes des Neurons j
Δ	Abtastintervall im Ortsraum
$\hat{\Delta}$	Abtastintervall im Frequenzraum
\vec{k}	Mittelfrequenz des Gaborwavelets $\hat{\psi}_{\vec{k}}(\vec{\omega})$ im kontinuierlichen Frequenzraum
$\vec{\kappa}$	Mittelfrequenz des Gaborwavelets $\check{\psi}_{\vec{\kappa}}(\vec{\nu})$ im kontinuierlichen Frequenzraum der DSFT
$\vec{\eta}$	Mittelfrequenz des Gaborwavelets $\check{\psi}_{\vec{\eta}}(\vec{\rho})$ im diskreten Frequenzraum der DFT für $N_1 = N_2$
a_0	Abstand zweier benachbarter Skalen (Quotient) bei DPSGWT und DGWT
$\mathcal{P}(N, K)$	Menge aller Polynome von Gesamtgrad N und Dimension K
$\mathcal{P}_z(N_1, N_2)$	Menge aller Signale, die auf dem endlichen Träger $0, \dots, N_j - 1, j = 1, 2$ definiert sind.
P_{k_0, \vec{x}_0}	Patch

Abkürzungen

AAF	Anti-Aliasing-Filter
AIF	Anti-Imaging-Filter
BP	Bandpaß
DFT	Diskrete Fouriertransformation
DGWT	Diskrete Gaborwavelettransformation
DPSGWT	Phasenraumdiskrete Gaborwavelettransformation
DSFT	Ortsdiskrete Fouriertransformation
FFT	Schnelle Fouriertransformation
FT	Fouriertransformation
GWT	Gaborwavelettransformation
IDFT	Inverse diskrete Fouriertransformation
IDGWT	Inverse diskrete Gaborwavelettransformation
IDPSGWT	Inverse phasenraumdiskrete Gaborwavelettransformation
IDSFT	Inverse ortsdiskrete Fouriertransformation
IFFT	Inverse Schnelle Fouriertransformation
IFT	Inverse Fouriertransformation
IGWT	Inverse Gaborwavelettransformation
INI	Institut für Neuroinformatik, Ruhr-Universität Bochum
IPT	Inverse Patchtransformation
LSI	linear und verschiebungsinvariant
MSA	Multiskalenanalyse
PT	Patchtransformation
TP	Tiefpaß
ZT	z-Transformation

Einleitung

Thema und Arbeitsgebiet

Das Thema dieser Diplomarbeit ist die Kodierung reeller Bildsignale mit einfachen und komplexen Zellen und die *Rekonstruierbarkeit* des ursprünglichen Signals aus dem transformierten. Dabei kann „Rekonstruierbarkeit“ folgendes bedeuten:

- Im strengen mathematischen Sinn reduziert sich „Rekonstruierbarkeit“ auf das Problem, ob zu der Abbildung

$$G : \mathcal{A} \rightarrow \mathcal{C}$$

eine Umkehrung existiert, wobei \mathcal{A} das Eingangsbildsignal und \mathcal{C} eine Transformierte enthalten kann. Ist die Darstellung durch ein Element aus \mathcal{C} (über)vollständig, ist G *injektiv*, wie es auf die sog. Frametransformation zutrifft. In diesem Fall sind bei einer Rekonstruktion unendlich viele Linearkombinationen der Framefunktionen möglich, um wieder das eine richtige Element aus \mathcal{A} zu bestimmen. Dazu verwendet man einen allgemeineren Begriff der Inversen von G .

- In einem eher praktischen Sinne bedeutet Rekonstruierbarkeit, ob man einen Algorithmus angeben kann, der numerisch stabil einem Element aus \mathcal{C} genau eines aus \mathcal{A} zuordnet. Die numerische Stabilität ist ein wichtiges Kriterium für die Auswahl einer Linearkombination aus den im vorigen Punkt erwähnten vielen Möglichkeiten zu treffen. Dies geschieht bei der inversen Frametransformation.
- Interessant wird der Begriff „Rekonstruierbarkeit“ erst bei komplexen Zellen, da in diesem Fall hinter dem injektiven ein nichtlinearer und nichtinjektiver Operator H geschaltet ist:

$$H : \mathcal{C} \rightarrow \mathcal{D}, \quad HG : \mathcal{A} \rightarrow \mathcal{D}.$$

Hierbei stellt sich die Frage ob der Operator HG injektiv ist, d.h. ob trotz der Probleme, die H bereitet, das richtige Element aus \mathcal{A} wiedergefunden werden kann.

Die Inhalte dieser Arbeit sind – wie das gesamte Gebiet der Neuroinformatik – interdisziplinär ausgerichtet, wobei der Schwerpunkt auf Aspekten der Signalverarbeitung liegt. Es werden aber auch anwendungsrelevante mathematische und biologische Zusammenhänge behandelt, wobei die letzteren für die Modellbildung stark schematisiert werden müssen. Diese Interdisziplinarität ist für das Gebiet des *Künstlichen Sehens (Computer Vision)* notwendig, da das menschliche visuelle System eine Vielzahl von extrem leistungsfähigen Strategien zur Signal- und Informationsverarbeitung entwickelt hat, und die Erkenntnisse darüber wertvolle Beiträge zur Entwicklung von Systemen zur vollautomatischen Bildauswertung liefern. Ein Bauteil aus dieser Maschinerie sind die komplexen Zellen, deren Funktionalität neben einer speziellen neuronalen Netzwerkarchitektur am Lehrstuhl für Systembiophysik zur technischen Umsetzung in einem System zur robusten Erkennung von Gesichtern geführt hat, s. z.B. (Lades et al., 1993). Die Anwendungsmöglichkeiten sind aber vielfältig.

Geleitet wurde die Aufgabestellung für diese Diplomarbeit von der Frage, ob komplexe Zellen in der Lage sind, reelle Bildsignale verlustfrei zu kodieren, wie es die einfachen tun. Um zur Beantwortung dieser Frage im Teil III Eindeutigkeitsbeweise führen und über Rekonstruktionsalgorithmen nachdenken zu können, unterliegt die gesamte Arbeit einer detaillierten mathematischen Darstellung der Sachverhalte, die auch von Lesern, die mit der Materie weniger vertraut sind, ohne zusätzliches Literaturstudium verstanden werden können. Einfache und wichtige Beweise sind der Literatur entnommen und aufbereitet worden, sofern sie Einblicke in die mathematischen Zusammenhänge gewähren. Eine oberflächlicherer Stil wäre weniger zugänglich und für den Algorithmenentwurf nicht sehr hilfreich.

Das Gebiet der Wavelettransformation ist noch relativ jung und in Ausbreitung in verschiedenste Zweige von Naturwissenschaft und Technik begriffen. Deshalb ist deutsche Fachliteratur sehr selten. Es wurden daher viele englische Fachausdrücke belassen und nur in wenigen Fällen übersetzt. Die deutschen wavelet- und frametheoretischen Begriffe sind (Louis et al., 1994) entnommen.

Da die Fachterminologie nicht standardisiert ist, sind einige in dieser Arbeit verwendete Begriffe in der Literatur unter einem anderen Namen zu finden.

Der Begriff der Phase ist in dieser Arbeit allgemein zu sehen. Er bezieht sich nicht nur auf die Polarwinkel der Fouriertransformierten in der komplexen Zahlenebene sondern auf die sämtlicher komplexwertiger Funktionen unabhängig davon, ob sie den Orts- oder den Frequenzraum beschreiben. Vor allem bei der DGWT ist die Rekonstruktion, insbesondere die des Mittelwertes, nicht exakt, weshalb das „=-Zeichen aus dem Kontext heraus zu verstehen ist.

Der Text ist in $\text{\LaTeX} 2_{\epsilon}$ geschrieben. Die Zeichnungen sind mit Xfig entstanden und die Daten und Funktionen sind von Programmen in C mit Unterstützung von (Press et al., 1992) berechnet und mit MATLAB visualisiert worden.

Gliederung

Diese Arbeit besteht aus drei Hauptteilen und dem Anhang mit Zusammenfassung. Im einzelnen sind dies folgende:

Im Teil I werden ein paar Grundlagen über kontinuierliche und diskrete Darstellungen der Fouriertransformation zusammengetragen, wobei die breite Darstellung solcher sehr elementarer Sachverhalte gleichzeitig eine Einführung in die hier durchgängig verwendete Nomenklatur darstellt. Zunächst wird aber kurz auf die biologischen Motivationen, vor allem der Gaborfunktionen, dieser Arbeit eingegangen.

Teil II führt darauf in die zweidimensionale Wavelettransformation ein, wobei natürlich die Gaborwavelets im Vordergrund stehen, was bedeutet, daß nicht alle Betrachtungen auf andere Waveletfamilien übertragen werden können. Diese Art der Bildsignaldarstellung entspricht der Kodierung durch sogenannte einfache Zellen im primären visuellen Kortex. Die sich daran anschließenden Abschnitte beschäftigen sich mit der Rekonstruktion aus Gaborantworten bei abzählbar und schließlich endlich vielen Kernen. Ein zentraler Begriff für diese Untersuchungen ist der des Frames, der definiert und auf die Gaborwavelettransformation angewandt wird. Das Ergebnis sind numerische Verfahren zur Analyse und Synthese. Ein Vergleich zwischen der behandelten und der Multiskalenanalyse mit orthogonalen Wavelet-Basen schließt diesen Teil ab.

Teil III befaßt sich mit der Bildkodierung durch komplexe Zellen, wobei die Eindeutigkeit einer Rekonstruktion aus den Amplituden der Gaborwavelettransformierten im Vordergrund steht. Dieser Teil besteht aus zwei Hälften. In der ersten geht es um Rekonstruktionen aus einem oder wenigen Jets, d.h. aus Amplituden lokaler Gaborantworten, in der zweiten um die Wiedergewinnung des gesamten Bildsignals. Entsprechende Algorithmen sind vorgeschlagen, aber nicht implementiert worden.

Danksagung

Mein Dank gilt Prof. Dr. C. von der Malsburg, der mir diese interessante und interdisziplinäre Diplomarbeit am Lehrstuhl für Systembiophysik gestellt hat, und Prof. Dr.-Ing. J. F. Böhme (Lehrstuhl für Signaltheorie), der mir als Prüfer der Fakultät für Elektrotechnik und Informationstechnik die Durchführung der Arbeit am Institut für Neuroinformatik ermöglicht hat.

Ebenfalls möchte ich mich bei Dr. R. Würtz, meinem Betreuer, bedanken, der jederzeit für Anregungen und Diskussionen zur Verfügung stand und somit wesentlich zum Gelingen dieser Arbeit beigetragen hat.

Ich bedanke mich auch bei unserem Systemadministrator Michael Neef für die Bereitstellung zuverlässiger Hard- und Software sowie die geduldige Beratung.

Ingo Wundrich, 26. Oktober 1998

Teil I

Grundlagen

- Das Konzept des rezeptiven Feldes
- Kurze Beschreibung der unteren Verarbeitungsstufen des visuellen Systems
- Mathematische Modellierung einfacher und komplexer Zellen
- Festlegung der Notation für die FT
- Definition von DSFT, DFT und ZT

1 Das visuelle System: Biologie und Modell

1.1 Neuronen und rezeptive Felder

Im Mittelpunkt dieser Arbeit stehen Modelle für rezeptive Felder im primären visuellen Kortex, deren biologische Motivation im folgenden kurz zusammengefaßt werden soll.

Das *rezeptive Feld* eines Neurons ist definiert als die Menge aller Punkte in der Sensorschicht (beim Sehen die *Retina*), die für ebendieses Neuron Systemeingaben liefern.

Das Modellneuron aus Abb. 1 bekommt Eingaben aus seinem rezeptiven Feld. Diese werden summiert oder – bei kontinuierlicher Beschreibung – integriert und ausgegeben,

$$\mathcal{I}_j = \int_{\mathcal{R}_j} d^2x I(\vec{x})h_j(\vec{x}), \quad (1.1-1)$$

wodurch ein solches Neuron mathematisch durch ein Funktional, im linearen Fall als Skalarprodukt, charakterisiert wird. Dies entspricht auch seinem groben Aufbau: Signale werden von außen über die Synapsen an den Dendriten elektrisch und chemisch in den *Soma*, den eigentlichen Zellkörper, eingespeist, zur Verarbeitung weitergeleitet und über das Axon an eine Vielzahl weiterer Neuronen übertragen. Die Summe bzw. das Integral kann dabei noch einer nichtlinearen Transformation unterworfen werden, was bei den Modellneuron aus Abb. 1 durch das Kästchen hinter dem Summierer symbolisiert ist. In den in diesem Abschnitt benutzten Formeln ist dies nicht berücksichtigt. Es gibt zwar eine Vielzahl verschiedener Zellen, die optische Reize von der Retina erhalten, es soll hier aber vorerst nur von einem Zelltyp die Rede sein. Haben diese Neuronen das gleiche Antwortverhalten, so unterscheiden sie sich nur durch ihre Position \vec{x}_0 auf der Retina. Nimmt man ein Kontinuum von rezeptiven Feldern an, so wird aus (1.1-1)

$$\mathcal{I}(\vec{x}_0) = \int_{\mathcal{R}_{\vec{x}_0}} d^2x I(\vec{x})h_{\vec{x}_0}(\vec{x}) = \int_{\mathbb{R}^2} d^2x I(\vec{x})h(\vec{x} - \vec{x}_0) = I(\vec{x}_0) * h(-\vec{x}_0), \quad (1.1-2)$$

wobei vorausgesetzt wird, daß $h(\vec{x})$ über einen endlichen Träger verfügt oder einen schnellen Abfall hat. Des weiteren ist aus (1.1-2) ersichtlich, daß die Neuronenschicht der $h_{\vec{x}_0}$ als LSI-System wirkt. Diese Betrachtungen sind nur in frühen Stufen visueller Verarbeitung sinnvoll, in denen die Darstellung der aufgenommenen Bilddaten *retinotop* ist. Das bedeutet, daß die Lichtreize von der Retina nachbarschaftserhaltend in weiteren Verarbeitungsstufen kodiert sind.

Die einen Typ rezeptiver Felder charakterisierende Funktion $h(\vec{x})$ kann sowohl positive als auch negative Werte annehmen. Wird das rezeptive Feld an einer Stelle $\vec{x}_1 \in \mathcal{R}_j$ mit $h(\vec{x}_1) > 0$ stimuliert, so erhöht sich die Feuerungsrate (Frequenz eines Spannungspulses über dem Axon) von einer Ruheaktivität ausgehend. Im entgegengesetzten Fall wird die Feuerungsrate gesenkt. Daher spricht man von einem *exzitatorischen* bzw. von einem *inhibitorischen* Stimulus des angeschlossenen Neurons. Diese durch Messungen an Versuchstieren belegten Vorgänge motivieren die Interpretation der Ausgabe \mathcal{I}_j der j -ten Zelle als mittlere Feuerungsrate („*integrate-and-fire neuron*“).

1.2 Frühe visuelle Verarbeitung

Der Vorgang des Sehens beginnt auf der *Retina*, wo die Bildaufnahme erfolgt und die neuronale Verarbeitung beginnt. Die Lichtreize werden dabei von sog. *Photorezeptoren* in elektrische und chemische Impulse umgewandelt. Dabei sind die *Zapfen* für das Tag- und Farbsehen, die *Stäbchen* für das Sehen bei geringen Beleuchtungsintensitäten verantwortlich. Die Aufgabe der Netzhaut beschränkt sich dabei nicht nur auf bloße Bildakquisition, sondern es finden auch schon die ersten Verarbeitungsprozesse statt. Beispielsweise haben Ganglien- und Bipolarzellen rezeptive Felder, die nach dem sog. *Center-Surround-Mechanismus* funktionieren, was bedeutet, daß $h(\vec{x})$ punktsymmetrisch mit $h(\vec{x}) > 0$ in der Mitte („*On-Center-Zelle*“) und $h(\vec{x}) < 0$ in der Umgebung ist. Es gibt auch den entgegengesetzten Fall der „*Off-Center-Zelle*“. Diesen Zellen wird die Funktion der lokalen Kontrastmessung zugeschrieben, wodurch unter anderem die Pupillenapertur geregelt wird.

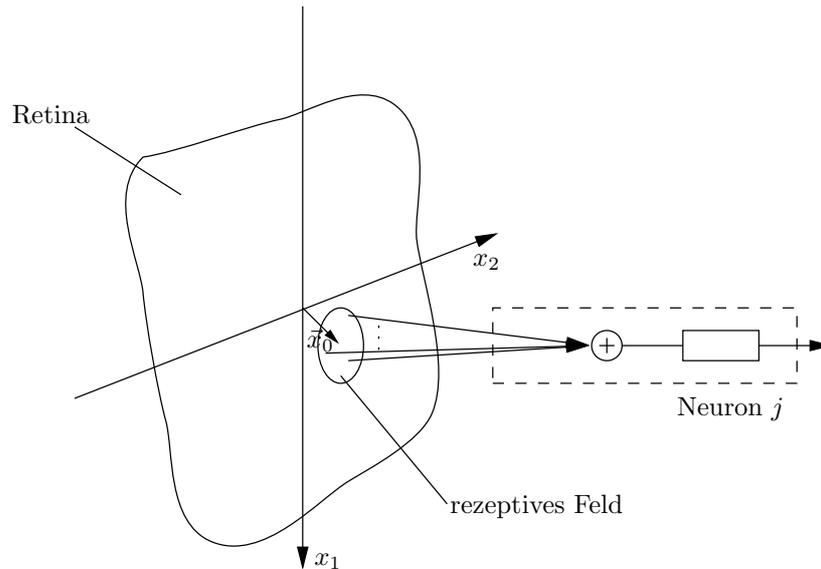


Abbildung 1: Einfaches Modell für das receptive Feld eines Neurons.

Von der Retina ausgehend bilden die Ganglienzellen den Sehnerv, dessen Axone sich kurz nach dem Eintritt in die Schädelkapsel im *Chiasma opticum* überkreuzen, und zwar derart, daß die rechte Hälfte der Retina, auf die die linke Welthälfte abgebildet wird, mit der linken Gehirnhälfte verschaltet ist. Der Sehnerv mündet in das *Corpus geniculatum laterale* (auch *LGN – Lateral geniculate nucleus* oder *seitlicher Kniehöcker* genannt), eine Durchgangsstation für die Information von der Retina zur Hirnrinde, (Dudel et al., 1996), (Würtz, 1995). Den groben Aufbau des visuellen Systems zeigt Abb. 2.

1.3 Bildverarbeitung im primären visuellen Kortex

Die nächste Verarbeitungsstufe nach dem LGN ist der *primäre visuelle Kortex*, der bei Affen und Menschen auch *V1* genannt wird.

Die Großhirnrinde, auch cerebraler Kortex genannt, besteht aus einer stark gefalteten, etwa 2 mm bis 4 mm dicken Schicht aus Nervengewebe und bildet die äußere Hülle der beiden Gehirnhälften.

In V1 lassen sich sechs Schichten unterscheiden, wobei die in weitere Unterschichten einteilbare Schicht IVc am auffälligsten ist. Sie ist für V1 charakteristisch und bei ihr findet die Eingabe visueller Informationen in den primären visuellen Kortex statt. Senkrecht zu den Schichtbegrenzungen existiert noch eine Unterteilung in Säulen, die allerdings nicht durch alle Schichten verlaufen. In der bereits erwähnten Schicht IVc von V1 läßt sich die getrennte Informationsverarbeitung aus dem rechten und linken Auge in diskreten anatomischen Bereichen, den Augendominanzbändern, nachweisen (Pöttsch, 1994), (Dudel et al., 1996).

Hier existiert eine Vielzahl verschiedenartiger Zellen, von denen mehrere Typen in Versuchen an Katzen von Hubel und Wiesel (Hubel und Wiesel, 1962) beschrieben wurden. Dabei werden in V1 grob folgende Zelltypen unterschieden:

- *Center-Surround-Zellen*
- *Einfache Zellen*
- *Komplexe Zellen*

Center-Surround-Zellen (s.o.) wurden in der Retina, im LGN und in V1 gefunden. Die rezeptiven

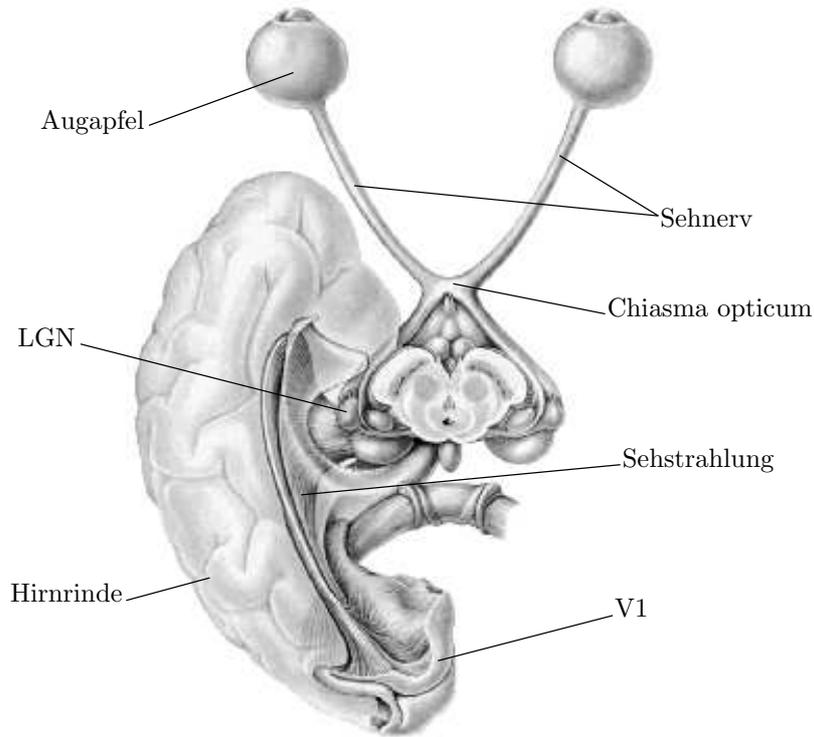


Abbildung 2: Das visuelle System. Aus (Hirsch, 1998).

Felder werden nach (Würtz, 1995) als

$$h(\vec{x}) = \sigma_-^2 e^{-\frac{\|\vec{x}\|^2}{2\sigma_+^2}} - \sigma_+^2 e^{-\frac{\|\vec{x}\|^2}{2\sigma_-^2}} \quad (1.3-1)$$

modelliert, wobei der Fall $\sigma_- > \sigma_+$ für eine On-center-off-surround-Zelle steht.

1.3.1 Einfache Zellen

Nach (Hubel und Wiesel, 1962) sind die rezeptiven Felder *einfache Zellen* bezüglich ihres exzitatorischen oder inhibitorischen Verhaltens nicht rotationssymmetrisch, sondern die Bereiche sind vielmehr nebeneinander angeordnet, wobei die Verteilung dieser länglichen Bereiche und deren Orientierung unterschiedlich ausfallen. Einige von ihnen sind symmetrisch, andere antisymmetrisch bezüglich einer Achse. Die einfachste Form, diese Ergebnisse in eine mathematische Beschreibung zu übertragen, ist

$$h_{\vec{k}}^g(\vec{x}) = e^{-\frac{\vec{x}^T D \vec{x}}{2\sigma^2}} \cos \vec{k}^T \vec{x} \quad (1.3-2)$$

für rezeptive Felder gerader Symmetrie und

$$h_{\vec{k}}^u(\vec{x}) = e^{-\frac{\vec{x}^T D \vec{x}}{2\sigma^2}} \sin \vec{k}^T \vec{x} \quad (1.3-3)$$

für die ungerader Symmetrie, wobei die Orientierung des rezeptiven Feldes mit \vec{k} gesteuert werden kann. D ist eine Diagonalmatrix, die dem rezeptiven Feld eine elliptische Form ermöglicht. Diese beiden Punktantworten werden nun zu der Punktantwort eines Zweikanalfilters zusammengefaßt,

$$g_{\vec{k}}(\vec{x}) = h_{\vec{k}}^g(\vec{x}) + i h_{\vec{k}}^u(\vec{x}) = e^{-\frac{\vec{x}^T D \vec{x}}{2\sigma^2}} e^{i\vec{k}^T \vec{x}} = g(\vec{x}) e^{i\vec{k}^T \vec{x}}, \quad (1.3-4)$$

wobei Real- und Imaginärteil jeweils einen Kanal repräsentieren. Versuche von Pollen und Ronner (Pollen und Ronner, 1981) zeigen, daß benachbarte einfache Zellen ein Quadraturpaar wie in (1.3-2) und (1.3-3) bilden, wobei wiederum eine Zelle die negativen und eine andere Zelle die positiven Anteile des rezeptiven Feldes ausgeben. Die Zelle mit dem rezeptiven Feld (1.3-4) besteht also eigentlich aus mindestens vier Neuronen in V1. Benutzt man (1.1-2), so liefert eine Schicht nach (1.3-4) modifizierter einfacher Zellen die Ausgabe

$$\mathcal{I}_g(\vec{x}_0, \vec{k}) = \int_{\mathbb{R}^2} d^2x I(\vec{x}) g_{\vec{k}}(\vec{x} - \vec{x}_0) = \int_{\mathbb{R}^2} d^2x I(\vec{x}) g_{\vec{k}}^*(\vec{x}_0 - \vec{x}) = \langle I, g_{\vec{k}}(\vec{x}_0 - \cdot) \rangle. \quad (1.3-5)$$

Die Funktionen (1.3-4) wurden eindimensional von D. Gabor zur Kodierung und Übertragung von Signalen in der Nachrichtentechnik vorgeschlagen (Gabor, 1946). Sie heißen deshalb *Gaborfunktionen*, die Vorschrift (1.3-5) *Gabortransformation*. Ein System von Gaborfunktionen, auch *Weyl-Heisenberg-System* genannt, ist, wenn \vec{x}_0 und \vec{k} diskretisiert sind, definiert als

$$g_{\vec{x}_0, \vec{k}}(\vec{x}) = g(\vec{x} - \vec{x}_0) e^{i\vec{k}^T \vec{x}}. \quad (1.3-6)$$

$g(\vec{x})$ braucht dabei keine Gaußfunktion zu sein, sondern ein Fenster, das das Eingabesignal in Orts- und Frequenzraum genügend lokalisiert. Ist dies der Fall, können Gaborfunktionen für eine *Zeit-Frequenzanalyse* herangezogen werden. Ist $g(\vec{x})$ eine Gaußfunktion, kann gezeigt werden, daß die entsprechenden Gaborfunktionen die beste Lokalisierung in Orts- und Frequenzraum haben (Würtz, 1995). Dies ist in der *Computer Vision* besonders wichtig, da die zu erkennenden Objekte unter anderem gedreht und verschoben wieder auftauchen können. Die Extraktion von Antworten von skalierten und gedrehten Gaborfunktionen (s. Teil II) bedeutet nach (Daugman, 1993) anhand einzelner Merkmale zu bestimmen, *was* für ein Objekt vorliegt. Die Translationen geben darüber Auskunft, *wo* sich das Objekt befindet. Zusammenfassend nennt man solche Ansätze *affin-invariante Objekterkennung*. Eine detaillierte Diskussion verschiedener Eigenschaften von Gaborssystemen findet sich in (Feichtinger und Strohmer, 1998).

Daß diese Modellierung einfacher Zellen tatsächlich sinnvoll ist, haben unter anderen J. P. Jones und L. A. Palmer gezeigt (Jones und Palmer, 1987). Sie haben Gaborfunktionen im Sinne kleinster Quadrate an die Meßdaten rezeptiver Felder sowohl im Orts- als auch im Frequenzraum angepaßt, d.h. die freien Parameter geschätzt. Das Ergebnis war, daß die Gaborfunktionen die rezeptiven Felder erstaunlich gut approximieren, obwohl der Fehler zwischen Modell und Daten nicht immer als stationäres Rauschsignal interpretierbar ist.

Die Modellierung von rezeptiven Feldern einfacher Zellen ist im übrigen nicht frei von Widersprüchen. Stork und Wilson (Stork und Wilson, 1990) haben einige Argumente zusammengetragen, die die biologische Plausibilität der Bildkodierung durch komplexwertige Gaborfunktionen und deren experimentelle Evaluierung in Frage stellen. Trotzdem ist das beschriebene Modell für einfache Zellen als komplexes Quadratur-Paar weitgehend anerkannt, s. (Daugman, 1993).

Schon Hubel und Wiesel (Hubel und Wiesel, 1962) haben beobachtet, daß mit konstanter Intensität beleuchtete einfache Zellen keine Antwort ausgeben, was für eine möglichst biologienaher Modellierung bedeutet, daß die Funktion des rezeptiven Feldes mittelwertfrei sein muß. Der einfachste Ansatz ist es, zur Mittelwertkorrektur eine Funktion abzuziehen. Sollen rezeptive Felder verschiedener Größe berücksichtigt werden, wird zusätzlich die Fensterbreite an die Frequenz \vec{k} angepaßt. Diese Einschränkungen erlauben es, eine derart modifizierte Gaborfunktion als *Wavelet* zu betrachten:

$$\tilde{g}_{\vec{x}_0, \vec{k}}(\vec{x}) = e^{-\frac{\|\vec{k}\|^2 (\vec{x} - \vec{x}_0)^T \mathbf{D} (\vec{x} - \vec{x}_0)}{2\sigma^2}} \left(e^{i\vec{k}^T (\vec{x} - \vec{x}_0)} - e^{-\frac{\sigma^2}{2}} \right). \quad (1.3-7)$$

Aus Gründen der Einfachheit beschränkt man sich auf kreisförmige Gaußfenster und sorgt mit einem Vorfaktor $\|\vec{k}\|^2$, der den z.B. in (van der Schaaf und van Hateren, 1996) an natürlichen Bildern beobachteten $\frac{1}{\|\omega\|^2}$ -Abfall kompensieren soll, für konstante Maxima dieser *Gaborwavelets* $\psi_{\vec{k}}(\vec{x} - \vec{x}_0)$ im Frequenzraum:

$$\psi_{\vec{k}}(\vec{x}) = \frac{\|\vec{k}\|^2}{\sigma^2} e^{-\frac{\|\vec{k}\|^2 \|\vec{x}\|^2}{2\sigma^2}} \left(e^{i\vec{k}^T \vec{x}} - e^{-\frac{\sigma^2}{2}} \right) \quad (1.3-8)$$

Auf die Theorie von Wavelets und deren Eigenschaften bei der Darstellung und Rekonstruktion wird im Teil II dieser Arbeit eingegangen.

1.3.2 Komplexe Zellen

Neben den Center-Surround und einfachen Zellen haben Hubel und Wiesel (Hubel und Wiesel, 1962) auch die sog. *komplexen Zellen* gefunden. Diese zeigten in den Versuchen ein sehr unterschiedliches Verhalten. Einige antworten beispielsweise besonders stark auf Kanten, andere auf Bewegung. Zusammenfassend werden komplexe Zellen, im Gegensatz zu den einfachen, nach (De Valois und De Valois, 1988) als nicht phasensensitiv bei einem harmonischen Amplitudengitter als Stimulus angesehen. Deshalb wird die Antwort als der Betrag der Antwort einfacher Zellen nach (1.3-8) (von der Malsburg und Shams, 1998) modelliert:

$$|\mathcal{J}(\vec{x}_0, \vec{k})| = \left| \int_{\mathbb{R}^2} d^2x I(\vec{x}) \psi_{\vec{k}}(\vec{x} - \vec{x}_0) \right|. \quad (1.3-9)$$

Da es aber psychophysische Anzeichen gibt, daß in die menschliche Wahrnehmung auch die Verarbeitung räumlicher Phaseninformationen einfließt, ist nach (De Valois und De Valois, 1988) die Population einfacher Zellen nicht nur eine Zwischenstation für die Antworten komplexer Zellen, sondern stellt ein separates System dar, das ebenfalls Ausgaben aus dem Areal V1 liefert. Die Phasenblindheit ist im allgemeinen bei den komplexen Zellen nicht gegeben, wohl aber exakt für komplexe und näherungsweise auch für reelle harmonischen Gitter mit der Frequenz $\vec{\omega}_0$ und einer Verschiebung \vec{y} :

$$\begin{aligned} I(\vec{x}) &= \frac{1}{\pi} I_0 \cos \vec{\omega}_0^T (\vec{x} - \vec{y}) \iff I_0 [\delta(\vec{\omega} - \vec{\omega}_0) + \delta(\vec{\omega} + \vec{\omega}_0)] e^{-i\vec{\omega}^T \vec{y}}, \quad (1.3-10) \\ |\mathcal{J}(\vec{x}_0, \vec{k})|^2 &= \left| \int_{\mathbb{R}^2} d^2\omega I_0 [\delta(\vec{\omega} - \vec{\omega}_0) + \delta(\vec{\omega} + \vec{\omega}_0)] e^{-i\vec{\omega}^T \vec{y}} \hat{\psi}_{\vec{x}_0, \vec{k}}^*(\vec{\omega}) \right|^2 \\ &= I_0^2 \left[\left| \hat{\psi}_{\vec{x}_0, \vec{k}}(\vec{\omega}_0) \right|^2 + \left| \hat{\psi}_{\vec{x}_0, \vec{k}}(-\vec{\omega}_0) \right|^2 + 2\text{Re} e^{-2i\vec{\omega}_0^T \vec{y}} \hat{\psi}_{\vec{x}_0, \vec{k}}(\vec{\omega}_0) \hat{\psi}_{\vec{x}_0, \vec{k}}^*(-\vec{\omega}_0) \right] \quad (1.3-11) \\ &= I_0^2 \left[\left| \hat{\psi}_{\vec{x}_0, \vec{k}}(\vec{\omega}_0) \right|^2 + \left| \hat{\psi}_{\vec{x}_0, \vec{k}}(-\vec{\omega}_0) \right|^2 + 2\hat{\psi}_{\vec{x}_0, \vec{k}}(\vec{\omega}_0) \hat{\psi}_{\vec{x}_0, \vec{k}}^*(-\vec{\omega}_0) \cos 2\vec{\omega}_0^T \vec{y} \right] \\ &\approx I_0^2 \left[\left| \hat{\psi}_{\vec{x}_0, \vec{k}}(\vec{\omega}_0) \right|^2 + \left| \hat{\psi}_{\vec{x}_0, \vec{k}}(-\vec{\omega}_0) \right|^2 \right], \end{aligned}$$

da $\hat{\psi}_{\vec{x}_0, \vec{k}}(\vec{\omega}_0) \hat{\psi}_{\vec{x}_0, \vec{k}}^*(-\vec{\omega}_0)$ vernachlässigbar klein ist. Die kleine Rechnung zeigt, daß die Antworten komplexer Zellen weniger empfindlich gegenüber \vec{y} sind. Während die Modellierung einfacher Zellen mit Gaborfunktionen weitestgehend anerkannt ist, findet man in der Literatur bei komplexen Zellen verschiedene Ansätze. Beispielsweise werden in (Spitzer und Hochstein, 1985) die Antworten von zwei Untereinheiten, die jeweils aus einer Zeile von Center-Surround-Zellen bestehen, gleichgerichtet und addiert.

Vielfältige Beobachtungen des Antwortverhaltens und die darin auftretenden Nichtlinearitäten sind die Hauptschwierigkeiten beim Finden eines adäquaten Modells für komplexe Zellen.

1.4 Technische Anwendung

Gaborfunktionen des Typs in (1.3-8) unterstützen leistungsfähige Objekterkennungsverfahren, wie das aus (Lades et al., 1993). Hierbei wird eine Gaboranalyse auf einigen wenigen Punkten \vec{x}_0 des aufgenommenen Bildsignals durchgeführt. Diese Punkte liegen aber nicht völlig unabhängig voneinander verteilt, sondern gehören zu einem Graphen. Die Beträge der lokalen Gabortransformierten werden nach sämtlichen Werten von \vec{k} in Vektoren, den sog. *Jets*, gespeichert. Auf gleiche Weise sind auch die Modelle, mit denen die aufgenommenen Bilddaten verglichen werden sollen, in einer Datenbank repräsentiert. Aus den Daten von Knoten und Kanten des Graphen und des Modells wird eine Kostenfunktion synthetisiert, die anschließend mit einer stochastischen Suche minimiert wird. Die Verwendung komplexer Zellen hat den Vorteil, daß die Antwort einfacher Zellen in beiden Kanälen bei für die Objekterkennung wichtigen hochfrequenten Signalanteilen Oszillationen in der Umgebung dieser Merkmale enthält. Ein weiterer Grund sind leichte Verschiebungen, die die Kamera und damit das Eingabebild gegenüber den Knoten des Graphen erfahren und bei komplexen

Zellen kleinere Veränderungen des lokalen Merkmals bewirken. Diese Eigenschaft läßt sich auch zu einem Kriterium verarbeiten. Geht man von einem leicht verschobenen Bildsignal $I(\vec{x} - \vec{x}_1)$ aus, so verschiebt sich auch die Gabortransformierte, $\mathcal{J}_{\vec{k}}(\vec{x}_0 - \vec{x}_1)$. Die oben erwähnte Eigenschaft komplexer Zellen lautet als Formel:

$$\left| \frac{\mathcal{J}_{\vec{k}}(\vec{x}_0 - \vec{x}_1) - \mathcal{J}_{\vec{k}}(\vec{x}_0)}{\mathcal{J}_{\vec{k}}(\vec{x}_0)} \right|^2 \geq \left[\frac{|\mathcal{J}_{\vec{k}}(\vec{x}_0 - \vec{x}_1)| - |\mathcal{J}_{\vec{k}}(\vec{x}_0)|}{|\mathcal{J}_{\vec{k}}(\vec{x}_0)|} \right]^2. \quad (1.4-1)$$

Der Nachweis kann durch langwieriges Nachrechnen erfolgen.

2 Bildverarbeitung

Wie in jeder Arbeit, die von der Darstellung oder Verarbeitung von Signalen handelt, sind Frequenzraumdarstellungen unverzichtbar. Deshalb werden hier zur Einführung der Nomenklatur die elementaren Transformationen für Signale in kontinuierlichen und diskreten Hilberträumen zusammengestellt.

2.1 Kontinuierliche Transformationen

2.1.1 Die Fouriertransformation

Die in dieser Arbeit durchgängig benutzte Notation für die zweidimensionale kontinuierliche *Fouriertransformation* ist

$$\hat{f}(\vec{\omega}) = \text{FT}\{f(\vec{x})\}(\vec{\omega}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} d^2x f(\vec{x}) e^{-i\vec{x}^T \vec{\omega}}. \quad (2.1-1)$$

Die Transformation vom Frequenz- in den Ortsraum lautet dann entsprechend

$$f(\vec{x}) = \text{IFT}\{\hat{f}(\vec{\omega})\}(\vec{x}) = \frac{1}{2\pi} \int_{\mathbb{R}^2} d^2\omega \hat{f}(\vec{\omega}) e^{i\vec{x}^T \vec{\omega}} \quad (2.1-2)$$

Einige Eigenschaften der Fouriertransformation sind im Anhang A.1 zusammengestellt.

2.1.2 Die Gabortransformation

Für die Darstellung und Verarbeitung lokaler Bildmerkmale sind *Gaborwavelets* ein leistungsfähiges Werkzeug. Sie sind durch

$$\psi_{\vec{k}}(\vec{x}) := \frac{\|\vec{k}\|^2}{\sigma^2} e^{-\frac{\|\vec{k}\|^2 \|\vec{x}\|^2}{2\sigma^2}} \left(e^{i\vec{k}^T \vec{x}} - e^{-\frac{\sigma^2}{2}} \right) \quad (2.1-3)$$

definiert. Ihre Frequenzraumdarstellung

$$\hat{\psi}_{\vec{k}}(\vec{\omega}) = e^{-\frac{\sigma^2 \|\vec{\omega} - \vec{k}\|^2}{2\|\vec{k}\|^2}} - e^{-\frac{\sigma^2 (\|\vec{k}\|^2 + \|\vec{\omega}\|^2)}{2\|\vec{k}\|^2}} \quad (2.1-4)$$

macht deutlich, daß ein solches System bei festgehaltenem \vec{k} eine Gaußglocke im Frequenzraum positioniert, wie in Abb. 3 zu sehen ist. Der zweite Term in (2.1-4) dient zur Unterdrückung des Gleichanteils, was eine wichtige Voraussetzung dafür ist, daß die Gaborwavelets zulässig² sind:

$$\hat{\psi}_{\vec{k}}(\vec{0}) = 0. \quad (2.1-5)$$

Die Zerlegung des Bildes $I(\vec{x})$ nach Gaborwavelets wird in Form einer Faltungsoperation im Ortsraum

$$\mathcal{J}_{\vec{k}}(\vec{x}) = \int_{\mathbb{R}^2} d^2x' I(\vec{x}') \psi_{\vec{k}}(\vec{x} - \vec{x}') = I(\vec{x}) * \psi_{\vec{k}}(\vec{x}) \quad (2.1-6)$$

²Was „zulässig“ bedeutet, s.u.

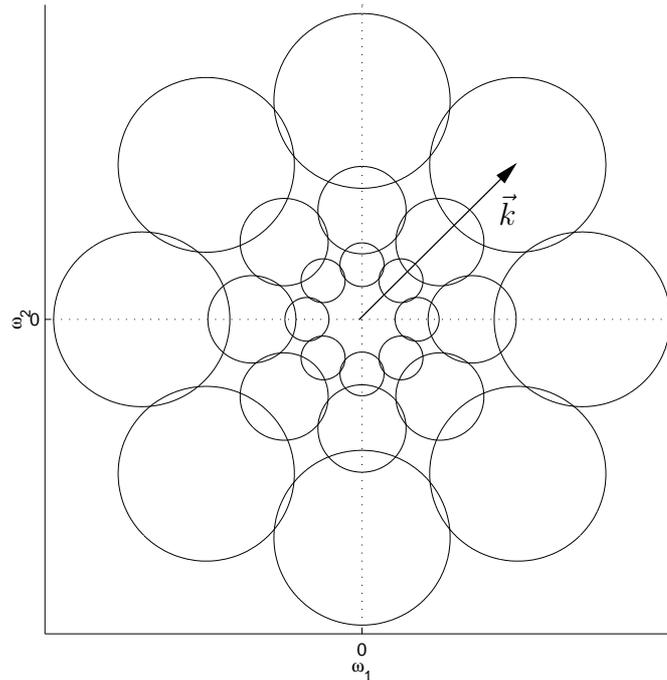


Abbildung 3: Teilweise Abdeckung des Frequenzraums durch Gaborfunktionen. Die Kreise stellen Gaußfunktionen dar, die jeweils um \vec{k} vom Nullpunkt verschoben sind. Dargestellt sind 8 Orientierungen und 3 Skalen.

und einer Multiplikation im Frequenzraum

$$\hat{\mathcal{J}}_{\vec{k}}(\vec{\omega}) = 2\pi \hat{I}(\vec{\omega}) \hat{\psi}_{\vec{k}}(\vec{\omega}) \quad (2.1-7)$$

geschrieben. Für eine endliche Menge an \vec{k} kann die Wavelet-Transformation (2.1-6) als Ausgabe einer Filterbank angesehen werden, deren Einzelfilter verschiedene Bereiche im Frequenzraum überdecken.

