

Lehrstuhl für Nachrichtentechnik

**Nichtlineare Informationsverarbeitung im visuellen System:
Analyse, Synthese und Identifikation**

Ulrich Nuding

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktor-Ingenieurs

genehmigten Dissertation.

Vorsitzender:

Univ.-Prof. Dr. rer. nat. Bernhard Wolf

Prüfer der Dissertation:

1. Univ.-Prof. Dr.-Ing. Gert Hauske, i.R.
2. Univ.-Prof. Dr. techn. Josef A. Nossek

Die Dissertation wurde am 14.06.2006 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 15.01.2007 angenommen.

Vorwort

Diese Arbeit entstand während meiner dreijährigen Mitgliedschaft im Graduiertenkolleg der Deutschen Forschungsgemeinschaft *Sensorische Interaktion in biologischen und technischen Systemen* (DFG-GRK 267). Dieses stark transdisziplinär ausgerichtete Kolleg ermöglichte eine Arbeit, die in den Kontext der biologischen Signalverarbeitung eingebettet ist. Dank gilt daher Frau Dr. Isolde von Bülow, der wissenschaftlichen Koordinatorin des Kollegs, die den Kollegiaten stets ein offenes Ohr und eine helfende Hand schenkte.

Besonderer Dank gilt meinem Doktorvater, Prof. Dr.-Ing. Gert Hauske, der mir, nach meinem Studium der Nachrichtentechnik an der TU München, den Zugang zur biologisch motivierten Signalverarbeitung im visuellen System und der Kybernetik im Allgemeinen erschloß. Einen besonderen Dank spreche ich auch meinem Betreuer Dr. Christoph Zetsche (Universität Bremen) aus, der mich über all die Jahre hervorragend anleitete. Durch die vielen fruchtbaren Diskussionen trug er einen wesentlichen Teil zum Gelingen dieser Arbeit bei. Nicht zuletzt war es seiner Gastfreundschaft, und der seiner Frau, Prof. Dr. Kerstin Schill (Universität Bremen), zu verdanken, daß auch trotz einer großen räumlichen Trennung über einen substantiellen Zeitraum der Kontakt nie abriß, und stets ein harmonisches und kollegiales Arbeitsverhältnis bestehen blieb. Der größte Teil der Arbeit wurde am Institut für medizinische Psychologie der LMU München angefertigt, dessen Ordinarius, Prof. Dr. Ernst Pöppel, ich herzlich für die angenehme und abwechslungsreiche Arbeitsatmosphäre an seinem sehr vielseitig orientierten Institut danke. Abschließender Dank gebührt meiner Familie und meiner langjährigen Freundin Anna Gurtner, die mich stets aufs Neue unterstützt und motiviert haben.

Zusammenfassung

Zellen in höheren visuellen Hirnregionen von Primaten (visuelle Kortizes V1–V4) weisen einige nichtlineare Merkmale, wie beispielsweise Phaseninvarianz, Endinhibition und Selektivität für komplexe Formen (Ecken, Krümmungen, etc.) auf. Um zu einem Verständnis für diese Verarbeitungseigenschaften zu gelangen, bedarf es der *Analyse* der Informationsverarbeitung im visuellen System, der *Modellierung* nichtlinearer kortikaler Phänomene sowie der *Synthese* biologisch motivierter technischer Systeme zur Kodierung natürlicher Bilder.

Hierzu werden unterschiedliche Ansätze evaluiert. Unter Verwendung von polynomialen Ansätzen wird die Statistik natürlicher Bilder mittels *Spektren höherer Ordnung* analysiert, nichtlineare neuronale Systeme mit Hilfe der *Volterra-Wiener Systemtheorie* modelliert und die Anpassung dieser Systeme an die statistischen Eigenschaften der Umwelt untersucht. Zentrale Frage ist, wie nichtlineare neuronale Systeme zu einer effizienten Repräsentation natürlicher Bilder beitragen und wie diese unter biologisch realistischen Bedingungen *identifiziert* werden können.

Ein komplementärer Ansatz ist das Lernen in neuronalen Netzen unter Verwendung von Optimalitätskriterien. Es wird eine mehrstufige neuronale Architektur implementiert, deren Gewichte aus einer Optimierung der Statistik natürlicher Bilder hervorgehen. Die Verarbeitungseigenschaften dieses Systems werden ausführlich untersucht und in Relation mit der biologischen Informationsverarbeitung gesetzt. Durch die Implementierung eines Codecs zur Kodierung natürlicher Bilder wird die technische Relevanz dieses Ansatzes demonstriert. Die Einführung einer vereinheitlichenden Darstellung in Zustandsraumkoordinaten erlaubt einen qualitativen Vergleich dieser unterschiedlichen Ansätze und eine weitere Interpretation der nichtlinearen neuronalen Informationsverarbeitung.

Inhaltsverzeichnis

1	Einleitung	1
1.1	Motivation	1
1.2	Ziel der Arbeit	2
1.3	Überblick	3
1.4	Anatomische und neurophysiologische Grundlagen des visuellen Systems	4
2	Informationsverarbeitung im visuellen System	11
2.1	Einführung	11
2.2	Das Konzept der intrinsischen Dimensionalität	13
2.3	Die Shannon'sche Informationstheorie	14
2.4	Statistik natürlicher Bilder	16
2.4.1	Signalraumdarstellung kontinuierlicher Signale	16
2.4.2	Natürliche Bilder im Signalraum	17
2.4.3	Statistiken zweiter Ordnung natürlicher Bilder	18
2.5	Optimale Transformation natürlicher sensorischer Stimuli	21
2.5.1	Quellenkodierung im visuellen System	21
2.5.2	Optimale lineare Transformationsverfahren	23
2.5.3	Spärliche Kodierung	30
2.6	Der klassische Ansatz zur Erklärung der Funktionsweise des visuellen Systems	32
2.6.1	Spatiotemporale Tiefpaß-Filterung in der Retina	34
2.6.2	Laterale Inhibition der Ganglienzellen	34
2.6.3	Selbstähnliche und lokalisierte rezeptive Felder	36
2.6.4	Das Mehrkanalmodell	39
2.7	Grenzen des Standardansatzes	40
3	Nichtlineare Ansätze zur Modellierung neuronaler Informationsverarbeitung	45
3.1	Explizite, implizite und polynomiale Ansätze	45
3.1.1	Expliziter Ansatz: Zustandsraumdarstellung	45
3.1.2	Impliziter Ansatz: Neuronale Netze	49
3.1.3	Polynomialer Ansatz: Polyspektren und Volterra-Wiener-Systeme	50
3.2	Extra-klassische Effekte im Zustandsraum	52

3.3	Nichtlineare Partitionierung des Zustandsraums	54
4	Synthese eines nichtlinearen LNL-Systems zur Bildkodierung	57
4.1	Statistische Abhängigkeiten zwischen linearen Wavelet-Koeffizienten	58
4.2	Nichtlineare Transformation der statistischen Abhängigkeiten	59
4.3	Eine invertierbare nichtlineare Architektur	62
4.3.1	Prinzipieller Aufbau	63
4.3.2	Statistische Optimierung der Ausgangsschicht	66
4.3.3	Quantisierung	68
4.4	Rekonstruktion und Fehler	69
4.4.1	Rekonstruktion	69
4.4.2	Vergleich mit linearem Referenz-Codec	71
4.4.3	Klassifizierung lokaler Merkmale und assoziierte Rekonstruktionsfehler	72
4.5	Nichtlineare Frequenzinteraktionen	76
4.6	Biologische Relevanz	77
5	Polynomiale Ansätze	82
5.1	Spektren höherer Ordnung	83
5.1.1	Momente und Kumulanten	83
5.1.2	Kumulantenspektren	89
5.1.3	Bispektrale Analyse natürlicher Bilder	100
5.2	Volterra-Wiener-Systeme	101
5.2.1	Kurzer geschichtlicher Überblick	102
5.2.2	Die Volterra-Reihenentwicklung	104
5.2.3	Synthese und Approximation nichtlinearer neuronaler Systeme	116
5.2.4	Modellierung nichtlinearer kortikaler Phänomene	126
5.2.5	Identifikation von Volterra-Kernen zweiter Ordnung	129
5.2.6	Dekorrelation höherer Ordnung	138
5.3	Grenzen des quadratischen Ansatzes	149
6	Diskussion	153
7	Literaturverzeichnis	155

1 Einleitung

Der Beginn ist der wichtigste Teil der Arbeit

Platon

1.1 Motivation

Das visuelle System des Menschen stellt ein Sinnesorgan dar, das innerhalb kürzester Zeit immense Datenmengen korrekt und effizient verarbeiten kann. Wesentliche ethologisch relevante Prozesse werden schnell und ohne ersichtliche Mühe, sozusagen im Hintergrund, bearbeitet, und optimale Lösungen für den Organismus bereitgestellt. Ob dies nun die urzeitliche Flucht vor überlegenen Gegnern, das Auffinden von Nahrung oder das Erkennen von Mitmenschen ist, es werden ständig Aufgaben bewältigt, mit welchen heutige technische Bildverarbeitungs- und Objekterkennungssysteme im Allgemeinen überfordert sind. Es ist daher nicht verwunderlich, daß das visuelle System schon zur Zeit der Antike im Zentrum des Interesses stand. Hier glaubte man, daß Licht vom Auge ausgesendet wird und sich mit den von den Objekten ausgesandten Repräsentationen vermischt. Die Epikuräer meinten beispielsweise, daß Objekte Abbilder aussenden, die in das Auge gelangen und dort zur Wahrnehmung des Objektes führen (Boring, 1942). Die traditionelle Ansicht zur Physiologie des Auges stammt vom Araber ALHAZEN (ca. 1039), wonach die Linse das perzeptive Organ darstellt (Lindberg, 1976). KEPLER und LEONARDO DA VINCI betonten die Ähnlichkeit des Auges mit einer *camera obscura*; KEPLER folgerte, daß die Retina das perzeptive Organ sei, da durch die Linse auf der Retina ein (spiegelverkehrtes) Abbild zu sehen ist (Kepler, 1604; Fishman, 1973). 1681 beschrieb Mariotte den blinden Fleck an der Stelle, an der der optische Nerv aus der Retina austritt (Mariotte, 1681). Daraus folgte logisch, daß die Retina das perzeptive Organ sein muß. Im weiteren Verlauf trugen SIR ISAAC NEWTON, THOMAS YOUNG, JOHANN WOLFGANG GOETHE, JOHANNES EVANGELISTA PURKINJE und HERMANN VON HELMHOLTZ zur weiteren Klärung der Funktionsweise des visuellen Systems bei (Sabra, 1981; Young, 1801; Goethe, 1953; Purkinje, 1825; Helmholtz, 1850).

Wesentliche Arbeiten zum visuellen System entstammen den Gebieten der *Neurophysiologie*, *Psychophysik* und der *Nachrichtentechnik*. Vom technischen Standpunkt aus

offeriert das Sehsystem beim Menschen zwei grundlegende Leistungen. Zum Einen ermöglicht es die Identifikation spezifischer Objekte, zum Anderen ist es invariant gegenüber nicht-relevanten Transformationen, wie zum Beispiel Rotation, Veränderung der Beleuchtung, Skalierung oder partielle Verdeckung und ermöglicht so eine Generalisierung oder auch Klassifikation. Es ist allgemein akzeptiert, daß sich wesentliche Aspekte der Leistung dieses Systems auf einer evolutionären Zeitskala durch Anpassung an die strukturellen Eigenschaften unserer Umwelt entwickelt haben. Diese strukturellen Eigenschaften basieren auf der Tatsache, daß unsere Umgebung aus kompakten Objekten besteht, die homogen sind und sich gegenseitig verdecken können. MACH postulierte die Ökonomie des Denkens, Objekte werden mit geringster Anstrengung erkannt, entsprechend ihrer sparsamsten und wahrscheinlichsten Interpretation (Mach, 1922). Er erklärt, daß sensorische Mechanismen als Ergebnis einer Darwinschen Anpassung an die Umgebung verstanden werden können. ATTNEAVE beschreibt die Hauptfunktion der perzeptuellen Systeme als eine Redundanzreduktion, um eine effizientere oder *ökonomischere* Darstellung der Information zu erreichen: „*encode incoming information in a form more economical than that which impinges on the receptors*“ (Attneave, 1954).

Ein optimales System muß in der Lage sein, diese der Umwelt inhärenten strukturellen Eigenschaften zu nutzen, respektive sie zu transformieren, um effizient arbeiten zu können. Diese Art der Problemstellung ist eng verwandt mit der Übertragung von Information über bandbegrenzte Kanäle, ein Problem, dem sich SHANNON gewidmet, und so die *Informationstheorie* eingeführt hat (Shannon, 1948). Unglücklicherweise ermöglicht die Informationstheorie zwar die Berechnung wesentlicher Größen der Information, wie beispielsweise Informationsgehalt, Transinformation oder Kanalkapazität, sagt jedoch nichts darüber aus *wie* eine optimale Repräsentation erreicht werden kann. Im Falle von speziellen, vereinfachenden Annahmen existieren wohlbekannte Lösungen, vgl. Abschnitt 2.6; im allgemeinen, komplexeren Fall müssen allerdings explizite Lösungen für dieses Problem gefunden werden.

1.2 Ziel der Arbeit

Die im letzten Abschnitt erwähnten vereinfachenden Annahmen bei der Analyse der Informationsverarbeitung im visuellen System sind auf Seite der Signalverarbeitung die linearen, zeit- bzw. ortsinvarianten Systeme und auf Seite der Analyse die Betrachtung von Statistiken zweiter Ordnung (Autokorrelation bzw. Leistungsdichtespektrum). Durch die Analyse der statistischen Eigenschaften natürlicher Bilder konnte gezeigt werden, daß die Beschränkung auf Statistiken zweiter Ordnung bei der Informationsverarbeitung im visuellen System unzulänglich ist (Zetzsche und Barth, 1990a; Wegmann und Zetzsche, 1990a; Zetzsche und Krieger, 2001b).

Wie ich im Detail in Abschnitt 2.7 herausarbeiten werde, sind wesentliche Merkmale

des visuellen Systems durch diesen Standardansatz nicht erklärbar. Diese in experimentellen Untersuchungen gefundenen Aspekte sind jedoch nach heutigem Kenntnisstand essentiell für die Funktionsweise des Systems und eine Näherung durch lineare Systeme oder Statistiken zweiter Ordnung somit unzulässig. Es müssen daher generische Ansätze verwendet werden, um den neuronalen *Kode* zu entschlüsseln und um ein Verständnis dafür zu entwickeln, wie sensorische Prozesse eine effiziente Repräsentation der Außenwelt erreichen. Dabei können drei komplementäre Ansätze identifiziert werden, um die unterschiedlichen Aspekte der neuronalen Informationsverarbeitung zu beschreiben:

1. Direkte Beschreibung von Wahrscheinlichkeitsdichten und deren Transformationen in repräsentationalen Zustandsraumkoordinaten,
2. indirekte Methoden im Sinne von neuronalen Netzen und unüberwachtem Lernen sind mächtig hinsichtlich den Ergebnissen einer globalen Optimierung, liefern jedoch keine direkte Erkenntnis über statistische Eigenschaften oder einfache funktionale Modellstrukturen,
3. funktionale Approximation mit Polynomsystemen und Potenzreihenentwicklung. Auf Seite der Analyse der Statistik natürlicher Bilder resultiert dies in Kumulantenfunktionen und Spektren höherer Ordnung (Polyspektren), auf Seite der Synthese von Modellen in der Volterra-Wiener Systemtheorie.

Mit Hilfe der obigen Ansätze sollen in dieser Arbeit die multivariate Wahrscheinlichkeitsdichtefunktion natürlicher Bilder, die nichtlinearen Eigenschaften visueller Neurone sowie deren Zusammenhang analysiert werden. Ziel hierbei ist einerseits die Herleitung nichtlinearer Transformationen, basierend auf grundlegenden Optimierungsannahmen, andererseits die Modellierung kortikaler Phänomene (wie beispielsweise *Complex-Zellen*, *Verstärkungsregelung*, *Endinhibition*, *extra-klassische rezeptive Felder*, *nichtlineare Selektivität*) und die Analyse der statistischen Abhängigkeiten in natürlichen Bildern.

1.3 Überblick

Im Detail wird hierzu in den Kapiteln 1 und 2 ein Überblick über das visuelle System des Menschen gegeben, kurz die anatomischen und physiologischen Grundlagen dargestellt, und schließlich wesentliche Aspekte der Informationsverarbeitung in biologischen Sehsystemen eruiert. In Kapitel 3 wird die Zustandsraumdarstellung eingeführt und mit deren Hilfe die oben aufgeführten Ansätze in einem einheitlichen Rahmen dargestellt. Somit können die Vor- und Nachteile der verschiedenen Ansätze verglichen und die jeweiligen Applikationsdomänen definiert werden. Ferner zeige ich, daß verschiedene Aspekte der nichtlinearen neuronalen Informationsverarbeitung in dieser Darstellung als direktes Resultat der nichtlinearen Antwortflächen im Zustandsraum verstanden werden können. Dabei sind zwei grundlegende Effekte klassifizierbar: (*i*) Generalisierung, dar-

stellbar als eine ODER-Verknüpfung der Zustandsgrößen und (ii) Selektivität, darstellbar als eine ebensolche UND-Verknüpfung. Es folgt dann in Kapitel 4 die Applikation des indirekten Ansatzes der neuronalen Netze zur Optimierung der Statistik natürlicher Bilder unter Verwendung eines biologisch plausiblen, mehrschichtigen nichtlinearen Filtersystems. Hier führe ich aus, inwiefern dieses Mehrschicht-System den klassischen Ansätzen überlegen ist, indem ich die nichtlinearen Verarbeitungseigenschaften des Systems herausarbeite, und die technische Bedeutung dieses Ansatzes verdeutliche, welche in der effizienten Kodierung natürlicher Bilder liegt. Hierfür wird ein Vergleich mit einem linearen Pendant durchgeführt und die Bedeutung der Rekonstruktionsfehler im Sinne unterschiedlicher Fehlerklassen analysiert. Kapitel 5 beschäftigt sich mit polynomialen Ansätzen und stellt eine Erweiterung der bisherigen Ergebnisse zur Anwendung von Volterra-Wiener Systemen und Polyspektren auf die kortikale Informationsverarbeitung dar. Wesentliche Merkmale dieses Ansatzes sind die polyspektrale Analyse natürlicher Bilder, die Synthese nichtlinearer Effekte in kortikalen Neuronen sowie die Identifikation nichtlinearer Systeme zweiter Ordnung. Geschlossen wird in Kapitel 6 mit einer Diskussion zu dieser Arbeit.

1.4 Anatomische und neurophysiologische Grundlagen des visuellen Systems

Das Ziel dieses Abschnitts ist es, einen kleinen Einblick in die biologischen Grundlagen des peripheren visuellen Systems zu gewähren. Dafür werden die folgenden Instanzen der Verarbeitung im visuellen System vorgestellt:

- Optische Abbildung durch die Linse
- Retinale Verarbeitung, laterale Inhibition
- Corpus Geniculatum Laterale (CGL)
- Primärer visueller Kortex (V1)

Das Auge. Die erste Verarbeitungsinstanz ist durch die optischen Abbildungseigenschaften von Hornhaut, Linse und Augeninnenflüssigkeit gegeben (schematische Zeichnung in Abb. 1.1). Das einfallende Licht wird durch die in der Retina befindlichen Rezeptoren in zelluläre elektrische Potentiale umgesetzt. Es existieren zwei Arten dieser Photorezeptoren: Die *Zapfen*, welche für das farbempfindliche, *photopische* Tagsehen zuständig sind, und die *Stäbchen*, die eine höhere Lichtempfindlichkeit besitzen und das Sehen im Dämmerlicht, das sogenannte *skotopische* Sehen, ermöglichen. Diese zwei Arten von Rezeptoren sind nicht gleichmäßig auf der Retina angeordnet, sondern die Zapfendichte nimmt außerhalb des Bereichs des scharfen Sehens, der *Fovea*, stark ab, wohingegen die Stäbchendichte in der Peripherie bis zu einem gewissen Winkel zunimmt.

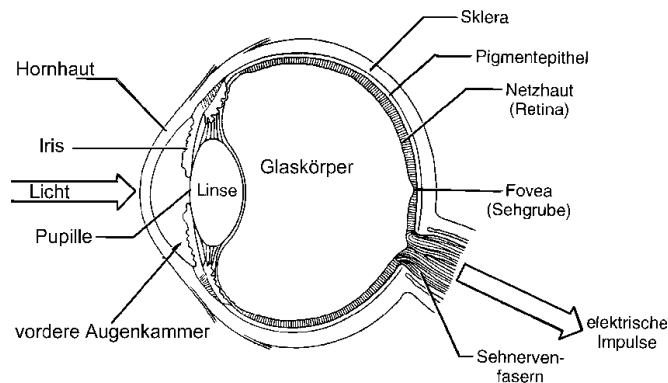


Abbildung 1.1: Schematischer Schnitt durch das menschliche Auge.

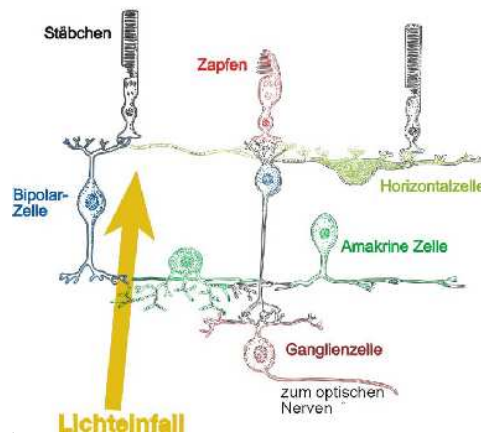


Abbildung 1.2: Schematisierter Schnitt durch die Vertebratenretina.

Die absolute Anzahl der in der Retina befindlichen Rezeptoren beträgt ca. $5 \cdot 10^6$ Zapfen und ca. 10^8 Stäbchen. Die Zapfen gliedern sich weiter in drei Subkategorien, den L-, M- und S-Typus. Diese unterscheiden sich in ihrer spektralen Empfindlichkeit und ermöglichen somit das Farbempfinden.

Die Ausgänge der Photorezeptoren sind über ein Netzwerk mit weiteren Zellen, den *Bipolar*-, *Amakrin*-, *Horizontal*- und *Ganglienzellen*, verbunden. Funktionale Eigenschaften dieser Verschaltung werden in Kap. 2 näher erläutert. Eine schematisierte Darstellung findet sich in Abb. 1.2. Zu beachten ist hier, daß bereits auf dieser frühen Stufe des Sehens eine starke Konvergenz von 1:100 auftritt, da die ca. 10^8 Ausgänge der Photorezeptoren auf ca. 10^6 Ganglienzellen abgebildet werden. Zu Beachten sind auch die verschiedenartigen Imperfektionen des Abbildungsapparats Auge. Die Ursachen dieser Störungen sind einerseits struktureller Natur, d.h. eine falsche Abstimmung der Lin-

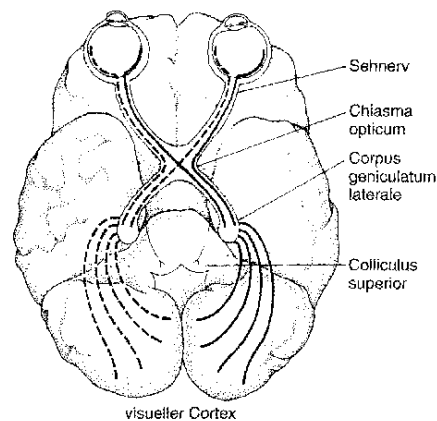


Abbildung 1.3: Darstellung des visuellen Pfades des frühen Sehens.

senbrennweite mit der Geometrie des Glaskörpers (Nah- bzw. Kurzsichtigkeit) oder eine anisotrope Krümmung der Kornea (Hornhaut). Der im letzten Fall auftretende Astigmatismus führt zu einer elliptischen Verzerrung des retinalen Bildes. Andere Ursachen für die Imperfektionen des Abbildungsapparates sind physikalischer Natur, gegeben durch Lichtbeugung und -streuung sowie nicht zuletzt durch Quantenfluktuationen. Desweiteren ist die Retina invertiert, wie auch anhand der Abb. 1.2 zu sehen ist, d.h. die photoempfindlichen Rezeptoren befinden sich auf der Rückseite der Retina, was zur Folge hat, daß das einfallende Licht sowohl die verschiedenen Schichten der neuronalen Elemente als auch die Zellkörper der Rezeptoren durchqueren muß, bevor es die Photopigmente an den äußeren Segmenten der Rezeptoren erreicht. Der *blinde Fleck* an der Stelle, an der die optischen Nervenfasern die Retina verlassen ist eine Konsequenz dieser Architektur. Erwähnenswert ist an dieser Stelle, daß die Bezeichnung „optischer Nerv“ nicht ganz zutreffend ist, da es sich bei der Retina um einen Teil des Thalamus handelt und diese sozusagen eine Ausstülpung des Gehirns darstellt. Für weitere Informationen bezüglich der Struktur und Anatomie der Retina sei auf das Werk von DOWLING verwiesen (Dowling, 1987).

Der Corpus Geniculatum Laterale. Die retinalen Ganglienzellen projizieren in den Corpus Geniculatum Laterale (seitlicher Kniehöcker), welcher ca. $1,3 \cdot 10^6$ Neurone enthält. Für den Verlauf des Pfades der visuellen Signale bis hin zum Kortex siehe Abb. 1.3. Der CGL besteht aus sechs Schichten, vier parvo- und zwei magnozellularen Schichten. Diese Bezeichnungen resultieren aus der unterschiedlichen Größe der Neurone im CGL und stellen separate Verarbeitungspfade dar, da sich diese unterschiedlichen Populationen sowohl bei den afferenten Ganglienzellen als auch bei weiteren verarbeitenden Instanzen im visuellen Kortex finden. Die parvozellulare Verarbeitung besitzt im Gegensatz zur magnozellularen eine wesentlich höhere örtliche Auflösung, wohingegen die zeitliche Auflösung im magnozellularen Pfad höher ist. Man findet, daß die P-Zellen

als zuständig für die Verarbeitung von Farbe und Form und die M-Zellen für temporale Information interpretiert werden können (Livingstone und Hubel, 1988).

Der primäre visuelle Kortex. Eine detaillierte Betrachtung der Anatomie des visuellen Kortex (Sehrinde) findet sich in (DeValois und DeValois, 1990); an dieser Stelle sollen nur einige wenige interessante Aspekte aufgezeigt und anschließend auf die Physiologie von retinalen und kortikalen rezeptiven Feldern eingegangen werden.

Als primärer visueller Kortex wird derjenige Teil der Sehrinde bezeichnet, der direkt vom CGL gespeist wird. Bei Primaten ist dies nur V1, während es bei Katzen V1, V2 und V3 sind. Eine wesentliche Eigenschaft des visuellen Kortex ist die topographische Bindung der Zellen, die sich bis hierher fortsetzt: Nervenfasern, welche an benachbarten retinalen Ganglienzellen beginnen, kontaktieren benachbarte CGL-Zellen und enden in benachbarten Punkten des V1 (Hubel und Wiesel, 1977). Ein Merkmal der kortikalen Informationsverarbeitung ist die hochgradige Spezialisierung der Neurone hinsichtlich Orientierung, Bewegungsrichtung und weiteren Eigenschaften des Eingangssignals. Diese hochgradig spezialisierte Darstellung setzt eine wesentlich höhere Anzahl Neurone voraus, was durch anatomische Untersuchungen auch bestätigt wurde (O’Kusky und Colonnier, 1982); 260 Millionen Neurone verarbeiten den Eingang von ca. 2 Millionen CGL-Fasern. Aufgrund einer Klassifizierung anhand von Zellgröße und -dichte sowie der Dichte der afferenten Nervenleitungen läßt sich der Kortex in sechs Schichten aufteilen. Diese können grob folgendermaßen beschrieben werden: (i) Schicht 4 wird gespeist vom CGL und stellt die Eingangsschicht des Kortex dar. (ii) Die Schichten 2 und 3 erhalten Signale von Neuronen aus Schicht 4 und projizieren auf andere kortikale Bereiche. (iii) Die tiefliegenden Schichten 5 und 6 erhalten Signale von Schicht 4 und projizieren zurück auf den CGL und in andere Teile des Gehirns. Im Zuge der retinotopen Organisation des Kortex wird von einem kortikalen Vergrößerungsfaktor gesprochen, mit dem benachbarte Bereiche der Retina auf größere, ebenfalls zusammenhängende Bereiche im visuellen Kortex abgebildet werden. Dieser Faktor hängt stark von der Exzentrizität auf der Retina ab, wobei der foveale Bereich stark überrepräsentiert ist.

Rezeptive Felder der Ganglien- und Genuculatumzellen. Wenn retinale Ganglienzellen durch einen visuellen Stimulus erregt werden, feuern sie eine Abfolge elektrischer Potentialstöße ab (*Spike-Train*). Diese können durch Anbringen einer Elektrode in der Nähe des Zellkörpers oder des Axons gemessen werden. Erste Untersuchungen wurden an Katzen durchgeführt, wobei KUFFLER entdeckte, daß die retinalen Ganglienzellen lediglich sensitiv auf Stimuli innerhalb eines kleinen, kreissymmetrischen Bereichs reagieren (Kuffler, 1953). Dieser Bereich wird als *rezeptives Feld* (RF) bezeichnet. Weiterhin hängt die Antwort auf einen Lichtpunkt von der Lage des Punktes innerhalb des rezeptiven Feldes ab. Einige Neurone werden exzitiert, wenn sich der Lichtpunkt im Zentrum befindet und inhibiert, wenn der umgebende Bereich beleuchtet wird. Diese Ganglienzellen werden als ON-center Zellen bezeichnet. Vice versa existieren dementsprechende OFF-center Ganglienzellen. Die typische exzitatorische und inhibitorische Reaktion einer

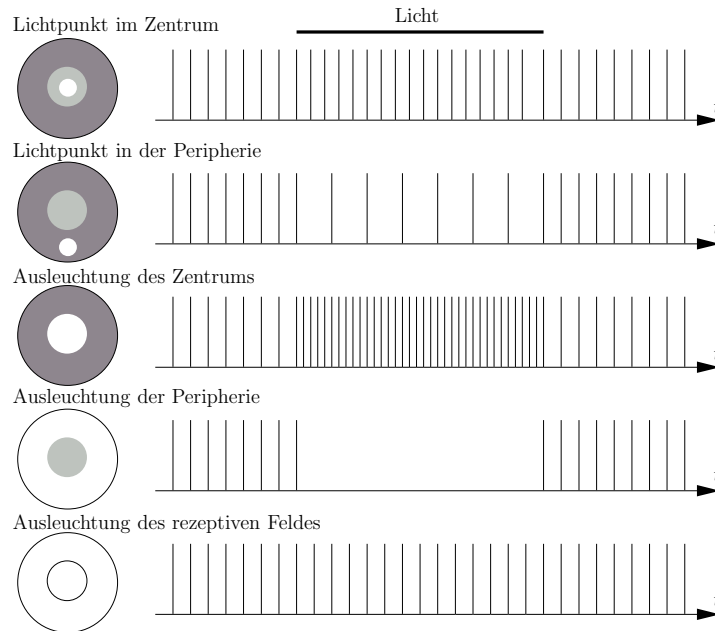


Abbildung 1.4: Antwortverhalten einer ON-center Ganglienzelle bei verschiedenen Stimuluskonfigurationen; die Form des rezeptiven Feldes wird aus offensichtlichen Gründen auch als *Mexican-Hat-Profil* bezeichnet.

ON-center Ganglienzelle ist in Abb. 1.4 gezeigt. Die Funktion dieser *lateralen Inhibition* wird in Abschnitt 2.6 genauer analysiert.

Die Profile der rezeptiven Felder der CGL-Neurone unterscheiden sich nur wenig von denen der Ganglienzellen, daher werden diese hier nicht näher besprochen. Für Details siehe beispielsweise (Kaplan und Shapley, 1982).

Rezeptive Felder im visuellen Kortex. Ähnlich den rezeptiven Feldern von retinalen Ganglienzellen antworten Neurone im visuellen Kortex nur auf Stimuli innerhalb eines gewissen Bereichs, dem sogenannten *klassischen rezeptiven Feld*. Ein wesentlicher Unterschied zu den isotropen rezeptiven Feldern der Ganglienzellen ist jedoch die von HUBEL und WIESEL entdeckte Orientierungselektivität (Hubel und Wiesel, 1968). Dies führt zu einer erhöhten Selektivität im Allgemeinen, da nur auf elongierte Stimuli mit bestimmter Orientierung eine Reaktion erfolgt. Klassische Versuche zur Bestimmung der Radial- und Orientierungsbandbreite verwenden Lichtbalken oder Sinusgitter. Gemessene Bandbreiten betragen für die Orientierung ca. 15° bis 30° und für die Radialbandbreite ca. 1-2 Oktaven. Ein typisches Ergebnis für die gemessene Parametrierung von 20 kortikalen Zellen im Mackaken zeigt Abb. 2.18 (siehe auch Abschnitt 2.6.3). Aus dieser Polardarstellung kann auch die Eigenschaft der Richtungsselektivität erkannt werden, die Zellen antworten auf bewegte Stimuli teilweise nur für eine bestimmte Richtung.

Abhängig von den Antworteigenschaften definierten HUBEL und WIESEL drei unterschiedliche Zelltypen im visuellen Kortex:

- Simple-Zellen,
- Complex-Zellen,
- Hypercomplex-Zellen.

Simple-Zellen zeichnen sich nach der klassischen Definition durch vier Charakteristika aus: *(i)* örtlich separate ON- und OFF-Regionen, *(ii)* lineare Summation in jeder Region, *(iii)* Antagonismus zwischen den ON- und OFF-Subregionen und *(iv)* die Antworteigenschaften können durch Analyse der rezeptiven Felder vorhergesagt werden¹. Eine andere Klassifizierungsmöglichkeit ist gegeben durch die Modulation der Antwort bei sinusförmigen Stimuli (Messung des f_1/f_0 -Verhältnisses, d.h. der Energie der ersten Harmonischen zur Energie des Gleichanteils). Die Zellpopulation in V1 weist eine bimodale Verteilung auf, entsprechend der Klassifikation in Simple- und Complex-Zellen; für einen Überblick siehe (DeValois und DeValois, 1990; Skottun

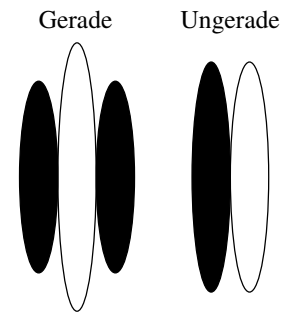


Abbildung 1.5: Stilisierte rezeptive Felder von Simple-Zellen im visuellen Kortex.

et al., 1991). Die empirisch bestimmten rezeptiven Felder weisen häufig eine gerade- und ungeradesymmetrische Struktur auf (Abb. 1.5)². Eine wesentliche Eigenschaft die sich hieraus ergibt ist die Phasenselektivität, bzw. Sensitivität gegenüber Lage und Orientierung des Stimulus innerhalb des RF. Häufig finden sich auch funktionale Bezeichnungen der Art Balken- und Kantendetektor, die sich aus den gezeigten symmetrischen rezeptiven Feldern ergeben. Complex-Zellen hingegen zeichnen sich durch eine Phaseninvarianz aus, d.h. sie antworten unabhängig von der Stimulusposition im rezeptiven Feld. Von einer funktionalen Betrachtungsweise aus ist diese Invarianzleistung für eine komplexe Objekterkennung notwendig, jedoch geht hierbei auch relevante Positionsinformation verloren. Nach einem hierarchischen Modell, erstmals propagiert von HUBEL und WIESEL erhalten Complex-Zellen Eingang von örtlich verschobenen Simple-Zellen und summieren über die Position. Alternative Modelle schlagen nichtlineare Verknüpfungen vor (Riesenhuber und Poggio, 1999). Eine weitere Klasse sind die Hyperkomplex-Zellen. Diese weisen ebenfalls eine Orientierungsabstimmung auf wie die Simple- oder Complex-Zellen, zeigen jedoch eine weitere nichtlineare Eigenschaft: sie reagieren nicht auf ausgedehnte Linien oder Gitter, sondern nur auf kurze Balken oder Ecken (diese stellen $i2D$ -Signale dar, siehe Abschnitt 2.2). Das Antwortverhalten einer derartigen Zelle ist in Abb. 1.6 dargestellt.

¹Complex-Zellen werden einfach dadurch definiert, daß sie die geforderten Bedingungen für Simple-Zellen nicht erfüllen.

²Die Ergebnisse hier sind nicht ganz eindeutig, (Jones und Palmer, 1987b) finden in den meisten Simple-Zellen keine Symmetrie, jedoch wurde in (Pollen und Ronner, 1981) gezeigt, daß benachbarte Simple-Zellen eine Phasenverschiebung von ca. 90° besitzen.

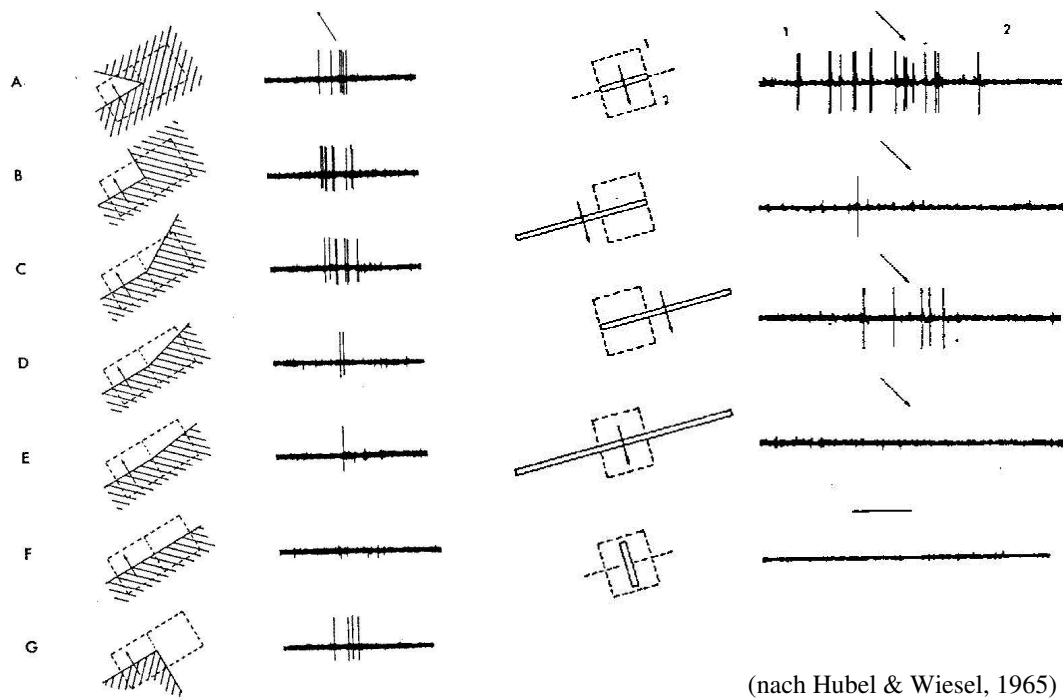


Abbildung 1.6: Antwortverhalten einer Hypercomplex-Zelle im visuellen Kortex einer Katze.

Wichtig ist zu erwähnen, daß sowohl die Annahmen der diskreten Klassifizierung unterschiedlicher Zelltypen als auch die konstruktiven hierarchischen Modelle nach wie vor diskutiert werden und eine ungeheure Anzahl an widersprüchlicher Literatur existiert. Die hier vorgestellten Methoden sind eher klassischer Natur und die Basis für die meisten Arbeiten auf diesem Gebiet.

2 Informationsverarbeitung im visuellen System

2.1 Einführung

Zum Einstieg in dieses Kapitel seien die zwei Bildpaare in Abb. 2.1 gezeigt. Leicht zu erkennen ist, daß es sich beim linken Paar um zwei Realisierungen natürlicher Bilder handelt, beim rechten um ein Rauschen. Die Frage jedoch, die an dieser Stelle von Interesse ist, lautet: sind die Bilder pro Paar identisch? Selbstverständlich handelt es sich beim Paar der Klasse der natürlichen Bilder um zwei völlig unterschiedliche Objekte, doch wie sieht unsere Diskriminationsfähigkeit für die Klasse der Rauschbilder aus? Ohne genaues Betrachten läßt sich keine Entscheidung treffen. Die Relevanz dieses „Experiments“ liegt darin, daß die rechts dargestellten Rauschbilder (*i*) objektiv (d.h. nach mathematischen Abstandsmaßen, vgl. Abschnitt 2.4) weitaus unterschiedlicher sind und (*ii*) die Klasse der Rauschbilder wesentlich größer ist als die Klasse der natürlichen Bilder. Dies wurde beispielsweise in (Zetzsche, 2002) gezeigt und wird in Abschnitt 2.4 noch ausführlicher erläutert. Derartige Experimente führen zu dem Schluß, daß sich das visuelle System, wie auch andere sensorische Systeme, im Laufe der Evolution hochgradig auf die Verarbeitung natürlicher Stimuli spezialisiert hat.

Bereits HELMHOLTZ postulierte, daß die *Apperzeption*, also das bewußte, willensgesteuerte Erfassen und Verarbeiten von Sinneseindrücken, aus der Interaktion der *Perzeption* (Vorgang des Auffassens und Erkennens ohne gedankliche Verarbeitung und

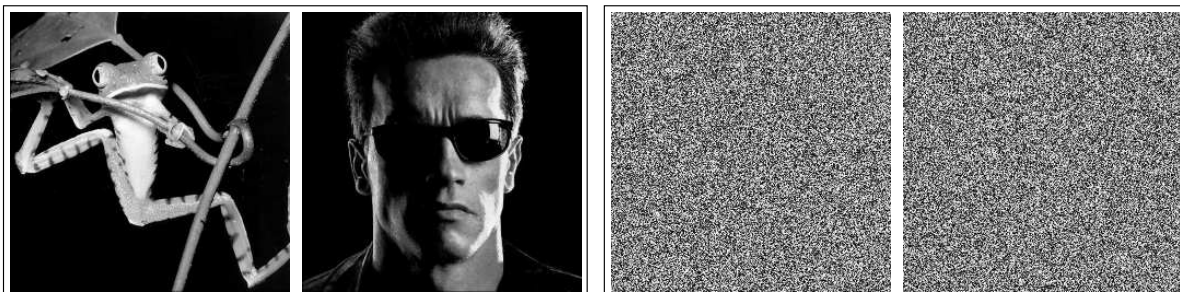


Abbildung 2.1: Jeweils zwei Realisierungen der Klasse der natürlichen Bilder (links) und der Rauschbilder (rechts).

Beurteilung) mit erinnerten Ideen, entstanden aus vergangener Erfahrung, resultiert (Helmholtz, 1910). Auch MACH (Mach, 1922) und PEARSON (Pearson, 1892) betonten die Relevanz der von der Umwelt a priori gegebenen Strukturen für die Verarbeitung visueller Information. Sie argumentierten, daß physikalische Konzepte und Gesetzmäßigkeiten unsere komplexe Erfahrung der Welt vereinfachen und daß sie eine *Ökonomie des Denkens* in unsere mentalen Prozesse bringen. Diese Hypothesen sind aufgrund ihres intuitiven Ansatzes attraktiv, jedoch können sie ohne quantifizierende Methoden nicht verifiziert werden. Primär gilt zu klären, *was* die Gesetzmäßigkeiten in unserer Umwelt sind, und *wie* sie beschrieben werden können. Desweiteren muß notwendigerweise nach den Kriterien gefragt werden, nach welchen eine (evolutionäre) Anpassung an diese Gesetzmäßigkeiten stattgefunden hat. Pionierarbeiten wurden hier von ATTNEAVE (Attneave, 1954) durchgeführt, der Konzepte aus der von SHANNON begründeten Informationstheorie (Shannon, 1948) in die Psychologie einbrachte. Diese befaßt sich grundsätzlich mit der Quantisierung und Übertragung von Information über bandbegrenzte Kanäle; eine kurze Einführung findet sich in Abschnitt 2.3. Die wesentlichen Konzepte, die von ATTNEAVE verwendet wurden, sind die Definitionen von *Information*, *Kanalkapazität* und *Redundanz*. Über einen Kommunikationskanal kann Information mit Raten bis zu einer bestimmten Grenze, der Kanalkapazität, übertragen werden. Nachrichten beinhalten jedoch häufig weniger als diese maximal erreichbare Information; die Differenz wird dann als Redundanz bezeichnet. Wichtig ist hierbei, daß jede Form der Regularität in Nachrichten eine Quelle von Redundanz darstellt, und daß diese Regularität durch statistische Abhängigkeiten zwischen den Symbolen der Nachricht beschrieben werden kann. Im Falle von natürlichen Bildern entspräche dies Abhängigkeiten zwischen benachbarten Bildregionen, die aufgrund von morphologischen Eigenschaften der Objekte in der Umwelt existieren.

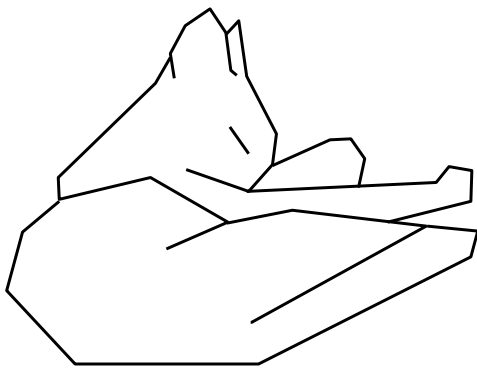


Abbildung 2.2: Attneavesche Katze.

Er illustrierte dies mit dem berühmten Beispiel eines Bildes einer schlafenden Katze, indem einfach die Punkte der Objektgrenze mit sich ändernder Orientierung gerade verbunden wurden (Abb. 2.2).

ATTNEAVE postulierte, daß natürliche Bilder ein hohes Maß an Redundanz beinhalten und daß das subjektive Hervortreten von Kanten bzw. Objektgrenzen ein Beispiel für einen Mechanismus darstellt, der diese Eigenschaft ausnutzt. Es könne ein Objekt effizienter repräsentiert werden, indem nur die Übergänge zwischen Objekten übertragen würden, da diese die unerwarteten, quasi informationsträchtigen, Signalanteile darstellten. Desweiteren seien gerade Kanten weniger informativ als Änderungen in der Orientierung der Objektgrenzen.

Auch BARLOW beschäftigte sich mit der Anwendung der Informationstheorie auf die Verarbeitung natürlicher sensorischer Signale. In einer seiner ersten Arbeiten (Barlow, 1961)

führte er das Prinzip der *spärlichen Kodierung* ein, d.h. die Ökonomie der Verarbeitung liegt in der Reduktion der Aktivität der Neurone begründet und nicht in der Anzahl der an der Verarbeitung beteiligten (siehe hierzu auch Abschnitt 2.5.3). Er beschäftigte sich auch mit dem experimentellen Nachweis, daß Neurone in den sensorischen Pfaden ein hochgradig selektives Antwortverhalten aufweisen (Barlow, 1972) und somit perzeptuelle Unterschiede auf die Aktivitätsänderung weniger Neurone anstatt auf die globale Veränderung einer großen Neuronenpopulation zurückzuführen sind. BARLOW betont jedoch, daß der direkte Vergleich mit ingenieursspezifischen Nachrichtenübertragungsprinzipien hinkt, da die Kosten in biologischen und technischen Systemen nicht notwendigerweise identisch sind (Barlow, 2001a) und die Kosten für effizientes Lernen durchaus eine übervollständige Repräsentation empfehlen (Gardner-Medwin und Barlow, 2001). Dies zeigt sich auch an der Tatsache, daß allein im primären visuellen Kortex (V1) des Menschen ungefähr drei Größenordnungen mehr Neurone ($\sim 10^9$) vorhanden sind, als Ganglienzellen in beiden Retinae ($\sim 2 \times 10^6$). Eine bloße Eliminierung von Redundanz hin zu einem kompakten Kode würde eine Reduktion der Anzahl der mit der visuellen Wahrnehmung beschäftigten Neurone verlangen. Redundanz kann, wenn sie in einer dem Organismus „bekannt“ Art vorliegt, vorteilhaft genutzt werden, beispielsweise zur Steigerung des Signal-zu-Rauschleistungsverhältnisses, für assoziatives Lernen oder Prädiktion (siehe (Barlow, 2001b) und Abschnitt 2.5). Die Bedeutung von spärlicher Kodierung und übervollständigen Repräsentationen für die Kodierung natürlicher Bilder im visuellen Kortex wurde auch von ZETZSCHE und Mitarbeitern (Zetzsche, 1990; Zetzsche und Krieger, 2001b; Zetzsche und Röhrbein, 2001; Zetzsche und Nuding, 2005b) und FIELD (Field, 1987; Field, 1999) untersucht. Ansätze unter Verwendung von neuronalen Netzen liefern weitere Evidenz für die Anpassung der visuellen Neurone an die Statistik natürlicher Bilder (Olshausen und Field, 1996; van Hateren und van der Schaaf, 1998). Eine detaillierte Beschreibung erfolgt in Abschnitt 2.5; eine eigene technische Anwendung zur Kodierung natürlicher Bilder wird in Kapitel 4 behandelt.

2.2 Das Konzept der intrinsischen Dimensionalität

Das Konzept der intrinsischen Dimensionalität in der Bildverarbeitung wurde in (Zetzsche und Barth, 1990a; Zetzsche und Barth, 1990b) eingeführt. Es setzt die Anzahl der Freiheitsgrade, die ein Signalraum zur Verfügung stellt, mit der mindestens benötigten Anzahl zur Beschreibung des Signals in Relation. Dies ist vergleichbar mit dem Rang einer Matrix und steht in einem engen Zusammenhang mit dem Begriff der Redundanz, der in Kap. 2 näher erläutert wird.

Nach diesem Konzept ist ein $i0D$ -Signal definiert durch einen lokal konstanten Luminanzverlauf:

$$u_{i0D}(x, y) = \text{const.} \quad (2.1)$$

Entsprechend ist ein *i1D*-Signal definiert als ein Signal mit einem effektiven Freiheitsgrad, d.h. das Luminanzprofil ist vollständig durch die Abhängigkeit in einer Variablen charakterisiert:

$$u_{i1D}(x, y) = u_{i1D}(x + \Delta x, y + \Delta y), \quad \Delta y = k \cdot \Delta x, k \in R, \forall \Delta x \in R. \quad (2.2)$$

Anschaulich entsprechen also *i1D*-Signale geraden Linien und Kanten, aber auch kontinuierlich modulierte Verläufe entlang einer Achse, wie beispielsweise Sinusgitter, die häufig in psychophysischen und neurophysiologischen Untersuchungen verwendet werden, gehören zu dieser Klasse.

Sämtliche übrigen Signale, also die Klasse der Signale, die zu ihrer Beschreibung die volle Anzahl Freiheitsgrade benötigen, werden dann dementsprechend als *i2D*-Signale bezeichnet. Diese sind dann typischerweise Ecken, Linienenden, Krümmungen (Kurven) und „Plaid-Muster“, also überlagerte Sinus-Gitter. Eine Modifikation dieses Konzepts der intrinsischen Dimensionalität findet sich in (Krüger und Felsberg, 2003); hier wird eine Generalisierung hin zu einer kontinuierlichen Klassifikation untersucht.

2.3 Die Shannon'sche Informationstheorie

Die Informationstheorie entstand in den 40er und 50er Jahren des 20. Jahrhunderts als Antwort auf den Bedarf der Elektro- und Nachrichtentechniker praktikable Kommunikationsgeräte zu entwickeln. Diese mathematisch fundierte Theorie wurde von CLAUDE E. SHANNON (Abb. 2.3) begründet und befaßt sich, vereinfacht gesagt,



Abbildung 2.3: Claude Elwood Shannon.

mit der Quantifizierung und Übertragung von Information über bandbegrenzte Kanäle (Shannon, 1948; Shannon und Weaver, 1976). Sie liefert quantitative Aussagen über die Effizienz einer Informationsrepräsentation und die Grenzen für die Informationsraten, um eine verlässliche Übertragung über eine rauschbehaftete Kommunikationsstrecke zu gewährleisten. In der Informationstheorie wird jedes Gerät, System oder jeder Prozeß, welcher Nachrichten oder Signale generiert, als Informationsquelle bezeichnet. Jede Informationsquelle hat als Ausgang Kombinationen von Symbolen, die entsprechend den *a priori* Wahrscheinlichkeiten der Quelle generiert werden. Eine Nachricht \vec{x} der Länge N kann also generell als eine Aneinanderreihung von N Zufallsgrößen beschrieben werden:

$$\vec{x} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}. \quad (2.3)$$

Der klassische Aufbau eines allgemeinen Kommunikationssystems nach SHANNON besteht aus Quelle, Sender, Empfänger und Informationssenke. In realita muß noch ein Übertragungsrauschen mit in Betracht gezogen werden, siehe Abb. 2.4.

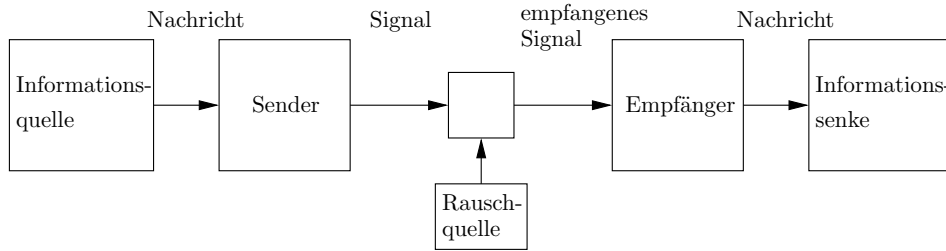


Abbildung 2.4: Schema eines generellen Kommunikationssystems nach (Shannon, 1948).

In dieser Arbeit ist vor allem der Begriff der Redundanz und die Transformation bzw. Eliminierung der Redundanz einer Quelle von Bedeutung. Wichtig ist die Differenzierung zweier Ursprünge von Redundanz: Man unterscheidet zwischen der Redundanz aufgrund der Auftretswahrscheinlichkeiten der einzelnen Symbole und der Redundanz aufgrund von statistischen Abhängigkeiten zwischen aufeinanderfolgenden Symbolen. Diese sogenannten *Intersymbolabhängigkeiten* stellen die hauptsächliche Quelle von Redundanz in natürlichen Signalen dar. Zum Beispiel wurde für geschriebenes Englisch erstmals von SHANNON die Entropie pro Zeichen geschätzt (Shannon, 1951). Bei dem von ihm verwendeten Paradigma wurden die statistischen Abhängigkeiten zwischen den Zeichen mitberücksichtigt. Er stellte fest, daß die Redundanz aufgrund der Ungleichverteilung der Symbole (Zeichen des Alphabets) ungefähr 15% ausmacht, wohingegen die Intersymbolabhängigkeiten mit ca. 55% deutlich mehr zur Redundanz beitragen. Insgesamt beträgt die Redundanz des geschriebenen Englisch also ca. 70%, das entspricht einer Entropie von 1,4 bit/Zeichen¹. Für den in dieser Arbeit hauptsächlich betrachteten Fall natürlicher Bilder, dargestellt als beispielsweise 8-bit kodierte Grauwertbilder (die Mächtigkeit der Menge aller möglichen Symbole beträgt $L = 2^8 = 256$), existieren ebenfalls massive statistische Abhängigkeiten zwischen den Symbolen (siehe auch die Abschnitte 2.4 und 5.1.3). Eine dem Shannon'schen Verfahren zur Schätzung der Entropie von geschriebenem Englisch analoge Untersuchung für die Entropie natürlicher Bilder wurde von KERSTEN durchgeführt (Kersten, 1987).

Diese Unterscheidung macht auch die unterschiedlichen Formen der Kodierung, die in späteren Abschnitten verwendet werden, explizit. Ein Kode, welcher die gesamte Redundanz minimiert, wird als *Minimum-Redundanz* Kode bezeichnet, während ein Kode, der lediglich die Intersymbolabhängigkeiten minimiert, als *Minimum-Entropie* oder auch als *faktorieller* Kode bezeichnet wird (siehe beispielsweise Abschnitt 2.5.2). Ein derartiger Kode generiert statistisch unabhängige Zufallsgrößen (ZG), so daß sich die gesamte Verbundwahrscheinlichkeitsdichtefunktion (VWDF) als dyadisches Produkt der Einzelwahrscheinlichkeitsfunktionen ergibt:

$$p(x_1, x_2, \dots, x_N) = p_{x_1}(x_1) p_{x_2}(x_2) \cdots p_{x_N}(x_N) \quad (2.4)$$

¹Für Schätzungen der Redundanz weiterer westlicher Sprachen siehe (Barnard, 1955).

Wichtig ist zu erwähnen, daß eine Umkodierung welche die Intersymbolabhängigkeiten eliminiert, per se nicht Redundanz reduzierend ist, sondern daß diese konstant gehalten wird, indem die Redundanz aufgrund von Intersymbolabhängigkeiten in eine Redundanz aufgrund von Ungleichverteilung transformiert wird. Die Bedeutung solcher faktorieller Codes liegt darin, daß in einem simplen zweiten Schritt die ZGen separat mit einem optimalen Redundanz reduzierenden Kode, wie beispielsweise dem Huffman-Kode, kodiert werden können. Wollte man denselben Effekt direkt erreichen, so bedarf es der Kenntnis der vollen, nicht separierbaren Verbundwahrscheinlichkeitsfunktion, deren Redundanz unter Verwendung von beispielsweise einem Vektor-Quantisierer ausgenutzt werden kann. Die Kenntnis dieser vollen VWF ist jedoch nur in Spezialfällen möglich und eine optimale Kodierung somit erschwert.

2.4 Statistik natürlicher Bilder

Im letzten Abschnitt wurden der Begriff der *Redundanz* eingeführt und erläutert, wie natürliche Quellen prinzipiell in eine effizientere Repräsentation überführt werden können. Essentiell für die Konstruktion eines derartig effizienten Codes ist die (partielle) Kenntnis der Statistik der Quelle. In diesem Abschnitt wird die Quellenstatistik für den in dieser Arbeit wesentlichen Fall der natürlichen Bilder behandelt. Einige relevante Ansätze und Paradigmen werden im Detail eruiert.

2.4.1 Signalraumdarstellung kontinuierlicher Signale

Im Kontext der digitalen Bildverarbeitung werden natürliche Bilder als diskrete Quellen behandelt. Die statistischen Verteilungsfunktionen und Auftrittswahrscheinlichkeiten entsprechen also denen diskreter Zufallsgrößen (Papoulis, 1991). Ein wesentlicher Aspekt hierbei ist die Darstellung im *Signalraum*, in welchem eine Realisierung eines natürlichen Bildes genau einem eindeutigen Punkt entspricht. Den Übergang von einem kontinuierlichen Signal in einen diskreten Signalraum bezeichnet man als *Signalraumdarstellung*. Der Fall eines einfachen eindimensionalen Zeitsignals, bestehend aus drei diskreten Abtastwerten ist in Abb. 2.5 illustriert. Ein kontinuierliches Signal kann entsprechend dem Abtasttheorem in eine diskrete Darstellung unter Erhaltung der vollständigen Information überführt werden, d.h. es kann ohne Fehler rekonstruiert werden (Haykin, 1978; Kammeyer, 1996). Bedingung hierfür ist, daß die Abtastrate (Nyquist-Frequenz) $1/\Delta t$ größer als die doppelte maximal im Signal vorhandene Frequenz f_{max} ist:

$$\Delta t < \frac{1}{2f_{max}}. \quad (2.5)$$

Wenn nun als eine Realisierung jeweils ein Ausschnitt der Länge T angenommen wird, so ist diese als ein $\frac{T}{\Delta t} =: N$ -dimensionaler Vektor in einem Koordinatensystem darstell-

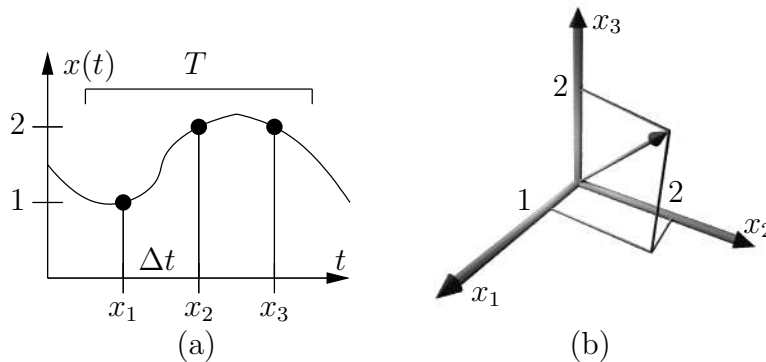


Abbildung 2.5: (a) Diskretisierung eines eindimensionalen Zeitsignals mit drei Abtastwerten. (b) Korrespondierende Signalraumdarstellung. Die Zuordnung ist ein-eindeutig.

bar. Die vollständige Kenntnis der Signalstatistik ist dann durch die N -dimensionale Verbundwahrscheinlichkeitsfunktion gegeben. Im Falle zweidimensionaler diskretisierter Bilder entspricht jeder Pixel einer Dimension im Signalraum.

2.4.2 Natürliche Bilder im Signalraum

Die Problematik, die sich für die Kenntnis der vollständigen Statistik natürlicher Bilder ergibt, kann leicht anhand der weiter oben beschriebenen Signalraumrepräsentation aufgezeigt werden: Betrachtet man für die Bildverarbeitung typische 8-bit kodierte Grauwertbilder der Größe 512×512 , so entspricht ein Bild einem 512^2 -dimensionalen Vektor. Die Mächtigkeit der zugrundeliegenden Gesamtmenge aller möglichen Realisierungen \mathcal{S}_{bild} ergibt sich zu $256^{512^2} = 256^{262144}$. Aus diesen Überlegungen wird ersichtlich, daß es weder möglich ist, Kenntnis über die vollständige Statistik natürlicher Bilder zu erlangen, noch daß eine genügend große Datenbasis verarbeitet werden kann, um diesen Raum zu füllen.

Wie am einführenden Beispiel dieses Kapitels gezeigt (Abb. 2.1), ist die Diskriminationsfähigkeit des visuellen Systems für die Klasse der natürlichen Bilder wesentlich besser als für die der Rauschbilder (Abb. 2.6). Mit Hilfe der Signalraumdarstellung kann diese Hypothese untermauert und quantifiziert werden. ZETZSCHE untersuchte in (Zetzsche, 2002) die Zufälligkeit natürlicher Bilder mit unterschiedlichen Ansätzen. Unter anderem kann unter Verwendung der Theorie der *algorithmischen Komplexität* (Chaitin, 1987; Li und Vitanyi, 1993) gezeigt werden, daß im Grenzfall nahezu alle möglichen Bilder zufällig sind. Nach dieser Theorie ist Zufälligkeit als Inkompressibilität definiert: Das kürzeste Programm, welches eine zufällige Folge beschreibt, bzw. generieren kann, darf nicht signifikant kompakter sein, als die Folge selbst. Natürliche Bilder, dementsprechend nicht-zufällig, da hochkompressibel, beschreiben also eine verschwindend geringe Menge, siehe Abb. 2.6. Dieses Ergebnis liefert weitere Evidenz dafür, daß das visuelle

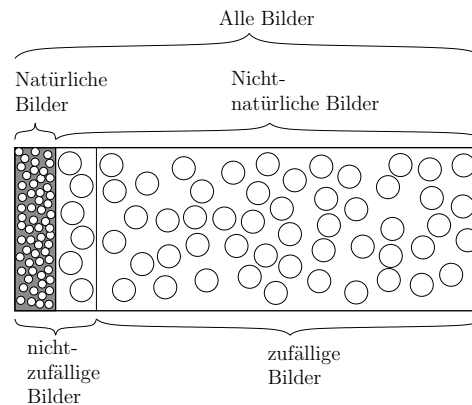


Abbildung 2.6: Die Diskriminationsleistung (Granularität) des visuellen Systems ist für natürliche Bilder deutlich höher als für die Klasse der Rauschbilder. Zudem ist die Klasse der natürlichen Bilder, verglichen mit der Gesamtheit aller möglichen Bilder, verschwindend gering.

System hochgradig an die statistischen Eigenschaften der Umwelt angepaßt ist, da innerhalb einer sehr unwahrscheinlichen Klasse, die jedoch unsere Umwelt konstituiert, die „Auflösung“ deutlich höher ist als für die typische Klasse.

2.4.3 Statistiken zweiter Ordnung natürlicher Bilder

Ein wesentliches Merkmal natürlicher Bilder ist die hohe Korrelation zweier benachbarter Bildpunkte. Dies kann anhand von sogenannten *Streudiagrammen* verdeutlicht werden (Abb. 2.7). Die Luminanzwerte benachbarter Pixel weisen häufig einen gleichen oder ähnlichen Wert auf, siehe Abb. 2.7(a). Mit zunehmendem Abstand der korrelierten Pixel nimmt diese Korrelation jedoch sehr schnell ab, Abb. 2.7(b). Vergleicht man dies hingegen mit der Korrelation zweier unmittelbar benachbarter Pixel eines statistisch unabhängigen Rauschens, so wird der Unterschied „zweiter Ordnung“ zwischen Rauschen und natürlichen Bildern schnell sichtbar; dies ist exemplarisch in Abb. 2.7(c) gezeigt.

Der Grund hierfür ist die morphologische Beschaffenheit der Objekte in unserer Umwelt. Objekte sind durch Objektgrenzen definiert, die typischerweise unetige Luminanzübergänge darstellen. Die interne Struktur der Objekte weist jedoch eine hohe Konstanz auf, gekennzeichnet durch stetige Übergänge im Luminanzprofil². In (Ruderman, 1997) wurde gezeigt, daß die Korrelation benachbarter Bildpunkte innerhalb eines Objekts im Mittel signifikant höher ist als zwischen Bildpunkten über Objektgrenzen hinweg. Die Autokorrelationsfunktion (AKF) spiegelt die Korrelationen zwischen Bildpunkten unterschiedlicher Distanz wieder. Es zeigt sich, daß die AKF natürlicher Bilder isotrop ist und einen radialsymmetrischen Abfall aufweist. Die korrespondierende Re-

²Texturähnliche Oberflächenstrukturen zeichnen sich durch andere statistische Eigenschaften aus.