2.2 Diskrete Transformationen

2.2.1 Signalabtastung

Für die Verarbeitung durch digitale Systeme werden die in den vorhergehenden Abschnitten kurz beschriebenen Filteroperationen auch diskretisiert betrachtet. Das Bild $I(\vec{x})$ wird daher durch äquidistante Gitterpunkte mit Abstand Δ , dem *Abtastintervall*, in beide Koordinatenrichtungen abgetastet, weshalb die Notation für die Bildsignale auf

$$I(\vec{n}) := I(\vec{n}\Delta) \quad \forall \vec{n} \in \mathbb{Z}^2 \quad (2.2-1)$$

geändert wird. Man geht dabei davon aus, daß bezüglich der Grenzfrequenzen $\vec{\omega}_G$ das *Nyquist-Kriterium*

$$\vec{\omega}_G^T \vec{e}_j \leq \frac{\pi}{\Delta} \quad j \in \{1, 2\} \quad (2.2-2)$$

erfüllt ist. Das eindimensionale *Nyquistintervall* sei in dieser Arbeit

$$\mathcal{N} = [-\pi; \pi[\quad (2.2-3)$$

Das Abtasttheorem bezeichnet die Rekonstruktionsformel kontinuierlicher Signale aus ihren Abtastwerten und ist im Anhang A.4 angegeben.

2.2.2 Die zweidimensionale ortsdiskrete Fouriertransformation

In völliger Analogie zu (2.1-1) kann die *ortsdiskrete Fouriertransformation* mit

$$\check{f}(\vec{\nu}) = \text{DSFT}\{f(\vec{n})\}(\vec{\nu}) = \frac{1}{2\pi} \sum_{\vec{n} \in \mathbb{Z}^2} f(\vec{n}) e^{-i\vec{n}^T \vec{\nu}} \quad (2.2-4)$$

und die inverse ortsdiskrete Fouriertransformation mit

$$f(\vec{n}) = \text{IDSFT}\{\check{f}(\vec{\nu})\}(\vec{n}) = \frac{1}{2\pi} \int_{\mathcal{N}^2} d^2\nu \check{f}(\vec{\nu}) e^{i\vec{n}^T \vec{\nu}} \quad (2.2-5)$$

angegeben werden. Bei solchen Darstellungen ist es üblich, den normierten Frequenzvektor

$$\vec{\nu} = \vec{\omega} \Delta \in \mathcal{N} \times \mathcal{N} \quad (2.2-6)$$

zu benutzen. $\check{f}(\vec{\nu})$ ist periodisch fortgesetzt, d.h. es gilt $\check{f}(\vec{\nu}) = \check{f}(\vec{\nu} + 2\pi j_1 \vec{e}_1 + 2\pi j_2 \vec{e}_2)$. Ist die Nyquistbedingung nicht erfüllt, d.h. die doppelte Grenzfrequenz größer als die vorgesehene Periodenlänge kommt es zum *Aliasing*. Die Abtastrate ist in diesem Fall nach oben zu korrigieren.

Entsprechend können auch die Gaborfunktionen und die zugehörigen Transformationen in abgetasteter Form als

$$\begin{aligned} \psi_{\vec{k}}(\vec{n}) &:= \psi_{\vec{k}}(\vec{n}\Delta) = \frac{\|\vec{k}\|^2}{\sigma^2} e^{-\frac{\|\vec{k}\|^2 \Delta^2 \|\vec{n}\|^2}{2\sigma^2}} \left(e^{i\vec{k}^T \Delta \vec{n}} - e^{-\frac{\sigma^2}{2}} \right) \\ &= \frac{\|\vec{k}\|^2}{\Delta^2 \sigma^2} e^{-\frac{\|\vec{k}\|^2 \|\vec{n}\|^2}{2\sigma^2}} \left(e^{i\vec{k}^T \vec{n}} - e^{-\frac{\sigma^2}{2}} \right) = \psi_{\frac{\vec{k}}{\Delta}}(\vec{n}) \end{aligned} \quad (2.2-7)$$

und im Frequenzraum bezüglich der DSFT (s. (A.2))

$$\check{\psi}_{\frac{\vec{k}}{\Delta}}(\vec{\nu}) = \frac{1}{\Delta^2} \left[e^{-\frac{\sigma^2 \|\vec{\nu} - \vec{k}\|^2}{2\|\vec{k}\|^2}} - e^{-\frac{\sigma^2 (\|\vec{k}\|^2 + \|\vec{\nu}\|^2)}{2\|\vec{k}\|^2}} \right] = \frac{1}{\Delta^2} \check{\psi}_{\vec{k}}(\vec{\nu}) \quad (2.2-8)$$

mit

$$\vec{k} = \Delta \vec{\vec{k}} \in \mathcal{N} \times \mathcal{N}. \quad (2.2-9)$$

dargestellt werden. Um sich des Vorfaktors zu entledigen, sollen die ortsdiskreten Gaborwavelets neu definiert werden:

$$\psi_{\vec{k}}(\vec{n}) = \frac{\|\vec{k}\|^2}{\sigma^2} e^{-\frac{\|\vec{k}\|^2 \|\vec{n}\|^2}{2\sigma^2}} \left(e^{i\vec{k}^T \vec{n}} - e^{-\frac{\sigma^2}{2}} \right) \quad (2.2-10)$$

$$\Downarrow \quad (2.2-11)$$

$$\check{\psi}_{\vec{k}}(\vec{\nu}) = e^{-\frac{\sigma^2 \|\vec{\nu} - \vec{k}\|^2}{2\|\vec{k}\|^2}} - e^{-\frac{\sigma^2 (\|\vec{k}\|^2 + \|\vec{\nu}\|^2)}{2\|\vec{k}\|^2}}. \quad (2.2-12)$$

Die diskrete Gaborwavelettransformation erfolgt nun durch

$$\mathcal{J}_{\vec{k}}(\vec{n}) = \sum_{\vec{n}' \in \mathbb{Z}^2} I(\vec{n}') \psi_{\vec{k}}(\vec{n} - \vec{n}') = I(\vec{n}) * \psi_{\vec{k}}(\vec{n}) \quad (2.2-13)$$

bzw.

$$\hat{\mathcal{J}}_{\vec{k}}(\vec{\nu}) = 2\pi \hat{I}(\vec{\nu}) \check{\psi}_{\vec{k}}(\vec{\nu}). \quad (2.2-14)$$

Einige Eigenschaften der DSFT und der Zusammenhang zur FT finden sich im Anhang A.2.

2.2.3 Die zweidimensionale Z-Transformation

Ausgehend von der expliziten Schreibweise der DSFT

$$f(\nu_1, \nu_2) = \frac{1}{2\pi} \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f(n_1, n_2) e^{-i\nu_1 n_1} e^{-i\nu_2 n_2} \quad (2.2-15)$$

setzt man den Exponenten des Transformationskerns holomorph für beide Variablen in die gesamte komplexe Zahlenebene fort

$$e^{-i\nu_j n_j} \rightarrow e^{-(\mu_j + i\nu_j)n_j} \quad \forall j \in \{1, 2\} \quad (2.2-16)$$

und führt die Variablen

$$z_j = e^{(\mu_j + i\nu_j)} \quad \forall j \in \{1, 2\} \quad (2.2-17)$$

ein. Somit wird ein Nyquist-Intervall einer Dimension in die ganze komplexe Zahlenebene konform abgebildet. Daraus ergibt sich die Definition der zweidimensionalen *Z-Transformation*:

$$\hat{f}(z_1, z_2) = ZT\{f(n_1, n_2)\}(z_1, z_2) = \frac{1}{2\pi} \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f(n_1, n_2) z_1^{-n_1} z_2^{-n_2}. \quad (2.2-18)$$

Da bei weitem nicht für alle Funktionen die Summen in (2.2-18) konvergieren, soll $f(\vec{n}) = I(\vec{n})$ über einen endlichen Träger verfügen, wie es bei Bildern der Fall ist:

$$I(\vec{n}) = 0 \quad \forall n_1 \notin [0, N_1 - 1] \vee n_2 \notin [0, N_2 - 1]. \quad (2.2-19)$$

$$\hat{I}(z_1, z_2) = ZT\{I(n_1, n_2)\}(z_1, z_2) = \frac{1}{2\pi} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} I(n_1, n_2) z_1^{-n_1} z_2^{-n_2} \quad (2.2-20)$$

kann daher als ein Polynom in den Variablen z_1^{-1} und z_2^{-1} interpretiert werden.

Die Z-Transformation wird im Abschnitt 7 zur Herausstellung der Polynomeigenschaft von DSFT und DFT Verwendung finden.

2.2.4 Die zweidimensionale diskrete Fouriertransformation

Für numerische Operationen im Frequenzraum muß dieser ebenfalls abgetastet werden, und zwar mit soviel Punkten, wie auch für das Bildsignal im Ortsraum vorgesehen sind:

$$\omega_A = \frac{2\pi}{\Delta} = N_1 \hat{\Delta}_1 = N_2 \hat{\Delta}_2, \quad (2.2-21)$$

$$\nu_1 = \Delta\omega_1 = \Delta \hat{\Delta}_1 \rho_1 = \frac{2\pi}{N_1} \rho_1, \quad (2.2-22)$$

$$\nu_2 = \Delta\omega_2 = \Delta \hat{\Delta}_2 \rho_2 = \frac{2\pi}{N_2} \rho_2, \quad (2.2-23)$$

wobei $\hat{\Delta}$ das Frequenzabtastrintervall und ω_A die Abtastrate bezeichnet. Da für eine verlustfreie Frequenzabtastrung der Ortsraum begrenzt sein muß, verfügt das Bildsignal über einen endlichen Träger. Aufgrund einer anderen Achsenskalierung im Frequenzraum ist dieser mit einem anderen Maß ausgestattet, weswegen der Vorfaktor korrigiert werden muß. Damit und mit (2.2-21) kann daher die *diskrete Fouriertransformation*

$$\hat{f}(\rho_1, \rho_2) = \text{DFT}_{N_1, N_2}\{f(n_1, n_2)\}(\rho_1, \rho_2) = \frac{1}{\sqrt{N_1 N_2}} \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} f(n_1, n_2) e^{-i \frac{2\pi}{N_1} n_1 \rho_1} e^{-i \frac{2\pi}{N_2} n_2 \rho_2} \quad (2.2-24)$$

und deren Umkehrung

$$f(n_1, n_2) = \text{IDFT}_{N_1, N_2}\{\hat{f}(\rho_1, \rho_2)\}(n_1, n_2) = \frac{1}{\sqrt{N_1 N_2}} \sum_{\rho_1=0}^{N_1-1} \sum_{\rho_2=0}^{N_2-1} \hat{f}(\rho_1, \rho_2) e^{i \frac{2\pi}{N_1} n_1 \rho_1} e^{i \frac{2\pi}{N_2} n_2 \rho_2} \quad (2.2-25)$$

definiert werden. Wichtig ist noch anzumerken, daß das Bildsignal sowohl im Orts- als auch im Frequenzraum als periodisch mit N_1 bzw. mit N_2 fortgesetzt angesehen werden muß. Die Durchführung der DFT oder der IDFT entspricht daher der Berechnung von Fourierkoeffizienten aus Fourierkoeffizienten. Im Anhang A.3 ist u.a. der Zusammenhang zwischen DFT und FT zu finden.

2.2.5 Angepaßte Abtastung von Gaborkernen

Im kontinuierlichen Fall ist die Übertragungsfunktion eines einzelnen Gabor-Filters $\hat{\psi}_{\vec{k}}(\vec{\omega})$ nicht bandbegrenzt, sondern außer bei $\vec{\omega} = \vec{0}$ von null verschieden. Da die Gaußglocken jedoch sehr schnell abfallen, kann für die numerische Umsetzung davon ausgegangen werden, daß dieses Filter einen Bereich aus dem Frequenzraum ausschneidet. Der Schwellwert, unterhalb dessen alle Amplituden zu null werden, sei β ($0 \leq \beta \leq 1$) bei den Frequenzen $\vec{\omega}_S$:

$$\hat{\psi}_{\vec{k}}(\vec{\omega}_S) = \beta = e^{-\frac{\sigma^2 \|\vec{\omega}_S - \vec{k}\|^2}{2\|\vec{k}\|^2}} - e^{-\frac{\sigma^2 (\|\vec{k}\|^2 + \|\vec{\omega}_S\|^2)}{2\|\vec{k}\|^2}}. \quad (2.2-26)$$

Abgeschnitten wird außerhalb von Kreisen mit dem Radius $r_{\vec{k}} = \|\vec{\omega}_S - \vec{k}\|$. Da das Wavelet nur an einer einzigen Stelle im Frequenzraum lokalisiert ist, kann außerhalb dieser Kreise die periodische Fortsetzung erfolgen. Als Abtastrate für die Punktantwort $\psi_{\vec{k}}(\vec{x})$ kann somit

$$\omega_A = \frac{2\pi}{\Delta} = 2r_{\vec{k}} \quad (2.2-27)$$

festgesetzt werden. Man vergleiche dazu Unterabschnitt 6.4.3. Dies stellt eine Verallgemeinerung der Nyquist-Bedingung dar, da die eigentliche Grenzfrequenz nicht mehr von Bedeutung ist. Um das Abtastintervall Δ in Abhängigkeit der Schwelle β angeben zu können, wird in (2.2-26) eine Näherung benutzt, ohne die das Auflösen dieser Gleichung analytisch nicht möglich wäre:

$$\begin{aligned} \hat{\psi}_{\vec{k}}(\vec{\omega}_S) = \beta &\approx e^{-\frac{\sigma^2 \|\vec{\omega}_S - \vec{k}\|^2}{2\|\vec{k}\|^2}} = e^{-\frac{\sigma^2 r_{\vec{k}}^2}{2\|\vec{k}\|^2}} \\ \Rightarrow \ln \frac{1}{\beta} &\approx \frac{\sigma^2 r_{\vec{k}}^2}{2\|\vec{k}\|^2} = \frac{\sigma^2 \pi^2}{2\|\vec{k}\|^2 \Delta^2} \\ \Rightarrow \Delta &\approx \frac{\sigma \pi}{\|\vec{k}\|} \frac{1}{\sqrt{2 \ln \frac{1}{\beta}}}. \end{aligned} \quad (2.2-28)$$

Eine genaue Angabe ist nicht möglich, also sollte β entsprechend niedrig gewählt werden, um vor bemerkbaren Aliasing-Effekten sicher zu sein.

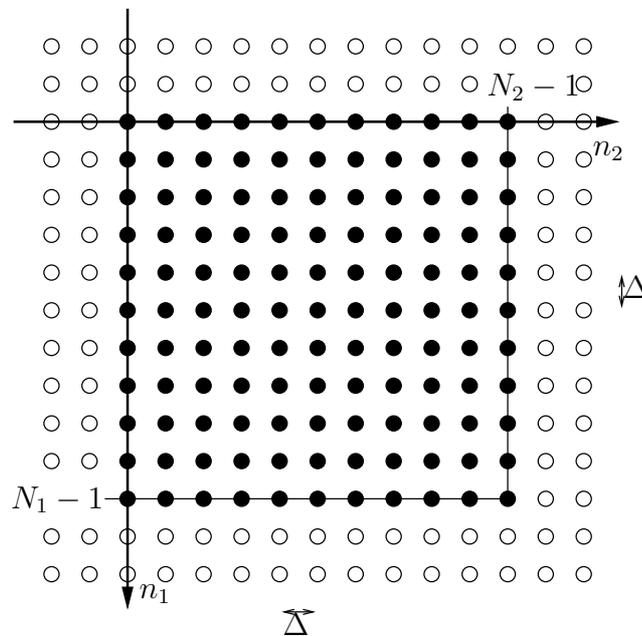


Abbildung 4: Das Koordinatensystem mit dem endlichen Träger B .

Teil II

Einfache Zellen

- Definition zweidimensionaler Wavelets.
- Definition der GWT.
- Berechnung der IGWT.
- Frames, Frameoperator, Metrischer Operator, Frameschranken, Übervollständigkeit.
- Abtastung des Phasenraums, also Diskretisierung von Translationen, Skalen und Orientierungen: DPSGWT und IDPSGWT.
- Zusätzliche Abtastung von Orts- und Frequenzraum: DGWT und IDGWT.
- Numerische Probleme bei der Inversion, Singulärwertzerlegung.
- Lokale Rekonstruktion.
- Pyramidale DGWT und IDGWT.
- Vergleich Wavelet-Frames – orthogonale Wavelet-Basen.
- Quantifizierung der Übervollständigkeit.

3 Die kontinuierliche Gaborwavelettransformation

3.1 Zweidimensionale kontinuierliche Gaborwavelets

In späteren Abschnitten gewinnt die Tatsache, daß man es hier nicht mit einer beliebigen Filterbank mit einander überlappenden Frequenzbändern sondern mit *Wavelets* zu tun hat, an Bedeutung. Zunächst muß aber der Begriff „Wavelet“ definiert werden.

Definition 1 Eine Funktion $\psi(\vec{x}) \neq 0$ mit $\vec{x} \in \mathbb{R}^2$ ist genau dann ein Wavelet, wenn

$$0 < \int_{\mathbb{R}^2} d^2\omega \frac{|\hat{\psi}(\vec{\omega})|^2}{\|\vec{\omega}\|^2} < \infty \quad (3.1-1)$$

gilt. (3.1-1) nennt man Zulässigkeitsbedingung³

Die Zulässigkeit kann daher durch zwei Bedingungen formuliert werden:

- Die Waveletfunktion hat keinen Mittelwert, $\hat{\psi}(\vec{0}) = 0$.
- Die Waveletfunktion hat im Frequenzraum einen starken Abfall für größer werdende $\|\vec{\omega}\|$.

Der Sinn dieser Definition wird in den nächsten Abschnitten geklärt. Von nun an werden zwei Notationen für Gaborwaveletfunktionen verwendet, nämlich ψ , das schon in (2.1-3) definiert wurde, und die im folgenden eingeführte χ -Bezeichnung. Zweidimensionale Wavelets gehen durch *Translationen* \vec{x}_0 , *Orientierungen* ϑ und *Skalen* a aus einem *Mutterwavelet* $\chi(\vec{x})$ hervor (Lee, 1996). Die Menge der auf diese Weise aus einem Mutterwavelet gebildeten Funktionen,

$$\chi_{x_{01}, x_{02}, a, \vartheta}(x_1, x_2) = \frac{1}{a} \chi \left(\frac{x_1 - x_{01}}{a} \cos \vartheta + \frac{x_2 - x_{02}}{a} \sin \vartheta, -\frac{x_1 - x_{01}}{a} \sin \vartheta + \frac{x_2 - x_{02}}{a} \cos \vartheta \right), \quad (3.1-2)$$

nennt man *Waveletfamilie*. Eine etwas kompaktere Schreibweise sei mit

$$\begin{aligned} \chi_{\vec{x}_0, a, \vartheta}(\vec{x}) &= \frac{1}{a} \chi \left(\frac{Q(\vartheta)(\vec{x} - \vec{x}_0)}{a} \right) \\ \Downarrow \\ \hat{\chi}_{\vec{x}_0, a, \vartheta}(\vec{\omega}) &= a \hat{\chi}[Q(\vartheta)\vec{\omega}a] e^{-i\vec{\omega}^T \vec{x}_0} \end{aligned} \quad (3.1-3)$$

angegeben, wobei die Matrix

$$Q(\vartheta) = \begin{bmatrix} \cos \vartheta & \sin \vartheta \\ -\sin \vartheta & \cos \vartheta \end{bmatrix} \quad (3.1-4)$$

die Drehung um den Orientierungswinkel ϑ durchführt. Das Mutterwavelet, welches den Satz von Gaborfiltern hervorbringt, wird als

$$\chi(\vec{x}) = \frac{1}{\sigma\tau} e^{-\frac{1}{2}\|S_{\sigma, \tau}\vec{x}\|^2} \left(e^{i\vec{x}^T \vec{e}_1} - e^{-\frac{\sigma^2}{2}} \right) \quad (3.1-5)$$

$$\Downarrow \\ \hat{\chi}(\vec{\omega}) = e^{-\frac{1}{2}\|S_{\sigma, \tau}^{-1}(\vec{\omega} - \vec{e}_1)\|^2} - e^{-\frac{1}{2}(\|S_{\sigma, \tau}^{-1}\vec{\omega}\|^2 + \sigma^2)} \quad (3.1-6)$$

definiert. Die Gaußglocke hat also elliptische Form, deren Hauptachsen durch

$$S_{\sigma, \tau} = \begin{bmatrix} \frac{1}{\sigma} & 0 \\ 0 & \frac{1}{\tau} \end{bmatrix}, \quad \tau \geq \sigma \quad (3.1-7)$$

³Diese Zuverlässigkeitsbedingung gilt bezüglich der durch die Translationen, Skalen und Orientierungen in (3.1-3) definierten Gruppe. In (Murenzi, 1989) ist ein gruppentheoretisch „sauberer“ Übergang von der affin-linearen Gruppe aus Skalen und Translationen auf die in dieser Arbeit benutzte. S. dazu auch (Louis et al., 1994). Eine gruppentheoretische Behandlung würde hier allerdings zu weit führen.

vorgegeben werden können, wobei die Ausbreitung der lokalisierten ebenen Welle immer in Richtung der kleineren Hauptachse erfolgt. Des weiteren wird die Abkürzung

$$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad (3.1-8)$$

verwendet. Die Analyse des Bildsignals $I(\vec{x})$ mit den Wavelets $\chi_{\vec{x}_0, a, \vartheta}(\vec{x})$ wird über ein Skalarprodukt berechnet:

$$\mathcal{I}(\vec{x}_0, a, \vartheta) = \text{GWT}\{I(\vec{x})\}(\vec{x}_0, a, \vartheta) = \langle I, \chi_{\vec{x}_0, a, \vartheta} \rangle = \int_{\mathbb{R}^2} d^2x I(\vec{x}) \chi_{\vec{x}_0, a, \vartheta}^*(\vec{x}). \quad (3.1-9)$$

Das Signal wird somit auf die Waveletfamilie projiziert. Die Menge aller Punkte \vec{x}_0, a, ϑ wird *Phasenraum* genannt. Dieses innere Produkt existiert, wenn

$$I(\vec{x}) \in L^2(\mathbb{R}^2) \quad \wedge \quad \chi_{\vec{x}_0, a, \vartheta}(\vec{x}) \in L^2(\mathbb{R}^2) \quad (3.1-10)$$

gilt. Bei $L^2(\mathbb{R}^2)$ handelt es sich um die Menge der energiebeschränkten Signale aus \mathbb{R}^2 :

$$f(\vec{x}) \in L^2(\mathbb{R}^2) \quad \Leftrightarrow \quad \int_{\mathbb{R}^2} d^2x |f(\vec{x})|^2 < \infty. \quad (3.1-11)$$

Die folgenden Überlegungen gelten der Umkehrung dieses Prozesses, d.h. der Transformation

$$I(\vec{x}) = \text{IGWT}\{\mathcal{I}(\vec{x}_0, a, \vartheta)\}(\vec{x}). \quad (3.1-12)$$

3.2 Signalrekonstruktion aus zweidimensionalen Gaborwavelets

3.2.1 Die inverse Gaborwavelettransformation

Die *Gaborwavelettransformation* kann auch folgendermaßen dargestellt werden:

$$\begin{aligned} \mathcal{I}(\vec{x}_0, a, \vartheta) &= \langle I, \chi_{\vec{x}_0, a, \vartheta} \rangle = \langle \hat{I}, \hat{\chi}_{\vec{x}_0, a, \vartheta} \rangle \\ &= \int_{\mathbb{R}^2} d^2\omega \hat{I}(\vec{\omega}) \hat{\chi}_{\vec{x}_0, a, \vartheta}^*(\vec{\omega}) \\ &= \int_{\mathbb{R}^2} d^2\omega \hat{I}(\vec{\omega}) a \hat{\chi}^*[\mathbf{Q}(\vartheta)\vec{\omega}a] e^{i\vec{\omega}^T \vec{x}_0} \\ &= 2\pi \text{IFT}\{a \hat{\chi}^*[\mathbf{Q}(\vartheta)\vec{\omega}a] \hat{I}(\vec{\omega})\}(\vec{x}_0). \end{aligned} \quad (3.2-1)$$

Eine direkte Folge daraus ist

$$\frac{1}{2\pi} \int_{\mathbb{R}^2} d^2x_0 \mathcal{I}(\vec{x}_0, a, \vartheta) e^{-i\vec{\omega}^T \vec{x}_0} = 2\pi a \hat{\chi}^*[\mathbf{Q}(\vartheta)\vec{\omega}a] \hat{I}(\vec{\omega}). \quad (3.2-2)$$

Mit einer Division durch $\hat{\chi}^*[\mathbf{Q}(\vartheta)\vec{\omega}a]$ und einer inversen Fouriertransformation ließe sich aus (3.2-2) eine Formel für das Ursprungsbildsignal $I(\vec{x})$ angeben. Allerdings liefert dies keine wirkliche Lösung, da Gaußglocken nach außen hin beliebig klein werden, weshalb deren Inversion aufgrund fehlender Konvergenz bei der Rücktransformation praktisch nicht durchführbar wäre. Gerade in der noch zu behandelnden abgetasteten Form auf einem endlichen Träger ist eine derartige inverse Filterung numerisch nicht sinnvoll. (3.2-2) mit $2\pi w(a, \vartheta) \hat{\chi}[\mathbf{Q}(\vartheta)\vec{\omega}a]$ multipliziert, ergibt

$$\int_{\mathbb{R}^2} d^2x_0 \mathcal{I}(\vec{x}_0, a, \vartheta) e^{-i\vec{\omega}^T \vec{x}_0} w(a, \vartheta) \hat{\chi}[\mathbf{Q}(\vartheta)\vec{\omega}a] = 4\pi^2 w(a, \vartheta) a |\hat{\chi}[\mathbf{Q}(\vartheta)\vec{\omega}a]|^2 \hat{I}(\vec{\omega}). \quad (3.2-3)$$

$w(a, \vartheta)$ ist eine Gewichtsfunktion, die zunächst nicht stört, sich aber später eher als nützlich herausstellen wird. Zur weiteren Verarbeitung wird (3.2-3) über die Variablen a und ϑ in ihren jeweiligen Grenzen integriert:

$$\begin{aligned} \int_{\mathbb{R}^+} da \int_0^{2\pi} d\vartheta w(a, \vartheta) \int_{\mathbb{R}^2} d^2x_0 \mathcal{I}(\vec{x}_0, a, \vartheta) e^{-i\vec{\omega}^T \vec{x}_0} \hat{\chi}[\mathbf{Q}(\vartheta)\vec{\omega}a] = \\ 4\pi^2 \int_{\mathbb{R}^+} da \int_0^{2\pi} d\vartheta w(a, \vartheta) a |\hat{\chi}[\mathbf{Q}(\vartheta)\vec{\omega}a]|^2 \hat{I}(\vec{\omega}) = Y(\vec{\omega}) \hat{I}(\vec{\omega}). \end{aligned} \quad (3.2-4)$$

Um aus dieser Gleichung das Bildsignal $I(\vec{x})$ rekonstruieren zu können, muß

$$Y(\vec{\omega}) = 4\pi^2 \int_{\mathbb{R}^+} da \int_0^{2\pi} d\vartheta w(a, \vartheta) a |\hat{\chi}[\mathbf{Q}(\vartheta)\vec{\omega}a]|^2 \quad (3.2-5)$$

die Bedingung

$$0 < A \leq Y(\vec{\omega}) \leq B < \infty \quad (3.2-6)$$

erfüllen. Die IGWT kann nun formal mit

$$\begin{aligned} \hat{I}(\vec{\omega}) &= Y^{-1}(\vec{\omega}) \int_{\mathbb{R}^+} da \int_0^{2\pi} d\vartheta w(a, \vartheta) \int_{\mathbb{R}^2} d^2x_0 \mathcal{I}(\vec{x}_0, a, \vartheta) e^{-i\vec{\omega}^T \vec{x}_0} \hat{\chi}[\mathbf{Q}(\vartheta)\vec{\omega}a] \\ &= Y^{-1}(\vec{\omega}) \int_{\mathbb{R}^+} \frac{da}{a} \int_0^{2\pi} d\vartheta w(a, \vartheta) \int_{\mathbb{R}^2} d^2x_0 \mathcal{I}(\vec{x}_0, a, \vartheta) \hat{\chi}_{\vec{x}_0, a, \vartheta}(\vec{\omega}) \\ &= \int_{\mathbb{R}^+} \frac{da}{a} \int_0^{2\pi} d\vartheta w(a, \vartheta) \int_{\mathbb{R}^2} d^2x_0 \mathcal{I}(\vec{x}_0, a, \vartheta) \hat{\chi}^{\vec{x}_0, a, \vartheta}(\vec{\omega}) \\ &\Downarrow \\ I(\vec{x}) &= \int_{\mathbb{R}^+} \frac{da}{a} \int_0^{2\pi} d\vartheta w(a, \vartheta) \int_{\mathbb{R}^2} d^2x_0 \mathcal{I}(\vec{x}_0, a, \vartheta) \chi^{\vec{x}_0, a, \vartheta}(\vec{x}) \end{aligned} \quad (3.2-7)$$

angegeben werden. $\chi^{\vec{x}_0, a, \vartheta}(\vec{x})$ sei eine durch

$$\begin{aligned} \hat{\chi}^{\vec{x}_0, a, \vartheta}(\vec{\omega}) &= Y^{-1}(\vec{\omega}) \hat{\chi}_{\vec{x}_0, a, \vartheta}(\vec{\omega}) \\ &\Downarrow \\ \chi^{\vec{x}_0, a, \vartheta}(\vec{x}) &= \frac{1}{2\pi} \int_{\mathbb{R}^2} d^2\omega Y^{-1}(\vec{\omega}) a \hat{\chi}[\mathbf{Q}(\vartheta)\vec{\omega}a] e^{i\vec{\omega}^T (\vec{x} - \vec{x}_0)} \end{aligned} \quad (3.2-8)$$

definierte Familie neuer Wavelets, von denen alle ohne Translation

$$\begin{aligned} \chi^{\vec{0}, a, \vartheta}(\vec{x}) &= \frac{1}{2\pi} \int_{\mathbb{R}^2} d^2\omega Y^{-1}(\vec{\omega}) a \hat{\chi}[\mathbf{Q}(\vartheta)\vec{\omega}a] e^{i\vec{\omega}^T \vec{x}} \\ &= \frac{1}{2\pi a} \int_{\mathbb{R}^2} d^2\omega Y^{-1} \left(\frac{\mathbf{Q}^T(\vartheta)\vec{\omega}}{a} \right) \hat{\chi}(\vec{\omega}) e^{i\vec{\omega}^T \frac{\mathbf{Q}(\vartheta)\vec{x}}{a}} \end{aligned} \quad (3.2-9)$$

sind. Um diese Funktionen tatsächlich als Wavelets interpretieren zu können, muß

$$\chi^{\vec{0}, a, \vartheta}(\vec{x}) = \frac{1}{a} \chi^{\vec{0}, 1, 0} \left(\frac{\mathbf{Q}(\vartheta)\vec{x}}{a} \right) \Leftrightarrow Y \left(\frac{\mathbf{Q}^T(\vartheta)\vec{\omega}}{a} \right) = Y(\vec{\omega}) \quad (3.2-10)$$

gelten. Diese Bedingung läßt für $Y(\vec{\omega})$ nur eine Wahl zu:

$$Y(\vec{\omega}) = C = \text{konst.} \quad (3.2-11)$$

Diese Vorgabe muß allerdings noch in (3.2-5) verwirklicht werden:

$$C = 4\pi^2 \int_{\mathbb{R}^+} da \int_0^{2\pi} d\vartheta w(a, \vartheta) a |\hat{\chi}[\mathbf{Q}(\vartheta)\vec{\omega}a]|^2. \quad (3.2-12)$$

Für eine bessere Übersichtlichkeit benutzt man die Substitution

$$\vec{\xi} = \mathbf{Q}(\vartheta)\vec{\omega}a, \quad dad\vartheta = \frac{a}{\|\vec{\xi}\|^2} d^2\xi, \quad (3.2-13)$$

woraus

$$C = 4\pi^2 \int_{\mathbb{R}^2} d^2\xi \frac{a^2}{\|\vec{\xi}\|^2} w(a, \vartheta) |\hat{\chi}(\vec{\xi})|^2 \quad (3.2-14)$$

folgt. Mit der Wahl $w(a, \vartheta) = \frac{1}{a^2}$ kann die Konstante C mit

$$C = 4\pi^2 \int_{\mathbb{R}^2} d^2\xi \frac{|\hat{\chi}(\vec{\xi})|^2}{\|\vec{\xi}\|^2} \quad (3.2-15)$$

angegeben werden, womit auch die Definition 1 gerechtfertigt ist. Um einen Zusammenhang zu den oben als $\psi_{\vec{k}}(\vec{x})$ definierten Gaborwavelets herzustellen, führt man in

$$\begin{aligned} I(\vec{x}) &= \text{IGWT}\{\mathcal{I}(\vec{x}_0, a, \vartheta)\}(\vec{x}) \\ &= C^{-1} \int_{\mathbb{R}^+} \frac{da}{a^3} \int_0^{2\pi} d\vartheta \int_{\mathbb{R}^2} d^2x_0 \mathcal{I}(\vec{x}_0, a, \vartheta) \chi_{\vec{x}_0, a, \vartheta}(\vec{x}) \\ &= C^{-1} \int_{\mathbb{R}^+} \frac{da}{a^4} \int_0^{2\pi} d\vartheta \int_{\mathbb{R}^2} d^2x_0 \mathcal{I}(\vec{x}_0, a, \vartheta) \chi \left(\frac{(x_1 - x_{01}) \cos \vartheta + (x_2 - x_{02}) \sin \vartheta}{a}, \right. \\ &\quad \left. \frac{-(x_1 - x_{01}) \sin \vartheta + (x_2 - x_{02}) \cos \vartheta}{a} \right) \end{aligned} \quad (3.2-16)$$

die Substitution

$$\vec{k} = \begin{bmatrix} k_1 \\ k_2 \end{bmatrix} = \frac{1}{a} \begin{bmatrix} \cos \vartheta \\ \sin \vartheta \end{bmatrix}, \quad da d\vartheta = a^3 dk^2 \quad (3.2-17)$$

durch:

$$\begin{aligned} I(\vec{x}) &= C^{-1} \int_{\mathbb{R}^2} d^2k \|\vec{k}\| \int_{\mathbb{R}^2} d^2x_0 \mathcal{I}(\vec{x}_0, \vec{k}_0) \cdot \\ &\quad \chi[k_1(x_1 - x_{01}) + k_2(x_2 - x_{02}), -k_2(x_1 - x_{01}) + k_1(x_2 - x_{02})]. \end{aligned} \quad (3.2-18)$$

Mit

$$\begin{aligned} &\chi[k_1(x_1 - x_{01}) + k_2(x_2 - x_{02}), -k_2(x_1 - x_{01}) + k_1(x_2 - x_{02})] = \\ &= \frac{1}{\sigma^2} e^{-\frac{\|\vec{k}\|^2 \|\vec{x} - \vec{x}_0\|^2}{2\sigma^2}} \left(e^{-i\vec{k}^T(\vec{x} - \vec{x}_0)} - e^{-\frac{\sigma^2}{2}} \right) = \frac{1}{\|\vec{k}\|^2} \psi_{\vec{k}}(\vec{x} - \vec{x}_0) \\ &= \frac{1}{\|\vec{k}\|^2} \psi_{\vec{k}}^*(\vec{x}_0 - \vec{x}) = \frac{1}{\|\vec{k}\|^2} \psi_{\vec{x}_0, \vec{k}}^*(\vec{x}) \end{aligned} \quad (3.2-19)$$

($\psi_{\vec{x}_0, \vec{k}}(\vec{x}) = \psi_{\vec{k}}(\vec{x})$) ergibt sich daraus

$$I(\vec{x}) = C^{-1} \int_{\mathbb{R}^2} d^2k \frac{1}{\|\vec{k}\|} \int_{\mathbb{R}^2} d^2x_0 \psi_{\vec{k}}(\vec{x} - \vec{x}_0) \mathcal{I}(\vec{x}_0, \vec{k}), \quad (3.2-20)$$

wenn

$$\sigma = \tau, \quad \mathbf{S}_{\sigma, \tau} = \frac{1}{\sigma} \mathbf{I}. \quad (3.2-21)$$

In gleicher Weise kann man in (3.1-9) den Vektor \vec{k} einführen und

$$\mathcal{J}_{\vec{k}}(\vec{x}_0) = \|\vec{k}\| \mathcal{I}(\vec{x}_0, \vec{k}) \quad (3.2-22)$$

setzen. Daraus läßt sich das Transformationspaar zur Analyse und Synthese kontinuierlicher Bildsignale mit Gaborwavelets angeben:

$$\mathcal{J}_{\vec{k}}(\vec{x}_0) = \text{GWT}\{I(\vec{x})\}(\vec{x}_0, \vec{k}) = \int_{\mathbb{R}^2} d^2x I(\vec{x}) \psi_{\vec{k}}(\vec{x}_0 - \vec{x}) = I(\vec{x}_0) * \psi_{\vec{k}}(\vec{x}_0), \quad (3.2-23)$$

$$\begin{aligned} I(\vec{x}) = \text{IGWT}\{\mathcal{J}_{\vec{k}}(\vec{x}_0)\}(\vec{x}) &= C^{-1} \int_{\mathbb{R}^2} d^2k \frac{1}{\|\vec{k}\|^2} \int_{\mathbb{R}^2} d^2x_0 \mathcal{J}_{\vec{k}}(\vec{x}_0) \psi_{\vec{k}}(\vec{x} - \vec{x}_0) \\ &= C^{-1} \int_{\mathbb{R}^2} d^2k \frac{1}{\|\vec{k}\|^2} \mathcal{J}_{\vec{k}}(\vec{x}) * \psi_{\vec{k}}(\vec{x}), \end{aligned} \quad (3.2-24)$$

$$C = 4\pi^2 \int_{\mathbb{R}^2} \frac{d^2\xi}{\|\vec{\xi}\|^2} |\hat{\psi}_{\vec{\xi}_1}(\vec{\xi})|^2. \quad (3.2-25)$$

Nach der neuen Notation ist $\psi_{\vec{e}_1}(\vec{x}) = \psi_{[0]}(\vec{x})$ das Mutterwavelet und $\psi^{\vec{k}}(\vec{x}) = C^{-1}\psi_{\vec{k}}(\vec{x})$ – zusammen mit sämtlichen Translationen – die zu $\psi_{\vec{k}}(\vec{x})$ *duale Waveletfamilie*. (3.2-25) zeigt, warum (2.1-5) unabdingbar ist: Bei einem von null verschiedenen Mittelwert würde die Konstante und somit die dualen Wavelets nicht existieren. Eine Rekonstruktion wäre dann nicht möglich.

Die Koexistenz der beiden Gaborwaveletfunktionen $\chi(\cdot)$ und $\psi(\cdot)$ trägt eigentlich nicht zur Übersichtlichkeit bei. Die ψ -Notation entspricht der aus (Lades et al., 1993), die Skalen und Orientierungen werden über den Vektor \vec{k} dargestellt. Der Vorfaktor $\|\vec{k}\|^2$ sorgt dafür, daß alle Gaborkerne im Frequenzraum dieselbe Höhe haben. Demgegenüber fallen sie bei der χ -Notation für steigende Frequenz $\|\vec{\omega}\|$ aufgrund des Vorfaktors im Ortsraum $\frac{1}{a}$ ab. Desweiteren sind Orientierungen und Skalen direkt mit ϑ und a parametrisiert, sowie auch elliptische Gaborfunktionen zugelassen. Der χ -Darstellung wird meistens der Vorrang eingeräumt.

3.2.2 Alternative Methoden der Signalrekonstruktion

Um auf analytische Weise das gaborwavelettransformierte Signal zu rekonstruieren, kann an der Stelle von (3.2-2) ein alternativer Weg beschritten werden. Statt mit $2\pi w(a, \vartheta)\hat{\chi}[Q(\vartheta)\vec{\omega}a]$ multipliziert man die Gleichung mit $2\pi w(a, \vartheta, p)\hat{\chi}[Q(\vartheta)\vec{\omega}a]^p$ ($p \in \mathbb{N}_0$).

$$Y_p(\vec{\omega}) = 4\pi^2 \int_{\mathbb{R}^+} da \int_0^{2\pi} d\vartheta w(a, \vartheta, p) a \hat{\chi}[Q(\vartheta)\vec{\omega}a]^{p+1} \quad (3.2-26)$$

steht dabei in Analogie zu (3.2-5). $Y_p(\vec{\omega})$ kann auch in diesem Fall nur eine Konstante sein, wenn $w(a, \vartheta, p) = \frac{1}{a^2}$ gilt:

$$C_p = Y_p(\vec{\omega}) = 4\pi^2 \int_{\mathbb{R}^2} d^2\xi \frac{\hat{\chi}(\vec{\xi})^{p+1}}{\|\vec{\xi}\|^2}, \quad C_1 = C. \quad (3.2-27)$$

Da $0 < C_p < \infty$ gelten muß, kann (3.2-27) nicht von jeder Waveletfunktion erfüllt werden, wohl aber von Gaborwavelets, da der Integrand in (3.2-27) bei $\vec{\xi} = \vec{0}$ aufgrund der Mittelwertfreiheit von $\hat{\chi}(\vec{\xi})$ keine Singularität hat, und die Gaußfunktionen nach außen hin schnell abfallen. Analog zu (3.2-5) kann das Eingabebild nach

$$\begin{aligned} \hat{I}(\vec{\omega}) &= C_p^{-1} \int_{\mathbb{R}^2} d^2x_0 \int_{\mathbb{R}^+} \frac{da}{a^2} \int_0^{2\pi} d\vartheta \mathcal{I}(\vec{x}_0, a, \vartheta) \hat{\chi}[Q(\vartheta)\vec{\omega}a]^p e^{-i\vec{\omega}^T \vec{x}_0} \\ &= C_p^{-1} \int_{\mathbb{R}^2} d^2x_0 \int_{\mathbb{R}^+} \frac{da}{a^{2+p}} \int_0^{2\pi} d\vartheta \mathcal{I}(\vec{x}_0, a, \vartheta) \hat{\chi}_{\vec{0}, a, \vartheta}(\vec{\omega})^p e^{-i\vec{\omega}^T \vec{x}_0} \end{aligned} \quad (3.2-28)$$

⌋

$$I(\vec{x}) = \begin{cases} \frac{1}{(2\pi)^{p-1} C_p} \int_{\mathbb{R}^2} d^2x_0 \int_{\mathbb{R}^+} \frac{da}{a^{2+p}} \int_0^{2\pi} d\vartheta \mathcal{I}(\vec{x}_0, a, \vartheta) \left(\prod_{j=1}^p * \chi_{\vec{0}, a, \vartheta} \right) (\vec{x} - \vec{x}_0) & \forall p \geq 1 \\ 2\pi C_0^{-1} \int_{\mathbb{R}^2} d^2x_0 \int_{\mathbb{R}^+} \frac{da}{a^2} \int_0^{2\pi} d\vartheta \mathcal{I}(\vec{x}_0, a, \vartheta) \delta(\vec{x} - \vec{x}_0) & p = 0 \\ = 2\pi C_0^{-1} \int_{\mathbb{R}^+} \frac{da}{a^2} \int_0^{2\pi} d\vartheta \mathcal{I}(\vec{x}, a, \vartheta) & \end{cases} \quad (3.2-29)$$

wie bei (Pötzsch, 1994) wiederhergestellt werden. Warum Rekonstruktionen mit $p \neq 1$ als suboptimal anzusehen sind, wird im Unterabschnitt 4.2.3 mit Hilfe der Frame-Theorie gezeigt.

4 Frames

4.1 Kontinuierliche Wavelet-Frames

Der Nachweis der eingangs geforderten Energiebeschränktheit soll nachgeholt werden:

$$\begin{aligned}
\|\psi_{\vec{k}}\|^2 &= \int_{\mathbb{R}^2} d^2x |\psi_{\vec{k}}(\vec{x})|^2 = \\
\|\hat{\psi}_{\vec{k}}\|^2 &= \int_{\mathbb{R}^2} d^2\omega |\hat{\psi}_{\vec{k}}(\vec{\omega})|^2 \\
&= \frac{\pi \|\vec{k}\|^2}{\sigma^2} \left(1 - 2e^{-\frac{3\sigma^2}{4}} + e^{-\sigma^2}\right) < \infty \quad \Leftrightarrow \quad \|\vec{k}\| < \infty \\
\Rightarrow \psi_{\vec{k}}(\vec{x}) &\in L^2(\mathbb{R}^2), \\
&\text{bzw.} \\
\|\chi_{\vec{x}_0, a, \vartheta}\|^2 &= \|\chi\|^2 = \frac{\pi}{\sigma\tau} \left(1 - 2e^{-\frac{3\sigma^2}{4}} + e^{-\sigma^2}\right). \tag{4.1-1}
\end{aligned}$$

Zur Motivation des Framebegriffs soll der wesentliche Unterschied der Kodierung analoger Bildsignale mit kontinuierlich parametrisierten Waveletfunktionen und beispielsweise und der durch die Fouriertransformation herausgearbeitet werden. Er wird im nächsten Abschnitt verallgemeinert.

Wenn man (3.2-23) in (3.2-24) einsetzt, so erhält man

$$\begin{aligned}
I(\vec{x}) &= C^{-1} \int_{\mathbb{R}^2} d^2k \frac{1}{\|\vec{k}\|^2} \int_{\mathbb{R}^2} d^2x_0 \int_{\mathbb{R}^2} d^2x' \psi_{\vec{k}}(\vec{x} - \vec{x}_0) \psi_{\vec{k}}(\vec{x}_0 - \vec{x}') I(\vec{x}') \tag{4.1-2} \\
\Rightarrow \delta(\vec{x} - \vec{x}') &\stackrel{\wedge}{=} C^{-1} \int_{\mathbb{R}^2} d^2k \frac{1}{\|\vec{k}\|^2} \int_{\mathbb{R}^2} d^2x_0 \psi_{\vec{k}}(\vec{x} - \vec{x}_0) \psi_{\vec{k}}(\vec{x}_0 - \vec{x}') \\
&\stackrel{\wedge}{=} C^{-1} \int_{\mathbb{R}^2} d^2k \frac{1}{\|\vec{k}\|^2} \int_{\mathbb{R}^2} d^2x_0 \psi_{\vec{x}_0, \vec{k}}(\vec{x}) \psi_{\vec{x}', \vec{k}}(\vec{x}_0) \\
&\stackrel{\wedge}{=} C^{-1} \int_{\mathbb{R}^2} d^2k \frac{1}{\|\vec{k}\|^2} \int_{\mathbb{R}^2} d^2x_0 \psi_{\vec{x}_0, \vec{k}}(\vec{x}) \psi_{\vec{x}_0, \vec{k}}^*(\vec{x}'). \tag{4.1-3}
\end{aligned}$$

Über diesen Ausdruck läßt sich jedes energiebeschränkte Signal darstellen, also auch die Wavelets $\psi_{\vec{x}_1, \vec{k}'}^*(\vec{x})$:

$$\begin{aligned}
\psi_{\vec{x}_1, \vec{k}'}(\vec{x}) &= C^{-1} \int_{\mathbb{R}^2} d^2k \frac{1}{\|\vec{k}\|^2} \int_{\mathbb{R}^2} d^2x_0 \langle \psi_{\vec{x}_1, \vec{k}'}, \psi_{\vec{x}_0, \vec{k}} \rangle \psi_{\vec{x}_0, \vec{k}}(\vec{x}) \\
&= \int_{\mathbb{R}^2} d^2k \frac{1}{\|\vec{k}\|^2} \int_{\mathbb{R}^2} d^2x_0 \langle \psi_{\vec{x}_1, \vec{k}'}, \psi_{\vec{x}_0, \vec{k}} \rangle \psi_{\vec{x}_0, \vec{k}}(\vec{x}). \tag{4.1-4}
\end{aligned}$$

Die Gaborwavelets sind nicht orthogonal zu ihren dualen Funktionen, d.h.

$$\frac{1}{\|\vec{k}\|^2} \langle \psi_{\vec{x}_1, \vec{k}'}, \psi_{\vec{x}_0, \vec{k}} \rangle = C^{-1} \frac{1}{\|\vec{k}\|^2} \langle \psi_{\vec{x}_1, \vec{k}'}, \psi_{\vec{x}_0, \vec{k}} \rangle \neq \delta(\vec{x}_0 - \vec{x}_1) \delta(\vec{k} - \vec{k}'). \tag{4.1-5}$$

Das bedeutet aber auch, daß ein Wavelet durch eine Linearkombination aus anderen Wavelets aufgebaut werden kann, weshalb sie linear abhängig sind. Daher handelt es sich bei den Gaborwavelets $\psi_{\vec{x}_0, \vec{k}}(\vec{x})$ nicht um eine Basis, sondern um einen *Frame*. Die vollständige mathematische Definition ist in Abschnitt 4.2 angegeben. Für die Fouriertransformation lautet die (4.1-5) entsprechende Beziehung

$$\frac{1}{4\pi^2} \int_{\mathbb{R}^2} d^2\omega e^{i\vec{\omega}^T \vec{x}} e^{-i\vec{\omega}^T \vec{x}'} = \delta(\vec{x} - \vec{x}'), \tag{4.1-6}$$

woran sich deutlich erkennen läßt, daß eine Orthonormalbasis vorliegt.

Die Gaborwavelets sind ein Frame, wenn

$$A\|I\|^2 \leq \|\mathcal{J}\|_{\mathcal{H}_{\mathcal{J}}}^2 \leq B\|I\|^2 \quad \Leftrightarrow \tag{4.1-7}$$

$$A \int_{\mathbb{R}^2} d^2x |I(\vec{x})|^2 \leq \int_{\mathbb{R}^2} d^2k \frac{1}{\|\vec{k}\|^2} \int_{\mathbb{R}^2} d^2x_0 |\mathcal{J}_{\vec{k}}(\vec{x}_0)|^2 \leq B \int_{\mathbb{R}^2} d^2x |I(\vec{x})|^2. \quad (4.1-8)$$

Die Konstanten A und B werden *Frameschranken* genannt und erfüllen

$$0 < A \leq B < \infty \quad (4.1-9)$$

und sind mit den Grenzen in (3.2-6) identisch, was sich zeigen läßt, wenn man (3.2-6) mit $|\hat{I}(\vec{\omega})|^2$ multipliziert und über den gesamten Frequenzraum integriert. Der Index $\mathcal{H}_{\mathcal{J}}$ verdeutlicht, daß die Norm über einen anderen Hilbertraum definiert ist. Für die hier betrachteten Gaborwavelets ergibt sich dadurch

$$A = B = C^{-1}. \quad (4.1-10)$$

Ein solcher Frame wird nach (Louis et al., 1994) fest genannt (tight frame).

4.2 Allgemeine Frames

Wie im vorangegangenen Abschnitt bereits beschrieben, ist das Funktionensystem der Gaborwaveletes aufgrund linearer Abhängigkeiten keine Basis. Jedoch ist eine solche Zerlegung vollständig, da, wie in 3.2 gezeigt wurde, eine exakte Rekonstruktion möglich ist. Die Theorie der Frames wurde von Duffin und Schaeffer 1952 in (Duffin und Schaeffer, 1952) begründet. Eine stellenweise brauchbare Einführung in Frames als Mengen von Vektoren liefert (Pei und Yeh, 1997)⁴. Mit Funktionen, die hier ohne Argument angegeben sind, soll wie mit Vektoren umgegangen werden, z.B. kann das innere Produkt auch $\langle f, g \rangle = g^+ f$ geschrieben werden.

Definition 2 (Kaiser, 1994) Sei \mathcal{H} ein Hilbertraum und sei M ein Maßraum mit dem Maß μ . Ein (allgemeiner) Frame in \mathcal{H} ist ein Funktionensystem $\mathcal{H}_M = \{h_{\vec{m}} \in \mathcal{H} \mid \vec{m} \in M\}$ mit folgenden Eigenschaften.

- Für jedes $f \in \mathcal{H}$ ist die Funktion $\mathcal{F} : M \rightarrow \mathbb{C}$,

$$\mathcal{F}(\vec{m}) = \langle f, h_{\vec{m}} \rangle_{\mathcal{H}} = h_{\vec{m}}^+ f, \quad (4.2-1)$$

meßbar.

- Es existieren zwei Konstanten $0 < A \leq B < \infty$, so daß für jedes $f \in \mathcal{H}$

$$A \|f\|_{\mathcal{H}}^2 \leq \|\mathcal{F}\|_{L^2(\mu)}^2 \leq B \|f\|_{\mathcal{H}}^2 \quad (4.2-2)$$

gilt.

A und B sind die *Frameschranken*. (4.2-2) nennt man *Framebedingung*.

Der Begriff des Frames beinhaltet auch den der Basis, d.h. wenn $A = B = 1$ gilt, ist (4.2-2) die Parsevalsche Gleichung. (4.2-2) ist auch zu entnehmen, daß die beiden Räume, also der Raum des Original- und des transformierten Signals, im allgemeinen mit unterschiedlichen Maßen ausgestattet sind. Die Normen werden somit für kontinuierliches M durch

$$\|\mathcal{F}\|_{L^2(\mu)}^2 = \int_M d\mu(\vec{m}) |\mathcal{F}(\vec{m})|^2 \quad (4.2-3)$$

und für diskretes M durch

$$\|\mathcal{F}\|_{L^2(\mu)}^2 = \sum_{\vec{m} \in M} \mu_{\vec{m}} |\mathcal{F}(\vec{m})|^2 \quad (4.2-4)$$

berechnet. Die bekanntesten Frames sind Wavelets und Gaborfunktionen.

⁴Aufgrund von Fehlern an wichtigen Stellen ist Vorsicht geboten.

4.2.1 Die Frametransformation als Operator

Zur Darstellung der *Frametransformation* eignet sich der lineare Operator

$$T : \mathcal{H} \rightarrow L^2(\mu), \quad Tf = \mathcal{F}, \quad (4.2-5)$$

auch *Frameoperator*⁵ genannt. Dazu existiert der adjungierte Operator

$$T^+ : L^2(\mu) \rightarrow \mathcal{H}, \quad \langle Tf, \mathcal{G} \rangle_{L^2(\mu)} = \langle f, T^+ \mathcal{G} \rangle_{\mathcal{H}} \Leftrightarrow \mathcal{G}^+ Tf = (T^+ \mathcal{G})^+ f. \quad (4.2-6)$$

Diese Darstellung ermöglicht es, (4.2-2) in

$$\begin{aligned} A \langle f, f \rangle_{\mathcal{H}} &\leq \langle Tf, Tf \rangle_{L^2(\mu)} \leq B \langle f, f \rangle_{\mathcal{H}} \\ A \langle f, If \rangle_{\mathcal{H}} &\leq \langle f, T^+ Tf \rangle_{\mathcal{H}} \leq B \langle f, If \rangle_{\mathcal{H}} \\ Af^+ If &\leq f^+ T^+ Tf \leq Bf^+ If \\ AI &\leq T^+ T \leq BI \end{aligned}$$

$$AI \leq G \leq BI \quad (4.2-7)$$

$$B^{-1}I \leq G^{-1} \leq A^{-1}I \quad (4.2-8)$$

umzuschreiben. Die Vergleichsoperatoren bezüglich linearer Operatoren sind im Sinne ihrer Definitheit zu verstehen. (4.2-7) bezeichnet die Framebedingung in Operatorform. $G = T^+T$ ist selbstadjungiert ($G = G^+$) und tätigt die Abbildung $\mathcal{H} \rightarrow \mathcal{H}$. Er wird in (Kaiser, 1994) *metrischer Operator* genannt. Um das ursprüngliche Signal f aus seiner transformierten Form wiederherzustellen, kann der einfache Zusammenhang

$$f = G^{-1}T^+ \mathcal{F} = (T^+T)^{-1}T^+ \mathcal{F} = (T^+T)^{-1}T^+ Tf = f \quad (4.2-9)$$

benutzt werden. Wie im diskreten und endlichen Fall mit Vektoren und Matrizen findet hier eine Approximation im quadratischen Mittel statt, so daß der Rekonstruktionsoperator $(T^+T)^{-1}T^+$ völlig analog zu klassischen Kleinste-Quadrate-Problemen der Moore-Penrose-Inversen entspricht. In gleicher Weise projiziert

$$P = TG^{-1}T^+ = T(T^+T)^{-1}T^+ \quad (4.2-10)$$

sämtliche Funktionen aus $L^2(\mu)$ orthogonal auf den Bildbereich \mathcal{F} von T . Die einzigen Bedingungen, die T dabei erfüllen muß, sind Linearität und Stetigkeit.

Um diese recht formale Operatornotation mit Leben zu füllen, muß eine Verbindung zur bisher benutzten Funktionenschreibweise hergestellt werden. Aber zuvor noch einige Definitionen: Die zu $h_{\vec{m}}$ *duale Familie* sei

$$h^{\vec{m}} = G^{-1}h_{\vec{m}} \quad (4.2-11)$$

und der *duale Frame*

$$\mathcal{H}^M = \{h^{\vec{m}} \mid \vec{m} \in M\}. \quad (4.2-12)$$

Aus (4.2-9) kann man nun direkt die Vollständigkeitsrelation ablesen und zu

$$\begin{aligned} G^{-1}G = I &= \int_M d\mu(\vec{m}) h^{\vec{m}} h_{\vec{m}}^+ = G^{-1} \int_M d\mu(\vec{m}) h_{\vec{m}} h_{\vec{m}}^+ \\ \Rightarrow G &= \int_M d\mu(\vec{m}) h_{\vec{m}} h_{\vec{m}}^+. \end{aligned} \quad (4.2-13)$$

weiterverarbeiten. Der zugehörige metrische Operator für den diskreten Fall lautet

$$G = \sum_{\vec{m} \in M} \mu_{\vec{m}} h_{\vec{m}} h_{\vec{m}}^+. \quad (4.2-14)$$

⁵Die Bezeichnung „Frameoperator“ für T stammt von (Kaiser, 1994). Andere Autoren benutzen ihn für T^+T .

Auch die *inverse Frametransformation* kann in Operatorschreibweise bequem formuliert werden:

$$f = \int_M d\mu(\vec{m}) h^{\vec{m}} \mathcal{F}(\vec{m}) \quad (4.2-15)$$

$$\Leftrightarrow f(\vec{x}) = \int_M d\mu(\vec{m}) h^{\vec{m}}(\vec{x}) \mathcal{F}(\vec{m}). \quad (4.2-16)$$

Der Beweis kann einfach durch Nachrechnen erbracht werden:

$$\begin{aligned} f &= (\mathbb{T}^+ \mathbb{T})^{-1} \mathbb{T}^+ \mathcal{F} = \int_M d\mu(\vec{m}) h^{\vec{m}} \mathcal{F}(\vec{m}) \\ &= \mathbb{G}^{-1} \int_M d\mu(\vec{m}) h_{\vec{m}} \mathcal{F}(\vec{m}) \\ \Leftrightarrow \mathbb{T}^+ \mathcal{F} &= \int_M d\mu(\vec{m}) h_{\vec{m}} \mathcal{F}(\vec{m}) \\ \Rightarrow f^+ \mathbb{T}^+ \mathcal{F} &= \int_M d\mu(\vec{m}) f^+ h_{\vec{m}} \mathcal{F}(\vec{m}) \\ \Leftrightarrow (\mathbb{T}f)^+ \mathcal{F} &= \int_M d\mu(\vec{m}) \mathcal{F}^*(\vec{m}) \mathcal{F}(\vec{m}) \\ \Leftrightarrow \mathcal{F}^+ \mathcal{F} &= \int_M d\mu(\vec{m}) |\mathcal{F}(\vec{m})|^2. \end{aligned} \quad (4.2-17)$$

Für die Signalsynthese aus diskreten Frames (genauer: Die Frames sind nach ihren Indizes \vec{m} abzählbar) gilt völlig analog

$$f = \sum_{\vec{m} \in M} \mu_{\vec{m}} h^{\vec{m}} \mathcal{F}(\vec{m}) \quad (4.2-18)$$

$$\Leftrightarrow f(\vec{x}) = \sum_{\vec{m} \in M} \mu_{\vec{m}} h^{\vec{m}}(\vec{x}) \mathcal{F}(\vec{m}). \quad (4.2-19)$$

Das größte Problem bei der Signalrekonstruktion ist das Finden des inversen metrischen Operators, der für die Berechnung des dualen Frames gebraucht wird. Ist, wie im Falle endlicher und abzählbarer Vektorräume, der Frameoperator eine Matrix, deren Spalten den Framevektoren $h_{\vec{m}}$ entsprechen, kann f mit den üblichen numerischen Methoden zur Lösung von Kleinste-Quadrate-Problemen berechnet werden. Bei kontinuierlichen Wavelet-Frames ist dieser Aufwand aufgrund ihrer besonderen Struktur nicht nötig, denn der metrische Operator ist hierbei $\mathbb{G} = C = \text{konst.}$

Um die Vollständigkeitsrelation herzuleiten, setzt man (4.2-1) in (4.2-16) ein und fordert, daß die resultierende Gleichung

$$f = \int_M d\mu(\vec{m}) h^{\vec{m}} h_{\vec{m}}^+ f \quad (4.2-20)$$

für alle f gilt, was auf

$$\int_M d\mu(\vec{m}) h^{\vec{m}} h_{\vec{m}}^+ = \mathbb{I} \quad (4.2-21)$$

und analog auf

$$\sum_{\vec{m} \in M} \mu(\vec{m}) h^{\vec{m}} h_{\vec{m}}^+ = \mathbb{I} \quad (4.2-22)$$

führt. Multipliziert man (4.2-21) mit der Framefunktion $h_{\vec{m}'}$, so erhält man

$$\begin{aligned} &\int_M d\mu(\vec{m}) h_{\vec{m}} h^{\vec{m}} h_{\vec{m}'}^+ = \int_M d\mu(\vec{m}) h_{\vec{m}} K(\vec{m}, \vec{m}')^+ = h_{\vec{m}'} \\ \Rightarrow &h_{\vec{m}'}^+ f = \int_M d\mu(\vec{m}) K(\vec{m}, \vec{m}') h_{\vec{m}}^+ f \\ \Leftrightarrow &\int_M d\mu(\vec{m}) K(\vec{m}, \vec{m}') \mathcal{F}(\vec{m}) = \mathcal{F}(\vec{m}') \end{aligned} \quad (4.2-23)$$

bzw.

$$\sum_{\vec{m} \in M} \mu(\vec{m}) K(\vec{m}, \vec{m}') \mathcal{F}(\vec{m}) = \mathcal{F}(\vec{m}'). \quad (4.2-24)$$

Die Bedeutung der beiden letzten Formeln ist offensichtlich: $K(\vec{m}, \vec{m}')$ ist ein reproduzierender Kern (s. (Kaiser, 1994)), durch den jeder Funktionswert der Frametransformierten \mathcal{F} aus anderen Werten durch eine Linearkombination bestimmt werden kann. Das funktioniert nur dann nicht, wenn

$$K(\vec{m}, \vec{m}') = h_{\vec{m}}^+ h_{\vec{m}}^{\vec{m}} = \gamma \delta_0(\vec{m} - \vec{m}') \quad (4.2-25)$$

gilt, d. h., wenn $h_{\vec{m}} \forall \vec{m} \in M$ eine orthogonale Basis ist. Gilt (4.2-25) nicht, so sind die $h_{\vec{m}}$ linear abhängig und bilden daher keine Basis.

4.2.2 Die Frameschranken

Nach (4.2-7) begrenzen die Frameschranken den metrischen Operator G . Dies läßt sich am besten anhand des Rayleigh-Quotienten $R(f)$ mit den Eigenwerten von G , den Quadraten der Singulärwerte des Frameoperators T , quantifizieren. Der Rayleigh-Quotient ist für die Eigenfunktionen des Operators ein stationäres Funktional, d.h., daß die erste Variation verschwindet, so daß

$$0 < A = \lambda_{\min} \leq R_G(f) = \frac{\langle f, Gf \rangle}{\langle f, f \rangle} = \frac{f^+ G f}{f^+ f} \leq \lambda_{\max} = B < \infty, \quad (4.2-26)$$

wobei λ_{\min} und λ_{\max} jeweils der kleinste und der größte Eigenwert des Operators G sind.

Solange das Funktionensystem $h_{\vec{m}}$ ein Frame ist, kann das Originalsignal vollständig aus $\mathcal{F}(\vec{m})$ rekonstruiert werden, wobei es für die Durchführung wesentlich günstiger ist, wenn ein tight Frame vorliegt:

$$G = AI = BI \quad (4.2-27)$$

$$\Rightarrow f(\vec{x}) = \frac{1}{A} \int_M d\mu(\vec{m}) h_{\vec{m}}(\vec{x}) \mathcal{F}(\vec{m}), \quad (4.2-28)$$

$$f(\vec{x}) = \frac{1}{A} \sum_{\vec{m} \in M} \mu_{\vec{m}} h_{\vec{m}}(\vec{x}) \mathcal{F}(\vec{m}). \quad (4.2-29)$$

Der Vorteil hierbei ist, den metrischen Operator weder bestimmen noch invertieren zu müssen, um die Signalsynthese aus der Frametransformierten zu erreichen. Vor allem in der digitalen Bildverarbeitung ist ein unbekannter konstanter Vorfaktor in vielen Anwendungen unproblematisch. Dabei ist es oft ausreichend einen Frame zu benutzen, der nur näherungsweise fest ist. Dies soll mit der Größe $v = \frac{B}{A} \geq 1$ quantifiziert werden. Ist v nur unwesentlich größer als 1, benutzt man das arithmetische Mittel der Frameschranken als metrischen Operator, was zu einer Rekonstruktion

$$\tilde{f}(\vec{x}) = \frac{2}{A+B} \int_M d\mu(\vec{m}) h_{\vec{m}}(\vec{x}) \mathcal{F}(\vec{m}) \approx f, \quad (4.2-30)$$

$$\tilde{f}(\vec{x}) = \frac{2}{A+B} \sum_{\vec{m} \in M} \mu_{\vec{m}} h_{\vec{m}}(\vec{x}) \mathcal{F}(\vec{m}) \approx f \quad (4.2-31)$$

führt. Das bedeutet, daß der Frame bei der Rekonstruktion wie eine Basis behandelt werden darf, was als erhebliche Vereinfachung anzusehen ist. Die Frage, ob es sich tatsächlich lohnt, kann nur mit Hilfe eines globalen Fehlerkriteriums

$$\begin{aligned} S(v) &= \|f - \tilde{f}\|^2 \\ &= \left\| (T^+ T)^{-1} T^+ \mathcal{F} - \frac{2}{A+B} T^+ \mathcal{F} \right\|^2 \\ &= \left\| f - \frac{2}{A+B} Gf \right\|^2 \end{aligned}$$

$$\begin{aligned}
&= f^+ \left(\mathbf{I} - \frac{2}{A+B} \mathbf{G} \right) \left(\mathbf{I} - \frac{2}{A+B} \mathbf{G} \right) f \\
&= f^+ \left(\mathbf{I} - \frac{4}{A+B} \mathbf{G} + \frac{4}{(A+B)^2} \mathbf{G} \mathbf{G} \right) f
\end{aligned} \tag{4.2-32}$$

beantwortet werden. Mit Hilfe von

$$\begin{aligned}
\frac{4}{v+1} \mathbf{I} &\leq \frac{4}{A+B} \mathbf{G} \leq \frac{4v}{v+1} \mathbf{I} \\
\frac{4}{v+1} \mathbf{I} &\leq \frac{\frac{4}{A} \xi(f)}{v+1} \mathbf{I} \leq \frac{4v}{v+1} \mathbf{I}, & \forall \xi(f) \in [A, B] \\
\frac{4}{v+1} \mathbf{I} &\leq \frac{4\eta(f)}{v+1} \mathbf{I} \leq \frac{4v}{v+1} \mathbf{I}, & \forall \eta(f) = \frac{\xi(f)}{A} \in [1, v]
\end{aligned} \tag{4.2-33}$$

und

$$\begin{aligned}
\frac{4}{(v+1)^2} \mathbf{I} &\leq \frac{4}{(A+B)^2} \mathbf{G} \mathbf{G} \leq \frac{4v^2}{(v+1)^2} \mathbf{I} \\
\frac{4}{(v+1)^2} \mathbf{I} &\leq \frac{4\eta^2(f)}{(v+1)^2} \mathbf{I} \leq \frac{4v^2}{(v+1)^2} \mathbf{I} & \forall \eta(f) \in [1, v]
\end{aligned} \tag{4.2-34}$$

(aus (4.2-7)) kann eine explizite Fehlerabschätzung

$$\begin{aligned}
S_{\eta(f)}(v) &= f^+ \left(\mathbf{I} - \frac{4\eta(f)}{v+1} \mathbf{I} + \frac{4\eta^2(f)}{(v+1)^2} \mathbf{I} \right) f \\
&= \|f\|^2 \left(1 - \frac{4\eta(f)}{v+1} + \frac{4\eta^2(f)}{(v+1)^2} \right)
\end{aligned} \tag{4.2-35}$$

angegeben werden. Bei den in (4.2-33) und (4.2-34) eingeführten Ersatzoperatoren ist nur ihre Definitheit von Interesse, so daß es möglich ist, ein ganzes Fehlerintervall zu vorgegebenen Frameschranken zu bestimmen. Die klassische Extremwertrechnung liefert ein Minimum in der Intervallmitte und zwei gleich große Maxima an den Intervallrändern:

$$S_{\min}(v) = S_{\frac{1}{2}(v+1)}(v) = 0 \tag{4.2-36}$$

$$S_{\max}(v) = S_1(v) = S_v(v) = \|f\|^2 \left(\frac{v-1}{v+1} \right)^2. \tag{4.2-37}$$

(4.2-36) zeigt, daß es vorkommen kann, daß der Fehler bei einer Tight-Frame-Rekonstruktion auch verschwindet, wenn

$$\frac{f^+ \mathbf{G} f}{f^+ f} = \frac{1}{2}(A+B) \tag{4.2-38}$$

gilt. Damit kann im allgemeinen natürlich nicht gerechnet werden. Für einen echten tight Frame gilt auch hier

$$S_{\eta(f)}(1) = S_{\max}(1) = 0. \tag{4.2-39}$$

Diese Zusammenhänge sind auch in Abb. 5 dargestellt.

Allein anhand der Frameschranken kann im allgemeinen nicht bestimmt werden, ob ein Funktionensystem eine Basis darstellt oder nicht. Jedoch können sie zu folgenden Aussagen herangezogen werden, die in Tabelle 1 beschrieben sind.

Ist \mathbf{G} eine Matrix, gibt es noch einen weiteren Grund, weshalb v nahe bei 1 liegen sollte, da für die Konditionszahl von \mathbf{G} der Zusammenhang

$$1 \leq k(\mathbf{G}) = \|\mathbf{G}\| \|\mathbf{G}^{-1}\| \leq v \quad \forall v \geq 1 \tag{4.2-40}$$

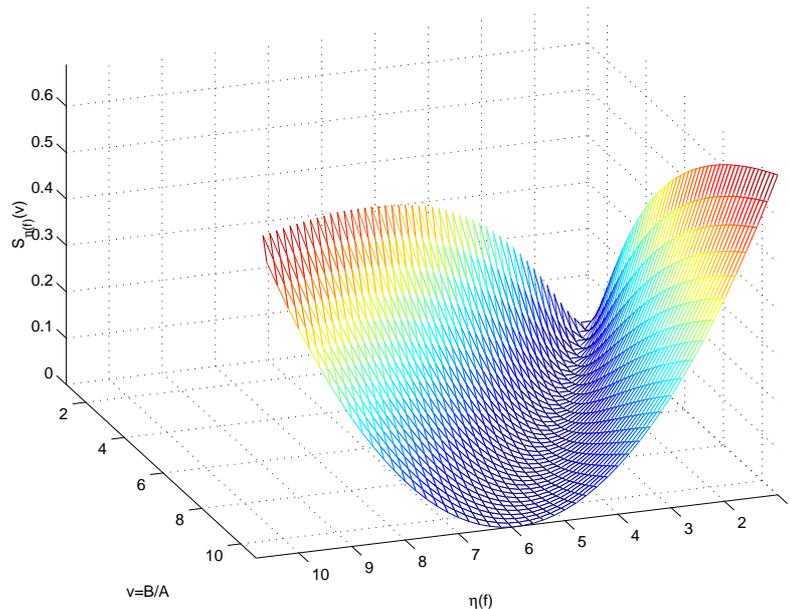


Abbildung 5: Darstellung von $S_{\eta(f)}(v)$. Die Größe $\eta(f)$ ist bestimmt durch die tatsächliche, normierte Gesamtleistung (4.2-26) eines frametransformierten Signals.

$0 < A = B = 1$	orthonormale Basis
$0 < A = B \neq 1, < \infty$	nichtnormierte orthogonale Basis oder tight frame
$0 < A < B < \infty$	nichtorthogonale Basis (exact frame, s. (Benedetto et al., 1998)) oder überhaupt keine Basis
$0 < A \approx B < \infty$	snug frame (s. (Daubechies, 1990))
$B < \infty$	Bessel-Folge (s. (Benedetto et al., 1998))

Tabelle 1: Aussagen der Frameschranken über die Vollständigkeit und einige zusätzliche Begriffe aus der Frametheorie.

gilt. Über Konditionierung von Gleichungssystemen s. (Golub und van Loan, 1996) oder (Schwarz, 1993).

Auch eine andere Interpretation der Frameschranken ist in Bezug auf die GWT möglich: Für einen Ort des Bildsignals besteht die Gaborwavelettransformierte aus mehreren (aber endlich vielen) Gaußglocken im Frequenzraum. Theoretisch haben diese keinen endlichen Träger, weshalb man das Gesamtbild mit einer einzigen Gaborantwort darstellen und auch, bis auf Frequenzen $\vec{\omega}$ auf der durch $\vec{\omega}^T \vec{k} = 0$ charakterisierten Geraden, rekonstruieren kann. Dies ist aber bei numerischer Anwendung aufgrund des schnellen Abfalls der Gaußfunktionen ein schlecht konditioniertes Problem, da nicht berücksichtigte Frequenzanteile im Quantisierungsrauschen versinken und bei Inversion große Fehler verursachen. Deshalb erfolgt die Phasenraumabtastung mit mehreren Gaborfiltern auf verschiedenen Frequenzpositionen. Werden nun deren Betragsquadrate wie bei einer Leistungsmessung aufsummiert, ergeben die untere bzw. obere Grenze dieser Hügelandschaft die Frameschranken. Sie liegen weiter auseinander, wenn das Bildsignal mit nur wenigen Gaborwavelets kodiert wird.

Diese Beispiele zeigen, daß es sehr hilfreich sein kann, die Frameschranken zu kennen, wengleich deren Abschätzung mit einem gewissen Aufwand verbunden ist.

4.2.3 Rekonstruktionsmöglichkeiten

Da bei einer Frametransformierten im allgemeinen von einer übervollständigen Darstellung auszugehen ist, hat man beliebig viele Möglichkeiten, durch eine Linearkombination wie in (4.2-16)

oder (4.2-19) das ursprüngliche Signal wiederzugewinnen, was bezüglich der Gaborwavelets schon in Abschnitt 3.2.2 thematisiert worden ist. In diesem idealen Fall rekonstruieren alle beschriebenen Möglichkeiten für $p \geq 0$ das Bildsignal $I(\vec{x})$ exakt. Die Lösung mit $p = 1$ entspricht der Rekonstruktion mit der Pseudoinversen (Moore-Penrose-Inversen, s. (Penrose, 1955)) $f = T^\dagger \mathcal{F}$. Ist die Frametransformierte Störungen unterworfen, so kann die Rekonstruktion nur approximativ erfolgen und die beste Approximation liefert T^\dagger , (Penrose, 1956):

$$\|Tf - \mathcal{F}\| \geq \|T(T^\dagger \mathcal{F}) - \mathcal{F}\|. \quad (4.2-41)$$

Nimmt man im Rahmen eines stochastischen Modells die Störung n als additiv, erwartungswertfrei und stationär an,

$$\mathcal{F} = Tf + n, \quad (4.2-42)$$

so kann das Ursprungssignal aus den transformierten und verrauschten Daten durch

$$\tilde{f} = T^\dagger \mathcal{F} = (T^+ T)^{-1} T^+ \mathcal{F} \quad (4.2-43)$$

geschätzt werden. Der Kleinste-Quadrate-Schätzer hat sehr günstige Eigenschaften, wie Erwartungstreue und Wirksamkeit (s. (Böhme, 1993)).

5 Die phasenraumdiskrete Gaborwavellettransformation

5.1 Abtastung des Phasenraums

Bei der kontinuierlichen Gaborwavellettransformation wird ein Bildsignal aus \mathbb{R}^2 in \mathbb{R}^4 transformiert, d.h. die Transformierte ist in vier Dimensionen kontinuierlich. In diesem Abschnitt geht es darum, inwieweit kontinuierliche Bilddaten durch eine abzählbare Anzahl von Translationen, Skalen und Orientierungen, die zusammengefaßt *Phasenraum* genannt werden, darstellbar sind, und welche Fehler dabei gemacht werden.

Um den theoretischen Unterbau für (2.2-13) zu liefern, und in dem Zusammenhang eine Rekonstruktionsformel herzuleiten, soll nun die betrachtete Abtastung des Phasenraums schrittweise durchgeführt werden. Ziel ist es, Translationen, Skalierungen und Drehungen von $\chi(\vec{x})$ zu diskretisieren:

$$\vec{x}_0 = \vec{n}_0 \Delta = a \vec{n}_0 \Delta_0 = a_0^m \vec{n}_0 \Delta_0, \quad (5.1-1)$$

$$a = a_0^m, a_0 > 1, \quad (5.1-2)$$

$$\vartheta = \vartheta_l = \frac{2\pi l}{L} \quad (5.1-3)$$

Die Abtastung von \vec{x}_0 ist nur dann exakt richtig, wenn $\chi_{\vec{x}_0, a, \vartheta}(\vec{x})$ bandbegrenzt sind, als was die Gaborfunktionen aufgrund der Schwellwerte aus Unterabschnitt 2.2.5 angenommen werden.

In einem ersten Schritt werden die Translationen diskretisiert, wobei $\chi_{\vec{x}_0, a, \vartheta}(\vec{0})$ näherungsweise als bandbegrenzt angenommen werden soll. Innerhalb eines Frequenzbandes gilt daher analog zu (3.2-1)

$$\mathcal{I}(\vec{n}_0 a \Delta_0, a, \vartheta) = \int_{\mathbb{R}^2} d^2 \omega \hat{I}(\vec{\omega}) a \hat{\chi}^* [Q(\vartheta) \vec{\omega} a] e^{i \vec{\omega}^T \vec{n}_0 a \Delta_0}. \quad (5.1-4)$$

Da die Translationen im Ortsraum diskretisiert und im Frequenzraum periodisch fortgesetzt sind, wird auf (5.1-5) die IDSFT angewandt, anschließend mit $\frac{2\pi}{a} \chi[Q(\vartheta) \vec{\omega} a]$ multipliziert und über a und ϑ integriert:

$$\frac{1}{4\pi^2} a^2 \Delta_0^2 \sum_{\vec{n}_0 \in \mathbb{Z}^2} \mathcal{I}(\vec{n}_0 a \Delta_0, a, \vartheta) e^{-i \vec{\omega}^T \vec{n}_0 a \Delta_0} = \hat{I}(\vec{\omega}) a \hat{\chi}^* [Q(\vartheta) \vec{\omega} a] \quad (5.1-5)$$

$$\Rightarrow \hat{\chi}[Q(\vartheta) \vec{\omega} a] \Delta_0^2 \sum_{\vec{n}_0 \in \mathbb{Z}^2} \mathcal{I}(\vec{n}_0 a \Delta_0, a, \vartheta) e^{-i \vec{\omega}^T \vec{n}_0 a \Delta_0} = \frac{4\pi^2}{a} \hat{I}(\vec{\omega}) |\hat{\chi}[Q(\vartheta) \vec{\omega} a]|^2 \quad (5.1-6)$$

$$\begin{aligned}
&\Rightarrow \Delta_0^2 \int_{\mathbb{R}^+} da \int_0^{2\pi} d\vartheta \hat{\chi}[\mathbf{Q}(\vartheta)\vec{a}a] \cdot \\
&\quad \cdot \sum_{\vec{n}_0 \in \mathbb{Z}^2} \mathcal{I}(\vec{n}_0 a \Delta_0, a, \vartheta) e^{-i\vec{\omega}^T \vec{n}_0 a \Delta_0} = \hat{I}(\vec{\omega}) 4\pi^2 \int_{\mathbb{R}^+} \frac{da}{a} \int_0^{2\pi} d\vartheta |\hat{\chi}[\mathbf{Q}(\vartheta)\vec{a}a]|^2 \\
&\quad = \hat{I}(\vec{\omega}) Y(\vec{\omega}) \tag{5.1-7}
\end{aligned}$$

Aus den Aussagen aus 3.2 folgt direkt

$$\begin{aligned}
\hat{I}(\vec{\omega}) &= Y(\vec{\omega})^{-1} \int_{\mathbb{R}^+} da \int_0^{2\pi} d\vartheta \sum_{\vec{n}_0 \in \mathbb{Z}^2} \mathcal{I}(\vec{n}_0 a \Delta_0, a, \vartheta) \hat{\chi}[\mathbf{Q}(\vartheta)\vec{a}a] e^{-i\vec{\omega}^T \vec{n}_0 a \Delta_0} \\
&= Y(\vec{\omega})^{-1} \int_{\mathbb{R}^+} \frac{da}{a} \int_0^{2\pi} d\vartheta \sum_{\vec{n}_0 \in \mathbb{Z}^2} \mathcal{I}(\vec{n}_0 a \Delta_0, a, \vartheta) \hat{\chi}_{\vec{n}_0 a \Delta_0}(\vec{\omega})(\vec{\omega}) \\
&= \int_{\mathbb{R}^+} \frac{da}{a} \int_0^{2\pi} d\vartheta \sum_{\vec{n}_0 \in \mathbb{Z}^2} \mathcal{I}(\vec{n}_0 a \Delta_0, a, \vartheta) \hat{\chi}^{\vec{n}_0 a \Delta_0}(\vec{\omega})(\vec{\omega}). \tag{5.1-8}
\end{aligned}$$

Bei der zusätzlichen Diskretisierung von Skalen und Orientierungen stellt man die Integration für $Y(\vec{\omega})$ als Grenzwerte Riemannscher Näherungssummen

$$\begin{aligned}
Y(\vec{\omega}) &= 4\pi^2 \lim_{a_0 \rightarrow 1^+} \lim_{L \rightarrow \infty} \frac{2\pi}{L} (a_0 - 1) \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \left| \hat{\chi} \left[\mathbf{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \right|^2 \\
&= \lim_{a_0 \rightarrow 1^+} \lim_{L \rightarrow \infty} Y'_{a_0, L}(\vec{\omega}), \\
Y'_{a_0, L}(\vec{\omega}) &= \frac{8\pi^3}{L} (a_0 - 1) \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \left| \hat{\chi} \left[\mathbf{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \right|^2 \\
&= \frac{8\pi^3}{L} (a_0 - 1) Y_{a_0, L}(\vec{\omega}) \tag{5.1-9}
\end{aligned}$$

dar, wobei $Y_{a_0, L}$ gegenüber $Y'_{a_0, L}$ auf eine einfachere Darstellung bei der Rekonstruktion führt. Dies hat darüberhinaus eine Reskalierung der Frameschranken zur Folge:

$$A_{a_0, L} = \inf_{\vec{\omega} \in \mathbb{R}^2} Y_{a_0, L}(\vec{\omega}) = \frac{8\pi^3}{L} (a_0 - 1) A'_{a_0, L}, \tag{5.1-10}$$

$$B_{a_0, L} = \sup_{\vec{\omega} \in \mathbb{R}^2} Y_{a_0, L}(\vec{\omega}) = \frac{8\pi^3}{L} (a_0 - 1) B'_{a_0, L}. \tag{5.1-11}$$

Der Quotienten der Frameschranken $v_{a_0, L}$, die wichtigste Kenngröße eines Frames, bleibt davon natürlich unberührt.