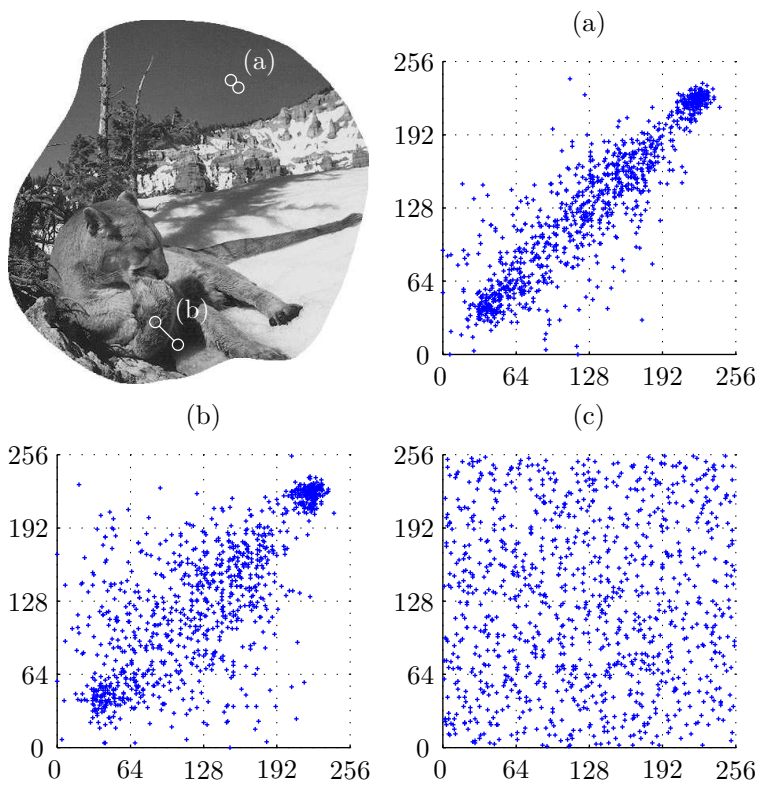


Abbildung 2.7: Streudiagramme zweier benachbarter Punkte eines natürlichen Bildes (a,b) und eines statistisch unabhängigen Rauschens (c). Die Achsen bezeichnen jeweils die Luminanzwerte eines Pixels. Die Distanz ist für (b) größer als für (a), was in einer erhöhten Streuung resultiert. Bei unabhängigem Rauschen ist bereits bei minimaler Distanz keine Korrelation mehr zu erkennen (c).

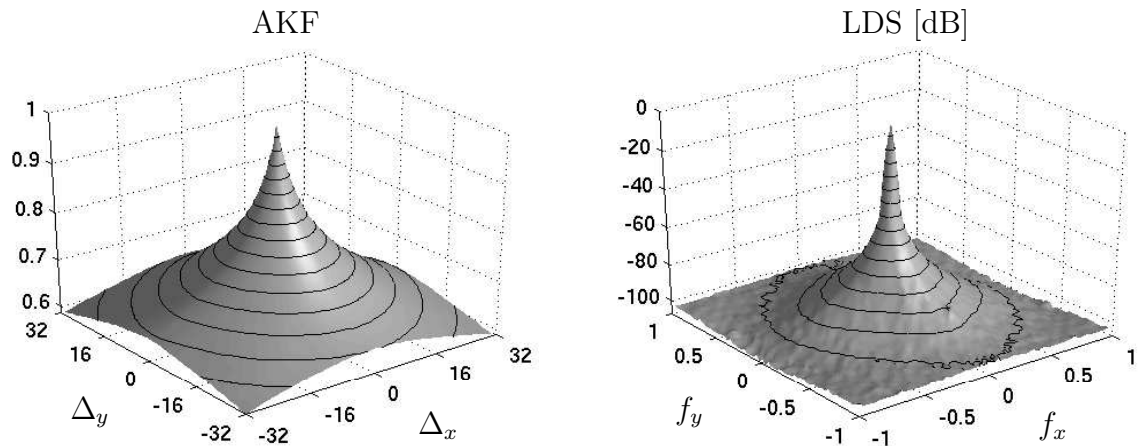


Abbildung 2.8: Autokorrelationsfunktion (links) und Leistungsdichtespektrum (rechts) einer Menge von natürlichen, kalibrierten Bildern (die Frequenzachsen sind auf die Nyquist-Frequenz normiert).

präsentation im Frequenzbereich ist das (Auto-) Leistungsdichtespektrum (LDS) $C_2(\vec{f})$. Dieses kann für eine große Klasse natürlicher Bilder durch ein Potenzgesetz modelliert werden (Field, 1987; Ruderman, 1997):

$$C_2(\vec{f}) = \frac{A}{|\vec{f}|^{2-\xi}}, \quad (2.6)$$

worin A eine Konstante entsprechend dem mittleren Kontrast und ξ den „Anomalie-Exponent“ darstellt, ein kleiner Wert, um statistischen Abweichungen gerecht zu werden. Als Ursache für dieses Potenzgesetz wird unter anderem die fraktale Natur der Umwelt oder die inhärente Konsistenz der Objekte, zusammen mit der Eigenschaft, daß Objekte voneinander statistisch unabhängig sind, genannt. Die in Abb. 2.8 gezeigte Autokorrelationsfunktion und das Leistungsdichtespektrum sind auf der Basis natürlicher Bilder der kalibrierten Bilddatenbank von HANS VAN HATEREN (van Hateren und van der Schaaf, 1998) berechnet.

Die Statistiken zweiter Ordnung natürlicher Bilder sind gut untersucht. Optimallösungen bzgl. der zweiten Ordnung zur effizienten Kodierung sind weithin bekannt und häufig auch analytisch lösbar. Mit Hilfe von Statistiken zweiter Ordnung können beispielsweise Aufgaben im Bereich Signaldetektion (Matched-Filter), Rauschunterdrückung (Wiener-Filter) und Redundanzreduktion gelöst werden (vgl. Karhunen-Loève-Transformation, Abschnitt 2.5.2). In Abschnitt 2.6 wird erläutert, inwiefern diese Lösungen in biologischen Sehsystemen implementiert sind, bzw. wie funktionelle Eigenschaften des visuellen Systems mit Hilfe von Statistiken zweiter Ordnung verstehbar sind. Den Grenzen und Problemen dieses Ansatzes widmet sich Abschnitt 2.7.

2.5 Optimale Transformation natürlicher sensorischer Stimuli

Wie bereits weiter oben erläutert, bestehen natürliche sensorische Stimuli, insbesondere natürliche visuelle Stimuli, nicht aus unstrukturierten und zufälligen Mustern, sondern zeichnen sich durch starke statistische Abhängigkeiten und demzufolge Redundanz aus. Die essentielle Frage ist nun: wie wird diese Redundanz im Sinne einer generellen effizienten Verarbeitung in einem optimalen „Signalprozessor“, nämlich dem visuellen Pfad, verwendet?

2.5.1 Quellenkodierung im visuellen System

Zunächst sollte die Dichotomie des Begriffs „effiziente Kodierung“, wie in Abschnitt 2.3 kurz dargestellt, für biologische sensorische Systeme näher betrachtet werden. Diese Dichotomie, also die Reduzierung der Redundanz zum Einen aufgrund der Ungleichverteilung der einzelnen Symbole, zum Anderen aufgrund von Intersymbolabhängigkeiten, entspricht der Kodierung in einzelnen Neuronen, bzw. der Kodierung in Neuronenverbänden. Betrachtet sei nun also die Aktivitätsverteilung der Antwort eines einzelnen sensorischen Neurons auf natürliche Stimuli. Um zu entscheiden, ob das Neuron im Mittel die maximal mögliche Information kodiert, bedarf es der Kenntnis der Verteilung mit maximaler Entropie. Im einfachen Falle diskreter ZGen ist dies eine Gleichverteilung (Shannon und Weaver, 1976). Bei neuronalen Aktionspotentialen oder Feuerraten handelt es sich jedoch um kontinuierliche ZGen, wobei die Verteilung mit maximaler Entropie von einer Randbedingung abhängt. Wenn beispielsweise die Antwort beschränkt ist auf x_{max} , so ist die Verteilung mit maximaler Entropie eine Gleichverteilung im Intervall $[0, x_{max}]$. Ist jedoch die Randbedingung, daß die Varianz der Antwort einen bestimmten Wert hat, so ergibt sich eine Gaußverteilung. Bei definiertem Mittelwert ist die informationsmaximierende Verteilung eine Exponentialfunktion³. Evidenz für die Behauptung, daß einzelne Neuronen ihre Antwortstatistik „optimieren“, findet man bereits in (Laughlin, 1981). Hierin wurde die Kontrast-Sensitivitätsfunktion der *LMC-Zelle* einer Fliege gemessen und gezeigt, daß diese tatsächlich die Kontrastverteilung der natürlichen Umgebung in eine ungefähre Gleichverteilung transformiert. Ebenso konnte in (Baddeley et al., 1998) gezeigt werden, daß die spontanen Feuerraten von Neuronen im primären und inferotemporalen visuellen Kortex von Katzen und Affen exponentiell verteilt sind. Dies ist konsistent mit einer optimalen Kodierung unter der Randbedingung einer konstanten mittleren Feuerrate.

Betrachtet man nun allerdings die Kodierung sensorischer Stimuli durch eine ganze Po-

³Im Allgemeinen ergibt sich für eine Randbedingung der Form $E\{\Phi(x)\} = c$, wobei Φ eine Randbedingungsfunktion und c eine Konstante ist, eine Verteilung $f_{H_{max}}(x) \propto e^{-\lambda\Phi(x)}$, $\lambda = const.$

pulation von Neuronen, so verlangt ein mögliches Optimalitätskriterium, daß die Antworten der Neurone voneinander statistisch unabhängig sein müssen. In anderen Worten bedeutet dies, daß bei Beobachtung der Antwort eines bestimmten Neurons keinerlei Rückschlüsse auf die Antworten der anderen Neurone in der Population gemacht werden können⁴. Der Vorteil dieses Kriteriums ist, daß keine weiteren Randbedingungen angenommen werden müssen. Wichtig ist zu erwähnen, daß der häufig propagierte Fall einer spärlichen und verteilten Kodierung im visuellen Kortex zu starken statistischen Abhängigkeiten der Einheiten führt. Dies liegt begründet im Antwortverhalten, da im Grenzfall für jeden möglichen Stimulus nur ein Neuron antwortet, und somit die Aktivität der anderen gleich Null ist (analog zu dem Verfahren der Vektorquantisierung). Dieser offenbare Konflikt soll nun kurz eruiert werden.

Die Optimierungskriterien in biologischen Systemen sind mannigfaltig. Eine Betrachtung der Signalverarbeitung lediglich aus der Perspektive der absoluten Reduktion der Redundanz simplifiziert das gegebene Problem unangemessen. Eine rein technische Anwendung der Quellencodierung würde einen *reversiblen* Redundanz reduzierenden Kode fordern. Hier muß jedoch ein Unterschied gemacht werden zwischen *Reversibilität* und *Selektivität* eines Kodes. Für biologische Organismen stellt sich häufig weniger die Frage nach der perfekten Rekonstruktion des Eingangssignals unter Verwendung eines möglichst redundanzfreien Kodes, sondern vielmehr nach der möglichst schnellen Detektion eines gewissen Stimulus und einer daraus ableitbaren überlebensrelevanten Reaktion. Ein klassisches Beispiel für einen derartigen Mechanismus sind die sogenannten „Bug-detectors“ in der Retina des Frosches, die auf Reize, die Futter bedeuten reagieren, und den Frosch zu Fangreaktionen veranlassen (Lettvin et al., 1961).

Eine Reduzierung der Redundanz und eine Verwendung eines möglichst kompakten Kodes ist dann sinnvoll, wenn Engpässe in der maximal möglichen Übertragungsrate im visuellen Pfad auftreten. Dies ist durchaus nicht unwahrscheinlich, unter Betracht der Tatsache, daß reale, visuelle Neurone einen beschränkten dynamischen Antwortbereich besitzen und die interneuronalen Verknüpfungen bandbegrenzt sind (Shapley und Enroth-Cugell, 1984). Eine andere Ursache der Limitation können aber auch „rechen-technische“ Beschränkungen in höheren Ebenen der visuellen Verarbeitungshierarchie sein. Studien betreffend der Geschwindigkeit der visuellen Wahrnehmung, vgl. beispielsweise (Sziklai, 1956), ergaben konsistente Schätzungen für die Kanalkapazität von ca. 40-50 bit/s. Diese Zahl kann als Maß für die Information genommen werden, die in den höheren, kognitiven Bereichen des visuellen Systems verarbeitet werden kann. Im Gegensatz hierzu ist die Datenrate, die von den Photorezeptoren aufgenommen wird, über 5×10^6 bit/s (Jacobson, 1951). Von diesem Standpunkt aus betrachtet, ist eine Datenkompression natürlich sinnvoll und notwendig.

BARLOW postulierte, daß die Art der Repräsentation der Umwelt von essentieller Bedeu-

⁴Dies ist äquivalent mit der Forderung, daß die Transinformation zwischen den Neuronen Null sein muß.

tung für biologische Organismen ist, um assoziativ zu lernen. Die Überlegungen basieren auf BAYES'SCHEN Lernregeln. Wenn der Organismus eine neue Assoziation zwischen zwei Ereignissen lernen soll, so bedarf es der Kenntnis gewisser a priori Wahrscheinlichkeiten. Denn ohne die Kenntnis der gegebenen Koinzidenzen zwischen den beiden Ereignissen läßt sich nicht sagen, ob nun ein Ereignis ein guter Prädiktor für das zweite ist. Da dies auch generell für die Assoziation vieler Ereignisse zutrifft, ist der Organismus gezwungen, eine hochdimensionale Wahrscheinlichkeitsdichte zu schätzen, und ist somit in genau der prekären Situation einer kombinatorischen Explosion, wie in Abschnitt 2.4 für die WDF natürlicher Bilder erläutert. Eine Umkodierung der Stimuli im Sinne eines faktoriellen Codes ermöglicht die Schätzung der Verbunddichte aufgrund der viel einfacher zu erhaltenden Randdichten, vergleiche Gl. (2.4).

Weitere Kriterien, nach denen sich ein biologischer Bildverarbeitungsmechanismus richten muß, sind beispielsweise Signal-zu-Rauschleistungsverhältnisse, oder die Minimierung des Energiebedarfs. Dies führt dann zum Begriff der *spärlichen Kodierung*, wie weiter unten detailliert beschrieben.

2.5.2 Optimale lineare Transformationsverfahren

Die Transformation redundanzbehafteter Signale in eine effiziente, statistisch unabhängige oder auch faktorielle Repräsentation ist nicht-trivial und hängt substantiell von den praktischen Anforderungen und dem daraus resultierenden Optimierungskriterium ab. Der allgemeine Fall einer (nicht-)linearen Transformation einer statistisch abhängigen Nachrichtenquelle \vec{x} mit VWDF $p(x_1, \dots, x_N)$ zum Erreichen der vollen statistischen Unabhängigkeit kann geschrieben werden als:

$$\vec{y} = \vec{f}(\vec{x}), \quad (2.7)$$

wobei $\vec{x} \in R^N$ die empfangene Nachricht, \vec{f} eine vektorwertige Funktion $R^N \rightarrow R^M$ und $\vec{y} \in R^M$ die transformierte Nachricht darstellt, für deren VWDF nun gilt $p(y_1, \dots, y_M) = p(y_1)p(y_2) \cdots p(y_M)$. Dieser allgemeine Fall erfordert im *worst-case* die Kenntnis der vollen Verbundstatistik, die jedoch praktisch nur selten zu erreichen ist. Eine in der Praxis häufig verwendete Vereinfachung besteht darin, die Klasse der Transformationen einzugrenzen, indem man sich auf lineare Abbildungen beschränkt.

Definition. Eine Abbildung $\vec{f}: R^N \rightarrow R^M$ heißt linear, wenn gilt:

1. \vec{f} ist homogen, d.h. $\vec{f}(\alpha\vec{x}) = \alpha\vec{f}(\vec{x})$ für alle $\alpha \in R, \vec{x} \in R^N$
2. \vec{f} ist additiv, d.h. $\vec{f}(\vec{u} + \vec{v}) = \vec{f}(\vec{u}) + \vec{f}(\vec{v})$ für alle $\vec{u}, \vec{v} \in R^N$.

□

Lineare Transformationen lassen sich dann mit Hilfe der Matrix-Notation folgendermaßen ausdrücken:

$$\vec{f}: R^N \rightarrow R^M, \quad \vec{f}(\vec{x}) = A\vec{x}, \quad (2.8)$$

oder einfach

$$\vec{y} = A\vec{x}. \quad (2.9)$$

Die Spalten der Matrix $B = A^{-1}$ bilden die Basisvektoren \vec{b}_i des transformierten Raumes. Im folgenden wird von dem wichtigen Fall der linearen Abbildung des R^N in sich ausgegangen, das bedeutet $A \in R^{N \times N}$. In diesem Fall spricht man von einer vollständigen linearen Transformation (die Basisvektoren bilden ein vollständiges Koordinatensystem), wenn gilt $\text{Rang}(A) = N$. Ein beobachtetes Signal \vec{x} läßt sich somit als Linearkombination der Basisvektoren \vec{b}_i , gewichtet mit den Koeffizienten y_i darstellen:

$$\vec{x} = \sum_{i=1}^N y_i \vec{b}_i = B\vec{y}. \quad (2.10)$$

Erfüllt A zudem die Bedingung der Orthogonalität, d.h. es gilt $A^T A = I_N$, wobei I_N die $N \times N$ -Einheitsmatrix bezeichnet, so spricht man von einer Orthogonaltransformation.

Karhunen-Loève-Transformation

Eine weitere Möglichkeit, die Komplexität des Problems einzuschränken, ist, nicht die volle Verbundstatistik zu berücksichtigen, sondern sich auf statistische Bindungen zweiter Ordnung (Korrelationen) zu beschränken und die Daten zu dekorrelieren, d.h. die Kovarianzmatrix zu diagonalisieren. Diese Lösung wird als Karhunen-Loève-Transformation (KLT), Hauptachsentransformation oder auch Principal Component Analysis (PCA) bezeichnet. Hierbei berechnen sich die Basisvektoren der Transformation (Spalten von B) als die Eigenvektoren der Kovarianzmatrix (Jayant und Noll, 1984). Diese ist definiert als der Erwartungswert über das dyadische Produkt der vom Mittelwert \vec{m} befreiten Zufallsvektoren:

$$C = E\{(\vec{x} - \vec{m})(\vec{x} - \vec{m})^T\}. \quad (2.11)$$

Eine wesentliche Eigenschaft der KLT ist, daß die ZGen nach der Transformation garantiert *dekorreliert* sind, d.h. die Kovarianzmatrix geht in der Transformation in eine Diagonalmatrix über:

$$\Lambda = B^T C B = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_N \end{pmatrix}. \quad (2.12)$$

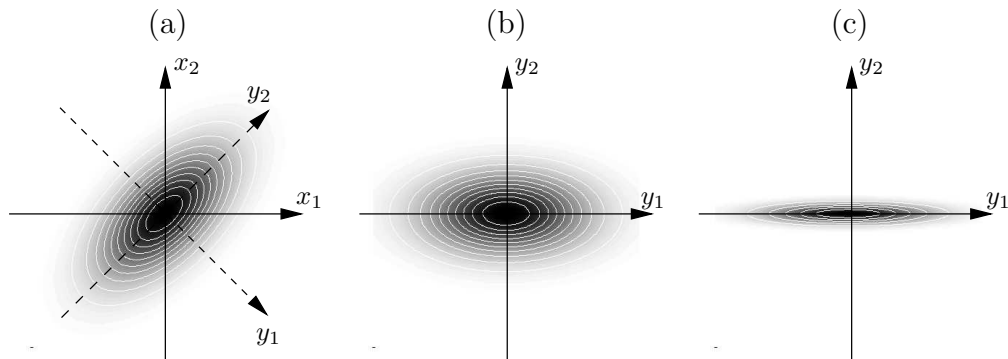


Abbildung 2.9: Karhunen-Loève-Transformation einer zweidimensionalen gaußschen Quelle. (a) Verbunddichte im Originalraum zusammen mit den Hauptachsen, (b) Verbunddichte im transformierten (Bild-) Raum. Im Falle stark korrelierter Größen können einzelne Merkmale (Dimensionen) ohne relevanten Rekonstruktionsfehler vernachlässigt werden (c).

Daran kann man erkennen, daß die Eigenwerte λ_i gleich den Varianzen $\sigma_{ii} := \sigma_i^2$ in den Randverteilungen der transformierten Achsen sind. Im Falle von gaußschen Verteilungen ist diese Transformation optimal, da diese vollständig durch statistische Bindungen zweiter Ordnung beschrieben werden und somit die Dekorrelation äquivalent zur statistischen Unabhängigkeit der transformierten Koeffizienten ist. Ein Beispiel für die Dekorrelation einer zweidimensionalen gaußschen Verbunddichte ist in Abb. 2.9 gezeigt. Eine weitere Eigenschaft der KLT ist, daß der mittlere quadratische Abbruchfehler minimiert wird, wenn zur Repräsentation des Signals nicht alle Freiheitsgrade (Dimensionen) genutzt werden, d.h. \vec{x} wird approximiert durch

$$\hat{\vec{x}} = \sum_{i=1}^r y_i \vec{b}_i \quad \text{mit } r < N. \quad (2.13)$$

Damit geht die Eigenschaft einher, daß das geometrische Mittel der Varianzen σ_i^2 minimiert wird:

$$\left[\prod_{i=1}^N \sigma_i^2 \right]^{\frac{1}{N}} \rightarrow \min. \quad (2.14)$$

Im Falle stark korrelierter Größen kann der Abbruchfehler sehr gering werden, wie in Abb. 2.9(c) illustriert. Die KLT ist daher geeignet, Quellen, deren Eigenschaften sich hinreichend genau durch statistische Momente zweiter Ordnung beschreiben lassen, in eine effiziente Repräsentation zu überführen. Für stationäre Prozesse ergibt die KLT eine Fourier-Basis. Dies ergibt sich aus deren FOURIER-STIELTJES Repräsentation als Summe unkorrelierter Frequenzkomponenten mit statistisch verteilter Amplitude und gleichverteilter Phase (Papoulis, 1991), also

$$E\{X(f_1)X^*(f_2)\} \equiv 0 \quad \forall f_1 \neq f_2. \quad (2.15)$$

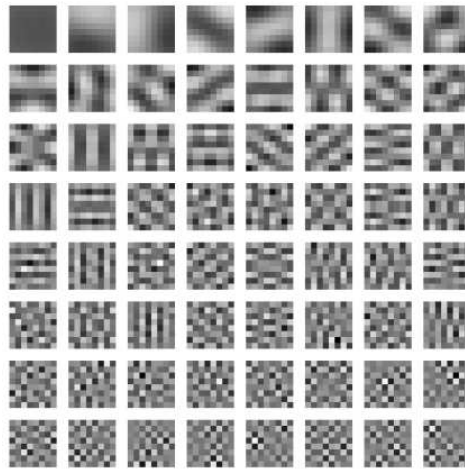


Abbildung 2.10: Hauptkomponenten natürlicher Bilder, berechnet auf der Basis von 8x8 Pixel Bildausschnitten.

Die Varianzen der transformierten Größen (Fourier-Komponenten) entsprechen hierbei genau den Einträgen im Leistungsdichtespektrum:

$$\sigma_i^2 = E\{X(f_i)X^*(f_i)\} = E\{|X(f_i^2)|\} = C_2^x(f_i). \quad (2.16)$$

Die Basisfunktionen, die sich für die Dekorrelation natürlicher Bilder ergeben, sind in Abb. 2.10 dargestellt (Olshausen und Field, 1996). Die Basisfunktionen sind entsprechend ihrer Varianz von links oben nach recht unten in absteigender Reihenfolge angeordnet. Dieses Ergebnis spiegelt auch das Leistungsdichtespektrum natürlicher Bilder wider (Abb. 2.8), da die Energie in den Komponenten mit niedriger Frequenz konzentriert ist. Eine Rekonstruktion entsprechend Gl. (2.13) ergibt Bilder mit fehlendem hochfrequentem Anteil, also Bilder, die unscharf erscheinen.

Independent Component Analysis

Wenn die Näherung der Quellenstatistik durch Momente zweiter Ordnung nicht mehr ausreicht, so bedarf es Verfahren, welche Statistiken höherer Ordnung berücksichtigen. Ein häufig zitierter Fall sind sogenannte spärliche, unabhängige Quellen (sparse independent sources, SIS). Die Annahme hierbei ist, daß die beobachteten Signale aus einer linearen Mixtur von nicht-gaußschen Quellen entstehen, die ihrerseits eine *spärliche* (s.u.) Verteilung aufweisen. Formal kann dies mithilfe eines statistischen Modells mit verborgenen Variablen definiert werden (Comon, 1994): Es seien $\mathbf{s}_1, \dots, \mathbf{s}_N$ statistisch unabhängige ZGen. Diese seien nicht direkt beobachtbar (daher verborgen), sondern es werden N lineare Überlagerungen $\mathbf{x}_1, \dots, \mathbf{x}_N$ beobachtet:

$$\mathbf{x}_j = a_{j1}\mathbf{s}_1 + a_{j2}\mathbf{s}_2 + \dots + a_{jN}\mathbf{s}_N; \quad j \leq 1 \leq N. \quad (2.17)$$

In Matrix-Vektor Notation läßt sich dies schreiben als

$$\vec{x} = A\vec{s}, \quad (2.18)$$

wobei A dann die sogenannte Mischungs-Matrix bezeichnet. Dieses statistische Modell wird *Independent Component Analysis* (ICA) Modell genannt (Hyvärinen et al., 2001). Die Aufgabe besteht nun darin, aus den beobachteten \mathbf{x}_j die Mischungs-Matrix A und die unabhängigen Komponenten \mathbf{s}_i zu bestimmen. Wenn A bekannt ist, so kann ihre Inverse W (Gewichts-Matrix) und über die Beziehung

$$\vec{s} = W\vec{x} \quad (2.19)$$

die unabhängigen Komponenten berechnet werden. Zur Lösung dieses Problems existieren viele unterschiedliche Implementierungsmöglichkeiten, siehe beispielsweise (Hurri, 1997; Hyvärinen und Oja, 2000; Fiori, 2003). Ein möglicher Ansatz hierbei ist, daß in einem erlaubten Subraum⁵ nach Achsen gesucht wird, deren Randverteilungen eine hohe *Kurtosis* aufweisen. Dieses Paradigma liegt begründet im zentralen Grenzwertsatz der Wahrscheinlichkeitstheorie: Hiernach ergibt die lineare Kombination vieler nicht-gaußscher unabhängiger Komponenten im Limit eine Gauß-Verteilung. Die gesuchten unabhängigen Komponenten sollten also möglichst weit von einer Gauß-Verteilung „entfernt“ sein. Dies ist typischerweise dann der Fall, wenn sie eine spärliche Verteilung aufweisen, d.h. im Vergleich zur Gauß-Verteilung eine Konzentration der Masse um die Null und längere Ausläufer besitzen. Als Spärlichkeitsmaß wird häufig die normalisierte Kurtosis

$$\kappa = \frac{E\{x^4\}}{E\{x^2\}^2} - 3 \quad (2.20)$$

verwendet. Als zentriertes und normalisiertes Moment vierter Ordnung nimmt dieses Maß große Werte an bei einer Verteilung mit einer Konzentration der Masse um die Null und wenigen relativ großen Werten. Der Vorteil dieser Kurtosis ist, daß sie als eine Art parameterunabhängige Formgröße verstanden werden kann. Beispielsweise ergibt sich für eine Laplace-Verteilung, gegeben durch

$$f^{(L)}(x) = \frac{1}{2b} e^{-|x-\mu|/b}, \quad (2.21)$$

eine parameterabhängige Varianz $\sigma^2 = 2b^2$ und eine parameterunabhängige Kurtosis $\kappa = 3$. Analoges gilt für eine Gauß-Verteilung, hier ist die Kurtosis identisch Null. Daher wird die Laplace-Verteilung auch als *supergaußsch* bezeichnet. Flache Verteilungen, wie beispielweise eine Gleichverteilung, besitzen negative Kurtosis und werden somit auch als *subgaußsch* bezeichnet.

Die ICA ist beispielhaft in Abb. 2.11 dargestellt: Die gezeigte Verteilung ist nicht korreliert, d.h. die KLT würde dieses Koordinatensystem mit der gezeigten Randverteilung bereits als Optimum sehen. Die wahre unabhängige Verbunddichte ergibt sich jedoch in einem um 45 Grad rotierten Raum, die Randverteilung hierfür ist spärlicher. In Abb.

⁵Eine mögliche Einschränkung ist Orthogonalität, dies ist jedoch nicht notwendig.

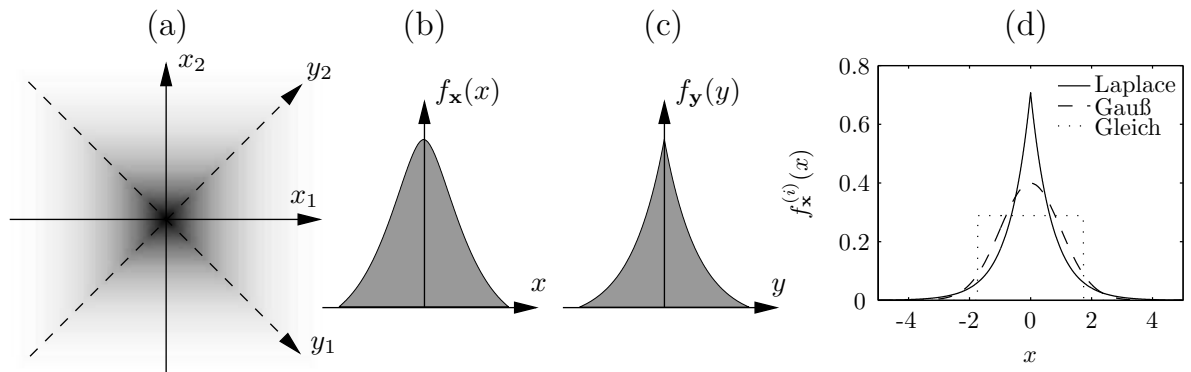


Abbildung 2.11: (a) Zweidimensionale Verbunddichte welche die lineare Mixtur zweier nicht-gaußscher Quellen beschreibt. Diese ist nicht korreliert, im statistischen Sinne zweiter Ordnung also bereits optimal. Beobachtete Randdichten (b) sowie Randdichten der unabhängigen Komponenten (c). (d) Laplace-, Gauß- und Gleichverteilung jeweils mit Varianz eins und Kurtosis $\kappa_{La} = 3, \kappa_{Ga} = 0, \kappa_{Gl} = -1, 2$.

2.11(d) sind die drei erwähnten Verteilungen, jeweils normiert auf Varianz eins, gezeigt.

Andere Ansätze minimieren die Transinformation zwischen den Komponenten oder verwenden die *Negentropie* (Comon, 1994; Bell und Sejnowski, 1997). Allen Implementierungsvarianten sind jedoch einige Aspekte gemein, die hier kurz aufgeführt werden:

- Der Begriff *Independent Component Analysis* ist irreführend, da die Unabhängigkeit der transformierten Komponenten nicht gewährleistet ist. Dies ist ein wesentlicher Unterschied zur KLT, da bei dieser das Optimierungskriterium (Dekorrelation) analytisch garantiert ist.
- Die ICA ist ein heuristisches Verfahren welches numerischer Optimierungsalgorithmen bedarf, es unterliegt also im Allgemeinen den prinzipiellen Konvergenz- und Stabilitätsproblemen numerischer Verfahren.
- Es können keine Varianzen (Energien) der unabhängigen Komponenten bestimmt werden. Dadurch, daß sowohl \vec{s} als auch A als unbekannt vorausgesetzt werden, kann jeder Skalierungsfaktor in den \mathbf{s}_j durch eine Division der zugehörigen Spalte \mathbf{a}_j mit dem gleichen Faktor ausgeglichen werden. Somit bleibt auch das Vorzeichen eine Unbekannte.
- Es kann keine dezidierte Reihenfolge der unabhängigen Komponenten bestimmt werden. Im ICA-Modell, Gl. (2.18), können die Terme beliebig permutiert werden, ohne daß sich etwas an der Beobachtung ändert. Formal kann dies mit einer Permutationsmatrix P ausgedrückt werden, indem man schreibt: $\vec{x} = AP^{-1}P\vec{s}$. Die Elemente von $P\vec{s}$ sind die originalen unabhängigen Variablen \mathbf{s}_j , allerdings in vertauschter Reihenfolge und die Matrix AP^{-1} ist eine neue, unbekannte Mischungs-Matrix.

- Die unabhängigen Komponenten dürfen nicht gaußsch sein. Diese Problematik ergibt sich aufgrund des häufig verwendeten Vorverarbeitungsschritts des *Whitenings*, was dann zu einer rotationssymmetrischen Verteilung führt und diese keinen Hinweis auf die Richtung der Spalten der Mischung-Matrix A geben kann (im Falle gaußscher Größen ist die KLT die optimale Transformation).

Nichtlineare Transformationen

Der allgemeine Fall der nichtlinearen Transformationen (illustriert in Abb. 2.12) ist mit derartigen expliziten Ansätzen nur in Sonderfällen möglich, vor allem, wenn das Problem durch geeignete a priori Annahmen in nur wenigen Dimensionen formuliert werden kann. Auf zunehmendes Interesse stoßen nichtlineare Varianten der PCA. Diese gehen zurück auf die Pionierarbeiten von OJA (Oja, 1982), welche eine Vielzahl unterschiedlicher Algorithmen für neuronale Netze zur Berechnung von Hauptkomponenten initiiert haben. Verglichen mit dem Standardansatz haben derartige Implementierungen gewisse Vorteile, v.a. Adaptationsfähigkeit bei nichtstati-

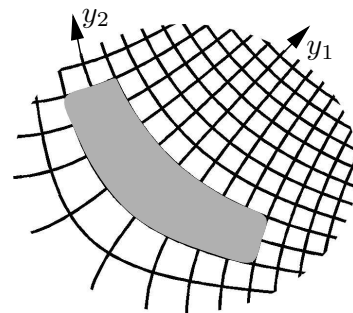


Abbildung 2.12: Beispiel einer nichtlinearen Koordinatentransformation.

onären Daten. Nichtlineare Varianten dieser Algorithmen können durch das Hinzufügen nichtlinearer Aktivierungsfunktionen erhalten werden. Weitere Ansätze verwenden Approximationen, um die nichtlinearen Zusammenhänge der Daten zu erfassen, beispielsweise lokal lineare PCA. Für eine ausführliche Zusammenstellung der aktuellen Algorithmen siehe (Diamantaras und Kung, 1996). Ein relativ neues Verfahren ist die sogenannte *kernel PCA* (Schölkopf et al., 1998). Hierbei werden die Hauptkomponenten nicht im Originalraum, sondern in einem hochdimensionalen Merkmalsraum berechnet, typischerweise der Raum aller Monome im R^N . Die kombinatorische Explosion der Dimensionalität verbietet es, mit herkömmlichen Varianten der PCA zu arbeiten. Statt dessen kann unter gewissen Annahmen (der Raum, den die Monome aufspannen, muß gewisse Eigenschaften besitzen) die Dimensionalität proportional zur Anzahl der beobachteten Stichproben gemacht werden. Dies ist dann möglich, wenn der Algorithmus im Merkmalsraum nur durch Verwendung von Skalarprodukten ausgedrückt werden kann (wie dies bei der PCA der Fall ist). Diese müssen dann nicht explizit berechnet werden, sondern können im Originalraum als Funktionsevaluationen dargestellt werden (Vapnik, 1995; Schölkopf und Smola, 2002). Die Proportionalität der Dimension zur Anzahl der erlangten Daten beinhaltet jedoch ihrerseits Schwierigkeiten, d.h. komplexe Probleme, die viele Daten für eine konvergente Schätzung benötigen, bedürfen weiterer Vereinfachungen. Eine Lösung ist die Verwendung iterativer Algorithmen, dies ist jedoch noch „ongoing research“ (Kim et al., 2003).

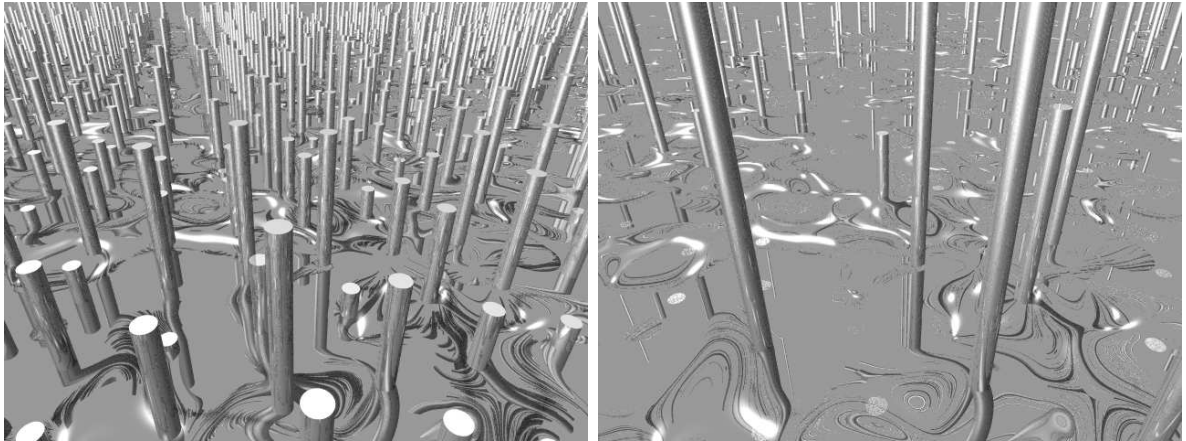


Abbildung 2.13: Illustration einer „Neuronenlandschaft“ für eine verteilte Entropie-Kodierung (links) und eine spärliche Kodierung (rechts). Jeder Zylinder soll ein Neuron und seine Länge das entsprechende postsynaptische Potential zu einem bestimmten Zeitpunkt darstellen.

Kapitel 4 bzw. 5 dieser Arbeit beschäftigt sich mit nichtlinearen Transformationen unter Verwendung von mehrschichtigen neuronalen Systemen, bzw. polynomialer Filterung.

2.5.3 Spärliche Kodierung

Der Begriff *spärliche Repräsentation* beschreibt eine Art der Kodierung, bei welcher die einzelnen Einheiten (Neurone) ein hochgradig selektives Antwortverhalten aufweisen. Im Gegensatz zu einer reinen Entropie-Kodierung, bei der im Limit vom gesamten Dynamikbereich gleichermaßen Gebrauch gemacht werden sollte, sind bei dieser Repräsentation jeweils nur sehr wenige Einheiten aktiv, diese jedoch mit großer Amplitude (vgl. Abb. 2.13). Hier gelangt man an einen Punkt, wo in der Literatur häufig falsche Assoziationen zu finden sind. Daher hier nun eine etwas ausführlichere Betrachtung dieses Themas.

Spärliche Kodierung in *einzelnen* Neuronen entspricht einer hohen Kurtosis der jeweiligen neuronalen Aktivitätsverteilung (WDF). Betrachtet man nun eine ganze Population derartiger spärlicher Neurone, so ergäbe sich in etwa ein Bild wie in Abb. 2.13 rechts illustriert. Dabei ist noch nichts über etwaige interneuronale statistische Abhängigkeiten gesagt. Annahmen hierüber werden häufig implizit gemacht, was dann zu einer Unklarheit des Begriffs spärliche Kodierung führt. So wird beispielsweise bei den generativen ICA-Modellen für natürliche Bilder angenommen, daß diese sich als eine lineare Superposition von (nicht notwendigerweise orthogonalen) Basisfunktionen $\Phi_i(x, y)$ modellieren lassen:

$$I(x, y) = \sum_{(i)} a_i \Phi_i(x, y). \quad (2.22)$$

Desweiteren wird angenommen, daß diese elementaren Bildmerkmale ihrerseits eine spärliche Verteilung aufweisen und voneinander statistisch unabhängig sind⁶. Die Optimallösung der ICA findet nun diese unabhängigen Komponenten, womit die resultierende Aktivitätsverteilung der Neurone eine spärliche und unabhängige Statistik aufweist. Ein häufig als einfache Erweiterung bezeichneter, jedoch konzeptuell deutlich unterschiedlicher Ansatz ist die spärliche und *überevullständige* Kodierung. Hierbei werden mehr Basisvektoren mit hoher Kurtosis gesucht, als für eine eindeutige Repräsentation notwendig wären (Olshausen und Field, 1997). Dies resultiert in einer nicht-orthogonalen Basis, in der dann selbstverständlich statistische Abhängigkeiten zwischen den Einheiten bestehen. Der Grenzfall dieses Konzepts wäre mit einer Art Vektor-Quantisierer vergleichbar, und jede Einheit würde genau einen bestimmten Stimulus repräsentieren. Diese Repräsentation weist eine hohe Redundanz auf, da bei Beobachtung eines aktiven Neurons mit Gewißheit die anderen nicht aktiv sind. Dies ist aus Sicht der klassischen Entropiekodierung suboptimal, jedoch müssen bei der Analyse der Funktionsweise des visuellen Kortex andere, nämlich biologisch relevante, Optimierungskriterien mit in Betracht gezogen werden. So kann beispielsweise gezeigt werden, daß eine derartige Kodierung die Effizienz von Assoziativspeichern erhöht (Zetzsche, 1990) und den Verdrahtungsaufwand minimiert (Foldiak, 1995). Desweiteren wird hierbei relativ viel Information mit nur wenigen Aktionspotentialen (Spikes) übertragen, was in einem biologischen Organismus wesentliche metabolische Vorteile mit sich bringt (Baddeley, 1996). Eine überevullständige, spärliche Repräsentation erhöht auch das SNR in rauschbehafteten Systemen, wo Signale einer Degradation unterliegen und ist somit für eine zuverlässige Mustererkennung von großem Vorteil. Man kann also sagen, daß in einem biologischen Organismus die reine Eliminierung der Redundanz nicht das höchste Ziel ist, sondern Ziel ist, die in den Strukturen der Umwelt gegebenen impliziten Redundanzen in eine dem Organismus zugängliche, explizite Form zu transformieren.

Einen physiologischen Nachweis, daß der visuelle Kortex eine derartige Strategie bei der Kodierung natürlicher Bilder verfolgt, liefert (Vinje und Gallant, 2000). Darin wurde gezeigt, daß Neurone im primären visuellen Kortex von Affen bei der Betrachtung natürlicher Bilder eine hohe Spärlichkeit aufweisen. Vergleiche hierzu auch Abb. 2.15 und die zugehörige Diskussion.

OLSHAUSEN und FIELD implementierten einen Algorithmus, der nach Basisfunktionen sucht, die entsprechend Gl. (2.22) natürliche Bilder möglichst gut beschreiben, gleichzeitig aber ein spärliches Antwortverhalten aufweisen (Olshausen und Field, 1996). Dies wurde in einem neuronalen Netz realisiert, welches die folgende Kostenfunktion minimiert:

$$E(a, \Phi) = \sum_{(x,y)} [I(x, y) - \sum_{(i)} a_i \Phi_i(x, y)]^2 + \beta \sum_{(i)} S\left(\frac{a_i}{\sigma_i}\right), \quad (2.23)$$

wobei $\sigma_i^2 = E\{a_i^2\}$. Der erste Term mißt den Rekonstruktionsfehler, also wie gut der

⁶Dies ist dann eine a priori Eigenschaft der Natur.

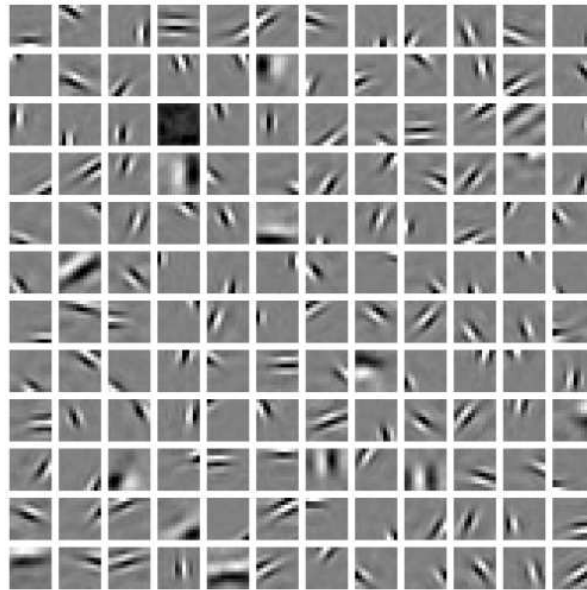


Abbildung 2.14: Unabhängige Komponenten natürlicher Bilder, berechnet auf der Basis von 12x12 Pixel Bildausschnitten (Olshausen und Field, 1996).

Kode natürliche Bilder beschreibt, und der zweite Term bestraft die Aktivität der einzelnen Einheiten, so daß eben der Zustand weniger aktiver Einheiten pro Realisierung favorisiert wird. Das Ergebnis dieser Optimierung ist in Abb. 2.14 gezeigt. Im Gegensatz zur Fourier-Basis, die unter Verwendung der KLT gefunden wird (Abb. 2.10), sind diese Basisfunktionen lokalisiert. Zudem sind sie orientiert und frequenzselektiv. Auch andere Ansätze aus der Klasse der ICA-Algorithmen liefern ähnliche Ergebnisse (Bell und Sejnowski, 1997; van Hateren und van der Schaaf, 1998).

Der Zusammenhang zwischen spärlicher Kodierung, übervollständiger Repräsentation und einer nichtlinearen Partitionierung des Signalraums wird in Kapitel 3 diskutiert.

2.6 Der klassische Ansatz zur Erklärung der Funktionsweise des visuellen Systems

Es ist bekannt, daß die Repräsentation natürlicher Stimuli im visuellen System zunehmend spärlicher wird und damit die Einzelentropien stetig abnehmen (Abb. 2.15). Es folgt nun eine Beschreibung des klassischen Ansatzes, der in der Vergangenheit einige wesentliche Aspekte dieser Informationsverarbeitung erklären konnte. Dieser zerfällt in zwei Untergruppen: Zum Einen die Analyse der Eingangssignale unter Verwendung von Statistiken zweiter Ordnung (Korrelationsfunktion und Leistungsdichtespektrum, vgl. Abschnitt 2.4.3) sowie zum Anderen die Synthese bzw. Modellierung mittels li-

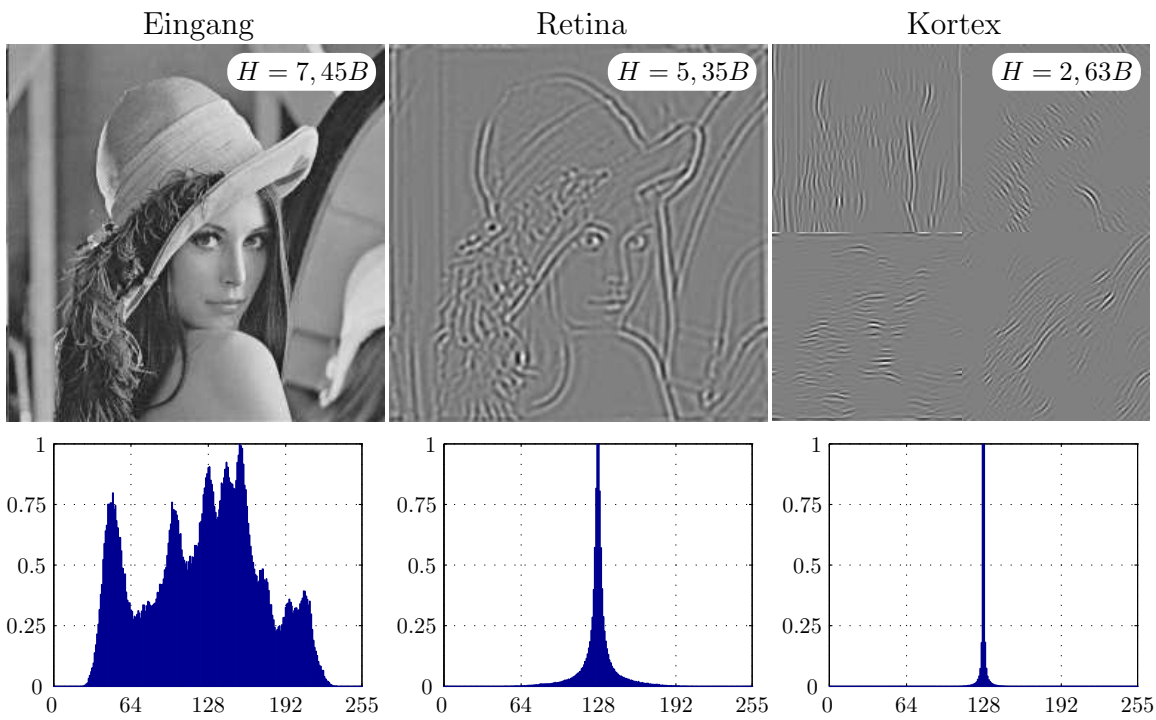


Abbildung 2.15: Die Transformation der Statistik natürlicher Bilder auf unterschiedlichen Ebenen des visuellen Systems. Links: Aktivitätsverteilung des Eingangs auf Ebene der Photorezeptoren. Mitte: Eine isotrope Bandpaß-Filterung (laterale Inhibition der Ganglienzellen) reduziert die Entropie und führt zu einer spärlichen Verteilung. Rechts: Eine weitere orientierungselektive Zerlegung (Simple-Zellen im visuellen Kortex) verstärkt diesen Effekt noch.

nearen Systemen. Es wird nun gezeigt, welche Aspekte der Verarbeitung mit diesem Ansatz widerspruchsfrei erklärt werden können. Dabei ist das Augenmerk einerseits auf die Modellierung bekannter Eigenschaften und andererseits auf die Untersuchung der Optimalität aus Sicht der Statistiken zweiter Ordnung gerichtet. Aspekte für die dieser Ansatz ungenügend ist, werden im nächsten Abschnitt erläutert.

2.6.1 Spatiotemporale Tiefpaß-Filterung in der Retina

Bereits in den 70er Jahren wurde durch Untersuchungen an Schildkröten entdeckt, daß die Photorezeptoren der Retina nicht als isolierte Sensoren arbeiten, sondern miteinander verschaltet sind (DeValois und DeValois, 1990). Diese elektrotonische Kopplung verursacht funktional eine örtliche Tiefpaß-Filterung. Ebenso ist bekannt, daß die Photorezeptoren nicht instantan auf eintreffende Photonen reagieren, sondern daß hier über einen Bereich von mehreren Millisekunden integriert wird. Zusammen ergibt sich im systemtheoretischen Sinn also eine spatiotemporale Tiefpaß-Filterung, die das Eingangssignal zu hohen Frequenzen hin stärker dämpft. Unter dem Gesichtspunkt der Signal-zu-Rauschleistung (SNR) ist dies sinnvoll, da im Bereich der hohen Frequenzen stärkere Quantenfluktuationen relativ zum Signal herrschen. Das ist eine Strategie, die das lineare *Wiener-Filter* verfolgt und das auf der Basis von Korrelationen zweiter Ordnung ableitbar ist. Somit liegt dieser Effekt innerhalb des Erklärungshorizonts des klassischen Ansatzes. Im Bereich der Fovea ist die Anzahl der miteinander gekoppelten Rezeptoren geringer als im peripheren Bereich. Hier wird das „Empfängerrauschen“ minimiert, indem die vor den Rezeptoren liegenden Versorgungsgefäße zur Seite gedrängt sind und somit keine Photonen absorbieren. Interessant ist auch, daß der Grad der örtlichen Verschaltung mit zunehmender Umgebungsleuchtdichte abnimmt, d.h. die Grenzfrequenz des Tiefpasses wird erhöht. Auch dies ist aus der Sicht des SNR ein Positivum, da bei hoher Leuchtdichte das SNR auch in höheren Frequenzbereichen zunimmt. Man könnte also durchaus sagen, die Retina ist ein adaptives Wiener-Filter.

2.6.2 Laterale Inhibition der Ganglienzellen

Die Bipolarzellen erhalten ihre Speisung von benachbarten Rezeptoren, dies ist der sogenannte *R-B-Kreislauf*. Zudem erhalten sie noch Afferenzen von den Horizontal-Zellen, welche selbst wiederum von einem weiteren Rezeptorbereich gespeist werden. Dies wird der *indirekte R-H-B-Kreislauf* genannt. Die Gesamtheit aller direkt oder indirekt empfangener Rezeptoren ist dann das rezeptive Feld (vgl. Kap. 1). Der direkte und der indirekte Weg wirken auf die Bipolarzelle mit unterschiedlichen Vorzeichen und unterschiedlicher Amplitude. Dieses Prinzip der *lateralen Inhibition* ist in Abb. 2.16 als Blockschaltbild zusammen mit dem resultierenden isotropen rezeptiven Feld gegeben. Das

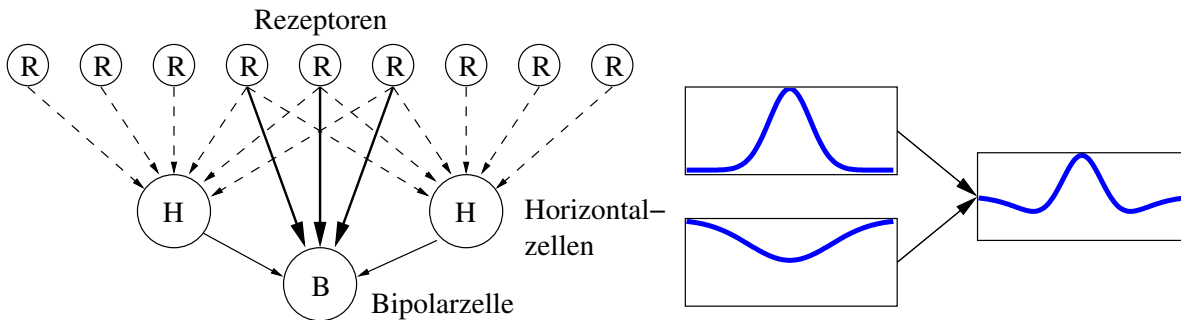


Abbildung 2.16: Links: Direkter und indirekter Weg des R-H-B-Kreislaufs. Rechts: Die Überlagerung der Verschaltungen unterschiedlicher Polarität resultiert in einem differenzierenden rezeptiven Feld.

Gewicht mit dem ein Rezeptor auf eine Bipolarzelle Einfluß nimmt, hängt von deren Distanz relativ zum Rezeptor ab. Diese Gewichtungsfunktion ist näherungsweise gaußsch, in der Überlagerung der exzitatorischen und inhibitorischen Anteile ergibt sich somit ein sogenanntes *Mexican-Hat-Profil*, das einem differenzierenden Element entspricht. Aus signaltheoretischer Sicht korrespondiert dieses Verhalten der retinalen Ganglienzellen mit einem linearen Übertragungssystem mit Bandpaß-Charakter.

Wie bereits in Abschnitt 2.5 besprochen, besteht der erste Schritt einer optimalen Signalverarbeitung darin, die statistischen Abhängigkeiten der Quellensymbole zu eliminieren. Aus Sicht der Abhängigkeiten zweiter Ordnung bedeutet dies, das Signal zu dekorrelieren. Diese Forderung ist äquivalent mit dem Erreichen eines möglichst konstanten (weißen) Leistungsdichtespektrums. Aus der Theorie der linearen Systeme ist bekannt, daß zwischen dem LDS des Eingangs $C_2^{U_1}(\vec{f})$, der Systemübertragungsfunktion $H(\vec{f})$ und dem LDS des Ausgangs $C_2^{U_2}(\vec{f})$ folgender Zusammenhang besteht:

$$C_2^{U_2}(\vec{f}) = |H(\vec{f})|^2 \cdot C_2^{U_1}(\vec{f}). \quad (2.24)$$

Aus dieser Gleichung kann leicht das inverse Filter für eine Äquivalenzklasse bzgl. der statistischen Abhängigkeiten zweiter Ordnung abgeleitet werden. Da das Ausgangs-LDS mit einem inversen Filter näherungsweise konstant wird, wird dieses auch häufig als *Whitening-Filter* bezeichnet.

In Abb. 2.17 ist der Effekt der lateralen Inhibition auf das Leistungsdichtespektrum natürlicher Bilder gezeigt. Es ist evident, daß dieser Mechanismus keine vollständige Dekorrelation erreicht, da das Spektrum nicht konstant ist. Es muß allerdings angemerkt werden, daß eine derartige Lösung unter SNR Gesichtspunkten suboptimal wäre, da in den hochfrequenten Bereichen relativ gesehen mehr Rauschanteile vorhanden sind. Eine Optimallösung besitzt daher einen bandpaß-artigen Charakter, da somit ein optimaler Kompromiß zwischen Dekorrelation und Signal-zu-Rauschleistung erreicht wird.

Eine experimentelle Überprüfung dieser Dekorrelations-Hypothese wurde in (Atick, 1992)

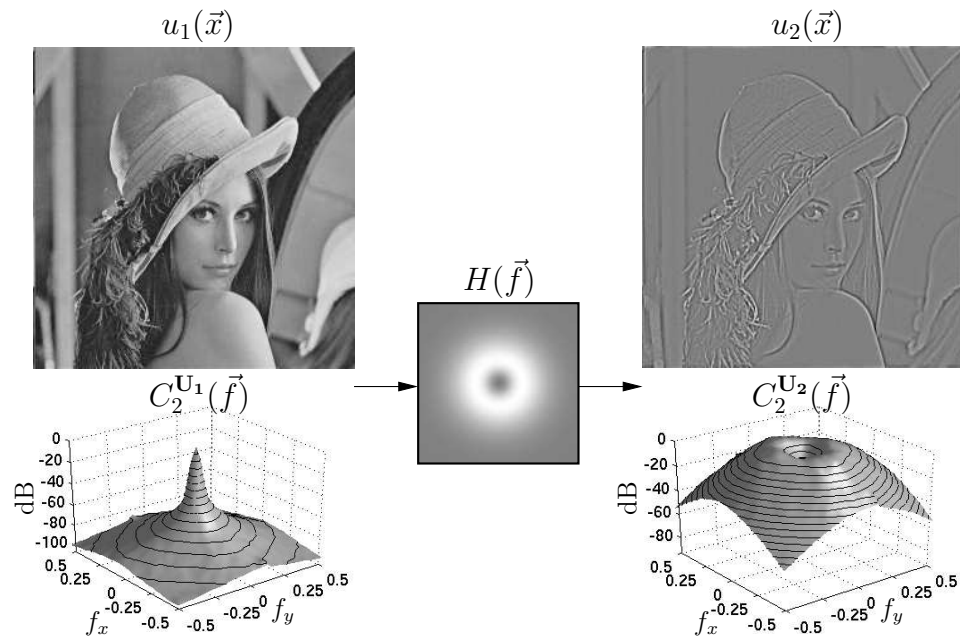


Abbildung 2.17: Effekt der lateralen Inhibition auf das Leistungsdichtespektrum natürlicher Bilder. Als Modell wurde ein *Difference-of-Gaussian* (DOG) Filter verwendet, ein Standardmodell der lateralen Inhibition, vgl. auch Abb. 2.16.

durchgeführt. Hier wurde unter den Gesichtspunkten SNR und Dekorrelation gezeigt, daß die hypothetisch optimale retinale Signalverarbeitung in der Gegenwart von Rauschen in der Tat mit der lateralen Inhibition vereinbar ist, und daß diese mit der empirisch gemessenen Kontrastsensitivitätsfunktion in Mackaken und Katzen übereinstimmt. Ein ähnlicher Test bezüglich der temporalen Dekorrelation findet sich in (Dan et al., 1996). Es konnte gezeigt werden, daß das temporale LDS, gemessen im CGL, für natürliche Bildsequenzen weiß wird. Unter Einbettung neurophysiologischer Untersuchungen konnte dann in (Stanley et al., 1999) ein lineares Modell der Verarbeitung im CGL aufgestellt werden, mit dem es möglich war, die von einer Katze fixierten bewegten Objekte aufgrund elektrophysiologischer Ableitungen zu rekonstruieren. Damit ist das Phänomen der lateralen Inhibition sowohl optimal im Sinne einer statistischen Optimierung zweiter Ordnung, als auch exakt modellierbar unter Verwendung linearer Systeme.

2.6.3 Selbstähnliche und lokalisierte rezeptive Felder

Rezeptive Felder im visuellen Kortex weisen zwei distinguierte Eigenschaften auf. Zum Einen sind sie selbstähnlich, d.h. durch Skalierung und Rotation ineinander überführbar. Eine Konsequenz dieser Selbstähnlichkeit ist die konstante Bandbreite auf einer logarithmischen Skala. Zum Anderen sind sie lokalisiert, d.h. die örtliche Ausdehnung der

rezeptiven Felder ist begrenzt, siehe Abb. 2.18. Dies steht im Gegensatz zu frühen Annahmen, daß das visuelle System eine reine Fourier-Analyse durchführt, da die hierbei zu verwendenden Basisfunktionen (komplexe Exponentialschwingungen) theoretisch eine unendliche Ausdehnung besitzen. Nach der von SHANNON entwickelten *Rate-Distortion-Theorie* (Berger, 1979) kann der Kodiergewinn des bestmöglichen Kodiersystems durch die *Spectral Flatness Measure*

$$\gamma_x^2 = \frac{\exp \left[\frac{1}{2\pi} \int_{-\pi}^{\pi} \ln(C_2^{\mathbf{U}_1}(e^{j\omega})) d\omega \right]}{\frac{1}{2\pi} \int_{-\pi}^{\pi} C_2^{\mathbf{U}_1}(e^{j\omega}) d\omega} \quad (2.25)$$

bestimmt werden. Unter realistischen Bedingungen mit endlichem SNR und begrenztem Dekodierpotential nachfolgender Verarbeitungsinstanzen ergibt sich aus der R-D-Theorie eine Aufspaltung in mehrere Kanäle, siehe auch (Zetzsche und Krieger, 2001b). Für die Selbstähnlichkeit der Kanäle findet sich ohne Weiteres keine explizite Erklärung. Eine Möglichkeit ist der $1/f^2$ -Abfall des LDS natürlicher Bilder sowie die Verwandtschaft zu selbstähnlichen fraktalen Prozessen. Durch den $1/f^2$ -Abfall ergibt sich bei einer logarithmischen Bandbreitenaufteilung der einzelnen Kanäle eine konstante Leistung in jedem Kanal, was für das Optimalitätskriterium aus Gl. (2.25) günstig ist. Anhand einer detaillierten Rate-Distortion Analyse konnte in (Röhrbein und Zetzsche, 1998) gezeigt werden, daß das tatsächliche Optimum für natürliche Bilder kleine, aber dennoch signifikante systematische Abweichungen hin zu einer Abnahme der Bandbreiten bei höheren Frequenzen aufweist.

Eine Möglichkeit zur mathematischen Modellierung derartiger rezeptiver Felder wurde in (Marčelja, 1980) vorgeschlagen. Dabei wurden Funktionen der Form

$$f(x) = \sin(2\pi f_0 x + \theta) e^{-x^2/2\sigma^2} \quad (2.26)$$

betrachtet. Diese wurden erstmals in GABORS „Theory of Communication“ zur Analyse zeitveränderlicher Funktionen etabliert (Gabor, 1946). MARČELJA zeigte einige wesentliche Gemeinsamkeiten dieser Funktionen mit den rezeptiven Feldern kortikaler Neuronen auf. Diese Funktionsgruppe wird nun gemeinhin als *Gabor-Funktionen* bezeichnet und dient in vielen Arbeiten als basales Modell für kortikale Zellen. Frühe experimentelle Überprüfungen zeigten, daß diese Funktionen in der Tat eine exzellente Übereinstimmung mit gemessenen rezeptiven Feldern aufweisen (Jones und Palmer, 1987a). DAUGMAN verallgemeinerten das bisherige Modell auf zweidimensionale Gabor-Funktionen, diese sind dann beschrieben als Produkt eines Sinus mit einer zweidimensionalen Gaußfunktion (Abb. 2.19). Auch diese weisen eine hohe Ähnlichkeit mit der zweidimensionalen Struktur rezeptiver Felder kortikaler Simple-Zellen auf (Jones und Palmer, 1987a). Es gibt noch weitere Modelle, die teilweise eine bessere Übereinstimmung aufweisen, jedoch sind die Unterschiede gering. Die relevanten Hauptmerkmale, Lokalisation im Ort und der Frequenz sowie Selbstähnlichkeit, sind allen Modellen gemein. Ein weiterer

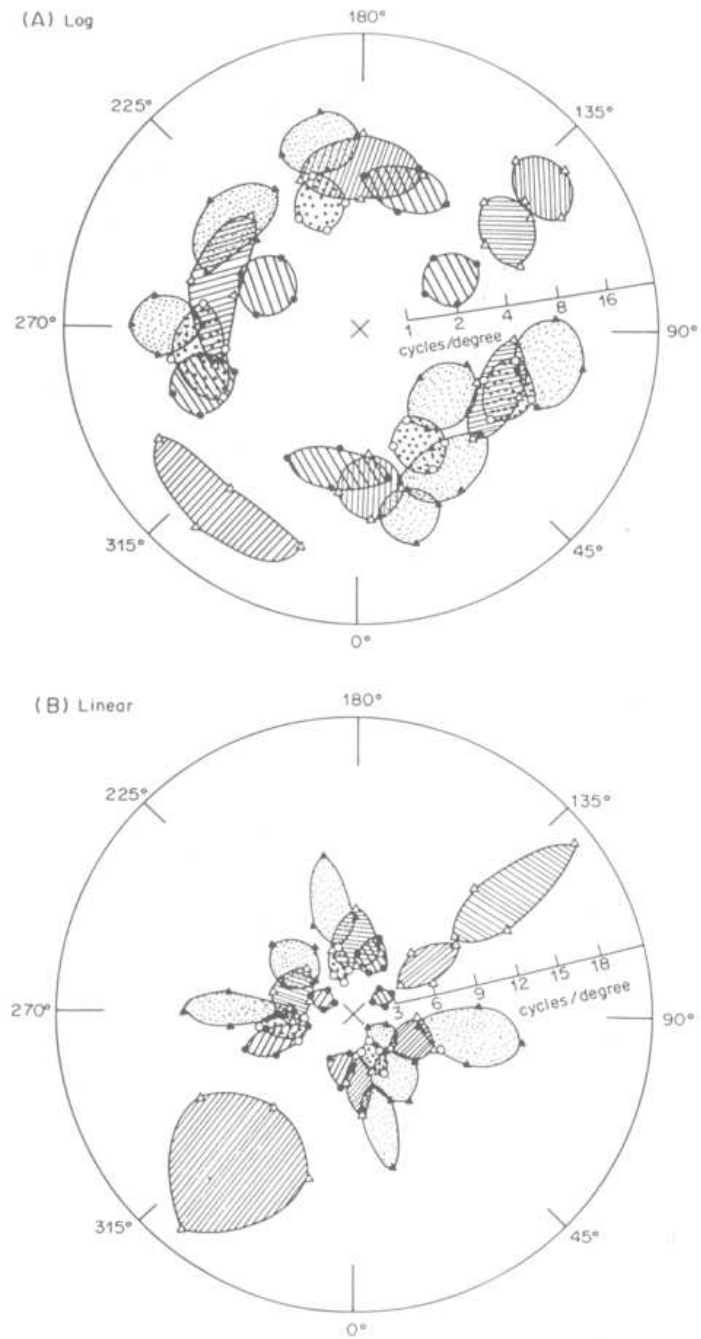


Abbildung 2.18: Orientierungs- und Radialfrequenzselektivität kortikaler Simple-Zellen, aus (DeValois und DeValois, 1990). (A) Logarithmische Skalierung, (B) lineare Skalierung.

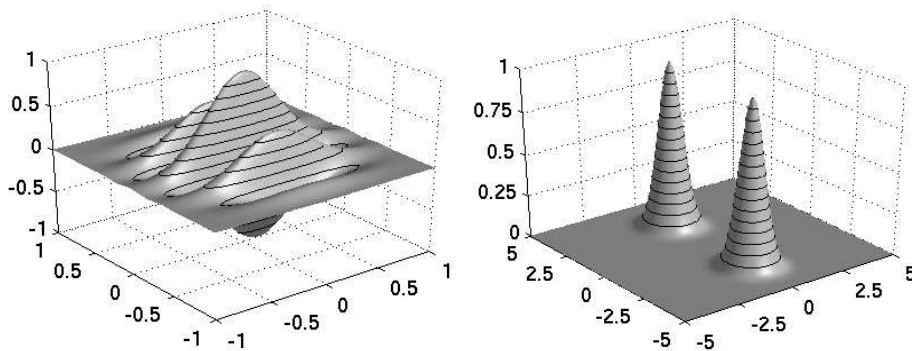


Abbildung 2.19: Ein geradesymmetrischer Vertreter der Familie der 2D-Gabor Funktionen (links) und seine Fourier-Transformierte (rechts).

Aspekt ist die Orientierungsselektivität, diese wird in Abschnitt 2.7 behandelt. Mittlerweile erlangten derlei Funktionen große Popularität im Zuge der Wavelet-Analyse oder Wavelet-Transformation.

Damit kann also auch die Existenz von selbstähnlichen und lokalisierten rezeptiven Feldern und mehreren visuellen Kanälen als erklärbar im Sinne zweiter Ordnung betrachtet werden; ein vereinheitlichendes Modell ist durch das Mehrkanalkonzept der visuellen Signalverarbeitung gegeben, siehe unten und (Hauske, 1994).

2.6.4 Das Mehrkanalmodell

Die Theorie einer visuellen Verarbeitung mit mehreren (un-)abhängigen Kanälen wurde bereits Ende der 60er Jahre etabliert. Bis dahin dominierte die Idee, daß ein einziger breitbandiger Mechanismus, beobachtbar als Kontrastsensitivitätsfunktion, das frühe örtliche Sehen bestimmt. Erstmals wurde von CAMPBELL und ROBSON die Vermutung angestellt, daß das visuelle System so etwas wie Gruppen von unabhängigen, quasilinearen Bandpaßfiltern enthält, welche in einer von außen beobachtbaren Einhüllenden, der Kontrastsensitivitätsfunktion, resultieren (Campbell und Robson, 1968). Abgesehen von den oben ausgeführten theoretischen Überlegungen die für eine derartige Modellstruktur sprechen, erlangte dieses Konzept primär dadurch Popularität, daß in vielen psychophysischen Experimenten diese Hypothese bestätigt wurde. Die Paradigmen hierbei

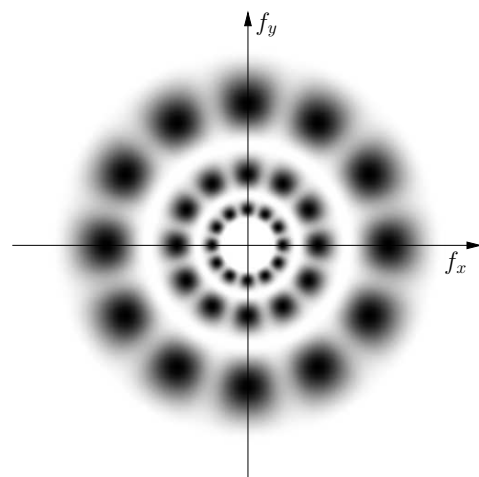


Abbildung 2.20: Partitionierung der Frequenzebene durch das Mehrkanalmodell.

waren vor allem die Detektion von Mustern auf der Basis des Amplitudenspektrums (Graham und Nachmias, 1971), die frequenzspezifische Adaptation (Pantle und Sekuler, 1968; Blakemore und Campbell, 1969), frequenzspezifische Nacheffekte (Blakemore und Sutton, 1969) und die Untersuchung von Maskierungseffekten, die auf Interaktionen der Kanäle schließen lassen (Carter und Henning, 1971). Eine typische Aufteilung der Frequenzebene in mehrere Kanäle ist schematisch in Abb. 2.20 gezeigt. Typische, d.h. physiologisch plausible Parameter der einzelnen Kanäle sind Radialbandbreiten von ca. ein bis zwei Oktaven und Orientierungsbandbreiten von 30 bis 40 Grad. Für einen Vergleich der Ergebnisse von neurophysiologischen Messungen unterschiedlicher Labors und Techniken siehe (DeValois und DeValois, 1990).

2.7 Grenzen des Standardansatzes

Wie anhand der vorangegangenen Abschnitte ausführlich gezeigt wurde, können wesentliche Aspekte des visuellen Systems unter Verwendung des klassischen Ansatzes, d.h. Statistiken zweiter Ordnung und lineare Systeme, hinreichend genau erklärt, bzw. beschrieben werden. Damit ist die Komplexität für die betrachteten Teilsysteme und -phänomene ebenfalls angemessen gewählt, ganz im Sinne WILLIAM VON OCKHAMS: *Entia non sunt multiplicanda praeter necessitatem*⁷. Dessen Postulat zur Denkökonomie verlangt, daß die Vielfalt und Komplexität von Denkansätzen nicht ohne Grund erhöht wird. Wenn jedoch Defizite vorhanden sind, die sich darin äußern, daß Sonderregeln und Zusätze verwendet werden müssen, damit Modelle mit der Wirklichkeit hinreichend genau übereinstimmen, kann unter Umständen ein komplexerer Modellansatz die Wirklichkeit besser erfassen und abbilden. In diesem Abschnitt soll daher eruiert werden, welche bekannten Attribute des visuellen Systems definitiv nicht innerhalb des Erklärungshorizonts des Standardansatzes liegen. Explizit sind dies:

1. Orientierungsselektivität kortikaler Neurone,
2. statistische Eigenschaften von gekrümmten Bildmerkmalen,
3. Phaseninvarianz kortikaler Complex-Zellen,
4. Endinhibition, oder allgemein Selektivität für $i2D$ -Signale,
5. kortikale Verstärkungsregelung,
6. statistische Abhängigkeiten *unabhängiger* Komponenten.

Ad 1. Wie bereits eingangs des letzten Abschnitts in Abb. 2.15 gezeigt wurde, erhöht eine orientierungsselektive Zerlegung des Eingangssignals die Spärlichkeit der Aktivitätsverteilung. Nichtsdestoweniger ist dies aus Sicht des klassischen Ansatzes nicht vorteil-

⁷Man nennt dieses Prinzip *Ockhamsches Rasiermesser*, weil es dazu dient Platons Bart abzuschneiden.

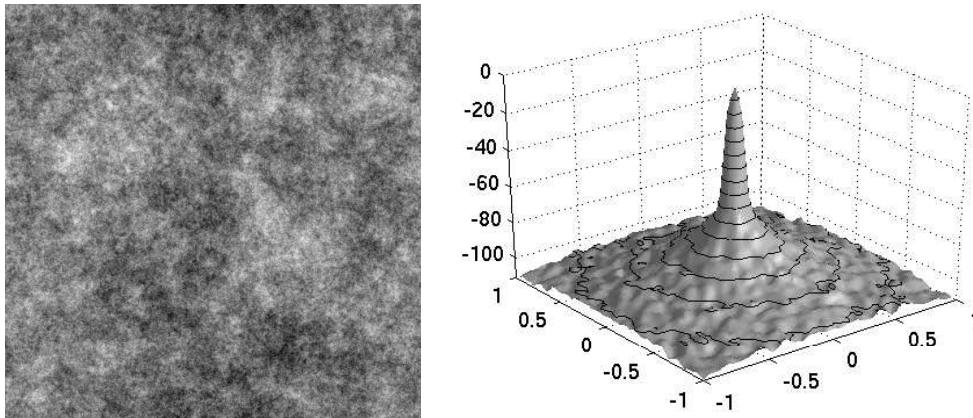


Abbildung 2.21: Ein gaußverteilter Prozeß (links) mit dazugehöriger spektraler Leistungsdichte, die nahezu identisch zu der natürlicher Bilder ist (rechts).

haft, hier gilt als relevantes Kriterium die SFM aus Gl. (2.25). Wie beispielsweise in (Zetzsche und Krieger, 1999) gezeigt wurde, ist aus Sicht der Statistik zweiter Ordnung eine orientierungsselektive Zerlegung sogar inferior gegenüber einer rein radialen Aufspaltung des Signals. Die Orientierungsselektivität ist jedoch notwendig für die Dekorrelation höherer Ordnung, siehe hierzu Abschnitt 5.2.6.

Ad 2. Die statistischen Eigenschaften zweiter Ordnung natürlicher Bilder sind mittlerweile hinreichend bekannt (siehe Abschnitt 2.4). Doch inwieweit sind diese auch ausreichend zur Beschreibung? Das in Abb. 2.21 gezeigte Rauschen besitzt das gleiche Leistungsdichtespektrum wie natürliche Bilder. Umgekehrt ist das LDS des invers gefilterten Bildes aus Abb. 2.17 relativ weiß. Das in (Zetzsche et al., 1993b) eingeführte Orientierungsrauschen verdeutlicht diesen Effekt noch. Es beinhaltet für das visuelle System deutliche orientierte Strukturen, die AKF zeigt jedoch eine nahezu vollständige Unkorreliertheit des Signals. Diese Beispiele offenbaren zweierlei: zum Ersten, daß Statistiken zweiter Ordnung „blind“ gegenüber orientierten Bildmerkmalen sind, was offensichtlich ist, da eine Zweipunktstatistik einfach nicht ausreicht, um Krümmungen zu erfassen; dafür bedarf es einer Analyse mit Spektren höherer Ordnung (Kap. 5). Zum Zweiten, daß das visuelle System Information jenseits der zweiten Ordnung nutzt (Julesz und Schumer, 1981; Yellott, 1993).

Ad 3. Complex-Zellen zeichnen sich in erster Linie durch ihre Phaseninvarianz aus, damit sind sie evident nichtlinear und somit nicht modellierbar mit dem klassischen Ansatz (eine Möglichkeit zur Modellierung findet sich in Kap. 5).

Ad 4. Endinhibition ist ein häufig auftretendes Phänomen in kortikalen visuellen Neuronen. Erste, klassische Vertreter sind die Hypercomplex-Zellen nach HUBEL und WIESEL (Abschnitt 1.4 und Abb. 1.6). Während diese Eigenschaft lange selbst für kortikale Zellen als „exotisch“ galt, findet sich mittlerweile zunehmend Evidenz, daß Endinhibition, oder genereller, $i2D$ -Selektivität einen generischen Mechanismus darstellt:

In recent years, it has become apparent that most V1 neurons are endstopped to some degree.

(Pack et al., 2003b)

Ein Versuch, *i2D*-Selektivität mit dem klassischen Ansatz zu erklären, scheitert einerseits daran, daß *i2D*-Signalanteile nicht mit Statistiken zweiter Ordnung erfaßt werden können (s.o.) und andererseits die Eigenfunktionen linearer Systeme eine derartige Selektivität a priori ausschließen (Zetzsche und Barth, 1990a). Die Modellierung dieses Effekts kann jedoch mit polynomialen Ansätzen erfolgen, wie in Kapitel 5 gezeigt wird.

Ad 5. Eine automatische Verstärkungsregelung (*gain control*) findet sich auf vielen Ebenen der visuellen Informationsverarbeitung. Diese beschreibt eine Adaptation der neuronalen Verarbeitung auf Änderungen des externen Stimulus und ist somit ein generelles Prinzip, welches in einer Vielzahl experimenteller Untersuchungen gefunden wurde (Yu und Lee, 2003).

Ein Standardmodell für die Verstärkungsregelung im primären visuellen Kortex und auch im auditorischen Nerv ist die divisive Normalisierung (Schwartz und Simoncelli, 2001). Dabei wird angenommen, daß der Ausgang eines (linearen) Neurons quadriert (gleichgerichtet) wird, und durch die gewichtete Summe der quadrierten Ausgänge benachbarter⁸ Neurone dividiert wird (siehe Kapitel 3 und Abbildung 3.2):

$$r_i = \frac{l_i^2}{\sum_{(j)} w_{ij} l_j^2 + c^2}. \quad (2.27)$$

Hier bezeichnet r_i den Ausgang eines nichtlinearen, verstärkungsgeregelten Neurons i und l_j den Ausgang des linearen neuronalen Subsystems j (lineares Filter). Die w_{ij} sind Gewichtungsfaktoren, die einem zu bestimmenden Optimalkriterium genügen, und c eine additive Konstante. Interessanterweise wurde in (Yu und Lee, 2003) das *Hodgkin-Huxley* (HH) sowie das *Leaky-Integrate-and-Fire* (LIF) Modell hinsichtlich ihres Antwortverhaltens auf Variationen in der Stimulusintensität untersucht. Dabei wurde festgestellt, daß diese Einzel-Neuronenmodelle ihrerseits bereits eine Verstärkungsregelung aufweisen. Dabei ist zu beachten, daß das LIF-Modell im Grunde linear ist, daher muß dieser Effekt auf die nichtlineare Aktionspotentialgenerierung zurückzuführen sein, was bei dieser Untersuchung auch gefunden wurde.

Ad 6. In den letzten Jahren ist die weiter oben beschriebene spärliche Kodierung und Independent Component Analysis zunehmend als Erklärungsansatz für die lokalisierte und orientierungsselektive Verarbeitung im visuellen Kortex herangezogen worden. In der Tat resultiert die Anwendung der ICA auf natürliche Bilder in derartig selektiven rezeptiven Feldern. Nichtsdestotrotz kann leicht gezeigt werden, daß diese Lösung zwar das

⁸Benachbart sei allgemein zu verstehen, es kann sich hierbei um eine örtliche Nachbarschaft handeln oder durchaus auch um eine funktionale (Orientierung, Radialfrequenz, etc.).

lineare Optimum darstellt, aber dennoch massive statistische Abhängigkeiten zwischen den Wavelet-Komponenten (den „unabhängigen“ Komponenten) bestehen, für Details siehe Kap. 4. Im Besonderen bestehen substantielle Abhängigkeiten zwischen Filtern mit unterschiedlichen Radialfrequenzen (Auflösungen); dies ist hauptsächlich auf das häufige Auftreten von Kanten in natürlichen Bildern zurückzuführen, da diese eine gekoppelte Aktivität von Filtern mit gleicher Orientierung verursachen. Man kann daraus schließen, daß eine lineare Verarbeitung nicht genügt, um die Struktur der statistischen Abhängigkeiten in natürlichen Bildern zu eliminieren. Die *voxata quaestio* gliedert sich nun in zwei Teilfragen: (i) welche nichtlineare Verarbeitung ist am Besten geeignet, um natürliche Bilder effizient zu verarbeiten, und (ii) wie manifestieren sich solche nichtlinearen neuronalen Systeme in experimentellen Untersuchungen?

Wie noch gezeigt wird, ist ein zentraler Teil dieser Arbeit die Verarbeitung natürlicher visueller Stimuli durch eine UND-Verknüpfung von Frequenzkomponenten. Dies stellt für die oben beschriebene statistische Frequenzkopplung eine adäquate Kodiervorschrift zur Ausnützung der statistischen Abhängigkeiten dar, da diese Einheiten nur bei signifikanter Antwort mehrerer Filter aktiv werden.

Trotz der oben beschriebenen Vielfältigkeit der Nichtlinearitäten im visuellen System findet man dennoch häufig die Meinung, daß lineare Systeme genügen, um die Verarbeitung prinzipiell zu beschreiben. Dies liegt unter anderem darin begründet, daß ein neuronales System seine Nichtlinearität nicht in trivialer Art und Weise „offenbart“. Der Artikel (Zetsche und Nuding, 2005a) beschäftigt sich eingehend mit diesem Problem; im Folgenden ein kurzes Exzerpt dieser Arbeit.

Gegeben sei ein eindimensionales, örtliches System⁹, bestehend aus dem Produkt zweier überlappender Bandpässe. Durch entsprechende Simulationen konnte gezeigt werden, daß

1. aufgrund der inhärenten Gleichrichtung in neuronalen Systemen kein Unterschied zu einem hypothetischen linearen System ersichtlich ist, wenn das System mit den üblicherweise verwendeten Sinusgittern getestet wird,
2. die somit erhaltene Übertragungsfunktion der linearen Hypothese mit der in experimentellen Untersuchungen gefundenen qualitativ übereinstimmt,
3. die Antwort des nichtlinearen Systems auf typische Teststimuli (Sprünge und Balken) ebenfalls typische Charakteristika aufweisen, die auch experimentell auftauchen,
4. die Antwort auf den optimalen sinusförmigen Stimulus durch die Addition eines zweiten Stimulus, der jedoch außerhalb des linear gemessenen Frequenzgangs liegt,

⁹Die Ortsdomäne birgt gegenüber zeitlichen Systemen den Vorteil, daß Kausalitätsüberlegungen marginal werden.

inhibiert werden kann¹⁰. Diese Inhibition, bzw. Frequenzinteraktion wurde experimentell in (DeValois und Tootell, 1983) nachgewiesen.

Aus diesen qualitativen Untersuchungen kann geschlossen werden, daß nichtlineare Systeme mit einem UND-Verhalten bezüglich der im Signal enthaltenen Frequenzkomponenten bei Verwendung von Standardtests durchaus ein lineares Verhalten aufweisen können. Nichtsdestoweniger besitzen diese Systeme ein großes Potential zur Redundanzreduktion in natürlichen Stimuli, welche zwingend eine substantielle nichtlineare Verarbeitung benötigt (siehe auch Kapitel 3). Ein minimaler kritischer Test für diese Systeme ist die Messung nicht mit einer Frequenz alleine, sondern mit Tupeln von Frequenzkomponenten. In der Tat ist dies auch eine exakte Methode, um nichtlineare Volterra-Wiener Systeme (Kapitel 5) zu charakterisieren.

¹⁰Genauer wird die Energie der Antwort bei der optimalen Frequenz reduziert, dies ist für die typischerweise verwendeten neurophysiologischen Verfahren jedoch äquivalent.

3 Nichtlineare Ansätze zur Modellierung neuronaler Informationsverarbeitung

In diesem Kapitel sollen die in dieser Arbeit verwendeten Ansätze zur Analyse und Modellierung der neuronalen Informationsverarbeitung im visuellen System kurz zusammengefaßt und verglichen werden. In Kap. 2 wurde die Bedeutung der effizienten Kodierung natürlicher Stimuli im visuellen System hauptsächlich von der Seite der Informationstheorie beleuchtet, d.h. die Analyse der hochdimensionalen Verbundwahrscheinlichkeitsdichte natürlicher Bilder. Dies ist die wohl allumfassendste, aber auch in praxi die unhandlichste Anschauungsweise, hier als *expliziter* Ansatz bezeichnet. Demgegenüber steht in Kap. 4 ein *impliziter* Ansatz, und zwar die Synthese eines bildverarbeitenden Systems auf der Basis der Optimierung der Gewichte eines neuronalen Netzes¹. Schließlich wird in Kap. 5 ein approximativer Ansatz, gegeben durch die Polynomapproximation der Verbundwahrscheinlichkeitsdichte natürlicher Bilder (Kumulanten), bzw. die Volterra-Reihenentwicklung, verwendet. Diese Ansätze haben einen komplementären Charakter, wie weiter unten genauer beschrieben wird, da ein jeder Ansatz zur Klärung bestimmter Fragestellungen geeignet ist. Die *Zustandsraumdarstellung*, welche bereits in Abschnitt 2.4.1 in ihrer speziellen Form, der Signalraumdarstellung, eingeführt wurde, eignet sich für eine gemeinsame Darstellung der verschiedenen Ansätze und ermöglicht somit einen direkten Vergleich.

3.1 Explizite, implizite und polynomiale Ansätze

3.1.1 Expliziter Ansatz: Zustandsraumdarstellung

Die Zustandsraumdarstellung wird hier als expliziter Ansatz eingeführt, da mir ihrer Hilfe nichtlineare Verarbeitungsinstanzen und statistische Abhängigkeiten explizit dar-

¹Zwar ist formal die Implementierung nicht die der in der Theorie der neuronalen Netze gebräuchlichen, jedoch ist das grundlegende Paradigma, nämlich die Optimierung von Gewichten ohne Berücksichtigung einer funktionalen Struktur, dieselbe.

gestellt werden können. Damit sind jedoch hohe Kosten verbunden, so daß eine Darstellung nur für reduzierte Konfigurationen praktikabel ist, wie weiter unten gezeigt wird. Der Vorteil dieser Darstellung liegt jedoch darin, daß einerseits spezifische Eigenschaften neuronaler Systeme leicht erklärt werden können, und daß andererseits die unten eingeführten Ansätze qualitativ miteinander verglichen werden können.

Die Zustandsraumdarstellung ist die allgemeine Form der Signalraumdarstellung. In letzterer wird prinzipiell die Abbildung eines Signals \vec{x} auf ein Bildsignal \vec{y} durch eine i. A. nichtlineare Transformation betrachtet (vgl. Kap. 2). Im neuronalen Substrat (beispielsweise im primären visuellen Kortex V1) entspräche dies der Abbildung der Luminanzwerte \vec{u}_1 auf eine neuronale Antwort u_2 . Damit kann also eine neuronale Repräsentation, bestehend aus den Antworten einer ganzen Population, als eine Koordinatentransformation des Originalraumes angesehen werden. Dabei ist von zentralem Interesse, wie die Quellenstatistik durch die Transformation verändert wird. Im idealen rauschfreien Fall sollte die Verbundstatistik der neuronalen Antworten separierbar sein, d.h. die Neurone antworten statistisch unabhängig voneinander. Dies führt dann zu den in Abschnitt 2.5.2 beschriebenen Koordinatentransformationen (für unterschiedliche Anforderungen an die Quellenstatistik).

Der einfachste Fall einer derartigen Koordinatentransformation ist eine lineare Transformation $\vec{y} = A^T \vec{x}$ mit einer orthonormalen Matrix A . Dies entspricht dann einer einfachen Rotation des Signalraums. So stellt beispielsweise

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad (3.1)$$

eine Rotation des Signalraums um 45° dar. Die Darstellung im Signalraum ist in Abb. 3.1 gezeigt. Im Rahmen der neuronalen Informationsverarbeitung entsprechen dann y_1 und y_2 den Antworten zweier unterschiedlicher Neurone auf den Eingang (x_1, x_2) . Die Antwortflächen der neuronalen Antworten bilden im Allgemeinen Hyperflächen im Zustandsraum (Abb. 3.1(a,b)), woran man auch sofort die starken Einschränkungen solcher linearer Orthogonaltransformationen erkennt. Das Ziel einer derartigen „Umkodierung“ des Eingangssignals ist eine Anpassung (Match) der (nicht-) linearen Antworten an die Verbundstatistik der Quelle. Damit ist klar, daß die Optimalität der linearen Transformationen essentiell von starken Forderungen an die Quellenstatistik abhängt. Dies wurde anhand der linearen Transformationen KLT und ICA bereits in Abschnitt 2.5.2 erläutert; die Notwendigkeit nichtlinearer Transformationen zur optimalen Transformation natürlicher Bilder wird in Kap. 4 gezeigt. Nichtdestoweniger können auch komplizierte nichtlineare Abbildungen als Transformation im Signalraum aufgefaßt werden; somit kann gezeigt werden, wie sich die Volterra'schen UND-Verknüpfungen (Kap. 5) und die nichtlineare Verstärkungsregelung (*gain control*) qualitativ von linearen Transformation unterscheiden.

Ein anderer Aspekt ist die Betrachtung der Verteilung der Wahrscheinlichkeitsmasse im

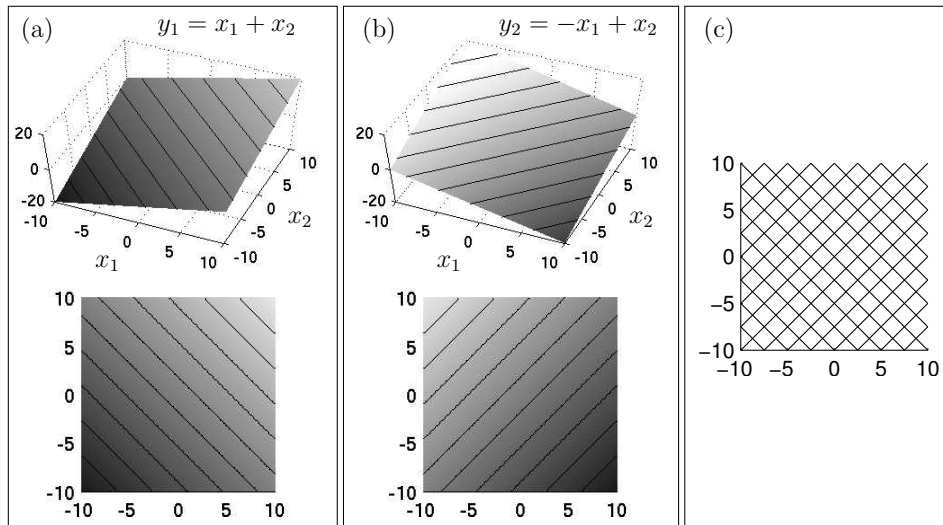


Abbildung 3.1: Lineare Transformation im Signalraum. (a), (b) Antwortflächen der transformierten Komponenten aus Gl. (3.1) in Abhängigkeit des Eingangs und ihre entsprechende Konturdarstellung. (c) Diese Transformation entspricht einer 90°-Drehung des Raumes.

Zustandsraum, speziell die Verbundwahrscheinlichkeitsdichte natürlicher Bilder. Wenn von einer Anpassung neuronaler Transformationen an die Quellenstatistik die Rede ist, so ist auch die Kenntnis der Quellenstatistik unabdingbar. Eine Möglichkeit den Aufbau dieser Dichte nachzuvollziehen ist gegeben durch die Parametrierung typischer Bildmerkmale wie z.B. Linien und Kanten. Eine kontinuierliche Änderung deren Parameter, beispielsweise Orientierung, Kontrast und Position, resultiert in Trajektorien durch den Signalraum, wie in (Zetsche und Krieger, 2001b) genauer beschrieben wird. Obwohl diese Zustandsraumdarstellung theoretisch die volle Information beinhaltet und somit eine absolute und vollständige Beschreibung ermöglicht, ist man in der Praxis durch den bereits in Abschnitt 2.4 beschriebenen „Fluch der Dimensionalität“ stark eingeschränkt. So ist eine qualitative Darstellung jenseits zwei oder drei Dimensionen bereits kaum mehr möglich und somit nicht viel informativer als klassische Zwei- oder Dreipunktstatistiken. Erkennbar ist damit hauptsächlich die bereits bekannte Tatsache, daß benachbarte Punkte hochgradig korreliert sind.

Eine Möglichkeit zur „Rettung“ dieses Ansatzes liegt in der bisher noch nicht erläuterten Dichotomie der Begriffe Zustandsraum und Signalraum. Diese soll verdeutlichen, daß es nicht notwendig ist, die Darstellung der Transformationen, bzw. Verteilungen auf die Ebene der Originaldaten (Luminanzwerte der Pixel) zu begrenzen, sondern daß bereits transformierte Größen ebenfalls als Basis dienen können. So wurde beispielsweise die multivariate WDF natürlicher Bilder in einem Wavelet-Zustandsraum gemessen (Wegmann und Zetsche, 1990a; Wegmann und Zetsche, 1990b; Zetsche und Krieger, 2001b). Dies entspricht der Verteilung der Koeffizienten orientierter Wavelets, analog

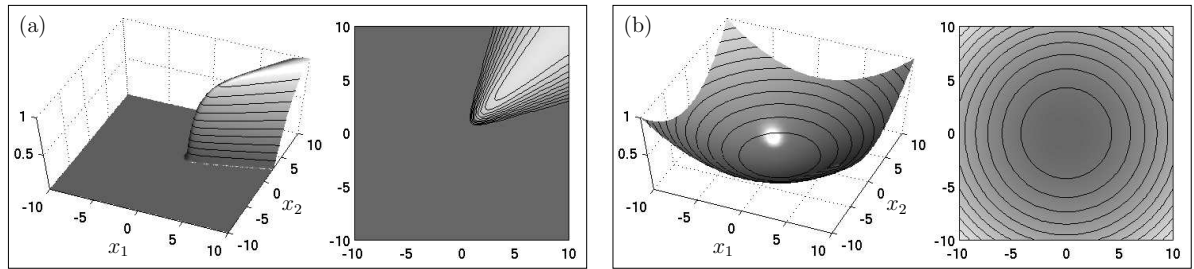


Abbildung 3.2: Illustration der nichtlinearen Selektivität und Generalisierung im Zustandsraum. (a) Selektivität entspricht einer eindeutigen Vorzugsrichtung, hier implementiert durch eine automatische Verstärkungsregelung. (b) Generalisierung entstanden aus dem Zusammenfassen quadrierter Komponenten, entsprechend dem Modell einer Complex-Zelle. Gezeigt sind jeweils die Antwortflächen der nichtlinearen Einheiten und ihre äquivalente Darstellung mittels Konturlinien. Nach (Zetzsche und Nuding, 2005b).

zu der Betrachtung in Abschnitt 4.1. Dies kann als die Verteilung zwischen bestimmten Unterräumen der wahren WDF verstanden werden und ermöglicht einen Einblick in strukturelle Eigenschaften derselben. Damit konnte gezeigt werden, daß eine Tendenz zu einer sternförmigen Struktur der WDF im Raum der orientierten Wavelets existiert, was in Übereinstimmung mit einem *i.i.d. sparse source*-Modell für natürliche Bilder ist (Wegmann und Zetzsche, 1990a). Mit einer genaueren Analyse konnte jedoch nachgewiesen werden, daß die Verteilung substantiell von einer linear separierbaren abweicht, was die Grundlage für die in Kap. 4 erläuterte Synthese eines bildkodierenden nichtlinearen LNL-Systems darstellt. Wie in (Zetzsche et al., 1999; Zetzsche und Röhrbein, 2001) gezeigt wurde, empfiehlt die genaue Struktur dieser Verteilung eine Kodierung in Polarkoordinaten. Die Phase dieser Repräsentation entspricht der „Art der Signaländerung“, und der Betrag kodiert die „Stärke der Signaländerung“ (Wegmann und Zetzsche, 1990a). Diese Merkmale können als die fundamentalen Eigenschaften einer neuronalen Repräsentation betrachtet werden. Einerseits als eine Art nichtlinearer *Selektivität*, gegeben durch eine bestimmte Vorzugsrichtung (Phase) im Zustandsraum, und andererseits eine *Generalisierung*, entsprechend dem Betrag des Vektors. Zwei mögliche Realisierungen dieser Selektivität und Generalisierung sind in der Zustandsraumdarstellung in Abb. 3.2 gezeigt. Die neuronale Einheit in Abb. 3.2(a) stellt eine hochgradig selektive Instanz bezüglich ihrer Eingangsgrößen dar, da sie nur auf eine ganz bestimmte Richtung im Zustandsraum antwortet. Die Selektivität wurde durch das Modell einer Verstärkungsregelung, entsprechend Gl. (2.27) erreicht. Die Einheit in Abb. 3.2(b) hingegen ist gekennzeichnet durch ein invariantes Verhalten, es wird lediglich die Energie des Signals kodiert, unabhängig von der Richtung im Raum. Dies wurde erreicht durch das Summieren quadrierter Komponenten ($y = x_1^2 + x_2^2$), entsprechend dem Modell der Complex-Zelle, das in Abschnitt 5.2.3 vorgestellt wird.

Dieses Konzept fügt sich hervorragend in die in dieser Arbeit besprochenen Ansätze zur Kodierung natürlicher Bilder ein. So ist die im visuellen Kortex implementierte

spärliche und übervollständige Repräsentation gekennzeichnet durch eine hochgradige Selektivität, wobei ebenfalls nachgewiesen ist, daß ein ebenso hohes Maß an Invarianz, bzw. Generalisation besteht; im einfachsten Fall die Phaseninvarianz kortikaler Complex-Zellen. Wie sich dieses Konzept mit dem Modell des LNL-Systems und der Volterra-Wiener-Systeme vereinbaren läßt, wird im Folgenden gezeigt.

3.1.2 Impliziter Ansatz: Neuronale Netze

Die direkte oder *explizite* Beschreibung der Verbundwahrscheinlichkeitsdichte natürlicher Bilder sowie deren Transformation beim Passieren nichtlinearer Systeme, wie im letzten Abschnitt beschrieben, kann auch implizit erfolgen. Algorithmen, die auf neuronalen Netzen oder ähnlichem basieren, können als derartige *implizite* Ansätze verstanden werden. Hierbei wird ein neuronales Netz an die Statistik der Quelle entsprechend eines Optimierungskriteriums angepaßt. So ist beispielsweise die KLT, bzw. PCA, ein impliziter Ansatz, bei welchem ein einstufiges neuronales Netz anhand der Quellenstatistik optimiert wird². Das Optimierungskriterium basiert hierbei auf Statistiken zweiter Ordnung und wird entsprechend den bekannten Eigenschaften der KLT gewählt, also beispielsweise Minimierung des geometrischen Mittels der Varianzen der transformierten Koeffizienten. Auch die ICA ist ein Angehöriger dieser Klasse. Hierbei wird jedoch als Optimierungskriterium häufig die Maximierung der Randentropien gewählt, wie in Abschnitt 2.5.2 bereits beschrieben. Mit dieser konnte gezeigt werden, daß die Annahme der statistischen Unabhängigkeit zu orientierten und lokalisierten rezeptiven Feldern ähnlich denen der Simple-Zellen im visuellen Kortex führt³.

Aufgrund der inhärenten Beschränkungen des einstufigen linearen Ansatzes ist eine Erweiterung hin zu nichtlinearen mehrstufigen neuronalen Netzen ein logischer nächster Schritt. Dies stellt die Motivation für die Synthese des zweischichtigen Linear-Nichtlinear-Linear (LNL) Systems aus Kap. 4 dar. Wie dort noch ausführlich gezeigt wird, ist dieses System in der Lage, die in den linearen Wavelet-Koeffizienten verbleibenden statistischen Abhängigkeiten höherer Ordnung durch eine simple Aufspaltung in *ON*- und *OFF*-Anteile in Korrelationen zweiter Ordnung zu transformieren, welche in einer einfachen linearen Transformation wiederum eliminiert werden können. Diese Lernprozedur führt nicht nur zu interessanten und vorteilhaften Merkmalen für die Kodierung und Kompression natürlicher Bilder, sondern auch zu einem Antwortverhalten, das deutliche Parallelen zur Verarbeitung im visuellen Kortex aufweist. Man kann beispielsweise die Phaseninvarianz kortikaler Complex-Zellen, die Längenabstimmung von endinhibierten Simple-Zellen, oder auch die Kombination dieser Effekte, entsprechend kortikaler Hypercomplex-Zellen beobachten (Zetsche und Röhrbein, 2001; Nuding, 2002). Ebenso

²Die KLT ist ein Sonderfall, da hier auch rein analytische Lösungen existieren, für praktische Zwecke wird aber häufig eine Implementierung mittels neuronalen Netzen gewählt (Hebb'sches Lernen).

³Auch wenn diese in der Tat überhaupt nicht erreicht wird.

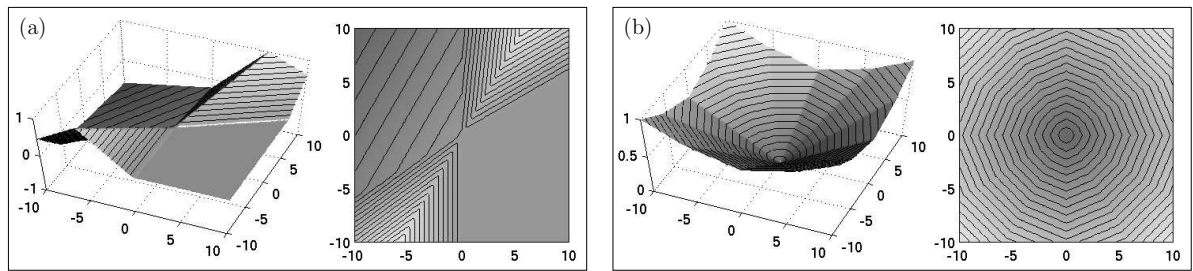


Abbildung 3.3: Zwei Einheiten einer reduzierten LNL-Architektur aus Kap. 4 im Zustandsraum. Die Antwortflächen bestehen aus zusammengesetzten (Hyper-) flächen, womit im hochdimensionalen beliebige Funktionen approximiert werden können. Somit sind auch die basalen Funktionen der Selektivität (a) und der Invarianz (b) realisierbar.

können die beschriebenen grundlegenden Funktionen der Selektivität und Generalisierung erreicht werden. Stellt man ein (stark reduziertes) LNL-System im Zustandsraum entsprechend Gl. (4.14) dar ($\vec{y} = \mathcal{LN}\mathcal{L}[\vec{x}]$), so kann für entsprechende Kombinationen der linearen Gewichte leicht ein derartiges Verhalten synthetisiert werden, siehe Abb. 3.3. Damit verwandt ist auch die *Slow Feature Analysis* (Berkes und Wiskott, 2004), sowie die zweischichtige Architektur, die in (Hyvärinen und Hoyer, 2001) vorgestellt wurde.

Ferner konnte ich in (Nuding und Zetzsche, 2006) zeigen, daß ein derartiges LNL-System zu vergleichbaren Selektivitäten gegenüber verschiedenen Stimulusklassen führt wie in neurophysiologischen Experimenten gefunden wurde. Ebenso führt dieser Ansatz zu substantiellen Interaktionen zwischen Frequenzen unterschiedlicher Skalen, ein Hinweis auf die Prädominanz von Kanten und Linien in natürlichen Bildern (siehe hierzu auch die Abschnitte 4.5 und 4.6).

Der Nachteil dieser Ansätze ist jedoch genau ihre implizite Natur, da einerseits nur Projektionen der wahren Verbunddichte analysiert werden und somit keine strukturellen Eigenschaften der Statistik erkennbar sind, und andererseits die Ergebnisse keine einfache funktionale Interpretation im Sinne realisierbarer Modelle gewähren.

3.1.3 Polynomialer Ansatz: Polyspektren und Volterra-Wiener-Systeme

Die beiden vorgestellten Ansätze der Zustandsraumbeschreibung und der neuronale Netze können als Extremfälle betrachtet werden, die sozusagen die beiden Enden eines Spektrums darstellen. Die in dieser Arbeit intensivsten verwendete Methode ist die der Approximation mittels Potenzreihen. Auf Seite der Analyse und Darstellung der Statistik natürlicher Bilder sind dies Spektren höherer Ordnung, im Besonderen das Bisppektrum. Der Begriff Potenzreihenentwicklung der Statistik liegt darin begründet,

daß die Kumulanten die Koeffizienten der Taylor-Reihenentwicklung der zweiten charakteristischen Funktion darstellen (vgl. Abschnitt 5.1). Im Gegensatz zu herkömmlichen Betrachtungsweisen (Autokorrelation und Leistungsdichtespektrum) ermöglichen Polyspektren prinzipiell eine vollständige Beschreibung der statistischen Eigenschaften. Doch auch hiermit kann aufgrund Stabilitäts- und Darstellungsbeschränkungen nur ein relativ niedriger Approximationsgrad betrachtet werden. Nichtsdestoweniger sind bereits im Spektrum dritter Ordnung (Bispektrum) relevante Informationen über die statistischen Bindungen *zwischen* unterschiedlichen Frequenzkomponenten enthalten, was eine wesentliche Generalisierung darstellt.

Auf Seite der Modellierung, bzw. Synthese nichtlinearer Systeme ist der polynomiale Ansatz durch die Volterra-Wiener Systemtheorie gegeben. Die Frage ist nun, welche Klasse von Flächen durch diese im Zustandsraum dargestellt werden können. Betrachtet man den einfachen Fall eines quadratischen Volterra-Systems, so ist klar, daß es sich aufgrund der quadratischen Form (siehe Gl. (5.65)) um eine Untergruppe der Flächen zweiter Ordnung handeln muß. Diese sind im Allgemeinen durch ihre Normalform definiert (Råde und Westergren, 1997):

$$l_1x_1^2 + l_2x_2^2 + l_3x_3^2 + 2dx_3 + e = 0. \quad (3.2)$$

Die Eingangs- Ausgangsrelation eines quadratischen Volterra-Systems nimmt für zwei Eingänge die Form

$$y = h_{11}x_1^2 + h_{22}x_2^2 + h_{12}x_1x_2 \quad (3.3)$$

an. Man beachte, daß dies nicht notwendigerweise direkt mit Luminanzwerten von Pixeln korrespondiert, da man hier vom allgemeinen Zustandsraum ausgeht, d.h. die x_i können beispielsweise auch die Koeffizienten von orientierten Filtern darstellen. Daraus kann die Matrix A sowie der Vektor \vec{b} der allgemeinen Form

$$Q(\vec{x}) = \vec{x}^T A \vec{x} + 2\vec{b}^T \vec{x} + c \quad (3.4)$$

einer Quadrik $Q(\vec{x})$ abgeleitet werden (x_3 dieser Form entspricht in Gl. (3.3) y):

$$A = \begin{pmatrix} h_{11} & \frac{h_{12}}{2} & 0 \\ \frac{h_{12}}{2} & h_{22} & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 0 \\ 0 \\ -\frac{1}{2} \end{pmatrix}, \quad c = 0. \quad (3.5)$$

Die Normalform kann aus dieser Darstellung durch die Berechnung der Eigenwerte und -vektoren erhalten werden. Bei dieser Berechnung erkennt man leicht, daß ein Eigenwert identisch Null ist. Man erhält also eine degenerierte Normalform, was darin begründet liegt, daß das Glied y^2 , bzw. x_3^2 fehlt. Damit können mittels eines Volterra-Operators genau zwei Quadriken realisiert werden:

- hyperbolische Paraboloiden und

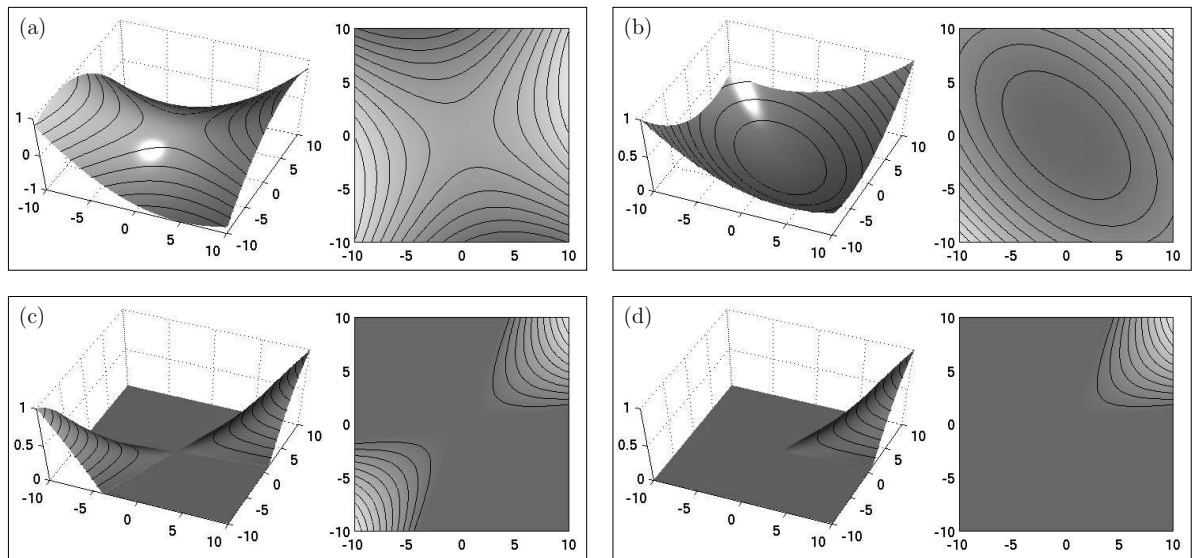


Abbildung 3.4: Mögliche Antwortflächen eines quadratischen Volterra-Operators. (a) Hyperbolisches Paraboloid, (b) elliptisches Paraboloid. (c) Hyperbolisches Paraboloid mit Ausgangsgerichtung und (d) zudem mit gleichgerichtetem Eingang.

- elliptische Paraboloid.

Diese sind beispielhaft in Abb. 3.4(a,b) mit ihren entsprechenden Konturliniendarstellungen abgebildet. Die hyperbolische Realisierung des quadratischen Volterra-Operators (a) entspricht genau der in Kap. 5 ausführlich beschriebenen UND-Selektivität, und gibt damit einer Richtung im Zustandsraum einen Vorzug. Dies korrespondiert also mit einer bezüglich des Zustandsraums erhöhten Selektivität, vergleichbar mit der Selektivität der Verstärkungsregelung (Abb. 3.2a) und LNL-Approximation (Abb. 3.3a). Zudem erkennt man sofort, daß auch die fundamentale Funktion der Generalisierung leicht zu erreichen ist, siehe Abb. 3.4(b). Speziell wenn die Achsen die Ausgänge eines Quadraturfilterpaares darstellen, erhält man genau das in Abschnitt 5.2.3 vorgestellte Modell einer Complex-Zelle. Die Eigenschaft der nichtlinearen Selektivität wird deutlicher, wenn zudem gleichrichtende Nichtlinearitäten mit in Betracht gezogen werden, was im Kontext der neuronalen Informationsverarbeitung durchaus gängig ist. Dies ist in Abb. 3.4(c,d) für den Fall einer einfachen Ausgangsgerichtung, bzw. einer zusätzlichen Eingangsgerichtung gezeigt.

3.2 Extra-klassische Effekte im Zustandsraum

Als *extra-klassisch* werden in der Neurophysiologie Effekte bezeichnet, bei welchen ein Stimulus außerhalb des klassischen rezeptiven Feldes die Antwort auf einen Stimulus innerhalb modulieren kann. Diese Effekte werden im Detail in Kapitel 5 beschrieben und

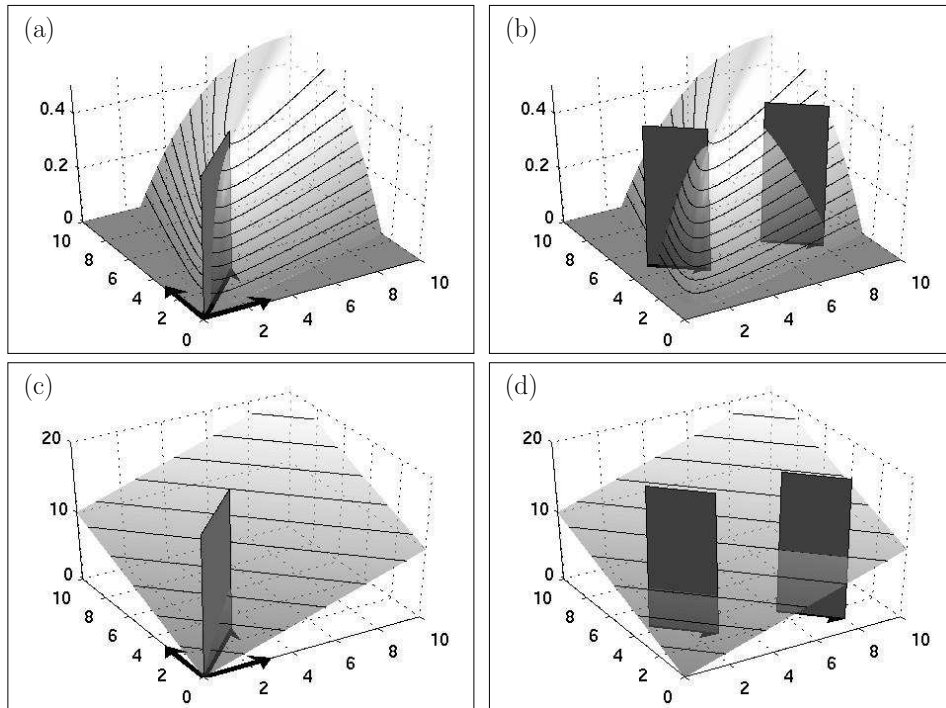


Abbildung 3.5: Darstellung extra-klassischer Effekte im Zustandsraum für nichtlineare selektive Mechanismen (a,b) und das Analogon für linearen Mechanismen (c,d). Details siehe Text.

mittels quadratischen Volterra-Systemen modelliert. Anhand einfacher Überlegungen im Zustandsraum wird klar, wie einfach diese „ominösen“ extra-klassischen Effekte bei einer zugrundeliegenden selektiven Nichtlinearität auftreten können, und wie sie zu deuten sind.

Dazu betrachtet man ein hypothetisches Neuron mit einer hochgradigen Selektivität, d.h. mit einem ähnlichen Antwortverhalten wie bereits in Abb. 3.2(a) dargestellt. Es sollen nun zwei Effekte anhand einer derartigen Antwortfläche erklärt werden. (i) Ein bilineares Antwortverhalten, bei welchem zwei einzelne Stimuli keine Antwort evozieren können, wohl aber deren Kombination, und (ii) die Suppression der Antwort auf einen optimalen Stimulus durch die Addition eines zweiten Stimulus, der für sich alleine keine Antwort verursacht⁴.

Betrachtet man Abb. 3.5(a), so erkennt man, daß zwei Stimuli, gegeben durch die schwarz eingezeichneten Vektoren, für sich alleine keine Antwort auf einer derartigen Antwortfläche hervorrufen können. Die lineare Summe der beiden Stimuli jedoch, gegeben durch die vektorielle Addition der Einzelstimuli, zeigt genau in die Vorzugsrichtung der nicht-linearen Selektivität, und stellt somit den optimalen Stimulus dar.

⁴Dies ist nicht ganz konsistent mit der in Kap. 5 verwendeten Bezeichnung, soll aber hier der Einfachheit halber unter dem Begriff „extra-klassisch“ subsumiert werden.

Eine etwas kompliziertere Darstellung ergibt sich, wenn beispielsweise die SILLITO- oder SHEVELEV-Effekte (vgl. Abschnitt 5.2.4) im Zustandsraum erklärt werden. Dies ist in Abb. 3.5(b) illustriert und zeigt, daß ein Stimulus, der für sich alleine keine Antwort hervorruft (nicht eingezeichnet, dies wären die in den Ursprung verschobenen Pfeile) durchaus die Antwort auf einen effektiven Stimulus modulieren kann. Gezeigt ist, daß im Extremfall ein und derselbe Stimulus, wenn er zu einem anderen Stimulus (also an einer anderen Stelle des Raums) aufaddiert wird, zu völlig unterschiedlichen Resultaten führen kann. Es kann zu einer Erhöhung (Exzitation) oder auch Unterdrückung (Inhibition) der Antwort führen. Obwohl trivialerweise eine derartige Selektivität für lineare Mechanismen per Definition nicht erreichbar ist, kann dies auch in der Zustandsraumdarstellung noch verdeutlicht werden. Daher sind in Abb. 3.5(c,d) die gleichen Stimuluskonfigurationen für eine lineare Antwortfläche gezeigt, woraus ersichtlich ist, daß die oben beschriebenen extra-klassischen Effekte hierbei nicht auftreten können.

3.3 Nichtlineare Partitionierung des Zustandsraums

Bereits in Kap. 2 wurde der Begriff der spärlichen und verteilten Kodierung eingeführt. Im Kontext neuronaler Netze, bzw. linearer Optimierung wird dies häufig als Motivation für ICA-ähnliche Algorithmen verwendet, da diese typischerweise die Kurtosis, ergo die Spärlichkeit, maximieren. Nun liegt aber, von formalen Definitionen abgesehen, die Spärlichkeit einer neuronalen Repräsentation nicht *per se* in der hohen Kurtosis der Antwortverteilung, sondern daran, daß die Neurone hochgradig selektiv antworten. Zudem findet sich in einigen ICA-Implementierungen das Prinzip der Übervollständigkeit, d.h. es wird von einer Repräsentation ausgegangen, die mehr neuronale Elemente (im linearen Fall sind das Basisvektoren) besitzt, als für eine ein-eindeutige Abbildung notwendig wären. Wenn aber nun dieses Problem in den Zustandsraum übertragen wird, so kann man sich leicht klar machen, daß eine tatsächlich spärliche Repräsentation mit linearen Transformationen nicht erreicht werden kann, und daß das Prinzip der Übervollständigkeit ein irreführender Begriff ist.

In Abb. 3.6(a) sind vier selektive nichtlineare Einheiten mit jeweils unterschiedlicher Vorzugsrichtung im Zustandsraum dargestellt. Bei Verwendung von genügend solcher Einheiten kann leicht der gesamte Raum abgedeckt werden und so eine ein-eindeutige Repräsentation erreicht werden (in der Abb. rechts dargestellt). Klassischerweise würde man in diesem Fall von einer übervollständigen Repräsentation sprechen, da zur Darstellung zweier Eingangsgrößen insgesamt 12 nichtlineare Einheiten verwendet werden. Wie allerdings bereits qualitativ der Abbildung zu entnehmen ist, kann hier nicht von Übervollständigkeit gesprochen werden, da aufgrund der hochgradigen Selektivität der Einheiten *alle* für eine ein-eindeutige Repräsentation benötigt werden. Wendet man eine ähnliche Konfiguration auf lineare Einheiten an, wie in Abb. 3.6(b) gezeigt ist, so ergibt

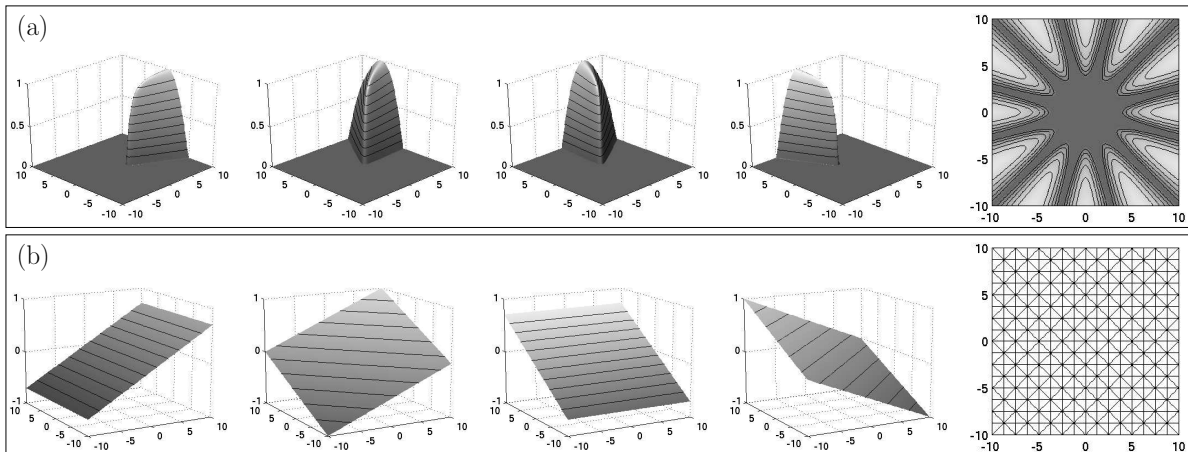


Abbildung 3.6: Darstellung einer Partitionierung des Zustandsraums mit nichtlinearen, hochgradig selektiven Einheiten (a), sowie mit linearen Einheiten (b). Nur bei der Konfiguration in (b) kann man tatsächlich von einer übervollständigen Repräsentation sprechen.

sich tatsächlich eine übervollständige und redundante Repräsentation. Dies liegt an der mangelnden Selektivität linearer Einheiten, wie bereits anderswo gezeigt (Zetzsche und Röhrbein, 2001; Zetzsche und Nuding, 2005b)

Im Allgemeinen kann das Kodierziel bei Verwendung einer selektiven nichtlinearen Repräsentation als die Anpassung an die Verbundwahrscheinlichkeitsdichte im Zustandsraum verstanden werden. Dabei werden Regionen mit hoher Wahrscheinlichkeitsdichte durch eine Vielzahl extrem selektiver Einheiten repräsentiert, wohingegen Regionen mit nahezu keiner Auftrittswahrscheinlichkeit durch Einheiten mit einer sehr breiten Abstimmung dargestellt werden können. Dies ist direkt verwandt mit dem Prinzip der Vektor-Quantisierung, die ebenfalls entsprechend der Auftrittswahrscheinlichkeit kodiert. Dieses Kodierschema ist in Abb. 3.7 illustriert, wobei die Wahrscheinlichkeitsdichte durch Luminanz kodiert ist.

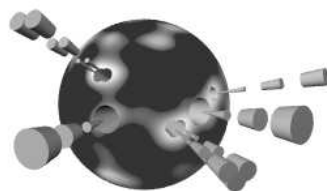


Abbildung 3.7: Nichtlineare Partitionierung des Zustandsraums in Abhängigkeit der Wahrscheinlichkeitsdichte. Regionen mit hoher Wahrscheinlichkeit werden höher aufgelöst durch selektivere und dementsprechend mehr Einheiten. Dieses Schema birgt eine hohe Ähnlichkeit zu einem Vektor-Quantisierer.

Als Realisierung solch selektiver Einheiten eignen sich, wie in den letzten Abschnitten gezeigt wurde, hyperbolische Volterra-Operatoren und eine Verstärkungsregelung entsprechend Gl. (2.27) und Abb. 3.2. In (Zetsche und Röhrbein, 2001; Zetsche und Nuding, 2005b) wird zudem gezeigt, wie eine Verstärkungsregelung im Detail zu einer erhöhten Selektivität führt und wie diese im Zusammenhang mit den UND-Operationen der Volterra-Systeme verstanden werden kann.

4 Synthese eines nichtlinearen LNL-Systems zur Bildkodierung

Ein lineares Mehrkanalmodell für die Merkmalsextraktion und Bildkodierung kann, wie bereits in Kapitel 2 gezeigt, durch die Anwendung statistischer Optimierungsalgorithmen, wie die *Principal Component Analysis* (PCA) oder die *Independent Component Analysis* (ICA), abgeleitet werden. Für typische natürliche Grauwertbilder resultiert die Applikation von ICA in Gabor-artigen Wavelets (Olshausen und Field, 1997). Diese Wavelet-Filter weisen eine hohe Ähnlichkeit zu Neuronen im visuellen Kortex auf (DeValois und DeValois, 1990; Olshausen und Field, 1997) und eine Vielzahl technischer Anwendungen hat bewiesen, daß diese auch für die Bilddatenkompression zu hervorragenden Ergebnissen führen (Vetterli und Kovacevic, 1995; Cosman et al., 1996; Marcellin et al., 2000). Nichtsdestotrotz besitzt ein linearer Ansatz inhärente Beschränkungen, weshalb es also durchaus gerechtfertigt ist, nach einer Erweiterung hin zu nichtlinearen Verfahren zur Bildkodierung zu suchen. Dabei sind vorrangig drei Punkte zu klären: *(i)* Existieren überhaupt substantielle statistische Abhängigkeiten im optimalen linearen Kode? *(ii)* Können adäquate nichtlineare Transformationen gefunden werden, die diese statistischen Abhängigkeiten eliminieren, respektive reduzieren? *(iii)* Ist es möglich, eine diese Transformationen realisierende Architektur zu finden, unter Erhaltung der Rekonstruktionsfähigkeit?

In diesem Kapitel werden diese Fragen im Detail eruiert. Dazu wird zuerst in Abschnitt 4.1 die Existenz von statistischen Abhängigkeiten zwischen den linearen Wavelet-Koeffizienten empirisch bewiesen, um anschließend in Abschnitt 4.2 zu demonstrieren, wie einfache rektifizierende Nichtlinearitäten in der Lage sind, diese zu transformieren. Genauer hat diese Rektifizierung der linearen Koeffizienten das Potential, statistische Abhängigkeiten höherer Ordnung (teilweise) in statistische Abhängigkeiten zweiter Ordnung (Korrelationen) zu transformieren. Diese Transformation ermöglicht es nachfolgenden linearen Transformationen, diese Korrelationen wiederum zu eliminieren. Dies motiviert den Aufbau einer zweischichtigen Architektur mit intermediärer Rektifizierung, erläutert in Abschnitt 4.3, sowie in (Nuding und Zetzsche, 2003; Nuding und Zetzsche, 2006). Die Ergebnisse der statistischen Optimierung der Ausgangsschicht mit PCA und ICA, sowie die korrespondierenden Basisfunktionen, sind ebenfalls in diesem Abschnitt gezeigt. Durch Implementierung eines einfachen Codecs ist es möglich, einerseits die Rekonstruktionsfähigkeit zu zeigen, und andererseits die Möglichkeit eines

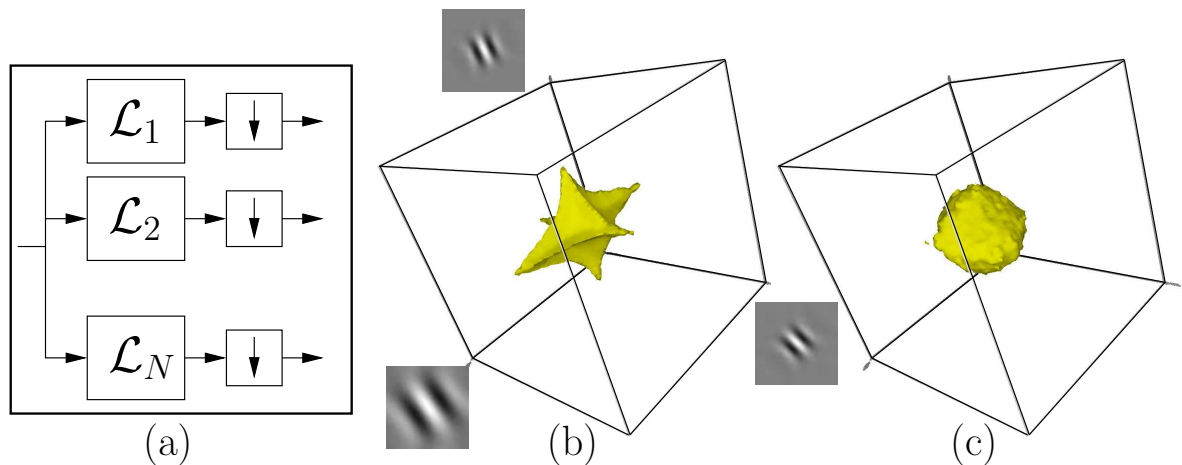


Abbildung 4.1: Statistische Abhängigkeiten zwischen den Koeffizienten einer linearen Filterbank. (a) Schema der linearen Filterbank, (b) Iso-Wahrscheinlichkeitsflächen der multivariaten WDF dreier hypothetisch statistisch unabhängiger Wavelet-Koeffizienten, (c) gemessene Verbund-WDF.

Kodierungsgewinns durch sukzessives Eliminieren von nichtlinearen Koeffizienten und einer varianzabhängigen Bitallokation zu demonstrieren. Der Vergleich mit einem linearen Referenz-Codec liefert weitere Evidenz für die Vorteile des nichtlinearen Zweischicht-Systems. Desweiteren wird die Rolle von unterschiedlichen lokalen Strukturklassen des Systems analysiert (Abschnitt 4.4). In Abschnitt 4.5 werden nichtlineare Interaktionen zwischen unterschiedlichen Frequenzkomponenten analysiert und die Möglichkeit der sukzessiven Erweiterung hin zu einem Mehrschicht-System aufgezeigt. Zum Abschluß werden in Abschnitt 4.6 einige Aspekte bezüglich der biologischen Relevanz des Systems diskutiert, vgl. hierzu (Nuding und Zetzsche, 2006).

4.1 Statistische Abhängigkeiten zwischen linearen Wavelet-Koeffizienten

Obwohl gemeinhin angenommen wird, daß die *Independent Component Analysis* auch tatsächlich statistisch unabhängige Koeffizienten liefert, gilt dies nicht a priori für den Fall allgemeiner multivariater Wahrscheinlichkeitsdichtefunktionen (vergleiche hierzu die Diskussion in Abschnitt 2.7). Die Abhängigkeiten zwischen den Koeffizienten über örtliche und spektrale Dimensionen hinweg wurde bereits von WEGMANN UND ZETZSCHE und anderen nachgewiesen (Wegmann und Zetzsche, 1990a; Zetzsche et al., 1993c; Schwartz und Simoncelli, 2001; Zetzsche und Krieger, 2001b). Dieser Effekt ist in Abb. 4.1 illustriert. Würde die Hypothese der statistischen Unabhängigkeit von orientierungsselektiven, selbstähnlichen Wavelet-Koeffizienten halten, so müßte notwendigerweise gel-

ten, daß die Verbundwahrscheinlichkeitsdichtefunktion gleich dem dyadischen Produkt der Randdichten ist:

$$f(x_1, x_2, \dots, x_N) = f(x_1) \cdot f(x_2) \cdots f(x_N). \quad (4.1)$$

In Abb. 4.1(b) ist die hypothetisch unabhängige Dichte (in der Darstellung als Iso-Wahrscheinlichkeitsflächen) für die Amplitudenverteilung dreier Wavelets gezeigt. Zwei sind aus einer höher auflösenden Schicht mit Orientierungen $+10^\circ$ und -10° , das dritte aus einer niedriger auflösenden Schicht mit Orientierung 0° ¹. Bei Betrachtung der empirisch bestimmten Verbunddichte, siehe Abb. 4.1(c), kann man leicht erkennen, daß diese deutlich vom hypothetischen Idealfall abweicht und somit substantielle statistische Abhängigkeiten in den Wavelet-Koeffizienten verbleiben. Dies impliziert, daß eine lineare Transformation für natürliche Bilder prinzipiell suboptimal ist, da die Wavelet-Basis bereits das lineare Optimum darstellt². Weitere lineare Transformationen dieser Koeffizienten können selbstverständlicherweise keinen weiteren Gewinn erzielen, da aufeinanderfolgende lineare Transformationen wiederum in einer linearen Transformation resultieren, und auch diese nicht besser als das Optimum sein kann. Dies kann auch direkt aus der Form der Verbunddichte geschlossen werden: Die ungefähre Rotations-symmetrie kann durch eine weitere lineare Transformation nicht erfaßt werden, da diese auf eine Rotation und Scherung des Koordinatensystems beschränkt ist.