Um die Nomenklatur noch einigermaßen übersichtlich zu halten, soll hier die Notation

$$X_{\vec{n}_0, m, l}(\vec{x}) = \chi_{\vec{n}_0 a_0^m \Delta_0, a_0^m, \frac{2\pi l}{L}}(\vec{x}) = a_0^{-m} \chi \left[\mathbf{Q} \left(\frac{2\pi l}{L} \right) (a_0^{-m} \vec{x} - \vec{n}_0 \Delta_0) \right] \tag{5.1-12}$$

⇕

$$\hat{X}_{\vec{n}_0, m, l}(\vec{\omega}) = \hat{\chi}_{\vec{n}_0 a_0^m \Delta_0, a_0^m, \frac{2\pi l}{L}}(\vec{\omega}) = a_0^m \hat{\chi} \left[\mathbf{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] e^{-i\vec{\omega}^T \vec{n}_0 \Delta_0 a_0^m} \tag{5.1-13}$$

für die im Phasenraum abgetasteten Wavelets eingeführt werden. Im folgenden werden sämtliche Integrationen durch die Trapezformel approximiert⁶. In (5.1-8) werden nun analog zu (5.1-9) die Integrationen durch ihre Riemannschen Näherungen ersetzt. Die Rekonstruktion des Originalsignals

⁶Man hätte auch von vornherein diskret rechnen können

$I(\vec{x})$ im Frequenzraum ist daher

$$\begin{aligned}\hat{I}(\vec{\omega}) &= Y_{a_0,L}^{-1}(\vec{\omega}) \frac{\Delta_0^2}{4\pi^2} \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \mathcal{I} \left(\vec{n}_0 a_0^m \Delta_0, a_0^m, \frac{2\pi l}{L} \right) \hat{X}_{\vec{n}_0,m,l}(\vec{\omega}) \\ &= \frac{\Delta_0^2}{4\pi^2} \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \mathcal{I} \left(\vec{n}_0 a_0^m \Delta_0, a_0^m, \frac{2\pi l}{L} \right) \hat{X}_{\vec{n}_0,m,l}(\vec{\omega}).\end{aligned}\quad (5.1-14)$$

Die Funktionen $X^{\vec{n}_0,m,l}(\vec{x}) = Y_{a_0,L}(\vec{\omega}) X_{\vec{n}_0,m,l}(\vec{x})$ bilden den zu $X_{\vec{n}_0,m,l}(\vec{x})$ dualen Frame. Das Transformationspaar für die phasenraumdiskrete Gaborwvetelettransformation lautet somit

$$\begin{aligned}\mathcal{I}(\vec{n}_0, m, l) &= \text{DPSGWT}\{I(\vec{x})\}(\vec{n}_0, m, l) \\ &= \langle I, X_{\vec{n}_0,m,l} \rangle = \int_{\mathbb{R}^2} d^2x I(\vec{x}) X_{\vec{n}_0,m,l}^*(\vec{x}),\end{aligned}\quad (5.1-15)$$

$$\begin{aligned}I(\vec{x}) &= \text{IDPSGWT}\{\mathcal{I}(\vec{n}_0, m, l)\}(\vec{x}) \\ &= \frac{\Delta_0^2}{4\pi^2} \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \mathcal{I}(\vec{n}_0, m, l) X^{\vec{n}_0,m,l}(\vec{x}).\end{aligned}\quad (5.1-16)$$

Um einen Ausdruck zu gewinnen, der die klassische Darstellung der Gaborwavelets $\psi(\cdot)$ enthält, ersetzt man formal die Variablen m und l durch

$$\vec{\kappa}_{m,l} = a_0^{-m} \begin{bmatrix} \cos \frac{2\pi l}{L} \\ \sin \frac{2\pi l}{L} \end{bmatrix} \quad (5.1-17)$$

($\|\vec{\kappa}_m\| = \|\vec{\kappa}_{m,l}\|$):

$$\begin{aligned}\hat{X}_{\vec{n}_0,m,l}(\vec{\omega}) &= a_0^m \hat{\chi} \left[\mathbf{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] e^{-i\vec{\omega}^T \vec{n}_0 \Delta_0 a_0^m} \\ &= a_0^m \left(e^{-\frac{\sigma^2}{2} \left[(a_0^m \omega_1 \cos \frac{2\pi l}{L} + a_0^m \omega_2 \sin \frac{2\pi l}{L} - 1)^2 + (-a_0^m \omega_1 \sin \frac{2\pi l}{L} + a_0^m \omega_2 \cos \frac{2\pi l}{L})^2 \right]} \right. \\ &\quad \left. - e^{-\frac{\sigma^2}{2} (a_0^{2m} \|\vec{\omega}\|^2 + 1)} \right) e^{-i\vec{\omega}^T \vec{n}_0 \Delta_0 a_0^m} \\ &= a_0^m \left(e^{-\frac{\sigma^2 a_0^{2m}}{2} (\|\vec{\omega}\|^2 - 2a_0^{-m} \omega_1 \cos \frac{2\pi l}{L} - 2a_0^{-m} \omega_2 \sin \frac{2\pi l}{L} + a_0^{-2m})} \right. \\ &\quad \left. - e^{-\frac{\sigma^2}{2} (a_0^m \|\vec{\omega}\|^2 + 1)} \right) e^{-i\vec{\omega}^T \vec{n}_0 \Delta_0 a_0^m} \\ &= a_0^m \left(e^{-\frac{\sigma^2 (\|\vec{\omega}\|^2 - 2\vec{\kappa}_{m,l}^T \vec{\omega} + \|\vec{\kappa}_{m,l}\|^2)}{2\|\vec{\kappa}_m\|^2}} - e^{-\frac{\sigma^2 (\|\vec{\omega}\|^2 + \|\vec{\kappa}_m\|^2)}{2\|\vec{\kappa}_m\|^2}} \right) e^{-i\frac{\vec{\omega}^T \vec{n}_0 \Delta_0}{\|\vec{\kappa}_m\|}} \\ &= a_0^m \left(e^{-\frac{\sigma^2 (\|\vec{\omega} - \vec{\kappa}_{m,l}\|^2)}{2\|\vec{\kappa}_m\|^2}} - e^{-\frac{\sigma^2 (\|\vec{\omega}\|^2 + \|\vec{\kappa}_m\|^2)}{2\|\vec{\kappa}_m\|^2}} \right) e^{-i\frac{\vec{\omega}^T \vec{n}_0 \Delta_0}{\|\vec{\kappa}_m\|}} \\ &= \frac{1}{\|\vec{\kappa}_m\|} \hat{\psi}_{\vec{\kappa}_{m,l}}(\vec{\omega}) e^{-i\frac{\vec{\omega}^T \vec{n}_0 \Delta_0}{\|\vec{\kappa}_m\|}}\end{aligned}\quad (5.1-18)$$

Wenn man wieder $\|\vec{\kappa}_m\| \mathcal{I}$ durch \mathcal{J} ersetzt, können

$$\begin{aligned}\hat{I}(\vec{\omega}) &= Y_{a_0,L}^{-1}(\vec{\omega}) \frac{\Delta_0^2}{4\pi^2} \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \frac{1}{\|\vec{\kappa}_m\|^2} \mathcal{J}_{\vec{\kappa}_{m,l}}(\vec{n}_0) \hat{\psi}_{\vec{\kappa}_{m,l}}(\vec{\omega}) e^{-i\frac{\vec{\omega}^T \vec{n}_0 \Delta_0}{\|\vec{\kappa}_m\|}} \\ &= \frac{\Delta_0^2}{4\pi^2} \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \frac{1}{\|\vec{\kappa}_m\|^2} \mathcal{J}_{\vec{\kappa}_{m,l}}(\vec{n}_0) \hat{\psi}_{\vec{\kappa}_{m,l}}(\vec{\omega}) e^{-i\frac{\vec{\omega}^T \vec{n}_0 \Delta_0}{\|\vec{\kappa}_m\|}}\end{aligned}\quad (5.1-19)$$

mit

$$\hat{\psi}^{\vec{k}_{m,l}}(\vec{\omega}) = Y_{a_0,l}^{-1}(\vec{\omega}) \hat{\psi}_{\vec{k}_{m,l}}(\vec{\omega}), \quad (5.1-20)$$

(5.1-15) und (5.1-16) in

$$\mathcal{J}_{\vec{k}_{m,l}}(\vec{n}_0) = \text{DPSGWT}\{I(\vec{x})\}(\vec{n}_0, \vec{k}_{m,l}) = \int_{\mathbb{R}^2} d^2x I(\vec{x}) \psi_{\vec{k}_{m,l}}\left(\vec{n}_0 \frac{\Delta_0}{\|\vec{k}_m\|} - \vec{x}\right), \quad (5.1-21)$$

$$\begin{aligned} I(\vec{x}) &= \text{IDPSGWT}\{\mathcal{J}_{\vec{k}_{m,l}}(\vec{n}_0)\}(\vec{x}) \\ &= \frac{\Delta_0^2}{4\pi^2} \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \frac{1}{\|\vec{k}_m\|^2} \mathcal{J}_{\vec{k}_{m,l}}(\vec{n}_0) \psi^{\vec{k}_{m,l}}\left(\vec{x} - \vec{n}_0 \frac{\Delta_0}{\|\vec{k}_m\|}\right) \end{aligned} \quad (5.1-22)$$

umgeschrieben werden.

5.2 Abschätzung der Frameschranken

(4.1-7) und (4.1-8) geben eine Bedingung an, unter der die Gaborwavelets einen Frame bilden. D.h.:

- Das aus dem Mutterwavelet gebildete Funktionensystem ist vollständig in dem Sinne, daß jedes $L^2(\mathbb{R}^2)$ -Signal durch eine Linearkombination aus Gaborwavelets dargestellt werden kann.
- Das Bildsignal kann exakt aus einem abgetasteten Phasenraum rekonstruiert werden.

Um sich für die weitere Rechnung Schreibarbeit zu ersparen, wird die Abkürzung $\mu = \frac{\Delta_0^2}{4\pi^2}$ benutzt. Für die beiden Notationen für Gaborwavelets (ψ und χ) bedeutet diesbezüglich der phasenraumdiskreten Gaborwavelettransformation, daß

$$A\|I\|^2 \leq \mu \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} |\langle I, X_{\vec{n}_0,m,l} \rangle|^2 \leq B\|I\|^2 \quad (5.2-1)$$

$$\|I\|^2 = \int_{\mathbb{R}^2} d^2x |I(\vec{x})|^2. \quad (5.2-2)$$

Bei den weiteren Untersuchungen, die von (Lee, 1996) motiviert sind, geht man am besten von (5.2-1) aus, da in der χ -Notation die Skalierung leichter abzulesen ist. Die obigen Gleichungen zeigen, daß die Wahl der Notation keinen Einfluß auf die Frameschranken hat.

$$\begin{aligned} &\sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \mu |\langle I, X_{\vec{n}_0,m,l} \rangle|^2 = \\ &= \mu \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} |\langle \hat{I}, \hat{X}_{\vec{n}_0,m,l} \rangle|^2 = \mu \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} |\langle \hat{I}, \hat{\chi}_{\vec{n}_0 a_0^m \Delta_0, a_0^m, \frac{2\pi l}{L}} \rangle|^2 \\ &= \mu \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \left| \int_{\mathbb{R}^2} d^2\omega I(\vec{\omega}) a_0^m \hat{\chi}^* \left[\mathbf{Q} \left(\frac{2\pi l}{L} \vec{\omega} a_0^m \right) \right] e^{i\vec{\omega}^T \vec{n}_0 \Delta_0 a_0^m} \right|^2 \\ &= \mu \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} a_0^{2m} \left| \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega_1 \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega_2 e^{i\vec{\omega}^T \vec{n}_0 \Delta_0 a_0^m} \right. \\ &\quad \cdot \left. \sum_{\vec{j} \in \mathbb{Z}^2} \hat{I}(\vec{\omega} - 2\pi \vec{j} a_0^{-1} \Delta_0^{-1}) \hat{\chi}^* \left[\mathbf{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega} a_0^m + 2\pi \vec{j} \Delta_0^{-1}) \right] \right|^2 \\ &= \mu \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} a_0^{2m} \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega_1 \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega_2 \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega'_1 \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega'_2 e^{i(\vec{\omega} - \vec{\omega}')^T \vec{n}_0 \Delta_0 a_0^m}. \end{aligned}$$

$$\begin{aligned}
& \cdot \sum_{\vec{j} \in \mathbb{Z}^2} \sum_{\vec{j}' \in \mathbb{Z}^2} \hat{I}(\vec{\omega} + 2\pi\vec{j}a_0^{-m}\Delta_0^{-1}) \hat{I}^*(\vec{\omega}' + 2\pi\vec{j}'a_0^{-m}\Delta_0^{-1}) \cdot \\
& \cdot \hat{\chi}^* \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega} a_0^m + 2\pi\vec{j}\Delta_0^{-1}) \right] \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega}' a_0^m + 2\pi\vec{j}'\Delta_0^{-1}) \right] \\
= & 4\pi^2 \mu \Delta_0^{-2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega_1 \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega_2 \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega'_1 \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega'_2 \delta(\vec{\omega} - \vec{\omega}') \cdot \\
& \cdot \sum_{\vec{j} \in \mathbb{Z}^2} \sum_{\vec{j}' \in \mathbb{Z}^2} \hat{I}(\vec{\omega} + 2\pi\vec{j}a_0^{-m}\Delta_0^{-1}) \hat{I}^*(\vec{\omega}' + 2\pi\vec{j}'a_0^{-m}\Delta_0^{-1}) \cdot \\
& \cdot \hat{\chi}^* \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega} a_0^m + 2\pi\vec{j}\Delta_0^{-1}) \right] \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega}' a_0^m + 2\pi\vec{j}'\Delta_0^{-1}) \right] \\
= & 4\pi^2 \Delta_0^{-2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega_1 \int_0^{\frac{2\pi}{a_0^m \Delta_0}} d\omega_2 \sum_{\vec{j} \in \mathbb{Z}^2} \sum_{\vec{j}' \in \mathbb{Z}^2} \hat{I}(\vec{\omega} + 2\pi\vec{j}a_0^{-m}\Delta_0^{-1}) \cdot \\
& \cdot \hat{I}^*(\vec{\omega}' + 2\pi\vec{j}'a_0^{-m}\Delta_0^{-1}) \hat{\chi}^* \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega} a_0^m + 2\pi\vec{j}\Delta_0^{-1}) \right] \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega}' a_0^m + 2\pi\vec{j}'\Delta_0^{-1}) \right] \\
= & 4\pi^2 \mu \Delta_0^{-2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \sum_{\vec{j} \in \mathbb{Z}^2} \sum_{\vec{j}' \in \mathbb{Z}^2} \int_{j_1 \frac{2\pi}{a_0^m \Delta_0}}^{(j_1+1) \frac{2\pi}{a_0^m \Delta_0}} d\omega'_1 \int_{j_2 \frac{2\pi}{a_0^m \Delta_0}}^{(j_2+1) \frac{2\pi}{a_0^m \Delta_0}} d\omega'_2 \hat{I}(\vec{\omega}') \cdot \\
& \cdot \hat{I}^*[\vec{\omega}' + 2\pi(\vec{j}' - \vec{j})a_0^{-m}\Delta_0^{-1}] \hat{\chi}^* \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega}' a_0^m \right] \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega}' a_0^m + 2\pi(\vec{j}' - \vec{j})\Delta_0^{-1}) \right] \\
= & 4\pi^2 \mu \Delta_0^{-2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \sum_{\vec{j} \in \mathbb{Z}^2} \sum_{\vec{j}'' \in \mathbb{Z}^2} \int_{j_1 \frac{2\pi}{a_0^m \Delta_0}}^{(j_1+1) \frac{2\pi}{a_0^m \Delta_0}} d\omega_1 \int_{j_2 \frac{2\pi}{a_0^m \Delta_0}}^{(j_2+1) \frac{2\pi}{a_0^m \Delta_0}} d\omega_2 \hat{I}(\vec{\omega}) \cdot \\
& \cdot \hat{I}^*[\vec{\omega} + 2\pi\vec{j}''a_0^{-m}\Delta_0^{-1}] \hat{\chi}^* \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega} a_0^m + 2\pi\vec{j}''\Delta_0^{-1}) \right] \\
= & 4\pi^2 \mu \Delta_0^{-2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \sum_{\vec{j} \in \mathbb{Z}^2} \int_{\mathbb{R}^2} d^2\omega \hat{I}(\vec{\omega}) \hat{I}^*(\vec{\omega} + 2\pi\vec{j}a_0^{-1}\Delta_0^{-1}) \cdot \\
& \cdot \hat{\chi}^* \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega} a_0^m + 2\pi\vec{j}\Delta_0^{-1}) \right]. \tag{5.2-3}
\end{aligned}$$

Betrachtet man in der Summe über alle Translationen nur den Term mit $\vec{j} = \vec{0}$, dann ist

$$\begin{aligned}
P &= 4\pi^2 \Delta_0^{-2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \int_{\mathbb{R}^2} d^2\omega \left| \hat{I}(\vec{\omega}) \right|^2 \left| \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \right|^2 \\
&= 4\pi^2 \Delta_0^{-2} \int_{\mathbb{R}^2} d^2\omega \left| \hat{I}(\vec{\omega}) \right|^2 \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \left| \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \right|^2 \\
&= \int_{\mathbb{R}^2} d^2\omega \left| \hat{I}(\vec{\omega}) \right|^2 Y_{a_0, L}(\vec{\omega}). \tag{5.2-4}
\end{aligned}$$

$Y_{a_0, L}$ ist selbstähnlich („skalenerperiodisch“), d.h. es gilt

$$Y_{a_0, L}(a_0^l \vec{\omega}) = Y_{a_0, L}(\vec{\omega}). \tag{5.2-5}$$

Möchte man P für die Frameschrankenberechnung einsetzen, sind die Abschätzungen

$$P \leq \int_{\mathbb{R}^2} d^2\omega \left| \hat{I}(\vec{\omega}) \right|^2 \sup_{\vec{\omega}} Y_{a_0, L}(\vec{\omega}) = \|I\|^2 \sup_{\vec{\omega} \in \mathcal{S}} Y_{a_0, L}(\vec{\omega}) \tag{5.2-6}$$

$$P \geq \int_{\mathbb{R}^2} d^2\omega \left| \hat{I}(\vec{\omega}) \right|^2 \inf_{\vec{\omega}} Y_{a_0, L}(\vec{\omega}) = \|I\|^2 \inf_{\vec{\omega} \in \mathcal{S}} Y_{a_0, L}(\vec{\omega}) \tag{5.2-7}$$

nützlich. Aufgrund der Skalenperiodizität und der Periodizität des Orientierungswinkels brauchen die Extrema von $Y_{a_0,L}$ nicht im ganzen Frequenzraum, sondern nur im *Fundamentalsektor* \mathcal{S} gesucht zu werden. Einen Eindruck der Wirkungsweise der Konstanten a_0 und L gibt Abb. 6. Werden die Wavelets tatsächlich als bandbegrenzt angenommen, fallen in (5.2-3) alle Summenglieder mit $\vec{j} \neq \vec{0}$ weg, dann gilt nämlich

$$\begin{aligned} \|I\|^2 \inf_{\vec{\omega} \in \mathcal{S}} Y_{a_0,L}(\vec{\omega}) \leq P &= \sum_{\vec{n}_0 \in \mathbb{Z}^2} \sum_{m \in \mathbb{Z}} \sum_{l=0}^{L-1} \mu |\langle I, X_{\vec{n}_0,m,l} \rangle|^2 \leq \|I\|^2 \sup_{\vec{\omega} \in \mathcal{S}} Y_{a_0,L}(\vec{\omega}) \\ &\Rightarrow A = \inf_{\vec{\omega}} Y_{a_0,L}(\vec{\omega}), \quad B = \sup_{\vec{\omega} \in \mathcal{S}} Y_{a_0,L}(\vec{\omega}). \end{aligned} \quad (5.2-8)$$

Es ist wichtig anzumerken, daß diese Ergebnisse nicht vom Vorfaktor a^{qm} ($q \in \mathbb{Z}$) abhängen, sondern dessen Einflüsse sich durch das entsprechend andere Maß, das für (5.2-1) vorzusehen ist, aufheben. Im Grunde sind Gaborwavelets nicht bandbegrenzt, so müssen nach einem in (Lee, 1996) angegebenen Verfahren die Frameschranken nach außen korrigiert werden. Eine allgemeinere, auch für nicht bandbegrenzte Wavelets einsetzbare Methode zur Abschätzung der Frameschranken von Oktavwavelets hat I. Daubechies in (Daubechies, 1990) angegeben, eine Verallgemeinerung auf zweidimensionale und mehrstimmige Wavelets wurde von T. S. Lee in (Lee, 1996) vorgestellt. Die Anzahl der Stimmen Ξ ist die der Skalen pro Oktave. Die Frameschranken ergeben sich in der in dieser Arbeit benutzten Nomenklatur als

$$A = 4\pi^2 \Delta_0^{-2} \left[\inf_{\vec{\omega}} \sum_{\xi=0}^{\Xi-1} \sum_{\vec{m} \in \mathbb{Z}} \sum_{l=0}^{L-1} \frac{1}{2} \left(\left| \hat{\chi}^\xi \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \right|^2 + \left| \hat{\chi}^\xi \left[\mathbb{Q} \left(-\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \right|^2 \right) - \tilde{R} \right] \quad (5.2-9)$$

$$B = 4\pi^2 \Delta_0^{-2} \left[\sup_{\vec{\omega}} \sum_{\xi=0}^{\Xi-1} \sum_{\vec{m} \in \mathbb{Z}} \sum_{l=0}^{L-1} \frac{1}{2} \left(\left| \hat{\chi}^\xi \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \right|^2 + \left| \hat{\chi}^\xi \left[\mathbb{Q} \left(-\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \right|^2 \right) + \tilde{R} \right] \quad (5.2-10)$$

mit

$$\tilde{R} = \sum_{\epsilon \in \{+1, -1\}} \sum_{\xi=0}^{\Xi-1} \sum_{\vec{p} \in \mathbb{Z} \setminus \vec{0}} \sqrt{\beta_\epsilon^\xi \left(\frac{2\pi \vec{p}}{\Delta_0} \right) \beta_\epsilon^\xi \left(-\frac{2\pi \vec{p}}{\Delta_0} \right)}, \quad (5.2-11)$$

$$\begin{aligned} \beta_\epsilon^\xi(\vec{q}) &= \frac{1}{4} \sup_{\vec{\omega}} \sum_{\vec{\omega}} \sum_{l=0}^{L-1} \left| \hat{\chi}^\xi \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] + \epsilon \hat{\chi}^\xi \left[-\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] \right| \\ &\cdot \left| \hat{\chi}^\xi \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega} a_0^m + \vec{q}) \right] + \epsilon \hat{\chi}^\xi \left[-\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{\omega} a_0^m - \vec{q}) \right] \right| \end{aligned} \quad (5.2-12)$$

und

$$\hat{\chi}^\xi(\vec{\omega}) = \hat{\chi} \left(2^{-\frac{\xi}{\Xi}} \vec{\omega} \right). \quad (5.2-13)$$

Aufgrund der Periodizität in polarer und der Skalenperiodizität in radialer Richtung, kann die Suche nach den Suprema und Infima im Frequenzraum auf einen endlichen Bereich, in (Lee, 1996) Fundamentalsektor genannt, beschränkt werden. Abb. 6 veranschaulicht die Bedeutung der Frameschranken für die Qualität der Rekonstruktion.

6 Die diskrete Gaborwavelettransformation

6.1 Abtastung des Ortsraums bei diskretem Phasenraum

Ziel dieses Abschnittes ist der Entwurf numerischer Verfahren, die den Einsatz der bereits diskutierten Transformationen auf Digitalrechnern ermöglichen. Das bedeutet, daß nun auch alle restlichen Variablen und Parameter diskretisiert und auf eine Menge endlich vieler Elemente beschränkt werden müssen. Zu tun ist also folgendes:

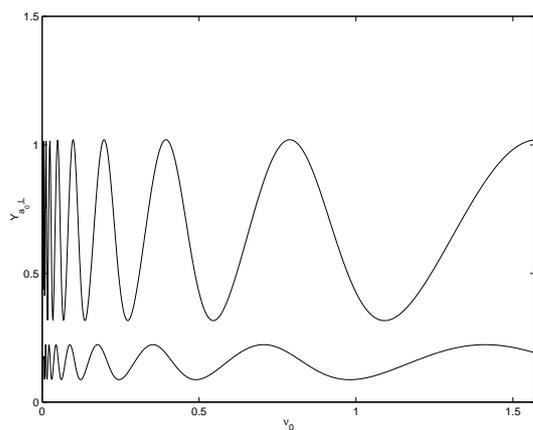
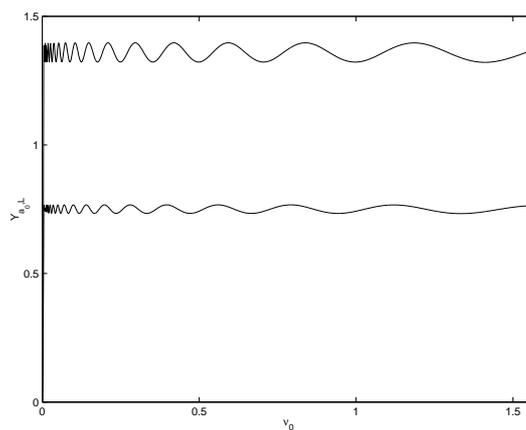
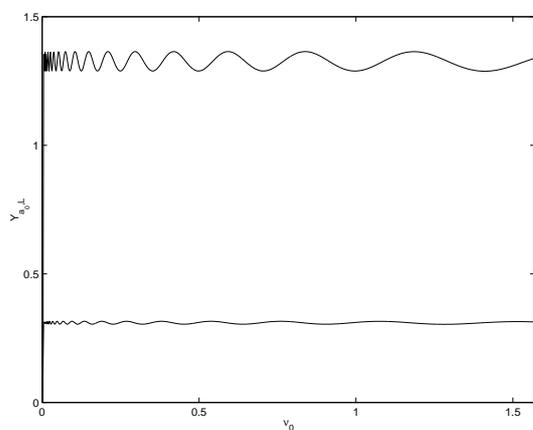
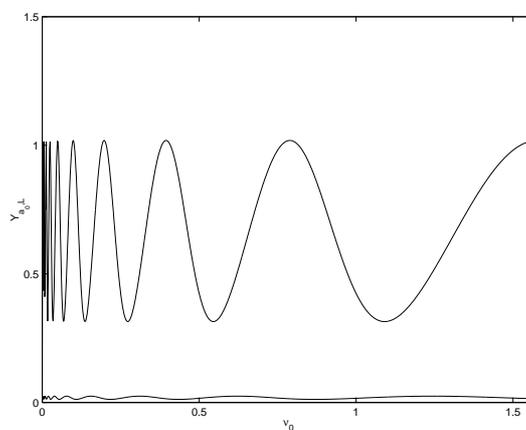
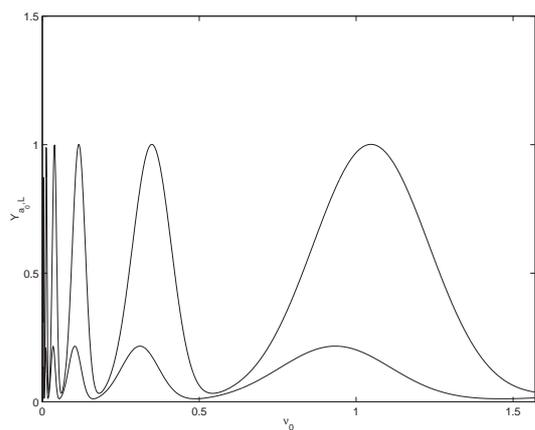
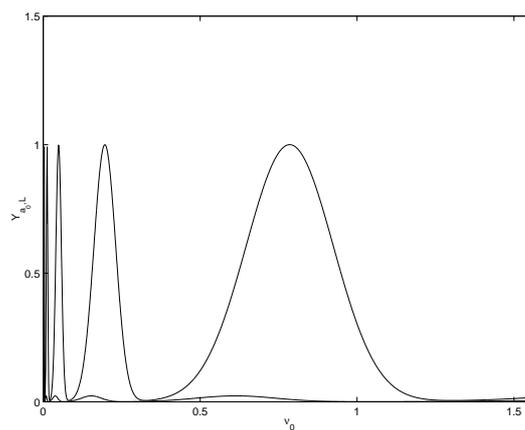
(a) $a_0 = 2, L = 12$ (b) $a_0 = \sqrt{2}, L = 16$ (c) $a_0 = \sqrt{2}, L = 12$ (d) $a_0 = 2, L = 8$ (e) $a_0 = 3, L = 12$ (f) $a_0 = 4, L = 8$

Abbildung 6: Dargestellt sind jeweils $Y_{a_0, L}[\nu_0 \vec{e}(0)]$, die $Y_{a_0, L}(\vec{\nu})$ nach oben, und $Y_{a_0, L}[\nu_0 \vec{e}(\frac{\pi}{L})]$, die $Y_{a_0, L}(\vec{\nu})$ nach unten begrenzt. Werden die Gaborwavelets als bandbegrenzt im Translationsraum angenommen, sind Infimum und Supremum der entsprechenden Funktion die Frameschranken A und B . Es wurden $\sigma = 4$ und $\tau = 6$ gewählt.

- Abtastung des Ortsraums,
- Abtastung des Frequenzraums, da im Ortsraum ein endlicher Träger vorliegt, d.h. Nutzbarkeit der DFT,
- Festlegen der *Frequenzebenen*⁷: Gaußlocken, die auf hohen Frequenzebenen (kleinen Skalen) liegen, werden an der Grenze des Nyquistintervalls einfach abgeschnitten. Somit kann strenggenommen von keiner Waveletfamilie mehr die Rede sein, da die Transformationskerne nicht mehr exakt durch Skalierung, Drehung und Translation auseinander hervorgehen.

Man sollte daher nur noch von Frames sprechen, obwohl dem nur im geringen Maße Rechnung getragen wird. Darüber hinaus werden Eigenschaften dieses Frames diskutiert.

6.2 Skalenunabhängige Abtastung des Ortsraums und der Translationen

Obwohl es nicht sehr ökonomisch ist wird die Abtastung von Orts- und Translationsraum nicht an die Skala angepaßt. Die entsprechenden ortskontinuierlichen Gaborwavelets lauten daher

$$\chi_{\vec{n}_0\Delta, a_0^m, \frac{2\pi l}{L}}(\vec{x}) = a_0^{-m} \chi \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{x} - \vec{n}_0\Delta) a_0^{-m} \right] \quad (6.2-1)$$

$$\begin{aligned} & \Downarrow \\ \hat{\chi}_{\vec{n}_0\Delta, a_0^m, \frac{2\pi l}{L}}(\vec{\omega}) &= a_0^m \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\omega} a_0^m \right] e^{-i\vec{\omega}^T \vec{n}_0\Delta}. \end{aligned} \quad (6.2-2)$$

Diese Wavelets werden nun nach $\vec{x} = \vec{n}\Delta$ abgetastet und mit der DFT in den Frequenzraum transformiert:

$$\chi_{\vec{n}_0\Delta, a_0^m, \frac{2\pi l}{L}}(\vec{n}\Delta) = a_0^{-m} \chi \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{n} - \vec{n}_0)\Delta a_0^{-m} \right] \quad (6.2-3)$$

$$\begin{aligned} & \Downarrow \\ \check{\chi}_{\vec{n}_0\Delta, a_0^m, \frac{2\pi l}{L}}(\vec{\rho}) &= \frac{2\pi}{\sqrt{N_1 N_2}} \frac{1}{\Delta^2} \hat{\chi}_{\vec{n}_0\Delta, a_0^m, \frac{2\pi l}{L}} \left(\frac{2\pi \vec{\rho}}{\Delta} \right) \\ &= \frac{2\pi}{\sqrt{N_1 N_2}} \frac{1}{\Delta^2} a_0^{-m} \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] e^{-i2\pi \vec{\rho}^T \vec{n}_0}. \end{aligned} \quad (6.2-4)$$

Der hier benutzte Hilfsvektor für den Frequenzraum bedeutet

$$\vec{\rho} = \begin{bmatrix} \frac{\rho_1}{N_1} \\ \frac{\rho_2}{N_2} \end{bmatrix}. \quad (6.2-5)$$

Dementsprechend kann völlig analog zu vorhergehenden Kapiteln die Transformation und die Rekonstruktion angegeben werden:

$$\begin{aligned} \mathcal{I}_0(\vec{n}_0, m, l) &= \sum_{\vec{n} \in \mathcal{B}} I(\vec{n}\Delta) \chi_{\vec{n}_0\Delta, a_0^m, \frac{2\pi l}{L}}^*(\vec{n}\Delta) \\ &= \sum_{\vec{n} \in \mathcal{B}} I(\vec{n}\Delta) a_0^{-m} \chi^* \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) (\vec{n} - \vec{n}_0)\Delta a_0^{-m} \right] \\ &= \frac{2\pi}{\sqrt{N_1 N_2}} \frac{1}{\Delta^2} \sum_{\vec{\rho} \in \mathcal{B}} \check{I}(\vec{\rho}) a_0^m \hat{\chi}^* \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] e^{i2\pi \vec{\rho}^T \vec{n}_0} \\ &\Leftrightarrow \frac{1}{\sqrt{N_1 N_2}} \sum_{\vec{n}_0 \in \mathcal{B}} \mathcal{I}_0(\vec{n}_0, m, l) e^{-i2\pi \vec{\rho}^T \vec{n}_0} = \frac{2\pi}{\Delta^2} a_0^m \check{I}(\vec{\rho}) \hat{\chi}^* \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] \end{aligned} \quad (6.2-6)$$

⁷Eine Frequenzebene entspricht einer reziproken Skala. Wird der Skalen-Orientierungsraum mit \vec{k} parametrisiert, ist eher von der Frequenzebene die Rede.

$$\begin{aligned} \Rightarrow & \frac{1}{\sqrt{N_1 N_2}} \frac{\Delta^2}{2\pi} a_0^{-m} \sum_{\vec{n}_0 \in \mathcal{B}} \mathcal{I}_0(\vec{n}_0, m, l) e^{-i2\pi \vec{\rho}^T \vec{n}_0} \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] \\ & = \check{I}(\vec{\rho}) \left| \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] \right|^2. \end{aligned} \quad (6.2-7)$$

Nach Summation über alle Skalen und Orientierungen ergibt sich daraus

$$\begin{aligned} \check{I}(\vec{\rho}) & = Y_{a_0, L}^{-1}(\vec{\rho}) \frac{1}{\sqrt{N_1 N_2}} \frac{\Delta^2}{2\pi} \sum_{m=m_0}^{M_0} \sum_{l=0}^{L-1} \sum_{\vec{n}_0 \in \mathcal{B}} \mathcal{I}_0(\vec{n}_0, m, l) a_0^{-m} \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] e^{-i2\pi \vec{\rho}^T \vec{n}_0} \\ & = \frac{\Delta^4}{4\pi^2} Y_{a_0, L}^{-1}(\vec{\rho}) \sum_{m=m_0}^{M_0} \sum_{l=0}^{L-1} \sum_{\vec{n}_0 \in \mathcal{B}} a_0^{-2m} \mathcal{I}_0(\vec{n}_0, m, l) \check{\chi}_{\vec{n}_0 \Delta, a_0^m, \frac{2\pi l}{L}}(\vec{\rho}) \\ & = Y_{a_0, L}^{-1}(\vec{\rho}) \frac{\Delta^2}{2\pi} \sum_{m=m_0}^{M_0} \sum_{l=0}^{L-1} \check{\mathcal{I}}_0(\vec{\rho}, m, l) a_0^{-m} \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] \end{aligned} \quad (6.2-8)$$

$$Y_{a_0, L}(\vec{\rho}) = \sum_{m=m_0}^{M_0} \sum_{l=0}^{L-1} \left| \hat{\chi} \left[\mathbb{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] \right|^2 \quad (6.2-9)$$

$$\Rightarrow I(\vec{n} \Delta) = \frac{\Delta^4}{4\pi^2} \sum_{m=m_0}^{M_0} \sum_{l=0}^{L-1} \sum_{\vec{n}_0 \in \mathcal{B}} a_0^{-2m} \mathcal{I}_0(\vec{n}_0, m, l) \chi^{\vec{n}_0 \Delta, a_0^m, \frac{2\pi l}{L}}(\vec{n} \Delta). \quad (6.2-10)$$

Somit wurde auch für den „ganz diskreten“ Fall eine Rekonstruktionsformel gewonnen. Der Abschnitt 6.3.1 verdeutlicht, daß die Benutzung dieser Inversionen nicht ganz unproblematisch ist.

In den bisherigen Herleitungen ist immer noch das räumliche Abtastintervall Δ vorhanden. Da der Mensch Bilder im Gegensatz zu akustischen Signalen auf verschiedenen Skalen wahrnehmen kann, spielt der Wert von Δ kaum eine Rolle. Mann könnte daher einfach $\Delta = 1$ setzen, aber in dieser Arbeit wird es um der Allgemeinheit willen bis zum Ende beibehalten.

6.3 DGWT und IDGWT in Vektornotation

Die Pixel des digitalisierten Bildsignals werden nun zeilenweise in einen Vektor geordnet. Die diskrete Gaborwvlettransformation kann als lineare Operation durch eine Matrix dargestellt werden, die wiederum aus den Matrizen der einzelnen Gaborfilter besteht. Da die Gabortransformation Faltungcharakter hat, haben diese Matrizen Block-Zirkulant-Struktur. Vektoren in höherdimensionalen Räumen werden hier mit einem Unterstrich gekennzeichnet.

$$\begin{aligned} \underline{\mathcal{I}}_0 & = \mathbf{X} \underline{\mathcal{I}} = \mathbf{A} \underline{\Psi} \underline{\mathcal{I}} \\ \Rightarrow \underline{\mathcal{I}} & = \mathbf{A}^{-1} \underline{\mathcal{I}}_0 = \underline{\Psi} \underline{\mathcal{I}} \end{aligned} \quad (6.3-1)$$

$$\mathbf{X} = \begin{bmatrix} \mathbf{X}_{m_0, 0} \\ \mathbf{X}_{m_0, 1} \\ \vdots \\ \mathbf{X}_{m, l} \\ \vdots \\ \mathbf{X}_{M_0, L-1} \end{bmatrix}, \quad \underline{\Psi} = \begin{bmatrix} \underline{\Psi}_{m_0, 0} \\ \underline{\Psi}_{m_0, 1} \\ \vdots \\ \underline{\Psi}_{m, l} \\ \vdots \\ \underline{\Psi}_{M_0, L-1} \end{bmatrix} \quad (6.3-2)$$

$$\underline{\mathcal{I}}_0 \in \mathbb{C}^{N_1 N_2 M L}, \quad \mathbf{X} \in \mathbb{C}^{N_1 N_2 M L \times N_1 N_2}$$

$$\mathbf{X}_{m, l} \in \mathbb{C}^{N_1 N_2 \times N_1 N_2}, \quad \underline{\mathcal{I}} \in \mathbb{R}^{N_1 N_2}$$

$$\mathbf{A} = \begin{bmatrix} a_0^{-2m_0} \mathbf{I} & \mathbf{0} & & & \\ \mathbf{0} & a_0^{-2m_0-2} \mathbf{I} & & & \\ & & \ddots & & \\ \mathbf{0} & & & \ddots & \\ & & & & a_0^{-2M_0} \mathbf{I} \end{bmatrix} \in \mathbb{R}^{N_1 N_2 M L \times N_1 N_2 M L} \quad (6.3-3)$$

$$(6.3-4)$$

Da bei der Wavelettransformation Faltungen ausgeführt werden, müssen die Bilder und Punktantworten der Gaborfilter auf das Format $N_1 N_2$ mit Nullen aufgefüllt werden („zero padding“). Im folgenden soll nun jedes einzelne Ausgabesignal der Gaborfilterbank diskret fouriertransformiert werden:

$$\mathbf{V}\underline{\mathcal{J}} = \mathbf{V}\Psi\mathbf{W}^+\mathbf{W}\underline{\mathcal{I}}, \quad (6.3-5)$$

$$\mathbf{V} = \begin{bmatrix} \mathbf{W} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{W} & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{W} \end{bmatrix} \in \mathbb{C}^{N_1 N_2 M L \times N_1 N_2 M L},$$

$$\begin{aligned} \mathbf{V}^+\mathbf{V} &= \mathbf{I} \\ \mathbf{W} &\in \mathbb{C}^{N_1 N_2 \times N_1 N_2}, \quad \mathbf{W}^+\mathbf{W} = \mathbf{I} \end{aligned} \quad (6.3-6)$$

$$\Rightarrow \underline{\hat{\mathcal{J}}} = \hat{\Psi}\hat{\underline{\mathcal{I}}} \quad (6.3-7)$$

$\hat{\Psi} = \mathbf{V}\Psi\mathbf{W}^+$ besteht aus Matrizen

$$\hat{\Psi}_{m,l} = \mathbf{W}\Psi_{m,l}\mathbf{W}^+ \quad (6.3-8)$$

($\mathbf{W}^+ = \mathbf{W}^{*T}$), die durch die Fouriermatrix \mathbf{W} ($\hat{\underline{\mathcal{I}}} = \mathbf{W}\underline{\mathcal{I}}$) transformiert werden. Wie schon angedeutet, haben die $\Psi_{m,l}$ Block-Zirkulant-Struktur, da sie auf die Bildvektoren als lineare verschiebungsinvariante (LSI-) Systeme wirken. Da monofrequente komplexe Exponentialschwingungen, d.h. die Fourierkerne $e^{-i\tilde{\omega}_0^T \tilde{x}}$, Eigenfunktionen von LSI-Systemen sind, sind die Spalten von \mathbf{W} die Eigenvektoren zu Block-Zirkulant-Matrizen. Bei den $\hat{\Psi}_{m,l}$ sind daher nur die Diagonalelemente besetzt, und zwar mit den Eigenwerten. Bei diesen handelt es sich um den Vektor der konjugiert komplexen Übertragungsfunktion, aus deren Punktantwort $\Psi_{m,l}$ gebildet wurde.

Bei (6.3-7) handelt es sich um ein überbestimmtes Gleichungssystem, das im Sinne kleinster Quadrate mit der *Pseudoinversen* $\hat{\Psi}^\dagger$ (*Moore-Penrose-Inverse*, (Penrose, 1955)) invertiert werden kann:

$$\hat{\underline{\mathcal{I}}} = \left(\hat{\Psi}^+\hat{\Psi}\right)^{-1} \hat{\Psi}^+\hat{\underline{\mathcal{J}}} = \hat{\Psi}^\dagger\hat{\underline{\mathcal{J}}}. \quad (6.3-9)$$

Aufgrund der besonderen Struktur ist noch eine Vereinfachung möglich:

$$\hat{\Psi}^+\hat{\Psi} = \sum_{m=m_0}^{M_0} \sum_{l=0}^{L-1} \hat{\Psi}_{m,l}^+ \hat{\Psi}_{m,l} = \mathbf{Y}. \quad (6.3-10)$$

Da die $\hat{\Psi}_{m,l} = \hat{\Psi}_{m,l}^+$ Diagonalmatrizen sind, trifft dies ebenso auf \mathbf{Y} zu. Die inverse diskrete Wavelettransformation läßt sich daher als

$$\begin{aligned} \Rightarrow \hat{\underline{\mathcal{I}}} &= \mathbf{Y}^{-1}\hat{\Psi}^+\hat{\underline{\mathcal{J}}} \\ \underline{\mathcal{I}} &= \mathbf{W}^+\hat{\underline{\mathcal{I}}} = \mathbf{W}^+\mathbf{Y}^{-1}\mathbf{W}\mathbf{W}^+\hat{\Psi}^+\mathbf{V}\mathbf{V}^+\hat{\underline{\mathcal{J}}} \\ &= \mathbf{W}^+\mathbf{Y}^{-1}\mathbf{W}\Psi^+\underline{\mathcal{J}} \\ &= \mathbf{W}^+\mathbf{Y}^{-1}\mathbf{W}\mathbf{X}^+(\mathbf{A}\mathbf{A})^{-1}\underline{\mathcal{I}}_0 \end{aligned} \quad (6.3-11)$$

formulieren.

Die Parallelen zu den Betrachtungen aus 4.2 sind unverkennbar, sie beziehen sich allerdings auf den Frequenzraum:

$$\begin{aligned} \hat{\Psi} &\hat{=} \mathbf{T} \\ \hat{\underline{\mathcal{I}}} &\hat{=} \mathbf{f} \\ \hat{\underline{\mathcal{J}}} &\hat{=} \mathcal{F} \\ \mathbf{Y} &\hat{=} \mathbf{G} \end{aligned} \quad (6.3-12)$$

Diese Vergleiche ermöglichen die anschauliche Einbindung der DGWT in die allgemeine Frame-Theorie (s. 4.2). Die zentrale Problematik bei der Anwendung wird im folgenden Abschnitt diskutiert.

6.3.1 Die Inversion des Metrischen Operators

Bei der kontinuierlichen Wavelettransformation aus 3.2 ist der metrische Operator gleich dem Einheitsoperator mit einer multiplikativen Konstanten (CI). Mit ihm wird in der diskreten Form die Matrix \mathbf{Y} identifiziert. Diese ist aber nur auf der Diagonalen mit von null verschiedenen Elementen besetzt und kann daher ebenfalls leicht invertiert werden. Diese Diagonale besteht aus der zeilenweise geordneten Funktion $Y_{a_0,L}(\vec{\rho})$ (s. 6.2-9). Notwendige Bedingung für die Zulässigkeit einer Funktion als Wavelet ist deren Mittelwertfreiheit, weshalb für $Y_{a_0,L}(\vec{\omega}) = Y_{a_0,L}(\vec{\omega}a_0^k) \forall \vec{\omega} \neq \vec{0}$ (s. 5.1-9), wenn man unendlich viele Skalen und Orientierungen vorsieht, bei $\vec{\omega} = \vec{0}$ unstetig ist. Bei den jeweiligen Herleitungen der IGWT, der IDPSGWT und der IDGWT ist die Inversion dieser Nullstelle unproblematisch, da die Gaborkerne und die transformierten Bilddaten bezüglich ihrer Translationen ebenfalls mittelwertfrei sind. So gelingt es zumindest theoretisch, auch den Mittelwert zu rekonstruieren, obwohl er bei der Gabortransformation – scheinbar – vernichtet wird.