Um diese statistischen Abhängigkeiten für einen Kodierungsgewinn zugänglich zu machen, muß ein nichtlineares System folgende Bedingungen erfüllen: (i) Transformation oder Eliminierung der Abhängigkeiten und (ii) Bewahrung der Invertierbarkeit. Während die zweite Bedingung für lineare Systeme (außer in degenerierten Sonderfällen) *in rerum natura* gegeben ist, muß dies im allgemeinen Fall nichtlinearer Systeme explizit bewiesen werden. Daher wird im Folgenden eine einfache nichtlineare Architektur untersucht, die beide Bedingungen erfüllt.

4.2 Nichtlineare Transformation der statistischen Abhängigkeiten

Die hier verwendete bipolare Einweggleichrichtung läßt sich beschreiben als:

$$v_i \rightarrow \{v_i^{ON}, v_i^{OFF}\} = \{\mathcal{N}^+[v_i], \mathcal{N}^-[v_i]\}, \quad (4.2)$$

¹Die Absolutwerte der Orientierung sind irrelevant, da von einer Isotropie der Statistik natürlicher Bilder ausgegangen werden kann.

²Wie bereits erwähnt, entspricht die Wavelet-Basis der auf Basis von natürlichen Bildern optimierten ICA-Basis, kann also als das lineare Optimum angesehen werden

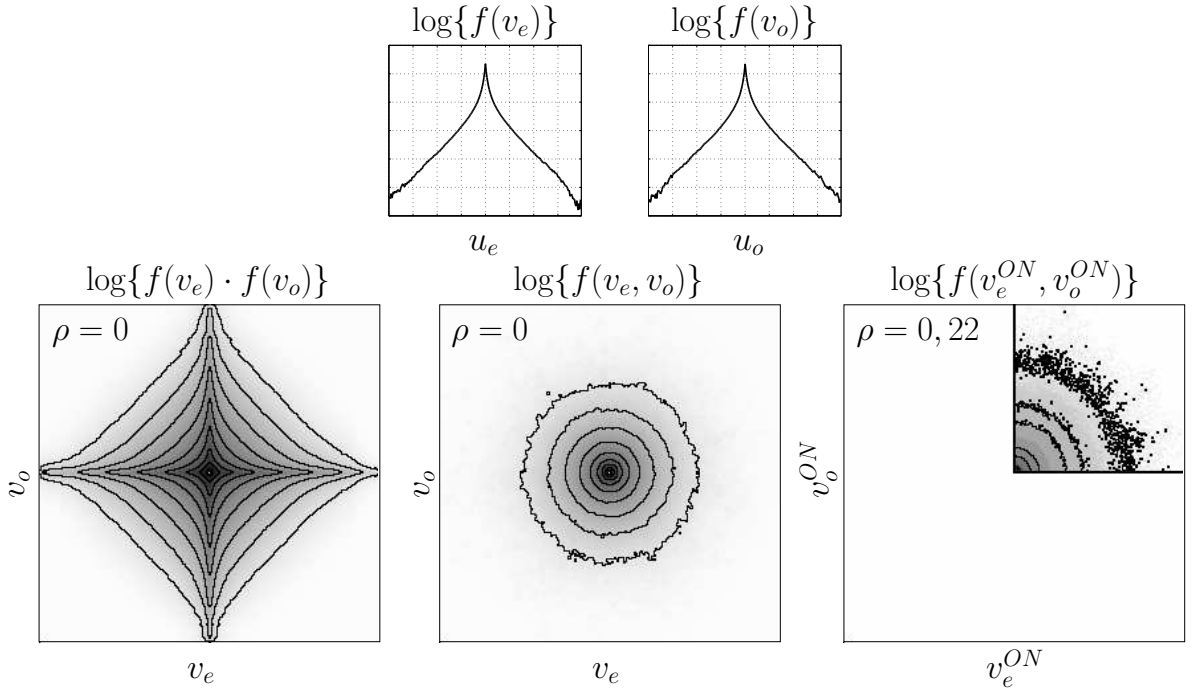


Abbildung 4.2: Transformation von statistischen Abhängigkeiten höherer Ordnung in Korrelationen zweiter Ordnung. Durch Vergleich der unkorrelierten gemessenen Verbunddichte (mittig) mit der statistisch unabhängigen (links) erkennt man die Abhängigkeiten höherer Ordnung. Diese können teilweise durch eine einfache Halbweg-Gleichrichtung in Korrelation zweiter Ordnung transformiert werden (rechts).

mit den gleichrichtenden Nichtlinearitäten

$$\mathcal{N}^+[v_i] = \mathcal{H}[v_i] \cdot v_i \quad (4.3)$$

$$\mathcal{N}^-[v_i] = \mathcal{H}[-v_i] \cdot (-v_i). \quad (4.4)$$

\mathcal{H} bezeichnet hier den Einheitssprung, oder auch HEAVISIDE-Funktion. Im Folgenden wird gezeigt, daß diese simplen gedächtnislosen Nichtlinearitäten das Potential besitzen, Korrelationen höherer Ordnung teilweise in Korrelationen zweiter Ordnung zu transformieren. Dazu betrachtet man den Effekt dieser einfachen Halbweg-Gleichrichtung auf die Verbundwahrscheinlichkeitsdichte zweier Wavelet-Koeffizienten. Als kanonisches Beispiel werden hier die Koeffizienten eines gerade- und eines ungeradesymmetrischen Filters (v_e, v_o) untersucht, da diese per Definition unkorreliert sind (Zetzsche et al., 1999).

In Abb. 4.2 ist zu erkennen, daß die gemessene VWDF $f(v_e, v_o)$ zwar unkorreliert (in diesem Fall rotationssymmetrisch) ist, aber dennoch deutlich von der hypothetisch unabhängigen abweicht, signalisierend, daß hier statistische Abhängigkeiten höherer Ordnung zwischen den Koeffizienten existieren müssen. Nach der Gleichrichtung, also $\{v_e; v_o\} \rightarrow \{v_e^{ON}, v_e^{OFF}; v_o^{ON}, v_o^{OFF}\}$ sind diese teilweise in Korrelationen zweiter Ordnung

transformiert, wie beispielhaft für $f(v_e^{ON}, v_o^{ON})$ gezeigt ist³. *Notabene:* Für den Fall einer tatsächlich unabhängigen Verteilung resultiert die Gleichrichtung in einem Korrelationskoeffizient (KK) identisch Null, d.h. die nach der Transformation auftretenden Korrelationen sind kein Artefakt der Verarbeitung. Für eine ausführlichere Aufstellung über die Zusammenhänge zwischen den Korrelationskoeffizienten unterschiedlicher Konfigurationen sei der Leser auf (Zetzsche und Röhrbein, 2001) sowie (Nuding, 2002) verwiesen. Den prinzipiellen Mechanismus, der zu diesen Transformationen der statistischen Abhängigkeiten führt, kann man anhand eines simplen diskreten Beispiels annähernd erklären.

Beispiel. Gegeben sei eine diskrete Verbunddichte $p(x, y)$, entsprechend Abb. 4.3. Diese hat als freie Parameter die Massen auf den Achsen (b), sowie die Massen auf den Diagonalen (a). Der Wert im Ursprung ergibt sich aus der Forderung für WDFen,

$$\int \int p_{XY}(x, y) \cdot dx dy = 1. \quad (4.5)$$

Damit ergibt sich die erste Randbedingung, da die Masse im Ursprung positiv sein muß:

$$a + b \leq \frac{1}{4}. \quad (4.6)$$

Diese Verteilung ist statistisch unabhängig, wenn folgende Bedingung hält:

$$a = 4a^2 + 4ab + b^2. \quad (4.7)$$

Rein rechnerisch ergibt sich durch Lösen der quadratischen Gleichung bei freier Wahl des Parameters b :

$$a_{1,2} = \frac{1 - 4b \pm \sqrt{1 - 8b}}{8}. \quad (4.8)$$

Also gilt als Nebenbedingung für die statistische Unabhängigkeit, daß b in einem gewissen Intervall zu liegen hat:

$$0 \leq b \leq \frac{1}{8}. \quad (4.9)$$

Ausgedrückt mit den Massen auf den Diagonalen und Achsen, a, b ergibt sich für den Korrelationskoeffizienten nach der Einweggleichrichtung der folgende Ausdruck, dargestellt in Abb. 4.3:

$$\rho_{XY} = \frac{a - 4a^2 - 4ab - b^2}{2a - 4a^2 - 4ab + b - b^2} \quad (4.10)$$

Dieser läßt sich erheblich vereinfachen, wenn zwei Hilfsgrößen eingeführt werden:

$$\epsilon = a + b \quad (4.11)$$

$$\gamma = a - 4a^2 - 4ab - b^2. \quad (4.12)$$

Dabei kann ϵ als reziproker Spärlichkeitsparameter betrachtet werden, für kleine Werte liegt der Großteil der Masse im Ursprung und für große Werte im Außenbereich. Die Größe γ ergibt

³Der Korrelationskoeffizient liegt bei anderen Konfigurationen (ON-OFF, unterschiedliche örtliche Positionen) deutlich höher, dieses Beispiel wurde aber aus Gründen der Klarheit gewählt.

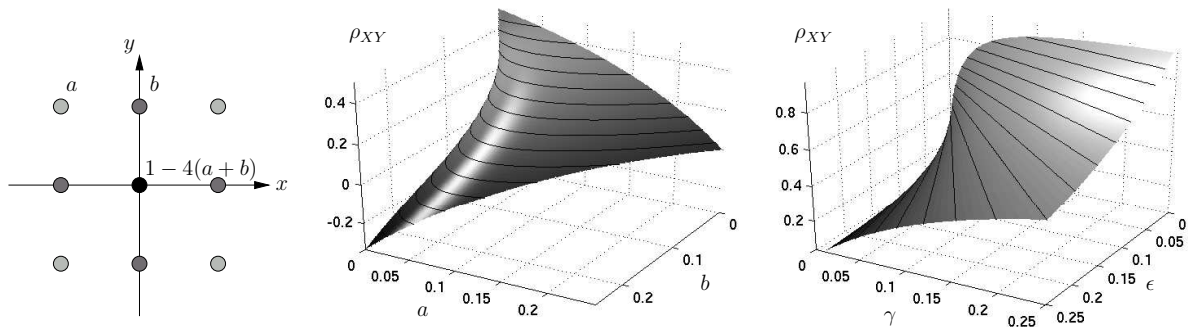


Abbildung 4.3: Links: Diskrete Verbunddichte zur Illustration der Transformation der statistischen Abhängigkeiten durch die bipolare Rektifizierung. Mitte: KK nach der Gleichrichtung in Abhängigkeit von a und b , rechts in Abhängigkeit von γ und ϵ .

sich aus Gl. (4.7) und kann als Abweichung von der statistischen Unabhängigkeit interpretiert werden. Daraus ergibt sich nun der folgende, vereinfachte Zusammenhang zwischen dem Korrelationskoeffizient nach der Gleichrichtung, dem Grad der Spärlichkeit und der Abweichung der tatsächlichen Verbunddichte von der hypothetisch unabhängigen Verteilung:

$$\rho_{XY} = \frac{\gamma}{\epsilon + \gamma}. \quad (4.13)$$

Dieses Ergebnis läßt die folgenden Implikationen zu:

- Für kleines ϵ , d.h. große Spärlichkeit, geht der Korrelationskoeffizient nach der Gleichrichtung gegen 1.
- Für großes γ , d.h. starke Abweichung von der statistischen Unabhängigkeit, geht der Korrelationskoeffizient ebenfalls gegen 1.

□

4.3 Eine invertierbare nichtlineare Architektur

Die Invertierbarkeit ist im Allgemeinen für nichtlineare Signalverarbeitungsmechanismen nicht gegeben. Verwendet man beispielsweise eine gradesymmetrische Nichtlinearität ($v_i^{NL} = v_i^{2n}, n \in N \setminus \{0\}$) oder das Produkt einiger Koeffizienten ($v_k^{NL} = v_i \cdot v_j$), so ist die Information über das Vorzeichen der einzelnen Koeffizienten verloren. Mit der hier verwendeten Architektur ist es jedoch möglich, ohne zusätzliche Regeln auf das Eingangssignal zurückzurechnen, weshalb sich diese auch für die Applikationsdomäne der Bildkodierung anbietet. Die einzige Bedingung, die für die Invertierbarkeit (oder Erhaltung der Information) gelten muß, ist die Vollständigkeit der linearen Transformationen.

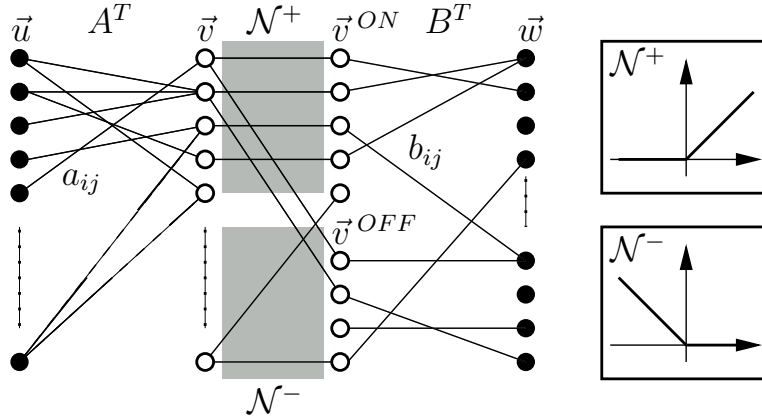


Abbildung 4.4: Netzwerk-Darstellung der LNL-Architektur. Notwendige und hinreichende Bedingung für die Invertierbarkeit ist die Vollständigkeit der Transformationsbasen in A und B .

4.3.1 Prinzipieller Aufbau

Formal kann die vollständige Transformation beschrieben werden als:

$$\begin{aligned}
 \vec{w} &= \mathcal{L}\mathcal{N}\mathcal{L}[\vec{u}] \\
 &= B^T \vec{v}^{\mathcal{NL}} = B^T \begin{bmatrix} \vec{v}^{ON} \\ \vec{v}^{OFF} \end{bmatrix} \\
 &= B^T \begin{bmatrix} \vec{\mathcal{N}}^+(\vec{v}) \\ \vec{\mathcal{N}}^-(\vec{v}) \end{bmatrix} = B^T \begin{bmatrix} \vec{\mathcal{N}}^+(A^T \vec{u}) \\ \vec{\mathcal{N}}^-(A^T \vec{u}) \end{bmatrix}.
 \end{aligned} \tag{4.14}$$

Hierbei bezeichnet $\vec{u} \in R^N$ das Eingangssignal, also die Luminanzwerte des zu kodierenden natürlichen Grauwertbildes, $\vec{w} \in R^{2N}$ die finalen nichtlinearen Koeffizienten, $A \in R^{N \times N}$ die Transformationsmatrix der initialen linearen Transformation und $B \in R^{2N \times 2N}$ die entsprechende Matrix der linearen Ausgangstransformation und es gilt weiterhin⁴

$$\begin{aligned}
 \text{Rang}(A) &= N \\
 \text{Rang}(B) &= 2N.
 \end{aligned} \tag{4.15}$$

$\vec{v}^{ON}, \vec{v}^{OFF}, \vec{\mathcal{N}}^+(\vec{v}), \vec{\mathcal{N}}^-(\vec{v})$ sind die intermediären nichtlinearen Koeffizienten sowie die bipolare Gleichrichtung wie in Abschnitt 4.2 beschrieben. Das zugehörige Netzwerk-Schema sowie die verwendeten Nichtlinearitäten sind in Abb. 4.4 dargestellt. Der Grundgedanke hinter dieser Architektur ist, daß sowohl die initiale als auch die finale lineare Transformation an die statistischen Eigenschaften der zu erwartenden Eingangssignale angepaßt sind. Als Initialtransformation wird hier eine Wavelet-Basis verwendet, da diese

⁴Dies gilt für vollrangige, nicht notwendigerweise orthonormale Transformationsmatrizen.

bekanntermaßen sehr gute Ergebnisse bezüglich Bildkodierung und -kompression liefert. Nach dieser festgelegten optimalen Transformation und Anwendung der Nichtlinearität werden Linearkombinationen dieser Koeffizienten, wie weiter unten beschrieben, mittels statistischer Optimierungsverfahren (PCA, ICA), ermittelt, welche die ordnungsreduzierten statistischen Abhängigkeiten weiter verringern.

Implementierung der Eingangsschicht

Da die Ergebnisse der ICA-Optimierung für natürliche Bilder hinreichend bekannt sind, wird hier auf eine explizite Neuberechnung der Gewichte verzichtet. Man beachte auch, daß eine blockweise lineare Transformation äquivalent zu einer Mehrkanalkodierung und anschließender Unterabtastung ist. Daher wird die initiale lineare Transformation aus Effizienzgründen als eine klassische Mehrkanalzerlegung basierend auf selbstähnlichen⁵, orientierungs- und frequenzselektiven Bandpaßfiltern mit anschließender Unterabtastung realisiert. Zur Vereinfachung der Abtastoperation wird mit dem *komplexen analytischen* Signal gerechnet um ein einseitiges Spektrum zu erhalten. Dazu werden Kombinationen aus gerade- und ungeradesymmetrischen Filtern verwendet, deren Ausgang dann als Real- und Imaginärteil angesehen werden können (Hauske und Zetzsche, 1990; Wegmann und Zetzsche, 1996). Die spektrale Übertragungsfunktion eines geradesymmetrischen Filters in Polarkoordinatendarstellung läßt sich schreiben als:

$$H_{even}(f_r, f_\phi) = \cos\left(\frac{\pi}{2} \frac{f_r - f_{r_0}}{2f_{r_{hbw}}}\right) \cos\left(\frac{\pi}{2} \frac{f_\phi - f_{\phi_0}}{2f_{\phi_{hbw}}}\right) \quad (4.16)$$

für:

$$\begin{aligned} f_{r_0} - 2f_{r_{hbw}} &\leq f_r \leq f_{r_0} + 2f_{r_{hbw}} \\ f_{\phi_0} - 2f_{\phi_{hbw}} &\leq f_\phi \leq f_{\phi_0} + 2f_{\phi_{hbw}} \\ f_{\phi_0} + \pi - 2f_{\phi_{hbw}} &\leq f_\phi \leq f_{\phi_0} + \pi + 2f_{\phi_{hbw}} \end{aligned}$$

sonst:

$$H_{even}(f_r, f_\phi) = 0$$

Hierbei sind f_{r_0} , f_{ϕ_0} die mittlere Radialfrequenz, bzw. mittlere Orientierung und $f_{r_{hbw}}$, $f_{\phi_{hbw}}$ die halbseitigen 3dB-Bandbreiten in Radial-, bzw. Orthogonalrichtung. Die Mittenfrequenz in Radialrichtung ist angegeben in Schwingungen/Pixel (*cycles/pixel*, c/p) und die Orientierung in Grad. Die Übertragungsfunktion für das im Ortsbereich ungeradesymmetrische Filter H_{odd} ist gegeben als die Hilbert-Transformierte der Übertragungsfunktion des geradesymmetrischen Filters. Diese Filterfunktionen sind, wie man Gl. (4.16) entnehmen kann, polar separierbar und besitzen eine cos-förmigen Abfall in

⁵Selbstähnlich bedeutet in diesem Kontext, daß sich die Filterfunktionen durch Rotation und Skalierung ineinander überführen lassen.

Schicht	$f_{r_0}/(c/p)$	$f_{r_{bw}}/(c/p)$	f_{ϕ_0}/deg	Abtastrate
1	0.25	0.08838828	-67.5 / -22.5 / 22.5 / 67.5	1:8
2	0.125	0.04419414	-67.5 / -22.5 / 22.5 / 67.5	1:32
3	0.0625	0.02209707	-67.5 / -22.5 / 22.5 / 67.5	1:128
4	0.03125	0.01104854	-67.5 / -22.5 / 22.5 / 67.5	1:512
5	0.015625	0.00552427	-67.5 / -22.5 / 22.5 / 67.5	1:2048
TP	0.015625	0.0055243	isotrop	1:1024

Tabelle 4.1: Parameter und Abtastraten der linearen Filterbank.

Radial- und Orthogonalrichtung. Ein Vorteil dieser Funktionen ist die Möglichkeit der exakten Kontrolle der Alias-Terme für örtlich abgetastete Signale, da die spektralen Übertragungsfunktionen außerhalb des $2f$ -Intervalls exakt zu Null werden.

Verwendet wird eine Filterbank mit fünf unterschiedlich auflösenden orientierungsselektiven Kanälen, mit jeweils einer Oktave Radialbandbreite und $22,5^\circ$ Orientierungsbandbreite. Die genauen Parameter sind Tabelle 4.1 zu entnehmen. Der tieffrequente Signalanteil inklusive Gleichanteil wird von einem Tiefpaß erfaßt, der wie die orientierungsselektiven Filter einen cos-Abfall in exzentrische Radialrichtung besitzt, zum Ursprung hin allerdings ab der Mittenfrequenz konstant auf eins gesetzt ist. Ein weiterer Aspekt dieser linearen Filterbank ist die anschließende Unterabtastung des analytischen Signals. Die Unterabtastung wird vektoriell durchgeführt, d.h. es werden 8×8 Pixel Blöcke mit einem speziellen Raster, gegeben durch eine Abtastmatrix, unterabgetastet. Gegenüber einer standardmäßigen kartesischen Abtastung birgt dies den Vorteil, daß die Wiederholerspektren dichter zu liegen kommen, ohne den Alias-Fehler zu erhöhen (Wegmann und Zetzsche, 1996). Die Unterabtastung erfolgt in der höchstauffösenden Schicht (1) mit einer Rate von 1 : 8. Die Raten der tieferliegenden Schichten ergeben sich aufgrund des Oktavabstandes der Mittenfrequenzen zu jeweils einer um den Faktor vier reduzierten Rate. Siehe dazu auch Tabelle 4.1. Die Signalverdopplung durch die Bildung des analytischen Signals hebt sich hier insofern wieder auf, als daß auch die Abtastraten halbiert werden können.

Verknüpfung der nichtlinearen Koeffizienten

Zur Optimierung der Basisvektoren der finalen linearen Transformation, also der Gewichte der nichtlinearen intermediären Koeffizienten, würden idealerweise sämtliche Koeffizienten aus einem Block über alle Skalen, örtliche Positionen, Orientierungen, Filtersymmetrien und Polaritäten hinweg zu einem Zufallsvektor zusammengefaßt werden. Dabei muß allerdings beachtet werden, daß die Dimensionalität des zu optimierenden Raumes immense Ausmaße annimmt. Bei fünf Skalen à vier Orientierungen, zwei Filtersymmetrien und der Aufspaltung in ON- und OFF-Anteile ergäbe sich somit die

Dimensionalität einer Realisierung zu

$$16 \sum_{i=0}^4 (4^i \cdot 8) = 16 \cdot 2728 = 43648.$$

Aus diesem Grund werden die nichtlinearen Koeffizienten von jeweils nur zwei Skalen zusammengefaßt. Es werden 32 örtliche Koeffizienten pro Filter in Schicht (1) und acht in Schicht (2) kombiniert. Zusammen mit vier Orientierungen, zwei Symmetrien, 2 Polaritäten ergibt dies eine Dimensionalität von $(32 + 8) \times 4 \times 2 \times 2 = 640$, welche mit standardmäßigen ICA-Algorithmen handhabbar ist. Ebenso wird mit den Schichten (3) und (4) Verfahren, Schicht (5) wird einzeln behandelt (512 Koeffizienten) und der Ausgang des Tiefpasses direkt skalar quantisiert.

4.3.2 Statistische Optimierung der Ausgangsschicht

Nach der oben beschriebenen Zusammenfassung der nichtlinearen Koeffizienten können weitere statistische Optimierungsverfahren angewendet werden, um die Abhängigkeiten zwischen diesen zu eliminieren, respektive zu reduzieren. Es wurden sowohl die Karhunen-Loève-Transformation (PCA) als auch unterschiedliche Implementierungen des ICA-Algorithmus untersucht, da diese jeweils spezifische Vor- und Nachteile mit sich bringen. Als für die Kodierung besser geeignet hat sich die PCA herausgestellt. Darum werden im Folgenden die Ergebnisse, die unter Verwendung der ICA-Algorithmen erhalten wurden, relativ kurz zusammengefaßt.

Optimierung mittels PCA

Zur Ausnutzung von statistischen Abhängigkeiten zweiter Ordnung ist die optimale Transformation durch die PCA gegeben. Diese transformiert die nichtlinearen Koeffizienten $v_i^{\mathcal{NL}}$ in Koeffizienten $w_k = \sum_{(j)} b_{kj} v_i^{\mathcal{NL}}$ welche vollständig dekorreliert sind. Die korrespondierenden Eigenvektoren \vec{b}_k und Eigenwerte λ_k der Kovarianzmatrix

$$C = E\{(\vec{v}^{\mathcal{NL}} - \vec{m}^{\mathcal{NL}})(\vec{v}^{\mathcal{NL}} - \vec{m}^{\mathcal{NL}})^T\}, \quad (4.17)$$

mit $\vec{m}^{\mathcal{NL}} = E\{\vec{v}^{\mathcal{NL}}\}$, legen die Linearkombinationen der Koeffizienten fest, die zu unkorrelierten Ausgangsgrößen führen und gleichzeitig die Verteilung der Gesamtvarianz auf wenige Koeffizienten konzentrieren (Papoulis, 1991). Wie bereits in (Nuding und Zetzsche, 2003) gezeigt, führt die Linearkombination von augenscheinlich „schwach nichtlinearen“ Koeffizienten $\vec{v}^{\mathcal{NL}}$ zu Ausgangskoeffizienten w_k mit ausgeprägten nichtlinearen Eigenschaften. Abbildung 4.5 zeigt die Struktur vier typischer Eigenvektoren und das assoziierte nichtlineare Antwortverhalten für die Anwendung der Optimierung anhand

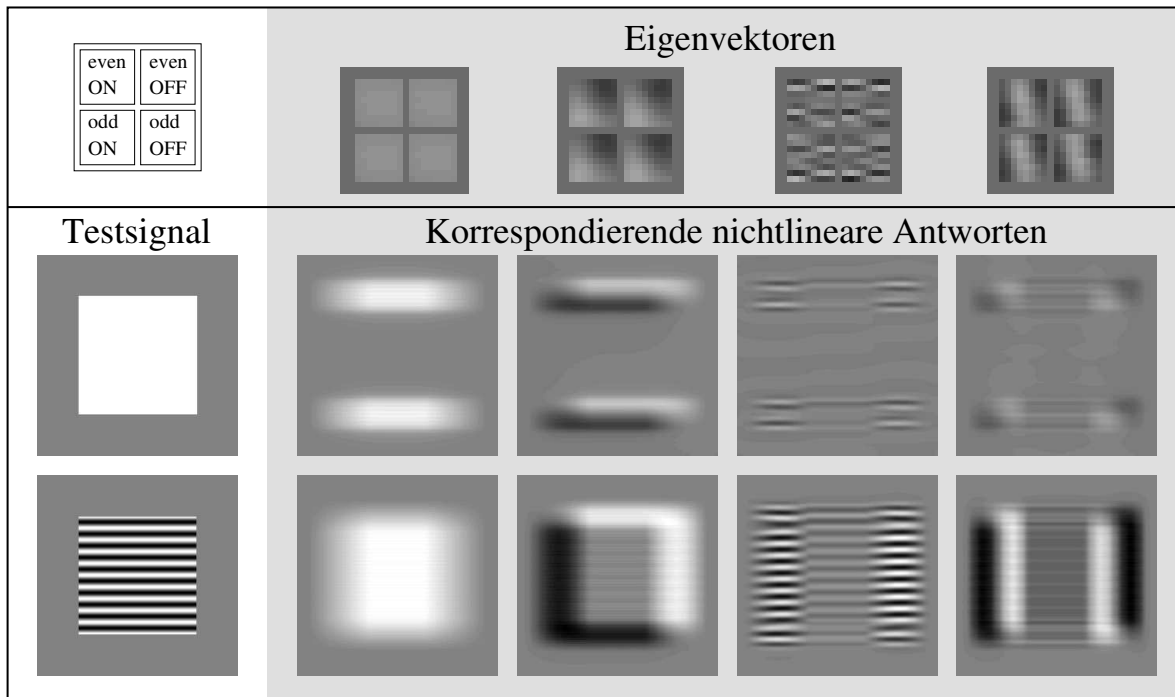


Abbildung 4.5: Beispiel zur Anwendung der PCA auf die gleichgerichteten Filterkoeffizienten einer reduzierten Konfiguration (Details siehe Text). Obere Reihe: Vier typische Eigenvektoren (vergrößert) und Darstellung der inhärenten Aufteilung der Gewichte, die auf die jeweiligen Symmetrien (even/odd) und Polaritäten (ON/OFF) appliziert werden. Untere Reihe: Die Antworten der nichtlinearen Einheiten auf die Testsignale links. Nach (Nuding und Zetsche, 2003).

einer reduzierten Konfiguration. Bei dieser wurden lediglich Koeffizienten einer Skala und Orientierung, aber von benachbarten örtlichen Positionen und von beiden Symmetrien (even/odd) und Polaritäten (ON/OFF) verknüpft. Die Antworten weisen substantielle nichtlineare Eigenschaften auf. Einige zeigen demodulierende Eigenschaften (konstante Antwort auf sinusförmigen Eingang) und extrahieren eine Art „lokaler Energie“. Andere wiederum besitzen eine erhöhte Selektivität für zweidimensionale im Gegensatz zu eindimensionalen Merkmalen, d.h. stärkere Antwort auf Linienenden und Ecken als auf gerade Kanten. Diese kann man somit als eine Art *i2D*-Operatoren beschreiben, vgl. Abschnitt 2.2 und (Krieger und Zetsche, 1996). Ein derartiger Operator ist notwendigerweise nichtlinear und kann nicht durch lineare Transformationen approximiert werden (Zetsche und Barth, 1990a). Ebenfalls zu beobachten sind Einheiten, die diese beiden Merkmale kombinieren (vgl. auch die Diskussion in Abschnitt 4.6).

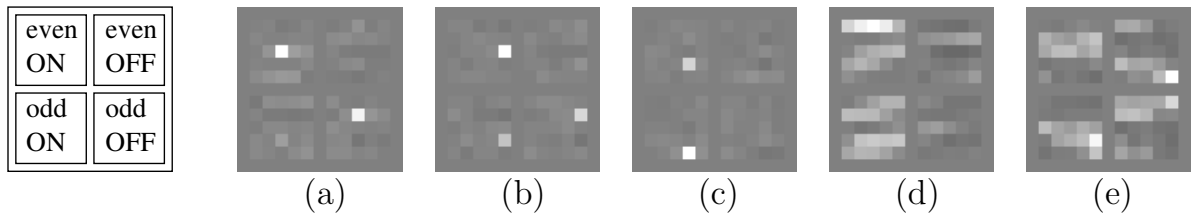


Abbildung 4.6: Typische ICA Basisfunktionen. Die meisten der gefundenen Funktionen weisen eine degenerierte Struktur auf (a-c) und lediglich ein kleiner Anteil zeigt ein gewisses Maß an Interaktion zwischen den nichtlinearen Koeffizienten (d,e).

Optimierung mittels ICA

Ein zweiter getesteter statistischer Optimierungsalgorithmus ist die ICA. Verwendet wurde der Algorithmus von OLSHAUSEN, modifiziert durch HOYER UND HYVÄRINEN (Hoyer und Hyvärinen, 2002). In den hier durchgeführten Untersuchungen hat sich allerdings kein Vorteil dieses Algorithmus für die Anwendung zur Bilddatenkompression gezeigt. Einige typische Basisfunktionen sind exemplarisch in Abb 4.6 dargestellt. Viele der gefundenen Basisfunktionen weisen eine starke Konzentration der Gewichte auf einzelne wenige nichtlineare Koeffizienten auf, was bedeutet, daß die Optimierung die Wavelet-Basis bereits als Optimum „sieht“ und die Koeffizienten durchreicht. Bei Verwendung einer übervollständigen, nichtorthogonalen Basis erhöht sich der Anteil der integrierenden Basisfunktionen, ein Hinweis darauf, daß die Verbunddicke durch hochspezialisierte Einheiten approximiert wird. Dieser Mechanismus hängt eng zusammen mit der in Kap. 3, Abschnitt 3.3 beschriebenen nichtlinearen Partitionierung des Signalraums.

4.3.3 Quantisierung

Der finale Verarbeitungsschritt des hier implementierten Bildkodierers ist die Quantisierung der Ausgangskoeffizienten. Die Quantisierung erfolgt nichtlinear, unter Verwendung einer quadratwurzelförmigen Kompander-Kennlinie. Da die Koeffizienten nicht notwendigerweise dieselben Statistiken haben, müssen die Quantisierungsparameter für jeden einzelnen der 640 Koeffizienten separat bestimmt werden. Dabei werden für jeden Koeffizienten w_i die positiven und negativen Quantisierungsbereiche q_{pos} und q_{neg} sowie die Anzahl der Quantisierungsstufen γ_{pos} und γ_{neg} in Abhängigkeit ihrer Eigenwerte λ_i bestimmt. Dies ist eine suboptimale Näherung, da die Eigenwerte, im Gegensatz zu linearen Systemen, nicht notwendigerweise identisch den Rekonstruktionsfehlern sind, die sich bei Vernachlässigung des entsprechenden Koeffizienten ergeben (siehe Abschnitt 4.4.1). Der Rekonstruktionsfehler wird bei der Bitallokation insofern berücksichtigt, als daß Koeffizienten mit sehr geringem assoziierten Fehler von vorneherein vernachlässigt werden. Für die Bestimmung der Quantisierungsparameter der einzelnen Koeffizienten

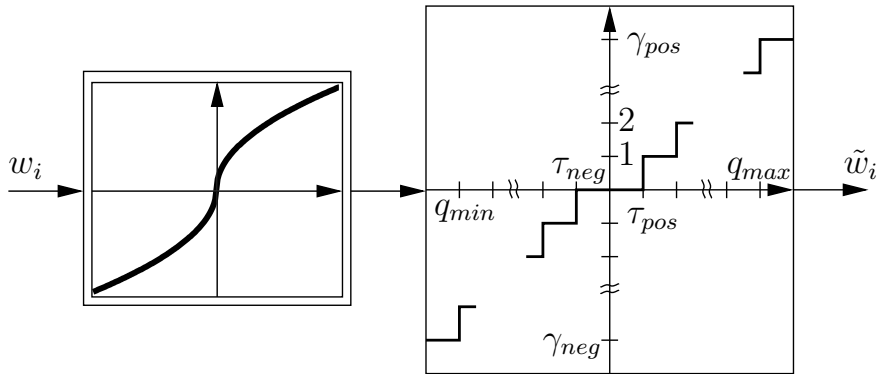


Abbildung 4.7: Schema des implementierten Quantisierers.

gilt im Detail also der folgende Zusammenhang:

$$\begin{aligned}
 \tilde{w} &= \mathcal{Q}[\vec{w}; q_{pos}, q_{neg}, \gamma_{pos}, \gamma_{neg}, \tau_{pos}, \tau_{neg}] \\
 q_{pos} &= c_1 * \sqrt[4]{\lambda_i} \\
 q_{neg} &= -q_{pos} \\
 \gamma_{pos} &= c_2 * \sqrt[4]{\lambda_i} \\
 \gamma_{neg} &= \gamma_{pos} \\
 \tau_{pos} &= c_3 * q_{max} \\
 \tau_{neg} &= c_3 * q_{min},
 \end{aligned} \tag{4.18}$$

wobei die c_i empirisch bestimmte Konstanten und τ_{pos}, τ_{neg} Quantisierungsschwellen um die Null sind. Ein Schema des Quantisierers ist in Abb. 4.7 dargestellt.

4.4 Rekonstruktion und Fehler

4.4.1 Rekonstruktion

Durch die oben beschriebene einfache Architektur des Systems ist eine Rekonstruktion des Signals \vec{u} aus den nichtlinearen Koeffizienten \vec{w} , bzw. den quantisierten Koeffizienten \tilde{w} möglich. Erstens ist die lineare Ausgangstransformation $B : \vec{v}^{\mathcal{NL}} \rightarrow \vec{w}$ für den Fall einer vollständigen Transformationsbasis perfekt invertierbar. Doch selbst wenn einige der finalen Koeffizienten vernachlässigt werden, erhält man Approximationen $\hat{v}_i^{ON}, \hat{v}_i^{OFF}$ der intermediären nichtlinearen Koeffizienten. Desweiteren können die linearen Wavelet-Koeffizienten aus diesen mit der Vorschrift $\hat{v}_i = \hat{v}_i^{ON} - \hat{v}_i^{OFF}$ berechnet werden. Schließlich kann der Eingang \hat{u} aus den geschätzten Wavelet Koeffizienten

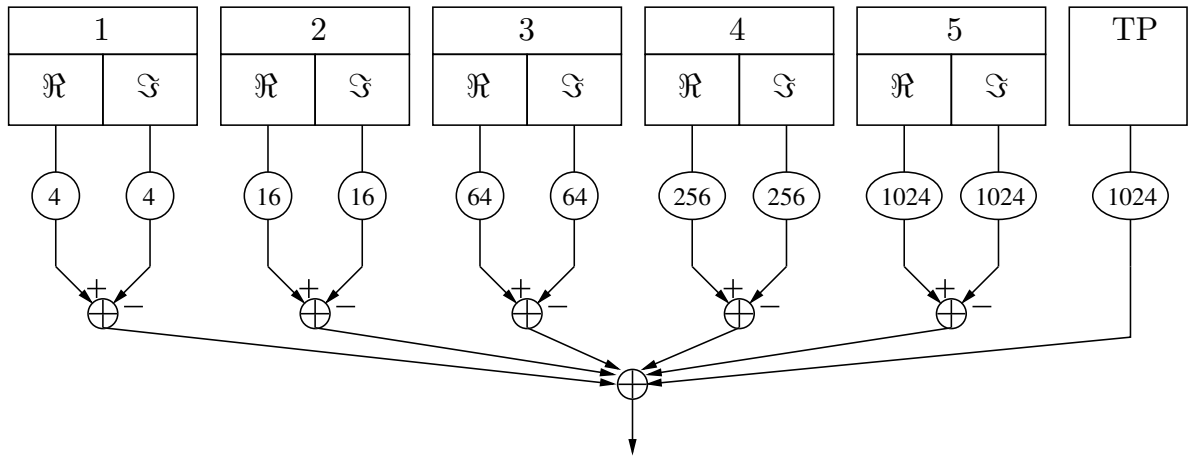


Abbildung 4.8: Schema der Rekonstruktion des Eingangssignals aus den linearen Wavelet-Koeffizienten. Diese einfache Vorschrift ergibt sich aus der speziellen Form der verwendeten Filterfunktionen und dem verwendeten vektoriellen Abtastverfahren.

\hat{v}_i mit typischen Verfahren rekonstruiert werden. Aufgrund der speziellen Implementierung der Eingangsschicht mit Filtern, die einen cos-förmigen Abfall in Radial- und Orthogonalrichtung besitzen, und dem vektoriellen Abtastverfahren, vereinfacht sich die Rekonstruktion zu einer einfachen gewichteten Summation der gerade- und ungeradesymmetrischen Subbänder (Real- und Imaginärteil), skizziert in Abb. 4.8.

Der zu erwartende Kodiergewinn kann bei linearen Systemen und Gaußschen Signalen aus der Struktur der Eigenwertverteilung mittels der *Spectral Flatness Measure* berechnet werden, siehe Gl. (2.25). Für eine möglichst steil abfallende Verteilung ist der zu erwartende Gewinn hoch, für eine flache Verteilung ist kein besonderer Kodiergewinn zu erwarten. Diese Regel kann jedoch nicht ohne weiteres auf das hier betrachtete nichtlineare System angewendet werden. Denn zum Einen korrespondieren die Eigenwerte nicht vollständig mit den assoziierten Rekonstruktionsfehlern, wie in Abb. 4.9 zu sehen ist, und zum Anderen gilt das Superpositionsprinzip der Fehler im nichtlinearen Fall nicht notwendigerweise. Das bedeutet, daß unter Vernachlässigung zweier Koeffizienten w_i und w_k mit zugehörigen Rekonstruktionsfehlern ϵ_i und ϵ_k der Gesamtfehler ϵ_{ik} im Allgemeinen ungleich der Summe der Einzelfehler ist:

$$\epsilon_{ik} \neq \epsilon_i + \epsilon_k. \quad (4.19)$$

Nichtsdestoweniger weisen beide Verteilungen einen steil abfallenden Verlauf auf⁶ und die Eigenwerte korrelieren mit den Rekonstruktionsfehlern mit einem Korrelationskoeffizienten von $\rho = 0.81$. Ein erster Test des nichtlinearen Codecs ist die sukzessive

⁶Es wurde auch der Verlauf der Rekonstruktionsfehler nach der ICA-Optimierung der Gewichte untersucht, hierbei ergab sich jedoch ein wesentlich flacherer Verlauf und somit ein geringer zu erwartender Kodiergewinn.

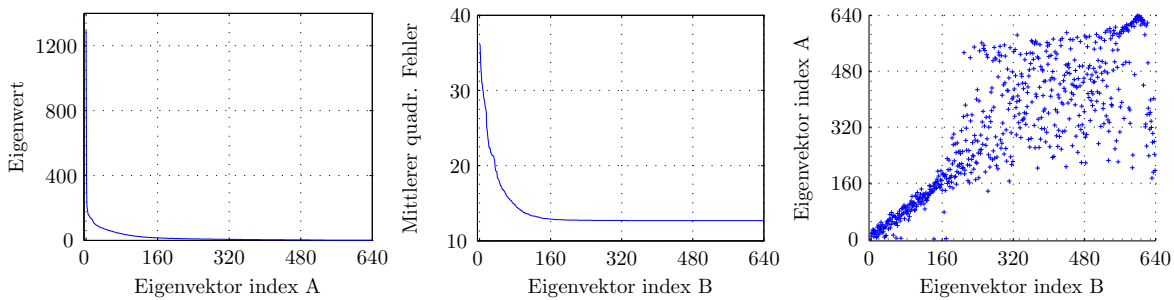


Abbildung 4.9: Eigenwertverteilung (links) und Verteilung der Rekonstruktionsfehler (mittig). Beide Verteilungen sind in absteigender Reihenfolge angeordnet, und zeigen einen starken Abfall, indizieren also einen möglichen Kodiergewinn. Im Gegensatz zum linearen Fall ist die Indizierung der beiden Verteilungen unterschiedlich, wie anhand des rechten Streudiagramms leicht zu erkennen ist.

Rekonstruktion aus den Koeffizienten w_i . In Abb. 4.10 ist das Ergebnis für drei unterschiedliche Anteile an verwendeten Koeffizienten gezeigt. Dieser Test indiziert, daß es möglich sein sollte, einen Kodiergewinn durch Vernachlässigung einiger Koeffizienten w_i und varianzabhängiger Bitallokation, wie in Abschnitt 4.3.3 beschrieben, zu erreichen.

4.4.2 Vergleich mit linearem Referenz-Codec

Das *experimentum crucis* für die hier vorgestellte Architektur besteht natürlich darin, einen *Rate-Distortion*-Vergleich mit einem linearen Referenz-Codec durchzuführen. Verwendet wird als Referenz eine lineare Mehrkanalzerlegung mit Wavelets, entsprechend der ersten Schicht der vorgestellten Architektur. Die Koeffizienten des linearen Codecs werden gleichförmig quantisiert und die Bitraten ergeben sich durch die Anzahl der Quantisierungsstufen multipliziert mit den Abtastraten. Um eine realistische Schätzung der endgültigen Bitraten zu erhalten, müßte abschließend ein Entropie-Koder implementiert werden. Doch da an dieser Stelle nur der Vergleich mit einem Referenz-Codec interessiert und nicht etwa ein quantitativer Vergleich mit populären Codecs wie JPEG-2000, genügt diese vereinfachte Schätzung der Bitraten völlig.

Es werden nun die beiden Extrema des Qualitätsspektrums untersucht, der Bereich hoher Qualität mit keinen oder kaum sichtbaren Fehlern und der Bereich extrem niedriger Bitraten. Es zeigt sich, daß der nichtlineare Codec im Hochqualitätsbereich eine niedrigere Bitrate zum Erreichen einer ähnlichen Güte des rekonstruierten Bildes benötigt als der lineare. Im Niedrigqualitätsbereich, bzw. Hochkompressionsbereich kann der lineare Codec eine gewisse Bitrate nur durch substantiellen Verlust der hochfrequenten Information erreichen, wohingegen der nichtlineare Codec selbst bei niedrigsten Bitraten einen Großteil dieser Information erhalten kann (Abb. 4.11). Die hier auftretenden Fehler ergeben ein vergleichbares *Peak Signal-to-Noise Ratio*, *PSNR*, der subjektive Qualitätsunterschied jedoch zeigt die Vorteile des nichtlinearen Codecs auf.



Abbildung 4.10: Sukzessive Rekonstruktion mit zunehmender Anzahl an nichtlinearen Koeffizienten. Die Koeffizienten sind entsprechend Abb. 4.9 (Mitte) nach absteigendem Rekonstruktionsfehler angeordnet. Man beachte, daß die Aufspaltung des Signals in ON- und OFF-Anteile zu einer Verdopplung der Koeffizienten führt und somit ein Rekonstruktionsanteil von 50% einer $n \rightarrow n$ Abbildung entspricht.

4.4.3 Klassifizierung lokaler Merkmale und assoziierte Rekonstruktionsfehler

Eine bislang noch nicht untersuchte Eigenschaft des Systems ist, daß die finalen nichtlinearen Koeffizienten einen größeren Bildbereich erfassen, als die linearen Wavelet-Koeffizienten. So ist der Support der nichtlinearen Koeffizienten, die sich auf die Wavelet-Schichten (1) und (2) beziehen, 16×16 Pixel. Eine interessante Frage hierbei ist, inwiefern dieser vergrößerte Support der Klassifizierung unterschiedlicher lokaler Bildstrukturen dienlich sein kann. Dafür wird zunächst ein Klassifikator implementiert, der jedem Block eine der drei Klassen $\{Null, Kontur, Textur\}$ zuordnet. Der hier verwendete Klassifikator arbeitet nach folgenden Regeln:

1. Markiere Block als *Null*, wenn gilt:
 - Es liegen genügend lineare Koeffizienten der Wavelet-Repräsentation in jeder Orientierungen und Skala unterhalb einer Schwelle.
2. Markiere Block als *Kontur*, wenn alle der folgenden Bedingungen erfüllt sind:
 - Die Anzahl der signifikanten Gradientenwerte liegt pro Block in einem gewissen Intervall. Diese wird berechnet, indem auf jeden Block eine binäre Schwellwertoperation angewendet und darauf der Gradient berechnet wird. Die Schwelle ergibt sich aus dem Median der Pixelwerte des jeweiligen Blocks. Für 16×16 Pixel Blöcke liegt die Anzahl für eine prototypische Kante bei exakt 16.
 - Die Anzahl der Nachbarn im oben beschriebenen Gradientenbild liegt pro



Abbildung 4.11: Vergleich der Kodierleistung des linearen Referenz-Codec mit der LNL-Architektur.

Block in einem gewissen Intervall. Dies ist notwendig, um unzusammenhängende Texturen von Kanten zu unterscheiden.

- Die Anzahl der aktiven Koeffizienten in mindestens einer Orientierung der höchstaflösenden Schicht der Wavelet-Repräsentation liegt in einem gewissen Intervall.

3. Markiere Block als *Textur*, wenn gilt:

- Die Anzahl der aktiven Koeffizienten in mindestens einer Orientierung und einer Auflösungsschicht übersteigt einen gewissen Wert.

Da bei diesem Klassifikationsschema auch mehrfache Klassifizierungen unter Umständen möglich sind, werden in einem finalen Schritt, unter Berücksichtigung der Prioritätenreihenfolge $Null \rightarrow Kontur \rightarrow Textur$, die Klassifikatorausgänge zusammengeführt.

Es können nun für jede Klasse unterschiedliche Kodierungs- und Quantisierungsparameter gewählt werden. Die typischen Fehler, die mit diesen Klassen assoziiert sind, sind in Abb. 4.12 gezeigt. Sowohl die *Null*- als auch die *Textur*-Klasse zeigen einen graduellen Übergang der Qualität bei zunehmend stärkerer Kompression. Bei den als *Null* klassifizierten Blöcken sind die nichtlinearen Koeffizienten nahe bei Null, können also zu Kompressionszwecken auf Null gesetzt werden. Bei Erhöhung der Toleranzgrenze werden zunehmend mehr Blöcke als *Null* klassifiziert, damit verschwinden feingranulare Strukturen und schließlich auch Kanten mit niedrigem Kontrast (Abb. 4.12(a)). Auch die Quantisierung der *Textur*-Blöcke erweist sich als robust gegenüber Rekonstruktionsfehlern. Obwohl in dieser Klasse mehr signifikant von Null verschiedene Koeffizienten als in der *Kontur*-Klasse sind, kann hier eine wesentlich gröbere Quantisierung verwendet werden. Erst bei einer sehr groben Quantisierung werden die sichtbaren Strukturen merklich beeinträchtigt, siehe Abb. 4.12(b). Ein eher bistabiles Verhalten zeigt die *Kontur*-Klasse. Zwar sind weniger Koeffizienten signifikant von Null verschieden als in der *Textur*-Klasse, jedoch erweist sich diese Klasse gegenüber Kompression, also Quantisierung und Vernachlässigung einiger Koeffizienten, als wenig fehlertolerant. Bereits bei geringer Kompression verschwinden kontrastarme Konturen und Rekonstruktionsfehler entstehen, die dazu führen, daß konturartige Strukturen stark beeinträchtigt werden und ein texturähnliches Aussehen annehmen.

Wenn nun diese empirischen Ergebnisse dazu verwendet werden, den Codec zu optimieren, indem Quantisierungsparameter entsprechend den Klassen angepaßt werden, so kann eine weitere Verbesserung des *Rate-Distortion*-Verhältnisses erreicht werden. Beispielsweise kann die subjektive Qualität der Hochqualitätsbilder aus Abb. 4.11 mit einer signifikant niedrigeren Bitrate erreicht werden (Abb. 4.13).

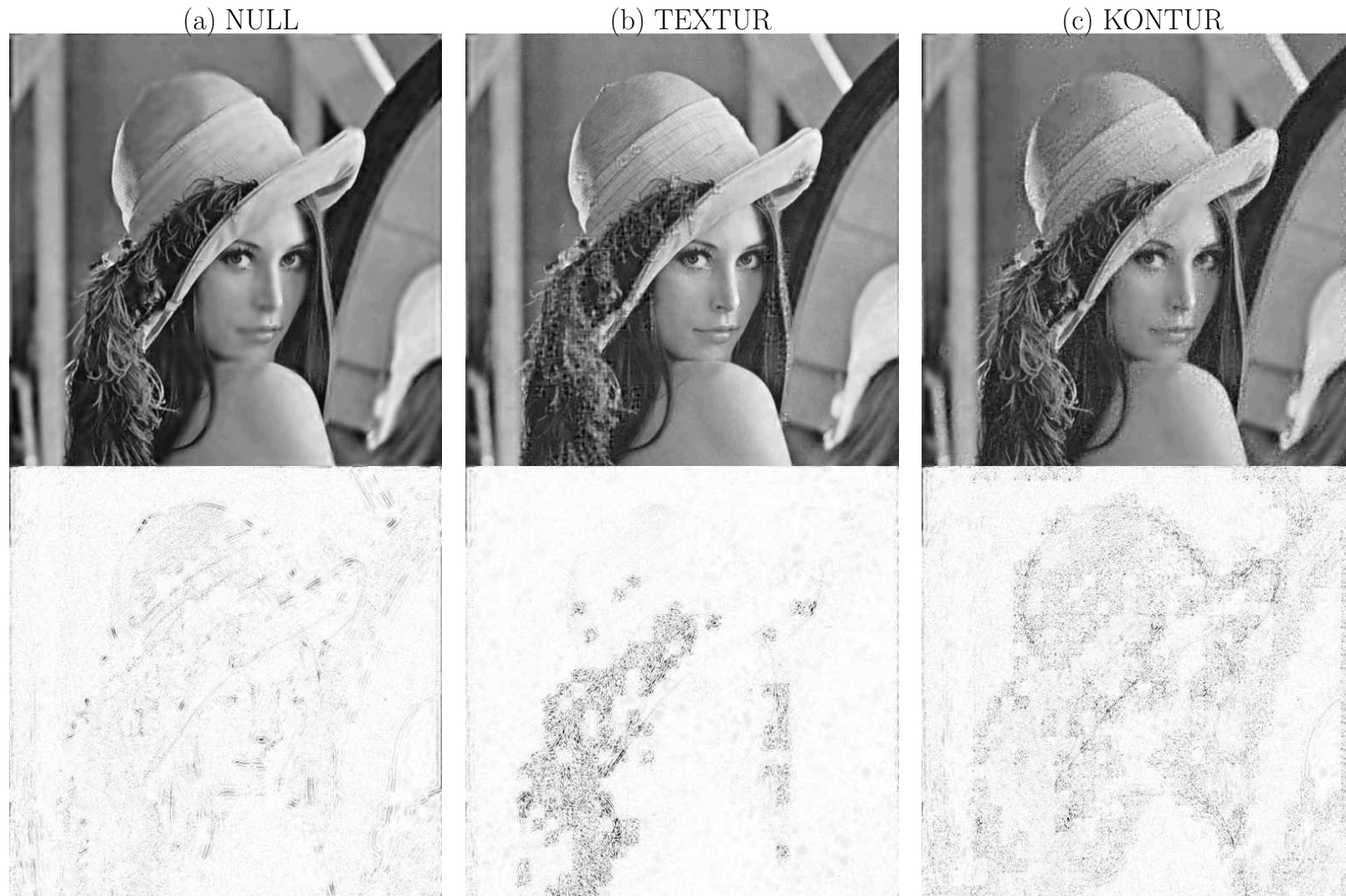


Abbildung 4.12: Typische Rekonstruktionsfehler der drei Merkmalsklassen. Es wurde die jeweils benannte Klasse stark komprimiert und die anderen fehlerfrei kodiert, um die Effekte zu trennen. Oben: Rekonstruierte Bilder, unten: Differenzbilder zwischen Original und Rekonstruktion.



Abbildung 4.13: Nichtlinearer Codec mit zusätzlichem Klassifikations-Schema. Im Vergleich zu den Hochqualitätsbildern aus Abb. 4.11 kann die Bitrate ohne merklichen Qualitätsverlust um ca. 50% weiter reduziert werden.

4.5 Nichtlineare Frequenzinteraktionen

Wie bereits in den vorigen Kapiteln ausgeführt wurde, ist ein linearer Ansatz nicht ausreichend um die Redundanz in natürlichen Bildern zu eliminieren. Basierend auf einer detaillierten Analyse mittels Spektren höherer Ordnung und der Anwendung der Volterra-Reihenentwicklung (Kapitel 5) kann gezeigt werden, daß eine optimale Anpassung an die Statistik natürlicher Bilder eine UND-artige Verknüpfung von Frequenzkomponenten benötigt (Zetzsche und Nuding, 2005b). Dementsprechend ist an dieser Stelle von Interesse, ob das hier vorgestellte Netzwerk in der Lage ist, Frequenzinteraktionen aus der Statistik heraus zu *lernen*.

Eine direkte Methode um die Frequenzinteraktionen zu quantifizieren besteht darin, das gemeinsame Gewichten (*Pooling*) von Frequenzkomponenten der Basisvektoren der zweiten Schicht zu analysieren. Daher werden hier Frequenz-Pooling Indizes (*FPI*) definiert, gegeben als das Produkt der Energien aus der hochfrequenten Skala (E_{hi}) und der niederfrequenten Skala (E_{lo}) für unterschiedliche Orientierungen θ_i :

$$FPI_{iso} = 4 \cdot \left(\sum_{\Theta=\theta_1}^{\theta_4} E_{lo,\Theta} \cdot E_{hi,\Theta} \right) \quad (4.20)$$

$$FPI_{cross} = 4 \cdot \left(\sum_{\Theta_1=\theta_1}^{\theta_4} \sum_{\substack{\Theta_2=\theta_1 \\ \Theta_2 \neq \Theta_1}}^{\theta_4} E_{lo,\Theta_1} \cdot E_{hi,\Theta_2} \right). \quad (4.21)$$

Diese Indizes sind eins, wenn genau zwei Orientierungen von unterschiedlichen Skalen mit der gleichen Energie $E_{hi} = E_{lo} = 0.5$ gewichtet werden und Null, wenn kein Zusammenfassen stattfindet. Gl. (4.20) berücksichtigt nur Pooling von unterschiedlichen Radialfrequenzen mit gleicher Orientierung (*Iso-Orientierung*), wohingegen Gl. (4.21) ausschließlich Pooling über unterschiedliche Orientierungen hinweg berechnet (*Kreuz-Orientierung*). Man beachte, daß dieser Index ein *orientierungsselektives* Pooling bevorzugt, d.h. gleichartiges Gewichten mehrerer Orientierungen führt zu einer Reduktion des Index. Diese Indizes werden für jeden Ausgangskoeffizienten berechnet und die entsprechenden Energieterme eines Koeffizienten i sind gegeben durch

$$E_{lo|hi}^{(i)} = \sum_{x,y,P,S} [b_i(x,y,P,S, f_{lo|hi})]^2, \quad (4.22)$$

wobei b_i den Basisvektor des Koeffizienten i bezeichnet und die Summation über sämtliche örtlichen Positionen x, y , Polaritäten P und Symmetrien S erfolgt. f_{hi}, f_{lo} im Argument unterscheidet, ob es sich um Gewichte der hoch-, bzw. niederfrequenten Skala handelt. In Abb. 4.14(a,b) ist das Iso- und Kreuz-Orientierungs Pooling der gelernten Einheiten als „Populationshistogramme“ gezeigt. Wie diese Histogramme zeigen, lernt das System gemeinsames Gewichten für die meisten Einheiten in einer Orientierung, als auch zwischen unterschiedlichen Orientierungen. Dies ist in Übereinstimmung mit dem Konzept von *i2D*-selektiven Einheiten, da diese Selektivität Interaktionen zwischen unterschiedlichen Orientierungen fordert (siehe auch Kap. 5). Nichtsdestotrotz kann dieses Maß nur eine ungefähre Vorstellung über die wahren Frequenzinteraktionen geben, da diese auch von der exakten Konfiguration der örtlichen Anordnung abhängen. Eine aussagekräftige Analyse diesbezüglich ist die Messung der Antworten auf zwei Sinusgitter mit unterschiedlicher Frequenz und Orientierung sowie auf die Kombination (*Plaid-Muster*). Der normalisierte Frequenz-Interaktions Index für jeden Koeffizienten ergibt sich somit zu

$$FII = \left| \frac{R_{12}}{R_1 + R_2} - 1 \right|, \quad (4.23)$$

wobei R_1 und R_2 die Einzelantworten auf die Gitter und R_{12} die Antwort auf die Überlagerung darstellen. Diese Analyse zeigt, daß die meisten Einheiten substantielle Frequenzinteraktionen aufweisen, Abb. 4.14(c). Interessanterweise ist dieser Effekt für die PCA-optimierten Einheiten deutlicher als für die ICA-Optimierung.

4.6 Biologische Relevanz

Das hier vorgestellte und untersuchte System erhebt nicht den Anspruch, ein vollständig realistisches Modell des visuellen Systems zu sein. Jedoch sind wesentliche Merkmale und Prinzipien implementiert, die, sofern sie nicht bereits durch empirische neurophysiologische Untersuchungen verifiziert wurden, weitere Evidenz für eine Anpassung des

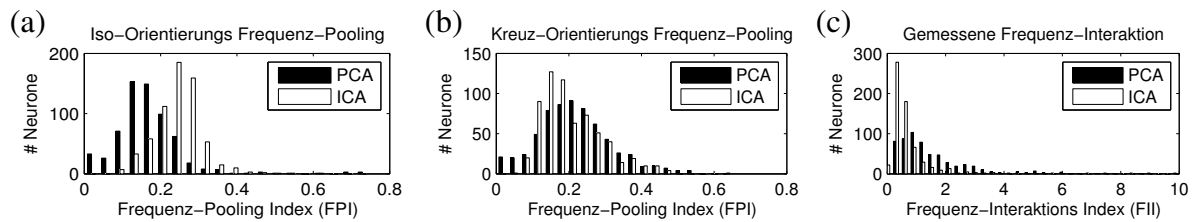


Abbildung 4.14: Nichtlineare Frequenzinteraktionen. Gezeigt sind die Populationshistogramme der pooling-basierten (a,b) als auch der gemessenen Frequenzinteraktionen (c). Die hier verwendete ICA ist aus rechentechnischen Gründen die schnelle FastICA 2.5 (Hyvärinen, 1999).

visuellen Systems an die der Umwelt inhärenten Strukturen liefern. Über die genaue Implementierung der hier vorgeschlagenen Funktionalität in der kortikalen „Hardware“ kann das System keine Aussage treffen. Die Überlegenheit einer mehrschichtigen *nicht*-linearen Verarbeitung gegenüber dem linearen Standardansatz konnte in diesem Kapitel allerdings gezeigt werden.

Die Eingangsschicht des Systems besteht aus linearen Wavelets, ein weithin verbreitetes Modell für kortikale *Simple*-Zellen, wie bereits in Abschnitt 2.6 besprochen. Die Aufspaltung in ON- und OFF-Anteile ist ein in der Biologie grundlegendes Prinzip, da Neurone keine negativen Signalanteile kodieren können und somit gezwungen sind, das Vorzeichen separat zu kodieren⁷. Dies geschieht mit den Mechanismen der Inhibition und Exzitation, einer biologischen Implementierung der bipolaren Einweggleichrichtung. Die Verwendung des analytischen Signals und damit einhergehend Quadratur-Filterpaaren ist insofern von biologischer Relevanz, als daß bereits empirisch nachgewiesen wurde, daß benachbarte Simple-Zellen im visuellen Kortex genau diese Eigenschaft besitzen (Pollen und Ronner, 1981).

Eine Architektur des visuellen Kortex im Sinne eines hierarchischen, mehrschichtigen Systems wurde bereits von HUBEL UND WIESEL (Hubel und Wiesel, 1965) vorgeschlagen. Dabei entstehen komplexe Zellen mit zunehmend ausgedehnten rezeptiven Feldern und erhöhter Merkmalsselektivität durch sukzessive Verschaltung von einfacheren Einheiten. Ein derartiges Prinzip wurde auch im HMAX-Modell (Riesenhuber und Poggio, 1999) zur invarianten Objekterkennung implementiert. Hierzu wurden modellierte *Simple*-Zellen, selektiv für unterschiedliche örtliche Positionen und Skalen, mittels eines Maximum-Prinzips kombiniert und resultierten in *Complex*-Zellen mit Invarianzeigenschaften bezüglich Größe und Position. Das hier vorgestellte LNL-System läßt wesentlich mehr Freiheitsgrade offen, da lediglich die basale Architektur und das Optimierungskriterium, nicht jedoch die Verschaltungen per se festgesetzt sind. Die resultierenden „Zellen“, wie die in Abb. 4.5 exemplarisch dargestellten, weisen ähnliche Eigenschaften auf

⁷Typischerweise kodieren Neurone das Signal in Feuerraten, also die Frequenz aufeinanderfolgender Aktionspotentiale.

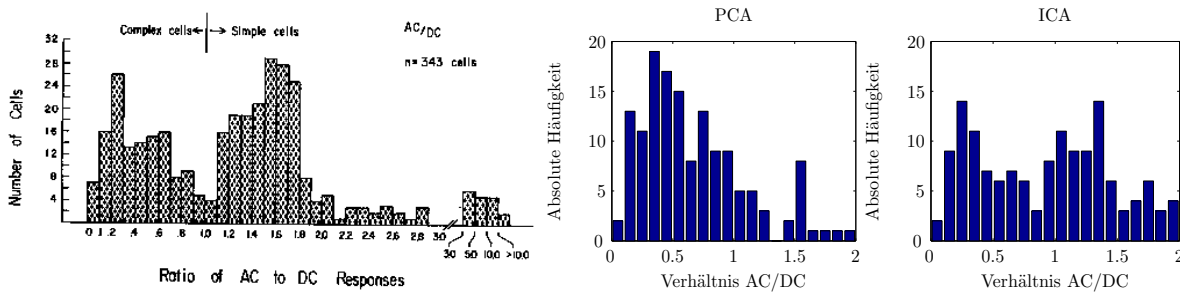


Abbildung 4.15: Verteilung der gemessenen Einheiten über das Verhältnis der Energie der ersten harmonischen zur Energie des Gleichanteils bei Erregung mit einer Sinus-schwingung. Links: Experimentelle Daten, gemessen im primären visuellen Kortex eines Mackacken. Mitte: Analoge Daten, gemessen an der bereits erläuterten reduzierten Konfiguration und PCA-Optimierung, rechts: dito für die ICA-Optimierung.

wie typische nichtlineare Zellen im visuellen Kortex. Die lokalen Energiedetektoren, die auf ein sinusförmiges Eingangssignal konstant antworten (Demodulation) lassen sich als prototypische *Complex*-Zellen beschreiben (DeValois und DeValois, 1990; Wegmann und Zetzsche, 1990b). Ein neurophysiologischer Ansatz zur Klassifikation von *Simple*- und *Complex*-Zellen im visuellen Kortex ist das Verhältnis der Energie der ersten Harmonischen zur Energie des Gleichanteils bei Erregung mit sinusförmigen Eingangssignalen (Kato et al., 1978). Die Verteilung der Zellen im primären visuellen Kortex zeigt eine bimodale Struktur, eine ähnliche Verteilung ergibt sich auch bei der analogen Bestimmung für die nichtlinearen Einheiten w_i , siehe auch Abb. 4.15 und (Nuding, 2002). Desweiteren führt die Optimierung zu Einheiten, die eine erhöhte Selektivität für zweidimensionale im Gegensatz zu eindimensionalen Merkmalen aufweisen, und somit kortikalen *Endstopped-Zellen* ähneln, bzw. *Hypercomplex-Zellen*, wenn beide Merkmale kombiniert auftreten (Kato et al., 1978; Saito et al., 1988).

Die Anforderungen an ein höheres bildverarbeitendes System sind zweierlei: erstens sollten die Einheiten sehr selektiv bezüglich des Eingangssignalraums sein, also nur auf einen Bruchteil aller möglichen Eingänge antworten und zweitens sollten die Einheiten invariant gegenüber irrelevanten Transformationen sein. Wie in Abschnitt 3.1.2 gezeigt wurde, ist dieses System prinzipiell in der Lage, selektives wie auch invariantes Antwortverhalten zu generieren. Durch Messen des Antwortverhaltens auf eine Menge von Testmustern, analog zu neurophysiologischen Untersuchungen (Hegd  und van Essen, 2003), kann  berpr ft werden, ob dieses System dieses Antwortverhalten auch anhand der Optimierung der Statistik nat rlicher Bilder *lernen* kann. Abb. 4.16 zeigt die Antworten einiger exemplarischer gelernter Einheiten. Das Netzwerk weist einige sehr selektive, als auch einige mehr invariante oder „breitbandigere“ Einheiten auf. Eine Methode zur Quantifizierung der Selektivit t ist durch den Stimulus Selektivit ts Index (*SSI*) (Hegd  und van Essen, 2003)

$$SSI = 1 - \frac{R_{mean}}{R_{peak}} \quad (4.24)$$

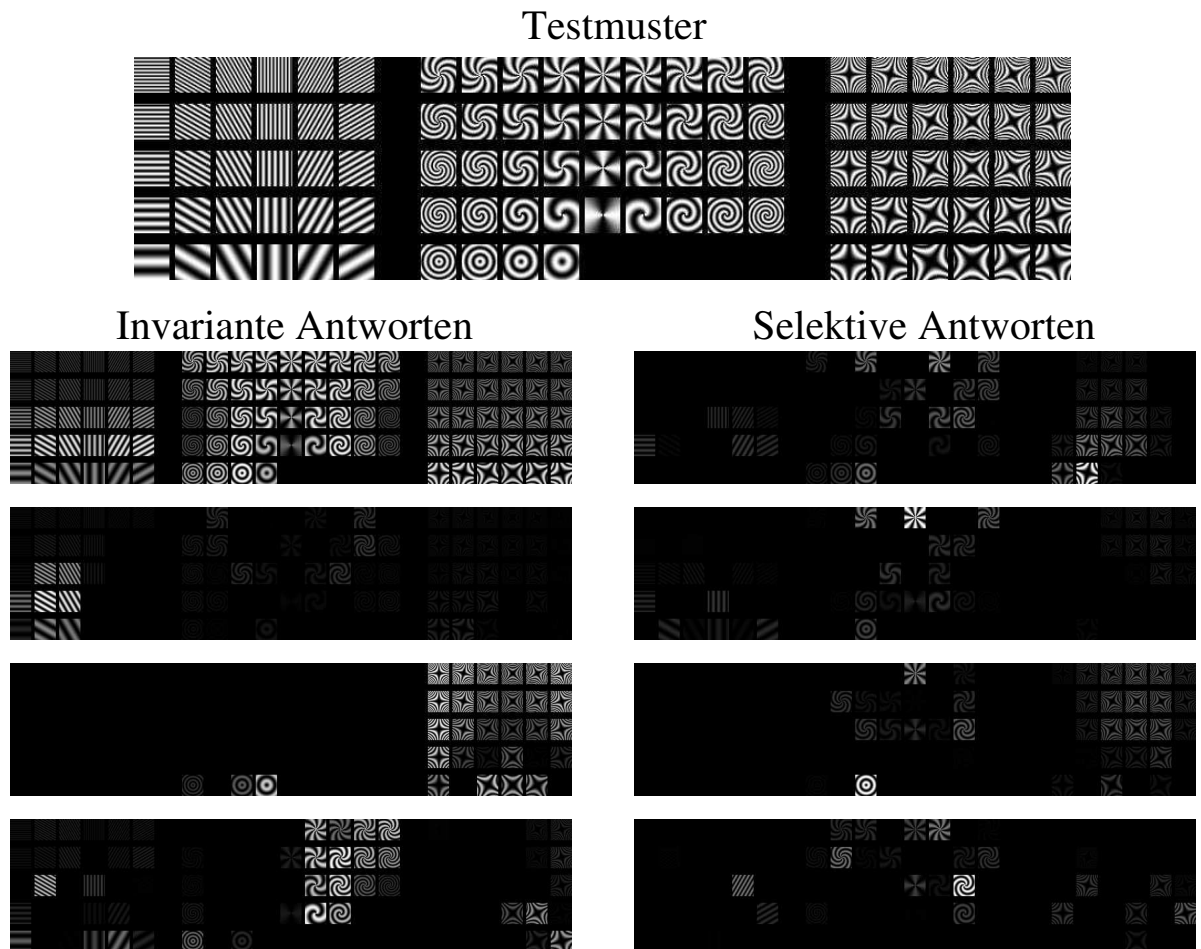


Abbildung 4.16: Antwortverhalten von acht exemplarischen Einheiten der gelernten LNL-Architektur auf eine Menge von Testmustern. Die Antworten einer jeden Einheit sind in der Form von skalierten Testmustern dargestellt, wobei jedes Muster proportional zur Antwort der Einheit auf dieses Muster skaliert ist.

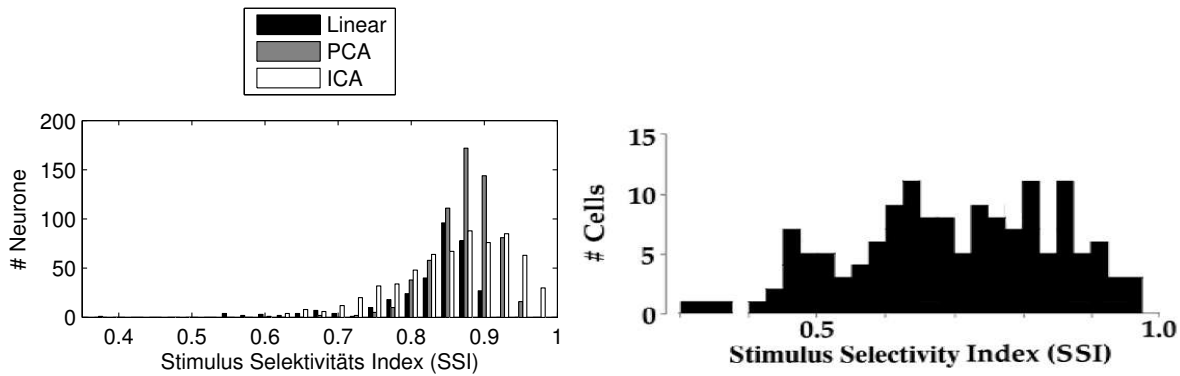


Abbildung 4.17: Stimulus Selektivitäts Indizes des gelernten Netzwerks und Vergleich mit neurophysiologischen Daten. Links: gezeigt sind die SSIs für die lineare Eingangsschicht als Vergleich, sowie für das mittels PCA und ICA gelernte Netzwerke. Rechts: physiologische Daten aus V2 von Mackaken, modifiziert nach (Hegd  und van Essen, 2003).

f r jede Einheit gegeben, wobei R_{mean} und R_{peak} die mittlere, bzw. die maximale Antwort auf die Menge der Testmuster bezeichnen. Die Indizes f r die gelernte LNL-Architektur (PCA- und ICA-Optimierung) sind zusammen mit einer neurophysiologischen Messung in der Sehrinde von Mackaken (Visueller Kortex V2) in Abb. 4.17 gezeigt. Es ist ersichtlich, da  die SSIs des gelernten Systems  ber die Population h her sind als f r die lineare Kontrolle, wof r nur die linearen Koeffizienten \vec{v} verwendet wurden.

Das vorgestellte Klassifikationsschema (*Null*, *Kontur*, *Textur*) wurde bereits in  hnlicher Form zur Sch tzung der wahrgenommenen Bildqualit t verwendet (Xu und Hauske, 1995; Hauske, 1997). Hier wurde gezeigt, da  eine derartige Klassifikation sehr gut mit der subjektiv empfundenen Qualit t eines Bildes korreliert.

In diesem Kapitel wurde also gezeigt, da  ein spezifisches \mathcal{LNL} -System in einer biologisch plausiblen Art implementiert werden kann, und f r technische Anwendungen, wie die Bilddatenkompression und Merkmalsextraktion, sowohl theoretisch als auch praktisch, Ergebnisse liefert, die ihren linearen Pendanten  berlegen sind.

5 Polynomiale Ansätze

Wie in Kapitel 2 demonstriert, stellen statistische Analysen zweiter Ordnung sowie lineare Systeme überaus mächtige Werkzeuge dar, um funktionelle Eigenschaften biologischer Sehsysteme zu analysieren und zu modellieren. Ebenfalls konnte gezeigt werden, daß einige relevante Aspekte der Informationsverarbeitung im visuellen System jedoch nicht innerhalb des Horizonts dieses klassischen Ansatzes liegen. Auf Seite der Analyse ist die Autokorrelation, bzw. das Leistungsdichtespektrum, weder ausreichend, die Diskriminationsfähigkeit für natürliche Stimuli zu erklären, noch die Optimalität basaler funktioneller Eigenschaften (z.B. die Orientierungsselektivität der Neurone im visuellen Kortex) zu beweisen. Auf der Seite der Modellierung wurde ebenfalls eine Vielzahl bekannter Phänomene angesprochen, die mit der linearen Systemtheorie definitiv nicht in Einklang zu bringen sind. Die logische Konsequenz ist dann die sukzessive Erhöhung der Komplexität der Ansätze, ganz im Sinne ALBERT EINSTEINS:

Man sollte die Dinge so einfach wie möglich gestalten, aber nicht einfacher.

Für die statistische Analyse bedeutet dies den Übergang von Zwei- zu Mehrpunktkorrelationen, genauer, *Kumulanten* höherer Ordnung und deren Spektren, *Polyspektren*. Dies kann als eine Reihenentwicklung der multivariaten Wahrscheinlichkeitsdichtefunktion natürlicher Signale betrachtet werden. Für die Modellsysteme stellt der Übergang zur Volterra-Wiener Theorie nichtlinearer Systeme die direkte Erweiterung der linearen Systeme dar; diese approximieren nichtlineare Eigenschaften mittels Polynomreihen. Es werden im Folgenden jeweils kurz die Grundlagen des jeweiligen Ansatzes referiert und gezeigt, welches Potential diese Ansätze für die Analyse und Modellierung des visuellen Systems bergen. Im Detail wird hierzu einerseits das Bispektrum natürlicher Bilder analysiert, sowie andererseits eine Reihe von bekannten empirischen nichtlinearen Phänomenen modelliert. Ein wesentlicher Aspekt ist hierbei die Entwicklung und Anpassung von Verfahren zur Identifikation nichtlinearer Systeme unter Verwendung artifizieller und natürlicher Stimuli. Desweiteren wird gezeigt, daß nichtlineare Systeme mit einem hochgradig selektiven Antwortverhalten bezüglich ihres Eingangs wesentlich günstigere Eigenschaften für die Identifikation aufweisen, als üblicherweise verwendete „schwach“ nichtlineare Kaskadensysteme. Zudem wird ein heuristischer Beweis zum Nachweis der Optimalität von $i2D$ -selektiven Systemen bezüglich der Eliminierung statistischer Abhängigkeiten höherer Ordnung geführt.

5.1 Spektren höherer Ordnung

Polyspektren sind die Spektren von Momenten- und Kumulantenfunktionen höherer Ordnung und sind sowohl für deterministische Signale als auch für stochastische Prozesse definiert. Die Einführung und die Terminologie gehen zurück auf JOHN W. TUKEY, der unter anderem durch die Entwicklung der schnellen Fourier-Transformation zu einer wichtigen Person der modernen Signalverarbeitung geworden ist (Brillinger, 2002). Seit der Einführung ist eine immense Zahl Publikationen zu diesem Thema verfaßt worden und eine Vielzahl von Anwendungen aus den unterschiedlichsten Domänen ist bekannt, so z.B. Anwendungen auf den Gebieten der Ozeanographie, Geophysik, Sonar, Astronomie, Bildverarbeitung, Strömungslehre, Plasmaphysik, und viele mehr. Erstmals auf dem Gebiet der Sehforschung wurden Korrelationsfunktionen höherer Ordnung von KLEIN und TYLER etabliert (Klein und Tyler, 1986), die Anwendung von Polyspektren jedoch ist bislang eher die Ausnahme; eine Anwendung ist die bispektrale Analyse natürlicher Bilder, siehe hierzu (Krieger et al., 1995; Krieger et al., 1997) und Abschnitt 5.1.3.

Kumulantenspektren höherer Ordnung zeichnen sich durch einige wesentliche Vorteile gegenüber dem bekannten Leistungsdichtespektrum aus¹. Diese sind im Speziellen: *(i)* Unempfindlichkeit gegenüber gaußschen Rauschprozessen mit unbekannter spektraler Charakteristik bei Detektion, Parameterschätzung und Klassifikation, *(ii)* Erhaltung der Phaseninformation, welche die Rekonstruktion des Signals ermöglicht, was für Statistiken zweiter Ordnung nur bei Minimalphasen-Signalen möglich ist und *(iii)* die Möglichkeit zur Detektion und Charakterisierung von Nichtlinearitäten in stochastischen Prozessen. Im Folgenden werden die grundlegenden Eigenschaften der Kumulanten und Polyspektren kurz angeführt, für weiterführende Literatur sei der Leser auf (Mendel, 1991) und (Nikias und Petropulu, 1993) verwiesen.

5.1.1 Momente und Kumulanten

Momente und Kumulanten von Zufallsgrößen

Definition. Gegeben sei eine Menge von n reellen Zufallsgrößen (ZG)

$$\vec{\mathbf{u}} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\} \quad (5.1)$$

mit n -dimensionaler Verbundwahrscheinlichkeitsdichtefunktion (VWDF)

$$f_{\vec{\mathbf{u}}}(u_1, u_2, \dots, u_n), \quad (5.2)$$

¹Wenn im Folgenden von Polyspektren die Rede ist, so sind damit immer die Spektren von Kumulanten gemeint, da diese i.A. andere Eigenschaften besitzen als die der Momente.

dann sind ihre Verbundmomente der Ordnung $r = k_1 + k_2 + \dots + k_n$ gegeben durch

$$\begin{aligned} Mom[\mathbf{u}_1^{k_1}, \mathbf{u}_2^{k_2}, \dots, \mathbf{u}_n^{k_n}] &= E\{\mathbf{u}_1^{k_1} \mathbf{u}_2^{k_2} \dots \mathbf{u}_n^{k_n}\} \\ &= (-j)^r \frac{\partial^r \Phi_{\bar{\mathbf{u}}}(\omega_1, \omega_2, \dots, \omega_n)}{\partial \omega_1^{k_1} \partial \omega_2^{k_2} \dots \partial \omega_n^{k_n}} \Bigg|_{\omega_1=\omega_2=\dots=\omega_n=0}. \end{aligned} \quad (5.3)$$

□

Hierbei stellt

$$\Phi_{\bar{\mathbf{u}}}(\omega_1, \omega_2, \dots, \omega_n) = E\{e^{j(\omega_1 u_1 + \omega_2 u_2 + \dots + \omega_n u_n)}\} \quad (5.4)$$

die charakteristische Funktion dar (Papoulis, 1991). Diese wird auch die *momentenerzeugende* Funktion genannt, da die Momente die Koeffizienten der Taylor-Reihenentwicklung derselben repräsentieren; dies kann direkt aus Gl. (5.3) entnommen werden. Beispielsweise sind die Momente zweiter Ordnung der ZGen $\{\mathbf{u}_1, \mathbf{u}_2\}$ gegeben durch $Mom[\mathbf{u}_1, \mathbf{u}_2] = E\{\mathbf{u}_1 \cdot \mathbf{u}_2\}$, $Mom[\mathbf{u}_1^2] = E\{\mathbf{u}_1^2\}$ und $Mom[\mathbf{u}_2^2] = E\{\mathbf{u}_2^2\}$. Völlig analog hierzu sind die Kumulanten als die Koeffizienten der Taylor-Reihenentwicklung um die Null der *zweiten charakteristischen Funktion*

$$\Psi_{\bar{\mathbf{u}}}(\omega_1, \omega_2, \dots, \omega_n) = \ln\{\Phi_{\bar{\mathbf{u}}}(\omega_1, \omega_2, \dots, \omega_n)\} \quad (5.5)$$

definiert:

$$Cum[\mathbf{u}_1^{k_1}, \mathbf{u}_2^{k_2}, \dots, \mathbf{u}_n^{k_n}] = (-j)^r \frac{\partial^r \Psi_{\bar{\mathbf{u}}}(\omega_1, \omega_2, \dots, \omega_n)}{\partial \omega_1^{k_1} \partial \omega_2^{k_2} \dots \partial \omega_n^{k_n}} \Bigg|_{\omega_1=\omega_2=\dots=\omega_n=0}. \quad (5.6)$$

Daraus läßt sich schließen, daß die Kumulanten einer Menge von ZGen durch ihre Momente ausgedrückt werden können. Beispielsweise sind die Momente

$$m_i = Mom[\mathbf{u}_1^i] = E\{\mathbf{u}_1^i\} \quad \forall i \in \{1..4\}$$

der ZG \mathbf{u}_1 mit ihren Kumulanten verknüpft über die Beziehungen

$$\begin{aligned} c_1 &= Cum[\mathbf{u}_1] = m_1 \\ c_2 &= Cum[\mathbf{u}_1^2] = m_2 - m_1^2 \\ c_3 &= Cum[\mathbf{u}_1^3] = m_3 - 3m_2 m_1 + 2m_1^3 \\ c_4 &= Cum[\mathbf{u}_1^4] = m_4 - 4m_3 m_1 - 3m_2^2 + 12m_2 m_1^2 - 6m_1^4. \end{aligned} \quad (5.7)$$

Dies kann leicht gezeigt werden, indem die Taylor-Entwicklung der charakteristischen Funktion

$$\Phi_{\mathbf{u}_1}(\omega_1) = 1 + j\omega_1 m_1 - \frac{\omega_1^2}{2!} m_2 \dots + \frac{(j\omega_1)^k}{k!} m_k + \dots \quad (5.8)$$

in Gl. (5.5) eingesetzt und die Kumulanten mit Hilfe Gl. (5.6) berechnet werden. So ergibt sich beispielsweise für den Kumulanten zweiter Ordnung

$$\begin{aligned}
 c_2 &= Cum[\mathbf{u}_1^2] = (-j)^2 \frac{\partial^2 \Psi(\omega_1)}{\partial \omega_1^2} \Big|_{\omega_1=0} \\
 &= -\frac{\partial^2}{\partial \omega_1^2} \left[\ln \left\{ \sum_{k=0}^{\infty} \frac{(j\omega_1)^k}{k!} m_k \right\} \right] \Big|_{\omega_1=0} \\
 &= -\frac{\partial}{\partial \omega_1} \left[\frac{\sum_{k=1}^{\infty} \frac{j^k \omega_1^{(k-1)}}{(k-1)!} m_k}{\sum_{k=0}^{\infty} \frac{(j\omega_1)^k}{k!} m_k} \right] \Big|_{\omega_1=0} \\
 &= \left[\frac{\sum_{k=2}^{\infty} \frac{j^k \omega_1^{(k-2)}}{(k-2)!} m_k \cdot \sum_{k=0}^{\infty} \frac{(j\omega_1)^k}{k!} m_k - \left(\sum_{k=1}^{\infty} \frac{j^k \omega_1^{(k-1)}}{(k-1)!} m_k \right)^2}{\left(\sum_{k=0}^{\infty} \frac{(j\omega_1)^k}{k!} m_k \right)^2} \right] \Big|_{\omega_1=0} \\
 &= m_2 - m_1^2. \tag{5.9}
 \end{aligned}$$

Zusammenhang zwischen Momenten und Kumulanten

Der allgemeine Zusammenhang zwischen den Momenten $Mom[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$ und Kumulanten $Cum[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n]$ der Ordnung $r = n$ geht zurück auf ein Theorem von LEONOV und SHIRYAEV, das später auch für die Identifikation nichtlinearer Systeme relevant wird (Leonov und Shiryayev, 1959):

$$Cum[\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n] = \sum_{(p)} (-1)^{(p-1)} (p-1)! \cdot E\left\{ \prod_{i \in s_1} \mathbf{u}_i \right\} \cdot E\left\{ \prod_{i \in s_2} \mathbf{u}_i \right\} \cdots E\left\{ \prod_{i \in s_p} \mathbf{u}_i \right\}, \tag{5.10}$$

wobei die Summation über alle Partitionen (s_1, s_2, \dots, s_p) , $p = 1, 2, \dots, n$ der Menge der Zahlen $\{1, 2, \dots, n\}$ ausgeführt wird. Beispielsweise kann die Menge $\{1, 2, 3\}$ folgendermaßen partitioniert werden:

$$\begin{aligned}
 p = 1 & \quad s_1 = \{1, 2, 3\} \\
 p = 2 & \quad s_1 = \{1\}, \quad s_2 = \{2, 3\} \\
 & \quad s_1 = \{2\}, \quad s_2 = \{1, 3\} \\
 & \quad s_1 = \{3\}, \quad s_2 = \{1, 2\} \\
 p = 3 & \quad s_1 = \{1\}, \quad s_2 = \{2\}, \quad s_3 = \{3\},
 \end{aligned}$$

somit ergibt sich aus Gl. (5.10) für den Kreuzkumulanten dritter Ordnung:

$$\begin{aligned}
 Cum[\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3] &= E\{\mathbf{u}_1 \mathbf{u}_2 \mathbf{u}_3\} - E\{\mathbf{u}_1\} \cdot E\{\mathbf{u}_2 \mathbf{u}_3\} - E\{\mathbf{u}_2\} \cdot E\{\mathbf{u}_1 \mathbf{u}_3\} \\
 &\quad - E\{\mathbf{u}_3\} \cdot E\{\mathbf{u}_1 \mathbf{u}_2\} + 2E\{\mathbf{u}_1\} \cdot E\{\mathbf{u}_2\} \cdot E\{\mathbf{u}_3\}. \tag{5.11}
 \end{aligned}$$

Hinweis: Die Beziehung (5.10) impliziert, daß die Berechnung von Verbundkumulanten der Ordnung r die Kenntnis sämtlicher Momente bis zur Ordnung r voraussetzt.

Eigenschaften

Die grundlegenden Eigenschaften von Momenten und Kumulanten können kurz in fünf Punkten zusammengefaßt werden (Brillinger, 1965; Nikias und Petropulu, 1993):

1. Für beliebige Konstanten $a_i \in \mathcal{R}, i = 1 \dots n$ gilt

$$\begin{aligned} Mom[(a_1 \mathbf{u}_1)^{k_1}, (a_2 \mathbf{u}_2)^{k_2}, \dots, (a_n \mathbf{u}_n)^{k_n}] &= \left(\prod_{i=1}^n a_i^{k_i} \right) Mom[\mathbf{u}_1, \dots, \mathbf{u}_n] \text{ und} \\ Cum[(a_1 \mathbf{u}_1)^{k_1}, (a_2 \mathbf{u}_2)^{k_2}, \dots, (a_n \mathbf{u}_n)^{k_n}] &= \left(\prod_{i=1}^n a_i^{k_i} \right) Cum[\mathbf{u}_1, \dots, \mathbf{u}_n]. \end{aligned} \quad (5.12)$$

Diese Identität folgt direkt aus den Gln. (5.3) und (5.10).

2. Da die Reihenfolge der partiellen Differentiation in Gl. (5.3) bzw. (5.6) beliebig vertauscht werden kann, sind die Momente und Kumulanten symmetrisch bezüglich ihrer Argumente, also beispielsweise

$$Mom[x_1, x_2, x_3] = Mom[x_2, x_1, x_3] = Mom[x_3, x_2, x_1], \text{ etc.}$$

3. Kann die Menge der ZGen $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ in mindestens zwei voneinander statistisch unabhängige nichtleere Teilmengen $\{\mathbf{u}_{i_1}, \dots, \mathbf{u}_{i_k}\}, \{\mathbf{u}_{i_{k+1}}, \dots, \mathbf{u}_{i_n}\}$ separiert werden, so sind alle Kreuzkumulanten der Teilmengen identisch Null. Da die VWDF $f_{\bar{\mathbf{u}}}(u_1, \dots, u_n)$ für statistisch unabhängige Teilmengen faktorisiert werden kann:

$$f_{\bar{\mathbf{u}}}(u_1, \dots, u_n) = g(u_{i_1}, \dots, u_{i_k}) \cdot h(u_{i_{k+1}}, \dots, u_{i_n}), \quad (5.13)$$

läßt sich nach den Gesetzen der Fourier-Transformation auch die erste charakteristische Funktion separieren. Dementsprechend gilt dann für die zweite charakteristische Funktion:

$$\Psi_{\bar{\mathbf{u}}}(\omega_1, \dots, \omega_n) = \ln\{G(\omega_{i_1}, \dots, \omega_{i_k})\} + \ln\{H(\omega_{i_{k+1}}, \dots, \omega_{i_n})\}. \quad (5.14)$$

Für einen Kreuzkumulanten der beiden Teilmengen

$$Cum[\dots, \mathbf{u}_m^{k_m}, \dots, \mathbf{u}_n^{k_n} \dots], \quad m \in \{i_1, \dots, i_k\} \wedge n \in \{i_{k+1}, \dots, i_n\} \quad (5.15)$$

wird der gemischte Differentialquotient $\frac{\partial \Psi}{\partial \omega_m \partial \omega_n}$ evaluiert, wobei jeweils ein Term in Gl. (5.14) unabhängig von ω_m , bzw. ω_n ist und somit Null wird. *Notabene:* Für Momente gilt diese Eigenschaft im Allgemeinen nicht.

4. Folgen die ZGen $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ verbundweise einer Gaußverteilung, so ist die vollständige Information über ihre Verteilung in den Momenten der Ordnung $r \leq 2$

enthalten. Das bedeutet, daß sämtliche Kumulanten der Ordnung $r = \sum_{i=1}^n k_i > 2$ verschwinden:

$$Cum[\mathbf{u}_1^{k_1}, \dots, \mathbf{u}_n^{k_n}] = 0 \quad \text{für} \quad f_{\vec{\mathbf{u}}}(\vec{\mathbf{u}}) = \mathcal{N}(\vec{\mu}, C) \quad \text{und} \quad r > 2, \quad (5.16)$$

$\vec{\mu}$ und C bezeichnen den Mittelwertsvektor respektive die Kovarianzmatrix der multivariaten Normalverteilung. Das bedeutet, daß Kumulanten höherer Ordnung sozusagen die „Nicht-Gaußförmigkeit“ einer Verteilung messen.

5. Sind die ZGen $\vec{\mathbf{u}} = \{\mathbf{u}_1, \dots, \mathbf{u}_n\}$ und $\vec{\mathbf{v}} = \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ voneinander statistisch unabhängig, so sind die Kumulanten der Summe gleich der Summe der Kumulanten, d.h. Kumulanten sind additiv bezüglich ihrer Argumente²:

$$Cum[(\mathbf{u}_1 + \mathbf{v}_1)^{k_1}, \dots, (\mathbf{u}_n + \mathbf{v}_n)^{k_n}] = Cum[\mathbf{u}_1^{k_1}, \dots, \mathbf{u}_n^{k_n}] + Cum[\mathbf{v}_1^{k_1}, \dots, \mathbf{v}_n^{k_n}]. \quad (5.17)$$

Dies liegt darin begründet, daß sich für statistisch unabhängige ZGen die Verteilung der Summe als Faltung ergibt. Somit ist die erste charakteristische Funktion nach den Gesetzen der Fourier-Transformation als Produkt und die zweite charakteristische Funktion als Summe darstellbar. Diese Eigenschaft gilt wiederum nicht für Momente und ist besonders für die Systemidentifikation in Abschnitt 5.2.5 relevant, da diese, zusammen mit Eigenschaft (4) impliziert, daß additives Gaußsches Rauschen keinen Einfluß auf die Kumulanten nimmt.

Momente und Kumulanten stationärer Prozesse

Wie bereits in Abschnitt 2.4 erläutert, kann ein kontinuierlicher stochastischer Prozeß durch eine geeignet gewählte Abtaststrategie diskretisiert werden. Die Abtastwerte können dann in einem Vektor zusammengefaßt und dieser als Zufallsvektor interpretiert werden. Für den Fall mehrdimensionaler Signale ergibt sich unter Verwendung von Abtastpositionen \vec{x}_i

$$\vec{\mathbf{u}} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\} = \{\mathbf{u}(\vec{x}_1), \mathbf{u}(\vec{x}_2), \dots, \mathbf{u}(\vec{x}_n)\}, \quad \vec{x}_i \in \mathcal{R}^N. \quad (5.18)$$

Somit sind die eben eingeführten grundlegenden Definitionen und Eigenschaften von Momenten und Kumulanten direkt auf stochastische Prozesse übertragbar. Sämtliche statistische Kenngrößen besitzen dann als Argumente die Abtastpositionen $\vec{x}_i, i \in \{1 \dots n\}$. Ist ein Prozeß streng stationär, so gehen die Argumente über in $n - 1$ Differenzvektoren. Insbesondere kann dann für die Momente und Kumulanten die folgende Schreibweise verwendet werden:

$$\begin{aligned} m_n^{\mathbf{u}}(\vec{x}_1, \dots, \vec{x}_{n-1}) &:= Mom[\mathbf{u}(\vec{x}), \mathbf{u}(\vec{x} + \vec{x}_1), \dots, \mathbf{u}(\vec{x} + \vec{x}_{n-1})] \\ c_n^{\mathbf{u}}(\vec{x}_1, \dots, \vec{x}_{n-1}) &:= Cum[\mathbf{u}(\vec{x}), \mathbf{u}(\vec{x} + \vec{x}_1), \dots, \mathbf{u}(\vec{x} + \vec{x}_{n-1})]. \end{aligned} \quad (5.19)$$

Wichtige Sonderfälle ergeben sich für

²Daher auch der Terminus *Kumulant*.

- $n = 1$:

$$c_1^{\mathbf{u}} = m_1^{\mathbf{u}} = E\{\mathbf{u}(\vec{x})\}, \quad (5.20)$$

dies ist der Mittelwert des Prozesses $\mathbf{u}(\vec{x})$.

- $n = 2$:

$$c_2^{\mathbf{u}}(\vec{x}_1) = m_2^{\mathbf{u}}(\vec{x}_1) - (m_1^{\mathbf{u}})^2, \quad (5.21)$$

dies ist die bekannte Gleichung der Autokovarianzfunktion des stationären Prozesses $\mathbf{u}(\vec{x})$. Für $\vec{x}_1 = 0$ ergibt sich die bekannte Varianz oder auch Wechselleistung des Signals $\gamma_2^{\mathbf{u}} = c_2^{\mathbf{u}}(0)$.

- $n = 3$:

$$c_3^{\mathbf{u}}(\vec{x}_1, \vec{x}_2) = m_3^{\mathbf{u}}(\vec{x}_1, \vec{x}_2) - m_1^{\mathbf{u}}(m_2^{\mathbf{u}}(\vec{x}_1) + m_2^{\mathbf{u}}(\vec{x}_2) + m_2^{\mathbf{u}}(\vec{x}_2 - \vec{x}_1)) + 2(m_1^{\mathbf{u}})^3, \quad (5.22)$$

wobei man für $\vec{x}_1 = \vec{x}_2 = 0$ die *Schiefe* $\gamma_3^{\mathbf{u}} = c_3^{\mathbf{u}}(0, 0)$ der Verteilung des Prozesses $\mathbf{u}(\vec{x})$ erhält.

- $n = 4$:

$$\begin{aligned} c_4^{\mathbf{u}}(\vec{x}_1, \vec{x}_2, \vec{x}_3) &= m_4^{\mathbf{u}}(\vec{x}_1, \vec{x}_2, \vec{x}_3) - m_2^{\mathbf{u}}(\vec{x}_1) \cdot m_2^{\mathbf{u}}(\vec{x}_3 - \vec{x}_2) - m_2^{\mathbf{u}}(\vec{x}_2) \cdot m_2^{\mathbf{u}}(\vec{x}_3 - \vec{x}_1) \\ &\quad - m_2^{\mathbf{u}}(\vec{x}_3) \cdot m_2^{\mathbf{u}}(\vec{x}_2 - \vec{x}_1) - m_1^{\mathbf{u}}(m_3^{\mathbf{u}}(\vec{x}_2 - \vec{x}_1, \vec{x}_3 - \vec{x}_1) \\ &\quad + m_3^{\mathbf{u}}(\vec{x}_2, \vec{x}_3) + m_3^{\mathbf{u}}(\vec{x}_1, \vec{x}_3) + m_3^{\mathbf{u}}(\vec{x}_1, \vec{x}_2)) \\ &\quad + (m_1^{\mathbf{u}})^2(m_2^{\mathbf{u}}(\vec{x}_1) + m_2^{\mathbf{u}}(\vec{x}_2) + m_2^{\mathbf{u}}(\vec{x}_3) + m_2^{\mathbf{u}}(\vec{x}_3 - \vec{x}_1) \\ &\quad + m_2^{\mathbf{u}}(\vec{x}_3 - \vec{x}_2) + m_2^{\mathbf{u}}(\vec{x}_2 - \vec{x}_1)) - 6(m_1^{\mathbf{u}})^4, \end{aligned} \quad (5.23)$$

setzt man bei dieser Kumulantenfunktion vierter Ordnung die Argumente zu Null, d.h. wird diese im Ursprung ausgewertet, so erhält man die *Kurtosis* $\gamma_4^{\mathbf{u}} = c_4^{\mathbf{u}}(0, 0, 0)$ der Verteilung des Prozesses $\mathbf{u}(\vec{x})$. Diese kann, wie die Schiefe $\gamma_3^{\mathbf{u}}$ als eine Art Formparameter der Verteilung gedeutet werden; häufig wird die auf die Varianz normalisierte Kurtosis zur Beschreibung der Spärlichkeit einer Verteilung verwendet, siehe Gl.(2.20) und Abb. 2.11(d).

Für einen mittelwertfreien Prozeß folgt direkt aus den Gln. (5.21) und (5.22), daß die Kumulanten bis zur dritten Ordnung identisch den entsprechenden Momenten sind. Desweiteren folgt aus Gl. (5.23), daß zur Berechnung der Kumulanten vierter Ordnung die Kenntnis der Momente vierter sowie zweiter Ordnung vonnöten sind. Abb. 5.1 zeigt verschiedene univariate Verteilungen sowie ihre entsprechenden Momente und Kumulanten.

Völlig analog zur bekannten Kreuzkorrelationsfunktion können formal auch Kreuzkumulanten definiert werden. So sind beispielsweise die Kreuzkumulanten dritter Ordnung der stationären Prozesse $\{\mathbf{u}(\vec{x})\}$, $\{\mathbf{v}(\vec{x})\}$, $\{\mathbf{w}(\vec{x})\}$ gegeben durch

$$\begin{aligned} c_3^{\mathbf{u}\mathbf{v}\mathbf{w}}(\vec{x}_1, \vec{x}_2) &= Cum[\mathbf{u}(\vec{x}), \mathbf{v}(\vec{x} + \vec{x}_1), \mathbf{w}(\vec{x} + \vec{x}_2)] \\ &= E\{(\mathbf{u}(\vec{x}) - m_1^{\mathbf{u}})(\mathbf{v}(\vec{x}) - m_1^{\mathbf{v}})(\mathbf{w}(\vec{x}) - m_1^{\mathbf{w}})\}. \end{aligned} \quad (5.24)$$

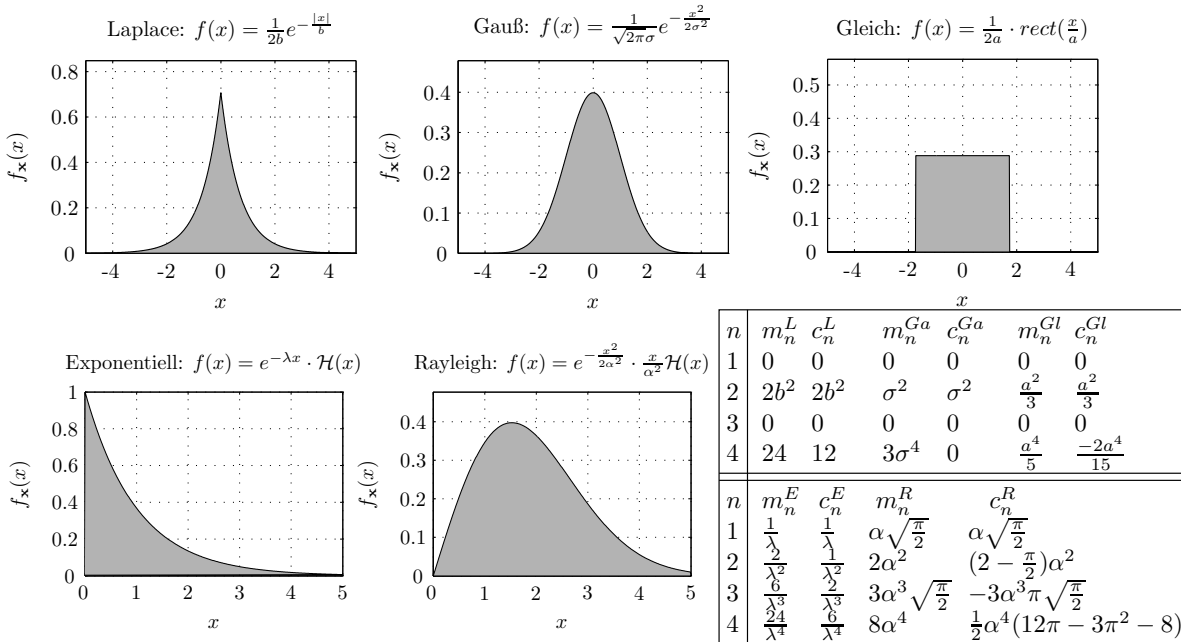


Abbildung 5.1: Verschiedene Beispiele symmetrischer (oben) und asymmetrischer (unten) univariater Verteilungen und ihre Momente und Kumulanten.

Selbstverständlich existieren noch weitere Kreuzkumulanten dieser drei Prozesse, z.B. $c_3^{\mathbf{u}\mathbf{u}\mathbf{v}}(\vec{x}_1, \vec{x}_2)$ oder $c_3^{\mathbf{v}\mathbf{v}\mathbf{v}}(\vec{x}_1, \vec{x}_2)$, etc. Allgemein ist die Kreuzkumulantenfunktion n -ter Ordnung der stationären Prozesse $\{\mathbf{u}_i(\vec{x})\}, i = 1 \dots n$ definiert als:

$$c_n^{\mathbf{u}_1\mathbf{u}_2 \dots \mathbf{u}_n}(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{n-1}) = \text{Cum}[\mathbf{u}_1(\vec{x}), \mathbf{u}_2(\vec{x} + \vec{x}_1), \dots, \mathbf{u}_n(\vec{x} + \vec{x}_{n-1})]. \quad (5.25)$$

5.1.2 Kumulantenspektren

Für einen reellen, streng stationären Prozeß $\{\mathbf{u}(\vec{x})\}$ mit einer Kumulantenfunktion n -ter Ordnung $c_n^{\mathbf{u}}(\vec{x}_1, \dots, \vec{x}_{n-1})$ kann das assoziierte Kumulantenspektrum wie folgt definiert werden.

Definition. Es sei die Kumulantenfunktion absolut integrierbar, d.h.

$$\int_{R^N} \dots \int_{R^N} |c_n^{\mathbf{u}}(\vec{x}_1, \dots, \vec{x}_{n-1})| d\vec{x}_1 \dots d\vec{x}_{n-1} < \infty, \quad (5.26)$$

dann ist das Kumulantenspektrum der Ordnung n eines N -dimensionalen Prozesses definiert als die $(n - 1)N$ -dimensionale Fourier-Transformation

$$\begin{aligned}
 C_n^{\mathbf{u}}(\vec{f}_1, \dots, \vec{f}_{n-1}) &\bullet\text{---}\circ c_n^{\mathbf{u}}(\vec{x}_1, \dots, \vec{x}_{n-1}) \\
 &= \int_{R^N} \cdots \int_{R^N} c_n^{\mathbf{u}}(\vec{x}_1, \dots, \vec{x}_{n-1}) e^{-j2\pi(\vec{f}_1^T \vec{x}_1 + \cdots + \vec{f}_{n-1}^T \vec{x}_{n-1})} d\vec{x}_1 \cdots d\vec{x}_{n-1}.
 \end{aligned}
 \tag{5.27}$$

□

Im Allgemeinen (für $n > 2$) ist $C_n^{\mathbf{u}}(\vec{f}_1, \dots, \vec{f}_{n-1})$ komplex, ist also darstellbar mit Betrag und Phase:

$$\begin{aligned}
 C_n^{\mathbf{u}}(\vec{f}_1, \dots, \vec{f}_{n-1}) &= |C_n^{\mathbf{u}}(\vec{f}_1, \dots, \vec{f}_{n-1})| e^{j\angle C_n^{\mathbf{u}}(\vec{f}_1, \dots, \vec{f}_{n-1})} \\
 &= |C_n^{\mathbf{u}}(\vec{f}_1, \dots, \vec{f}_{n-1})| e^{j\Phi_n^{\mathbf{u}}(\vec{f}_1, \dots, \vec{f}_{n-1})}.
 \end{aligned}
 \tag{5.28}$$

Wie weiter unten gezeigt wird, sind die Kumulantenspektren auch analog für deterministische Signale definiert. Die Phaseninformation ermöglicht dann eine Rekonstruktion des ursprünglichen Signals; so können beispielsweise Schnitte durch das Bispektrum die Phase des Signals rekonstruieren (Bartelt et al., 1984; Holambe et al., 1996; Petropulu und Pozidis, 1998). Dies ist auch ein wesentlicher Unterschied zum bekannten Leistungsdichtespektrum, da dies rein reell ist, und somit keine Phaseninformation beinhaltet. Daher ist eine Rekonstruktion aus dem Spektrum zweiter Ordnung nur für Minimalphasensysteme möglich (Nikias und Petropulu, 1993).

Eine Anmerkung zur Terminologie: Polyspektren bezeichnen sowohl die Spektren von Momenten als auch von Kumulanten und ebenso die Spektren von deterministischen und stochastischen Signalen, also Energie- und Leistungsdichtespektren.

Alternative Definition

Die physikalische Signifikanz von Polyspektren wird deutlich, wenn die Definition über die FOURIER-STIELTJES Repräsentation stochastischer Prozesse erfolgt. Hiernach kann ein stochastischer Prozeß $\{\mathbf{u}(\vec{x})\}$ spektral dargestellt werden als Überlagerung von Exponentialschwingungen, deren Amplituden komplexe Zufallsgrößen sind (Papoulis, 1991):

$$\mathbf{u}(\vec{x}) = \int_{R^N} e^{j2\pi \vec{f}^T \vec{x}} d\mathbf{U}(\vec{f}).
 \tag{5.29}$$

Wenn der Prozeß stationär ist, so müssen für eine gegebene Frequenz \vec{f} sämtliche Phasen gleichwahrscheinlich vorkommen, was zur Folge hat, daß der Erwartungswert für $\vec{f} \neq 0$ verschwindet:

$$E\{d\mathbf{U}(\vec{f})\} = 0.
 \tag{5.30}$$

Ebenso folgt aus der Bedingung der Stationarität, daß der Erwartungswert über das Produkt zweier Komponenten verschwinden muß, solange $\vec{f}_1 \neq \vec{f}_2$ ist:

$$E\{d\mathbf{U}(\vec{f}_1) \cdot d\mathbf{U}(\vec{f}_2)\} = \begin{cases} C_2^{\mathbf{u}}(\vec{f}), & \text{falls } \vec{f}_1 = \vec{f}_2 = \vec{f} \\ 0, & \text{falls } \vec{f}_1 \neq \vec{f}_2 \end{cases} \quad (5.31)$$

Allgemein kann gezeigt werden, daß sich die Kumulanten des Prozesses $\{d\mathbf{U}(\vec{f})\}$ zu

$$Cum[d\mathbf{U}(\vec{f}_1), \dots, d\mathbf{U}(\vec{f}_n)] = \begin{cases} C_n^{\mathbf{u}}(\vec{f}_1, \dots, \vec{f}_{n-1}), & \text{falls } \vec{f}_1 + \dots + \vec{f}_n = 0 \\ 0, & \text{für } \vec{f}_1 + \dots + \vec{f}_n \neq 0 \end{cases} \quad (5.32)$$

ergeben. Daraus ist ersichtlich, daß das Polyspektrum der Ordnung n den Kumulantenbeitrag von n Fourier-Komponenten mißt, deren Summe Null ist. Obwohl das Kumulantenpektrum der Ordnung n eine Funktion von $n-1$ Variablen ist, sollte dennoch bedacht werden, daß es eine „verborgene“ Variable $\vec{f}_n = -\vec{f}_1 - \dots - \vec{f}_{n-1}$ gibt. Man kann also sagen, die Spektren höherer Ordnung messen die Interaktionen, bzw. die statistischen Bindungen *zwischen* verschiedenen Frequenzkomponenten.

Spezielle Polyspektren

Das Leistungsspektrum, Bispektrum und Trispektrum sind spezielle Varianten der Polyspektren aus Gl. (5.27) mit der Ordnung

- $n=2$: Leistungsspektrum

$$C_2^{\mathbf{u}}(\vec{f}) = \int_{R^N} c_2^{\mathbf{u}}(\vec{x}) e^{-j2\pi \vec{f}^T \vec{x}} d\vec{x}. \quad (5.33)$$

Für einen mittelwertfreien Prozeß stellt dies genau die Wiener-Khintchine Identität dar. Es gilt hierfür:

$$\begin{aligned} c_2^{\mathbf{u}}(\vec{x}) &= c_2^{\mathbf{u}}(-\vec{x}) \\ C_2^{\mathbf{u}}(\vec{f}) &= C_2^{\mathbf{u}}(-\vec{f}) \\ C_2^{\mathbf{u}}(\vec{f}) &\geq 0 \quad (\text{reelle, nichtnegative Funktion}). \end{aligned} \quad (5.34)$$

- $n=3$: Bispektrum

$$C_3^{\mathbf{u}}(\vec{f}_1, \vec{f}_2) = \int_{R^N} \int_{R^N} c_3^{\mathbf{u}}(\vec{x}_1, \vec{x}_2) e^{-j2\pi(\vec{f}_1^T \vec{x}_1 + \vec{f}_2^T \vec{x}_2)} d\vec{x}_1 d\vec{x}_2. \quad (5.35)$$

Das Bispektrum zeichnet sich durch eine hochgradige Symmetrie aus, die direkt auf die grundlegenden Eigenschaften der Kumulanten zurückzuführen sind (Nikias

und Petropulu, 1993). So besitzt beispielsweise das Bispektrum eines eindimensionalen, reellen Prozesses 12 Symmetrieregionen, die zur effizienten Schätzung und Rauschunterdrückung verwendet werden können.

- n=4: Trispektrum

$$C_4^{\mathbf{u}}(\vec{f}_1, \vec{f}_2, \vec{f}_3) = \int_{R^N} \int_{R^N} \int_{R^N} c_4^{\mathbf{u}}(\vec{x}_1, \vec{x}_2, \vec{x}_3) e^{-j2\pi(\vec{f}_1^T \vec{x}_1 + \vec{f}_2^T \vec{x}_2 + \vec{f}_3^T \vec{x}_3)} d\vec{x}_1 d\vec{x}_2 d\vec{x}_3. \quad (5.36)$$

Wie das Bispektrum, so zeichnet sich auch das Trispektrum durch eine Vielzahl an Symmetrien, bedingt durch die Eigenschaften der Kumulantenfunktionen, aus. Ein eindimensionaler, reeller Prozeß besitzt insgesamt 96 Symmetrieregionen.

Detektion und Identifikation von nichtlinearen Prozessen

Polyspektren bieten prinzipiell die Möglichkeit, Nichtlinearitäten in Prozessen zu detektieren, respektive zu identifizieren. Dabei muß unterschieden werden zwischen linearen *nicht-gaußschen* und *nichtlinearen* Prozessen. Für diese Unterscheidung wird häufig die *Kohärenzfunktion* verwendet, diese kombiniert das Polyspektrum der Ordnung n mit dem Leistungsspektrum (Nikias und Petropulu, 1993):

$$P_n^{\mathbf{u}}(\vec{f}_1, \dots, \vec{f}_{n-1}) = \frac{C_n^{\mathbf{u}}(\vec{f}_1, \dots, \vec{f}_{n-1})}{\sqrt{C_2^{\mathbf{u}}(\vec{f}_1) \cdot C_2^{\mathbf{u}}(\vec{f}_2) \cdots C_2^{\mathbf{u}}(\vec{f}_{n-1}) \cdot C_2^{\mathbf{u}}(\vec{f}_1 + \vec{f}_2 + \cdots + \vec{f}_{n-1})}}. \quad (5.37)$$

Desweiteren läßt sich leicht zeigen, daß für die Polyspektren des Ausgangs eines linearen Systems mit spektraler Übertragungsfunktion $H(\vec{f})$ gilt:

$$C_n^{\mathbf{u}_2}(\vec{f}_1, \dots, \vec{f}_{n-1}) = H(\vec{f}_1)H(\vec{f}_2) \cdots H(\vec{f}_{n-1})H^*(\vec{f}_1 + \cdots + \vec{f}_{n-1}) \cdot C_n^{\mathbf{u}_1}(\vec{f}_1, \dots, \vec{f}_{n-1}), \quad (5.38)$$

hier bezeichnen \mathbf{u}_1 und \mathbf{u}_2 den Ein- bzw. Ausgang des Systems und $C_n^{\mathbf{u}_1}, C_n^{\mathbf{u}_2}$ die entsprechenden Polyspektren der Ordnung n . Handelt es sich beim Eingangssignal um ein nicht-gaußsches weißes Rauschen (*Higher-Order White Noise*, HOWN), so ergibt sich aus Gl. (5.38):

$$C_n^{\mathbf{u}_2}(\vec{f}_1, \dots, \vec{f}_{n-1}) = \gamma_n^{\mathbf{u}_1} \cdot H(\vec{f}_1)H(\vec{f}_2) \cdots H(\vec{f}_{n-1})H^*(\vec{f}_1 + \cdots + \vec{f}_{n-1}). \quad (5.39)$$

Wenn nun also ein Prozeß linear ist, kann dieser mithilfe von Gl. (5.39) modelliert werden. Damit ergibt sich die Kohärenzfunktion dieses Prozesses zu

$$P_n^{\mathbf{u}_2}(\vec{f}_1, \dots, \vec{f}_{n-1}) = \frac{\gamma_n^{\mathbf{u}_1}}{(\gamma_2^{\mathbf{u}_1})^{n/2}} \frac{H(\vec{f}_1)H(\vec{f}_2) \cdots H(\vec{f}_{n-1})H^*(\vec{f}_1 + \cdots + \vec{f}_{n-1})}{|H(\vec{f}_1)||H(\vec{f}_2)| \cdots |H(\vec{f}_{n-1})||H^*(\vec{f}_1 + \cdots + \vec{f}_{n-1})|}, \quad (5.40)$$

woran man erkennen kann, daß in diesem Fall der Betrag der Kohärenzfunktion

$$|P_n^{\mathbf{u}_2}(\vec{f}_1, \dots, \vec{f}_{n-1})| = \frac{\gamma_n^{\mathbf{u}_1}}{(\gamma_2^{\mathbf{u}_1})^{n/2}} \quad (5.41)$$

konstant für alle Frequenzen ist. Sollte der Prozeß zudem gaußsch sein, so ist $\gamma_n^{\mathbf{u}_1} \equiv 0$ für alle $n > 2$ und damit verschwindet auch die Kohärenzfunktion. Weist sie jedoch eine hohe Variabilität bezüglich der Frequenzen auf, so ist der Prozeß $\{\mathbf{u}_2(\vec{x})\}$ substantiell nichtlinear.

Nichtlineare Prozesse entstehen häufig beim Passieren von nichtlinearen Systemen, dies wird in Abschnitt 5.2 genauer erläutert. Ein Phänomen, das hierbei typischerweise auftritt, und welches nur mit Spektren *höherer* Ordnung detektierbar ist, ist die sogenannte *Phasenkopplung*. Dies wird im Folgenden anhand zweier einfacher Beispiele demonstriert.

Beispiel. Betrachtet wird ein einfacher Quadrierer $u_2(x) = (u_1(x))^2$. Der Eingang bestehe aus der Überlagerung zweier Prozesse mit den Frequenzen f_a, f_b und zufälligen Phasen ϕ_a, ϕ_b :

$$\mathbf{u}_1(x) = \cos(2\pi f_a x + \phi_a) + \cos(2\pi f_b x + \phi_b). \quad (5.42)$$

Der Ausgang des Systems ergibt sich wie folgt:

$$\begin{aligned} \mathbf{u}_2(x) &= \left(\cos(2\pi f_a x + \phi_a) + \cos(2\pi f_b x + \phi_b) \right)^2 \\ &= \left(\frac{1}{2} (e^{j(2\pi f_a x + \phi_a)} + e^{-j(2\pi f_a x + \phi_a)}) + \frac{1}{2} (e^{j(2\pi f_b x + \phi_b)} + e^{-j(2\pi f_b x + \phi_b)}) \right)^2 \\ &= \frac{1}{4} \left[e^{j(4\pi f_a x + 2\phi_a)} + 2e^{j(2\pi(f_a + f_b)x + \phi_a + \phi_b)} + 2e^{j(2\pi(f_a - f_b)x + \phi_a - \phi_b)} \right. \\ &\quad + e^{-j(4\pi f_a x + 2\phi_a)} + 2e^{-j(2\pi(f_a - f_b)x + \phi_a - \phi_b)} + 2e^{-j(2\pi(f_a + f_b)x + \phi_a + \phi_b)} \\ &\quad \left. + e^{j(4\pi f_b x + 2\phi_b)} + e^{-j(4\pi f_b x + 2\phi_b)} + 4 \right] \\ &= 1 + \frac{1}{2} \cos(4\pi f_a x + 2\phi_a) + \cos(2\pi(f_a + f_b)x + \phi_a + \phi_b) \\ &\quad + \cos(2\pi(f_a - f_b)x + \phi_a - \phi_b) + \frac{1}{2} \cos(4\pi f_b x + 2\phi_b). \end{aligned} \quad (5.43)$$

□

Am obigen einfachen Beispiel kann man erkennen, daß die Antwort eines Quadrierers auf die lineare Überlagerung zweier harmonischer aus einem Gleichanteil, einem Anteil bei den doppelten Frequenzen und Anteilen bei den Summen- und Differenzfrequenzen besteht. Ersichtlich ist auch, daß die Anteile bei den Summen-, bzw. Differenzfrequenzen auch in einer bestimmten Phase zueinander stehen. Diese Phasenrelation ist charakteristisch für nichtlineare Systeme und wird in diesem Fall *quadratische Phasenkopplung* genannt. Das Besondere dabei ist, daß genau diese Phasenkopplung von Spektren höherer Ordnung detektiert wird, wie im folgenden Beispiel verdeutlicht.

Beispiel. Betrachtet werden nun zwei Prozesse, bestehend aus jeweils drei harmonischen Schwingungen mit stochastisch verteilten Phasen im Intervall $[0, 2\pi]$:

$$\begin{aligned} \mathbf{u}_1(x) &= \cos(2\pi f_a x + \phi_a) + \cos(2\pi f_b x + \phi_b) + \cos(2\pi f_c x + \phi_c) \\ \mathbf{u}_2(x) &= \cos(2\pi f_d x + \phi_d) + \cos(2\pi f_e x + \phi_e) + \cos(2\pi f_f x + \phi_f), \end{aligned} \quad (5.44)$$

wobei gilt: $f_c = f_a + f_b, f_f = f_d + f_e, \phi_f = \phi_d + \phi_e$. Die Frequenzkomponenten in Signal \mathbf{u}_1 sind harmonisch verwandt, da sich die dritte als Summe der beiden ersten ergibt. In Signal \mathbf{u}_2 hingegen sind zudem auch die Phasen in entsprechender Relation zueinander, dies wird als quadratisch phasengekoppelt bezeichnet. Das Spektrum der Signale ist trivial, bestehend aus Impulsen an den jeweiligen Frequenzen, Abb. 5.2(a). Wie leicht berechnet werden kann, ergeben sich die Kumulanten dritter Ordnung zu:

$$\begin{aligned} c_3^{\mathbf{u}_1}(x_1, x_2) &= 0 \\ c_3^{\mathbf{u}_2}(x_1, x_2) &= \frac{1}{4} \left[\cos(f_e x_1 + f_d x_2) + \cos(f_f x_1 - f_d x_2) + \cos(f_d x_1 + f_e x_2) \right. \\ &\quad \left. + \cos(f_f x_1 - f_e x_2) + \cos(f_d x_1 - f_f x_2) + \cos(f_e x_1 - f_f x_2) \right], \end{aligned} \quad (5.45)$$

dies ist in Abb. 5.2(b) gezeigt. □

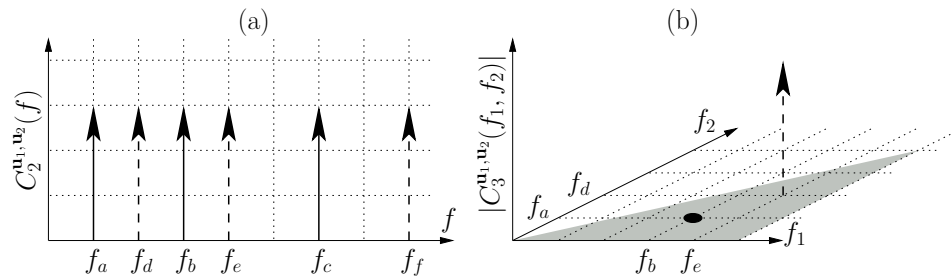


Abbildung 5.2: Der Effekt der Phasenkopplung im (a) Leistungsspektrum und (b) Bisppektrum. Erkennbar ist, daß nur das Bisppektrum sensitiv gegenüber einer quadratischen Phasenkopplung ist. Hinweis: Das Bisppektrum ist nur im dunkel hinterlegten Bereich gezeichnet, alle anderen Oktanten sind aufgrund der Symmetrie redundant.

Aus den obigen Beispielen geht deutlich hervor, daß ein typisches Charakteristikum eines nichtlinearen System, die Generierung von Phasenkopplungen, nur mit Spektren höherer Ordnung identifizierbar ist. Auch für stationäre Prozesse, welche Abhängigkeiten zwischen den Frequenzkomponenten aufweisen, ist eine polyspektrale Analyse notwendig.

Praktische Schätzung des Bisppektrums

Den in den Gln. (5.27) und (5.32) angegebenen Definitionen von Polyspektren liegt die Annahme zugrunde, daß theoretisch beliebig viele Realisierungen vorhanden sind, bzw. daß die Kumulanten exakt bekannt sind. Unter realistischen Bedingungen jedoch, müssen

die Statistiken höherer Ordnung aus einer endlichen Anzahl Realisierungen geschätzt werden. Selbst wenn, wie hier angenommen, die Prozesse als stationär und ergodisch betrachtet werden können, so bleibt immer noch die Anzahl der Realisierungen ein kritischer Aspekt. Zur Schätzung der Statistiken höherer Ordnung können zwei basale Verfahren unterschieden werden: (i) die konventionellen Methoden, die auf der Mittelung verschiedener Realisierungen und der schnellen Fourier-Transformation basieren, und (ii) die parametrischen Methoden, welche die Parameter von Modellen zur Generierung einer entsprechenden Statistik anpassen. Diese Modelle sind häufig vom Typus *moving average (MA)*, *autoregressive (AR)* oder *ARMA*. Während die parametrischen Methoden eine relativ schnelle und hochauflösende Schätzung liefern, so ist die Generalität bei den konventionellen Verfahren höher. Insbesondere, wenn keine *a priori* Annahmen über die Natur der Quelle gemacht werden können, bzw. wollen.

Die konventionellen Methoden wiederum gliedern sich in zwei Subgruppen, entsprechend den angegebenen Definitionen. Die *indirekte Klasse* bezieht sich auf die Definition in Gl. (5.27); hier werden zuerst die Kumulanten des Prozesses geschätzt und anschließend durch Fourier-Transformation die Spektren berechnet. Somit ist diese Methode ähnlich zur bekannten BLACKMAN-TUKEY-Methode zur Schätzung der Autokovarianz, bzw. des Leistungsdichtespektrums (Kay, 1988). Die in dieser Arbeit berechneten Spektren höherer Ordnung sind nach der *direkten Methode* geschätzt. Dabei werden zuerst die Fourier-Koeffizienten der Realisierungen geschätzt, und ihre gemittelten Tripel-Produkte ergeben dann die Schätzung des Bispektrums. Dies ist vergleichbar mit dem Verfahren der *gemittelten Periodogramme*, oder auch WELCH-Methode (Kay, 1988). Diese Methode zeichnet sich durch eine schnellere Konvergenz und vor allem durch eine bessere Berechenbarkeit aus, da bei der indirekten Methode zum einen über sehr viele Datenpunkte gemittelt werden muß und abschließend eine hochdimensionale Fourier-Transformierte berechnet wird.

Die grundlegenden Schritte zur Berechnung des Polyspektrums der Ordnung n können wie folgt zusammengefaßt werden; für eine ausführliche Darstellung, siehe z.B. (Nikias und Petropulu, 1993):

1. Segmentiere die gegebenen Daten in Blöcke mit konstanter Anzahl Abtastwerte,
2. befreie jeden Block vom Mittelwert,
3. multipliziere ggf. jeden Block mit einer geeigneten Fensterfunktion zur Glättung des Spektrums (dies geht zu Lasten der Auflösung, kann jedoch die Interpretation der Ergebnisse substantiell vereinfachen),
4. berechne für jeden Block die Fourier-Transformierte (mittels FFT),
5. bilde von jedem Block die n -tupel Produkte

$$U(f_1)U(f_2)\cdots U(f_{n-1})U(-f_1 - f_2 - \cdots - f_{n-1}),$$

d.h. bilde die Periodogramme,

6. mittele über sämtliche Periodogramme.

Die statistischen Eigenschaften der konventionellen Verfahren wurden studiert in (Hinich, 1982; Rao und Gabr, 1984; Rosenblatt und Van Ness, 1965). Es wurde gezeigt, daß die Verfahren für hinreichend große Blöcke und Blockanzahl konsistente und erwartungstreue Schätzer darstellen, und daß die Schätzwerte asymptotisch normalverteilt sind.

Es seien $C_2^{\mathbf{u}}$, bzw. $C_3^{\mathbf{u}}(f_1, f_2)$ das wahre Leistungsdichtespektrum, bzw. Bispektrum eines streng stationären, mittelwertfreien Prozesses $\mathbf{u}(x)$. Das nach obigem Verfahren geschätzte Bispektrum sei $\hat{C}_3^{\mathbf{u}}(f_1, f_2)$. Dieser Schätzer stellt dann mit hinreichend großen und einer genügenden Anzahl Blöcken einen erwartungstreuen Schätzer dar:

$$E\{\hat{C}_3^{\mathbf{u}}(f_1, f_2)\} \approx C_3^{\mathbf{u}}(f_1, f_2). \quad (5.46)$$

Dieser ist asymptotisch gaußsch mit den Varianzen

$$\begin{aligned} \text{var}\{\Re[\hat{C}_3^{\mathbf{u}}(f_1, f_2)]\} &\approx \text{var}\{\Im[\hat{C}_3^{\mathbf{u}}(f_1, f_2)]\} \\ &\propto \frac{1}{\Delta^2 N} C_2^{\mathbf{u}}(f_1) C_2^{\mathbf{u}}(f_2) C_2^{\mathbf{u}}(f_1 + f_2), \end{aligned} \quad (5.47)$$

wobei Δ den Abstand im Frequenzbereich darstellt (Auflösung) und N die gesamte Länge der verfügbaren Daten. Für eine ausführlichere Darstellung und einen Vergleich zwischen direkter und indirekter Methode, siehe (Petropulu, 1998).

Beispiele

Um die Darstellungen in den nächsten Abschnitten zu vereinfachen, und um die bispekturale Analyse etwas intuitiv gestalten zu können, werden in diesem Abschnitt exemplarisch zwei Bispektren berechnet. Eine erforderliche Voraussetzung für die Diskussion von Bispektren natürlicher Bilder, bzw. nichtlinearer Systemausgänge ist eine geeignet gewählte Darstellung. Da in dieser Arbeit zweidimensionale Signale und Systeme behandelt werden, ergibt sich die Dimensionalität der Bispektren zu vier. Die Darstellung derartiger Signale erfolgt durch gerade Schnitte des kartesischen Raumes und Anordnung dieser entsprechend ihrer Schnittposition. Abb. 5.3 zeigt das kartesische Koordinatensystem dieser Darstellung. Sämtliche gezeigten Bispektren sind entsprechend diesem System dargestellt³.

Beispiel. Gegeben sei ein deterministisches, energiebegrenzt und reelles Signal

$$k_R(x, y) = \delta(\sqrt{x^2 + y^2} - R) \quad (\text{Kreisring mit Radius } R). \quad (5.48)$$

³Das kartesische System wird aufgrund seiner „Vertrautheit“ gewählt, z.T. kann aber eine andere Darstellung, wie etwa Polarkoordinaten, die Zusammenhänge im vierdimensionalen besser aufdecken.

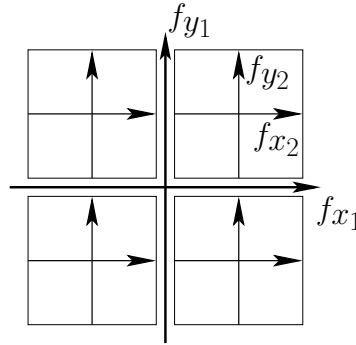


Abbildung 5.3: Koordinatensystem zur Darstellung der vierdimensionalen Bispektren.