In der Literatur wird im phasenraumdiskreten und ortsraumkontinuierlichen Fall bei bandbegrenzten Wavelets die untere Frameschranke als essentielles Infimum $A = \text{ess inf } Y_{a_0,L}(\vec{\omega})$, das Mengen vom Maß Null ignoriert, angegeben. Ansonsten sind die Frameschranken nur für $\vec{\omega} \neq \vec{0}$ zu suchen.

Problematisch wird aber die Inversion der DGWT in Form eines numerischen Algorithmus. Die Matrix $\hat{\Psi}$, die die Frequenzraumdarstellung der benutzten Gaborwavelets in Form von Diagonalmatrizen enthält, hat eine Spalte (die der Mittelwerte), deren Einträge allesamt null sind. Deshalb ist $\hat{\Psi}$ nicht mehr spaltenregulär. Der Mittelwert kann aus $\hat{\mathcal{J}}$ nicht mehr rekonstruiert werden, das zu lösende Gleichungssystem ist diesbezüglich unterbestimmt. Der Lösungsraum ist somit eine Gerade im $\mathbb{C}^{N_1 N_2}$. Ψ ist somit auch kein Frameoperator mehr. Hierin liegt offensichtlich ein Paradoxon: Die Mittelwertfreiheit garantiert gemeinsam mit einem schnellen Abfallen nach außen die Zulässigkeit einer kontinuierlichen Funktion als Wavelet, zugleich sorgt sie bei Diskretisierung für eine Verletzung der Framebedingung, da hier $A = 0$ gilt. Ein solches Funktionensystem, für dessen obere Frameschranke $B < \infty$ gilt, heißt nach (Benedetto et al., 1998) *Bessel-Folge*.

6.3.2 Die Singulärwertzerlegung von Ψ

Eine Möglichkeit bei der Inversion unterbestimmter Systeme ist die Entscheidung für die Lösung mit der kleinsten Euklidischen Norm, d.h. mit der geringsten Signalenergie. Genau das leistet die *Singulärwertzerlegung* bei der Berechnung der Moore-Penrose-Inversen.

Hierbei findet eine Zerlegung der Modellmatrix $\hat{\Psi} \in \mathbb{C}^{MLN_1 N_2 \times N_1 N_2}$, die sämtliche Gaborkerne enthält, in drei Matrizen statt:

$$\hat{\Psi} = \bar{\mathbf{U}} \bar{\mathbf{S}} \bar{\mathbf{V}}^+ \quad (6.3-13)$$

Die Matrizen $\bar{\mathbf{U}} \in \mathbb{C}^{MLN_1 N_2 \times MLN_1 N_2}$ und $\bar{\mathbf{V}} \in \mathbb{C}^{N_1 N_2 \times N_1 N_2}$ sind unitär, d.h. es gilt $\bar{\mathbf{U}}^+ \bar{\mathbf{U}} = \mathbf{I}$ und $\bar{\mathbf{V}}^+ \bar{\mathbf{V}} = \mathbf{I}$. $\bar{\mathbf{S}} \in \mathbb{C}^{MLN_1 N_2 \times N_1 N_2}$ ist folgendermaßen aufgebaut:

$$\begin{aligned} \bar{\mathbf{S}} &= \begin{bmatrix} \mathbf{S} \\ \mathbf{0} \end{bmatrix}, \\ \mathbf{S} &= \text{diag}(s_{j_1}, s_{j_2}, \dots, s_{j_i}, \dots, s_{j_{N_1 N_2}}) \in \mathbb{R}^{N_1 N_2 \times N_1 N_2}, \\ \mathbf{0} &\in \mathbb{R}^{(ML-1)N_1 N_2 \times N_1 N_2}. \end{aligned} \quad (6.3-14)$$

Die Singulärwerte sind nichtnegativ und können als die Quadratwurzeln der Eigenwerte von $\mathbf{Y} = \hat{\Psi}^+ \hat{\Psi}$ berechnet werden:

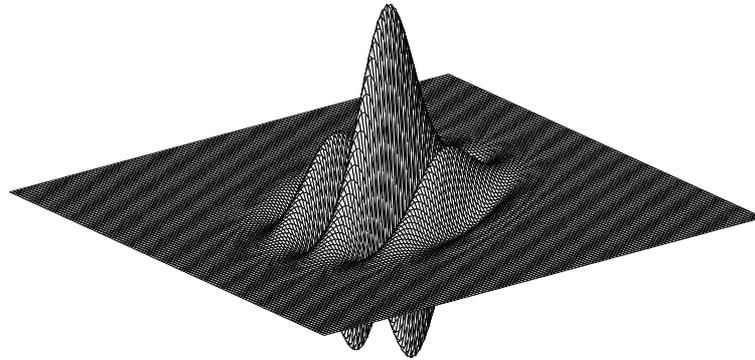
$$\mathbf{Y} \underline{x}_i = \hat{\Psi}^+ \hat{\Psi} \underline{x}_i = \lambda_i \underline{x}_i, \quad (6.3-15)$$

$$s_i = \sqrt{\lambda_i}. \quad (6.3-16)$$

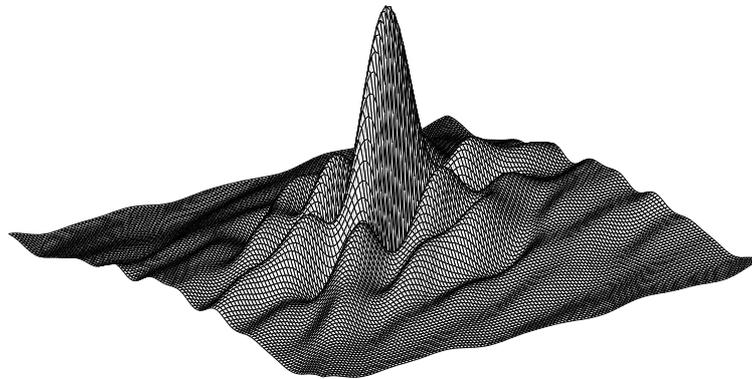
Ist $\hat{\Psi}$ nicht spaltenregulär, werden die kleinsten Singulärwerte 0.

Wendet man die Singulärwertzerlegung auf das Rekonstruktionsproblem an,

$$\hat{\underline{I}} = \left(\hat{\Psi}^+ \hat{\Psi} \right)^{-1} \hat{\Psi}^+ \hat{\underline{J}} = \hat{\Psi}^+ \hat{\underline{J}} = \bar{\mathbf{V}} \bar{\mathbf{S}}^+ \bar{\mathbf{U}}^+ \hat{\underline{J}} = \bar{\mathbf{V}} \begin{bmatrix} \mathbf{S}^+ & \mathbf{0}^+ \end{bmatrix} \bar{\mathbf{U}}^+ \hat{\underline{J}}, \quad (6.3-17)$$



(a) Realteil eines Gaborwavelets



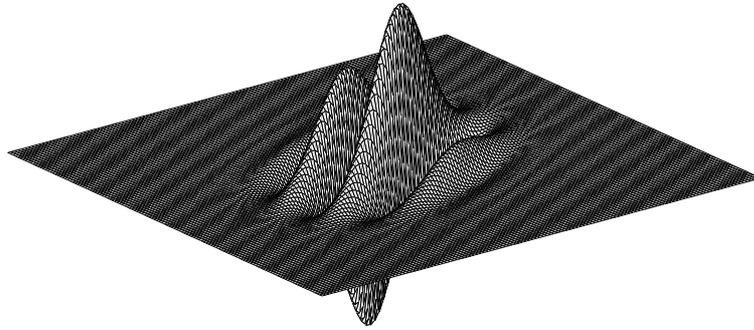
(b) Realteil des zugehörigen dualen Gaborwavelets

Abbildung 7: Realteil von Gaborwavelets zur Analyse und Synthese

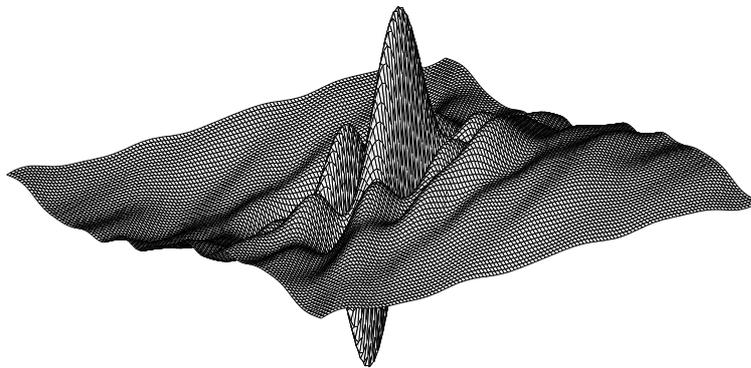
(\hat{I}_0 bezeichnet das rekonstruierte Bild) hat man die Berechnung der Moore-Penrose-Inversen von $\hat{\Psi}$ auf die von $\bar{\mathbf{S}}$ zurückgeführt. Definiert man sie als

$$\mathbf{S}^\dagger = \lim_{\epsilon \rightarrow 0} (\mathbf{S}^+ \mathbf{S} + \epsilon \mathbf{I})^{-1} \mathbf{S}^+, \quad (6.3-18)$$

so kann leicht nachgewiesen werden, daß \mathbf{S}^\dagger auf der Diagonalen mit den inversen Singulärwerten besetzt ist. Die, die gleich null sind, haben auch in der Moore-Penrose-Inversen den Wert null. Daraus wird gefolgert, daß die Moore-Penrose-Inverse immer existiert, auch dann, wenn $\hat{\Psi}$ nicht



(a) Imaginärteil eines Gaborwavelets



(b) Imaginärteil des zugehörigen dualen Gaborwavelets

Abbildung 8: *Imaginärteil von Gaborwavelets zur Analyse und Synthese*

den vollen Rang hat.

Aufgrund der besonderen Struktur der Transformationsmatrix $\hat{\Psi}$ und der Mittelwertfreiheit der Wavelets gilt $\mathbf{Y} = \mathbf{S}^+ \mathbf{S} = \text{diag}(s_1, \dots, s_{\min} = 0, \dots, s_{N_1 N_2})$. Für die weitere Verarbeitung von (6.3-17) wird $\bar{\mathbf{U}}$ nach

$$\bar{\mathbf{U}} = [\bar{\mathbf{U}}_1 \quad \bar{\mathbf{U}}_2] \quad (6.3-19)$$

in gleicher Weise wie $\bar{\mathbf{S}}$ partitioniert und eingesetzt:

$$\hat{\underline{\mathbf{I}}}_0 = \bar{\mathbf{V}} \mathbf{S}^+ \bar{\mathbf{U}}_1^+ \hat{\underline{\mathbf{J}}} = \bar{\mathbf{V}} \mathbf{Y}^{-1} \mathbf{S}^+ \bar{\mathbf{U}}_1^+ \hat{\underline{\mathbf{J}}} = \mathbf{Y}^{-1} \hat{\Psi}^+ \hat{\underline{\mathbf{J}}}. \quad (6.3-20)$$

Die unitären Transformationsmatrizen können aus (6.3-20) durch Vergleich gefunden werden:

$$\bar{\mathbf{V}} = \mathbf{I}, \quad (6.3-21)$$

$$\bar{\mathbf{U}}_1 \mathbf{S} = \hat{\Psi}, \quad \Rightarrow \quad \bar{\mathbf{U}}_1 = \hat{\Psi} \mathbf{S}^+ (\mathbf{S} \mathbf{S}^+)^{-1}. \quad (6.3-22)$$

Dabei muß $\bar{\mathbf{U}}_2$ noch die Bedingungen

$$\bar{\mathbf{U}}_2^+ \bar{\mathbf{U}}_2 = \mathbf{I}, \quad \bar{\mathbf{U}}_1^+ \bar{\mathbf{U}}_2 = \mathbf{0} \quad (6.3-23)$$

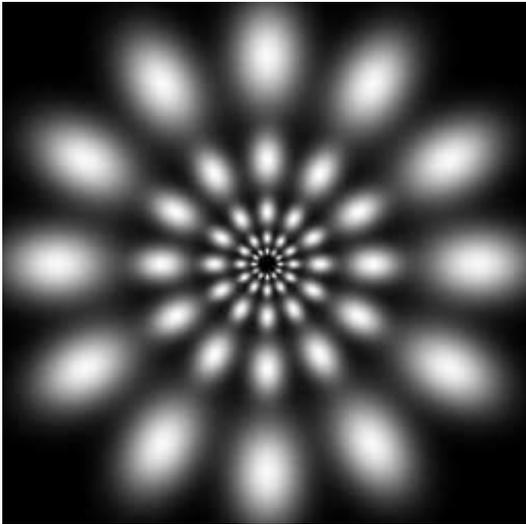
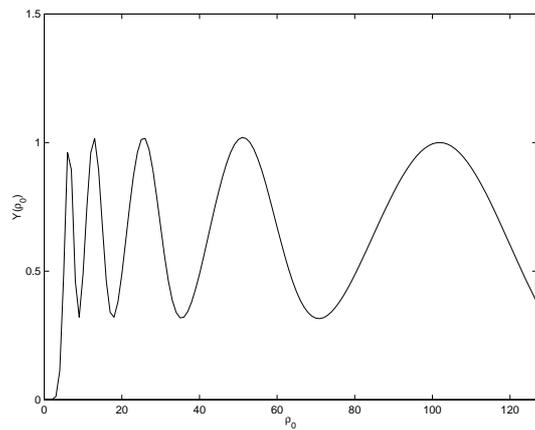
erfüllen. Somit kann im Sinne der Singulärwertzerlegung der Moore-Penrose-Inversen das dem Mittelwert entsprechende verschwindende Diagonalelement von \mathbf{Y} bei der Inversion wieder auf null gesetzt werden. Dieses Vorgehen muß auch bei Frequenzpunkten Anwendung finden, deren Matrixelemente von \mathbf{Y} nahe null sind, was bei nicht zu weit auseinanderliegenden Frameschranken auf die Rekonstruktion eines Tiefpaß- oder Bandpaßbildes hinausläuft. Im übrigen handelt es sich dabei nach (Louis, 1989) um das Verfahren der *abgeschnittenen Singulärwertzerlegung* in der Theorie der Regularisierung schlecht gestellter Probleme. Die Dimensionen des Vektorraums \mathbb{R} , die den zu null gesetzten Frequenzpunkten entsprechen, können entfernt werden, so daß die Gaborwavelet-funktionen im verbleibenden Unterraum einen Frame bilden. Dieser Vorgang ist unproblematisch, da jeder endlichdimensionale normierte Raum vollständig und bei Einführung eines zu der Norm passenden Skalarproduktes auch ein Hilbertraum ist.

Der Verlust des Gleichanteils hat für die meisten Bildverarbeitungsprobleme, wie die Objekterkennung, kaum nachteilige Auswirkungen, da die Intensitäten gleichmäßig zwischen 0 und beispielsweise 255 verteilt werden. Hinzu kommt, daß sich Bildmerkmale hauptsächlich durch ihre hochfrequenten Anteile, d.h. eine eventuelle Feinstruktur, auszeichnen.

In den Abb. 10,11 und 12 ist zu sehen, daß die Rekonstruierbarkeit in keinem der Fälle gefährdet ist. Soll für die IGWT die Tight-Frame-Approximation benutzt werden, kann es bei der Verwendung von nur wenigen Gaborkernen zu Problemen kommen (s. Abb. 12(d)). Die Programme, die diese Transformationen und Rekonstruktionen berechnet haben, sind mit Algorithmen aus (Press et al., 1992) in C implementiert.



Abbildung 9: Das Originalbild.

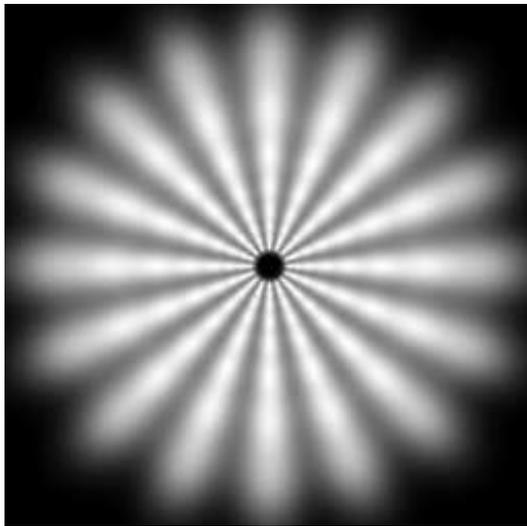
(a) Metrischer Operator $Y(\vec{\rho})$ im Frequenzraum(b) Schnittlinie $Y(\rho_0)$ mit $\rho_0 = \vec{\rho}^T \vec{e}(0) \geq 0$ 

(c) Rekonstruktion

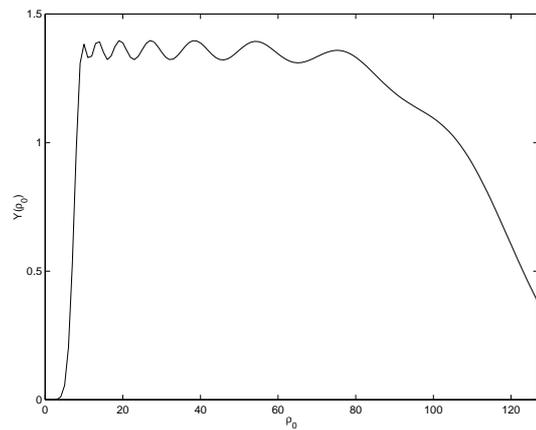


(d) Tight-Frame-Rekonstruktion

Abbildung 10: Rekonstruktionen mit $L = 12$ Orientierungen und $M = 5$ Skalen. Der Skalenabstand ist $a_0 = 2$ ($\sigma = 4$, $\tau = 6$).



(a) Metrischer Operator $Y(\vec{\rho})$ im Frequenzraum



(b) Schnittlinie $Y(\rho_0)$ mit $\rho_0 = \vec{\rho}^T \vec{e}(0) \geq 0$

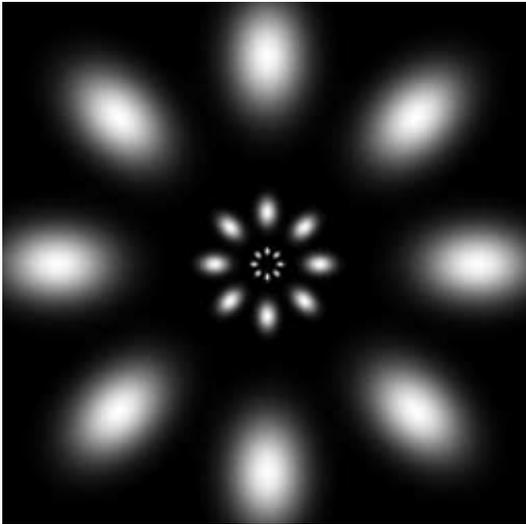
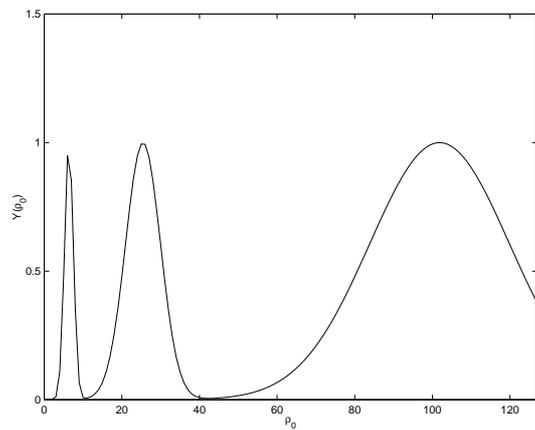


(c) Rekonstruktion

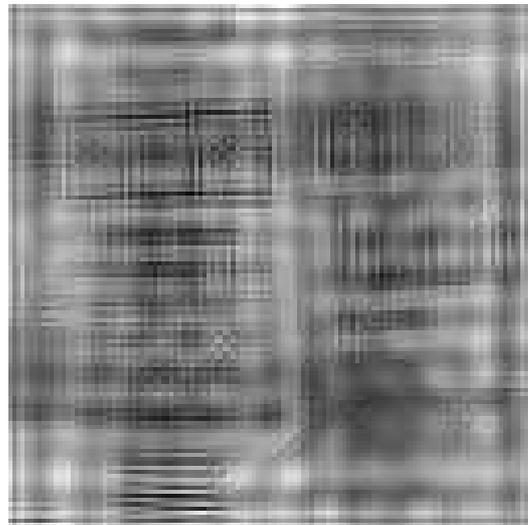


(d) Tight-Frame-Rekonstruktion

Abbildung 11: Rekonstruktionen mit $L = 16$ Orientierungen und $M = 8$ Skalen. Der Skalenabstand ist $a_0 = \sqrt{2}$ ($\sigma = 4$, $\tau = 6$).

(a) Metrischer Operator $Y(\vec{\rho})$ im Frequenzraum(b) Schnittlinie $Y(\rho_0)$ mit $\rho_0 = \vec{\rho}^T \vec{e}(0) \geq 0$ 

(c) Rekonstruktion



(d) Tight-Frame-Rekonstruktion

Abbildung 12: Rekonstruktionen mit $L = 8$ Orientierungen und $M = 3$ Skalen. Der Skalenabstand ist $a_0 = 4$ ($\sigma = 4$, $\tau = 6$).

6.4 Skalenabhängige Abtastung des Ortsraums und der Translationen

6.4.1 Die pyramidale Gaborwavellettransformation

Das Bild $I(\vec{x})$ wird nun im Ortsraum durch die Funktion

$$s_{\Delta_0 a_0^m}(\vec{x}) = \sum_{\vec{n} \in \mathbb{Z}^2} \delta(\vec{x} - \vec{n} a_0^m \Delta_0) \quad (6.4-1)$$

mit dem Abtastintervall Δ diskretisiert:

$$I^a(\vec{x}) = I(\vec{x}) s_{\Delta_0 a_0^m}(\vec{x}) = I(\vec{n} \Delta_0 a_0^m) s_{\Delta_0 a_0^m}(\vec{x}). \quad (6.4-2)$$

$I^a(\vec{x})$ ist das abgetastete, aber kontinuierliche Bildsignal $I(\vec{n} \Delta)$, eine Folge (diskrete Funktion), die nur auf den Abtastgitterpunkten definiert ist. Setzt man $I^a(\vec{x})$ in (5.1-15) ein, erhält man

$$\begin{aligned} \mathcal{I}(\vec{n}_0, m, l) &= \int_{\mathbb{R}^2} d^2x I(\vec{x}) \sum_{\vec{n} \in \mathbb{Z}^2} \delta(\vec{x} - \vec{n} \Delta_0 a_0^m) X_{\vec{n}_0, m, l}^*(\vec{x}) \\ &= \sum_{\vec{n} \in \mathbb{Z}^2} I(\vec{n} \Delta_0 a_0^m) X_{\vec{n}_0, m, l}^*(\vec{n} \Delta_0 a_0^m) \\ &= \sum_{\vec{n} \in \mathbb{Z}^2} I(\vec{n} \Delta_0 a_0^m) a_0^{-m} \chi^* \left[\mathcal{Q} \left(\frac{2\pi l}{L} \right) (\vec{n} - \vec{n}_0) \Delta_0 \right] \end{aligned} \quad (6.4-3)$$

Bei der Berechnung der einzelnen Gaborwavellets ist man nicht an irgendeine Gitter gebunden, da die Funktionen insgesamt bekannt sind und exakt an jeder Stelle berechnet werden können. (6.4-3) ermöglicht dabei eine ökonomischere Speicherbenutzung, da zum einen die Ausgabebilder der Gaborfilterbank bei größeren Skalen eine entsprechend reduzierte Größe haben, und zum anderen nur noch L Kerne berechnet und gespeichert werden müssen. Diese pyramidale Struktur ist im Blockschaltbild von Abb. 13 visualisiert.

Damit für die kleinste Skala $a_0^{m_0}$, d.h. für die höchste Frequenzebene, jeder Gitterpunkt von den Translationen erreicht werden kann, setzt man

$$a_0^{m_0} \frac{\Delta_0}{\Delta} = 1 \quad \Rightarrow \quad \Delta_0 = a_0^{-m_0} \Delta. \quad (6.4-4)$$

Die größte Skala ist $a_0^{M_0}$, wobei M_0 an die Anzahl der Pixel $N_1 N_2$ gekoppelt sein sollte. Der endliche Träger des abgetasteten Bildsignals gibt die Menge

$$\mathcal{B} = \{\vec{n} \in \mathbb{Z}^2 \mid 0 \leq n_1 \leq N_1 - 1 \wedge 0 \leq n_2 \leq N_2 - 1\} \quad (6.4-5)$$

an.

Läßt man die DGWT im Frequenzraum berechnen, was auch sinnvoll ist, so entsprechen die Anti-Aliasing-Filter (AAF) mit den Abwärtstastern einem Entfernen hochfrequenter Anteile und einer entsprechenden Halbierung der Frequenzperiode.

Infolge dieser Vereinbarungen ist die Gaborwavellettransformation nun numerischen Berechnungen zugänglich:

$$\begin{aligned} \mathcal{I}(\vec{n}_0, m, l) &= \text{DGWT}\{I(\vec{n})\}(\vec{n}_0, m, l) \\ &= \sum_{\vec{n} \in \mathbb{Z}^2} I(\vec{n} \Delta a_0^{m-m_0}) a_0^{-m} \chi^* \left[\mathcal{Q} \left(\frac{2\pi l}{L} \right) (\vec{n} - \vec{n}_0) a_0^{-m_0} \Delta \right] \\ &= \sum_{\vec{n} \in \mathbb{Z}^2} I(\vec{n} \Delta a_0^{m-m_0}) a_0^{-m} \chi \left[\mathcal{Q} \left(\frac{2\pi l}{L} \right) (\vec{n}_0 - \vec{n}) a_0^{-m_0} \Delta \right] \\ &= \sum_{\vec{n} \in \mathbb{Z}^2} I^{(\Delta a_0^{m-m_0})}(\vec{n}) a_0^{-m} \tilde{\chi}_l(\vec{n}_0 - \vec{n}) \\ &= a_0^{-m} I^{(a_0^{m-m_0})}(\vec{n}_0) * \tilde{\chi}_l(\vec{n}_0) \end{aligned} \quad (6.4-6)$$

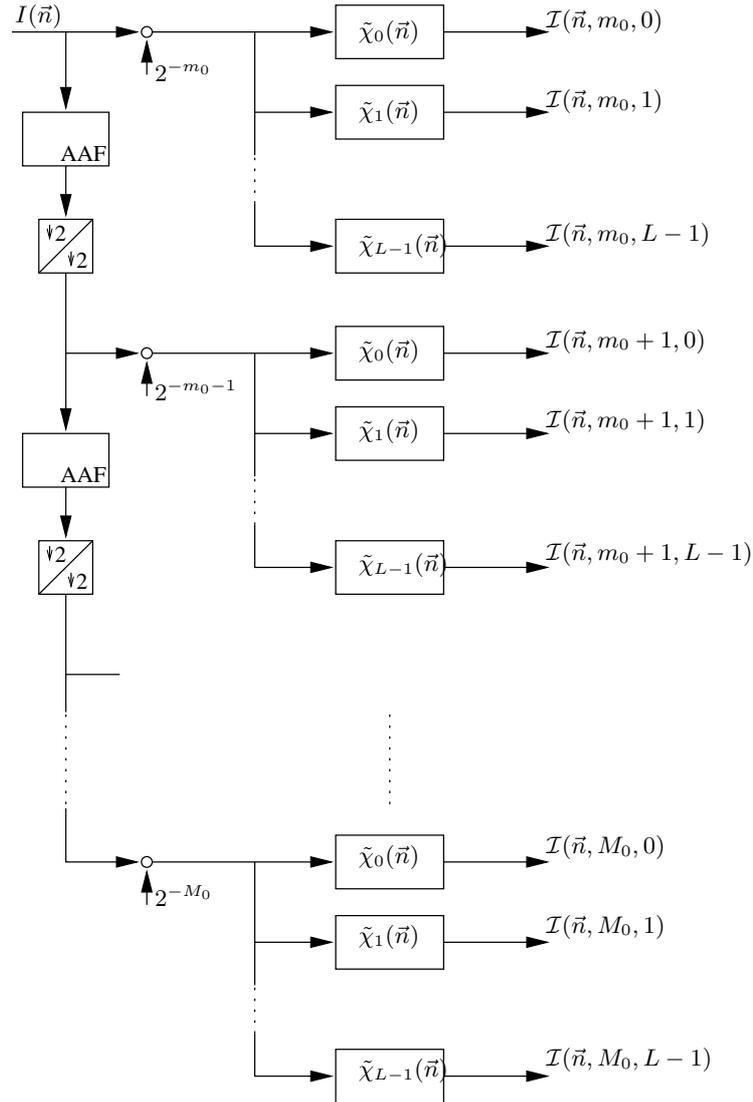


Abbildung 13: Struktur der pyramidalen Gaborwavellettransformation. Die Frequenzebenen liegen um Oktaven auseinander.

mit

$$\tilde{\chi}_l(\vec{n}) = \chi \left[\mathbf{Q} \left(\frac{2\pi l}{L} \right) \vec{n} a_0^{-m_0} \Delta \right]. \quad (6.4-7)$$

$I^{(a_0^{m-m_0})}(\vec{n})$ ist eine um $a_0^{m-m_0}$ abwärtsgetastete Version von $I(\vec{n}\Delta)$, so daß die diskrete Gaborwavellettransformation mit einem Satz Filter, deren Punktantworten sich nur durch die Orientierung unterscheiden, berechnet werden kann. Aufgrund der sukzessiven Abwärtstastung des Originalbildes bezeichnet man diese Form als *pyramidale Gaborwavellettransformation*. Eine effektive pyramidale Bildkodierung wurde erstmals in (Burt und Adelson, 1983) in Form der Laplaceschen Pyramide vorgeschlagen, wobei der Hochpaßanteil des Bildsignals jeweils festgehalten und kodiert wird. Nach diesem Prinzip arbeitet auch die pyramidale DGWT. Die Struktur wird von Abb. 13 und die Hierarchie der Frequenzebenen von Abb. 14 visualisiert. In den Abbildungsbeispielen wird von einem Oktavabstand zwischen den Skalen ausgegangen, was aber keineswegs notwendig ist. Da das Eingangsbildsignal dezimiert und nicht bloß abwärtsgetastet wird, und diese Operation im Frequenzraum stattfindet, kann a_0 jeden beliebigen Wert, der größer als eins ist, annehmen. Dies gilt selbstverständlich auch für die nächsten beiden Unterabschnitte.

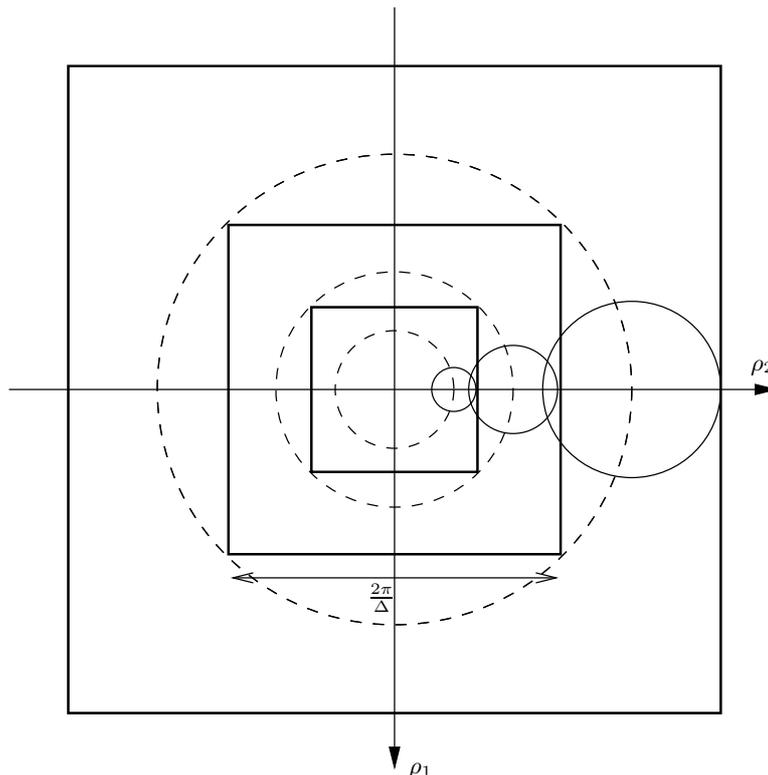


Abbildung 14: Halbierung der Nyquistintervalle bei der pyramidalen Gaborwavelettransformation (Tiefpaßabtastung). Hier ist nur eine Orientierung auf drei Frequenzebenen dargestellt und der Einfachheit halber $\sigma = \tau$ gesetzt.

6.4.2 Die inverse pyramidale Gaborwavelettransformation

Mit der gleichen Struktur wie in Abb. 13 kann das kodierte Bild zurückgewonnen werden. Man beginne dabei unten, d.h. bei der größten Skala. $\tilde{\chi}^l(\vec{n})$ ist die zu $\tilde{\chi}_l(\vec{n})$ duale Gaborfunktion. Wie bei der Signalanalyse ist es auch für die Synthese günstig, den Weg über den Frequenzraum zu nehmen. Anti-Imaging-Filter (AIF) und Aufwärtstaster haben dann die Wirkung einer Verdoppelung der Frequenzperiode und Nullsetzen der hinzugekommenen Frequenzpunkte. Die zugehörigen Grundlagen der Abstratenumsetzung können in (Fliege, 1993) nachgelesen werden.

6.4.3 DGWT mit Bandpaßabtastung

Die datenspeichergünstigste Implementierung beruht auf der separat optimalen Abtastung jedes einzelnen Teilbandes, wobei bei Gaborwavelets mit kreisförmiger Gaußfunktion (d.h. für $\sigma = \tau$) das Abtastintervall Δ bei konstanter Skala unverändert bleibt. Abb. 16 zeigt, wie die entsprechenden Intervalle im Frequenzraum ausgeschnitten werden, was einer Abwärtstastung um einen gebrochenen Faktor gleichkommt. Bei der Bildsignalsynthese beschreitet man den zur Tiefpaßabtastung umgekehrten Weg.

Im Falle einer Bandpaßabtastung kann ebenfalls eine pyramidale Gabortransformation nach Abb. 13 berechnet werden. Das Abtastintervall bleibt für alle Orientierungen einer konstanten Skala unverändert und halbiert sich bei Oktavwavelets beim Übergang auf die nächst höhere.

6.5 Lokale Rekonstruktion

Aufgrund der Tatsache, daß die Gabortransformation eine Lokalisierung in Orts- und Frequenzraum leistet, muß keine Gesamtbildrekonstruktion durchgeführt werden, wenn nur einzelne Bereiche

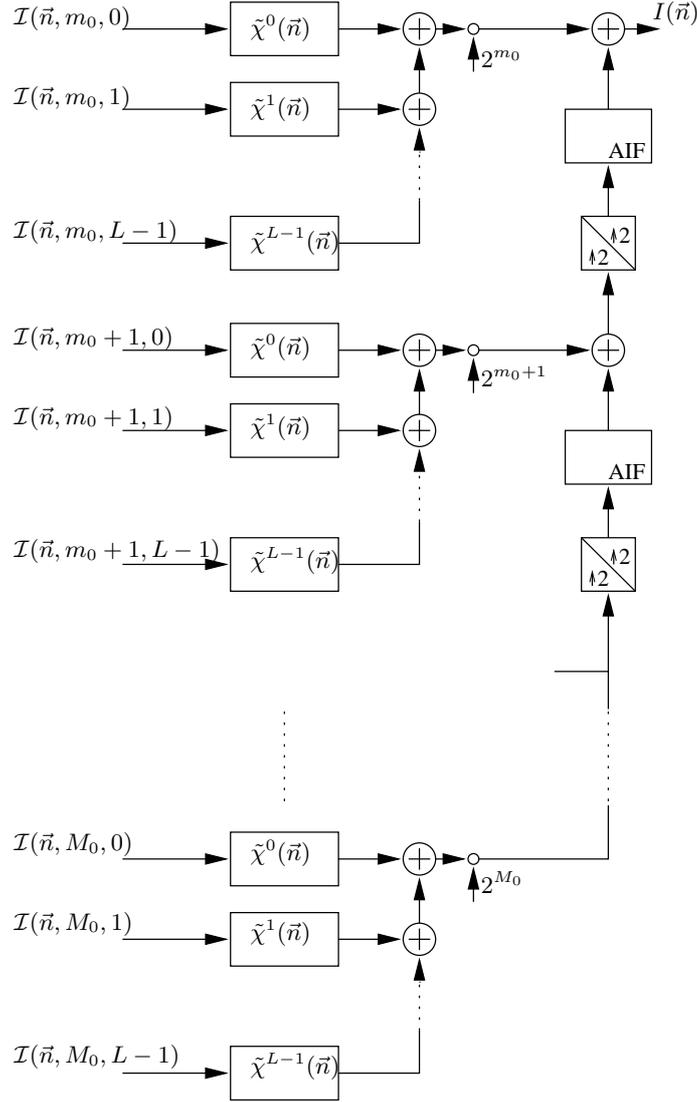


Abbildung 15: Struktur der inversen pyramidalen Gaborwvetelettransformation. Die Skalen liegen um Oktaven auseinander. Bei der Anwendung der dualen Gaborfilter ist zu beachten daß der Vektorraum mit einem anderen Maß ausgestattet ist: $\mu = 2^{-4m}$, vgl. (4.2-16)

interessieren. Das geht soweit, daß eine lokale Merkmalsrekonstruktion für einen einzelnen Translationspunkt \vec{n}_1 , aber aus allen Skalen und Orientierungen berechnet werden kann. Zu diesem Zweck setzt man die Daten aller anderen Translationen zu null:

$$\mathcal{I}_0(\vec{n}_0, m, l) \rightarrow \mathcal{I}_0(\vec{n}_0, m, l) \delta_0(\vec{n}_0 - \vec{n}_1). \quad (6.5-1)$$

Die entsprechende Rekonstruktionsformel lautet somit

$$I_{\vec{n}_0}(\vec{n}\Delta) = \frac{\Delta^4}{4\pi^2} \sum_{m=m_0}^{M_0} \sum_{l=0}^{L-1} a_0^{-2m} \mathcal{I}_0(\vec{n}_1, m, l) \chi^{\vec{n}_0\Delta, a_0^m, \frac{2\pi l}{L}}(\vec{n}\Delta). \quad (6.5-2)$$

Das Ergebnis ist zwar ein Bild der ursprünglichen Größe, die Grauwerte der einzelnen Bildpunkte fallen aber mit steigender Entfernung von \vec{n}_1 stark ab.

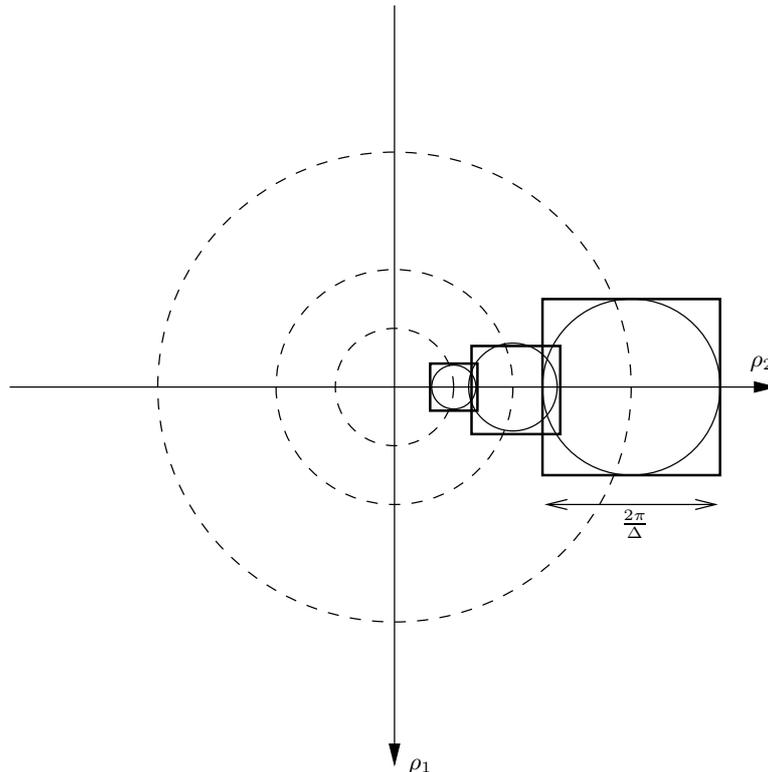


Abbildung 16: Halbierung der Nyquistintervalle bei der Gaborwavellettransformation (Bandpaßabtastung). Hier ist nur eine Orientierung auf drei Frequenzebenen dargestellt und der Einfachheit halber $\sigma = \tau$ gesetzt.

6.6 Wavelet-Frames und orthogonale Wavelet-Basen

In diesem Abschnitt soll kurz auf die wichtigsten Eigenschaften und Unterschiede von Signaldarstellungen mit den bisher behandelten *Wavelet-Frames* und *orthogonalen Wavelets* eingegangen werden.

Der Begriff „Wavelet“ wurde erstmals von den Geophysikern A. Grossmann und J. Morlet im Jahre 1984 veröffentlicht (Grossmann und Morlet, 1984). Diese „wavelets of constant shape“, wie die Autoren sie zunächst nannten, waren übervollständige Frames, ohne daß sie so bezeichnet wurden. 1984 gelang Y. Meyer die Konstruktion orthonormaler Wavelet-Basen und 1989 in den Arbeiten von Y. Meyer und S. Mallat die schnelle Transformation mit der sog. *Multiskalenanalyse* („*multiresolution analysis*“, (Mallat, 1989)), die der Wavelet-Analyse zum Durchbruch verhalf.

6.6.1 Analyse und Synthese: Filterbänke und Multiskalenanalyse

Die Multiskalenanalyse ist eine effektive und numerisch stabile Methode zur Berechnung der diskreten Wavelettransformation. Eine grundlegende Einführung findet sich in (Fliege, 1993). Der Raum der energiebeschränkten Signale ist eine Verschachtelung von Unterräumen, von denen jeder einer Auflösungsstufe entspricht, d.h. jeder entspricht dem des um 2 abwärtsgetasteten Signals aus dem übergeordneten Raum. Jeder dieser Signalkäume wird von den orthogonalen *Skalierungsfunktionen* aufgespannt, die nächste Auflösungsstufe kann durch Hinzunahme des entsprechenden Wavelets, des orthogonalen Komplements, erreicht werden. Die gesamte Transformation geht im Falle von Oktavwavelets auf diese Weise durch eine Kaskade von Zweikanalfilterbänken vorstatten, an deren Ende ein *Skalierungskoeffizient* den Tiefpaßanteil und viele *Waveletkoeffizienten* die Hochpaßanteile tragen. Die MSA ist aus diesem Grund auch in der Lage, den Gleichanteil eines Signals zu kodieren, da die Skalierungsfunktionen nicht mittelwertfrei sind. Die Zweige dieser Filterbank enthalten ein LSI-System, deren Impulsantworten aus den Wavelets und Skalierungsfunktionen berechnet werden können, und einen Abwärtstaster (s. Abb. 17). Diese Thematik kann in dieser Arbeit nicht vertieft

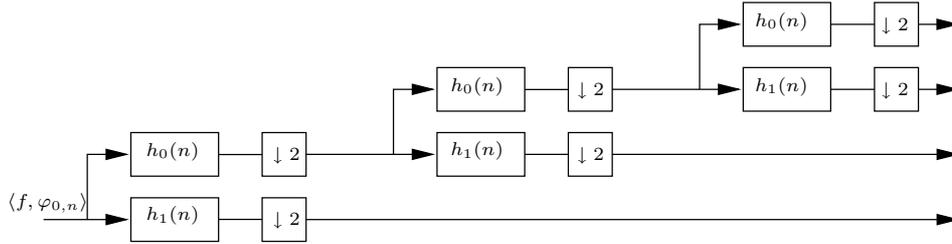


Abbildung 17: Struktur einer eindimensionalen Multiskalenanalyse mit drei Zweikanal-Oktavfilterbänken. $\varphi_{0,n}$ sind die Skalierungsfunktionen, die mit ihren Translationen den Originalsignalraum aufspannen.

werden, weshalb auf (Mallat, 1989), (Fliege, 1993), (Louis et al., 1994), (Kaiser, 1994) verwiesen sei.