Die Fourier-Transformierte läßt sich aufgrund der Ausblendeigenschaft der Dirac-Linie schreiben als

$$K_R(f_x, f_y) = \frac{1}{4\pi^2} \int \int e^{j2\pi(f_x x + f_y y)} dx dy, \quad (5.49)$$

wobei die Integration über den infinitesimalen Bereich des Kreisringes ausgeführt wird. Nach Übergang zu Polarkoordinaten (R, θ) und Anwendung der Amplituden-Phasen-Polarform ergibt sich

$$K_R(f_x, f_y) = \frac{R}{4\pi^2} \int_0^{2\pi} e^{j2\pi R f_r \cos(\theta - \phi)} d\theta, \quad \text{mit } f_r = \sqrt{f_x^2 + f_y^2}, \phi = \arctan \frac{f_y}{f_x}. \quad (5.50)$$

Verwendet man eine bekannte Entwicklung der Besselfunktionen erster Art (Abramowitz und Stegun, 1970), so ergibt sich schlußendlich als Spektrum des Ringes

$$K_R(f_x, f_y) = \frac{R}{2\pi} J_0(2\pi R \sqrt{f_x^2 + f_y^2}), \quad (5.51)$$

wobei J_0 die Besselfunktion erster Art der Ordnung 0 darstellt; siehe Abb. 5.4(a). Sein Bispektrum berechnet sich dann zu

$$C_3^k(\vec{f}_1, \vec{f}_2) = K_R(\vec{f}_1) K_R(\vec{f}_2) K_R(-\vec{f}_1 - \vec{f}_2). \quad (5.52)$$

Genauer betrachtet ist dies das Momenten-Bispektrum des deterministischen Signals $k_R(\vec{x})$; i.A. bringen Kumulantenspektren bei deterministischen Signalen keinen Gewinn. Man sieht, daß dieses Bispektrum vollständig durch die Fourier-Transformierte $K_R(\vec{f})$ beschrieben ist. \square

Beispiel. Das Bispektrum eines linearen Exponentialprozesses mit Leistungsdichte

$$C_2^e(\vec{f}) = e^{-|\vec{f}|}. \quad (5.53)$$

Äquivalent kann dieser Prozeß als ein weißes Rauschen höherer Ordnung, gefiltert mit der linearen Übertragungsfunktion

$$H(\vec{f}) = e^{-|\vec{f}|} \quad (5.54)$$

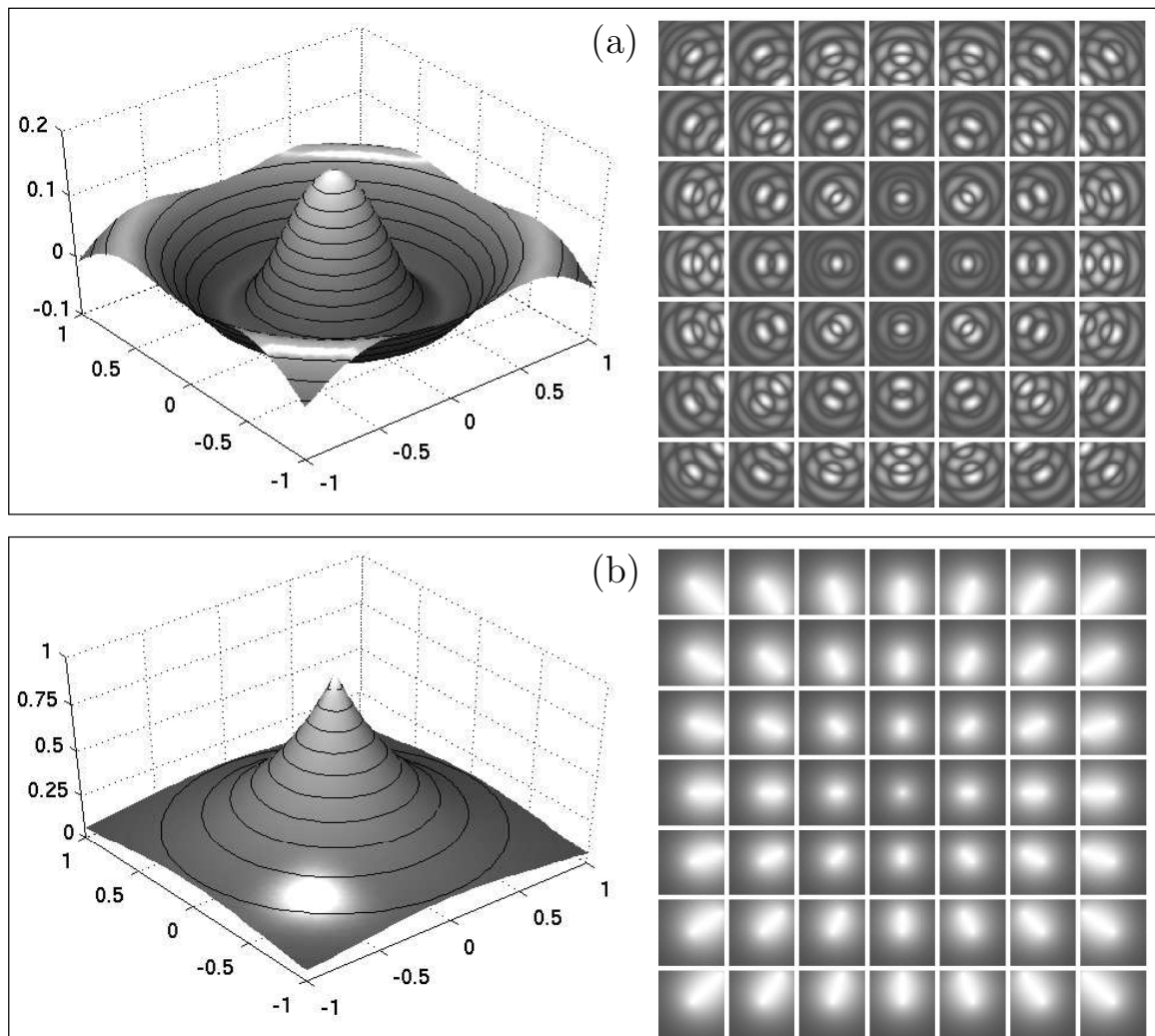


Abbildung 5.4: (a) Fourier-Transformierte und Bisppektrum des deterministischen Energiesignals aus Gl. (5.48). (b) Leistungsdichtespektrum entsprechend Gl. (5.53) und zugehöriges Bisppektrum eines linearen Prozesses.

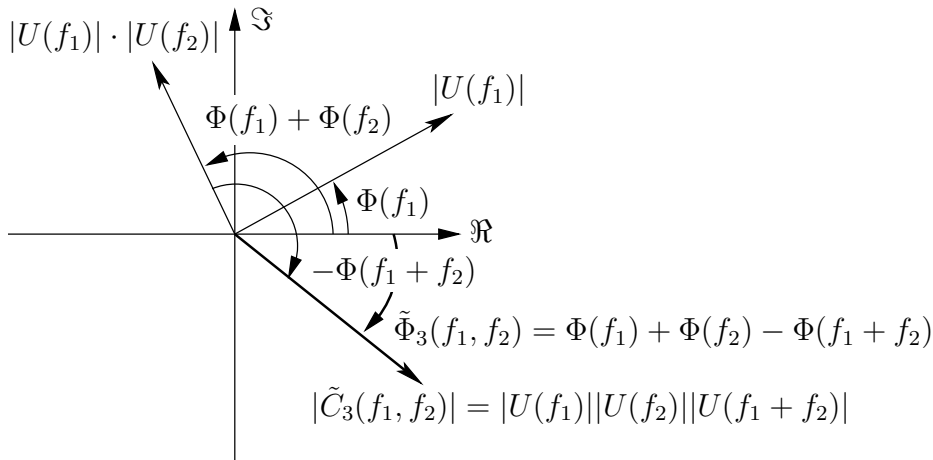


Abbildung 5.5: Sukzessive Berechnung einer Realisierung des Bisppektrums für eine bestimmte Frequenzkombination (f_1, f_2) in der komplexen Ebene. Das finale Bisppektrum ergibt sich durch Mittelung des fett eingezeichneten Vektors.

betrachtet werden. Damit ergibt sich das Bisppektrum entsprechend Gl. (5.39) zu:

$$C_3^e(\vec{f}_1, \vec{f}_2) = e^{-(|\vec{f}_1| + |\vec{f}_2| + |\vec{f}_1 + \vec{f}_2|)}, \quad (5.55)$$

siehe Abb. 5.4(b). Hieran offenbart sich auch eine fundamentale Eigenschaft der Spektren höherer Ordnung. Wird ein Prozeß mit einem linearen Filter manipuliert, so können definierte Abhängigkeiten zweiter Ordnung erzeugt werden. Es werden allerdings damit auch gleichzeitig die Bindungen höherer Ordnung beeinflusst, eine Eigenschaft, die im Allgemeinen nicht wünschenswert ist. \square

Anschauliche Deutung der bisppektralen Schätzung

In den beiden obigen Beispielen handelt es sich sowohl um ein deterministisches Signal als auch um einen stationären Prozeß. Den beiden Bisppektralen gemein ist, daß sie in der Form $U(f_1)U(f_2)U^*(f_1 + f_2)$ separierbar sind. Im ersten Fall ist dies einfach die Vorschrift zur Berechnung des Bisppektrums eines deterministischen Signals, wohingegen im zweiten Fall diese Eigenschaft in der *Linearität* des Prozesses begründet liegt, wie auch bereits mit Gl. (5.39) gezeigt. Wichtig ist hierbei, daß die für uns interessanten Eigenschaften (bei der Analyse der Statistik natürlicher Bilder oder der Identifikation nichtlinearer Systeme) in der *nicht*-Separierbarkeit der Bisppektralen liegen. Diese ergibt sich durch die statistische Mittelung über viele Realisierungen mit bestimmten Phasenkopplungen, wie in Abb. 5.5 illustriert. Dies kann veranschaulicht werden, indem man sich die Berechnung des Betrags und der Phase

$$|C_3(f_1, f_2)| = |U(f_1)| \cdot |U(f_2)| \cdot |U(f_1 + f_2)| \quad (5.56)$$

$$\Phi_3(f_1, f_2) = \Phi(f_1) + \Phi(f_2) - \Phi(f_1 + f_2) \quad (5.57)$$

für eine ganz bestimmte Frequenzkombination (f_1, f_2) schrittweise betrachtet. Abb. 5.5 stellt die Messung einer Realisierung in der komplexen Ebene dar. Wenn die Phase an der dritten Frequenz, $(f_1 + f_2)$ unabhängig ist, so ist der resultierende (fett eingezeichnete) Vektor gleichverteilt in der Ebene⁴, d.h. der Mittelwert ist Null. Ist die Phase jedoch an die Phasen bei f_1 und f_2 gekoppelt (mit einer erlaubten additiven Konstante), so fallen *im Mittel* die resultierenden Vektoren auf eine bestimmte Lage in der komplexen Ebene und bestimmen somit Betrag und Phase des Bispektrums an (f_1, f_2) .

5.1.3 Bispektrale Analyse natürlicher Bilder

Wie weiter oben erläutert, mißt das Bispektrum die Bindungen zwischen Frequenzkomponenten. Es stellt sich nun die Frage, wie das Bispektrum für die Klasse der natürlichen Bilder aussieht. Eine Hypothese ist, daß sich die Statistik unserer Umwelt signifikant aus der Existenz von orientierten Kanten manifestiert. So konnte gezeigt werden, daß die lokale WDF einer simulierten spärlichen Verteilung, bestehend aus orientierten Kanten, eine große Ähnlichkeit mit der real gemessenen aufweist (Zetzsche und Krieger, 2001b). Dies ist auch kohärent mit den eingangs in Kap. 2 ausgeführten Überlegungen zur Redundanz in natürlichen Bildern. Die nicht-redundanten, also informationsträchtigen Signalanteile sind eben *nicht* die häufig vorkommenden Kanten und Linien (Objektgrenzen), sondern v.a. die Änderungen in deren Orientierung, also *i2D*-Signale.

Analog dazu kann im Bispektrum eine sternförmige Struktur entdeckt werden (Krieger et al., 1995; Krieger et al., 1997; Krieger, 1999), was bedeutet, daß vermehrt Frequenzen mit gleicher oder zumindest ähnlicher Orientierung gekoppelt sind. Dies ist ebenfalls durch Kanten zu erklären, da die Fourier-Transformierte einer Kante viele unterschiedliche Radialfrequenzen mit definierter Phasenlage zueinander enthält. In Abb. 5.6 ist das Bispektrum natürlicher Bilder und dessen quantitative Analyse gezeigt. Dazu wurde ein hochauflösendes Bispektrum berechnet⁵ und die „Masse“ innerhalb einer bestimmten Orientierungsdifferenz aufintegriert (vgl. Abb. 5.6):

$$\xi_{nat} = \int_{D_{\Delta\phi}} \int C_3^u(\vec{f}_1, \vec{f}_2) d\vec{f}_1 d\vec{f}_2, \quad (5.58)$$

wobei für den Integrationsbereich $D_{\Delta\phi}$ gilt:

$$D_{\Delta\phi} = \left\{ \vec{f}_1, \vec{f}_2 \in R \mid \left| \arctan \frac{f_{y1}}{f_{x1}} - \arctan \frac{f_{y2}}{f_{x2}} \right| \leq \frac{\Delta\phi}{2} \right\}, \quad \vec{f}_i = (f_{x_i}, f_{y_i})^T. \quad (5.59)$$

⁴Dies geht aus der Fourier-Stieltjes Repräsentation hervor, vgl. Gl. (5.30)

⁵Das bedeutet 64 Frequenzstützstellen pro Dimension, und entspricht einem Speicherbedarf von $64^4 \times (32 \cdot 2) \text{ bit} = 128 \text{ MByte}$ (je eine Gleitkommazahl mit 32 bit für Real- und Imaginärteil). Eine höherer Auflösung verlangt exponentiell mehr Speicher- und Rechenressourcen und erhöht zudem die Varianz des Schätzers.

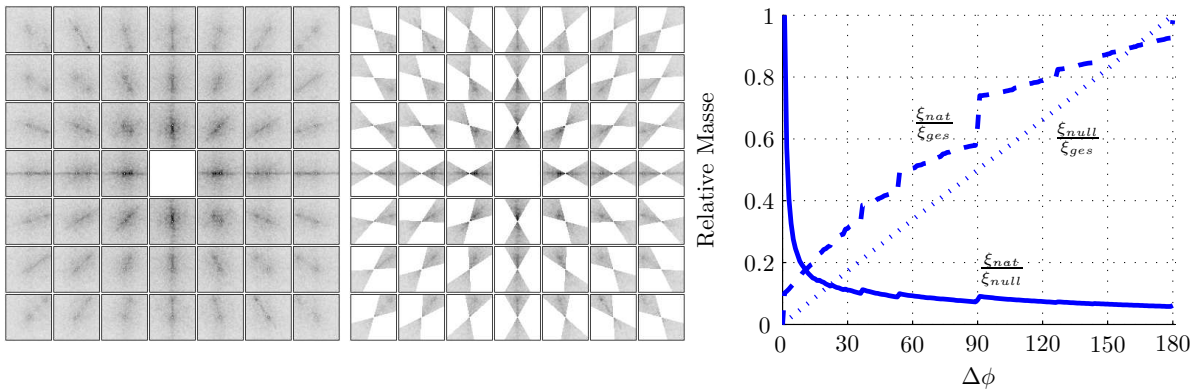


Abbildung 5.6: Analyse des Bisppektrums natürlicher Bilder. Links: gemessenes hochauflösendes Bisppektrum, mittig: Integrationsbereich für $\Delta\phi = 60^\circ$, rechts: relative Massen in Abhängigkeit von der Orientierungsdifferenz (Details siehe Text).

Diese Masse in Abhängigkeit von der Orientierungsdifferenz wird ins Verhältnis mit der Gesamtmasse ξ_{ges} gesetzt. Erkennbar ist ein konkaver Verlauf (gestrichelt), der über der Nullhypothese ξ_{null} eines konstanten Bisppektrums liegt (gepunktet). Das Verhältnis dieser beiden Kurven (durchgezogen) offenbart die starken Bindungen benachbarter Orientierungen, da für kleine Orientierungsunterschiede $\Delta\phi$ relativ betrachtet am meisten Masse im Bisppektrum zu finden ist.

Diese Untersuchung stellt somit eine konsequente Erweiterung der bisherigen Erkenntnisse im Sinne einer quantitativen Analyse dar. Die Ergebnisse liefern weitere Evidenz, daß lokale Strukturen in natürlichen Bildern überwiegend aus gerichteten Kanten und anderen *1D*-Signalanteilen bestehen.

5.2 Volterra-Wiener-Systeme

Models should not be true but it is important that they are applicable.

John W. Tukey

Nichtlineare Systeme sind alle Systeme, die *nicht* linear sind. In dieser augenscheinlich trivialen Feststellung sind jedoch bereits die fundamentalen Schwierigkeiten, die der Umgang mit nichtlinearen Systemen bereithält, aggregiert. Alle *linearen* Systeme gehorchen dem Superpositionsprinzip, also gilt für beliebige Konstanten $\alpha, \beta \in \mathbb{R}$:

$$\mathcal{L}\{\alpha u_1(\vec{x}) + \beta u_2(\vec{x})\} = \alpha \mathcal{L}\{u_1(\vec{x})\} + \beta \mathcal{L}\{u_2(\vec{x})\}, \quad (5.60)$$

wobei $\mathcal{L}\{\cdot\}$ den Ausgang eines linearen Operators und $u_i(\vec{x})$ verschiedene Eingangssignale bezeichnen. Das bedeutet, die Antwort auf die Summe mehrerer Signale ist gleich

der Summe der Antworten auf die einzelnen Signale. Dies ist eine schwerwiegende Einschränkung und ermöglicht es, sämtliche lineare Systeme in einem fundierten, geschlossenen mathematischen Rahmen zu behandeln. Nichtsdestotrotz gehorchen nichtlineare Systeme dieser Einschränkung nicht, und da *alle* Systeme, die diesem Prinzip nicht gehorchen, nichtlinear sind, gibt es keine Möglichkeit, alle derartige Systeme in einem einheitlichen Rahmen zu behandeln. Daher werden traditionellerweise jeweils unterschiedliche Klassen von nichtlinearen Systemen mit gemeinsamen Eigenschaften definiert, und für diese dann explizite Ansätze, bzw. Formulierungen angegeben. Polynomiale Systeme bilden eine derartige Klasse nichtlinearer Systeme. Diese können generell wie folgt definiert werden.

Definition. Es sei $u_1(\vec{x})$ ein bandbegrenzter Prozeß der Dimensionalität N . Bei geeignet gewählter Abtastrate $1/\vec{\Delta}_x$ (siehe Gl. (2.5)) kann dieser als diskrete Folge

$$u_1(\vec{x}) = \{\dots, u_1(\vec{x}_0 - \vec{\Delta}_x), u_1(\vec{x}_0), u_1(\vec{x}_0 + \vec{\Delta}_x), \dots\} \quad (5.61)$$

repräsentiert werden. Polynomiale Systeme der Ordnung P sind dann definiert durch Eingangs-Ausgangs-Relationen der Form (Mathews und Sicuranza, 2000)⁶

$$\begin{aligned} u_2(\vec{x}_0) &= \mathcal{P}\{u_1(\vec{x})\} \\ &= \sum_{i=0}^P f_i\{u_1(\vec{x}_0 - \frac{K}{2}\vec{\Delta}_x), \dots, u_1(\vec{x}_0), \dots, u_1(\vec{x}_0 + \frac{K}{2}\vec{\Delta}_x), \\ &\quad u_2(\vec{x}_0 - \frac{L}{2}\vec{\Delta}_x), \dots, u_2(\vec{x}_0), \dots, u_2(\vec{x}_0 + \frac{L}{2}\vec{\Delta}_x)\}, \end{aligned} \quad (5.62)$$

wobei $u_2(\vec{x}_0)$ den Ausgang darstellt, $f_i(\dots)$ ein Polynom der Ordnung i ist und K und L die Größe des Definitionsbereichs (Support) des Systems im Vorwärts-, bzw. Rückwärtszweig bestimmen. □

Wenn das so in Gl. (5.62) definierte System *BIBO*-stabil ist (bounded-input bounded-output), so kann dies in eine Volterra-Reihe entwickelt werden. Der Begriff *Volterra-Wiener-Systeme* bezeichnet die Klasse nichtlinearer Systeme, die aus der Volterra-Reihenentwicklung hervorgehen.

5.2.1 Kurzer geschichtlicher Überblick

VITO VOLTERRA, ein italienischer Mathematiker (*1860, †1940), studierte die nach ihm benannte Funktionalreihe als eine Erweiterung der bekannten Taylor-Reihenentwicklung. Seine ersten Ergebnisse publizierte er im Alter von 27 Jahren (Volterra, 1887). FRÉCHET konnte bald darauf zeigen, daß die Volterra-Funktionale vollständig sind (Fréchet, 1910).

⁶Die hier verwendete Definition ist akausal, da in dieser Arbeit örtliche Signale und Systeme betrachtet werden, und Kausalität hierbei irrelevant ist.

Sein Theorem impliziert, daß jedes kontinuierliche Funktional eines Signals $x(t)$ mit beliebiger Genauigkeit durch die Summe einer endlichen Anzahl Volterra-Funktionale in $x(t)$ approximiert werden kann. Dies ist eine Weiterentwicklung des Theorems von WEIERSTRASS, das besagt, daß jede kontinuierliche Funktion mit beliebiger Genauigkeit durch eine endliche Anzahl Polynome in x angenähert werden kann. Die erste Verwendung in einer nichtlinearen Systemtheorie fand die Volterra'sche Arbeit zu Beginn der 1940er Jahre. NORBERT WIENER, der berühmte Systemtheoretiker und Begründer der Kybernetik, analysierte nichtlineare Systeme, indem er die Volterra-Reihe für einen weißen, gaußschen Eingang orthogonalisierte (Wiener, 1958). Die resultierenden, von ihm benannten *G-Funktionale* haben den Vorteil, daß die Terme unterschiedlicher Ordnung im Antwortverhalten separiert werden können (siehe Abschnitt 5.2.5).

Viele darauffolgende Arbeiten in den 60er und 70er Jahren mußten sich der Rechenintensivität der Volterra-Systeme beugen, da die zu dieser Zeit existierenden Rechenmaschinen dies nicht bewältigen konnten. Erst mit dem Aufkommen schneller Rechner in den 80er Jahren erfuhr die Modellierung und Synthese mit Volterra-Reihen eine Renaissance. Zu nennen sind hier Arbeiten auf den Gebieten der adaptiven Echo- und Rauschunterdrückung, blinden Entzerrung von Übertragungskanälen und Modellierung digitaler Satellitenkanäle. Auch turbulente Strömungen und das Treibverhalten von Schiffen im Meer konnten mit Hilfe der Volterra-Wiener-Theorie modelliert werden. Weitere Arbeitsgebiete betreffen Populationsstudien, Kernspaltung und Immunologie.

Auf dem Gebiet der biologischen Systeme finden sich ebenfalls einige Anwendungen der Volterra-Modelle. Einen Überblick kann man in (Marmarelis und Marmarelis, 1978) erhalten. Die Anwendung beschränkt sich hierbei hauptsächlich auf die Analyse zeitlicher neuronaler Information und auf sensorische und motorische Neurone. Ein häufig verwendeter Ansatz sind reduzierte Spezialfälle von Volterra-Systemen, die nichtlinearen Kaskadensysteme, vorwiegend bestehend aus linearen Systemen in Verbindung mit einer statischen Nichtlinearität (Hunter und Korenberg, 1986; Korenberg und Hunter, 1986). Insbesondere im visuellen System wurden nichtlineare Interaktionen mit sogenannten „Wiener-like kernels“ gemessen und beschrieben. So wurden beispielsweise die Kerne zweiter Ordnung von Complex-Zellen gemessen (Emerson et al., 1987; Szulborski und Palmer, 1990) oder die Interaktionen von Bewegungsrichtungen im mediotemporalen Kortex (MT), dem die Verarbeitung von Bewegung zugesprochen wird (Livingstone et al., 2001; Pack et al., 2003a).

Funktionale Beschreibungen von visuellen Neuronen unter Zuhilfenahme von Volterra-Reihen werden von ZETZSCHE und Mitarbeitern vorgeschlagen. Ein Schlüsselkonzept hierbei ist die Möglichkeit einer UND-artigen Verknüpfung von Frequenzkomponenten; diese stellt einen fundamentalen Unterschied zu den klassischen linearen Systemen dar. Diese Eigenschaft öffnet das Tor zu einer großen Klasse von funktionalen Zusammenhängen, die zur Analyse und Modellierung einer Vielzahl nichtlinearer kortikaler Phänomene verwendet werden können, siehe hierzu Abschnitt 5.2.3.

5.2.2 Die Volterra-Reihenentwicklung

In diesem Abschnitt erfolgt eine kurze Einführung in die Volterra-Reihenentwicklung und die damit verbundenen Signalverarbeitungseigenschaften. In dieser Arbeit liegt die Betonung auf quadratischen Volterra-Systemen für zweidimensionale Signale, daher werden diese etwas ausführlicher dargestellt. Im Gegensatz zu der klassischen Modellierung und Analyse im Zeit, bzw. Ortsbereich werden die Systeme hier nahezu ausschließlich im Frequenzbereich dargestellt. Desweiteren wird aus Gründen der Einfachheit grundsätzlich von translationsinvarianten und im Allgemeinen akausalen Systemen ausgegangen. Insbesondere die Kausalität ist für örtliche Systeme ohnehin irrelevant.

Eine grundlegende Ausarbeitung der Theorie der Volterra-Wiener Systeme findet sich in (Schetzen, 1980; Rugh, 1981). Für ausführliche Beispiele zu quadratischen Volterra-Systemen für zweidimensionale Signale siehe (Krieger, 1999; Nuding, 2002).

Die Volterra-Reihe im Ortsbereich

Gegeben sei ein Operator $\mathcal{H}\{\cdot\}$, der ein (mehrdimensionales) Eingangssignal $u_1(\vec{x})$ auf ein Ausgangssignal $u_2(\vec{x})$ abbildet:

$$u_2(\vec{x}) = \mathcal{H}\{u_1(\vec{x})\}, \quad (5.63)$$

wobei $\vec{x} \in \mathbb{R}^N$ von der Natur, bzw. Dimensionalität N des Prozesses abhängt, also ergibt sich für Zeitreihen $N = 1, \vec{x} = t$ oder für Bilder $N = 2, \vec{x} = (x, y)^T$. Dieser Operator sei zudem *translationsinvariant*, d.h.

$$u_2(\vec{x} + \vec{\Delta}_x) = \mathcal{H}\{u_1(\vec{x} + \vec{\Delta}_x)\}. \quad (5.64)$$

Dann kann die Volterra-Reihenentwicklung für diesen Operator wie folgt angegeben werden:

$$\begin{aligned} u_2(\vec{x}) = & h_0 \\ & + \int_{\mathbb{R}^N} h_1(\vec{x}_1) \cdot u_1(\vec{x} - \vec{x}_1) \cdot d\vec{x}_1 \\ & + \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} h_2(\vec{x}_1, \vec{x}_2) \cdot u_1(\vec{x} - \vec{x}_1) u_1(\vec{x} - \vec{x}_2) \cdot d\vec{x}_1 d\vec{x}_2 \\ & + \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} h_3(\vec{x}_1, \vec{x}_2, \vec{x}_3) \cdot u_1(\vec{x} - \vec{x}_1) u_1(\vec{x} - \vec{x}_2) u_1(\vec{x} - \vec{x}_3) \cdot d\vec{x}_1 d\vec{x}_2 d\vec{x}_3 \\ & + \dots \end{aligned} \quad (5.65)$$

Diese Reihe stellt im Prinzip eine Taylor-Reihenentwicklung für Systeme mit Gedächtnis dar. Die $h_i(\vec{x}_1, \dots, \vec{x}_i)$ werden hierbei als Volterra-Kerne der Ordnung i bezeichnet. Es ist

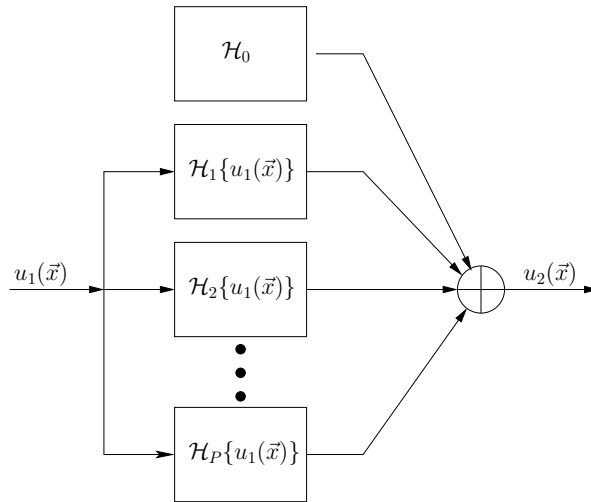


Abbildung 5.7: Darstellung der Gesamtantwort eines nichtlinearen Systems der Ordnung P als lineare Überlagerung der Einzelantworten der homogenen Volterra-Kerne.

offensichtlich, daß sich Gl. (5.65) aus der linearen Superposition von Volterra-Operatoren unterschiedlicher Ordnung ergibt, wenn man formal definiert

$$\mathcal{H}_p\{u_1(\vec{x})\} := \int_{R^{pN}} h_p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) \cdot \prod_{i=1}^p u_1(\vec{x} - \vec{x}_i) \cdot d\vec{x}_i. \quad (5.66)$$

Somit kann die Gesamtantwort des Operators $\mathcal{H}\{u_1(\vec{x})\}$ als Summe von im Grenzfall unendlich vielen Teiloperatoren $\mathcal{H}_p\{u_1(\vec{x})\}$ der Ordnung p ausgedrückt werden (vgl. auch Abb. 5.7):

$$\mathcal{H}\{u_1(\vec{x})\} = \sum_{p=0}^{\infty} \mathcal{H}_p\{u_1(\vec{x})\}. \quad (5.67)$$

Die Antwort von \mathcal{H}_0 ist unabhängig vom Eingang und stellt einen systeminhärenten Gleichanteil dar. Der Beitrag von \mathcal{H}_1 berechnet sich entsprechend Gl. (5.65) identisch einem linearen System als bekannte Faltungsoperation und stellt den linearen Anteil dar. Dementsprechend indiziert \mathcal{H}_2 ein homogenes quadratisches und \mathcal{H}_3 ein kubisches System.

Wie man Gl. (5.66) entnehmen kann, ist die Volterra-Reihe linear in ihren Kernen. Dies wird besonders dann evident, wenn die Operation in einem expandierten Raum betrachtet wird. Dazu definiert man das expandierte Eingangssignal als dyadisches Produkt

$$\tilde{u}_1(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = u_1(\vec{x}_1) \cdot u_1(\vec{x}_2) \cdots u_1(\vec{x}_p), \quad (5.68)$$

setzt dies in Gl. (5.66) ein und vergleicht das Ergebnis mit der bekannten mehrdimensionalen Faltungsoperation, so erkennt man, daß diese identisch sind für den Fall

$\vec{x}_1 = \vec{x}_2 = \dots = \vec{x}_n =: \vec{x}$:

$$\mathcal{H}_p\{u_1(\vec{x})\} = h_p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) \overset{\vec{x}_1 \vec{x}_2}{\star \star} \dots \overset{\vec{x}_p}{\star} \tilde{u}_1(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) \Big|_{\vec{x}=\vec{x}_1=\vec{x}_2=\dots=\vec{x}_n}, \quad (5.69)$$

wobei \star die Faltungsoperation symbolisiert. In anderen Worten bedeutet dies, daß die Antwort eines (homogenen) Volterra-Systems in drei Schritten berechnet werden kann:

1. Bilde das expandierte Eingangssignal $\tilde{u}_1(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = u_1(\vec{x}_1) \cdot u_1(\vec{x}_2) \cdot \dots \cdot u_1(\vec{x}_p)$,
2. berechne die Faltung zwischen dem expandierten Eingangssignal und dem entsprechenden Volterra-Kern im expandierten Raum,
3. schneide entlang der Hauptdiagonalen $\vec{x}_1 = \vec{x}_2 = \dots = \vec{x}_n =: \vec{x}$ im expandierten Raum.

Da das Produkt in Gl. (5.68) sowie die Faltung in Gl. (5.69) kommutativ sind, ist die Systemantwort $\mathcal{H}_p\{u_1(\vec{x})\}$ invariant gegenüber einer Permutation der Argumente im Kern $h_p(\vec{x}_1, \dots, \vec{x}_p)$. Damit existiert für jeden Kern h_p eine große Äquivalenzklasse von Kernen mit identischen Systemantworten. Diese Ambiguität kann überwunden werden, indem über sämtliche $p!$ Kerne mit permutierten Argumenten summiert und anschließend durch die Anzahl der Permutationen dividiert wird:

$$h_p^{sym}(\vec{x}_1, \dots, \vec{x}_p) = \frac{1}{p!} \sum_{\{\varphi_i\}} h_p(\varphi_i(\vec{x}_1, \dots, \vec{x}_p)). \quad (5.70)$$

Dabei stellt h_p^{sym} den eindeutigen symmetrischen Kern zu allen äquivalenten h_p dar und $\{\varphi_i\}$ die Menge aller möglichen Permutationen der Argumente $(\vec{x}_1, \dots, \vec{x}_p)$.

Im Folgenden ein paar einfache eindimensionale Beispiele zur Illustration der Volterra-Reihe.

Beispiel. Gegeben sei ein Quadrierer im Zeitbereich $u_2(t) = u_1^2(t)$. Dies kann unter Verwendung der Ausblendeigenschaft der Dirac-Funktion leicht auf eine Integralform gebracht werden:

$$u_2(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(t_1)\delta(t_2) \cdot u_1(t-t_1)u_1(t-t_2) \cdot dt_1 dt_2. \quad (5.71)$$

Durch Vergleich mit Gl. (5.66) kann sofort erkannt werden, daß für die Volterra-Reihenentwicklung des Quadrierers ausschließlich ein homogenes System zweiter Ordnung genügt⁷:

$$h_2^{Quad}(t_1, t_2) = \delta(t_1) \cdot \delta(t_2).$$

□

⁷Dies ist zugegebenermaßen trivial.

Beispiel. Gegeben sei ein homogenes Volterra-System zweiter Ordnung, definiert durch den Kern aus Abb. 5.8

$$h_2(t_1, t_2) = e^{-(t_1^2+t_2^2)} \cdot (3t_1^2 + t_2^2). \quad (5.72)$$

Der symmetrische Kern ergibt sich zu

$$\begin{aligned} h_2^{sym}(t_1, t_2) &= \frac{1}{2} [h_2(t_1, t_2) + h_2(t_2, t_1)] \\ &= 2e^{-(t_1^2+t_2^2)} \cdot (t_1^2 + t_2^2). \end{aligned} \quad (5.73)$$

Es werde nun das System mit einem Gauß-Puls

$$u_1(t) = e^{-3t^2} \quad (5.74)$$

beaufschlagt. Das expandierte Ausgangssignal erhält man wie oben beschrieben durch die Faltung des Kerns mit dem expandierten Eingangssignal:

$$\tilde{u}_2(t_1, t_2) = 2 \cdot \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(\tau_1^2+\tau_2^2)} (\tau_1^2 + \tau_2^2) \cdot e^{-3(t_1-\tau_1)^2} e^{-3(t_2-\tau_2)^2} \cdot d\tau_1 d\tau_2, \quad (5.75)$$

was sich nach einer etwas länglichen Rechnung schreiben läßt als

$$\tilde{u}_2(t_1, t_2) = \frac{\pi}{32} e^{-\frac{3}{4}(t_1^2+t_2^2)} \cdot (9[t_1^2 + t_2^2] + 4). \quad (5.76)$$

Damit ergibt sich schließlich als Endresultat (siehe Abb. 5.8)

$$u_2(t) = \tilde{u}_2(t_1, t_2) \Big|_{t_1=t_2=t} = \frac{\pi}{16} e^{-\frac{3}{2}t^2} \cdot (9t^2 + 2). \quad (5.77)$$

□

Beispiel. Wiener-System. Gegeben sei eine Kaskadenstruktur, bestehend aus einem linearen Filter $h(t)$ und einer statischen (gedächtnislosen) Nichtlinearität $v = f(w)$. Derartige Strukturen werden gemeinhin als *Wiener-System* bezeichnet. Die Antwort u_2 eines Systems mit der Nichtlinearität $v = w^2 + 5w - 4$ auf einen Eingang u_1 berechnet sich in Abhängigkeit der linearen Impulsantwort zu (vgl. Abb. 5.9(a)).

$$\begin{aligned} u_2(t) &= u_m^2(t) + 5u_m(t) - 4 \\ &= \left(\int_{-\infty}^{\infty} h(t_1)u_1(t-t_1) \cdot dt_1 \right)^2 + 5 \int_{-\infty}^{\infty} h(t_1)u_1(t-t_1) \cdot dt_1 - 4 \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \underbrace{h(t_1)h(t_2)}_{h_2(t_1, t_2)} \cdot u_1(t-t_1)u_1(t-t_2) \cdot dt_1 dt_2 + \int_{-\infty}^{\infty} \underbrace{5h(t_1)}_{h_1(t_1)} u_1(t-t_1) \cdot dt_1 - \underbrace{4}_{h_0}. \end{aligned}$$

Allgemein gilt, daß für ein Wiener-System sämtliche Kerne linear separierbar sind und sich diese als äußeres Produkt des linearen Kerns, d.h.

$$h_p(t_1, \dots, t_p) = h(t_1) \cdots h(t_p),$$

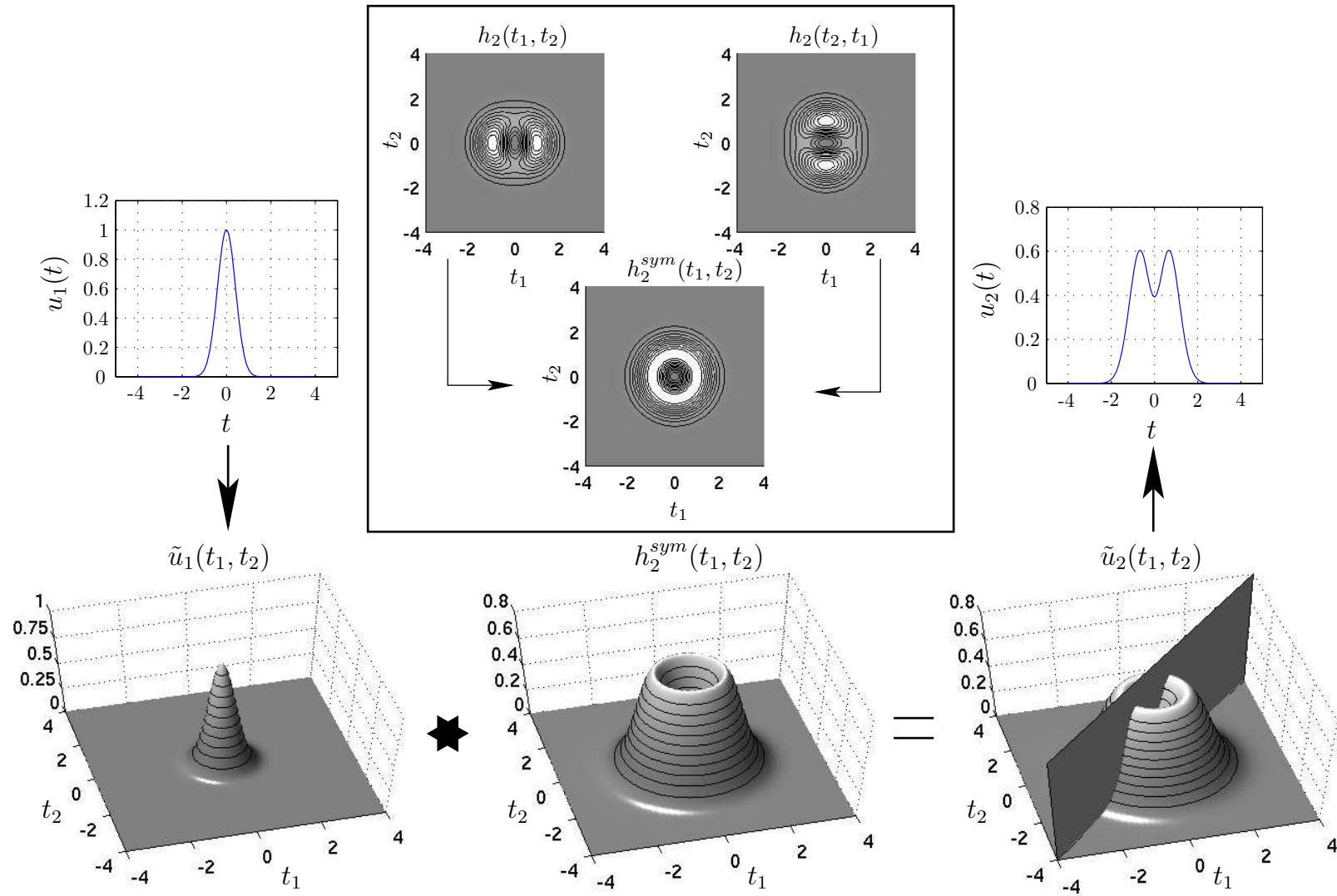


Abbildung 5.8: Illustration der Symmetrisierung des Kerns sowie der Berechnungsvorschrift eines homogenen quadratischen Volterra-Systems anhand des Beispiels auf Seite 107.

gewichtet mit den Koeffizienten der Polynomentwicklung der Nichtlinearität, berechnen. Der quadratische Kern dieses Systems ist in Abb. 5.9(c) für eine exemplarische Wahl von $h(t)$ gezeigt. \square

Beispiel. Hammerstein System. Die umgekehrte Struktur des Wiener-Systems, also statische Nichtlinearität gefolgt von einem linearen System, wird als *Hammerstein-System* bezeichnet (vgl. Abb. 5.9(b)). Wenn als Beispiel wiederum die oben angegebene Nichtlinearität verwendet wird, so berechnet sich die Antwort analog zu oben, und ebenfalls unter Verwendung der Ausblendeigenschaft der Dirac-Funktion, zu

$$\begin{aligned}
 u_2(t) &= \int_{-\infty}^{\infty} h(t_1) u_m(t-t_1) \cdot dt_1 \\
 &= \int_{-\infty}^{\infty} h(t_1) \left(u_1^2(t-t_1) + 5u_1(t-t_1) - 4 \right) \cdot dt_1 \\
 &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \underbrace{\sqrt{h(t_1)h(t_2)}\delta(t_1-t_2)}_{h_2(t_1,t_2)} \cdot u_1(t-t_1)u_1(t-t_2) \cdot dt_1 dt_2 \\
 &\quad + \int_{-\infty}^{\infty} \underbrace{5h(t_1)}_{h_1(t_1)} u_1(t-t_1) \cdot dt_1 - \underbrace{\int_{-\infty}^{\infty} 4h(t_1) \cdot dt_1}_{h_0}.
 \end{aligned}$$

Das bedeutet, daß sich die Kerne höherer Ordnung eines Hammerstein-Systems als eine Art Streckung des Betrags des linearen Kerns $h(t)$ entlang der Hauptdiagonalen $t_1 = \dots = t_n$ ergeben. Der quadratische Kern des Hammerstein-Systems ist für eine exemplarische Wahl von $h(t)$ in Abb. 5.9(d) gezeigt \square

Eigenschaften der Volterra-Reihenentwicklung

An den vorangegangenen Beispielen konnten bereits eine Reihe grundlegender Eigenschaften der Volterra-Reihe identifiziert werden:

1. Linearität bezüglich der Kerne,
2. Faltungseigenschaft im expandierten Raum,
3. Symmetrie der Kerne.

Eine weitere fundamentale Eigenschaft ist, daß die Impulsantwort, im Gegensatz zu linearen Systemen, nicht genügt, um ein Volterra-System höherer Ordnung zu charakterisieren. Die Impulsantwort eines Systems der Ordnung P ergibt sich nämlich zu

$$h(\vec{x}) = h_0 + h_1(\vec{x}) + h_2(\vec{x}, \vec{x}) + \dots + h_P(\vec{x}, \dots, \vec{x}), \quad (5.78)$$

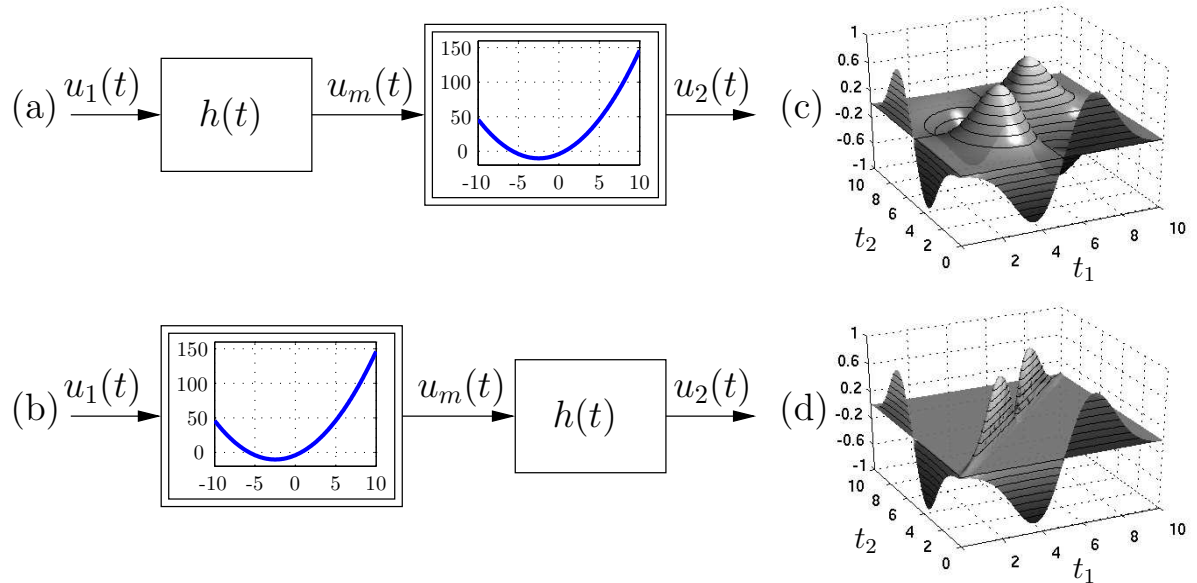


Abbildung 5.9: Beispiele für (a) Wiener- und (b) Hammerstein-Modelle unter Verwendung einer statischen Nichtlinearität $v = w^2 + 5w - 4$. (c,d) die zugehörigen rein quadratischen Kerne bei einem $h(t) = (t - 5) \cdot \exp(-0.25(t - 5)^2)$.

was bedeutet, daß *nur* die Einträge auf den Hauptdiagonalen *aller* Kerne zur Impulsantwort beitragen. Soll ein System mit Hilfe von Impulsantworten charakterisiert werden, so müssen n-tupel von gepaarten Dirac-Impulsen verwendet werden, dies kann für ein rein quadratisches System leicht gezeigt werden, indem man zuerst als Eingang ein zusammengesetztes Signal definiert:

$$u_1(\vec{x}) = u_{1,a}(\vec{x}) + u_{1,b}(\vec{x}). \quad (5.79)$$

Die Antwort eines homogenen quadratischen Systems kann nun angegeben werden als

$$u_2(\vec{x}) = u_{2,a}(\vec{x}) + u_{2,b}(\vec{x}) + 2\bar{h}_{ab}\{u_{1,a}(\vec{x}), u_{1,b}(\vec{x})\}, \quad (5.80)$$

wobei $u_{2,a}(\vec{x})$ und $u_{2,b}(\vec{x})$ die Antworten auf die jeweils einzelnen Signale $u_{1,a}(\vec{x})$, bzw. $u_{1,b}(\vec{x})$ darstellen und

$$\bar{h}_{ab}\{u_{1,a}(\vec{x}), u_{1,b}(\vec{x})\} = \int_{R^N} \int_{R^N} h_2(\vec{x}_1, \vec{x}_2) \cdot u_{1,a}(\vec{x} - \vec{x}_1) u_{1,b}(\vec{x} - \vec{x}_2) \cdot d\vec{x}_1 d\vec{x}_2 \quad (5.81)$$

als sogenannter bilinearer Operator bezeichnet wird. Dieser stellt also den Interaktionsanteil zwischen den Teilsignalen dar. Wenn man nun annimmt, daß der Eingang aus der Summe zweier Dirac-Impulse an unterschiedlichen Stellen besteht, also

$$\begin{aligned} u_{1,a}(\vec{x}) &= \delta(\vec{x} - \vec{x}_a) \\ u_{1,b}(\vec{x}) &= \delta(\vec{x} - \vec{x}_b), \end{aligned} \quad (5.82)$$

so kann man durch Auswertung von Gl. (5.81) leicht erkennen, daß hierfür gilt:

$$\bar{h}_{ab}\{\delta(\vec{x} - \vec{x}_a), \delta(\vec{x} - \vec{x}_b)\} = h_2(\vec{x} - \vec{x}_a, \vec{x} - \vec{x}_b). \quad (5.83)$$

Dieses Ergebnis stellt also ein Identifikationsverfahren für quadratische Systeme dar. Für den Fall von Systemen höherer Ordnung müssen dementsprechend n-tupel von gepaarten Dirac-Impulsen verwendet werden, dies ist ausführlich in (Schetzen, 1980) dargestellt. Im Allgemeinen Fall muß zudem von inhomogenen Systemen ausgegangen werden, hierbei muß dann beachtet werden, daß die Anteile der Kerne höherer Ordnung von der aktuell gemessenen Antwort subtrahiert werden müssen. Allerdings ist dieses Vorgehen extrem unpraktisch, v.a. wegen der exponentiell steigenden Komplexität der Volterra-Kerne. Effiziente Verfahren zur Identifikation von Volterra-Systemen sind in Abschnitt 5.2.5 behandelt.

Außerdem ist als eine weitere Eigenschaft der Volterra-Reihe die Ähnlichkeit zu multidimensionalen linearen Systemen zu nennen. Wie in Gl. (5.69) zum Ausdruck kommt, kann ein N -dimensionales homogenes Volterra-System der Ordnung P äquivalent als ein NP -dimensionales lineares System betrachtet werden. Als zusätzliche Randbedingung gilt dann für den Eingang, daß sich dieser aus dem äußeren Produkt von P identischen N -dimensionalen Signalen ergibt. Das bedeutet, daß bekannte Verfahren für das Design und die Implementierung von multidimensionalen linearen Filtern direkt auf nichtlineare Volterra-Systeme übertragbar sind.

Abschließend sollte noch angemerkt werden, daß die Modellierung von nichtlinearen Systemen mittels Volterra-Reihen auch einige Schwierigkeiten bereithält. So unterliegt die Konvergenzeigenschaft von Volterra-Reihen den gleichen mathematischen Bedingungen wie die der Taylor-Reihe. Das bedeutet, daß harte Nichtlinearitäten im mathematischen Sinn nicht exakt modellierbar sind. Abhängig von der Applikationsdomäne kann dies dann zu einem kritischen Problem werden (Mathews und Sicuranza, 2000).

Die Volterra-Reihe im Frequenzbereich

Ein wesentlicher Aspekt dieser Arbeit ist die Betrachtung und Modellierung von nichtlinearen neuronalen Systemen im Frequenzbereich. Einige wichtige Eigenschaften visueller Neurone können relativ leicht im Fourier-Transformationsbereich offenbart werden, wohingegen die örtliche Darstellung relevante Merkmale „verschleiert“. Dabei wird die folgende Notation verwendet:

$$u_i(\vec{x}) \quad \circ \text{---} \bullet \quad U_i(\vec{f}) \quad (5.84)$$

$$\tilde{u}_i(\vec{x}_1, \dots, \vec{x}_p) := u_i(\vec{x}_1) \cdots u_i(\vec{x}_p) \quad \circ \text{---} \bullet \quad U_i(\vec{f}_1) \cdots U_i(\vec{f}_p) =: \tilde{U}_i(\vec{f}_1, \dots, \vec{f}_p) \quad (5.85)$$

$$h_p(\vec{x}_1, \dots, \vec{x}_p) \quad \circ \text{---} \bullet \quad H_p(\vec{f}_1, \dots, \vec{f}_p), \quad (5.86)$$

wobei $\circ \rightarrow \bullet$ die Fourier-Transformierte kennzeichnet, also stellt beispielsweise

$$H_p(\vec{f}_1, \dots, \vec{f}_p) = \int_{R^N} \cdots \int_{R^N} h_p(\vec{x}_1, \dots, \vec{x}_p) e^{-j2\pi(f_1^T \vec{x}_1 + \cdots + f_p^T \vec{x}_p)} \cdot d\vec{x}_1 \cdots d\vec{x}_p. \quad (5.87)$$

den Volterra-Kern im Frequenzbereich oder auch die *verallgemeinerte Übertragungsfunktion* dar. Bei der Betrachtung der Signale im expandierten Raum ergibt sich die Antwort eines homogenen Teiloperators \mathcal{H}_p als Faltung zwischen dem Kern der Ordnung p und dem expandierten Eingangssignal (vergleiche auch die Gln. (5.68) und (5.69)):

$$\tilde{u}_2(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) = h_p(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p) \overset{\vec{x}_1 \vec{x}_2}{\star \star} \cdots \overset{\vec{x}_p}{\star} \tilde{u}_1(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_p). \quad (5.88)$$

Nach grundlegenden Korrespondenzen der Fourier-Transformation entspricht dies im Frequenzbereich einer einfachen Multiplikation:

$$\tilde{U}_2(\vec{f}_1, \dots, \vec{f}_p) = H_p(\vec{f}_1, \dots, \vec{f}_p) \cdot \tilde{U}_1(\vec{f}_1, \dots, \vec{f}_p). \quad (5.89)$$

Wie bereits anhand Gl. (5.69) erläutert, erhält man die Antwort im Originalraum durch einen Schnitt entlang der Hauptdiagonalen, was gemäß des Zentralschnitt-Theorems im Frequenzbereich mit einer dazu orthogonalen Projektion korrespondiert (Marko, 1994). Damit erhält man das Spektrum des Ausgangssignal $U_2(\vec{f})$ aus dem expandierten Ausgangsspektrum $\tilde{U}_2(\vec{f}_1, \dots, \vec{f}_p)$ durch die Projektion auf die Hauptdiagonale⁸:

$$\begin{aligned} U_2(\vec{f}) &= \int_{R^N} \cdots \int_{R^N} \tilde{U}_2(\vec{f}_1, \dots, \vec{f}_{p-1}, \vec{f} - \vec{f}_1 - \cdots - \vec{f}_{p-1}) \cdot d\vec{f}_1 \cdots d\vec{f}_{p-1} \\ &= \int_{R^N} \cdots \int_{R^N} H_p(\vec{f}_1, \dots, \vec{f}_{p-1}, \vec{f} - \vec{f}_1 - \cdots - \vec{f}_{p-1}) \cdot \\ &\quad \cdot \tilde{U}_1(\vec{f}_1, \dots, \vec{f}_{p-1}, \vec{f} - \vec{f}_1 - \cdots - \vec{f}_{p-1}) \cdot d\vec{f}_1 \cdots d\vec{f}_{p-1}. \end{aligned} \quad (5.90)$$

Das bedeutet, daß sich das Ausgangsspektrum für jede Frequenz \vec{f} dadurch berechnet, daß sämtliche Werte des expandierten Spektrums in der Ebene

$$\{\vec{f}_1, \dots, \vec{f}_{p-1}, \vec{f} - \vec{f}_1 - \cdots - \vec{f}_{p-1}\}$$

aufintegriert werden. Daher wird diese Operation alternativ auch als Kontraktion bezeichnet. Dies ist in Abb. 5.10 für ein eindimensionales Volterra-System 3. Ordnung illustriert.

Während bei linearen Systemen die Übertragungsfunktion die Gewichte der spektralen Komponenten darstellt, so bestimmt die verallgemeinerte Übertragungsfunktion nichtlinearer Systeme die Gewichtung von bestimmten Frequenzkombinationen. Dies läßt sich

⁸Im Falle mehrdimensionaler Signale, d.h. $\vec{f} \in R^N$, $N > 1$ ist die Hauptdiagonale als höherdimensionale Mannigfaltigkeit zu verstehen, was auch für die bereits eingeführte örtliche Darstellung gilt.

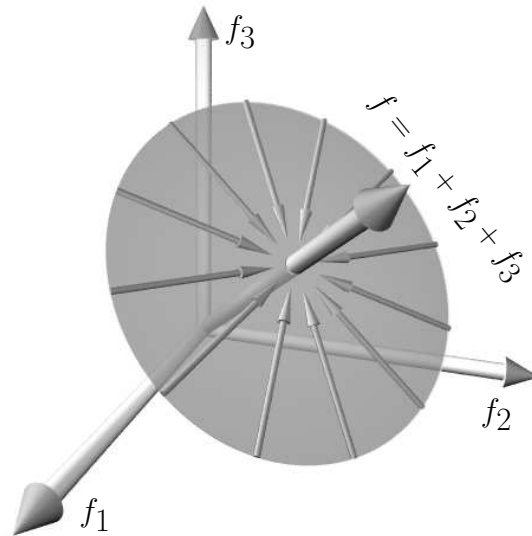


Abbildung 5.10: Darstellung der Kontraktion des expandierten Ausgangsspektrums eines eindimensionalen kubischen Volterra-Systems auf die Hauptdiagonale. Gezeigt ist der Integrationsbereich für einen bestimmten Wert von f ; dieser bestimmt die Lage der Ebene entlang des Vektors $f_1 + f_2 + f_3$.

zeigen, indem zuerst die Antwort eines homogenen (eindimensionalen) quadratischen Systems $H_2(f_1, f_2)$ auf eine einzelne harmonische Schwingung

$$\begin{aligned} u_1(t) &= A \cos(2\pi f_0 t + \lambda_0) \\ &= \frac{A}{2} \left[e^{j(2\pi f_0 t + \lambda_0)} + e^{-j(2\pi f_0 t + \lambda_0)} \right] \end{aligned} \quad (5.91)$$

berechnet wird. Wird die komplexe Darstellung in den quadratischen Teil von Gl. (5.65) eingesetzt und die Klammer ausmultipliziert, so erhält man:

$$\begin{aligned} u_2(t) = \frac{A^2}{4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_2(t_1, t_2) &\left(e^{j4\pi f_0 t} e^{-j2\pi f_0(t_1+t_2)} e^{j2\lambda_0} + \right. \\ &\left. + e^{j2\pi f_0(t_2-t_1)} + e^{j2\pi f_0(t_1-t_2)} + e^{-j4\pi f_0 t} e^{j2\pi f_0(t_1+t_2)} e^{-j2\lambda_0} \right) \cdot dt_1 dt_2. \end{aligned} \quad (5.92)$$

Durch Umordnen der Terme zu Fourier-Integralen erhält man sodann:

$$\begin{aligned} u_2(t) = \frac{A^2}{4} \left[e^{j4\pi f_0 t} e^{j2\lambda_0} H_2(f_0, f_0) + H_2(f_0, -f_0) + \right. \\ \left. + H_2(-f_0, f_0) + e^{-j4\pi f_0 t} e^{-j2\lambda_0} H_2(-f_0, -f_0) \right]. \end{aligned} \quad (5.93)$$

Unter der Voraussetzung von physikalisch realisierbaren Systemen, also Systemen mit einem rein reellen Kern h_2 gilt im Frequenzbereich

$$H_2(f_1, f_2) = H_2^*(-f_1, -f_2), \quad (5.94)$$

wobei * die konjugiert Komplexe symbolisiert. Schreibt man nun den Kern in Gl. (5.93) jeweils mit Betrag und Phase und verwendet die Identität für komplexe Zahlen

$$z + z^* = 2\Re\{z\}, \quad (5.95)$$

so erhält man schlußendlich als Ergebnis

$$u_2(t) = \frac{A^2}{2}|H_2(f_0, f_0)| \cos(4\pi f_0 t + 2\lambda_0 + \phi_0) + \frac{A^2}{2}\Re\{H_2(f_0, -f_0)\}, \quad (5.96)$$

wobei $\phi_0 = \arg\{H_2(f_0, f_0)\}$. Daran kann man erkennen, daß ein quadratisches System, bei Beaufschlagung mit einer harmonischen Schwingung, als Ausgang eine Schwingung mit doppelter Frequenz sowie einen Gleichanteil besitzt⁹. Wie jedoch an der Darstellung mit den komplexen Exponentialschwingungen ersichtlich ist, werden die Anteile der Antwort mit der doppelten Frequenz, entstehend aus der Interaktion der Schwingungen gleicher Frequenz, gewichtet mit dem Kern an der Stelle (f_0, f_0) . Der Gleichanteil geht hervor aus der Summation der Frequenzanteile mit unterschiedlichem Vorzeichen, gewichtet mit dem Kern an der Stelle $(f_0, -f_0)$. Allgemein läßt sich zeigen, daß sich die Antwort eines quadratischen Volterra-Systems für ein aus m komplexen Exponentialschwingungen zusammengesetztem Eingangssignal

$$u_1(t) = \sum_{k=1}^m A_k \cdot e^{j2\pi f_k t}, \quad (5.97)$$

mit Amplituden A_k und Frequenzen f_k , schreiben läßt als

$$u_2(t) = \sum_{k=1}^m \sum_{l=1}^m A_k A_l \cdot H_2(f_k, f_l) \cdot e^{j2\pi(f_k+f_l)t}. \quad (5.98)$$

Da sich nach dem Fourier-Theorem nahezu jede beliebige Funktion als eine Summe komplexer Exponentialschwingungen schreiben läßt, wenn nur die Frequenzunterschiede hinreichend klein und die Gesamtzahl hinreichend groß gewählt werden, kann dies auch leicht für Signale mit einem kontinuierlichen Spektrum verallgemeinert werden. Das bedeutet also, daß ein quadratisches Volterra-System jedes Eingangssignal in seine Fourier-Komponenten zerlegt, alle möglichen Produkte der Frequenzen berechnet und diese mit den entsprechenden Kerneinträgen gewichtet. Diese Art der *UND-Verknüpfung* von Frequenzkomponenten stellt den fundamentalen Unterschied zu rein linearen Systemen im Frequenzbereich dar; diese sind eher mit einer *ODER-Verknüpfung* der Frequenzkomponenten zu beschreiben, da hierbei jede Frequenz unabhängig gewichtet wird, vgl. Abb. 5.11.

⁹Dies ist prinzipiell trivial, da bereits die elementaren trigonometrischen Gleichungen dasselbe Ergebnis liefern.

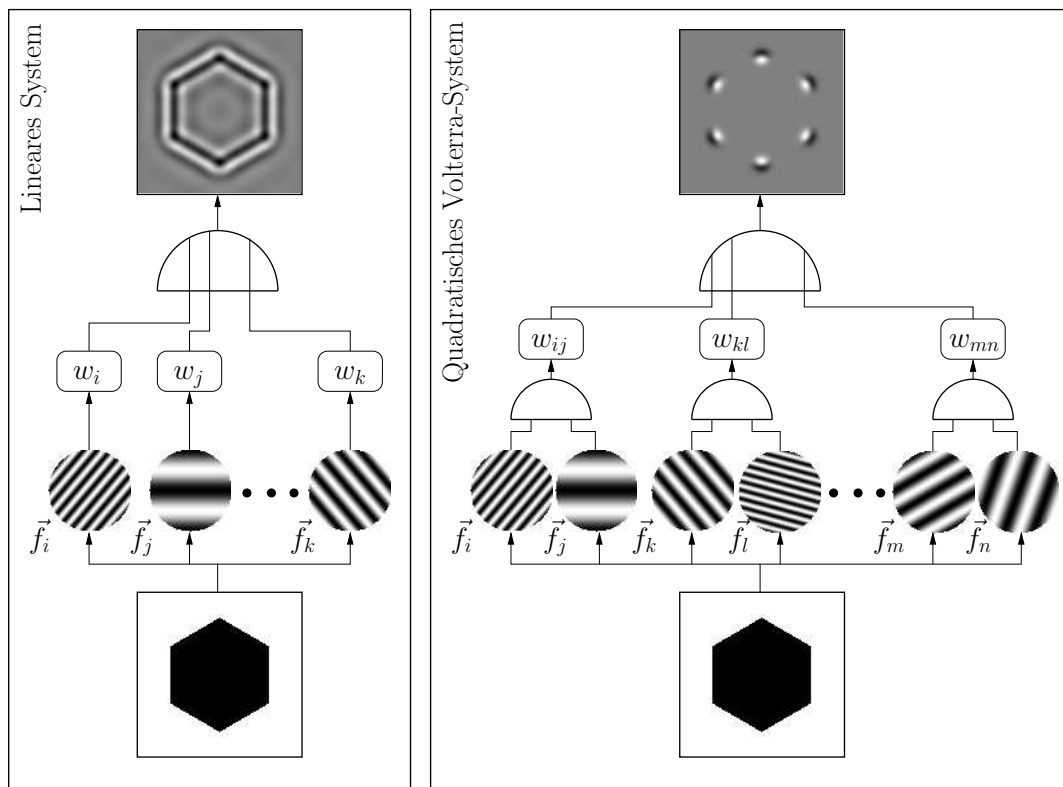


Abbildung 5.11: Ein lineares System gewichtet die Frequenzkomponenten des Eingangs unabhängig und kann somit als *ODER-Verknüpfung* betrachtet werden. Ein quadratischer Volterra-Operator gewichtet die Produkte der Frequenzkomponenten und führt somit eine *UND-ähnliche Verknüpfung* durch.

5.2.3 Synthese und Approximation nichtlinearer neuronaler Systeme

Bei der in dieser Arbeit durchgeführten Untersuchung der nichtlinearen Informationsverarbeitung im visuellen System wird nun der im letzten Abschnitt erläuterte Formalismus für nichtlineare Systeme zur Modellierung kortikaler Phänomene verwendet. Dabei gibt es in der Sehrinde von Säugetieren wahrlich einen ganzen „Zoo“ unterschiedlicher Neurone, deren Antwortverhalten bis heute eher im Dunkeln liegt. Im Folgenden werden vier unterschiedlicher Vorgehensweisen im Detail eruiert. *(i)* Synthese generischer Prinzipien durch konstruktives Vorgehen, d.h. es werden grundlegende nichtlineare Eigenschaften systemtheoretisch *konstruiert* und die Eigenschaften in der Struktur des Volterra-Kerns identifiziert (s.u.). *(ii)* Diese Prinzipien können verwendet werden, um ein bestimmtes Antwortverhalten kortikaler Neurone zu modellieren, Abschnitt 5.2.4. *(iii)* Im Allgemeinen Fall besteht die Möglichkeit, die Neuronenaktivität durch das klassische Verfahren der *Systemidentifikation* auf nichtlineare Systeme abzubilden, dies wird im Detail in Abschnitt 5.2.5 besprochen. *(iv)* Ist man in Kenntnis der modellierenden Kerne, so stellt sich die Frage, wie diese durch einfache Strukturen implementiert sein können. Im Zentrum des Interesses steht hierbei die Suche nach physiologisch plausiblen „Architekturen“.

Wie bereits in Abschnitt 2.7 dargelegt, ist die Modellierung kortikaler Phänomene mit rein linearen Systemen ungenügend. Das zentrale Konzept, das hier bei der Modellierung verwendet wird, ist die UND-ähnliche Verschaltung von Frequenzkomponenten, eine Eigenschaft, die erst bei Volterra-Systemen mit Ordnung ≥ 2 auftritt. Unglücklicherweise sind die Kerne von zweidimensionalen Volterra-Systemen höherer Ordnung sehr unhandlich was Darstellung und Berechnung betrifft. Daher werden im Folgenden als Kompromiß lediglich Volterra-Systeme der Ordnung 2 betrachtet, deren vierdimensionale Kerne noch relativ übersichtlich entsprechend dem Koordinatensystem für Bispektren (Abb. 5.3) darstellbar sind, und für die effiziente und gut verstandene Implementierungen existieren.

Synthese generischer Prinzipien kortikaler Neurone

Es werden nun einige grundlegende Effekte, die in neurophysiologischen Untersuchungen gefunden werden, aufgeführt und dargelegt, wie diese mit einem quadratischen Volterra-System synthetisierbar sind. Diese sind

1. Suppression isolierter Frequenzkomponenten und bilineares Antwortverhalten,
2. Demodulation, bzw. Phaseninsensitivität,
3. Endinhibition, *i2D*-Selektivität,
4. Extra-Klassische Effekte.

Ad 1. Der Effekt, daß einzelne harmonische Schwingungen keine Antwort eines Systems hervorrufen führt bei einer linearen Annahme zu der trivialen Übertragungsfunktion $H(f) \equiv 0$. Dies liegt darin begründet, daß harmonische Schwingungen die Eigenfunktionen von linearen Systemen darstellen, siehe auch Abb. 5.11. Aus Sicht der quadratischen Volterra-Systeme ist dies jedoch eine grundlegende Eigenschaft, da dies genau der UND-Verknüpfung zweier Frequenzkomponenten entspricht. Realisiert werden kann dies trivialerweise durch eine Multiplikation zweier linearer Systeme mit nichtüberlappender Übertragungsfunktion. In Abb. 5.12 ist ein derartiges System gezeigt, wobei die Übertragungsfunktionen schmalbandige Bandpässe sind. Wird das System nun mit einer einzelnen Frequenz beaufschlagt, so ist immer ein Multiplikand Null, was dazu führt, daß das System auf einzelne Sinusschwingungen keine Antwort liefert. Sind jedoch zwei (oder mehr) Frequenzkomponenten enthalten, von denen mindestens eine im Durchlaßbereich je eines Filters liegt, so wird ein von Null verschiedener Ausgang erzeugt. Trotz der Trivialität dieses Beispiels zeigt sich hier die Wirkung der fundamentalen UND-Verknüpfung von Frequenzkomponenten. Sämtliche folgende Ansätze zur Modellierung können auf dieses Prinzip reduziert werden, lediglich die *Arten* der kombinierten Frequenzen sind unterschiedlich, wie z.B. Orientierung oder Phase. Ein weiterer Aspekt, der hiermit direkt zusammenhängt, ist das sogenannte bilineare Antwortverhalten. Dies bedeutet lediglich eine Verletzung des Überlagerungsprinzips, so daß die Antwort auf die Summe zweier Stimuli ungleich der Summe der Einzelantworten ist. Im gezeigten Beispiel ist dies in extremer Form vorhanden, im Allgemeinen sind die Unterschiede jedoch subtiler. Die Differenz zwischen der Antwort auf die Überlagerung und die Summe der Einzelantworten wird als bilinearer Anteil bezeichnet, ein Effekt, der auch in sensorischen Neuronen gefunden wird (siehe (Krieger, 1999) und die darin enthaltenen Referenzen).

Ad 2. Bereits im primären visuellen Kortex finden sich Zellen, die durch eine *Phaseninvarianz* gekennzeichnet sind (Kato et al., 1978; Skottun et al., 1991), und damit auch gerne als positionsinvariant bezeichnet werden. Dies entspricht im nachrichtentechnischen Sinn einer klassischen Demodulation und kann durch einen Quadrierer mit nachfolgendem Tiefpaß realisiert und somit als homogenes quadratisches Volterra-System dargestellt werden. Ebenfalls möglich ist die Verwendung des Konzepts des analytischen Signals, das die Darstellung als lokalen Betrag und Phase erlaubt (vgl. Kap. 4). Der lokale Betrag, oder auch die lokale Energie ergibt sich zu

$$u_2(\vec{x}) = \left(u_1(\vec{x}) \star h_1^{(even)}(\vec{x}) \right)^2 + \left(u_1(\vec{x}) \star h_1^{(odd)}(\vec{x}) \right)^2, \quad (5.99)$$

wobei $h_1^{(even)}(\vec{x})$ und $h_1^{(odd)}(\vec{x})$ die Impulsantworten von zwei gerade-, respektive ungeradesymmetrischen linearen Filtern bezeichnen, deren Übertragungsfunktionen betragsmäßig identisch sind: $|H_1^{(even)}(\vec{f})| = |H_1^{(odd)}(\vec{f})|$. Im Frequenzbereich schreibt sich der quadratische Volterra-Kern als (vgl. (Zetzsche und Krieger, 2001a)):

$$H_2(\vec{f}_1, \vec{f}_2) = H_1^{(even)}(\vec{f}_1) \cdot H_1^{(even)}(\vec{f}_2) + H_1^{(odd)}(\vec{f}_1) \cdot H_1^{(odd)}(\vec{f}_2). \quad (5.100)$$

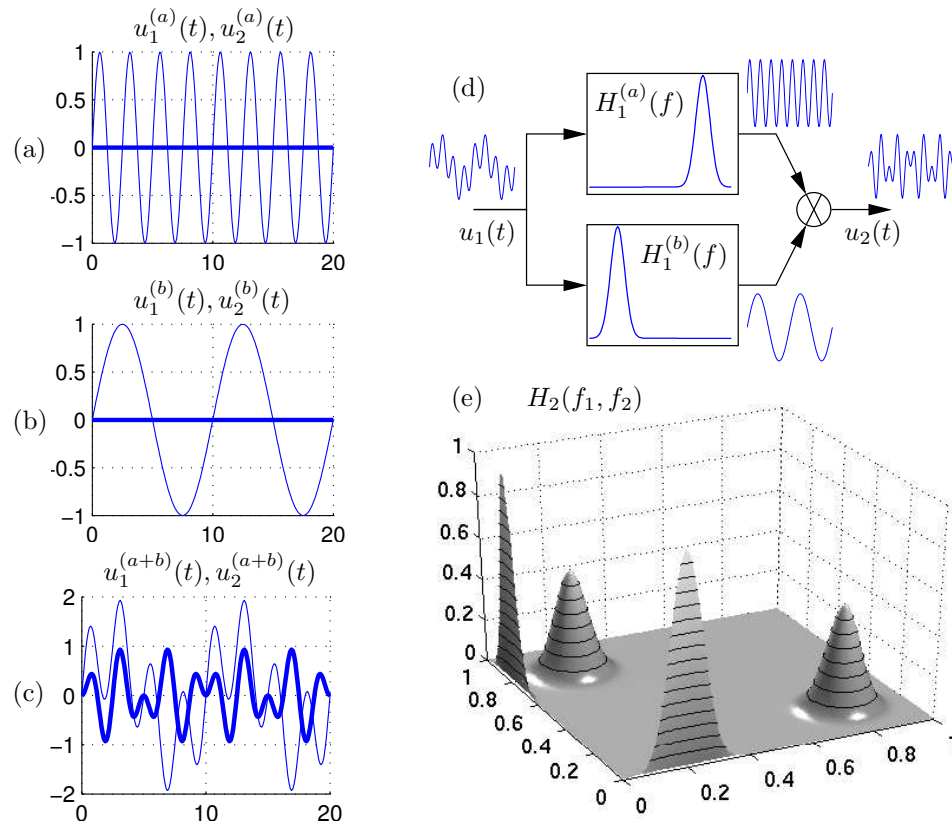


Abbildung 5.12: Suppression einzelner Frequenzen durch ein quadratisches Volterra-System. (a) Einzelner Sinus mit optimaler Frequenz bezüglich des linearen Filters $H_1^{(a)}$ sowie Antwort des Gesamtsystems (fett), (b) dito für Filter $H_1^{(b)}$. (c) Summe der einzelnen Harmonischen sowie Ausgangssignal (fett). (d) Realisierung des Gesamtsystems als Produkt der linearen Filter $H_1^{(a)}$ und $H_1^{(b)}$. Eingezeichnet sind zudem die Signalanteile für das zusammengesetzte Eingangssignal, die das System passieren. (e) Symmetrischer Kern zweiter Ordnung des Gesamtsystems aus (d).

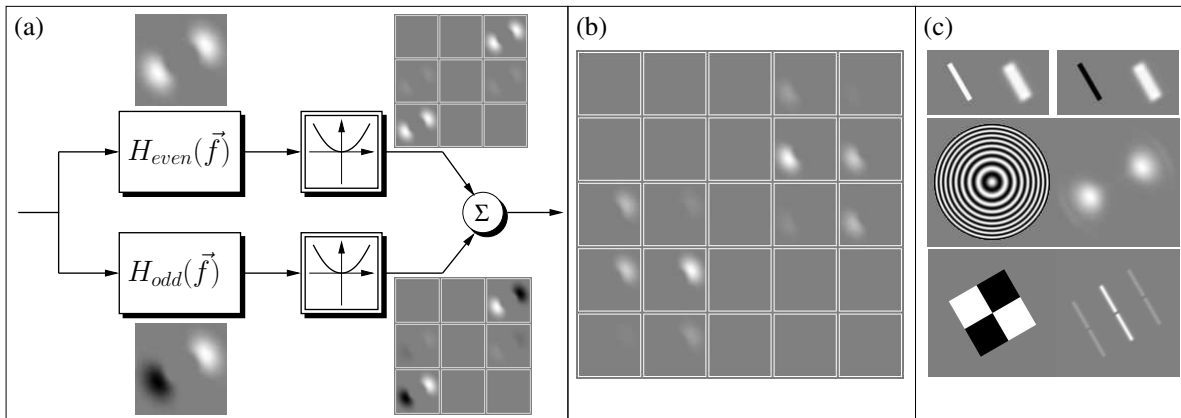


Abbildung 5.13: Synthese der Phaseninvarianz kortikaler Neurone durch einen Demodulierer. (a) Blockschaltbild des Systems, bestehend aus der Summe der quadrierten Filterausgänge zweier linearer Filter $H_{even}(\vec{f})$ und $H_{odd}(\vec{f})$. Dazu sind die Teilkerns der linearen sowie der quadratischen Anteile gezeigt. Notabene: $H_{even}(\vec{f})$ ist rein reell und $H_{odd}(\vec{f})$ rein imaginär. (b) Quadratischer Kerns des Gesamtsystems. (c) Typisches Antwortverhalten.

Eine Implementierung des Systems ist in Abb. 5.13 (a) gezeigt. Beim Passieren eines quadratischen Systems wird das Spektrum eines Signals dahingehend verändert, daß von sämtlichen Frequenzkomponenten die Summen und Differenzen gebildet werden, siehe Gl. (5.98). Anhand des quadratischen Volterra-Kerns (Abb. 5.13 (b)) ist ersichtlich, daß lediglich Paarungen mit unterschiedlichem Vorzeichen gewichtet werden, womit die hochfrequenten Anteile eliminiert und nur die niederfrequenten Anteile, bzw. der Gleichanteil, erhalten bleiben. Eine weitere Eigenschaft, die anhand der Struktur des Kerns identifiziert werden kann, ist die der Orientierungsselektivität. Diese ergibt sich durch die bereits orientierungsselektive lineare Filterung. Typische Antworten auf Testmuster, wie sie auch gerne in neurophysiologischen Untersuchungen verwendet werden, sind in Abb. 5.13 (c) gezeigt.

Ad 3. Wie bereits in Abschnitt 2.7 erwähnt, handelt es sich bei der Endinhibition um ein bislang unterschätztes Phänomen, da sich zunehmend zeigt, daß der Großteil der Neurone im visuellen Kortex zu einem gewissen Grad endinhibiert sind (Pack et al., 2003b). Endinhibition ist dadurch charakterisiert, daß Neurone nur auf kurze Stimuli, bzw. auf Linienenden oder Ecken, antworten. Die Funktion dieses Mechanismus ist nicht völlig klar, eine Erklärung ist die Überwindung des Aperturproblems, das sich ergibt, wenn sich ein ausgedehnter Stimulus durch ein rezeptives Feld bewegt und dabei der Geschwindigkeitsvektor geschätzt werden soll (Zetzsche et al., 1993a; Pack et al., 2003b). Unter Verwendung des Konzepts der intrinsischen Dimensionalität entsprechen die endinhibierten Zellen den *2D-Detektoren*, die sensitiv auf Krümmungen reagieren und somit zur Redundanzreduktion im Sinne von Attneave beitragen. Eine einfache Implementierung ist durch das Produkt zweier nicht überlappender orientierungsselektiver Filter gegeben. Dieses System kann nur von einem Eingangssignal passiert werden das beide

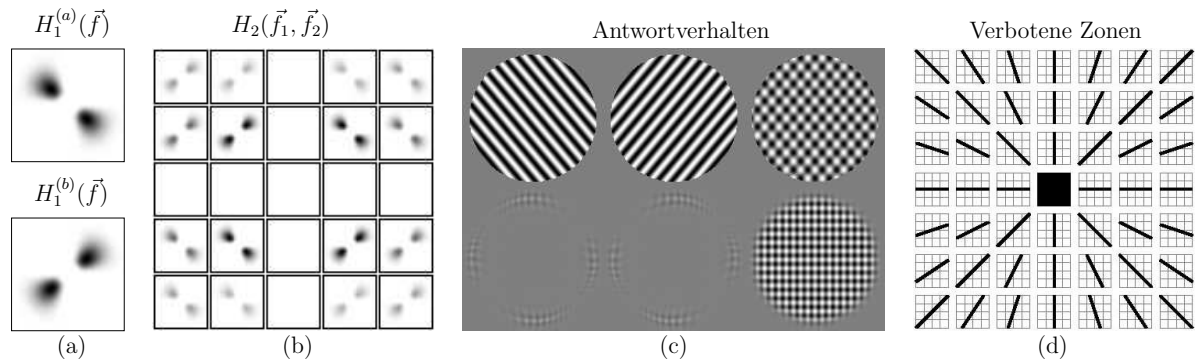


Abbildung 5.14: Prototypisches $i2D$ -System, bestehend aus der multiplikativen Verknüpfung zweier linearer orientierungsselektiver Bandpaß-Filter $H_1^{(a)}(\vec{f}), H_1^{(b)}(\vec{f})$ (a), mit quadratischem Volterra-Kern (b). Dieses System antwortet nicht auf einzelne Schwingungen im Durchlaßbereich nur eines Filters, da der jeweils andere Multiplikand Null ist (c). Allgemein ist ein $i2D$ -selektives System durch die notwendige und hinreichende Bedingung in Gl. (5.101) charakterisiert (d).

Orientierungen beinhaltet, also vom Typus $i2D$ ist, siehe Abb. 5.14. Im Allgemeinen kann die $i2D$ -Selektivität anhand des quadratischen Kerns definiert werden. Die notwendige und hinreichende Bedingung ist hierfür, daß der Kern an sämtlichen Stellen \vec{f}_1 und \vec{f}_2 mit identischer Orientierung Null sein muß, also (siehe beispielsweise auch (Krieger und Zetzsche, 1996))

$$H_2(f_{x_1}, f_{y_1}, f_{x_2}, f_{y_2}) \equiv 0 \quad \forall \quad \frac{f_{y_1}}{f_{x_1}} = \frac{f_{y_2}}{f_{x_2}}. \quad (5.101)$$

Diese „verbotenen Zonen“ des quadratischen Volterra-Kerns sind in Abb. 5.14 (d) gezeigt. Es ist klar, daß das einfache Beispielsystem diese Bedingung erfüllt und somit $i2D$ -selektiv ist. Im nächsten Abschnitt werden Design-Prinzipien für die Konstruktion komplexerer Systeme behandelt.

Ad 4. Ein ebenfalls zunehmend beobachtetes Phänomen im visuellen Kortex ist das Auftreten sogenannter *extra-klassischer rezeptiver Felder* (ECRF). Lax formuliert bedeutet dies, daß rezeptive Regionen existieren, die für sich alleine betrachtet keine Antwort evozieren können, also nicht zum klassischen rezeptiven Feld (CRF) gehören. Bei Stimulation des CRF kann jedoch die Antwort durch einen passenden suppressiven Stimulus im ECRF inhibiert werden (Sillito et al., 1995; Walker et al., 2000). Diese extra-klassischen rezeptiven Felder sind eine genuine Eigenschaft nichtlinearer, im Besonderen UND-verknüpfender Systeme. Neurophysiologische Untersuchungen verwenden häufig ausgedehnte Lichtpunkte oder Balken, um das klassische rezeptive Feld des Neurons abzubilden. Bei einer linearen Hypothese approximiert dieses Verfahren dann die charakteristische Impulsantwort. Nun kann ein länglicher Stimulus innerhalb des rezeptiven Feldes verwendet werden, um die Orientierungsabstimmung zu messen. Betrachtet man beispielsweise das $i2D$ -selektive System zweiter Ordnung, dessen Kern in Abb. 5.15 (a) dargestellt ist, so erhält man die ebenfalls gezeigte Impulsantwort. Bei Verwendung ty-

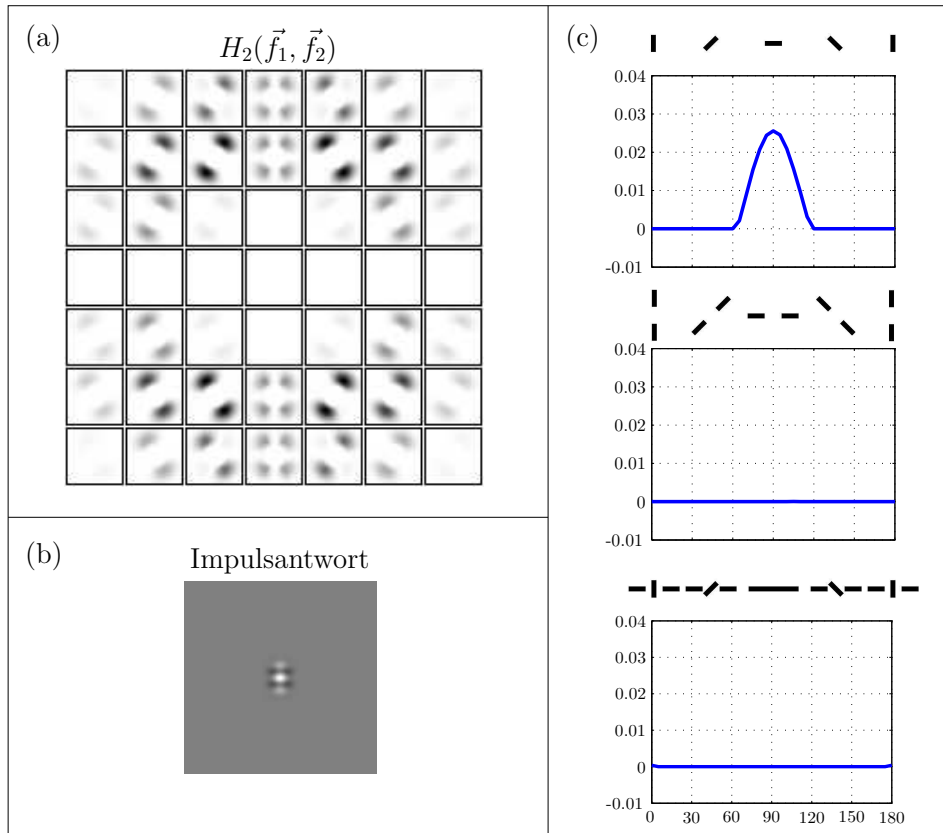


Abbildung 5.15: Modell zur Synthese extra-klassischer Effekte. (a) Quadratischer Kern. (b) Nichtlineare Impulsantwort des Systems. (c) Gemessene Abstimmungskurven für unterschiedlich orientierte Stimuli innerhalb (oben) und außerhalb (mittig) des rezeptiven Feldes. Bei Addition eines isoliert unwirksamen äußeren Stimulus und erneutem Messen der Orientierungsabstimmung wird diese vollständig unterdrückt.

pischer experimenteller Stimuli kann dann entsprechend die Orientierungsabstimmung gemessen werden. Ein entsprechender Stimulus außerhalb des CRF kann per Definition keine Antwort verursachen; simulierte Messungen zeigt Abb. 5.15 (c). Überraschenderweise jedoch kann ein äußerer Stimulus mit der bevorzugten Orientierung des rezeptiven Feldes die Antwort auf den inneren völlig unterdrücken. Dieser des öfteren berichtete Effekt ist jedoch nur dann überraschend, wenn implizit lineare Annahmen gemacht werden und das Vorhandensein von UND-artigen Kombinationen ignoriert wird. Mit quadratischen Systemen hingegen ist dieses physiologische Phänomen leicht und direkt modellierbar, während das Fehlen eines entsprechenden Konzepts zur Verwirrung über die Funktion des ECRF führt:

The nature and role of CRF surrounds is an enigma.

(Walker et al., 2000)

Effiziente Implementierung und Approximation quadratischer Volterra-Systeme

Weiter oben wurde beschrieben, wie generische Eigenschaften kortikaler Neurone durch a priori Wissen über das Systemverhalten *konstruktiv* modelliert werden können. Diese recht qualitativen Überlegungen verlangen konsequenterweise nach einer Möglichkeit, funktionale Eigenschaften durch gezieltes Variieren der Filterparameter abzustimmen, völlig analog dem Frequenz-tuning linearer Übertragungsfunktionen. So wurde beispielsweise in (Krieger, 1999) ausführlich dargelegt, wie die Eigenschaft der *i2D*-Selektivität von quadratischen Volterra-Systemen gezielt im Frequenzbereich abgestimmt werden kann. Dies ist im allgemeinen nichttrivial und zudem mit einem enormen Rechenaufwand verbunden, da die Antwort direkt über den vierdimensionalen Kern berechnet werden muß¹⁰. Wesentlich effizienter und anschaulicher gestaltet sich die Abstimmung der Systeme, wenn kanonische Strukturen zur Realisierung verwendet werden. Bereits vorgestellt wurden die Kaskadenstrukturen der *Wiener*- und *Hammerstein*-Modelle, die einfach zu verstehen und abzustimmen sind, deren nichtlineare Verarbeitungsleistung aber auch stark eingeschränkt ist. Wie man leicht zeigen kann ergibt sich der quadratische Kern dieser Systeme bei Verwendung eines linearen Filters $H_1(\vec{f})$ zu:

$$H_2(\vec{f}_1, \vec{f}_2) = H_1(\vec{f}_1) \cdot H_1(\vec{f}_2) \quad (\text{Wiener}) \quad (5.102)$$

$$H_2(\vec{f}_1, \vec{f}_2) = H_1(\vec{f}_1 + \vec{f}_2) \quad (\text{Hammerstein}) \quad (5.103)$$

Eine weitaus mächtigere Struktur sind Produkte von Systemen, wie bereits im Beispiel zu Abb. 5.12 verwendet. Die Einschränkung dieser Systeme liegt in der Separierbarkeit des Kerns, der sich zu

$$H_2(\vec{f}_1, \vec{f}_2) = H_1^{(a)}(\vec{f}_1) \cdot H_1^{(b)}(\vec{f}_2) \quad (5.104)$$

ergibt, wobei $H_1^{(a)}(\vec{f})$, bzw. $H_1^{(b)}(\vec{f})$ die Kerne der linearen Subsysteme sind. Dies wiederum bedeutet, daß eine isotrope *i2D*-selektive Filterung damit nicht möglich ist. Allerdings kann das Verhalten dieser Systeme noch recht einfach bestimmt werden, womit bereits eine Vielzahl nichtlinearer Effekte darstellbar sind. Die Selektivität dieser Struktur kann weiter erhöht werden, wenn ein linearer Filter (Postfilter) $H_1^{(P)}(\vec{f})$ nachgeschaltet wird. Die verallgemeinerte Übertragungsfunktion hierbei ergibt sich dann zu

$$H_2(\vec{f}_1, \vec{f}_2) = H_1^{(a)}(\vec{f}_1) \cdot H_1^{(b)}(\vec{f}_2) \cdot H_1^{(P)}(\vec{f}_1 + \vec{f}_2). \quad (5.105)$$

Mit dieser relativ allgemeinen Kaskadenstruktur kann bereits eine große Klasse quadratischer Systeme abgedeckt werden; detaillierte Untersuchungen zur Selektivität verschiedener Strukturen finden sich in (Zetzsche et al., 2001; Nuding, 2002).

Komplexere Strukturen, die mit einfachen kanonischen Modellen nicht erreichbar sind, bedürfen mehrerer Subsysteme, deren Teilkern addiert werden. Damit kann beispielsweise eine Architektur, bestehend aus (beliebig) vielen parallelen quadrierten linearen Filtern jeden möglichen Kern approximieren¹¹, wie man sich anhand der Illustration in

¹⁰Also Expansion, Produkt und Kontraktion.

¹¹Von pathologischen Sonderfällen einmal abgesehen.

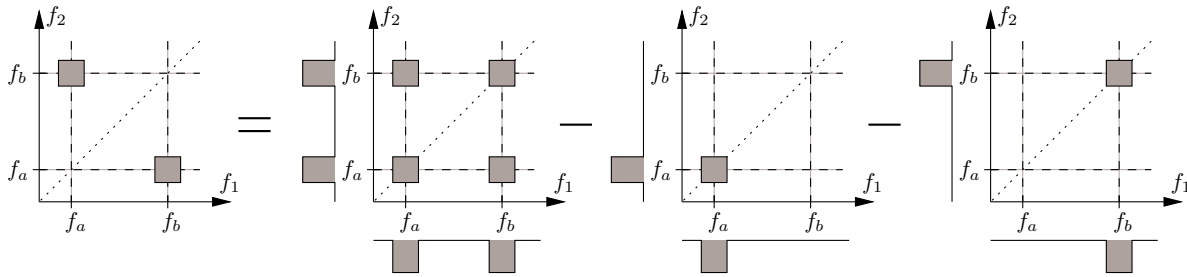


Abbildung 5.16: Realisierung beliebiger nicht separierbarer Volterra-Kerne durch die Summe quadrierter linearer Filterausgänge. Anhand dieser Skizze ist ersichtlich, daß sich jeder quadratische Kern prinzipiell approximieren läßt, obwohl die konkrete Realisierung mit schmalbandigen Filtern nicht effizient ist.

Abb. 5.16 verdeutlichen kann. Formal kann dies leicht über fundamentale Sätze der linearen Algebra bewiesen werden. So ist beispielsweise bekannt, daß jede symmetrische Matrix $H_2 \in R^{N \times N}$ mit Rang r in eine endliche Summe von r Rang-1 Matrizen zerlegt werden kann (Gantmacher, 1960)¹²:

$$H_2 = \sum_{i=1}^r q_i \vec{v}_i \vec{v}_i^T, \tag{5.106}$$

wobei die q_i skalare Konstanten und die \vec{v}_i N -dimensionale Vektoren darstellen. Wie in Abschnitt 5.2.5 gezeigt wird, kann jedes diskrete homogene quadratische Volterra-System leicht in Vektor-Matrix-Notation angegeben werden, wobei die Kernkoeffizienten in der Matrix H_2 angeordnet werden. Damit können also entsprechend Gl. (5.106) quadratische Volterra-Kerne durch die äußeren Produkte von Vektoren dargestellt werden, was funktional einer Quadrierung linearer Filter entspricht. Zur Realisierung dieser Matrix-Zerlegung existieren mehrere Möglichkeiten mit unterschiedlichen Vor- und Nachteilen. Eine attraktive Lösung bietet die Singulärwertzerlegung (*Singular Value Decomposition, SVD*), welche approximative Lösungen findet, die aber häufig mit einer stark reduzierten Komplexität bei nahezu identischer Leistung einhergehen (Gantmacher, 1960; Mathews und Sicuranza, 2000). Hierbei wird die Matrix $H_2 \in R^{N \times N}$ in das Produkt einer orthogonalen Matrix U , einer Diagonalmatrix¹³ W und der Transponierten einer ebenfalls orthogonalen Matrix V zerlegt:

$$H_2 = U \cdot W \cdot V^T. \tag{5.107}$$

¹²Prinzipiell kann jede Rang- r Matrix mittels Eigen-, bzw. Singulärwertzerlegung diagonalisiert werden; da sich die betrachtete Klasse jedoch aus den Koeffizienten der Volterra-Kerne konstituiert, sind die Matrizen immer symmetrisch, siehe Gl. (5.70).

¹³Die Einträge in dieser Matrix sind ≥ 0 und werden als *Singulärwerte* bezeichnet.

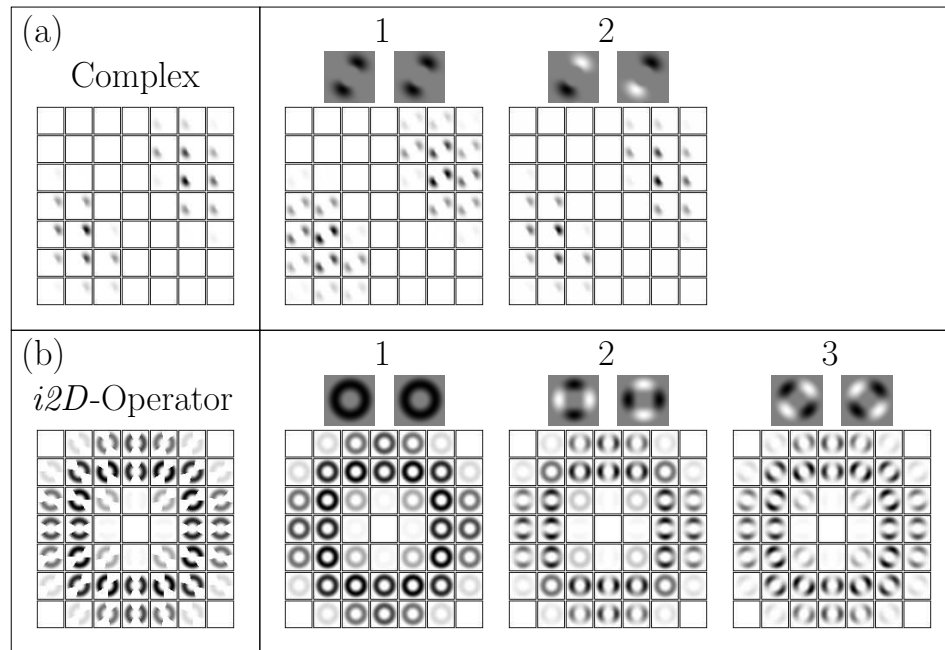


Abbildung 5.17: Approximation quadratischer Volterra-Systeme mittels Singulärwertzerlegung. Gezeigt ist jeweils links der korrekte und rechts die sukzessive approximierten Kerne, sowie die jeweiligen linearen Teilfilter \vec{u}_i, \vec{v}_i . (a) Das bereits in Abb. 5.13 gezeigte System zur Demodulierung wird (bis auf die ambigen Vorzeichen) korrekt zerlegt. (b) Bei einem „harten“, nicht separierbaren $i2D$ -Operator werden die Fehler sichtbar.

Dies läßt sich auch schreiben als

$$\begin{aligned}
 H_2 &= \left(\vec{u}_1, \vec{u}_2, \dots, \vec{u}_N \right) \cdot \begin{pmatrix} w_1 & & & \\ & w_2 & & \\ & & \ddots & \\ & & & w_N \end{pmatrix} \cdot \begin{pmatrix} \vec{v}_1^T \\ \vec{v}_2^T \\ \vdots \\ \vec{v}_N^T \end{pmatrix} \\
 &= \sum_{i=1}^N w_i \vec{u}_i \vec{v}_i^T.
 \end{aligned} \tag{5.108}$$

Funktional entspricht dies der Parallelschaltung von N linear separierbaren quadratischen Volterra-Systemen, die sich aus dem Produkt zweier linearer Filter mit den Übertragungsfunktionen \vec{u}_i , bzw. \vec{v}_i ergeben. Es kann gezeigt werden, daß bei der SVD, analog zur Rekonstruktion bei der KLT, die bestmögliche Approximation für eine Repräsentation mit Rang $k \leq N$ gefunden wird (Mathews und Sicuranza, 2000). Abb. 5.17 zeigt zwei Beispiele für die Approximation mittels der SVD. Die besondere Attraktivität dieser Zerlegung liegt nicht nur in der effizienten Approximation und der Möglichkeit zur Reduzierung der Komplexität der Systeme, sondern auch darin, daß sie sich hervorragend zur Implementierung in VLSI Schaltungen eignet (Mathews und Sicuranza, 2000).

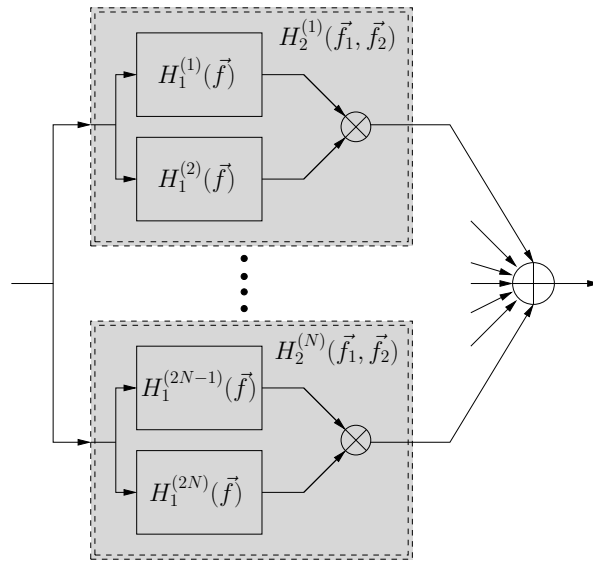


Abbildung 5.18: Schema der quadratischen $\Sigma\Pi$ Architektur.

Die Systemstruktur, die sich aus der SVD ergibt, wird verallgemeinert auch als $\Sigma\Pi$ -Architektur bezeichnet. Für quadratische Systeme ergibt sich diese eben aus der Summe von N Volterra-Operatoren $H_2^{(i)}(\vec{f}_1, \vec{f}_2)$, $i \in 1..N$, die ihrerseits linear zerlegbar sind (siehe Abb. 5.18), d.h.

$$\begin{aligned}
 H_2(\vec{f}_1, \vec{f}_2) &= \sum_{i=1}^N H_2^{(i)}(\vec{f}_1, \vec{f}_2) \\
 &= \sum_{i=1}^N H_1^{(2i-1)}(\vec{f}_1) H_1^{(2i)}(\vec{f}_2).
 \end{aligned} \tag{5.109}$$

Dies kann auch leicht für Systeme höherer Ordnung verallgemeinert werden. Trotz der augenscheinlichen Komplexität dieser Architektur können auch hierfür dezidierte Eigenschaften kontrolliert werden. In (Zetsche und Barth, 1990a; Zetsche und Barth, 1990b) wird die Kompensationsgleichung hergeleitet, eine notwendige und hinreichende Bedingung zum Erreichen der $i2D$ -Selektivität. Diese besagt, daß die Summe über die Produkte für eine bestimmte Orientierung identisch Null sein muß, um unerwünschte UND-Kombinationen für Frequenzen mit gleicher Orientierung zu unterdrücken, siehe auch (Zetsche und Krieger, 2001b). Dies ist vor allem für den Aufbau eines parametrierbaren $i2D$ -selektiven Systems interessant. Hiervon wird in Abschnitt 5.2.6 Gebrauch gemacht, um ein entsprechendes System an die Statistik natürlicher Bilder anzupassen.

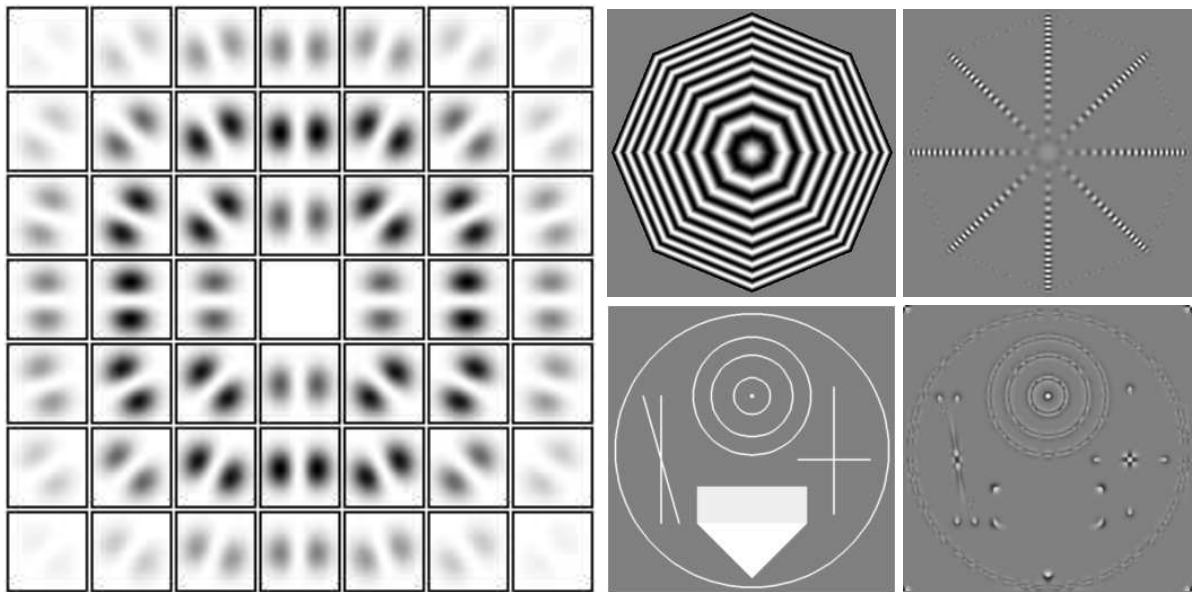


Abbildung 5.19: Modell einer Dot-Responsive-Zelle. Anhand des Kerns (links) erkennt man sowohl die $i2D$ -Selektivität, als auch die Isotropie. Die Antworten auf die Testbilder (rechts) lassen zudem auf die Abstimmung schließen, da lediglich Signalanteile mit starker Krümmung passieren können.

5.2.4 Modellierung nichtlinearer kortikaler Phänomene

Im letzten Abschnitt wurde gezeigt, wie prinzipiell nichtlineare neuronale Eigenschaften modelliert und effizient implementiert werden können. Hier soll nun anhand einiger exemplarischer Beispiele verdeutlicht werden, daß damit in der Tat eine Vielzahl bekannter kortikaler Phänomene erfaß- und erklärbar sind. Im Besonderen werden hier bestimmte, v.a. $i2D$ -selektive Neuronenklassen modelliert und diese bezüglich ihrer Merkmalsselektivität analysiert. Letzteres ist im Lichte einer übervollständigen und spärlichen Repräsentation entsprechend Abschnitt 2.5.3 von Relevanz. Als prototypische Neurone für die Verarbeitung von Form im visuellen Kortex sollen dabei fungieren: (i) Complex-Zellen, charakterisiert durch ihre Phaseninvarianz und Orientierungsselektivität, (ii) Dot-Responsive-Zellen, diese entsprechen $i2D$ -Operatoren mit breitem Sperrbereich, also nur erregbar von starken Krümmungen (Punkte, Ecken und Kreise mit kleinem Durchmesser) und (iii) Hypercomplex-Zellen, welche phaseninsensitiv und endinhibiert sind und damit am Besten auf kurze Linienstücke antworten. Für physiologische Befunde zu diesen Zelltypen siehe beispielsweise (Kato et al., 1978; Saito et al., 1988).

Das Modell der Complex-Zelle wurde bereits in Abschnitt 5.2.3, bzw. in Abb. 5.13 gezeigt. Die vierdimensionalen Kerne sowie das typische Antwortverhalten der Modelle der beiden $i2D$ -selektiven Zelltypen zeigen die Abbildungen 5.19 und 5.20.

Im Folgenden wird die Merkmalsselektivität der Modellneurone untersucht. Sowohl theo-

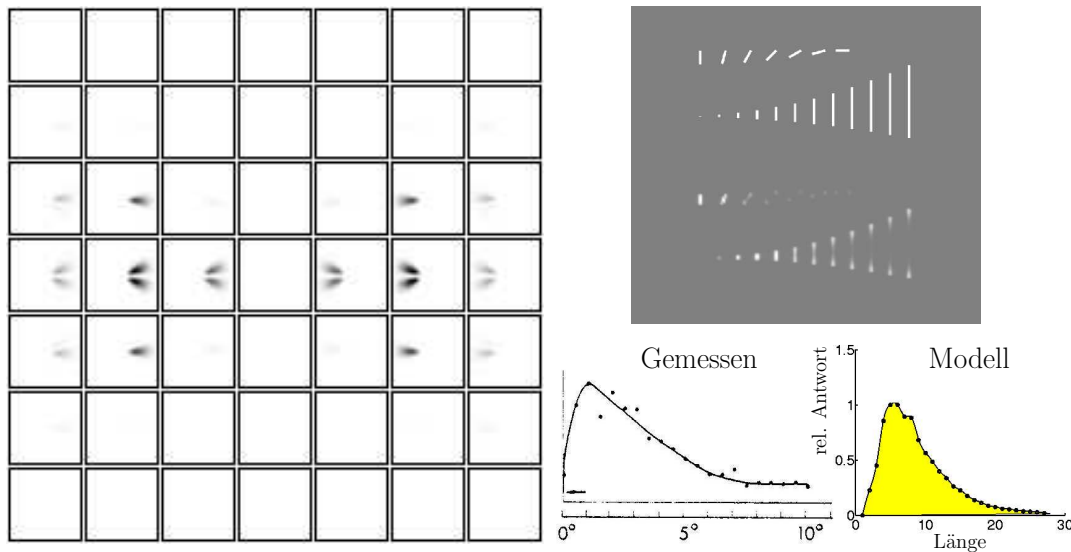


Abbildung 5.20: Modell einer Hypercomplex-Zelle. Der Kern weist typische Eigenschaften einer Complex-Zelle auf, im Besonderen Orientierungsselektivität und Demodulation. Zudem ist er durch einen Kompensationsmechanismus auf den verbotenen Zonen inhibiert, was zu einer *i2D*-Selektivität führt. Dieser Operator antwortet am Besten auf kurze Linienstücke; die gemessene Längenabstimmung stimmt qualitativ mit physiologischen Befunden überein. Die empirischen Daten stammen aus (Kato et al., 1978).

retische als auch neurophysiologische Befunde legen nahe, daß visuelle Neurone hochgradig auf die Verarbeitung spezifischer Bildmerkmale ausgerichtet sind, was in einer spärlichen und übervollständigen Repräsentation resultiert (dies wurde bereits ausführlich in Kap. 2 dargelegt). Zellen in einer derartigen Repräsentation werden auf die meisten Stimuli nicht antworten, und nur auf wenige ausgewählte Stimuli eine starke Antwort geben. Zellen in den visuellen Kortizes V2 (Hegdé und van Essen, 2003) und V4 (Gallant et al., 1996) in Primaten weisen diese hochgradige Selektivität für spezifische Klassen komplexer Stimuli auf. Auch die Antworten von V2-Neurone auf *Chevron*-Stimuli mit unterschiedlichem Öffnungswinkel und Orientierung zeigen diese Selektivität (Ito und Komatsu, 2004).

Abb. 5.21 illustriert die Antwortselektivität des Hypercomplex-Modells auf ein Stimulus-Testset analog zu der neurophysiologischen Untersuchung in (Hegdé und van Essen, 2003). Wie in diesem und anderen Experimenten gezeigt, reagiert das Neuron nur auf spezifische Subklassen der Teststimuli. Vergleicht man diese Abstimmung auf komplexe Muster mit der eines linearen Gabor-Modells, so sieht man, daß dieses eine weitaus breitere Abstimmung aufweist und somit weitaus weniger selektiv auf die angebotenen Teststimuli reagiert. Ein anderes *i2D*-selektives Modellneuron, ähnlich der Dot-Responsive-Zelle¹⁴ antwortet lediglich auf *Chevron*-Stimuli mit bestimmter Orientierung

¹⁴Der Unterschied liegt in der zusätzlichen Orientierungsselektivität des Modells.

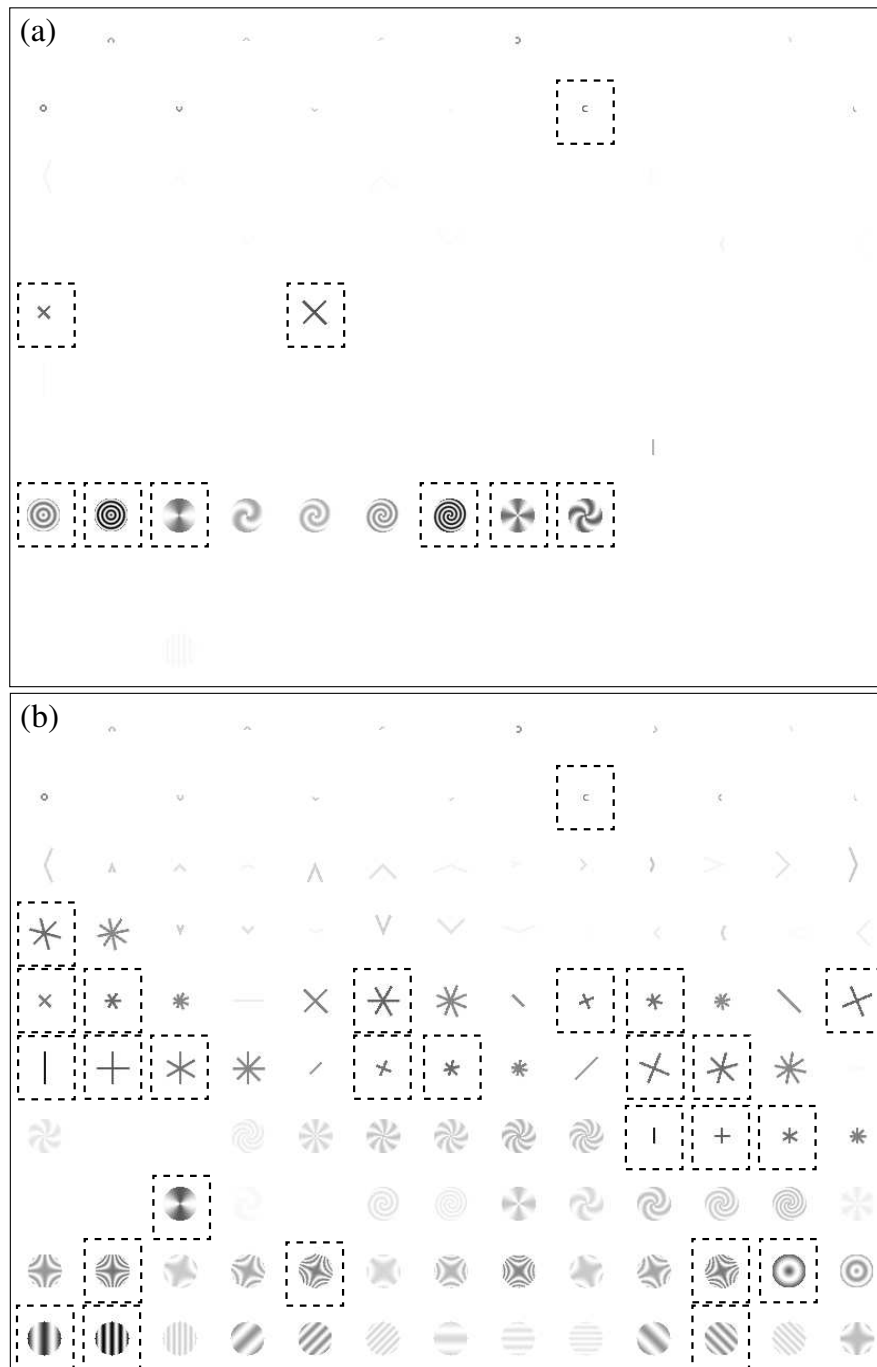


Abbildung 5.21: (a) Antwort des Hypercomplex-Modells auf eine Menge von Teststimuli, vergleichbar mit den in (Hegd  und van Essen, 2003) verwendeten. Es fallt auf, da das Modell in einer stark selektiven Weise antwortet. (b) Ein lineares Gabor-Modell als Vergleich kann diese Selektivitat nicht erreichen. Die gestrichelten Kastchen weisen die 3dB-Antworten aus (Stimuli, die mindestens 50% der Maximalantwort hervorrufen). Siehe auch (Nuding und Zetzsche, 2004).

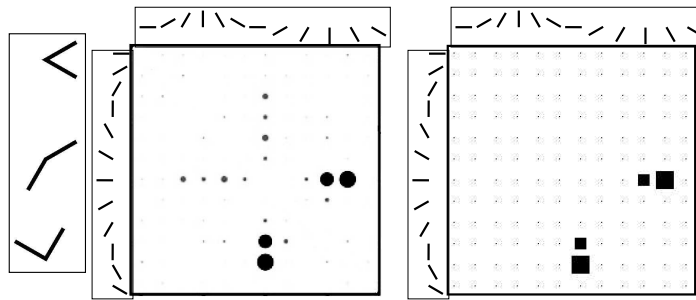


Abbildung 5.22: Links: Antworten eines V2-Neurons auf Chevron-Stimuli mit unterschiedlichen Öffnungswinkeln und Orientierungen (nach (Ito und Komatsu, 2004)). Drei Stimuli sind beispielhaft links außen gezeigt. Rechts: Antworten eines $i2D$ -selektiven Modells. Die Fläche der Kreise, bzw. Quadrate sind proportional zu den Feuerraten.

und Öffnung, siehe Abb. 5.22 und (Ito und Komatsu, 2004).

Ein weiteres bekanntes Phänomen sind nichtlineare Interaktionseffekte, bei welchen Stimuli, die außerhalb Selektivität eines Neurons liegen, signifikant die Antwort auf den optimalen Stimulus beeinflussen können. Dies ist ein klassisches bilineares Antwortverhalten, das im letzten Abschnitt mit Hilfe Abb. 5.12 für den einfachen Fall der Frequenzselektivität eingeführt wurde. Empirisch wurde dies sehr deutlich in (Shevelev et al., 1998) nachgewiesen. Bereits ein einfaches Produktsystem bestehend aus relativ breitbandigen Filtern kann dieses Verhalten sehr gut modellieren (Zetsche et al., 2001; Zetsche und Nuding, 2005b).

Die ebenfalls im letzten Abschnitt besprochenen extra-klassischen Effekte sind für einfache Konfigurationen, wie in Abb. 5.15 gezeigt, leicht zu modellieren. Aber auch komplexere Messungen, wie etwa das Antwortverhalten auf sinusförmige Annuli (Sillito et al., 1995) und assoziierte Suppressionsindizes sind bereits mit quadratischen Systemen hervorragend modellierbar. Für Details siehe hierzu ebenfalls (Zetsche et al., 2001; Zetsche und Nuding, 2005b).

5.2.5 Identifikation von Volterra-Kernen zweiter Ordnung

Die konstruktive Generierung von neuronalen Eigenschaften sowie die Anpassung bestehender Modelle an das neuronale Antwortverhalten stellen eine geeignete Methode dar, um neuronale Systeme zu modellieren. Dies impliziert jedoch, daß bestimmte strukturelle Eigenschaften als gegeben angenommen werden. Im Allgemeinen entzieht sich dieses *a priori* Wissen der Kenntnis, womit die Modellierung unter falschen Randbedingungen erfolgt und zu suboptimalen Ergebnissen führt. Desweiteren besteht die Gefahr, daß durch die einschränkenden Randbedingungen wesentliche funktionale Eigenschaften der zu modellierenden Systeme verborgen bleiben und die Systeme unangemessen simplifiziert werden. Wenn man beispielsweise annimmt, daß das wahre System

aus dem einfachen Produkt zweier linearer Filter mit nicht oder nur schwach überlappender Orientierungsabstimmung besteht, so kann ein Wiener-Modell durchaus einige relevante Aspekte dieses Systems modellieren (z.B. Orientierungsabstimmung für Balken oder das Auftreten im Eingangssignal nicht vorhandener Radialfrequenzkomponenten). Der wesentliche Aspekt dieses Systems, nämlich das Produkt zweier Orientierungen, was zu einer Hervorhebung von $i2D$ -Signalen führt, ist damit allerdings nicht sichtbar. In der Literatur werden Identifikationsverfahren, die eine gewisse Modellstruktur implizieren, häufig als *parametrische Identifikation* bezeichnet; es werden hierbei die Koeffizienten bestimmter Systemparameter gelernt. Dies ist dann von Vorteil, wenn eine hohe Auflösung gewünscht ist, oder das Optimierungsproblem im Allgemeinen zu komplex ist (in Abschnitt 5.2.6 wird ein System mittels eines parametrischen Verfahrens an ein Optimalitätskriterium angepaßt).

Allgemeine Verfahren, die neuronale Eigenschaften direkt auf Systemkerne abbilden, sind mit den Methoden der blinden Systemidentifikation gegeben. Diese ermöglichen eine Identifikation des Systems, indem bei definiertem Eingangssignal der Ausgang beobachtet wird¹⁵. Man muß allerdings anmerken, daß die in dieser Arbeit verwendeten Volterra-Systeme zweiter Ordnung selbst starken Einschränkungen unterliegen, da diese nur quadratische Polynomapproximationen darstellen. Diese Einschränkung ist allerdings aufgrund von Komplexitätsüberlegungen notwendig; die postulierte hochgradig relevante Eigenschaft der $i2D$ -Selektivität ist damit allerdings experimentell verifizierbar.

Es werden im Folgenden vier Verfahren vorgestellt, mit welchen der quadratische Volterra-Kern von nichtlinearen Systemen gemessen werden kann. Dabei sind die ersten beiden Verfahren mehr akademischer Natur und für die Praxis eher ungeeignet (Identifikation mit gepaarten Dirac-Impulsen und Kombination von harmonischen Schwingungen). Weitaus mehr Relevanz für die praktische Analyse und Identifikation von visuellen Neuronen bergen die anderen zwei Verfahren. Das erste basiert auf der erstmals von TICK vorgeschlagenen Methode (Tick, 1961) und stellt die direkte Erweiterung der Kreuzkorrelationsmethode für lineare Systeme mit gaußischem Eingang dar. Im Kontext dieser Arbeit wird die Methodik für die Messung vierdimensionaler Kerne im Frequenzbereich generalisiert und hinsichtlich des Konvergenzverhaltens und der Rauschstabilität für unterschiedliche Rauscharten analysiert. Bekannte Anwendungen sind bisher beschränkt auf die Identifikation eindimensionaler neuronaler Systeme im Zeitbereich (Marmarelis und Marmarelis, 1978), wobei die relevanten funktionalen Eigenschaften (s.o.) als integrale Bereiche abgebildet werden, und somit „by inspection“ nicht sichtbar sind. Zudem wurde typischerweise *weißes* gaußisches Rauschen als Stimulus verwendet, da dies leichter zu generieren ist und das Meßverfahren schneller konvergiert. Allerdings hat sich gezeigt, daß v.a. visuelle Neurone in der Regel nur sehr schwach oder gar nicht auf wei-

¹⁵Es existieren auch Verfahren, die nur aufgrund von Beobachtungen des Ausgangs Rückschlüsse auf das System zulassen, diese werden hier aber nicht behandelt.

bes Rauschen antworten. Daher wird die allgemeine Form der Methodik verwendet, die theoretisch beliebige Bindungen zweiter Ordnung im Eingangssignal erlaubt (farbiges Rauschen) und daher besser geeignet ist, um reale Neurone zu treiben. Das letzte untersuchte Verfahren geht aus der Motivation hervor, daß auch farbiges gaußsches Rauschen keinen optimalen Stimulus für visuelle Neurone darstellt, und daß diese am Besten auf natürliche Stimuli antworten, bzw. eine Charakterisierung bezüglich der Antwort auf natürliche Stimuli sinnvoll erscheint (Theunissen et al., 2001). Dieses Verfahren basiert auf der von KIM und POWERS entwickelten Methode zur Identifikation quadratischer Systeme mit nicht-gaußschem Eingang (Kim und Powers, 1988) und wird hier auf die Analyse zweidimensionaler Signale erweitert und für die Verwendung von natürlichen Stimuli untersucht.

Identifikation mit gepaarten Dirac-Impulsen

Ausgehend von den Gln. (5.80 – 5.83) kann der Kern eines quadratischen Volterra-Systems mithilfe zweier Dirac-Impulse bei $\vec{x} = \vec{0}$ und $\vec{x} = \Delta\vec{x}$ bestimmt werden; siehe auch (Krieger, 1999):

$$h_2(\vec{x}, \vec{x} - \Delta\vec{x}) = \frac{1}{2} \left(u_2(\vec{x}; \Delta\vec{x}) - h_2(\vec{x}, \vec{x}) - h_2(\vec{x} - \Delta\vec{x}, \vec{x} - \Delta\vec{x}) \right), \quad (5.110)$$

wobei $u_2(\vec{x}; \Delta\vec{x})$ die Systemantwort auf den gepaarten Dirac-Impuls und $h_2(\vec{x}, \vec{x})$, bzw. $h_2(\vec{x} - \Delta\vec{x}, \vec{x} - \Delta\vec{x})$ die Antworten auf die Einzelimpulse darstellen. Man erhält also einen Schnitt durch den Kern an der Stelle $\vec{x}_2 = \vec{x}_1 - \Delta\vec{x}$, indem man von der Antwort auf den Doppelimpuls die Summe der Einzelantworten abzieht. Wie auch bei der Impulsantwort für lineare Systeme eignet sich dieses Verfahren in praxi kaum zur Identifikation des Kerns, hauptsächlich aufgrund der „mathematischen“ Beschaffenheit des Dirac-Impulses (unendliche Höhe bei infinitesimaler Dauer, bzw. Ausdehnung).

Identifikation mit Kombinationen von harmonischen Schwingungen

Das Verfahren zur Messung des quadratischen Volterra-Kerns im Frequenzbereich mit Kombinationen von harmonischen Schwingungen¹⁶ wurde bereits in (Nuding, 2002) hergeleitet und implementiert, daher folgt hier nur eine kurze Zusammenfassung.

Man definiert eine Systemantwort \mathcal{H}_{+S+S} , bestehend aus der Antwort eines quadratischen Systems auf die Summe zweier Sinusschwingungen unterschiedlicher Frequenz, minus den Einzelantworten:

$$\mathcal{H}_{+S+S} = \mathcal{H}\{\sin(2\pi\vec{f}_1^T\vec{x}) + \sin(2\pi\vec{f}_2^T\vec{x})\} - \mathcal{H}\{\sin(2\pi\vec{f}_1^T\vec{x})\} - \mathcal{H}\{\sin(2\pi\vec{f}_2^T\vec{x})\}. \quad (5.111)$$

¹⁶Im Fall zweidimensionaler Systeme sind dies Sinusgitter.

Analog definiert man eine Systemantwort \mathcal{H}_{+C-C} , welche sich aus der subtraktiven Kombination zweier Cosinusschwingungen ergibt:

$$\mathcal{H}_{+C-C} = \mathcal{H}\{\cos(2\pi\vec{f}_1^T\vec{x}) - \cos(2\pi\vec{f}_2^T\vec{x})\} - \mathcal{H}\{\cos(2\pi\vec{f}_1^T\vec{x})\} - \mathcal{H}\{-\cos(2\pi\vec{f}_2^T\vec{x})\}. \quad (5.112)$$

Berechnet man nun unter Zuhilfenahme der Gln. (5.97) und (5.98) die Summe dieser Systemantworten, so ergibt sich:

$$\mathcal{H}_{+S+S} + \mathcal{H}_{+C-C} = -2\Re\{H_2^{sym}(\vec{f}_1, \vec{f}_2)e^{j2\pi(\vec{f}_1+\vec{f}_2)^T\vec{x}}\} =: \mathcal{M}_{\Re}(\vec{f}_1, \vec{f}_2, \vec{x}). \quad (5.113)$$

Wertet man die Antwort an der Stelle $\vec{x} = 0$ aus, so erhält man bis auf einen Skalierungsfaktor direkt den Realteil des Kerns an der Stelle (\vec{f}_1, \vec{f}_2) :

$$\mathcal{M}_{\Re}(\vec{f}_1, \vec{f}_2, \vec{x})\Big|_{\vec{x}=0} = -2\Re\{H_2^{sym}(\vec{f}_1, \vec{f}_2)\}. \quad (5.114)$$

Ebenso läßt sich zeigen, daß der Imaginärteil über die Antwort auf gemischte Kombinationen von Sinus und Cosinus berechnet werden kann,

$$\mathcal{H}_{+S+C} + \mathcal{H}_{+C+S} = 2\Im\{H_2^{sym}(\vec{f}_1, \vec{f}_2)e^{j2\pi(\vec{f}_1+\vec{f}_2)^T\vec{x}}\} =: \mathcal{M}_{\Im}(\vec{f}_1, \vec{f}_2, \vec{x}), \quad (5.115)$$

wobei der Imaginärteil des Systemkerns wiederum durch Auswerten der Antwort an der Stelle Null erhalten werden kann:

$$\mathcal{M}_{\Im}(\vec{f}_1, \vec{f}_2, \vec{x})\Big|_{\vec{x}=0} = 2\Im\{H_2^{sym}(\vec{f}_1, \vec{f}_2)\}. \quad (5.116)$$

Dieses Verfahren wäre prinzipiell geeignet, um realistische Messungen durchzuführen, leider aber unter dem Nachteil, daß für jede Frequenzkombination explizit eine Messung durchgeführt werden muß, bzw. mehrere, um gegebenenfalls auftretendes Meßrauschen auszumitteln. Zudem reagiert das Verfahren sensibel auf die Phasenlage des Meßpunkts, wie direkt den Gln. (5.113), bzw. (5.115) entnommen werden kann.

Identifikation mit Gaußschem Eingang

Basierend auf dem Theorem von LEONOV UND SHIRYAEV in der Theorie von Kumulantenpektren (siehe (Leonov und Shiryayev, 1959) und Gl. (5.10)) konnte TICK erstmals zeigen, wie sich der quadratische Kern eines inhomogenen Volterra-Systems zweiter Ordnung bei gaußschen Eingangssignalen über das Kreuz-Bispektrum zwischen Ein- und Ausgang schätzen läßt. Im Allgemeinen wird hierbei von dekorrelierten (weißen) Eingangssignalen ausgegangen, was Vorteilhaft ist, da diese einfacher zu generieren sind und der Schätzer schneller konvergiert. Für die Messung der Kerne von visuellen Neuronen muß aber beachtet werden, daß diese nicht oder nur sehr schwach auf weißes Rauschen antworten (s.o.). Daher wird im Folgenden gezeigt, daß das Verfahren auch für farbiges gaußsches Rauschen gültig ist.

Die Herleitung erfolgt über die Kreuzkumulante dritter Ordnung zwischen dem Eingang des Systems, $u_1(\vec{x})$, und dem Ausgang, $u_2(\vec{x})$:

$$c_3^{\mathbf{u}_1 \mathbf{u}_1 \mathbf{u}_2}(\vec{x}_1, \vec{x}_2) = E\{\mathbf{u}_1(\vec{x} + \vec{x}_1)\mathbf{u}_1(\vec{x} + \vec{x}_2)\tilde{\mathbf{u}}_2(\vec{x})\}. \quad (5.117)$$

Hierbei ist $\tilde{\mathbf{u}}_2(\vec{x}) = \mathbf{u}_2(\vec{x}) - m_{u_2}$ der vom Mittelwert befreite Ausgangsprozess. Wenn man nun die Vorschrift zur Berechnung des Ausgangs eines homogenen¹⁷ quadratischen Volterra-Systems, also

$$u_2(\vec{x}) = \int \int h_2(\vec{x}_1, \vec{x}_2)u_1(\vec{x} - \vec{x}_1)u_1(\vec{x} - \vec{x}_2) \cdot d\vec{x}_1 d\vec{x}_2, \quad (5.118)$$

in Gl. (5.117) einsetzt, und für den Mittelwert des Ausgangs

$$m_{u_2} = \int \int h_2(\vec{x}_1, \vec{x}_2)E\{u_1(\vec{x} - \vec{x}_1)u_1(\vec{x} - \vec{x}_2)\} \cdot d\vec{x}_1 d\vec{x}_2 \quad (5.119)$$

verwendet, so erhält man einen relativ komplexen Ausdruck, welcher sich aber deutlich vereinfachen läßt, wenn als Eingangssignal ein mittelwertfreier Gaußscher Rauschprozeß angenommen wird. Bei diesem lassen sich Momente vierter Ordnung vereinfacht als Produkte von Momenten zweiter Ordnung schreiben, womit das Ergebnis folgende Form annimmt:

$$c_3^{\mathbf{u}_1 \mathbf{u}_1 \mathbf{u}_2}(\vec{x}_1, \vec{x}_2) = 2 \int \int h_2(\vec{\alpha}_1 - \vec{x}_1, \vec{\alpha}_2 - \vec{x}_2) \cdot c_2^{\mathbf{u}_1}(\vec{\alpha}_1)c_2^{\mathbf{u}_1}(\vec{\alpha}_2) \cdot d\vec{\alpha}_1 d\vec{\alpha}_2. \quad (5.120)$$

Dieses Ergebnis wurde in ähnlicher Form bereits in (Schetzen, 1980) hergeleitet. Durch eine einfache weitere Integralumformung und Fourier-Transformation folgt schließlich das Endergebnis:

$$H_2(-\vec{f}_1, -\vec{f}_2) = \frac{C_3^{\mathbf{u}_1 \mathbf{u}_1 \mathbf{u}_2}(\vec{f}_1, \vec{f}_2)}{2C_2^{\mathbf{u}_1}(\vec{f}_1)C_2^{\mathbf{u}_1}(\vec{f}_2)}, \quad (5.121)$$

wobei $C_3^{\mathbf{u}_1 \mathbf{u}_1 \mathbf{u}_2}(f_1, f_2)$ das Kreuz-Bispektrum zwischen dem Eingang \mathbf{u}_1 und dem Ausgang \mathbf{u}_2 darstellt und $C_2^{\mathbf{u}_1}(\vec{f})$ das Leistungsdichtespektrum des Eingangs bezeichnet. Das bedeutet also, daß das Eingangssignal zwar gaußsch sein muß, die spektralen Eigenschaften, also die statistischen Bindungen zweiter Ordnung, weitestgehend frei wählbar sind. Dieses Ergebnis kann auch leicht auf homogene Systeme beliebiger Ordnung verallgemeinert werden; für inhomogene Systeme wurde eine ähnliche Vorschrift in (Koukoulas und Kalouptsidis, 1995) abgeleitet.

Beim Messen von echten Neuronen müssen zwei ganz wesentliche Faktoren beachtet werden: (i) Die Messung ist extrem rauschbehaftet und (ii) Neurone tendieren zu Instationaritäten aufgrund von Adaptationsvorgängen. Aus diesem Grund wird im Folgenden das

¹⁷Bei Verwendung eines inhomogenen Systems zweiter Ordnung ändert sich nichts, da die entsprechend auftretenden Kumulanten bei einem gaußschem Eingang zu Null werden; allerdings wird die Rechnung umständlicher.