6.6.2 Die Wavelettransformierte: übervollständige und kritische Abtastung

Bei der Gaborwavelettransformation nach (6.2-6) handelt es sich um eine Bildsignalzerlegung durch eine $M \cdot L$ -Kanalfilterbank. Die Gesamtausgabe umfaßt daher $D_{FB} = MLN_1N_2$, bei der Multiskalenanalyse $D_{MSA} = N_1N_2 = D_I$ Datenpunkte, wobei D_I die Anzahl der Datenpunkte des Eingabebildes ist. $\mathcal{I}_0(\vec{n}_0, m, l)$ ist daher ML -fach überabgetastet, was bei den üblichen 12 Orientierungen und 5 Skalen einem Faktor 60 entspricht. Etwas ökonomischer verhält sich dagegen die pyramidale Tiefpaßabtastung, bei der die Anzahl der Ausgabeabtastwerte

$$D_{pyrTP} = L \sum_{m=0}^{M-1} \frac{N_1N_2}{a_0^{2m}} = N_1N_2L \frac{1 - a_0^{-2M}}{1 - a_0^{-2}} \quad (6.6-1)$$

beträgt, was für $L = 12$, $M = 5$ und $a_0 = 2$ auf einen Überabtastungsfaktor

$$\ddot{U}_{pyrTP} = \frac{D_{pyrTP}}{D_I} = L \frac{1 - a_0^{-2M}}{1 - a_0^{-2}} \approx 15,984 \quad (6.6-2)$$

führt. Benutzt man eine Bandpaßabtastung, beschränkt sich der Nyquist-Bereich nur auf die Umgebung der Gaußglocke. Für $\tau = \sigma$ ergibt sich somit die Anzahl der Bildpunkte zu

$$D_{pyrBP} = \bar{N}_1\bar{N}_2L \frac{1 - a_0^{-2M}}{1 - a_0^{-2}} = \frac{N_1N_2L}{\alpha^2} \frac{1 - a_0^{-2M}}{1 - a_0^{-2}}. \quad (6.6-3)$$

α ist hier der Dezimationsfaktor für die transformierten Bilder auf der kleinsten Skala. Voraussetzung für diese Überlegungen ist allerdings, daß kein Teilband der kleinsten Skala von der Begrenzung des Nyquistbereichs abgeschnitten wird.

6.6.3 Das Problem der „Verschiebbarkeit“

Den oben bereits beschriebenen Vorteilen einer Wavelettransformation mit der Multiskalenanalyse soll ein Nachteil gegenübergestellt werden. Dabei handelt es sich um die Instabilität der Ausgabe-signale in den einzelnen Teilbändern unter Translationen der Systemeingabe. In (Simoncelli et al., 1992) ist dieses Phänomen untersucht und veranschaulicht worden. Demnach kann sich bei einer pyramidalen Transformation mit orthogonalen Waveletbasen der Signalgehalt völlig ändern, wenn das Eingabebild um nur einen Pixel verschoben wird. Dafür sind weitgehend Blöcke von dem in Abb. 18 dargestellten Typ, den man auch bei der Multiskalenanalyse in Abb. 17 benutzt, verantwortlich. Ein ähnliches Verhalten ist in zwei Dimensionen bezüglich Drehungen und Subpixeltranslationen zu erwarten. Die Transformation ist nicht verschiebbar, wenn in

$$I_{2,\vec{K}}(n_1, n_2) = I_2(K_1n_1, K_2n_2) = \sum_{m_1=0}^{M_1-1} \sum_{m_2=0}^{M_2-1} h(K_1n_1 - m_1, K_2n_2 - m_2) I_1(m_1, m_2) \quad (6.6-4)$$

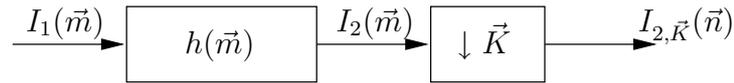


Abbildung 18: LSI-System mit Abwärtstaster

$I_2(n_1, n_2)$ aus $I_{2,\vec{K}}(n_1, n_2)$ nicht durch eine Interpolation rekonstruiert werden kann. Nichtverschiebbarkeit ist somit die Folge von Aliasing beim Abwärtstasten.

Bei der nicht skalenangepaßten diskreten Gaborwavellettransformation aus Abschnitt 6.2 ist eine Verschiebung um einen oder mehrere Pixel trivial: Die Ausgabe in sämtlichen Teilbändern ist ebenfalls um ebensoviele Pixel verschoben. Eventuell ist die Periodizitätseigenschaft der DFT zu beachten. Da die Abtastung so vorgenommen wird, daß in keinem Teilband wesentliche Aliasing-Effekte auftreten, bewirkt auch eine Subpixelverschiebung nur leichte Veränderungen, weil die Gaborwavellettransformation aus LSI-Systemen besteht. Wird ein Bildsignal mit der Pyramide aus Abb. 13 kodiert, so kann dank der AAF der Tiefpaßanteil der nächst höheren Stufe durch Interpolation wiederhergestellt werden. Daher ist auch hierbei kein instabiles Verhalten eines Teilbandbildes zu erwarten.

Das Problem der Verschiebbarkeit läßt sich auch auf die Orientierungen einer zweidimensionalen Wavellettransformation erweitern, wobei allerdings nur darauf zu achten ist, daß alle Winkel im Frequenzraum einigermaßen gut abgedeckt sind, was bei den Gaborwavellet-Frames immer möglich ist. Ist dies der Fall, kann aus der Transformierten jedes beliebige Teilband durch Rekonstruktion aus allen Teilbändern und Projektion des Ergebnisses auf eine entsprechende Wavelletfunktion erfolgen. Die numerische Stabilität bei diesem Vorgehen ist natürlich durch die Frameschranken vorgegeben. Statt dessen kann aber auch die Redundanz der Frametransformierten ausgenutzt werden. Das Verhalten zweidimensionaler orthogonaler Wavelets ist dagegen in dieser Hinsicht problematisch. Nach (Louis et al., 1994) können sie nur die Orientierungen in Richtung der kartesischen Koordinatenachsen und der Diagonalen annehmen. Dagegen gibt es aber auch (nichtorthogonale) Kerne, die sich durch eine Linearkombination endlich vieler Funktionen in jede beliebige Richtung drehen lassen. Solche Filter wurden von W. Freeman und H. Adelson entworfen und untersucht (s. (Freeman und Adelson, 1991)).

Zusammenfassend sei für diesen Abschnitt angemerkt, daß das Problem der Verschiebbarkeit durch die Übervollständigkeit der Frametransformierten weitgehend entschärft ist. Denn man kann nach (4.2-24) die Frametransformierte bezüglich eines hinzugefügten Indexvektors \vec{m}' aus den bisher bekannten transformierten Daten berechnen.

Teil III

Komplexe Zellen

- Polynome und Fourieramplituden.
- Die Hayesschen Theoreme.
- Patches.
- Mehrdeutigkeit der Jetrekonstruktion.
- Eindeutige Gesamtbildrekonstruktion.
- Probleme bei der numerischen Umsetzung.

7 Rekonstruierbarkeit lokaler Merkmale aus Gaboramplituden

7.1 Fourieramplituden

Die Fouriertransformierte ein- oder mehrdimensionaler Funktionen (2.1-1) kann nach Betrag (Amplitude) und Phase als

$$\hat{f}(\vec{\omega}) = |\hat{f}(\vec{\omega})| e^{i\Phi(\vec{\omega})} \quad (7.1-1)$$

dargestellt werden. Um $f(\vec{x})$ oder $f(\vec{n})$ wieder völlig nach (2.1-2) aus dem Frequenzraum rekonstruieren zu können, ist die Verfügbarkeit sowohl der Amplitude als auch der Phase nötig. Fourieramplituden spielen aufgrund ihrer Translationsinvarianz eine gewisse Rolle in der Mustererkennung, s. (Gardenier et al., 1986), dürften aber lokalisierbaren Transformationen mit Weyl-Heisenberg-Systemen oder Wavelets unterlegen sein.

7.1.1 Polynome in einer und mehreren komplexen Variablen

Polynome vom Grad $N - 1$, die aus der eindimensionalen Z-Transformation einer reellen Folge, d.h. einer diskreten Funktion, entstehen, lassen sich als Produkt aus Linearfaktoren darstellen:

Theorem 1 (Fundamentalsatz der Algebra)

$$\check{f}(z) = \sum_{n=0}^{N-1} f(n) z^{-n} = c \prod_{m=1}^{N-1} (z^{-1} - \alpha_m) \quad z, \alpha_m \in \mathbb{C}, f \in \mathbb{R}. \quad (7.1-2)$$

Das bedeutet, daß Polynome in einer Variablen immer im Sinne von Theorem 1 reduzierbar sind. Um den Beitrag der einzelnen Linearfaktoren zu Amplitude und Phase des eindimensionalen DSFT-Spektrums darzustellen, betrachtet man nur Punkte auf dem Einheitskreis $z = e^{i\nu}$ in der z-Ebene:

$$\begin{aligned} \check{f}(e^{i\nu}) &= \sum_{n=0}^{N-1} f(n) e^{-i\nu n} = c \prod_{m=1}^{N-1} (e^{-i\nu} - \alpha_m) \\ &= |c| e^{i\varphi_0} \prod_{m=1}^{N-1} |e^{-i\nu} - \alpha_m| e^{i\varphi_m(\nu)}. \end{aligned} \quad (7.1-3)$$

Es wird sich zeigen, daß die durch Theorem 1 festgelegte Reduzierbarkeit in Linearfaktoren bei Polynomen in mehreren Variablen nicht mehr gegeben ist. Die folgenden beiden Theoreme bringen diesen Sachverhalt zum Ausdruck.

Theorem 2 (Hayes und McClellan, 1982) Ist $f : \mathbb{R}^M \rightarrow \mathbb{R}^N$, $M < N$ eine stetig differenzierbare Abbildung, so hat $f(\mathbb{R}^M)$ in \mathbb{R}^N das Maß Null.

Man betrachte im folgenden Polynome in K Variablen vom Gesamtgrad $N = \sum_{j=1}^K N_j$ mit reellen Koeffizienten:

$$\check{f}_N(\vec{z}) = \sum_{n_1=0}^{N_1} \dots \sum_{n_K=0}^{N_K} f(\vec{n}) z_1^{-n_1} \dots z_K^{-n_K} \in \mathcal{P}(N, K). \quad (7.1-4)$$

Die Koeffizienten $f(\vec{n})$ eines Polynoms aus $\mathcal{P}(N, K)$ können als ein Punkt in einem $\alpha(N, K)$ -dimensionalen Vektorraum gesehen werden, d.h. jedes Polynom in K Dimensionen vom Gesamtgrad N ist durch $\alpha(N, K)$ Punkte eindeutig definiert. Jedem dieser Punkte ist eineindeutig ein Element aus $\mathcal{P}(N, K)$ zugeordnet. Für das folgende Theorem ist es günstig, auch komplexwertige Koeffizienten zuzulassen, was bedeutet, daß man es jetzt mit Vektoren im $\mathbb{R}^{2\alpha(N, K)}$, die die Real- und Imaginärteile enthalten, zu tun hat.

Theorem 3 (Hayes und McClellan, 1982) Sei $\mathcal{Q} \subset \mathcal{P}(N, K)$ die Teilmenge aller reduzierbaren Polynome aus $\mathcal{P}(N, K)$, dann entspricht \mathcal{Q} einer Teilmenge, die das Maß Null im $\mathbb{R}^{2\alpha(N, K)}$ hat.

Beweis: Sei ein reduzierbares Polynom der Form

$$\check{f}_N(\vec{z}) = \lambda \check{g}_L(\vec{z}) \check{h}_{N-L}(\vec{z}) \in \mathcal{Q}(L), \quad (7.1-5)$$

wobei $\hat{g}_L(\vec{z})$ und $\hat{h}_{N-L}(\vec{z})$ einen auf 1 normierten Koeffizienten zum nullten Grad haben. In der Vektordarstellung von (7.1-5) läßt sich die obige Zerlegung als stetig differenzierbare Abbildung

$$\mathbb{R}^2 \times \mathbb{R}^{2\alpha(L,K)-2} \times \mathbb{R}^{2\alpha(N-L,K)-2} \rightarrow \mathbb{R}^{2\alpha(N,K)} \quad (7.1-6)$$

beschreiben. Der Beweis ist nach Theorem 2 erbracht, wenn sich

$$\begin{aligned} 2 + 2\alpha(L, K) - 2 + 2\alpha(N - L, K) - 2 &< 2\alpha(N, K) \Leftrightarrow \\ \alpha(L, K) + \alpha(N - L, K) - 1 &< \alpha(N, K) \end{aligned} \quad (7.1-7)$$

als richtig herausstellt. Um dieses wiederum nachzuweisen, benötigt man zunächst den trivialen Zusammenhang

$$\alpha(N, K) > \alpha(N - 1, K) \quad \forall N \geq 1, K \geq 1. \quad (7.1-8)$$

Soll das Polynom aus (7.1-5) um eine Variable erweitert werden, so kann es eindeutig zum einen als

$$\check{f}_N(\vec{z}, z_{K+1}) = \sum_{j=0}^N \check{q}_{N-j}(\vec{z}) z_{K+1}^j \quad (7.1-9)$$

und zum anderen mit den Dimensionen der Koeffizientenvektoren formuliert werden:

$$\begin{aligned} 2\alpha(N, K + 1) &= 2 \sum_{j=0}^{N-1} \alpha(N - j, K) + 2 \Leftrightarrow \\ \alpha(N, K + 1) &= \sum_{j=0}^{N-1} \alpha(N - j, K) + 1 \\ &= \sum_{j=0}^{N-L-1} \alpha(N - j, K) + \sum_{j=N-L}^{N-1} \alpha(N - j, K) + 1 \\ &= \sum_{j=0}^{N-L-1} \alpha(N - j, K) + \sum_{j=0}^{L-1} \alpha(L - j, K) + 1 \\ &= \sum_{j=0}^{N-L-1} \alpha(N - j, K) + \alpha(L, K + 1) \Rightarrow \\ \alpha(N, K + 1) &> \sum_{j=0}^{N-L-1} \alpha(N - L - j, K) + \alpha(L, K + 1) \\ &= \alpha(N - L, K + 1) - 1 + \alpha(L, K + 1) \\ \alpha(N, K') &> \alpha(N - L, K') - 1 + \alpha(L, K') \end{aligned} \quad (7.1-10)$$

Dies gilt jedoch nur, wenn $K' = K + 1 \geq 2$. Aus (7.1-7) wird somit

$$\alpha(L, K') + \alpha(N - L, K') - 1 < \alpha(N, K'). \quad (7.1-11)$$

$K \geq 1$ ist als Dimension der Definitionsmenge des Polynoms immer gegeben, z.B. durch (7.1-4). $\alpha(\cdot, 0)$ würde daher keinen Sinn haben. \mathcal{Q} hat das Maß Null in $\mathcal{P}(N, K)$, so daß die endliche Vereinigung

$$\bigcup_{L=1}^{N-1} \mathcal{Q}(L) = \mathcal{Q} \quad (7.1-12)$$

in $\mathcal{P}(N, K)$ ebenfalls das Maß Null hat. \square

Für die Anwendung bedeutet dies, daß es sehr unwahrscheinlich ist, auf ein Bildsignal (bei $K' = 2$) zu treffen, dessen DSFT ein reduzierbares Polynom liefert. Trotzdem soll eine Faktorisierung der Z-Transformierten in den folgenden Betrachtungen möglich sein, wodurch die allgemeine Formulierung

$$\check{f}(\vec{z}) = az_1^{-n_{01}} z_2^{-n_{02}} \prod_{j=1}^M \check{f}_j(\vec{z}) \quad (7.1-13)$$

motiviert wird. $\check{f}_j(\vec{z})$ sind hierbei nichttriviale und nichtreduzierbare Polynome. *Nichttriviale Polynome* sind bei (Hayes, 1982) solche, die aus mehr als einem Summenglied bestehen.

7.1.2 Die Hayesschen Theoreme

In (Hayes, 1982) sind einige Eindeutigkeitsätze zu Nur-Amplituden- und Nur-Phasenrekonstruktionen nebst Algorithmen beschrieben. Für das Problem der Signalrekonstruktion aus Amplitudenspektren bzw. aus

$$\begin{aligned} |\check{f}(e^{i\nu_1}, e^{i\nu_2})|^2 &= \check{f}(e^{i\nu_1}, e^{i\nu_2}) \check{f}(e^{i\nu_1}, e^{i\nu_2})^* = \check{f}(e^{i\nu_1}, e^{i\nu_2}) \check{f}(e^{-i\nu_1}, e^{-i\nu_2}) \\ &= \check{f}(z_1, z_2) \check{f}(z_1^{-1}, z_2^{-1})|_{z_1=e^{i\nu_1}, z_2=e^{i\nu_2}} = \check{r}_f(z_1, z_2)|_{z_1=e^{i\nu_1}, z_2=e^{i\nu_2}} \end{aligned} \quad (7.1-14)$$

wendet man die Zerlegung (7.1-13) an:

$$\check{r}_f(z_1, z_2) = a^2 \prod_{j=1}^M \check{f}_j(z_1, z_2) \check{f}_j(z_1^{-1}, z_2^{-1}). \quad (7.1-15)$$

Dieser und allen nachfolgenden Polynomformulierungen liegt die Z-Transformation nach (2.2-20) zugrunde. Die Multiplikation mit $z_1^{-N_1+1} z_2^{-N_2+1}$ macht daraus ein Polynom vom Grad $2[(N_1 - 1) + (N_2 - 1)]$ in $z_1^{-1} z_2^{-1}$, nämlich

$$\check{q}_f(z_1, z_2) = z_1^{-N_1+1} z_2^{-N_2+1} \check{r}_f(z_1, z_2) = a^2 \prod_{j=1}^M \check{f}_j(z_1, z_2) \check{f}_j^-(z_1, z_2). \quad (7.1-16)$$

$\check{f}_j^-(z_1, z_2)$ ist die Z-Transformierte des am Mittelpunkt gespiegelten Bildsignals, d.h.

$$\check{f}_j^-(z_1, z_2) = z_1^{-N_1+1} z_2^{-N_2+1} \check{f}_j(z_1^{-1}, z_2^{-1}), \quad (7.1-17)$$

wenn $\check{f}(z_1, z_2)$ keine trivialen Faktoren hat. Mit dem Ziel, die Eindeutigkeit eines nur aus Amplituden rekonstruierten Bildes bis auf Negation und Punktspiegelung zu zeigen, werden mehrere Theoreme benötigt. $\mathcal{P}_z(N_1, N_2)$ sei die Menge aller reeller zweidimensionaler Signale auf einem endlichen Träger $[0; N_1 - 1] \times [0; N_2 - 1]$.

Theorem 4 (Hayes, 1982) Sei $f(\vec{n}) \in \mathcal{P}_z(N_1, N_2)$ mit einer Z-Transformierten der Form

$$\check{f}(z_1, z_2) = az_1^{-n_{01}} z_2^{-n_{02}} \prod_{j=1}^M \check{f}_j(z_1, z_2), \quad (7.1-18)$$

wobei $\forall j = 1, \dots, M$ $\check{f}_j(z_1, z_2)$ nichtreduzierbare und nichttriviale Polynome sind. Wenn $g(\vec{n}) \in \mathcal{P}(\vec{n})$ und $|\check{f}(e^{i\nu_1}, e^{i\nu_2})| = |\check{g}(e^{i\nu_1}, e^{i\nu_2})| \forall \vec{\nu}$, dann hat $\check{g}(e^{i\nu_1}, e^{i\nu_2})$ die Form

$$\check{g}(e^{i\nu_1}, e^{i\nu_2}) = \pm az_1^{-m_1} z_2^{-m_2} \prod_{j \in [1, M]} \check{f}_j(z_1, z_2) \cdot \prod_{j \notin [1, M]} \check{f}_j^-(z_1, z_2). \quad (7.1-19)$$

Beweis:

$$|\check{g}(e^{i\nu_1}, e^{i\nu_2})| = |\check{f}(e^{i\nu_1}, e^{i\nu_2})| \Leftrightarrow \check{f}(z_1, z_2) \check{f}(z_1^{-1}, z_2^{-1}) = \check{g}(z_1, z_2) \check{g}(z_1^{-1}, z_2^{-1}). \quad (7.1-20)$$

Ebenso wie $\check{f}(z_1, z_2)$ läßt sich auch $\check{g}(z_1, z_2)$ als

$$\check{g}(z_1, z_2) = bz_1^{-m_{01}} z_2^{-m_{02}} \prod_{j=1}^L \check{g}_j(z_1, z_2) \quad (7.1-21)$$

schreiben. Die Faktoren $\check{g}_j(z_1, z_2)$ sind nichttrivial und nichtreduzierbar. (7.1-21) und (7.1-18) in (7.1-20) eingesetzt und mit $z_1^{-N_1+1} z_2^{-N_2+1}$ multipliziert, liefert

$$a^2 z_1^{-n'_1} z_2^{-n'_2} \prod_{j=1}^M \check{f}_j(z_1, z_2) \check{f}_j^-(z_1, z_2) = b^2 z_1^{-n''_1} z_2^{-n''_2} \prod_{j=1}^L \check{g}_j(z_1, z_2) \check{g}_j^-(z_1, z_2). \quad (7.1-22)$$

Da auf beiden Seiten von (7.1-22) Polynome stehen, muß es zu jedem Faktor auf der rechten Seite einen Faktor auf der linken Seite geben, der sich von ihm nur durch eine multiplikative Konstante unterscheidet. Daraus folgt unmittelbar, daß $M = L$ und $\vec{n}' = \vec{n}''$ gilt:

$$\begin{aligned} a^2 \prod_{j=1}^M \check{f}_j(z_1, z_2) \check{f}_j^-(z_1, z_2) &= b^2 \prod_{j=1}^M \check{g}_j(z_1, z_2) \check{g}_j^-(z_1, z_2) \\ \Rightarrow g(z_1, z_2) &= cz_1^{-m_{01}} z_2^{-m_{02}} \prod_{j \in [1, M]} \check{f}_j(z_1, z_2) \cdot \prod_{j \notin [1, M]} \check{f}_j^-(z_1, z_2). \end{aligned} \quad (7.1-23)$$

Da weiterhin nach Voraussetzung $|\check{f}(e^{i\nu_1}, e^{i\nu_2})| = |\check{g}(e^{i\nu_1}, e^{i\nu_2})|$ ist, muß $c = \pm a$ folgen. \square

Für weitere Überlegungen soll $\check{f}(z_1, z_2)$ *symmetrisch* genannt werden, wenn für beliebige $m_1, m_2 > 0$

$$\check{f}(z_1, z_2) = \pm z_1^{-m_1} z_2^{-m_2} \check{f}(z_1^{-1}, z_2^{-1}) \quad (7.1-24)$$

gilt. In dem Zusammenhang ist $\check{f}(z_1, z_2)$ mit $f(\vec{n}) \in \mathcal{P}_z(N_1, N_2)$ ohne trivialen Faktoren,

$$\check{f}(z_1, z_2) = a \prod_{j=1}^M \check{f}_j(z_1, z_2), \quad (7.1-25)$$

symmetrisch, wenn

$$\check{f}(z_1, z_2) = \pm \check{f}^-(z_1, z_2) \quad (7.1-26)$$

gilt. Der allgemeine Ausschluß trivialer Faktoren gewährleistet, daß der bei (0, 0) beginnende endliche Träger nicht verschoben wird. Für die folgende Diskussion ist es hilfreich, mit einer Äquivalenzklasse der Sequenz $f(\vec{n})$ die Menge aller $g(\vec{n})$ zu bezeichnen, die für ein \vec{m} aus $f(\vec{n})$ durch

$$g(\vec{n}) = \pm f(\vec{m} \pm \vec{n}) \Leftrightarrow g(\vec{n}) \sim f(\vec{n}) \quad (7.1-27)$$

hervorgehen. Sämtliche Signale einer Äquivalenzklasse haben dasselbe Amplitudenspektrum.

Will man ein neues Signal mit dem gleichen Amplitudenspektrum generieren, so braucht man in (7.1-19) nur ein $\check{f}_j(z_1, z_2)$ durch ein $\check{f}_j^-(z_1, z_2)$ zu ersetzen. Sind die nichttrivialen und nichtreduzierbaren Faktoren $\check{f}_j(z_1, z_2)$ zudem noch symmetrisch, d.h.

$$\check{f}_j(z_1, z_2) = \pm \check{f}_j^-(z_1, z_2), \quad (7.1-28)$$

dann kann ein solcher Austausch nicht mehr als einen Vorzeichenwechsel bewirken, womit auch keine neuen Äquivalenzklassen erzeugt werden. Theorem 5 ist daher eine direkte Folge von Theorem 4, da mit einem nichtsymmetrischen und nichtreduzierbaren Faktor $\check{p}(z_1, z_2)$ genau eine Äquivalenzklasse übrigbleibt, wenn $\check{p}(z_1, z_2)$ durch $\check{p}^-(z_1, z_2)$ ersetzt wird.

Theorem 5 *Habe $\check{f}(z_1, z_2)$ mit $f(\vec{n}) \in \mathcal{P}_z(N_1, N_2)$ höchstens einen nichtreduzierbaren und unsymmetrischen Faktor, d.h. die Form*

$$\check{f}(z_1, z_2) = \check{p}(z_1, z_2) \prod_{j=1}^M \check{f}_j(z_1, z_2), \quad (7.1-29)$$

wobei $\check{p}(z_1, z_2)$ nichtreduzierbar und $\check{f}_j(z_1, z_2) \forall j = 1, \dots, M$ nichtreduzierbar und symmetrisch ist. Ist $\check{g}(\vec{n}) \in \mathcal{P}_\ddagger(N_1, N_2)$ mit $|\check{f}(e^{i\nu_1}, e^{i\nu_2})| = |\check{g}(e^{i\nu_1}, e^{i\nu_2})| \forall \vec{\nu}$, dann gilt

$$g(\vec{n}) \sim f(\vec{n}). \quad (7.1-30)$$

Theorem 6 (Hayes, 1982) Seien $f, g \in \mathcal{P}_z(N_1, N_2)$ definiert mit einem endlichen Träger

$$n_1 \in [0, N_1 - 1], n_2 \in [0, N_2 - 1], \quad (7.1-31)$$

und sei \mathcal{M} die Menge von M paarweise verschiedenen Punkten in $\mathcal{N} \times \mathcal{N}$ mit $M \geq 4(N_1 - 1)(N_2 - 1) + 1$. Hat $\check{f}(z_1, z_2)$ höchstens einen nichtreduzierbaren und nichtsymmetrischen Faktor und gilt

$$|\check{f}(e^{i\nu_1}, e^{i\nu_2})|_{\mathcal{M}} = |\check{g}(e^{i\nu_1}, e^{i\nu_2})|_{\mathcal{M}}, \quad (7.1-32)$$

dann

$$g(\vec{n}) \sim f(\vec{n}). \quad (7.1-33)$$

Beweis: M sei zunächst unbekannt. $f, g \in \mathcal{P}_z(N_1, N_2)$ haben den endlichen Träger (7.1-31). \mathcal{Z} sei hier die zu \mathcal{M} äquivalente Menge im \mathbb{C}^2 . Daher gelten die Beziehungen

$$\begin{aligned} |\check{f}(e^{i\nu_1}, e^{i\nu_2})|_{\mathcal{M}}^2 = \check{r}_f(z_1, z_2)|_{\mathcal{Z}} \quad \wedge \quad |\check{f}(e^{i\nu_1}, e^{i\nu_2})|_{\mathcal{M}}^2 &= |\check{g}(e^{i\nu_1}, e^{i\nu_2})|_{\mathcal{M}}^2 \\ \Rightarrow \check{r}_f(z_1, z_2)|_{\mathcal{Z}} &= \check{r}_g(z_1, z_2)|_{\mathcal{Z}}, \\ \check{q}_f(z_1, z_2)|_{\mathcal{Z}} &= \check{q}_g(z_1, z_2)|_{\mathcal{Z}}. \end{aligned} \quad (7.1-34)$$

$\check{q}_f(z_1, z_2)$ und $\check{q}_g(z_1, z_2)$ sind Polynome in z_1^{-1} und z_2^{-1} mit dem Summenglied höchsten Grades $\gamma_{z_1}^{-2(N_1-1)} \gamma_{z_2}^{-2(N_2-1)}$, daher gilt für die Gesamtzahl an Abtastpunkten

$$M \geq 2(N_1 - 1) \cdot 2(N_2 - 1) + 1 = 4N_1N_2 - 4N_1 - 4N_2 + 5 \quad (7.1-35)$$

(wegen des nullten Grades ist ein Polynom in einer Variablen vom Grad N durch $N + 1$ paarweise verschiedene Punkte bestimmt). Deshalb ist

$$\begin{aligned} \check{q}_f(\vec{z}) &= \check{q}_g(\vec{z}) \quad \forall \vec{z} \in \mathbb{C}^2 \\ \Rightarrow |\check{f}(e^{i\nu_1}, e^{i\nu_2})| &= |\check{g}(e^{i\nu_1}, e^{i\nu_2})| \quad \forall \vec{\nu} \in \mathbb{R}^2. \end{aligned} \quad (7.1-36)$$

Folglich gilt mit Theorem 5, wenn $\check{f}(z_1, z_2)$ höchstens einen nichtreduzierbaren und nichtsymmetrischen Faktor hat, daß

$$g(\vec{n}) \sim f(\vec{n}). \quad (7.1-37)$$

□

Damit gibt es für das Ergebnis einer Nur-Amplitudenrekonstruktion $g(\vec{n})$ eines reellen Bildsignals mit dem genannten endlichen Träger nur vier Möglichkeiten:

$$\begin{aligned} g(n_1, n_2) &= f(n_1, n_2) \quad \vee \\ g(n_1, n_2) &= -f(n_1, n_2) \quad \vee \\ g(n_1, n_2) &= f(N_1 - 1 - n_1, N_2 - 1 - n_2) \quad \vee \\ g(n_1, n_2) &= -f(N_1 - 1 - n_1, N_2 - 1 - n_2). \end{aligned} \quad (7.1-38)$$

Die wichtigste Voraussetzung für die bisherigen Überlegungen ist die Nichtreduzierbarkeit fast aller Polynome in zwei oder mehreren Variablen. Sollte sich das Polynom tatsächlich faktorisieren lassen, sollten die übrigen nichttrivialen Faktoren symmetrisch sein, um nicht noch einen höheren Grad an Mehrdeutigkeit hervorzurufen.

Es ist noch wichtig anzumerken, daß Theorem 6 keine äquidistanten Gitterpunkte verlangt, sondern nur eine Mindestzahl an paarweise verschiedenen Punkten im Frequenzraum. Dabei dient die hohe Zahl M nur dem Vergleich der Quadratamplituden, für die aufgrund der Autokorrelation, bei der sich der Träger auf $[0, 2N_1 - 1] \times [0, 2N_2 - 1]$ vergrößert, im Ortsraum entsprechend mehr Abtastwerte erforderlich sind, was einer Interpolation im Frequenzraum gleichkommt. Diese Erweiterung muß jedoch vor dem Wegwerfen der Phasen vorgenommen werden, da Abtastratenumsetzung und Betragsbildung nicht vertauscht werden dürfen.

7.2 Die Eindeutigkeit der Jetrekonstruktion

Zur Erkennung von Objekten, vor allem von Gesichtern, wird für den Vergleich lokaler Merkmale Gaboramplituden an eben diesen Stellen \vec{x}_0 herangezogen, $|\mathcal{J}_{\vec{k}}(\vec{x}_0)|$. Diese werden mit allen Skalen und Orientierungen zu einem Vektor zusammengefasst,

$$\underline{J}(\vec{n}_0) = \sum_{m=m_0}^{M_0} \sum_{l=0}^{L-1} |\mathcal{J}_{\vec{k}_{m,l}}(\vec{n}_0)| \underline{e}_{(m-m_0+1)L+l}^{ML}, \quad (7.2-1)$$

der nach (Lades et al., 1993) *Jet* genannt wird. \underline{e}_i^{ML} ist der i -te Einheitsvektor in \mathbb{R}^{ML} .

7.2.1 Gabor- und Fourieramplituden

Im folgenden soll der Zusammenhang zwischen einem Jet aller Orientierungen einer konstanten Wellenzahl k_0 hergestellt werden. Dazu wird das Bildsignal $I(\vec{x})$ mit einem isotropen Bandpaß $H_{k_0}(\vec{x})$, dessen Übertragungsfunktion beispielsweise die Form

$$\hat{H}_{k_0}(\vec{\omega}) = e^{-\frac{\sigma^2(\|\vec{\omega}\| - k_0)^2}{2k_0^2}} \quad (7.2-2)$$

haben könnte, gefiltert:

$$I_{k_0}^{BP}(\vec{x}) = I(\vec{x}) * H_{k_0}(\vec{x}). \quad (7.2-3)$$

Der nächste Schritt ist die Fensterung im Ortsraum mit der Gaußglocke

$$F_{k_0}(\vec{x}) = e^{-\frac{k_0^2 \|\vec{x}\|^2}{2\sigma^2}} \iff \hat{F}_{k_0}(\vec{\omega}) = \frac{\sigma^2}{k_0^2} e^{-\frac{\sigma^2 \|\vec{\omega}\|^2}{2k_0^2}} \quad (7.2-4)$$

an der Stelle \vec{x}_0 , wodurch das Bild

$$Q_{k_0, \vec{x}_0}(\vec{x}) = I_{k_0}^{BP}(\vec{x}) F_{k_0}(\vec{x} - \vec{x}_0) \quad (7.2-5)$$

entsteht. Dieses fouriertransformiert ergibt

$$\begin{aligned} \hat{Q}_{k_0, \vec{x}_0}(\vec{\omega}) &= \frac{1}{2\pi} \hat{I}_{k_0}^{BP}(\vec{\omega}) * \left[\hat{F}_{k_0}(\vec{\omega}) e^{-i\vec{\omega}^T \vec{x}_0} \right] \\ &= \frac{\sigma^2}{2\pi k_0^2} \int_{\mathbb{R}^2} d^2\omega' \hat{I}_{k_0}^{BP}(\vec{\omega}') e^{-\frac{\sigma^2 \|\vec{\omega} - \vec{\omega}'\|^2}{2k_0^2}} e^{i\vec{\omega}'^T \vec{x}_0} e^{-i\vec{\omega}^T \vec{x}_0}. \end{aligned} \quad (7.2-6)$$

Um die Mittelwertfreiheit des transformierten Bildes für den Vergleich mit Gaborkomponenten herzustellen, wird $\hat{Q}_{k_0, \vec{x}_0}(\vec{\omega})$ zu

$$\begin{aligned} \hat{P}_{k_0, \vec{x}_0}(\vec{\omega}) &= \hat{Q}_{k_0, \vec{x}_0}(\vec{\omega}) - \hat{Q}_{k_0, \vec{x}_0}(\vec{0}) e^{-\frac{\sigma^2 \|\vec{\omega}\|^2}{2k_0^2}} e^{-i\vec{x}_0^T \vec{\omega}} \\ &\Downarrow \\ P_{k_0, \vec{x}_0}(\vec{x}) &= Q_{k_0, \vec{x}_0}(\vec{x}) - \hat{Q}_{k_0, \vec{x}_0}(\vec{0}) \frac{k_0^2}{\sigma^2} F_{k_0}(\vec{x} - \vec{x}_0) \\ &= \left[I_{k_0}^{BP}(\vec{x}) - \frac{k_0^2}{\sigma^2} \hat{Q}_{k_0, \vec{x}_0}(\vec{0}) \right] F_{k_0}(\vec{x} - \vec{x}_0) \end{aligned} \quad (7.2-7)$$

korrigiert. Das auf diese Weise modifizierte Bildsignal $\hat{P}_{k_0, \vec{x}_0}(\vec{\omega})$ ist somit durch Bandpaßfilterung, Intensitätsabsenkung und Fensterung aus $I(\vec{x})$ hervorgegangen und soll als *Patch* bezeichnet werden. Nach Konstruktion ist $P_{k_0, \vec{x}_0}(\vec{x})$ mittelwertfrei:

$$\hat{P}_{k_0, \vec{x}_0}(\vec{0}) = 0. \quad (7.2-8)$$

Zur Weiterverarbeitung geht man nun von (7.2-7) aus:

$$\begin{aligned}
\hat{P}_{k_0, \vec{x}_0}(\vec{\omega}) &= \frac{\sigma^2}{2\pi k_0^2} \int_{\mathbb{R}^2} d^2\omega' \hat{I}_{k_0}^{BP}(\vec{\omega}') e^{-\frac{\sigma^2 \|\vec{\omega} - \vec{\omega}'\|^2}{2k_0^2}} e^{i\vec{\omega}'^T \vec{x}_0} e^{-i\vec{\omega}^T \vec{x}_0} \\
&\quad - \frac{\sigma^2}{2\pi k_0^2} e^{-\frac{\sigma^2 \|\vec{\omega}\|^2}{2k_0^2}} e^{-i\vec{x}_0^T \vec{\omega}} \int_{\mathbb{R}^2} d^2\omega' \hat{I}_{k_0}^{BP}(\vec{\omega}') e^{-\frac{\sigma^2 \|\vec{\omega}'\|^2}{2k_0^2}} e^{i\vec{\omega}'^T \vec{x}_0} \\
&= \frac{\sigma^2}{2\pi k_0^2} e^{-i\vec{\omega}^T \vec{x}_0} \cdot \\
&\quad \cdot \int_{\mathbb{R}^2} d^2\omega' \hat{I}_{k_0}^{BP}(\vec{\omega}') \left[e^{-\frac{\sigma^2 \|\vec{\omega} - \vec{\omega}'\|^2}{2k_0^2}} - e^{-\frac{\sigma^2 (\|\vec{\omega}\|^2 + \|\vec{\omega}'\|^2)}{2k_0^2}} \right] e^{i\vec{\omega}'^T \vec{x}_0}. \quad (7.2-9)
\end{aligned}$$

Hält man den Patch im Frequenzraum an den Punkten

$$\vec{\omega} = k_0 \begin{bmatrix} \cos \varphi \\ \sin \varphi \end{bmatrix} = k_0 \vec{e}(\varphi) \quad (7.2-10)$$

fest, so ergibt sich

$$\begin{aligned}
\hat{P}_{k_0, \vec{x}_0}[k_0 \vec{e}(\varphi)] &= \frac{\sigma^2}{2\pi k_0^2} e^{-ik_0 \vec{e}(\varphi)^T \vec{x}_0} \cdot \\
&\quad \cdot \int_{\mathbb{R}^2} d^2\omega' \hat{I}_{k_0}^{BP}(\vec{\omega}') \left[e^{-\frac{\sigma^2 \|\vec{\omega}' - k_0 \vec{e}(\varphi)\|^2}{2k_0^2}} - e^{-\frac{\sigma^2 (\|\vec{\omega}'\|^2 + k_0^2)}{2k_0^2}} \right] e^{i\vec{\omega}'^T \vec{x}_0} \\
&= \frac{\sigma^2}{2\pi k_0^2} e^{-ik_0 \vec{e}(\varphi)^T \vec{x}_0} \cdot \\
&\quad \cdot \int_{\mathbb{R}^2} d^2\omega' \hat{I}_{k_0}^{BP}(\vec{\omega}') \hat{\psi}_{k_0 \vec{e}(\varphi)}(\vec{\omega}') e^{i\vec{\omega}'^T \vec{x}_0} \\
&= \frac{\sigma^2}{2\pi k_0^2} e^{-ik_0 \vec{e}(\varphi)^T \vec{x}_0} \int_{\mathbb{R}^2} d^2\omega' \hat{\mathcal{J}}_{k_0 \vec{e}(\varphi)}^{BP}(\vec{\omega}') e^{i\vec{\omega}'^T \vec{x}_0} \\
&= \frac{\sigma^2}{k_0^2} e^{-ik_0 \vec{e}(\varphi)^T \vec{x}_0} \mathcal{J}_{k_0 \vec{e}(\varphi)}^{BP}(\vec{x}_0). \quad (7.2-11)
\end{aligned}$$

Das heißt, daß für den Jet des Bandpaßbildsignals an der Stelle \vec{x}_0

$$|\mathcal{J}_{k_0 \vec{e}(\varphi)}^{BP}(\vec{x}_0)| = \frac{k_0^2}{\sigma^2} |\hat{P}_{k_0, \vec{x}_0}[k_0 \vec{e}(\varphi)]| \quad (7.2-12)$$

gilt.

Die obigen Rechnungen hätten auch in abgetasteter Form vorgenommen werden können. Das Ergebnis (7.2-12) für ortsdiskrete Signale und Systeme lautet analog

$$|\mathcal{J}_{\kappa_0 \vec{e}(\varphi)}^{BP}(\vec{n}_0)| = \frac{\kappa_0^2}{\sigma^2} |\hat{P}_{\kappa_0, \vec{n}_0}[\kappa_0 \vec{e}(\varphi)]|. \quad (7.2-13)$$

Bei der diskreten Durchführung ist allerdings bei der Fensterung Vorsicht geboten, da die entsprechende Faltung im Frequenzraum Aliasing verursachen kann. Daher ist von vornherein ein verkleinertes Abtastintervall vorzusehen. Bei (7.2-12) oder (7.2-13) handelt es sich um eine Schnittstelle zwischen Fourier- und Gaboramplituden, die es ermöglicht, Aussagen über die Eindeutigkeit des aus dem Jet rekonstruierten Bildes in (7.2-13) mit Hilfe der Hayesschen Theoreme zu machen. Der Frequenzraum muß für die Eindeutigkeit nicht wie bei der diskreten Fouriertransformation mit einem Raster äquidistanter Gitterpunkte abgetastet werden, sondern die Punkte müssen nur in erforderlicher Anzahl auftreten und voneinander verschieden sein. In diesem Fall werden sie auf einem Kreis mit dem Radius κ_0 angeordnet.

Da das Gesamtbild im Ortsraum über seine endliche Pixelanzahl bestimmt ist, ist im Frequenzraum die gleiche Anzahl an Abtastwerten notwendig, wobei deren Anordnung nicht von Bedeutung

ist. Sie können also auch auf dem Kreis der Frequenzebene κ_0 liegen. Ein Patch verkleinert zwar theoretisch den Träger des Bildsignals nicht, aber die Betrachtung seiner Randbereiche ist numerisch aufgrund der Intensitätsquantisierung nicht mehr sinnvoll, weshalb der Definitionsbereich eines Patches $P_{\kappa, \vec{n}_0}(\vec{n})$ quadratisch mit der Seitenlänge N_P gegeben ist. Um einen solchen Patch aus (7.2-12) rekonstruieren zu können, muß die Gabortransformation auf der Frequenzebene κ_0 mit

$$L = 2N_P(N_P - 1) \quad (7.2-14)$$

Orientierungen durchgeführt worden sein. Da ein Patch jedoch kein allgemeines Bild ist, sondern außerhalb eines symmetrischen Frequenzbandes (nahezu) verschwindet, kann diese Zahl, wie in den folgenden Abschnitten gezeigt werden wird, weiter reduziert werden.

7.2.2 Anpassung der Abtastrate des Patches an die Bandpaßfrequenz

Aus (7.2-12) ergibt sich nun das Problem, den Patch (ein kleines Bild) im Ortsraum aus den Gaboramplituden auf der Frequenzebene k_0 zu gewinnen. Abgesehen von der eigentlichen Rekonstruktion, die in Abb. 20 erklärt ist, müssen die Abtastwerte eines Minibildes aus Punkten auf einem Frequenzring, die den Orientierungen der Wavelettransformation entsprechen. Es müßten sehr viele Orientierungen vorgesehen werden. Sie lassen sich aber durch eine Begrenzung des Bandes mit der Mittenfrequenz k_0 einschränken, damit nur die signifikanten Patchpixel berechnet werden müssen. Die Grenze wird bei einem Schwellwert der durch die Faltung in (7.2-6) verbreiterten Übertragungsfunktion \hat{H}_{k_0} gezogen (s. rechte Seite von Abb. 19). Die Fensterung bewirkt nämlich im Frequenzraum eine Glättung des Spektrums, d.h. näherungsweise verbreitert sich das Gaußprofil in radialer Richtung um $\sqrt{2}$, so daß man sich zum vorzugegebenden Schwellwert γ durch

$$\begin{aligned} \gamma &\approx e^{-\frac{\sigma^2(\omega_G - k_0)^2}{4k_0^2}} \\ \Rightarrow \omega_G &\approx k_0 \left(1 + \frac{2}{\sigma} \sqrt{\ln \frac{1}{\gamma}} \right) \\ \Rightarrow \Delta = \frac{\pi}{\omega_G} &\approx \frac{\pi\sigma}{k_0 \left(\sigma + 2\sqrt{\ln \frac{1}{\gamma}} \right)}. \end{aligned} \quad (7.2-15)$$

eine Vorstellung von der Grenzfrequenz und vom Abtastintervall des Patches \hat{P}_{k_0, \vec{x}_0} machen kann.

Um die nötige Anzahl Datenpunkte auf der Frequenzebene k_0 festzulegen, muß, abermals mit einem Schwellwert (ϵ), der auf die Fensterfunktion F_{k_0} angewandt wird, die Pixelanzahl pro Patch festgelegt werden. P_{k_0, \vec{x}_0} liegt hierbei in einem kleinem Quadrat mit der Seitenlänge a . D.h.

$$\begin{aligned} \epsilon &= e^{-\frac{\frac{a^2}{4} k_0^2}{2\sigma^2}} \\ \Rightarrow a &= \frac{2\sigma}{k_0} \sqrt{2 \ln \frac{1}{\epsilon}}. \end{aligned} \quad (7.2-16)$$

Für eine Koordinatenrichtung ergeben sich somit unabhängig von der Frequenzebene k_0

$$N_P = \frac{a}{\Delta} \approx \frac{2}{\pi} \sqrt{2 \ln \frac{1}{\epsilon}} \left(\sigma + 2\sqrt{\ln \frac{1}{\gamma}} \right) \quad (7.2-17)$$

Pixel, d.h. alle Patches speichern ihre Bildinformationen auf demselben Raster. Bei diesen Überlegungen ist vorauszusetzen, daß sich die Eingabebilder „vernünftig“ verhalten, so daß die Verteilung der Signalleistung über die Frequenzen in etwa von der erwähnten verbreiterten rotations-symmetrischen Gaußfunktion geformt wird.

7.3 Bildrekonstruktion aus Gaboramplituden

Will man ein ganzes Bild aus Gaboramplituden einer Frequenzebene rekonstruieren, müssen mehrere Patches an mehreren Punkten, z.B. auf einem Gitter, erzeugt werden. Die Rekonstruktion

jedes Patches liefert unabhängig von den anderen eine der vier durch die Hayesschen Theoreme gegebenen Möglichkeiten. Sind die einmal richtig gewählt, kann das Bandpaßsignal verlustfrei wiedergewonnen werden.

7.3.1 Rekonstruktion des Patches $P_{\kappa_0, \vec{n}_0}(\vec{n})$ aus dem Jet $|\mathcal{J}_{\kappa_0 \vec{e}(\varphi)}^{BP}(\vec{n}_0)|$

Gegeben ist der Jet an der Stelle \vec{x}_0 auf der Frequenzebene κ_0 . Er kann als Fourieramplitudenspektrum eines Patches interpretiert werden:

$$|\check{P}_{\kappa_0, \vec{n}_0}[\kappa_0 \vec{e}(\varphi)]| = \frac{\sigma^2}{\kappa_0^2} |\mathcal{J}_{\kappa_0 \vec{e}(\varphi)}^{BP}(\vec{n}_0)| \quad (7.3-1)$$

Aus (7.2-14) und (7.2-17) läßt sich die Anzahl der Orientierungen des Jets auf der Frequenzebene abschätzen. Es ist sicher günstig, diese gleichmäßig auf dem Kreis anzuordnen:

$$\varphi_j = \frac{2\pi j}{L} \quad \forall j = 0, \dots, L-1. \quad (7.3-2)$$

Das Ziel der folgenden Rechnungen ist die Extrapolation der Frequenzen auf dem Ring κ_0 in seine Umgebung, die vom Schwellwert γ (s. Abschnitt 7.2.2) nach außen begrenzt wird (s. Abb. 19). Zunächst müssen aus dem Jet die Fourieramplituden des Patches auf einem äquidistantem Gitter errechnet werden. Nach Quadrierung können die Patchamplituden aus (7.3-1) als DSFT einer Autokorrelationsfunktion auf der Frequenzebene κ_0 interpretiert werden:

$$|\check{P}_{\kappa_0, \vec{n}_0}[\kappa_0 \vec{e}(\varphi)]|^2 = \hat{r}_P[\kappa_0 \vec{e}(\varphi)] = \frac{1}{2\pi} \sum_{n_1=-N_P+1}^{N_P-1} \sum_{n_2=-N_P+1}^{N_P-1} r_P(\vec{n}) e^{-i\vec{n}^T \kappa_0 \vec{e}(\varphi)}. \quad (7.3-3)$$

Die Korrelationsfunktion im Ortsraum kann auf einem äquidistanten Raster aus $\check{r}_P(\vec{\rho})$ mit der DFT berechnet werden:

$$r_P(\vec{n}) = \frac{1}{2N_P-1} \sum_{\rho_1=-N_P+1}^{N_P-1} \sum_{\rho_2=-N_P+1}^{N_P-1} \check{r}_P(\vec{\rho}) e^{-i\frac{2\pi}{2N_P-1} \vec{\rho}^T \vec{n}}. \quad (7.3-4)$$

Hier kann ausgenutzt werden, daß eine punktsymmetrische und reelle Funktion im Ortsraum auch im Frequenzraum punktsymmetrisch und reell ist. Damit wird aus (7.3-4)

$$\begin{aligned} r_P(\vec{n}) &= \frac{1}{2N_P-1} \left(\sum_{\rho_1=-N_P+1}^{-1} + \sum_{\rho_1=1}^{N_P-1} \right) \sum_{\rho_2=-N_P+1}^{N_P-1} \check{r}_P(\vec{\rho}) e^{-i\frac{2\pi}{2N_P-1} \vec{\rho}^T \vec{n}} \\ &\quad + \frac{1}{2N_P-1} \sum_{\rho_2=-N_P+1}^{N_P-1} \check{r}_P(0, \rho_2) e^{-i\frac{2\pi}{2N_P-1} \rho_2 n_2} \\ &= \frac{1}{2N_P-1} \sum_{\rho_1=1}^{N_P-1} \sum_{\rho_2=-N_P+1}^{N_P-1} \check{r}_P(-\rho_1, \rho_2) e^{-i\frac{2\pi}{2N_P-1} (-\rho_1 n_1 + \rho_2 n_2)} \\ &\quad + \frac{1}{2N_P-1} \sum_{\rho_1=1}^{N_P-1} \sum_{\rho_2=-N_P+1}^{N_P-1} \check{r}_P(\rho_1, \rho_2) e^{-i\frac{2\pi}{2N_P-1} (\rho_1 n_1 + \rho_2 n_2)} \\ &\quad + \frac{1}{2N_P-1} \left(\sum_{\rho_2=1}^{N_P-1} + \sum_{\rho_2=-N_P+1}^{-1} \right) \check{r}_P(0, \rho_2) e^{-i\frac{2\pi}{2N_P-1} \rho_2 n_2} \\ &= \frac{1}{2N_P-1} \sum_{\rho_1=1}^{N_P-1} \sum_{\rho_2=-N_P+1}^{N_P-1} \check{r}_P(-\vec{\rho}) e^{i\frac{2\pi}{2N_P-1} \vec{\rho}^T \vec{n}} \\ &\quad + \frac{1}{2N_P-1} \sum_{\rho_1=1}^{N_P-1} \sum_{\rho_2=-N_P+1}^{N_P-1} \check{r}_P(\vec{\rho}) e^{-i\frac{2\pi}{2N_P-1} \vec{\rho}^T \vec{n}} \end{aligned}$$

$$\begin{aligned}
& + \frac{1}{2N_P - 1} \sum_{\rho_2=1}^{N_P-1} \check{r}_P(0, \rho_2) e^{-i \frac{2\pi}{2N_P-1} \rho_2 n_2} + \frac{1}{2N_P - 1} \sum_{\rho_2=1}^{N_P-1} \check{r}_P(0, \rho_2) e^{i \frac{2\pi}{2N_P-1} \rho_2 n_2} \\
= & \frac{2}{2N_P - 1} \sum_{\rho_1=1}^{N_P-1} \sum_{\rho_2=-N_P+1}^{N_P-1} \check{r}_P(\vec{\rho}) \cos \frac{2\pi \vec{\rho}^T \vec{n}}{2N_P - 1} \\
& + \frac{2}{2N_P - 1} \sum_{\rho_2=1}^{N_P-1} \check{r}_P(0, \rho_2) \cos \frac{2\pi \rho_2 n_2}{2N_P - 1} \tag{7.3-5}
\end{aligned}$$

($\check{r}_P(\vec{0}) = 0$). Darüber hinaus weiß man über die zu transformierende Funktion $\check{r}_P(\vec{\rho})$, daß sie außerhalb eines Rings verschwindet. Der Schwellwert, unterhalb welchem die Funktion zu null gesetzt wird, ist mit γ in (7.2-15) bereits festgesetzt. Dieser Ring kann nun als

$$\begin{aligned}
k_0 \left(1 - \frac{2}{\sigma} \sqrt{\ln \frac{1}{\gamma}}\right) &\leq \|\vec{\omega}\| \leq k_0 \left(1 + \frac{2}{\sigma} \sqrt{\ln \frac{1}{\gamma}}\right) &\Leftrightarrow \\
\kappa_0 \left(1 - \frac{2}{\sigma} \sqrt{\ln \frac{1}{\gamma}}\right) &\leq \|\vec{v}\| \leq \kappa_0 \left(1 + \frac{2}{\sigma} \sqrt{\ln \frac{1}{\gamma}}\right) &\Leftrightarrow \\
\rho_u = \rho_0 \left(1 - \frac{2}{\sigma} \sqrt{\ln \frac{1}{\gamma}}\right) &\leq \|\rho\| \leq \rho_0 \left(1 + \frac{2}{\sigma} \sqrt{\ln \frac{1}{\gamma}}\right) = \rho_o \tag{7.3-6}
\end{aligned}$$

und die Punktmenge, die darin definiert ist, mit

$$\mathcal{K} = \{\vec{\rho} \in \mathbb{Z}^2 \mid \rho_u \leq \|\vec{\rho}\| \leq \rho_o\} \tag{7.3-7}$$

beschrieben werden. (7.3-5) wird mit diesem Ergebnis zu

$$\begin{aligned}
r_P(\vec{n}) &= \frac{2}{2N_P - 1} \sum_{\vec{\rho} \in \mathcal{K}, \rho_1 \geq 1} \check{r}_P(\vec{\rho}) \cos \frac{2\pi \vec{\rho}^T \vec{n}}{2N_P - 1} \\
&+ \frac{2}{2N_P - 1} \sum_{\rho_u \leq \rho_2 \leq \rho_o} \check{r}_P(0, \rho_2) \cos \frac{2\pi \rho_2 n_2}{2N_P - 1} \tag{7.3-8}
\end{aligned}$$

umgeformt. Hierbei sei darauf hingewiesen, daß die Schwellwerte ϵ und γ nur relativ zu sehen sind, da sie sich nur auf die Höhe der Gaußglocke und nicht auf den tatsächlichen Signallhub beziehen. Um schließlich die Korrelationsfunktion des Patches auf einem äquidistanten Abtastraster zu erhalten, muß ein lineares Gleichungssystem, das sich durch Einsetzen von (7.3-8) in (7.3-3) ergibt,

$$\begin{aligned}
& \left| \check{P}_{\kappa_0, \vec{n}_0} [\kappa_0 \vec{e}(\varphi_j)] \right|^2 = \\
& \frac{1}{\pi(2N_P - 1)} \sum_{\vec{\rho} \in \mathcal{K}, \rho_1 \geq 1} \left| \check{P}_{\kappa_0, \vec{n}_0}^{rek}(\vec{\rho}) \right|^2 \sum_{n_1=-N_P+1}^{N_P-1} \sum_{n_2=-N_P+1}^{N_P-1} e^{-i \vec{n}^T \kappa_0 \vec{e}(\varphi_j)} \cos \frac{2\pi \vec{\rho}^T \vec{n}}{2N_P - 1} \\
& + \frac{1}{\pi(2N_P - 1)} \sum_{\rho_u \leq \rho_2 \leq \rho_o} \left| \check{P}_{\kappa_0, \vec{n}_0}^{rek}(0, \rho_2) \right|^2 \sum_{n_1=-N_P+1}^{N_P-1} \sum_{n_2=-N_P+1}^{N_P-1} e^{-i \vec{n}^T \kappa_0 \vec{e}(\varphi_j)} \cos \frac{2\pi \rho_2 n_2}{2N_P - 1} \\
& \forall j = 0, \dots, L' - 1, \tag{7.3-9}
\end{aligned}$$

gelöst werden. L' gibt somit die Anzahl der Orientierungen φ_j an, weshalb, um das System (7.3-9) eindeutig zu lösen,

$$\begin{aligned}
L' &= \left| \mathcal{K} \cap [\{\vec{\rho} \in \mathbb{Z}^2 \mid 1 \leq \rho_1 \leq N_P - 1 \wedge -N_P + 1 \leq \rho_2 \leq N_P - 1\} \right. \\
&\quad \left. \cup \{\vec{\rho} \in \mathbb{Z}^2 \mid \rho_1 = 0 \wedge 1 \leq \rho_2 \leq N_P - 1\}] \right| \tag{7.3-10}
\end{aligned}$$

gilt. Da die Systemmatrix aus (7.3-9) nur für $L' = \frac{1}{2}L$ vollen Rang hat, kann in diesem System nur die Hälfte der Orientierungen berücksichtigt werden, was mit der Ausnutzung der Symmetrie der

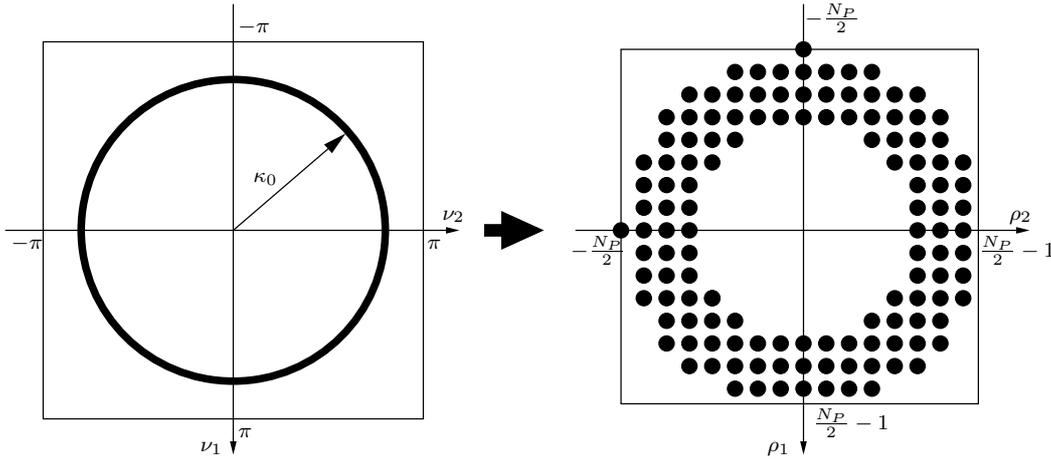


Abbildung 19: Wirkung des Gleichungssystems (7.3-9).

Korrelationsfunktionen im Frequenzraum zu tun hat. Die Operation $|\cdot|$ auf eine Menge angewandt, liefert die Anzahl ihrer Elemente. Die Elemente der Systemmatrix zum Gleichungssystem (7.3-9) lassen sich durch eine diskrete Kosinustransformation berechnen. Es bleibt noch anzumerken, daß die Summation im diskreten Frequenzraum sich nicht über das gesamte Intervall $[-N_P + 1; N_P - 1]$ erstreckt, sondern nur jeweils etwa bis zur Hälfte.

Hat man eine Lösung gewonnen, kann jeder Patch aus seinem Amplitudenspektrum mit Algorithmen, wie z.B. aus (Fienup, 1987), rekonstruiert werden. Ein solches „*phase retrieval*“ (s. Abb. 20) liefert näherungsweise ein Element der Menge

$$\{P_{\kappa_0, \vec{n}_0}(n_1, n_2), -P_{\kappa_0, \vec{n}_0}(n_1, n_2), P_{\kappa_0, \vec{n}_0}(N_P - 1 - n_1, N_P - 1 - n_2), -P_{\kappa_0, \vec{n}_0}(N_P - 1 - n_1, N_P - 1 - n_2)\}, \quad (7.3-11)$$

d. h. eine bandgefilterte, gefensterter, abgesenkte und dezimierte Version von $I(\vec{n})$. Die obigen Überlegungen sowie Theorem 6 ergeben, daß ein Jet ein durch eine Gaborwavellettransformation extrahiertes lokales Bildmerkmal an einer festen Stelle \vec{n}_0 und für eine Skala mit der Mehrdeutigkeit aus (7.3-11) kodiert.

Anzumerken bleibt noch, daß die Anzahl der verlangten Orientierungen gleich der Anzahl von Pixel für das gefensterter Bildsignal ist. L wird also direkt von σ vorgegeben.

7.3.2 Rekonstruktion des Bildes $I(\vec{n})$ aus den Patches $P_{\kappa_0, \vec{n}_0}^{rek}(\vec{n})$

Die oben erläuterte Rekonstruktion liefert auf einer Frequenzebene und an einem Ort unabhängig von den anderen einen Patch von vier möglichen Varianten (s. (7.3-11)). Patches zweier benachbarter Skalen kann ein gemeinsames Frequenzband im Überlappbereich extrahiert werden, um aus der Menge (7.3-11) die zu der anderen Skala kompatible Funktion auszugewählen. Dies kann durch

- einen Vergleich der Phasen im Frequenzraum
- oder nach einer Dezimation des Patches der kleineren Skala, so daß ein Vergleich von wenigen Pixeln im Überlappbereich zwischen den Frequenzebenen ausreicht,

erfolgen, wobei der zweite Vorschlag als der zuverlässigere anzusehen ist. Das Ergebnis besteht aus M (Anzahl der Skalen, Frequenzebenen) Patches, die sich auf eine Möglichkeit aus (7.3-11) geeinigt haben.

Da die Leistungsfähigkeit dieses Algorithmus stark bilddatenabhängig ist, kann ohne weiteres nicht bewiesen werden, daß ein Jet durch die Entfernung der Phase genau vier Möglichkeiten als lokales Merkmal zuläßt.

Der nächste Schritt wäre, jeden Patch mit seinen Nachbarn im Ortsraum zu vergleichen, und zwar auf allen M Skalen separat, wobei es sich empfiehlt, wegen des größeren Überlappbereichs

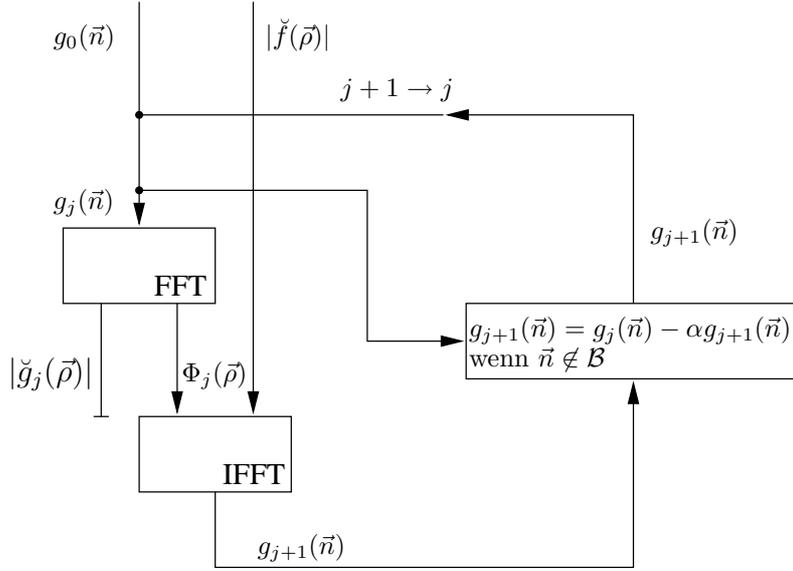


Abbildung 20: Der Phase-Retrieval-Algorithmus nach (Fienup, 1987). Eingegeben werden die Amplitudenfunktion und ein Initialisierungsbild (z.B. Realisierung eines stationären Rauschsignals). Diesem Bild wird in jedem Iterationsschritt die Phase extrahiert, und diese mit der Amplitudenfunktion kombiniert. Da im Ortsraum ein endlicher Träger eingehalten werden muß, werden die Signalanteile, die diesen überschreiten, korrigiert (in einer alternativen Form werden sie zu null gesetzt).

bei niedrigen Frequenzebenen anzufangen. Dies entspricht einer lokalen Optimierung, da nur die acht nächsten Nachbarn für das an einer bestimmten Stelle gefensterte Bild relevant sind. Liefert dies keine befriedigenden Ergebnisse, so muß ein globales Fehlerkriterium über alle Konfigurationen optimiert werden. Ist eine richtige „Ausrichtung“ gefunden, kann das Eingabebild $I(\vec{n})$ mit dem im folgenden beschriebenen Verfahren zurückgewonnen werden.

Um die Notation nicht allzu unübersichtlich werden zu lassen, werden hier Bildsignale mit quadratischem Format ($N_1 = N_2 = N$) angenommen. Eine völlige Diskretisierung wird erreicht, wenn nach (A.3-4) $\vec{\rho} := \frac{N}{2\pi} \vec{\nu}$ und analog $\vec{\eta} := \frac{N}{2\pi} \vec{\kappa}$ umgewandelt wird. Des weiteren soll nicht nur eine Skala betrachtet werden, weswegen der Frequenzebenenindex 0 durch $m = m_0, \dots, M_0$ ersetzt wird. Führt man die Rechnungen aus Abschnitt 7.2.1 konsequent mit der DFT durch, lautet der diskrete Patch im Frequenzraum

$$\check{P}_{\eta_m, \vec{n}_0}^{rek}(\vec{\rho}) = e^{-i \frac{2\pi}{N} \vec{\rho}^T \vec{n}_0} \frac{\sigma^2 N}{4\pi^2 \eta_m^2} \sum_{\vec{\rho}' \in \mathcal{B}} \check{I}_{\eta_m}(\vec{\rho}') \left[e^{-\frac{\sigma^2 \|\vec{\rho} - \vec{\rho}'\|^2}{2\eta_m^2}} - e^{-\frac{\sigma^2 (\|\vec{\rho}\|^2 + \|\vec{\rho}'\|^2)}{2\eta_m^2}} \right] e^{i \frac{2\pi}{N} \vec{\rho}'^T \vec{n}_0}. \quad (7.3-12)$$

Das weitere Vorgehen folgt weitgehend einer Framerekonstruktion. Zu diesem Zweck wird (7.3-12) mit $e^{-i \frac{2\pi}{N} \vec{\rho}^T \vec{n}_0} \eta_m^2$ multipliziert und anschließend einer DFT unterzogen:

$$\frac{\eta_m^2}{N} \sum_{\vec{n}_0 \in \mathcal{B}} \check{P}_{\eta_m, \vec{n}_0}^{rek}(\vec{\rho}) e^{i \frac{2\pi}{N} \vec{\rho}^T \vec{n}_0} e^{-i \frac{2\pi}{N} \vec{\rho}'^T \vec{n}_0} = \frac{\sigma^2 N^2}{4\pi^2} \check{I}(\vec{\rho}') \check{H}_{\eta_m}(\vec{\rho}') \left[e^{-\frac{\sigma^2 \|\vec{\rho} - \vec{\rho}'\|^2}{2\eta_m^2}} - e^{-\frac{\sigma^2 (\|\vec{\rho}\|^2 + \|\vec{\rho}'\|^2)}{2\eta_m^2}} \right] \quad (7.3-13)$$

Weiterhin wird (7.3-13) mit $\check{H}_{\eta_m}(\vec{\rho}')$ $\left[e^{-\frac{\sigma^2 \|\vec{\rho} - \vec{\rho}'\|^2}{2\eta_m^2}} - e^{-\frac{\sigma^2 (\|\vec{\rho}\|^2 + \|\vec{\rho}'\|^2)}{2\eta_m^2}} \right]$ multipliziert und über m und $\vec{\rho}$ summiert, wobei die Abkürzung

$$Y(\vec{\rho}') = \sum_{m=m_0}^{M_0} \sum_{\vec{\rho} \in \mathcal{B}} \check{H}_{\eta_m}(\vec{\rho}')^2 \left[e^{-\frac{\sigma^2 \|\vec{\rho} - \vec{\rho}'\|^2}{2\eta_m^2}} - e^{-\frac{\sigma^2 (\|\vec{\rho}\|^2 + \|\vec{\rho}'\|^2)}{2\eta_m^2}} \right]^2 \quad (7.3-14)$$

benutzt wird. Das Eingabebild kann nun nach

$$\check{I}(\vec{\rho}') = \frac{4\pi^2}{\sigma^2 N^3} Y^{-1}(\vec{\rho}') \sum_{\vec{n}_0 \in \mathcal{B}} \sum_{\vec{\rho} \in \mathcal{B}} \sum_{m=m_0}^{M_0} \eta_m^2 \check{P}_{\eta_m, \vec{n}_0}^{rek}(\vec{\rho}) e^{i \frac{2\pi}{N} \vec{\rho}^T \vec{n}_0} e^{-i \frac{2\pi}{N} \vec{\rho}'^T \vec{n}_0} \cdot \check{H}_{\eta_m}(\vec{\rho}') \left[e^{-\frac{\sigma^2 \|\vec{\rho} - \vec{\rho}'\|^2}{2\eta_m^2}} - e^{-\frac{\sigma^2 (\|\vec{\rho}\|^2 + \|\vec{\rho}'\|^2)}{2\eta_m^2}} \right] \quad (7.3-15)$$

⌋

$$I(\vec{n}) = \frac{4\pi^2}{\sigma^2 N^4} \sum_{\vec{n}_0 \in \mathcal{B}} \sum_{\vec{\rho} \in \mathcal{B}} \sum_{m=m_0}^{M_0} \eta_m^2 \check{P}_{\eta_m, \vec{n}_0}^{rek}(\vec{\rho}) e^{i \frac{2\pi}{N} \vec{\rho}^T \vec{n}_0} \cdot \sum_{\vec{\rho}' \in \mathcal{B}} Y^{-1}(\vec{\rho}') \check{H}_{\eta_m}(\vec{\rho}') \left[e^{-\frac{\sigma^2 \|\vec{\rho} - \vec{\rho}'\|^2}{2\eta_m^2}} - e^{-\frac{\sigma^2 (\|\vec{\rho}\|^2 + \|\vec{\rho}'\|^2)}{2\eta_m^2}} \right] e^{i \frac{2\pi}{N} \vec{\rho}'^T (\vec{n} - \vec{n}_0)} \quad (7.3-16)$$

$$= \text{IPT}\{\dots\}(\vec{n}) \quad (7.3-17)$$

wiedergewonnen werden.

Nach (7.3-12) kann ein Patch als eine Transformation über ein Skalarprodukt geschrieben werden:

$$\begin{aligned} \check{P}_{\eta_m, \vec{n}_0}(\vec{\rho}) &= \text{DFT}\{\text{PT}\{I(\vec{n}')\}(\vec{n}, \eta_m, \vec{n}_0)\}(\vec{\rho}, \eta_m, \vec{n}_0) \\ &= \left\langle \check{I}, \frac{\sigma^2 N}{4\pi^2 \eta_m^2} \check{H}_{\eta_m} \left[e^{-\frac{\sigma^2 \|\vec{\rho} - \cdot\|^2}{2\eta_m^2}} - e^{-\frac{\sigma^2 (\|\vec{\rho}\|^2 + \|\cdot\|^2)}{2\eta_m^2}} \right] e^{i \frac{2\pi}{N} (\cdot - \vec{\rho})^T \vec{n}_0} \right\rangle. \end{aligned} \quad (7.3-18)$$

Damit die Rekonstruktionsformel (7.3-15) tatsächlich eine vernünftige Lösung liefert, sollten kurz die Eigenschaften von $Y(\vec{\rho}')$ bezüglich einer Inversion betrachtet werden. Die einzelnen Summenglieder bestehen aus den Quadraten von leicht verzerrten Gaußfunktionen, multipliziert mit jeweils einem Ring, der ebenfalls das Profil einer Gaußfunktion hat. Da alle diese Funktionen nach oben beschränkt sind, gilt dies auch für $Y(\vec{\rho}')$. Auch die Beschränktheit nach unten kann durch Überlegungen plausibel gemacht werden: Die verzerrten Gaußglocken verschiedener Breite werden an jeden Frequenzpunkt verschoben und, mit dem Bandpaß gewichtet, der Summe (7.3-14) hinzugefügt. Diese Aussagen motivieren

$$0 < Y(\vec{\rho}') < \infty. \quad (7.3-19)$$

Es bleibt nur noch die Rolle der Bandpaßfilter $\check{H}_{\eta_m}(\vec{\rho})$ zu klären. Zu einer numerisch stabilen Rekonstruktion von $I(\vec{n})$ in (7.3-16) leisten sie offensichtlich keinen Beitrag. Ihre Aufgabe besteht vielmehr darin, für eine Einschränkung der Frequenzpunkte, die von einem linearen Gleichungssystem (7.3-9) berechnet werden, zu sorgen. Bei einem endlichen Signalträger im Ortsraum können theoretisch sämtliche Punkte von einem Ring im (kontinuierlichen) Frequenzraum berechnet werden. Dabei darf allerdings nicht außer acht gelassen werden, daß Frequenzpunkte, die von diesem Ring, d.h. einer festen Skala der Gaborwavelettransformierten, weit entfernt liegen, stark fehlerbehaftete Rekonstruktionen liefern, vor allem bei zusätzlichem Rauschen. Daher rührt die Einschränkung auf den in (7.3-6) angegebenen Bereich.

8 Bildrekonstruktion aus Antworten komplexer Zellen

8.1 Erweiterung der Hayesschen Theoreme auf komplexwertige Sequenzen

Nach Theorem 3 kommt es so gut wie überhaupt nicht vor, daß sich ein über mehrere Variablen definiertes Polynom faktorisieren läßt. Trotzdem werden in Theorem 4 mehrere Faktoren angesetzt. Diese Abschwächung der Voraussetzungen an die Bildsignale soll im folgenden fallengelassen werden.

Theorem 7 Seien $f, g \in \mathcal{P}_z^c(N_1, N_2)$, also komplexwertig, definiert auf einem endlichen Träger \mathcal{B} und seien $\check{f}(\vec{\nu})$ und $\check{g}(\vec{\nu})$ nur trivial faktorisiert sowie $|\check{f}(\vec{\nu})|^2$ und $|\check{g}(\vec{\nu})|^2$ bekannt auf der Menge \mathcal{M} von $(2N_1 - 1)(2N_2 - 1)$ Punkten mit

$$|\check{f}(\vec{\nu})|^2 = |\check{g}(\vec{\nu})|^2 \quad \forall \vec{\nu} \in \mathcal{M}, \quad (8.1-1)$$

dann gilt

$$g(n_1, n_2) = e^{in} f(n_1, n_2) \quad \vee \quad g(n_1, n_2) = e^{in} f^*(N_1 - 1 - n_1, N_2 - 1 - n_2). \quad (8.1-2)$$

Beweis: Es gelte

$$\begin{aligned} |\check{f}(\vec{\nu})|^2 &= |\check{g}(\vec{\nu})|^2 \quad \forall \vec{\nu} \in \mathcal{M} \\ \Leftrightarrow \check{f}(\vec{\nu})\check{f}^*(\vec{\nu}) &= \check{g}(\vec{\nu})\check{g}^*(\vec{\nu}) \quad \forall \vec{\nu} \in \mathcal{M}, \end{aligned} \quad (8.1-3)$$

$$\check{f}^c(\vec{\nu}) = \text{DSFT}\{f^c(\vec{n})\}(\vec{\nu}) = \text{DSFT}\{f^*(\vec{n})\}(\vec{\nu}) = \check{f}^*(-\vec{\nu}). \quad (8.1-4)$$

Analog zu Theorem 4 benutzt man hier als Ansatz

$$\begin{aligned} \check{f}(\vec{\nu}) &= |\alpha| e^{i\beta} e^{-i\nu_1 p_1} e^{-i\nu_2 p_2} \check{f}_1(\nu_1, \nu_2), \\ \check{g}(\vec{\nu}) &= |\gamma| e^{i\zeta} e^{-i\nu_1 q_1} e^{-i\nu_2 q_2} \check{g}_1(\nu_1, \nu_2). \end{aligned} \quad (8.1-5)$$

$\check{f}_1(\nu_1, \nu_2)$ sei nur trivial faktorisiert und normiert. Dasselbe gilt für $\check{g}_1(\nu_1, \nu_2)$. (8.1-5) in (8.1-3) eingesetzt liefert mit (8.1-4) den Zusammenhang

$$\begin{aligned} |\alpha| e^{i\beta} e^{-i\nu_1 p_1} e^{-i\nu_2 p_2} \check{f}_1(\nu_1, \nu_2) \cdot \\ \cdot |\alpha| e^{-i\beta} e^{i\nu_1 p_1} e^{i\nu_2 p_2} \check{f}_1^c(-\nu_1, -\nu_2) &= |\gamma| e^{i\zeta} e^{-i\nu_1 q_1} e^{-i\nu_2 q_2} \check{g}_1(\nu_1, \nu_2) \cdot \\ &\cdot |\gamma| e^{-i\zeta} e^{i\nu_1 q_1} e^{i\nu_2 q_2} \check{g}_1^c(-\nu_1, -\nu_2) \quad \forall \vec{\nu} \in \mathcal{M} \\ |\alpha|^2 \check{f}_1(\nu_1, \nu_2) \check{f}_1^c(-\nu_1, -\nu_2) &= |\gamma|^2 \check{g}_1(\nu_1, \nu_2) \check{g}_1^c(-\nu_1, -\nu_2) \quad \forall \vec{\nu} \in \mathcal{M}. \end{aligned} \quad (8.1-6)$$

Da die Polynome auf beiden Seiten normiert sind, muß

$$\Rightarrow \quad |\gamma| = |\alpha| \quad (8.1-7)$$

gelten. Aufgrund ihrer Nichtfaktorierbarkeit ist einem Polynom auf der linken eines auf der rechten Seite von (8.1-6) zuzuordnen. Daher sind zwei Fälle zu unterscheiden:

$$\begin{aligned} 1. \quad &\check{g}_1(\nu_1, \nu_2) = \check{f}_1(\nu_1, \nu_2) \\ \Rightarrow &\check{g}(\nu_1, \nu_2) = |\alpha| e^{i\zeta} e^{-i\nu_1 q_1} e^{-i\nu_2 q_2} \check{f}_1(\nu_1, \nu_2) \\ &\Downarrow \\ &g(n_1, n_2) = |\alpha| e^{i\zeta} f_1(n_1 - q_1, n_2 - q_2) \\ &= e^{i(\zeta - \beta)} f(n_1 - q_1 + p_1, n_2 - q_2 + p_2) \end{aligned} \quad (8.1-8)$$

$$\begin{aligned} 2. \quad &\check{g}_1(\nu_1, \nu_2) = \check{f}_1^c(-\nu_1, -\nu_2) \\ \Rightarrow &\check{g}(\nu_1, \nu_2) = |\alpha| e^{i\zeta} e^{-i\nu_1 q_1} e^{-i\nu_2 q_2} \check{f}_1^c(-\nu_1, -\nu_2) \\ &\Downarrow \\ &g(n_1, n_2) = |\alpha| e^{i\zeta} f_1^c(-n_1 + q_1, -n_2 + q_2) \\ &= |\alpha| e^{in} f_1^*(-n_1 + q_1, -n_2 + q_2) \\ &= e^{i(\zeta - \beta)} f^*(-n_1 + q_1 + p_1, -n_2 + q_2 + p_2) \end{aligned} \quad (8.1-9)$$

Daß diese beiden Möglichkeiten miteinander verträglich sind, kann durch Nachrechnen gezeigt werden. Als Einschränkung kann der endliche Träger, auf dem die Bilddaten definiert sind, hinzugezogen werden:

$$1. \quad g(n_1, n_2) = e^{in} f(n_1, n_2) \quad \vec{n} \in \mathcal{B} \quad (8.1-10)$$

$$2. \quad g(n_1, n_2) = e^{in} f^*(N_1 - 1 - n_1, N_2 - 1 - n_2) \quad \vec{n} \in \mathcal{B} \quad (8.1-11)$$

mit $\eta = \zeta - \beta$. □

Somit bestehen unendlich viele Rekonstruktionsmöglichkeiten, da η im Intervall $[-\pi; \pi[$ kontinuierlich gewählt werden kann. Die zusammenfassende Schreibweise für diese Freiheitsgrade ist

$$g(n_1, n_2) = \{f(n_1, n_2)\}_{b, \eta}. \quad (8.1-12)$$

b ist zweiwertig, d.h. es liegt für $b = 1$ die richtige und für $b = 0$ die konjugierte punktgespiegelte Version vor. Dieses Theorem setzt die Nichtzerlegbarkeit der Fouriertransformierten in nichttriviale Faktoren voraus. Theorem 3 stellt dies soweit sicher, daß es für die Anwendung ausreichen dürfte.

8.2 Die Eindeutigkeit der Bildrekonstruktion von Amplituden der DGWT

Das folgende Theorem zeigt, daß bis auf eine eventuelle Negation kein Informationsverlust bei der Bildkodierung mit komplexen Zellen stattfindet. Die Einschränkungen sind minimal, so daß natürliche Bilder davon weitgehend unberührt bleiben.

Theorem 8 *Bilden die diskreten Gaborwavelets $\chi_{\vec{n}_0 \Delta, a_0^m, \frac{2\pi l}{L}}(\vec{n})$ einen Frame im Hilbertraum $\bar{\mathcal{H}}_I$ und sei $\langle \check{I}, \check{\chi}_{\vec{n}_0 \Delta, a_0^m, \frac{2\pi l}{L}} \rangle$ ein nur trivial faktorisiertes Polynom, so kann das diskrete reelle Bildsignal $I(\vec{n}) \in \mathcal{H}_I$ bis auf das Vorzeichen eindeutig aus den Amplituden der Gaborwavelettransformierten*

$$|\mathcal{I}_0(\vec{n}_0, m, l)| = \left| \langle I, \chi_{\vec{n}_0 \Delta, a_0^m, \frac{2\pi l}{L}} \rangle \right| \quad (8.2-1)$$

rekonstruiert werden:

$$|\mathcal{I}_0(\vec{n}_0, m, l)| = |\mathcal{I}_0^{rec}(\vec{n}_0, m, l)| \quad \forall \vec{n}_0 \in \mathcal{B}^{int} \quad \Rightarrow \quad I^{rek}(\vec{n}) = \pm I(\vec{n}). \quad (8.2-2)$$

$\bar{\mathcal{H}}_I$ ist die Menge aller Bildsignale ohne Mittelwert $I(\vec{n})$ mit $n_1 \in [0; N_1 - 1] \subset \mathbb{Z}$ und $n_2 \in [0; N_2 - 1] \subset \mathbb{Z}$, \mathcal{B}^{int} der Träger des um 2 interpolierten Bildes. Dabei unterliegt $I(\vec{n})$ folgenden Einschränkungen:

- N_1 ist gerade, N_2 ist ungerade:

$$\check{I}(\rho_1, 0) \neq 0 \quad \text{für ein } \rho_1. \quad (8.2-3)$$

- N_1 ist ungerade, N_2 ist gerade:

$$\check{I}(0, \rho_2) \neq 0 \quad \text{für ein } \rho_2. \quad (8.2-4)$$

- N_1 und N_2 gerade:

$$\check{I}(\rho_1, \rho_1) \neq 0 \quad \text{für ein } \rho_1. \quad (8.2-5)$$

Beweis: Das Skalarprodukt zur Durchführung der diskreten Gabortransformation kann mit dem Plancherelschen Theorem (A.3-3) ebenso im Frequenzraum formuliert werden (6.2-6), wodurch die (6.2-6) die Gestalt einer IDFT hat. Nimmt man nur die Amplituden, so erhält man

$$|\mathcal{I}_0(\vec{n}_0, m, l)| = \left| \frac{2\pi}{\sqrt{N_1 N_2}} \frac{1}{\Delta^2} \sum_{\vec{\rho} \in \mathcal{B}} \check{I}(\vec{\rho}) a_0^m \hat{\chi}^* \left[\mathcal{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] e^{i2\pi \vec{\rho}^T \vec{n}_0} \right|. \quad (8.2-6)$$

Da auf dem üblichen DFT-Raster das Ortsraumsignal innerhalb $||$ in (8.2-6), also hier im Transformationsraum bei fester Skala und Orientierung, ein Polynom in \vec{n}_0 ist, kann Theorem 7 angewendet werden. Demnach ist die Wiedergewinnung von $I(\vec{n})$ aus Funktion (8.2-6) eindeutig bis auf eine

additive, kontinuierliche und konstante Phase und eine Punktspiegelung mit komplexer Konjugation:

$$\begin{aligned} \left\{ \check{\mathcal{I}}_0(\vec{\rho}, m, l) \right\}_{b(m,l), \eta(m,l)} &= \left\{ \frac{2\pi}{\Delta^2} \check{I}(\vec{\rho}) a_0^m \hat{\chi}^* \left[\mathcal{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] \right\}_{b(m,l), \eta(m,l)} \\ &= \left\{ \left| \check{\mathcal{I}}_0(\vec{\rho}, m, l) \right| \right\}_{b(m,l)} e^{i \left\{ \Phi_{\check{\mathcal{I}}_0}(\vec{\rho}, m, l) \right\}_{b(m,l), \eta(m,l)}}. \end{aligned} \quad (8.2-7)$$

Ziel der Überlegungen ist, ein $\check{\mathcal{I}}_0(\vec{\rho}, m, l)$ zu finden, mit dem durch eine IDGWT $\check{I}^{rek}(\vec{\rho})$ bzw. $I^{rek}(\vec{n})$ rekonstruiert werden kann. Im folgenden soll die Mehrdeutigkeit durch Anwendung von Vorwissen entfernt werden. Ausgenutzt werden Symmetrieeigenschaften, die mit der Voraussetzung, daß das Eingabebild $I(\vec{n})$ reell ist, zusammenhängen. In einem ersten Schritt wird jeweils das Element mit dem $b(m, l)$ entfernt, das der Bedingung bezüglich der Amplituden nicht genügt. Der zweite Schritt läßt durch Betrachtung der Phasenfunktion nur das passende Element zu $\eta(m, l)$ übrig. Dabei soll die Menge der zu transformierenden Bildsignale keine weiteren Einschränkungen erfahren.

- Der Betrag von (8.2-7) hängt natürlich nicht mehr von η ab:

$$\left\{ \left| \check{\mathcal{I}}_0(\vec{\rho}, m, l) \right| \right\}_{b(m,l)} = \left\{ \frac{2\pi}{\Delta^2} \left| \check{I}(\vec{\rho}) \right| a_0^m \left| \hat{\chi} \left[\mathcal{Q} \left(\frac{2\pi l}{L} \right) \vec{\rho} \frac{2\pi}{\Delta} a_0^m \right] \right| \right\}_{b(m,l)}. \quad (8.2-8)$$

$|\check{I}(\vec{\rho})|$ ist eine punktsymmetrische Funktion (von Zeilen und Spalten an den Rändern in manchen Fällen abgesehen). Zur Erkennung von Spiegelungen muß eine Fallunterscheidung vorgenommen werden:

1. N_1 und N_2 sind ungerade. In diesem Fall ist es günstig, den Ausschnitt des periodischen Frequenzraums so zu verschieben, daß die Frequenz $\vec{\rho} = \vec{0}$ genau in der Mitte liegt (\vec{B}). $|\check{I}(\vec{\rho})|$ ist für reelle Bilder $I(\vec{n})$ dann eine punktsymmetrische Funktion, d.h. ihre Spiegelung um den Mittelpunkt ändert nichts. Da m und l bekannt sind, wird nun die einseitige Energieverteilung von $|\check{\mathcal{I}}_0(\vec{\rho}, m, l)|$ dahingehend genutzt, daß das Maximum von (8.2-8) auf der richtigen Seite der Geraden $\vec{\rho}^T \vec{n}$ liegt:

$$\begin{aligned} \left\{ \left| \check{\mathcal{I}}_0(\vec{\rho}, m, l) \right| \right\}_{b(m,l)} &> \left\{ \left| \check{\mathcal{I}}_0(-\vec{\rho}, m, l) \right| \right\}_{b(m,l)} \quad ? \\ \text{ja:} \quad \left\{ \check{\mathcal{I}}_0^{rek}(\rho_1, \rho_2, m, l) \right\}_{\eta(m,l)} &:= \left\{ \left| \check{\mathcal{I}}_0(\rho_1, \rho_2, m, l) \right| \right\}_{b(m,l)} e^{i \left\{ \Phi_{\check{\mathcal{I}}_0}(\vec{\rho}, m, l) \right\}_{\eta(m,l)}} \\ \text{nein:} \quad \left\{ \check{\mathcal{I}}_0^{rek}(\rho_1, \rho_2, m, l) \right\}_{\eta(m,l)} &:= \left\{ \left| \check{\mathcal{I}}_0(N_1 - 1 - \rho_1, N_2 - 1 - \rho_2, m, l) \right| \right\}_{b(m,l)} \\ &\quad \cdot e^{-i \left\{ \Phi_{\check{\mathcal{I}}_0}(N_1 - 1 - \rho_1, N_2 - 1 - \rho_2, m, l) \right\}_{\eta(m,l)}}. \end{aligned} \quad (8.2-9)$$

Dieses Kriterium reicht aus, da das Quadrat der fouriertransformierten Waveletfunktion auf der einen Seite immer größer als die zu einem um $\vec{\rho} = \vec{0}$ gespiegelter Punkt auf der anderen Seite bezüglich eines Winkels ist, wie in Anhang B gezeigt ist. Aus der Menge $\{\cdot\}_{b(m,l)}$ wird nun das Element von beiden gewählt, für welches (8.2-9) richtig ist.

2. N_1 ist gerade, N_2 ist ungerade. Der betrachtete Frequenzraumausschnitt ist nun

$$\left[-\frac{1}{2}N_1; \frac{1}{2}N_1 - 1 \right], \quad \left[-\frac{1}{2}(N_2 - 1); \frac{1}{2}(N_2 - 1) \right]. \quad (8.2-10)$$

Bei einer Punktspiegelung wie in (8.1-11) verschiebt sich, im Gegensatz zum obigen Fall, der Punkt $\vec{\rho} = \vec{0}$, weshalb dieser Fall ein anderes Vorgehen erfordert. Gilt

$$\left\{ \left| \check{\mathcal{I}}_0(\vec{0}, m, l) \right| \right\}_{b(m,l)} \neq 0 \quad (8.2-11)$$

für das eine Mengenelement, so ist das andere das richtige. Gilt (8.2-11) nicht, so muß man sich für ebendieses entscheiden. Bei einer Punktspiegelung würde $\vec{\rho} = \vec{0}$ auf $\vec{\rho} = [-1 \ 0]^T$ abgebildet werden. Haben diese beiden Punkte den Wert 0, so muß aufgrund der Punktsymmetrie von $|\check{I}(\vec{\rho})|$ die gesamte senkrechte Gerade 0 sein:

$$|\check{I}(\rho_1, 0)| = 0 \quad \forall \rho_1 \in \mathbb{Z}. \quad (8.2-12)$$

Dieser Fall ist jedoch durch die Voraussetzungen ausgeschlossen. Diese Betrachtungen werden durch die Abb. 21(a) und 21(b) illustriert. Für $l = 0$ und $l = \frac{L}{2}$ muß daher das Kriterium

$$\begin{aligned} \max_{\vec{\rho} \in \check{\mathcal{B}} | \vec{\rho}^T \vec{e}(\frac{2\pi l}{L}) > 0} \left\{ |\check{\mathcal{I}}_0(\vec{\rho}, m, l)| \right\}_{b(m,l)} &> \max_{\vec{\rho} \in \check{\mathcal{B}} | \vec{\rho}^T \vec{e}(\frac{2\pi l}{L}) < 0} \left\{ |\check{\mathcal{I}}_0(\vec{\rho}, m, l)| \right\}_{b(m,l)} \quad ? \\ \text{ja: } \left\{ \check{\mathcal{I}}_0^{rek}(\rho_1, \rho_2, m, l) \right\}_{\eta(m,l)} &:= \left\{ |\check{\mathcal{I}}_0(\rho_1, \rho_2, m, l)| \right\}_{b(m,l)} e^{i\{\Phi_{\check{\mathcal{I}}_0}(\vec{\rho}, m, l)\}_{\eta(m,l)}} \\ \text{nein: } \left\{ \check{\mathcal{I}}_0^{rek}(\rho_1, \rho_2, m, l) \right\}_{\eta(m,l)} &:= \left\{ |\check{\mathcal{I}}_0(N_1 - 1 - \rho_1, N_2 - 1 - \rho_2, m, l)| \right\}_{b(m,l)} \cdot \\ &\cdot e^{-i\{\Phi_{\check{\mathcal{I}}_0}(N_1 - 1 - \rho_1, N_2 - 1 - \rho_2, m, l)\}_{\eta(m,l)}}. \end{aligned} \quad (8.2-13)$$

die Entscheidung über Punktspiegelung und komplexe Konjugation treffen.

3. N_1 ist ungerade, N_2 ist gerade. Das Frequenzraster wird nun auf den Punkten

$$\left[-\frac{1}{2}(N_1 - 1); \frac{1}{2}(N_1 - 1) \right], \quad \left[-\frac{1}{2}N_2; \frac{1}{2}N_2 - 1 \right] \quad (8.2-14)$$

auf Symmetrie analog zu 2. untersucht. Ebenfalls analog ist der Ausschluß von Bildern mit

$$|\check{I}(0, \rho_2)| = 0 \quad \forall \rho_2 \in \mathbb{Z}, \quad (8.2-15)$$

wie durch die Abb. 21(c) und 21(d) deutlich wird. Für Gabortransformierte bei $l = \frac{L}{4}$ und $l = \frac{3L}{4}$ muß (8.2-13) herangezogen werden.

4. N_1 und N_2 sind gerade. Das zweidimensionale Frequenzintervall ist hier

$$\left[-\frac{1}{2}N_1; \frac{N_1}{2} - 1 \right], \quad \left[-\frac{1}{2}N_2; \frac{1}{2}N_2 - 1 \right]. \quad (8.2-16)$$

Das Symmetrieargument funktioniert genauso wie bei 2. Ausgeschlossen werden müssen Bilder mit

$$|\check{I}(\rho_1, \rho_1)| = 0 \quad \forall \rho_1 \in \mathbb{Z}. \quad (8.2-17)$$

Die graphische Darstellung des Problems enthalten 21(e) und 21(f). (8.2-13) gewährleistet des weiteren die Eindeutigkeit für $l = \frac{L}{2\pi} \arctan \frac{N_2}{N_1}$ und $l = \frac{L}{2\pi} \arctan \frac{N_2}{N_1} \pm \frac{L}{2}$.

Nun sind die Gabortransformierten bezüglich einer eventuellen Punktspiegelung mit komplexer Konjugation festgelegt worden.

- Bis jetzt konnten die Fourieramplituden der Teilbandsignale eindeutig rekonstruiert werden:

$$|\mathcal{I}_0^{rek}(\vec{\rho}, m, l)| = |\mathcal{I}_0(\vec{\rho}, m, l)|. \quad (8.2-18)$$

Es bleibt die Aufgabe, die Fourierphasen der Teilbandbilder zu ermitteln, die zu diesem Zweck hier auch separat betrachtet werden:

$$\begin{aligned} \left\{ \Phi_{\check{\mathcal{I}}_0}^{rek}(\vec{\rho}, m, l) \right\}_{\eta(m,l)} &= \eta(m, l) + \Phi_{\check{I}}(\vec{\rho}) + \frac{\pi}{2} \left(1 - \text{sgn} \left[\vec{\rho}^T \vec{e} \left(\frac{2\pi l}{L} \right) \right] \right) \\ \left\{ \Phi_{\check{\mathcal{I}}_0}'^{rek}(\vec{\rho}, m, l) \right\}_{\eta(m,l)} &= \eta(m, l) + \Phi_{\check{I}}(\vec{\rho}). \end{aligned} \quad (8.2-19)$$

Da hier nur reelle Bildsignale vorgesehen sind, gilt $\Phi_I(\vec{\rho}) = -\Phi_I(-\vec{\rho})$. Diese Information ermöglicht das Entfernen von $\eta(m, l)$:

$$\eta(m, l) = \frac{1}{2} \left[\left\{ \Phi_{\check{I}_0}^{rek}(\rho, m, l) \right\}_{\eta(m, l)} + \left\{ \Phi_{\check{I}_0}^{rel}(-\rho, m, l) \right\}_{\eta(m, l)} \right] \quad (8.2-20)$$

Die Korrektur erfolgt nach

$$\Phi_{\check{I}_0}^{rek}(\vec{\rho}, m, l) = \left\{ \Phi_{\check{I}_0}^{rek}(\vec{\rho}, m, l) \right\}_{\eta(m, l)} - \eta(m, l). \quad (8.2-21)$$

Allerdings können Phasenverschiebungen um $\pm\pi$ nicht detektiert werden, da sich die Phasenfunktion nur in einer Wertemenge von beispielsweise $[-\pi; \pi[$ bewegt, die als periodisch fortgesetzt betrachtet werden kann. Zudem haben die Phasen auf der Geraden $\vec{\rho}^T \vec{e} \left(\frac{2\pi l}{L} \right)$ keine Bedeutung und können zu der Mittelung (8.2-20) auch keinen Beitrag liefern. Daher gilt

$$\Phi_{\check{I}_0}^{rek}(\vec{\rho}, m, l) \in \left\{ \Phi_{\check{I}_0}(\vec{\rho}, m, l), \Phi_{\check{I}_0}(\vec{\rho}, m, l) \pm \pi \right\}. \quad (8.2-22)$$

Die Berechnung von $\eta(m, l)$ braucht nur in einem Teilband stattzufinden, so daß für die anderen die Phase $\Phi_{\check{I}_0}^{rek}(\vec{\rho}, m, l)$ direkt ermittelt werden kann. Daher sind entweder alle Teilbandfourierphasen richtig oder alle um $\pm\pi$ verschoben.

Die benutzten Eigenschaften reeller Bilder erlauben nun, Fourieramplituden und -phasen zu der diskreten Gaborwavellettransformierten im Frequenzraum zusammenzusetzen:

$$\check{I}_0^{rek}(\vec{\rho}, m, l) = \left| \check{I}_0^{rek}(\vec{\rho}, m, l) \right| e^{i\Phi_{\check{I}_0}^{rek}(\vec{\rho}, m, l)} = \begin{cases} \left| \check{I}_0(\vec{\rho}, m, l) \right| e^{i\Phi_{\check{I}_0}(\vec{\rho}, m, l)} = \check{I}_0(\vec{\rho}, m, l) & \forall m, l \\ \text{oder} \\ \left| \check{I}_0(\vec{\rho}, m, l) \right| e^{i[\Phi_{\check{I}_0}(\vec{\rho}, m, l) \pm \pi]} = -\check{I}_0(\vec{\rho}, m, l) & \forall m, l \end{cases} \quad (8.2-23)$$

Bilden die Gaborfunktionen einen Frame in $\tilde{\mathcal{H}}_I$, so kann nach Definition 2 $\check{I}_0^{rek}(\vec{\rho}, m, l)$ eindeutig das Bild $I^{rek}(\vec{n})$ zugeordnet werden, woraus direkt

$$I^{rek}(\vec{n}) = \pm I(\vec{n}) \quad (8.2-24)$$

folgt. □

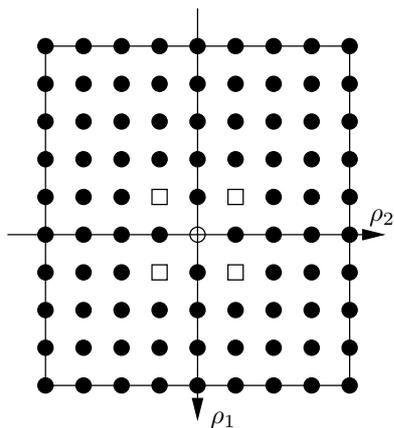
Sieht man von der Tatsache ab, daß die Bildsignale $I(\vec{n})$ als quantisiert angenommen werden müssen, haben die, für die das Theorem 8 bezüglich der Einschränkungen (8.2-3), (8.2-4) und (8.2-5) nicht gilt, das Maß Null. In der Praxis dürften digitale Bilder, die im Frequenzraum ganze „Nullstreifen“ haben, äußerst selten sein.

Die Eindeutigkeitsaussagen von Theorem 8 lassen sich allerdings nicht so einfach auf die Betragsausgabe der pyramidalen Gaborwavellettransformierten übertragen, da die Fourierphasen von Teilbändern verschiedener Skala nicht ohne weiteres auseinander berechnet werden können, was mit der Nichtvertauschbarkeit von Abstratenumsetzung und Betragsbildung zusammenhängt. Sollten demnach die Fourierphasen in nur jedem Orientierungsband konsistent sein, dann ergäßen sich 2^M Rekonstruktionsmöglichkeiten. Dazu wären allerdings noch einige Untersuchungen nötig.

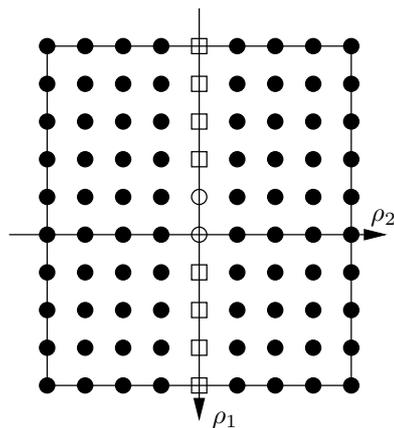
Die Polynomgestalt, die die bis auf das Vorzeichen eindeutige Rekonstruktion aus den Antworten komplexer Zellen erlaubt, findet sich ebenso bei Weyl-Heisenberg-Transformierten mit dem Kern (1.3-6) wieder. Da diese Gaborfunktionen als komplexwertig definiert sind, könnte Theorem 8 mit denselben Einschränkungen ebenso auf die Betragsantwort, auch *Spektrogramm* genannt, angewandt werden.

8.3 Ein Rekonstruktionsalgorithmus für die Antworten komplexer Zellen

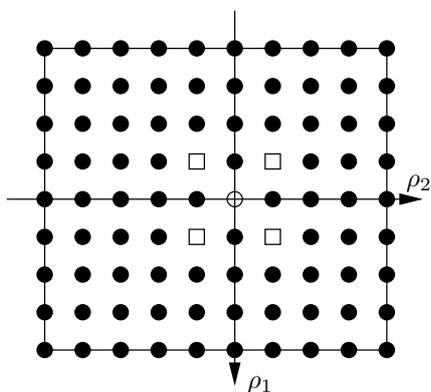
Der Beweis zu Theorem 8 gibt in groben Zügen schon einen Algorithmus vor, der die *LM* Ausgabebildsignale der Gaborwavelletfilterbank aus deren Beträgen wiedergewinnt. Bei der Entscheidung



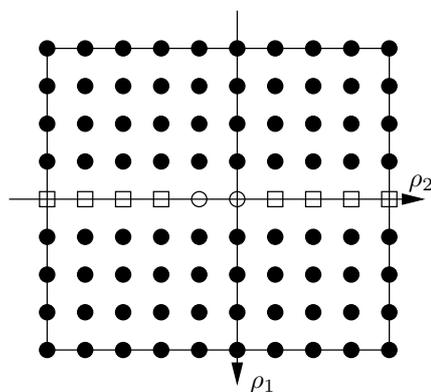
(a) Punkt $\vec{\rho} = \vec{0}$ eindeutig lokalisiert.



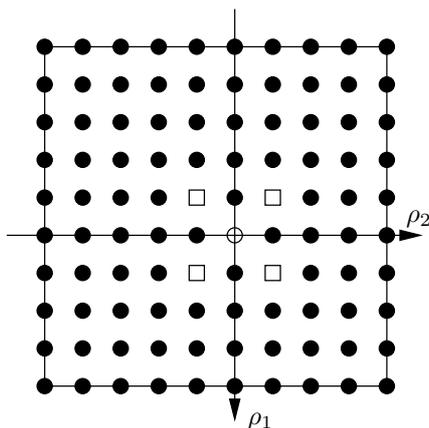
(b) Lokalisierung von $\vec{\rho} = \vec{0}$ nicht möglich.



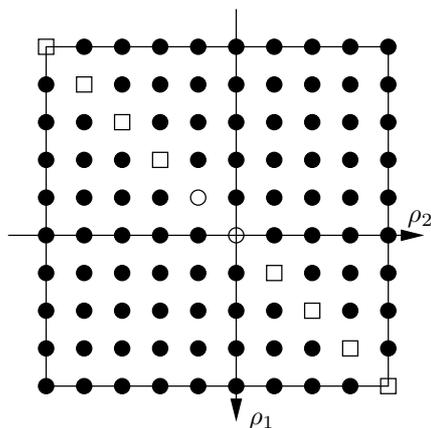
(c) Punkt $\vec{\rho} = \vec{0}$ eindeutig lokalisiert.



(d) Lokalisierung von $\vec{\rho} = \vec{0}$ nicht möglich.



(e) Punkt $\vec{\rho} = \vec{0}$ eindeutig lokalisiert.



(f) Lokalisierung von $\vec{\rho} = \vec{0}$ nicht möglich.

Abbildung 21: Erläuterungen zum Beweis von Theorem 8. „□“ steht für einen Punkt $\vec{\rho}$ mit $|\check{I}(\vec{\rho}, m, l)| = 0$, „•“ für einen Punkt $\vec{\rho}$ mit $|\check{I}(\vec{\rho}, m, l)| \neq 0$ und „○“ für den mutmaßlichen Punkt $\vec{\rho} = \vec{0}$. (a), (b): N_1 gerade, N_2 ungerade. (c), (d): N_1 ungerade, N_2 gerade. (e), (f): N_1 gerade, N_2 gerade.

über eine zusätzliche Punktspiegelung mit einer komplexen Konjugation kann die Symmetrie von „Nullmustern“ nicht mehr herangezogen werden, weil das Abfragen von als Gleitkommatyp formatierten Bilddaten auf null nicht praktikabel ist. Bis auf das Verbot der Nullstreifen im Frequenzraum sind die Aussagen von Theorem 8 bilddatenunabhängig. Diese Forderung muß für die Anwendung auf Bilder abgeschwächt werden, bei deren Filterbankausgaben im Frequenzraum die Signalleistung sich an derselben Stelle konzentriert, an der auch die Gaborfunktion lokalisiert ist. Da Skala und Orientierung bekannt sind, wird die Signalleistungskonzentration an einer bestimmten Stelle im Frequenzraum erwartet. Um zuverlässig zu entscheiden, ob das Frequenzbild um eine Punktspiegelung mit einer komplexen Konjugation korrigiert werden muß, brauchen die Amplituden auf beiden Seiten der Geraden $\vec{\rho}^T \vec{e}(\frac{2\pi l}{L}) = 0$ nur durch das Kriterium

$$\sum_{\vec{\rho} \in \{\vec{\rho} \mid \vec{\rho}^T \vec{e}(\frac{2\pi l}{L}) > 0\}} \left\{ \left| \tilde{\mathcal{I}}(\vec{\rho}, m, l) \right|^2 \right\}_{b(m,l)} > \sum_{\vec{\rho} \in \{\vec{\rho} \mid \vec{\rho}^T \vec{e}(\frac{2\pi l}{L}) < 0\}} \left\{ \left| \tilde{\mathcal{I}}(\vec{\rho}, m, l) \right|^2 \right\}_{b(m,l)} \quad (8.3-1)$$

verglichen zu werden. Auf beiden Seiten von (8.3-1) muß natürlich über die gleiche Anzahl von Frequenzpunkten summiert werden. Dieser Teil des Algorithmus dürfte ein sehr robustes Verhalten zeigen.

Die Probleme treten allerdings bei der eigentlichen Rekonstruktion des Frequenzbildes aus dem Betrag des gaborgefilterten Ortsraumsignals auf. Hierzu bietet es sich an, den Phase-Retrieval-Algorithmus aus (Fienup, 1987) unter Vertauschung von Orts- und Frequenzraum zu benutzen. In seiner klassischen Form aus Abb. 20 liefert er aus Fourieramplituden nach etwa 200 Iterationen brauchbare Ergebnisse, was mit der Glattheit natürlicher Bilder in Zusammenhang gebracht werden kann. Genauere Aussagen sind nicht möglich, da es dazu keine Konvergenztheorie gibt. In der vorgeschlagenen modifizierten Form kommt es zu Problemen, da die Frequenzbilder sehr fein strukturiert sind. Es ist zu beobachten, daß die Ergebnisse, die allesamt nach 200 und unverändert auch nach 2000 Iterationen immer noch unbrauchbar sind, stark vom initialisierten Eingabebild, für das bei einer Frequenzraum-Ortsraum-Rekonstruktion weißes Rauschen gewählt werden kann, abhängen. Der Algorithmus nach Abb. 20 erfordert im Ortsraum einen endlichen Träger, auf den die gewonnene Phasenfunktion angepaßt werden soll. Im hier betrachteten umgekehrten Fall müssen, um die Phase auf analoge Weise zu optimieren, Nullen bei hohen Frequenzen, die bei den meisten FFT-Algorithmen in der Mitte des zu verarbeitenden Feldes gespeichert sind, eingefügt werden. Das Gebiet, auf dem die Korrektur in jedem Iterationsschritt durchgeführt wird, hat hier also die Form eines Kreuzes. Eine Möglichkeit, das Verfahren zu stabilisieren, wäre, gleichzeitig einen Nullrahmen um das Bild im Ortsraum zu legen, um seine diskrete Fouriertransformierte zu glätten. Da Größenänderungen aufgrund der Vorgaben der FFT nur in Zweierpotenzen vorgenommen werden können, würde der Zeitaufwand weiter ansteigen, zumal mehrere hundert Iterationen bei ML Bildern gerechnet werden müssen. Zusammenfassend muß gesagt werden, daß eine numerisch stabile Rekonstruktion mit dem Phase-Retrieval-Algorithmus in dieser Form nicht möglich ist.

Sollte dies einmal doch funktionieren, bietet (8.3-1) die Möglichkeit, die falsche Punktspiegelung zu erkennen. Die additive konstante Phase kann danach, da nun der Punkt $\vec{\rho} = \vec{0}$ eindeutig gefunden ist, nach (8.2-20) berechnet werden. Ein Problem bei der Phasenberechnung könnten die sehr kleinen Amplituden im Ausgabebild sein, die unter Umständen durch Quantisierungsrauschen gestört sind, weshalb auch hier Vorsicht geboten ist. Ein numerisch stabiler und effizienter Rekonstruktionsalgorithmus, der mehrdeutige Ergebnisse verhindert, ist für Antworten komplexer Zellen auf der Basis von Theorem 8 nicht in Sicht. Gute Rekonstruktionsergebnisse sind in (von der Malsburg und Shams, 1998) dokumentiert. Der dort vorgestellte Algorithmus minimiert das Kriterium

$$S(\tilde{\mathcal{I}}) = \sum_{\vec{n}_0 \in \mathcal{B}} \sum_{m=m_0}^{M_0} \sum_{l=0}^{L-1} \mu_{m,l} \left[\left| \mathcal{I}(\vec{n}_0, m, l) \right|^2 - \left| \tilde{\mathcal{I}}(\vec{n}_0, m, l) \right|^2 \right]^2 \quad (8.3-2)$$

mit einem Gradientenabstieg, der aber initialisierungsabhängige Ausgaben liefert. Das in dieser Arbeit benutzte Maß ist $\mu_{m,l} = a_0^{-2m}$. Eventuell kann dieses Kriterium auch auf die Rekonstruktion der einzelnen ML Ausgabebilder angewandt werden und somit den modifizierten Phase-Retrieval ersetzen.

Teil IV
Ausblick und Anhang

A Formeln zu Fouriertransformation und Systemtheorie

A.1 FT

- Faltung im Ortsraum

$$I_1(\vec{x}) * I_2(\vec{x}) \iff 2\pi \hat{I}_1(\vec{\omega}) \hat{I}_2(\vec{\omega}) \quad (\text{A.1-1})$$

- Faltung im Frequenzraum

$$2\pi I_1(\vec{x}) I_2(\vec{x}) \iff \hat{I}_1(\vec{\omega}) * \hat{I}_2(\vec{\omega}) \quad (\text{A.1-2})$$

- Affine Transformation im Ortsraum

$$I(A\vec{x} + \vec{b}) \iff \frac{1}{|\det A|} \hat{I}(A^{-T}\vec{\omega}) e^{i(A^{-T}\vec{\omega})^T \vec{b}} \quad (\text{A.1-3})$$

- Affine Transformation im Frequenzraum

$$\frac{1}{|\det A|} I(A^{-T}\vec{x}) e^{-i\vec{b}^T A^{-T}\vec{x}} \iff \hat{I}(A\vec{\omega} + \vec{b}) \quad (\text{A.1-4})$$

- Plancherelsche Gleichung

$$\int_{\mathbb{R}^2} d^2x I_1(\vec{x}) I_2^*(\vec{x}) = \int_{\mathbb{R}^2} d^2\omega \hat{I}_1(\vec{\omega}) \hat{I}_2^*(\vec{\omega}) \quad (\text{A.1-5})$$

A.2 DSFT

- Plancherelsche Gleichung

$$\sum_{\vec{n} \in \mathbb{Z}^2} I_1(\vec{n}) I_2^*(\vec{n}) = \int_{\mathcal{N}^2} d^2\nu \check{I}_1(\vec{\nu}) \check{I}_2^*(\vec{\nu}) \quad (\text{A.2-1})$$

- Zusammenhang zur FT

$$\check{I}(\vec{\nu}) = \frac{1}{\Delta^2} \hat{I}\left(\frac{\vec{\nu}}{\Delta}\right) \quad (\text{A.2-2})$$

A.3 DFT

- Faltung in Ortsraum

$$I_1(\vec{n}) * I_2(\vec{n}) \iff \sqrt{N_1 N_2} \check{I}_1(\vec{\rho}) \check{I}_2(\vec{\rho}) \quad (\text{A.3-1})$$

- Faltung in Frequenzraum

$$\sqrt{N_1 N_2} I_1(\vec{n}) I_2(\vec{n}) \iff \check{I}_1(\vec{\rho}) * \check{I}_2(\vec{\rho}) \quad (\text{A.3-2})$$

- Plancherelsche Gleichung:

$$\sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} f_1(n_1, n_2) f_2^*(n_1, n_2) = \sum_{\rho_1=0}^{N_1-1} \sum_{\rho_2=0}^{N_2-1} \check{f}_1(\rho_1, \rho_2) \check{f}_2^*(\rho_1, \rho_2) \quad (\text{A.3-3})$$

- Zusammenhang zur FT

$$\check{I}(\rho_1, \rho_2) = \frac{2\pi}{\sqrt{N_1 N_2}} \frac{1}{\Delta^2} \hat{I}\left(\frac{2\pi\rho_1}{\Delta N_1}, \frac{2\pi\rho_2}{\Delta N_2}\right) \quad (\text{A.3-4})$$

- Zusammenhang zur DSFT

$$\check{I}(\rho_1, \rho_2) = \frac{2\pi}{\sqrt{N_1 N_2}} \check{I}\left(\frac{2\pi\rho_1}{N_1}, \frac{2\pi\rho_2}{N_2}\right) \quad (\text{A.3-5})$$

A.4 Das Abtasttheorem

$$I(\vec{x}) = \frac{1}{\Delta^2} \sum_{\vec{n} \in \mathbb{Z}^2} I(\vec{n}\Delta) \operatorname{si} \frac{\pi(x_1 - n_1\Delta)}{\Delta} \operatorname{si} \frac{\pi(x_2 - n_2\Delta)}{\Delta} \quad (\text{A.4-1})$$

B Die Waveletfunktionen $|\hat{\chi}_{\cdot, a, \vartheta}[\omega \vec{e}(\vartheta)]|$

Um die Eindeutigkeit der Gaboramplituden zu zeigen, benötigt man unter anderem den Zusammenhang

$$|\hat{\chi}_{\vec{x}_0, a, \vartheta}[\omega \vec{e}(\vartheta)]|^2 > |\hat{\chi}_{\vec{x}_0, a, \vartheta}[\omega \vec{e}(\vartheta + \pi)]|^2 \quad \forall a > 0, \omega > 0, \sigma > 0. \quad (\text{B.0-2})$$

Der Beweis erfolgt durch Nachrechnen:

$$\begin{aligned} |\hat{\chi}_{\vec{x}_0, a, \vartheta}[\omega \vec{e}(\vartheta)]|^2 &> |\hat{\chi}_{\vec{x}_0, a, \vartheta}[\omega \vec{e}(\vartheta + \pi)]|^2 &&\Leftrightarrow \\ a^2 \hat{\chi}(a\omega \vec{e}_1)^2 &> a^2 \hat{\chi}(-a\omega \vec{e}_1)^2 &&(a \neq 0) \Leftrightarrow \\ e^{-\sigma^2(a^2\omega^2 - 2a\omega + 1)} + e^{-\sigma^2(a^2\omega^2 + 1)} &&& \\ -2e^{-\sigma^2(a^2\omega^2 - a\omega + 1)} &> e^{-\sigma^2(a^2\omega^2 + 2a\omega + 1)} + e^{-\sigma^2(a^2\omega^2 + 1)} && \\ &&&-2e^{-\sigma^2(a^2\omega^2 + a\omega + 1)} \Leftrightarrow \\ e^{2\sigma^2 a\omega} - 2e^{\sigma^2 a\omega} &> e^{-2\sigma^2 a\omega} - 2e^{-\sigma^2 a\omega} &&\Leftrightarrow \\ 2(e^{\sigma^2 a\omega} - e^{-\sigma^2 a\omega}) &< e^{2\sigma^2 a\omega} - e^{-2\sigma^2 a\omega} &&\Leftrightarrow \\ 2 \sinh(\sigma^2 a\omega) &< \sinh(2\sigma^2 a\omega) && \\ &= 2 \sinh(\sigma^2 a\omega) \cosh(\sigma^2 a\omega) &&\Leftrightarrow \\ \cosh(\sigma^2 a\omega) &> 1 \quad \forall \sigma^2, a, \omega > 0 && \quad (\text{B.0-3}) \end{aligned}$$

Ausblick

Die eben erwähnten numerischen Probleme sind auf die Instabilität des Phase-Retrieval-Algorithmus nach (Fienup, 1987) von Ortsraumphasen zurückzuführen. Dieser könnte durch den in (von der Malsburg und Shams, 1998) vorgeschlagenen Gradientenabstieg in jedem einzelnen Teilband ersetzt werden. Obwohl dann auch noch weitere Probleme zu lösen sind, würde dies ein weiterer Schritt in Richtung einer eindeutigen Rekonstruktion ganzer Bilder aus den Antworten komplexer Zellen bedeuten.

Die zweite Möglichkeit bietet das Konzept der Patches, mit deren Hilfe Eindeutigkeitsaussagen über die lokale Bildkodierung durch Jets gewonnen werden können. Die Rekonstruktion einer ganzen Frequenzebene, die allerdings numerisch sehr aufwendig ist, ist vorgeschlagen worden, müsste aber erst getestet werden, um die Leistungsfähigkeit beurteilen zu können. Eine eindeutige Rekonstruktion kann nur durch den Abgleich mit anderen erfolgen. Zur Lösung dieses Optimierungsproblems werden Modelle aus der statistischen Mechanik vorgeschlagen, die aber noch formuliert werden müssen. Neben zahlreichen theoretischen Überlegungen müssen robuste Vergleichskriterien und effektive Algorithmen für die aufwendigen Berechnungen gefunden werden.

Es besteht im übrigen die Hoffnung, einmal eine auf Gaborwavelets zugeschnittene Frameschrankenabschätzung zu finden. Verfahren für allgemeine Wavelet- und Weyl-Heisenberg-Frames sind in (Daubechies, 1990) nachzulesen. Für die letzteren können die Frameschranken durch Berechnen von Infimum bzw. Supremum der Zaktransformierten bestimmt werden. Darüberhinaus werden in (Zeevi et al., 1998) Weyl-Heisenberg-Frames mit variable Fensterbreite und in (Ron und Shen, 1997) in mehrdimensionalen Funktionensystemen betrachtet. Möglicherweise liefert die stückweise Zaktransformation aus (Zeevi et al., 1998) eine Abschätzung der Frameschranken von Familien von Gaborwavelets. Hierbei sind allerdings sehr große theoretische Anstrengungen erforderlich.

Des Weiteren könnte man weitere Untersuchungen zur ökonomischen Darstellung (Einführung von Quantisierungsrauschen) der diskreten Gaborwavelettransformierten durchführen, da dieser aufgrund ihrer Übervollständigkeit sehr fehlertolerant ist, was mit den guten statistischen Eigenschaften der Kleinste-Quadrate-Schätzer zusammenhängt. Experimentelle Ergebnisse mit einer gröberen Rasterung, d.h. einer Darstellung der Graustufen mit weniger Bit als ursprünglich vorgesehen, sind in (Lee, 1996) erzielt worden, allerdings nur mit Tight-Frame-Rekonstruktionen.

Literatur

- Benedetto, J. J., Heil, C. und Walnut, D. F. (1998). Gabor systems and the Balian-Low Theorem. In Feichtinger, H. G. und Strohmer, T. (Hrsg.), *Gabor Analysis and Algorithms – Theory and Applications*, Applied and Numerical Harmonic Analysis, Seite 85 – 122. Birkhäuser.
- Böhme, J. F. (1993). *Stochastische Signale*. Teubner Studienbücher: Elektrotechnik. B. G. Teubner Stuttgart.
- Burt, P. J. und Adelson, E. H. (1983). The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications*, COM-31(4):532 – 540.
- Daubechies, I. (1990). The Wavelet Transform, Time-Frequency Localization and Signal Analysis. *IEEE Transactions on Information Theory*, 36(5):961 – 1005.
- Daugman, J. G. (1993). Quadrature-phase simple-cell pairs are appropriately described in complex analytic form. *Journal of the Optical Society of America*, 10(2):375 – 377.
- De Valois, R. L. und De Valois, K. K. (1988). *Spatial Vision*, Band 14 von *Oxford Psychology Series*. Oxford University Press.
- Dudel, J., Menzel, R. und Schmidt, R. F. (Hrsg.) (1996). *Neurowissenschaft – Vom Molekül zur Kognition*. Springer.
- Duffin, R. J. und Schaeffer, A. C. (1952). A class of nonharmonic Fourier series. *Transactions of the American Mathematical Society*, 72:341 – 366.
- Feichtinger, H. G. und Strohmer, T. (Hrsg.) (1998). *Gabor Analysis and Algorithms – Theory and Applications*. Applied and Numerical Harmonic Analysis. Birkhäuser.
- Fienup, J. R. (1987). Reconstruction of a complex-valued object from the modulus of its Fourier transform using a support constraint. *Journal of the Optical Society of America*, 4(1):118 – 123.
- Fliege, N. (1993). *Multiraten-Signalverarbeitung*. Informationstechnik. B. G. Teubner Stuttgart, erste Ausgabe.
- Freeman, W. T. und Adelson, E. H. (1991). The Design and Use of Steerable Filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(9):891 – 906.
- Gabor, D. (1946). Theory of Communication. *Journal IEE*, 93(III):429 – 457.
- Gardenier, P. H., McCallum, B. C. und Bates, R. H. T. (1986). Fourier Transform Magnitudes Are Unique Pattern Recognition Templates. *Biological Cybernetics*, 54:385 – 391.
- Golub, G. H. und van Loan, C. F. (1996). *Matrix Computations*. The Johns Hopkins University Press, dritte Ausgabe.
- Grossmann, A. und Morlet, J. (1984). Decomposition of Hardy functions into square integrable wavelets of constant shape. *SIAM Journal of Mathematical Analysis*, 15(4):723 – 736.
- Hayes, M. H. (1982). The Reconstruction of a Multidimensional Sequence from the Phase or Magnitude of Its Fourier Transform. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 30(2):140 – 154.
- Hayes, M. H. und McClellan, J. H. (1982). Reducible Polynomials in More Than One Variable. *Proceedings of the IEEE*, 70(2):197 – 198.
- Hirsch, M. C. (1998). *interBRAIN – Topographical Anatomy of the Human CNS (CD ROM)*. Springer.

- Hubel, D. H. und Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160:106 – 154.
- Jones, J. P. und Palmer, L. A. (1987). An Evaluation of the Two-Dimensional Gabor Filter Model of Simple Receptive Fields in Cat Striate Cortex. *Journal of Neurophysiology*, 58(6):1233 – 1258.
- Kaiser, G. (1994). *A Friendly Guide to Wavelets*. Birkhäuser.
- Lades, M., Vorbrüggen, J. C., Buhmann, J., Lange, J., von der Malsburg, C., Würtz, R. P. und Konen, W. (1993). Distortion Invariant Object Recognition in the Dynamic Link Architecture. *IEEE Transactions on Computers*, 42(3):300 – 311.
- Lee, T. S. (1996). Image Representation Using 2D Gabor Wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):959 – 971.
- Louis, A. K. (1989). *Inverse und schlecht gestellte Probleme*. Teubner Studienbücher: Mathematik. B. G. Teubner Stuttgart.
- Louis, A. K., Maaß, P. und Rieder, A. (1994). *Wavelets*. Teubner Studienbücher: Mathematik. B. G. Teubner Stuttgart.
- Mallat, S. (1989). A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7):674 – 693.
- Murenzi, R. (1989). Wavelet Transforms Associated to the n-Dimensional Euclidean Group with Dilations: Signal in More Than One Dimension. In Combes, J. M., Grossmann, A. und Tchamitchian, P. (Hrsg.), *Wavelets – Time-Frequency Methods and Phase Space*, Seite 239 – 246. Springer.
- Pei, S.-C. und Yeh, M.-H. (1997). An Introduction to Discrete Finite Frames. *IEEE Signal Processing Magazine*, 14(6):84 – 96.
- Penrose, R. (1955). A Generalized Inverse for Matrices. *Proceedings of the Cambridge Philosophical Society*, 51:406 – 413.
- Penrose, R. (1956). On Best Approximate Solutions of Linear Matrix Equations. *Proceedings of the Cambridge Philosophical Society*, 52:17 – 19.
- Pollen, D. A. und Ronner, S. F. (1981). Phase Relationships Between Simple Cells in the Visual Cortex. *Science*, 212(19):1409 – 1411.
- Pöttsch, M. (1994). Die Behandlung der Wavelet-Transformation von Bildern in der Nähe von Objektkanten. Internal Report 94-04, Institut für Neuroinformatik, Ruhr-Universität Bochum.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T. und Flannery, B. P. (1992). *Numerical Recipes in C – The Art of Scientific Computing*. Cambridge University Press, zweite Ausgabe.
- Ron, A. und Shen, Z. (1997). Weyl-Heisenberg frames and Riesz bases in $L_2(\mathbb{R}^d)$. *Duke Mathematical Journal*, 89(2):237 – 282.
- Schwarz, H. R. (1993). *Numerische Mathematik*. B. G. Teubner Stuttgart, dritte Ausgabe.
- Simoncelli, E. P., Freeman, W. T., Adelson, E. H. und Heeger, D. J. (1992). Shiftable Multiscale Transforms. *IEEE Transactions on Information Theory*, 38(2):587 – 607.
- Spitzer, H. und Hochstein, S. (1985). A Complex-Cell Receptive-Field Model. *Journal of Neurophysiology*, 53(5):1266 – 1286.
- Stork, D. G. und Wilson, H. R. (1990). Do Gabor functions provide appropriate descriptions of visual cortical receptive fields? *Journal of the Optical Society of America*, 7(8):1362 – 1373.

- van der Schaaf, A. und van Hateren, J. H. (1996). Modelling the Power Spectra of Natural Images: Statistics and Information. *Vision Research*, 36(17):2759 – 2770.
- von der Malsburg, C. und Shams, L. (1998). Making sense of complex cells.
- Würtz, R. P. (1995). *Multilayer dynamic link networks for establishing image point correspondences and visual object recognition*, Band 41 von *Reihe Physik*. Verlag Harri Deutsch.
- Zeevi, Y. Y., Zibulski, M. und Porat, M. (1998). Multi-window Gabor schemes in signal and image representations. In Feichtinger, H. G. und Strohmer, T. (Hrsg.), *Gabor Analysis and Algorithms – Theory and Applications*, Applied and Numerical Harmonic Analysis, Seite 381 – 408. Birkhäuser.