Konvergenzverhalten für sehr wenig Daten und sehr niedriger Signal-zu-Rauschleistung analysiert. Als Rauschquelle (Meßrauschen, neuronales Rauschen, etc.) wird zuerst additives weißes gaußsches Rauschen (AWGN) angenommen. Aus der Theorie der Polyspektren ist bekannt (vgl. Abschnitt 5.1), daß Kumulanten, bzw. Polyspektren invariant gegenüber additiven gaußschen Größen sind, insbesondere natürlich auch Rauschen. An dieser Stelle interessiert jedoch das tatsächliche Konvergenzverhalten im Grenzbereich für extrem niedrige SNR-Werte. Exemplarisch sind hier die Ergebnisse für die Modellneurone aus Abschnitt 5.2.4 (Complex, Dot-Responsive, Hypercomplex) in Abb. 5.23 (a) gezeigt. Es ist ersichtlich, daß das Konvergenzverhalten für eine verlässliche Schätzung der Kerne spricht, da der mittlere quadratische Fehler bereits für sehr wenig Daten und niedriges SNR schnell gegen Null konvergiert, vgl. (Nuding et al., 2004; Nuding und Zetzsche, 2004). Zur Verdeutlichung sind jeweils die geschätzten Kerne für $N = 640$ Datenblöcke und $SNR = 0.04$ gezeigt. Man beachte, daß dies ca. -14 dB entspricht und das Rauschen eine um den Faktor 25 höhere Leistung als das gemessene Signal hat. Die theoretische Invarianz der Kumulanten bestätigt sich also auch für extrem niedrige SNR-Werte bei bereits wenig Datenblöcken. Abb. 5.23 (b) zeigt Schnitte durch diese Fehlerflächen für die maximale Anzahl Datenblöcke ($N = 640$), aufgetragen als das *Peak Signal-to-Noise-Ratio* (PSNR) des Schätzers gegenüber dem SNR der Messung.

Wie bereits erläutert, ist dieses Meßverfahren für gaußsches Rauschen mit beliebigen spektralen Eigenschaften gültig. Daher wurde explizit untersucht, wie sich der Schätzer für einen Eingang verhält, der besser in der Lage ist, visuelle Neurone zu treiben. Die typischen spektralen Eigenschaften natürlicher Stimuli wurden bereits in Abschnitt 2.4.3 beschrieben, und sind mit einem $1/f^2$ -Abfall der Leistungsdichte modellierbar. Die entsprechenden Messungen der Modellneurone mit einem derartigen Eingang zeigt Abb. 5.23 (c). Man sieht, daß auch hierfür sehr gute Werte erreicht werden, die gemessenen Kerne entsprechen denen aus Abb. 5.23 (a). Wie bereits in Abb. 5.23 (b) zu sehen ist, lassen sich die Kerne der Wiener- und Hammerstein-Modelle schlechter schätzen, d.h. sie sind empfindlicher gegenüber Rauschen und konvergieren langsamer bzgl. der Datenmenge. Dieser Effekt liegt in der relativ höheren Selektivität der Modellneurone vs. den Kaskaden-Modellen begründet. Das wiederum bedeutet, daß die erhöhte Selektivität nicht nur zu einer robusten und effizienten Kodierung führt, wie in Kap. 2 beschrieben, sondern daß auch die Identifikations- und Klassifikationsleistung dadurch erhöht wird.

Ein weiterer Punkt der untersucht wurde, ist die Messung unter der biologisch plausibleren Annahme einer *multiplikativen* gaußschen Störung. Dies liegt darin begründet, daß die Feuerraten typischerweise als Poisson-verteilte Zufallsgrößen betrachtet werden, und die Varianz des Rauschens proportional zur mittleren Feuerrate ist. Damit erhält das Rauschen dann einen multiplikativen Effekt (Papoulis, 1991; Gottschalk et al., 2004). Auch für diese Art der Störung wurden qualitativ gleiche Ergebnisse gefunden, obwohl dies ad hoc nicht ohne Weiteres theoretisch zu begründen ist. Zusammenfassend läßt sich sagen, daß dieses Meßverfahren speziell für die erwartete hochselektive Klasse der kortikalen Neurone sehr gute Ergebnisse liefert, deren Qualität weiter verbessert werden

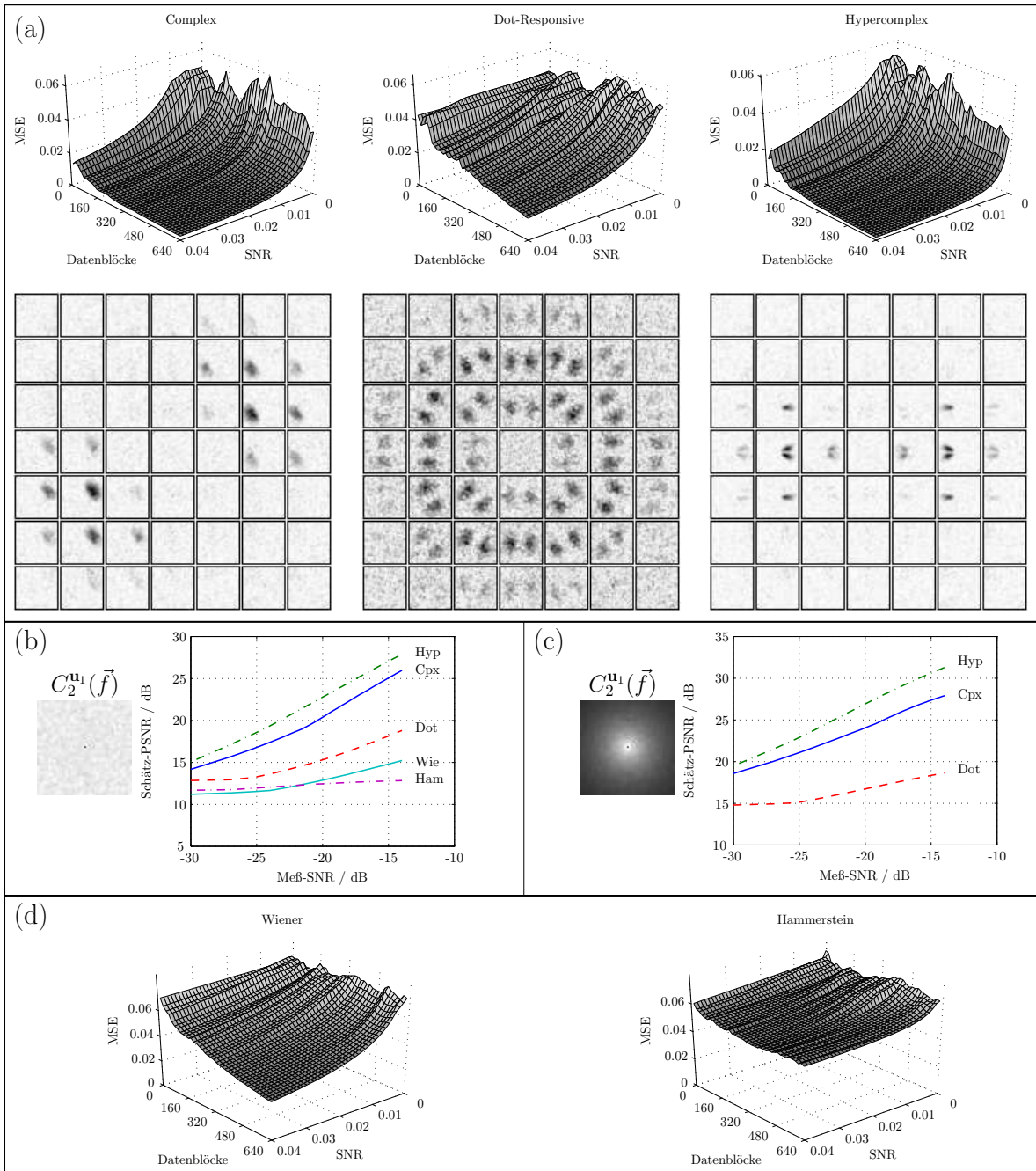


Abbildung 5.23: (a) Mittlerer quadratischer Schätzfehler (MSE) in Abhängigkeit der Anzahl der verfügbaren Daten sowie des Signal-Rauschleistungsverhältnisses (SNR) für WGN Eingang und AWGN-Rauschquelle. Gezeigt sind die jeweils geschätzten Kerne für $N = 640$ Datenblöcke und $SNR = 0.04$. (b) Peak Signal-to-Noise-Ratio Kurven der Modelle aus (a) sowie eines Wiener- und Hammerstein- Modells für die maximale Anzahl Daten ($N = 640$). (c) Dto., für $1/f^2$ -GN Eingang. (d) Zum Vergleich die Fehlerflächen entsprechend (a) für die Wiener- und Hammerstein-Modelle.

kann, wenn lediglich signifikante Strukturen in den Kernen identifiziert werden sollen (also beispielweise Demodulation oder *i2D*-Selektivität).

Identifikation mit natürlichem Eingang

Wünschenswert für die Analyse eines biologischen Systems ist es, das System in seinem natürlichen Arbeitspunkt zu beobachten. Für das visuelle System bedeutet das, während des freien Beobachtens natürlicher Bilder analysiert zu werden. Daher wird ein Meßverfahren benötigt, das in der Lage ist, Volterra-Kerne mit beliebigen Eingangssignalen zu messen. Obwohl dies bei linearen Systemen, zumindest theoretisch, leicht zu bewerkstelligen ist, muß bei Systemen höherer Ordnung ein etwas komplexerer Ansatz verfolgt werden. Hierbei wird ausgenutzt, daß Volterra-Systeme *linear* in ihren Kernkoeffizienten sind, wie weiter unten gezeigt wird. Unter Verwendung der Matrix-Vektor-Notation für *diskrete* Volterra-Systeme können dann die notwendigen Kerneinträge für einen bestimmten Ausgang berechnet werden.

Diskret kann der quadratische Volterra-Operator im Frequenzbereich wie folgt ausgedrückt werden:

$$U_2[\vec{k}] = \sum_{k_{x_1}=-\frac{M}{2}}^{\frac{M}{2}-1} \sum_{k_{y_1}=-\frac{M}{2}}^{\frac{M}{2}-1} H_2[\vec{k}_1, \vec{k} - \vec{k}_1] \cdot U_1[\vec{k}_1] U_1[\vec{k} - \vec{k}_1]. \quad (5.122)$$

Das Argument $\vec{k} = [k_x, k_y]^t$ ist der Frequenzindex der *diskreten Fourier-Transformation* (DFT) mit Segmentlänge, bzw. Blockgröße M . Bei der tatsächlichen Implementierung muß allerdings beachtet werden, daß die Summationsgrenzen entsprechend der Länge der Vektoren und der Werte k_x, k_y angepaßt werden müssen. Auf die explizite Formulierung wird hier aus Gründen der besseren Lesbarkeit verzichtet. Um die Idee zu verdeutlichen, wird Gl. (5.122) für den eindimensionalen Fall ausgeschrieben:

$$\begin{aligned} U_2[k] &= H_2\left[-\frac{M}{2}, k + \frac{M}{2}\right] U_1\left[-\frac{M}{2}\right] U_1\left[k + \frac{M}{2}\right] \\ &+ H_2\left[-\frac{M}{2} + 1, k + \frac{M}{2} - 1\right] U_1\left[-\frac{M}{2} + 1\right] U_1\left[k + \frac{M}{2} - 1\right] + \dots \\ &+ H_2\left[\frac{M}{2} - 1, k - \frac{M}{2} + 1\right] U_1\left[\frac{M}{2} - 1\right] U_1\left[k - \frac{M}{2} + 1\right]. \end{aligned} \quad (5.123)$$

Gl. (5.123) kann unter Verwendung der Matrix-Vektor-Notation auch geschrieben werden als

$$U_2[\vec{k}] = \vec{H}^t[\vec{k}] \vec{U}_1[\vec{k}], \quad (5.124)$$

wobei $\vec{H}^t[\vec{k}]$ und $\vec{U}_1[\vec{k}]$ entsprechend angeordnete Vektoren, bestehend aus den Kerneinträgen, bzw. aus Produkten der Eingangswerte, sind. Wie sich unschwer erkennen läßt,

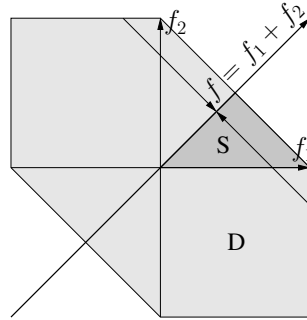


Abbildung 5.24: Projektionsschema und Support eines eindimensionalen quadratischen Volterra-Systems. Zur Berechnung einer spektralen Komponente der Antwort wird nur ein Bruchteil der Kerneinträge benötigt, weshalb die Komplexität, d.h. Größe der zu invertierenden Matrix in Gl. (5.125), stark abnimmt (S: Summeninteraktions-Region, D: Differenzinteraktions-Region).

ist Gl. (5.124) linear in dem generalisierten Übertragungsfunktionsvektor, somit kann für $\vec{H}[\vec{k}]$ der folgende Schätzer angegeben werden (Nuding et al., 2004):

$$\vec{H} = \left[E\{\vec{U}_1^* \vec{U}_1^t\} \right]^{-1} E\{\vec{U}_1^* U_2\}. \quad (5.125)$$

Obwohl dieses Verfahren eine direkte, und damit typischerweise eine sehr rechenintensive, Parameterschätzung darstellt, reduziert sich die Komplexität signifikant aufgrund der bei der Berechnung der Antwort eines Volterra-Operators auftretenden Projektion der expandierten Antwort auf die Hauptdiagonale (Abb. 5.24). Beispielsweise wurde in den Simulationen mit einem Support von 32 Stützstellen pro Dimension gerechnet; dies korrespondiert mit einem $32^4 = 1048576$ -dimensionalem Parameterraum. Damit wäre die Größe der zu invertierenden Matrizen bei einer *Least-Squares* Schätzung $32^4 \times 32^4$. Mit diesem Verfahren jedoch reduziert sich die Größe auf $32^2 \times 32^2$, was dann auch in vernünftiger Zeit berechenbar ist.

Man beachte, daß entsprechend Gl. (5.125) spektrale Momente vierter Ordnung geschätzt werden müssen; damit geht einher, daß der Schätzer eine größere Varianz besitzt als der zuletzt vorgestellte¹⁸. Allerdings weist der Algorithmus dennoch ähnliche Konvergenzeigenschaften und Fehlerrobustheit auf, wie der Ansatz über das Kreuz-Bispektrum, wenn WGN als Eingangssignal verwendet wird (vgl. Abb. 5.25). Hinweis: Die in dieser Abbildung gezeigten Kurven liegen über denen des Kreuz-Bispektrum-Schätzers aus Abb. 5.23(b). Dies liegt einerseits darin begründet, daß aufgrund der Rechenintensität der direkten Methode ein kleinerer Support der Kerne verwendet wurde, und andererseits daran, daß die Qualität des ersten Schätzers darunter leidet, daß der hypothetisch gaußsche Eingang nie ideal ist. Da die direkte Methode davon nicht explizit Gebrauch

¹⁸Dies ist offensichtlich, da bei dieser Methode keinerlei vereinfachende Annahmen gemacht werden, und somit vom *Worst-Case* ausgegangen werden muß.

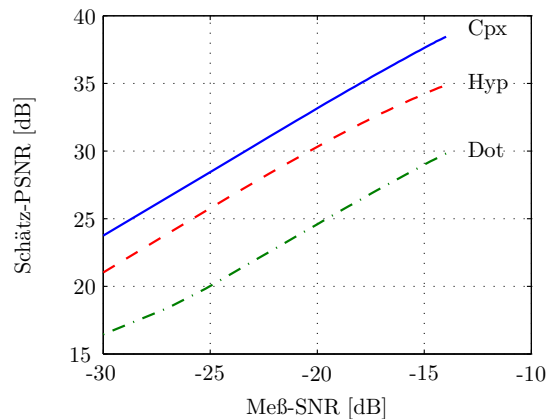


Abbildung 5.25: Peak Signal-to-Noise-Ratio Kurven der drei Modelle (Complex, Hypercomplex und Dot-Responsive) der Schätzung in Abhängigkeit des Meßrauschens (AWGN) in dB bei Verwendung eines gaußschen Eingangs.

macht, sind die Ergebnisse für einen nicht-idealen gaußschen Eingang typischerweise etwas besser (Kim und Powers, 1988).

Werden nun allerdings natürliche Bilder für die Identifikation verwendet, so ist die Fehlerempfindlichkeit dieses direkten Ansatzes relativ hoch. Hier wurden die Antworten auf 5120 Segmente natürlicher Bilder analysiert. Bei Verwendung einer idealen rauschfreien Messung ergeben sich hervorragende Ergebnisse (nicht gezeigt), jedoch ist hier in erster Linie das Verhalten unter Einfluß von Störuschen relevant. Daher wurde diese Messung mit natürlichen Bildern mit einem starken AWG-Rauschen beaufschlagt, so daß ein SNR von -6dB entsteht. Man sieht deutlich, daß die Messungen deutlich rauschbehaftet sind, dies liegt daran, daß in diesem allgemeinen Fall selbst gaußsche Störgrößen mitberücksichtigt werden müssen. Nichtsdestoweniger können durchaus relevante Strukturen in den Kernen identifiziert werden, wie in Abb. 5.26 gezeigt ist.

5.2.6 Dekorrelation höherer Ordnung

Wie in den Abschnitten 5.1.1 und 5.1.2 bereits erläutert wurde, erfassen Kumulantenpektren höherer Ordnung (Polyspektren) statistische Bindungen zwischen verschiedenen Frequenzkomponenten. Für einen nicht-gaußschen Rauschprozeß ohne Bindungen höherer Ordnung gilt für seine Polyspektren der Ordnung n

$$C_n^{HOWN}(f_1, \dots, f_{n-1}) = \gamma_n. \quad (5.126)$$

Diese sind unabhängig vom Argument und nehmen jeweils den konstanten Wert γ_n an. Ein derartiges Rauschen wird auch als *Higher-Order White Noise* (HOWN) bezeichnet. Damit geht einher, daß ein Prozeß ohne Bindungen dritter Ordnung ein weißes Bispektrum aufweist. Es kann also formal nach einer Transformation gesucht werden, welche

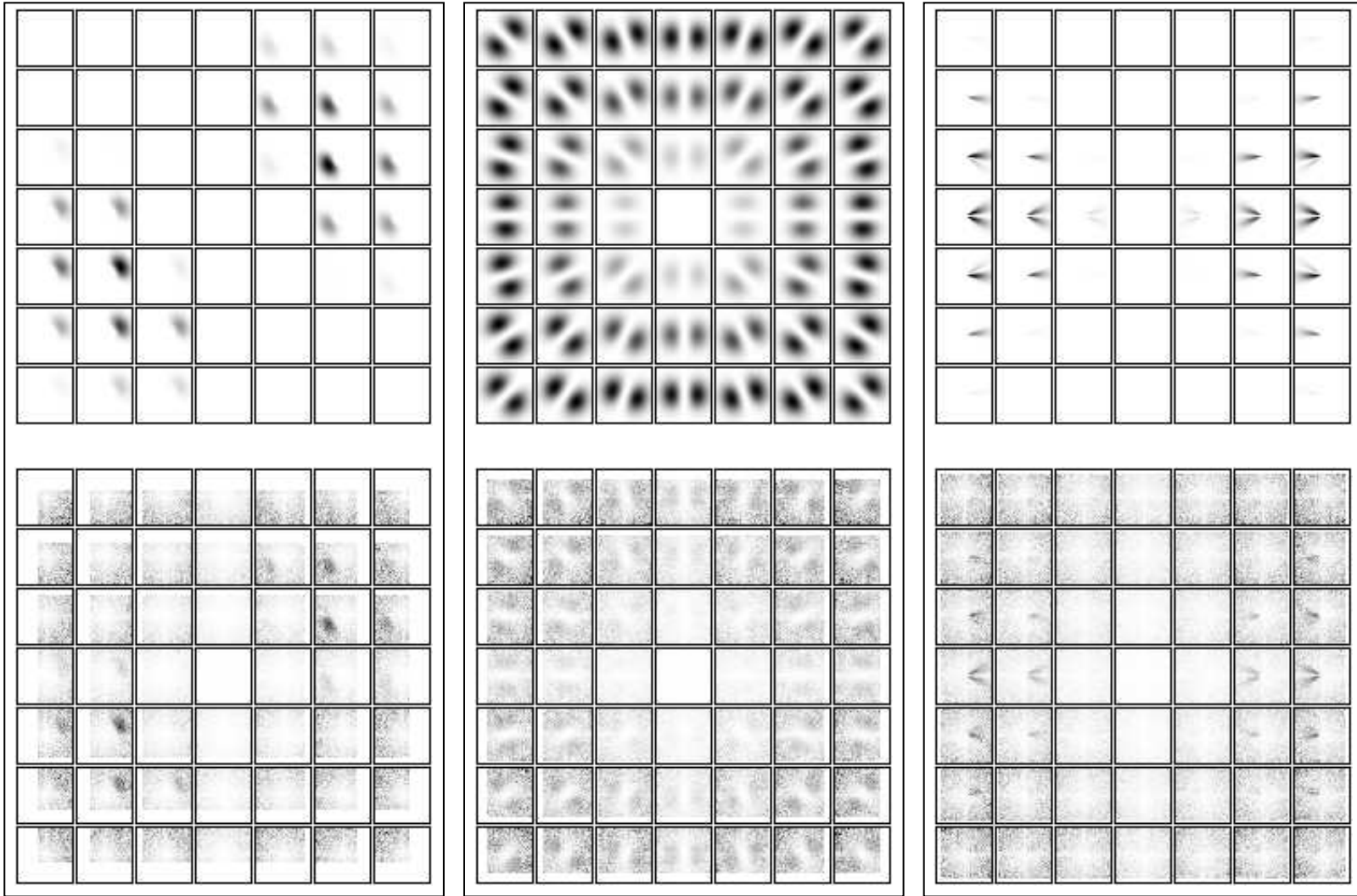


Abbildung 5.26: Modellkerne (oben) und gemessene Kerne mit der direkten Methode (unten). Als Stimulus wurden natürliche Bilder verwendet und die Messung mit einem additiven weißen gaußschen Rauschen beaufschlagt (SNR = -6dB).

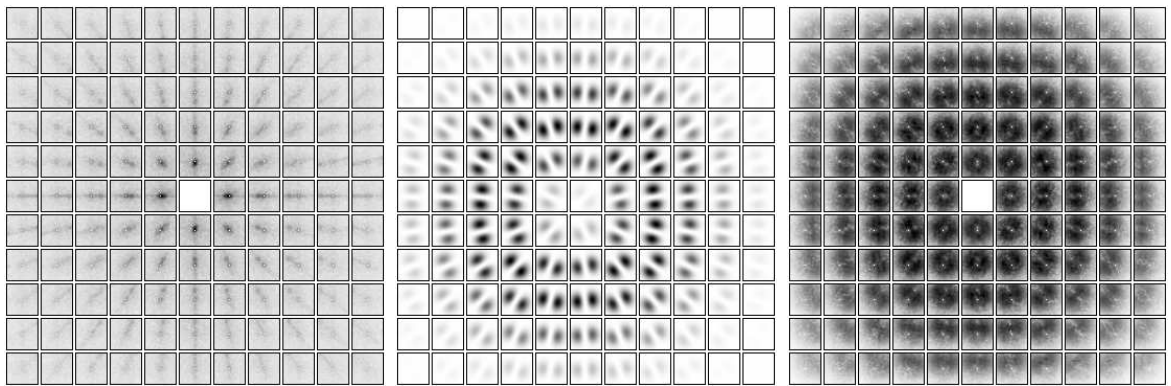


Abbildung 5.27: Links: Bispektrum natürlicher Bilder. Mitte: Kern eines $i2D$ -selektiven quadratischen Volterra-Operators. Rechts: Bispektrum natürlicher Bilder nach $i2D$ -selektiver Filterung.

das Bispektrum weiß macht. Dies ist somit eine direkte Erweiterung des Optimalitätskriteriums zweiter Ordnung zur Dekorrelation eines Signals, das mit der *Spectral Flatness Measure* aus Gl. (2.25) assoziiert ist.

Es ist bekannt, daß die Applikation eines $i2D$ -selektiven Filters auf natürliche Bilder das Bispektrum tendenziell weiß macht und somit statistische Abhängigkeiten dritter Ordnung reduziert, siehe hierzu beispielsweise (Zetzsche und Krieger, 2001a) und Abb. 5.27. Dieser heuristische Ansatz verlangt nach einer weiteren Klärung der Dekorrelation höherer Ordnung für natürliche Bilder. Prinzipiell besteht die Aufgabe darin, ein *inverses* Filter zu finden, dessen Ausgang für natürliche Bilder ein konstantes Bispektrum aufweist. Dieses Problem ist analog zum linearen Fall, mit dem kritischen Unterschied, daß hier im Allgemeinen keine geschlossenen Lösungen zwischen den Bispektren des Ein- und Ausgangs, sowie des charakterisierenden Kerns angegeben werden können (vgl. auch Abschnitt 5.3).

Es soll nun ein empirischer „Vorwärtsbeweis“ für das tatsächliche Dekorrelationspotential von $i2D$ -Systemen für natürliche Bilder geführt werden. Dazu wird ein Modellsystem gelernt indem dessen Parameterraum nach der Lösung mit dem möglichst flachsten assoziierten Bispektrum des Ausgangs abgesucht wird. Abb. 5.28 zeigt ein Schema des Lernverfahrens für ein quadratisches Volterra-System: auf das Eingangssignal (Prozeß der natürlichen Bilder) wird ein quadratischer Volterra-Operator angewendet. Der Ausgangsprozess wird hinsichtlich seines Bispektrums analysiert und die Parameter des Systems werden entsprechend des Optimierungskriteriums nachgeregelt. Dabei müssen, wie bei jedem Optimierungsverfahren, die folgenden Punkte mit Bedacht bearbeitet werden:

1. Wahl der Parametrierung,
2. Bestimmung einer Zielfunktion,
3. Wahl eines numerischen Optimierungsverfahrens.

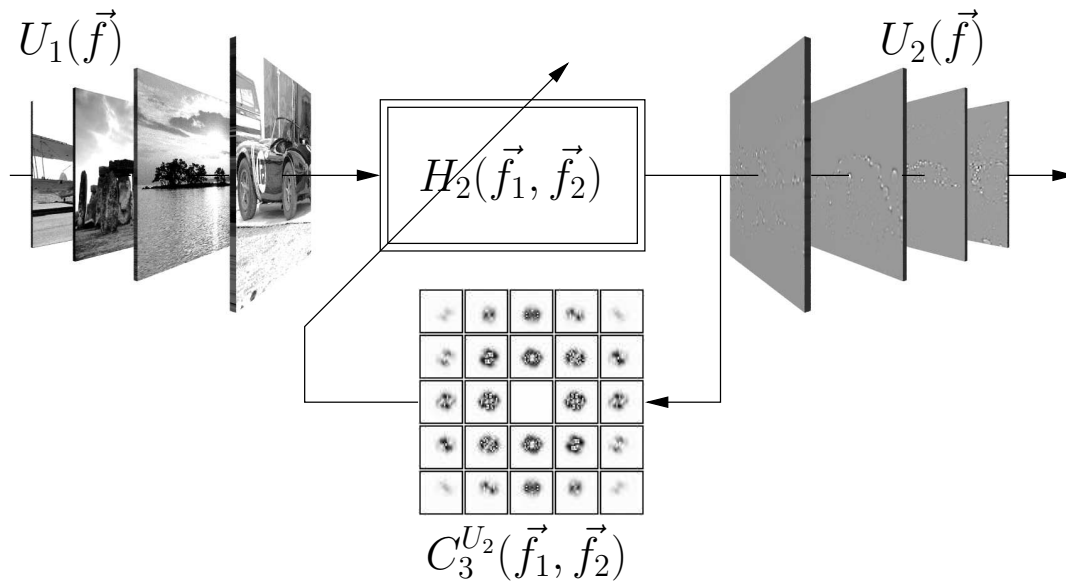


Abbildung 5.28: Schematische Darstellung der Optimierungs-Prozedur (Details siehe Text).

Ad 1. Die Parametrierung stellt häufig den kritischsten Aspekt für nichtlineare Optimierung dar, da hier a priori Wissen impliziert wird, und damit Annahmen über die inhärente Struktur des Systems gemacht werden. Im Falle konstruktiver Modelle, bei welchen die Architektur zu einem hohen Grad invariabel ist, ist dieses Problem marginal. Hier steht jedoch der Aspekt einer gewissen funktionalen Identifikation im Vordergrund, bei der es a priori nicht klar ist, welche Modellarchitektur optimal ist. So bleibt nur, empirisches Wissen entsprechend der Synthesemöglichkeiten aus Abschnitt 5.2.3 einzubringen. Die erste Einschränkung, die hierbei stattfindet ist zuerst einmal die Reduktion der Systemordnung, da hier prinzipiell Systeme zweiter Ordnung betrachtet werden. Man kann dann zwischen einer *direkten* und einer *indirekten* Parametrierung unterscheiden: Bei der direkten Parametrierung wird der vollständige Kern des Systems optimiert. Dies ist ein Verfahren, das bei Systemen mit sehr kleinem Support angewendet wird. Das Ziel ist hierbei häufig, ein Modell an ein unbekanntes System anzupassen, indem die Differenz (mittlerer quadratischer Fehler) zwischen dem Modellausgang und dem unbekanntem Prozeß minimiert wird. Dies entspricht dann allerdings einer Systemidentifikation, die an dieser Stelle nicht von Interesse ist. Die zentrale Frage in dieser Untersuchung ist die Struktur eines Systems, das natürliche Bilder im Sinne dritter Ordnung dekorreliert. Zu zeigen ist dabei, daß die verbotenen Zonen des Optimalsystems nicht belegt sind. Daher ist man gezwungen, Systeme mit größerem Support zu betrachten, was darin resultiert, daß eine direkte Parametrierung nicht angewendet werden kann. Konkret würde ein Volterra-System zweiter Ordnung mit einem Support von 16 Einheiten in einem $16^4 = 65536$ dimensionalen Parameterraum zu suchen sein. Aus diesem Grund wird hier eine indirekte Parametrierung durchgeführt, das bedeutet, daß

ein kanonisches System verwendet wird, bei dem der Parameterraum drastisch reduziert wird, also beispielsweise die Summe von n quadratischen Volterra-Systemen, die ihrerseits linear zerlegbar sind. Damit entsprechen die zu optimierenden Parameter genau den linearen Filterkernen. Eine weitere Reduktion kann erreicht werden, indem selbst die linearen Filterfunktionen parametrisiert werden. Somit können dann beispielsweise Bandpässe mit variabler Radialfrequenz, Orientierung und Bandbreite verwendet werden.

Ad 2. Die Zielfunktion (oder auch Kostenfunktion), die es gilt zu optimieren, sollte idealerweise eine monotone Abhängigkeit von den Parametern aufweisen und nur ein globales Extremum besitzen. Im Allgemeinen ist dies leider nicht der Fall, die Zielfunktion weist typischerweise eine „zerklüftete“ Struktur auf, mit vielen lokalen Extrema. Dies stellt das basale Problem für sämtliche Optimierungsalgorithmen dar, da diese nicht zwischen einem lokalen und dem globalen Extremum unterscheiden können. Möglichkeiten, diesen Fallen zu entkommen sind statistischer Art, beispielweise mehrfache Initialisierung des Algorithmus mit unterschiedlichen Startwerten oder Störung eines lokalen Extremums durch finite Abweichungen, um die Stabilität zu prüfen. Die Zielfunktion in dem hier betrachteten Optimierungsproblem soll ein Maß für die Flachheit des Ausgangsbispektrums $C_3^{U_2}(\vec{f}_1, \vec{f}_2)$ darstellen. Es wurden mehrere unterschiedliche Standardmaße verglichen, wobei sich gezeigt hat, daß die normierte Varianz des Bispektrums

$$\gamma = \frac{E\{(C_3^{U_2})^2\} - E^2\{C_3^{U_2}\}}{E^2\{C_3^{U_2}\}} \quad (5.127)$$

eine adäquate Wahl für die Zielfunktion darstellt. Hier ist allerdings zu beachten, daß dieses Maß eine *inverse* SFM darstellt, d.h. einem flach verlaufenden Bispektrum wird ein kleiner Wert zugeordnet. Wenn als exemplarische Bispektren Funktionen mit einem exponentiellen Abfall oder Gauß-Funktionen mit variabler Varianz betrachtet werden, so erkennt man einen monotonen Zusammenhang zwischen der Steilheit des Bispektrums und γ , siehe Abb. 5.29.

Ad 3. Die Wahl eines Optimierungsverfahrens hat hauptsächlich numerische Gründe. Prinzipiell muß man die Wahl treffen zwischen Verfahren, die lediglich die Zielfunktion evaluieren, und Verfahren, die zudem die partiellen Ableitungen der Zielfunktion evaluieren. Letztere sind im Allgemeinen etwas mächtiger hinsichtlich ihrer Konvergenzgeschwindigkeit, allerdings auf Kosten des numerischen Aufwands. Ist man in der Lage, die Ableitungen analytisch zu bestimmen, empfiehlt es sich in jedem Fall, ein ableitungsbasiertes Verfahren zu wählen. Klassische Vertreter hierbei sind die Gradienten- und konjugierte Gradientenverfahren (Press et al., 1992). Da dies im vorliegenden Fall nicht möglich ist, wird das sogenannte *Simplex*-Verfahren verwendet (Nelder und Mead, 1965). Dieses Verfahren minimiert die Zielfunktion, indem der direkte Weg „ins Tal“ gesucht wird, ohne dabei besondere Annahmen über die Funktion selbst zu machen. Es kann dadurch langsam konvergieren, ist aber tendenziell sehr robust und vor allem sehr

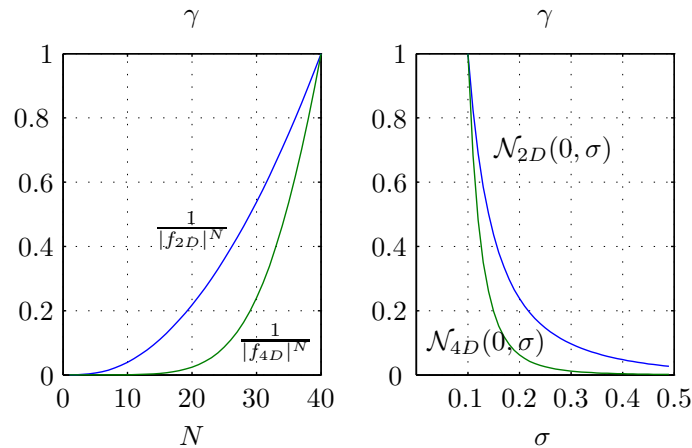


Abbildung 5.29: Abhängigkeit der inversen SFM γ aus Gl. (5.127) von der Steilheit zweier prototypischer Bispektren. Links für 2D- und 4D-Bispektrum mit variierendem exponentiellem Abfall, rechts für gaußförmiges 2D- und 4D-Bispektrum mit variierender Varianz.

kompakt. Der Speicheraufwand für ein N -dimensionales Optimierungsproblem ist in der Ordnung N^2 und es sind keine Ableitungen vonnöten.

Das Simplex-Verfahren

Ein Simplex ist eine geometrische Figur, bestehend aus, in N Dimensionen, $N + 1$ Eckpunkten und all ihren verbindenden Kanten und Polygonflächen. In zwei Dimensionen ergibt sich ein Dreieck, in drei ein Tetraeder. Im Allgemeinen betrachtet man nur nicht-degenerierte Simplexes, d.h. solche, die ein finites N -dimensionales Volumen einschließen. Wenn unter diesen Voraussetzungen ein Eckpunkt als Ursprung definiert wird, stellen damit die anderen Eckpunkte Richtungen im Raum dar, die den vollen Parameterraum aufspannen. Das Verfahren wird initialisiert mit einem (randomisierten oder deterministischen) Startwert \vec{p}_0 . Die anderen N Eckpunkte des Initial-Simplex ergeben sich dann aus der Vorschrift

$$\vec{p}_i = \vec{p}_0 + \lambda_i \vec{e}_i, \quad (5.128)$$

wobei die \vec{e}_i Einheitsvektoren des Parameterraums darstellen und die λ_i Konstanten, die charakteristisch für die Skalierung des jeweiligen Parameters sind. Der Algorithmus berechnet eine Reihe von Schritten, wobei typischerweise eine Reflexion vom höchsten Punkt durch die gegenüberliegende Fläche zum niedrigsten Punkt durchgeführt wird. Im Falle eines steilen Abstiegs wird die Reflexion zudem ausgedehnt (Expansion), um die Konvergenz zu beschleunigen. Wenn das Simplex ein Tal erreicht, kontrahiert es, d.h. die Schrittweite wird verringert. Im Falle eines „Nadelöhres“ kontrahiert es bezüglich aller Dimensionen um den niedrigsten Punkt. Diese grundlegenden Manöver sind in Abb. 5.30 zusammengefaßt. Die Abbruchbedingung ist erreicht, wenn die zurückgelegte relative vektorielle Distanz in einem Schritt eine gewisse Grenze unterschreitet.

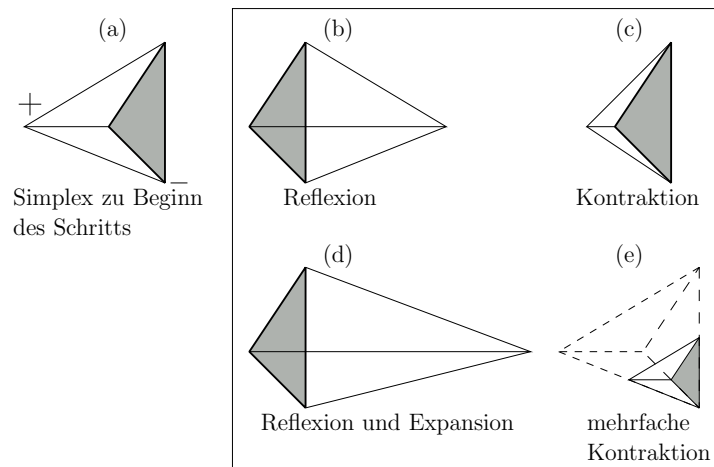


Abbildung 5.30: Beispiele für mögliche Schritte eines dreidimensionalen Optimierungsproblems unter Verwendung des Simplex-Verfahrens. (a) Simplex (hier ein Tetraeder) zu Beginn eines Schritts. (b-e) mögliche Transformationen: (b) Reflexion in Gegenrichtung vom höchsten Punkt (+), (c) Kontraktion entlang der Dimension bzgl. (+), (d) Reflexion und Expansion in Gegenrichtung von (+), (e) Kontraktion entlang sämtlicher Dimensionen hin zum niedrigsten Punkt (-). Eine geeignete Sequenz solcher Schritte wird immer zu einem Minimum einer Funktion konvergieren.

Modell und Parametrierung

Um zu einer möglichst verlässlichen Aussage zu kommen, wurden zweierlei Modelle untersucht. Zum Einen eine generelle $\Sigma\Pi$ -Architektur (siehe Gl. (5.109) und Abb. 5.18), bestehend aus der Summe von vier quadratischen Subsystemen, die ihrerseits linear zerlegbar sind, d.h. es sind acht unabhängige lineare Filter zu lernen. Die Übertragungsfunktionen dieser linearen Filter sind durch die polar separierbaren Filterfunktionen aus Gl. (4.16), allerdings mit \cos^2 -Abfall, gegeben. Die freien Parameter sind dann Amplitude, radiale Mittenfrequenz, radiale Bandbreite, Orientierung und Orientierungsbandbreite. Insgesamt sucht die Optimierung also in einem $8 \cdot 5 = 40$ -dimensionalen Parameterraum. Das Modell deckt eine relativ große Klasse ab, dies muß jedoch mit der hohen Zahl an Freiheitsgraden erkauft werden, weshalb sich die Optimierung dieses Systems als relativ komplexes Problem erweist. Im Speziellen hat sich empirisch gezeigt, daß die Zielfunktion viele lokale Minima aufweist, weshalb mit standardisierten Optimierungsalgorithmen kein zufriedenstellendes Ergebnis erreichbar ist. Um den Aufwand in vertretbaren Grenzen zu halten, wird das System nicht per Zufall initialisiert, sondern es werden feste Startwerte verwendet. Da gezeigt werden soll, daß das optimale System $i2D$ -Charakter besitzt, wird der Kern so initialisiert, daß die Durchlaßbereiche genau auf den verbotenen Zonen liegen. Die Aussagekraft des damit erreichbaren Ergebnisses (s.u.) ist natürlich eingeschränkt, daher wird zudem eine zweite Architektur optimiert, die einen stark reduzierten Parameterraum aufweist und somit für randomisierte Initialisierung geeignet ist.

Die zweite Architektur besteht aus der Summe von 18 nicht einfach separierbaren quadratischen Systemen, wobei das Wissen über die Isotropie der Statistik natürlicher Bilder ausgenutzt wird, und das System daher als Randbedingung Isotropie impliziert. Formal kann das System beschrieben werden als:

$$H_2(\vec{f}_1, \vec{f}_2; f_0, f_{hbw}, \phi_0, \phi_{hbw}) = \sum_{i=-9}^9 \left[H_2^{(p)}(\vec{f}_1, \vec{f}_2; f_0, f_{hbw}, \phi_0^{(i)}, \phi_{hbw}) - H_2^{(k)}(\vec{f}_1, \vec{f}_2; f_0, f_{hbw}, \phi_0^{(i)}, \phi_{hbw}) \right] \quad (5.129)$$

mit

$$H_2^{(p)}(\vec{f}_1, \vec{f}_2; f_0, f_{hbw}, \phi_0, \phi_{hbw}) = H_1^{(p)}(\vec{f}_1; f_0, f_{hbw}, 10i + \phi_{hbw}, \phi_{hbw}) \cdot H_1^{(p)}(\vec{f}_2; f_0, f_{hbw}, 10i - \phi_{hbw}, \phi_{hbw}) \quad (5.130)$$

und

$$H_2^{(k)}(\vec{f}_1, \vec{f}_2; f_0, f_{hbw}, \phi_0, \phi_{hbw}) = \frac{1}{4} \cdot H_1^{(k)}(\vec{f}_1; f_0, f_{hbw}, 10i, \frac{\phi_{hbw}}{2}) \cdot H_1^{(k)}(\vec{f}_2; f_0, f_{hbw}, 10i, \frac{\phi_{hbw}}{2}). \quad (5.131)$$

In jeder Orientierung besteht das System aus der Differenz eines quadratischen Primärfilters (p) und eines Kompensationsfilters (k), siehe Gl. (5.129). Das Primärsystem wiederum besteht aus dem Produkt zweier linearer Bandpaßfilter (Filterfunktionen entsprechend Gl. (4.16), aber wiederum mit \cos^2 -förmigem Abfall in Radial- und Orthogonalrichtung) mit unterschiedlicher Vorzugsrichtung $\Delta\phi_0 = 2\phi_{hbw}$, Gl. (5.130). Da die Übertragungsfunktionen erst nach $2\phi_{hbw}$ bzgl. der Vorzugsrichtung zu Null werden, überlappen sich die beiden Bandpässe in Orientierungsrichtung in einem Bereich von genau $2\phi_{hbw}$. Damit antwortet das Primärfilter also auch auf Stimuli mit nur einer Frequenzkomponente im Überlappungsbereich, ist also per Definitionem nicht $i2D$ -Selektiv. Um diese Selektivität erreichen zu können, bedarf es eines Kompensationsfilters (Zetzsche und Barth, 1990b), das mit halber Amplitude und Bandbreite in der Mitte der beiden Orientierungen liegt, Gl. (5.131). Für die speziellen \cos^2 -förmig abfallenden Orientierungsflanken der Primärfilter muß das Kompensationsfilter einen \cos -förmigen Flankenabfall aufweisen (Zetzsche und Barth, 1990b).

Da die Orientierungen durch die Isotropieannahme, bzw. die spezielle Kompensationsarchitektur bereits festgelegt sind, besitzt das Gesamtsystem lediglich drei freie Parameter: radiale Mittenfrequenz f_0 , radiale Bandbreite f_{hbw} , sowie Orientierungsbandbreite ϕ_{hbw} . Bleibt das System unterhalb einer gewissen Orientierungsbandbreite, so ist es durch die spezielle Architektur $i2D$ -selektiv; es kann allerdings auch leicht auf einen anderen Kern konvergieren.

Dieses zweite System stellt also bezüglich der Freiheitsgrade und Parameter das Gegenteil des ersten dar. Die Klasse der damit zu modellierenden Kerne ist wesentlich kleiner,

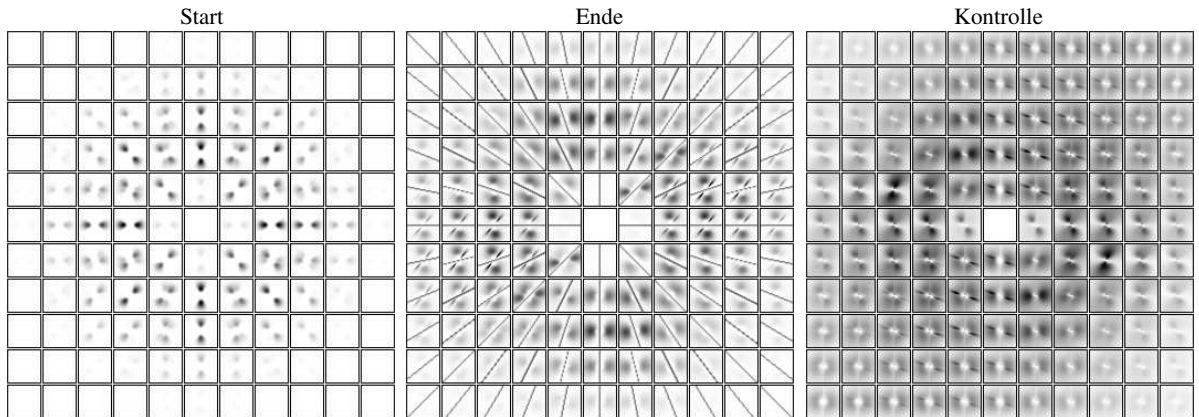


Abbildung 5.31: Der nicht- $i2D$ -selektiv initialisierte Kern (links) wird nach ca. 2500 Schritten des Simplex-Verfahrens $i2D$ -selektiv: die notwendige und hinreichende Bedingung $H_2(\vec{f}_1, \vec{f}_2)|_{f_{x_1} f_{y_2} = f_{x_2} f_{y_1}} \equiv 0$ ist erfüllt, wie an den leicht grau eingezeichneten verbotenen Zonen erkannt werden kann (Mitte). Wird zur Kontrolle weißes Rauschen als Eingang verwendet, ergibt sich keine erkennliche Struktur (rechts).

allerdings erlaubt die geringe Anzahl Freiheitsgrade, daß mit mehreren unterschiedlichen Startwerten initialisiert wird und das für diese Architektur *globale* Optimum gefunden werden kann.

Ergebnisse

Die Resultate für die volle Parametrierung des $\Sigma\Pi$ -Systems sind in Abb. 5.31 dargestellt. Das Initialsystem (links) hat die ungünstigste Lage für die Dekorrelation natürlicher Bilder da es voll auf den verbotenen Zonen $f_{x_1} f_{y_2} = f_{x_2} f_{y_1}$ liegt. Nach ca. 2500 Schritten des Simplex-Algorithmus konvergiert das Verfahren auf die ebenfalls in Abb. 5.31 gezeigte Lösung. Es ist deutlich zu erkennen, daß die notwendige $i2D$ -Bedingung erfüllt ist. Zudem ist eine Tendenz zu breitbandigen Filtern erkenntlich, die wie folgt gedeutet werden kann¹⁹

Die initialen linearen Übertragungsfunktionen $H_1^{(k)}(\vec{f})$ sind in dieser Untersuchung ad hoc so gewählt, daß diese den grundlegenden biologischen Anforderungen hinsichtlich Radial- und Orientierungsbandbreiten genügen²⁰. Läßt man nun aber die Hypothese zu, daß ein nichtlineares Subsystem $H_2^{(i)}(\vec{f}_1, \vec{f}_2)$ eine neuronale Einheit darstellt, die in klassischen Experimenten auf ein lineares Modell abgebildet wurde (was nach den Ausführungen in Abschnitt 2.7 und (Zetzsche und Nuding, 2005a) auch nicht beliebig unplausibel ist), so folgt daraus, daß die zugrundeliegenden „atomaren“ linearen Sy-

¹⁹Mit zugegebenermaßen hypothetischem Charakter.

²⁰Radialbandbreite von ca. 1-2 Oktaven und einseitige Orientierungsbandbreite von ca. 15-20 Grad, siehe auch Abb. 2.18 und Abschnitt 2.6.

steme breitbandigere Mechanismen aufweisen müssen. Dies kann leicht erklärt werden, wenn man bedenkt, daß $i2D$ -selektive Systeme konstruktiv aus nicht oder nur gering überlappenden orientierungsselektiven Bandpässen bestehen. Denn die Abbildung eines derartigen nichtlinearen Systems auf ein lineares Modell bei Messung mit beispielsweise Sinusgittern, resultiert im Produkt der linearen Teilkerne²¹, siehe hierzu (Zetzsche und Nuding, 2005a). Bei nur geringer Überlappung der Frequenzbänder „sieht“ die lineare Identifikation also einen schmalbandigen Mechanismus.

Fairerweise muß zu diesem Ergebnis gesagt werden, daß es nicht klar ist, ob hier das tatsächliche *globale* Optimum gefunden wurde. Aufgrund der Initialisierung und der schnellen Konvergenz davon weg, hin auf einen $i2D$ -selektiven Kern, liefert Evidenz, daß das Ergebnis durchaus relevant ist. Die Minimalaussage, also die *Worst-Case*-Betrachtung sagt allerdings lediglich, daß eine $i2D$ -selektive Struktur für die Dekorrelation höherer Ordnung natürlicher Bilder *besser* geeignet ist, als eine dazu komplementäre. Dies ist zugegebenermaßen nur beschränkt aussagekräftig, und daher rührt auch die Motivation zur Optimierung eines Systems mit mehr Randbedingungen (s.o.), für welches das globale Optimum gefunden werden kann.

Die Ergebnisse dieser Optimierung zeigt Abb. 5.32. Exemplarisch sind vier deutlich unterschiedliche zufällige Initialisierungen mit den jeweils konvergenten Endergebnissen abgebildet. Es ist deutlich zu erkennen, daß die optimierten Kerne die gleiche Struktur haben, kleinere Unterschiede, v.a. in der Bandbreite sind zurückzuführen auf das Abbruchkriterium, das auf Grund der Rechenzeit größer gewählt wurde, als die Maschinenpräzision, was im Limit die besten Ergebnisse liefert.

Betrachtet man nun die Ergebnisse gemeinsam, so erhält man doch eine starke Evidenz dafür, daß eine $i2D$ -selektive Filterung in der Tat optimal geeignet ist, um die statistischen Bindungen höherer Ordnung zu eliminieren. Das erste Ergebnis zeigt, daß eine Architektur mit vielen Freiheitsgraden, die in der Lage ist, eine große Klasse quadratischer Volterra-Systeme zu approximieren, besser geeignet ist, Korrelationen höherer Ordnung in natürlichen Bildern zu eliminieren, als eine komplementäre Struktur. Das zweite Ergebnis weist nach, daß eine eingeschränkte aber nicht notwendigerweise $i2D$ -selektive Architektur im Rahmen ihrer Möglichkeiten stets auf eine $i2D$ -selektive Lösung konvergiert. Selbstverständlich geht hier auch kritisch das Optimierungskriterium aus Gl. (5.127) mit ein, das zwar heuristisch gewählt wurde, aber für generische Funktionen einen stetigen und monotonen Verlauf aufweist und somit eine vernünftige Wahl darstellt. Damit ist gezeigt, oder zumindest Evidenz dafür erbracht, daß die propagierten UND-Kombinationen von Frequenzkomponenten, die ja eine Voraussetzung für die $i2D$ -Selektivität sind, nicht nur für die Modellierung und Synthese nichtlinearer kortikaler Effekte geeignet ist, sondern auch zur Eliminierung, bzw. Reduzierung von statistischen Abhängigkeiten höherer Ordnung in natürlichen Bildern.

²¹Die inhärente Gleichrichtung in den neuronalen Antworten führt dazu, daß eine Frequenzverdopplung leicht übersehen werden kann, da die Grundschwingung erhalten bleibt.

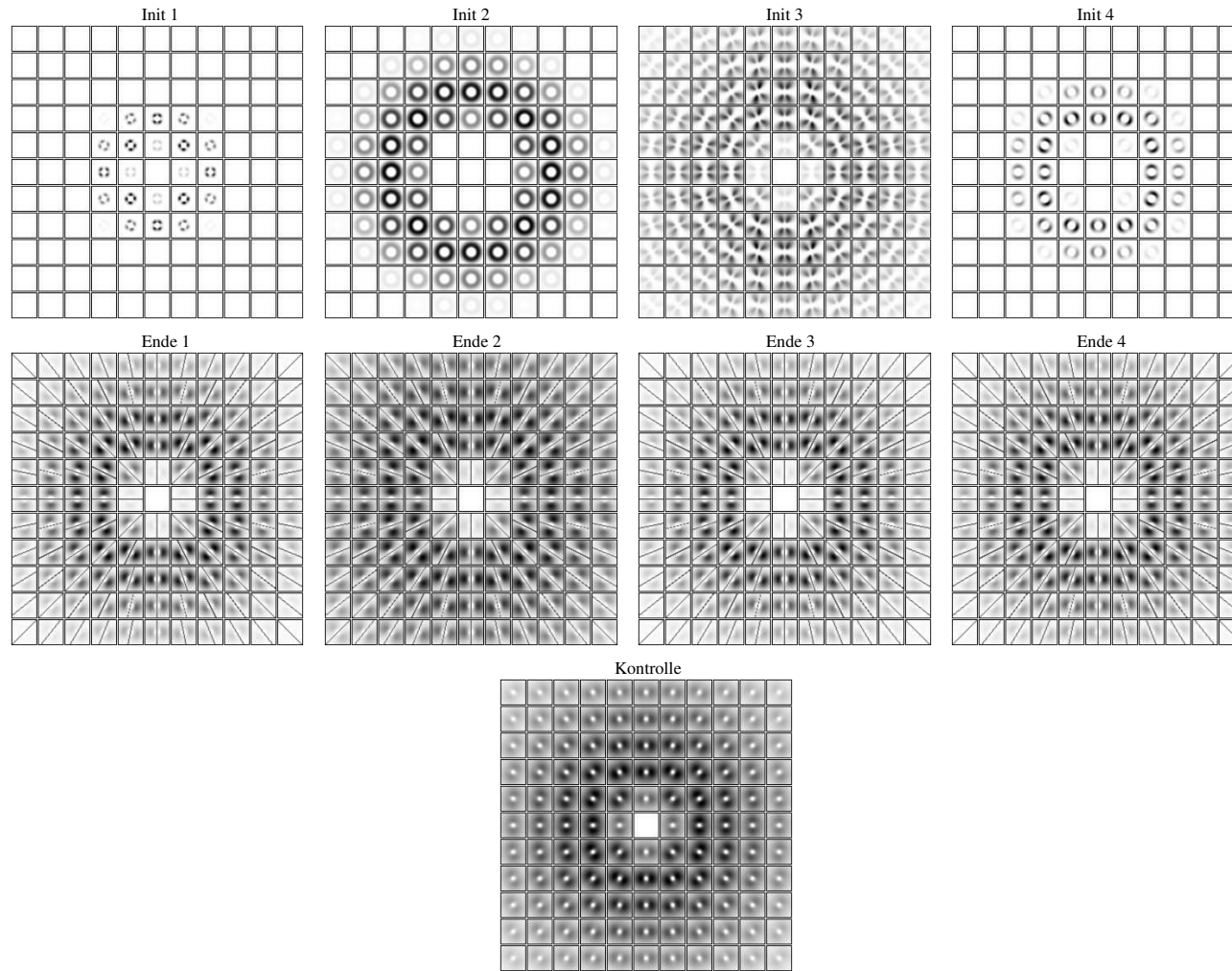


Abbildung 5.32: Optimierte Kerne zur Dekorrelation natürlicher Bilder des quadratischen Volterra-Systems aus den Gln. (5.129) – (5.131). Gezeigt sind die Initialkerne und die korrespondierenden gelernten Kerne für vier exemplarische zufällige Initialisierungen. Wird ein Gaußverteiltes Rauschen als Eingang verwendet, ergibt sich bei keiner Initialisierung eine dezidierte Struktur (Kontrolle).

5.3 Grenzen des quadratischen Ansatzes

Wie in den vorangegangenen Abschnitten gezeigt wurde, eignen sich Spektren und Systeme zweiter Ordnung (Bispektren und quadratische Volterra-Systeme) zur Analyse, Synthese und Modellierung von natürlichen Bildern und ihrer assoziierten optimalen Verarbeitung im visuellen System. Nichtsdestotrotz unterliegt natürlich auch dieser Ansatz gewissen Einschränkungen, die bei der Anwendung dieser Verfahren nicht außer Acht gelassen werden dürfen. Primär sind dies natürlich die Einschränkungen aufgrund der endlichen Ordnung der betrachteten Bindungen und Polynome. So sind Kopplungen mehrerer Frequenzkomponenten (≥ 3) mit quadratischen Systemen nicht analysierbar oder modellierbar, wie den Gl. (5.32, 5.98) direkt zu entnehmen ist. Speziell für Bispektren gilt zudem die fundamentale Eigenschaft, daß bei einer geradesymmetrischen Verteilung der Amplituden des analysierten Prozesses das Bispektrum verschwindet, da im Erwartungswert die Tripel-Produkte (Kumulanten) zu Null werden. In (Krieger, 1999) wurde dies anhand des Beispiels des *Konturrauschens* verdeutlicht und gezeigt, daß das Trispektrum die orientierten Bildmerkmale erfassen kann, wohingegen das Bispektrum aufgrund der Symmetrie der Verteilung verschwindet.

Ein besonderes Merkmal dieses Ansatzes ist, daß keine direkte Assoziation der analysierten Bindungen mit den multiplikativen Kopplungen der Volterra-Reihenentwicklung möglich ist. So *manipuliert* zum Beispiel ein Volterra-System der Ordnung p unendlich viele Kumulanten höherer Ordnung²² und im Allgemeinen auch niedrigerer Ordnung. Andererseits jedoch *generiert* ein derartiges System Phaseninteraktionen der Ordnung $p + 1$, aber im Allgemeinen auch niedrigerer Ordnung. Das bedeutet, daß die klassische Analyse der Bindungen höherer Ordnung keinen Rückschluß auf den Grad der multiplikativen Interaktionen zuläßt.

Einen Ansatz zur Überwindung dieses Problems wurde in (Franz und Schölkopf, 2004) vorgeschlagen. Dieser Ansatz besteht prinzipiell in der Ermittlung eines Wiener-Funktions zur (nichtlinearen) Prädiktion eines Bildpunktes in einer lokalen Umgebung. Die gelernten Kerne werden zur Analyse der multiplikativen Interaktionen von benachbarten Bildpunkten verwendet. Der Vorteil und die Hauptmotivation bei diesem Ansatz bestehen darin, daß die Wiener-Kerne *implizit* unter Verwendung des sogenannten „Kernel-Tricks“ gelernt werden können (Schölkopf und Smola, 2002). Da damit die Komplexität der Schätzung im Vergleich zu den klassischen Statistiken höherer Ordnung stark reduziert werden kann, kann der Grad der geschätzten multiplikativen Interaktionen erhöht werden. Damit wurde gezeigt, daß natürliche Bilder bereits durch multiplikative Interaktionen bis zur fünften Ordnung recht gut beschrieben werden können.

Eng verwandt mit dieser Thematik ist auch die gezielte Manipulation der Spektren höherer Ordnung mit Volterra-Systemen. Während für lineare Systeme und statisti-

²²Dies gilt bereits für ein lineares System, vgl Gl. (5.38).

sche Bindungen zweiter Ordnung eine eindeutige Abbildungsvorschrift existiert, ist dies für höhere Ordnungen nicht der Fall. So ist beispielsweise das Leistungsdichtespektrum $C_2^{\mathbf{u}_2}(f)$ des Ausgangs eines eindimensionalen linearen Systems $H(f)$ gegeben durch

$$C_2^{\mathbf{u}_2}(f) = |H(f)|^2 C_2^{\mathbf{u}_1}(f), \quad (5.132)$$

wenn $C_2^{\mathbf{u}_1}(f)$ das Eingangsleistungsdichtespektrum darstellt. Eine wünschenswerte Eigenschaft wäre nun, daß ähnliche geschlossene Relationen auch für Kerne und Spektren höherer Ordnung existierten. Im Folgenden wird hergeleitet, daß dies leider nicht der Fall ist.

Während sich die Berechnung der Spektren höherer Ordnung am Ausgang eines linearen Systems noch recht einfach gestaltet, siehe Gl. (5.38), so ergibt sich bereits für das Ausgangs-LDS eines quadratischen Volterra-Systems ein etwas unhandlicher Ausdruck. Ausgehend von der Frequenzbereichs-Berechnung des Ausgangs, Gl. (5.90), und unter Verwendung der Definition des Bispektrums, Gl. (5.32), läßt sich für das Ausgangs-LDS $C_2^{\mathbf{u}_2}(f)$ schreiben:

$$\begin{aligned} C_2^{\mathbf{u}_2}(f) &= E\{\mathbf{U}_2(f)\mathbf{U}_2^*(f)\} \\ &= E\left\{ \int H_2(f_a, f - f_a) \cdot \mathbf{U}_1(f_a)\mathbf{U}_1(f - f_a) \cdot df_a \right. \\ &\quad \left. \cdot \int H_2^*(f_b, f - f_b) \cdot \mathbf{U}_1^*(f_b)\mathbf{U}_1^*(f - f_b) \cdot df_b \right\}. \end{aligned} \quad (5.133)$$

Da in Gl. (5.133) bereits unterschiedliche Integrationsvariablen verwendet sind und die Integration eine lineare Operation ist, können die Integrale zusammengefaßt und der Erwartungswert in die Integrale hineingezogen werden:

$$\begin{aligned} C_2^{\mathbf{u}_2}(f) &= \int \int H_2(f_a, f - f_a)H_2^*(f_b, f - f_b) \cdot \\ &\quad \cdot E\left\{ \mathbf{U}_1(f_a)\mathbf{U}_1(f - f_a)\mathbf{U}_1^*(f_b)\mathbf{U}_1^*(f - f_b) \right\} \cdot df_a df_b. \end{aligned} \quad (5.134)$$

Interessant ist hierbei, daß der Ausdruck im Erwartungswert nun das *Trispektrum* des Eingangsprozesses darstellt. Dieses ist ja entsprechend Gl. (5.32) definiert als der Erwartungswert über die Summe vierer komplexer Amplituden, wobei deren Frequenzkomponenten in der Summe Null ergeben. Man kann leicht sehen, daß dies in Gl. (5.134) der Fall ist, daher hängt also das Ausgangs-LDS eines quadratischen Volterra-Systems von definierten Projektionen des Eingangs-Trispektrums ab. Dies ist ein relativ komplexer Zusammenhang, der sich vereinfachen läßt, wenn als Eingang weißes gaußsches Rauschen angenommen wird. Der naheliegende Gedanke, daß die Spektren höherer Ordnung für gaußsche Signale identisch Null sind, und damit das Ausgangs-LDS ebenfalls identisch Null sein muß, führt leicht in die Irre. In der Tat ist bekannt, daß gewisse infinitesimale Bereiche in der Definition der Spektren eine besondere Stellung haben. Genauer

gesagt sind dies die Punkte, bei denen sich jeweils die Phasen von zwei Komponenten auslöschen, d.h. es müssen alle Frequenzkomponenten zu Paaren mit unterschiedlichem Vorzeichen zusammenfaßbar sein; dann ergibt sich (für reelle Signale) das Betragsquadrat der komplexen Amplitude:

$$E\{\mathbf{U}(f_1)\mathbf{U}(f_2)\mathbf{U}(f_3)\mathbf{U}(f_4)\} \Big|_{f_3=-f_1; f_4=-f_2} = E\{|\mathbf{U}(f_1)|^2|\mathbf{U}(f_2)|^2\}. \quad (5.135)$$

Diese werden jedoch bei der diskreten Berechnung in praxi ausgeschlossen. Bei einem Projektionsintegral der Art aus Gl. (5.134) müssen diese dezidierten Punkte allerdings mit in Betracht gezogen werden.

In dem momentan betrachteten Problem ist also das Trispektrum des gaußschen Eingangsprozesses in dem Projektionsintegral nicht vernachlässigbar, wenn $f_b = f_a$ ist. Das Integral kann somit wie folgt vereinfacht werden:

$$C_2^{\mathbf{u}_2}(f) = \int |H_2(f_a, f - f_a)|^2 \cdot E\{|\mathbf{U}_1(f_a)|^2|\mathbf{U}_1(f - f_a)|^2\} \cdot df_a. \quad (5.136)$$

Aufgrund der Fourier-Stieltjes-Repräsentation stochastischer Prozesse sind die komplexen Amplituden für unterschiedliche Frequenzkomponenten unkorreliert und somit kann der Erwartungswert mit

$$E\{|\mathbf{U}_1(f_a)|^2|\mathbf{U}_1(f - f_a)|^2\} = E\{|\mathbf{U}_1(f_a)|^2\}E\{|\mathbf{U}_1(f - f_a)|^2\} = \gamma_2^2 \quad (5.137)$$

angegeben werden, wenn γ_2 die konstante Leistungsdichte des gaußschen Eingangsprozesses beschreibt. Damit ergibt sich also schlußendlich der Ausdruck für das Leistungsdichtespektrum eines weißen gaußschen Prozesses nach dem Passieren eines quadratischen Volterra-Systems:

$$C_2^{\mathbf{u}_2}(f) = \gamma_2^2 \int |H_2(f_a, f - f_a)|^2 \cdot df_a. \quad (5.138)$$

Analog zur Projektion des expandierten Ausgangsspektrums, wie in Gl. (5.90) angegeben, stellt dies die Projektion des Betragsquadrats des Kerns auf die Hauptdiagonale dar. Damit ist auch klar, daß aus einem weißen Rauschen ein Prozeß mit beliebiger Leistungsdichte generiert werden kann, da hierfür nur ein Kern konstruiert werden muß, der eben diesen LDS-Verlauf als Projektion auf die Hauptdiagonale besitzt.

Etwas anders ist die Lage, wenn man die Ordnung des betrachteten Spektrums erhöht. Wird wiederum von einem homogenen quadratischen System ausgegangen, so kann mit einer völlig analogen Rechnung zu oben gezeigt werden, daß für das Ausgangsbispektrum gilt:

$$C_3^{\mathbf{u}_2}(f_1, f_2) = \int \int \int H_2(f_a, f_1 - f_a)H_2(f_b, f_2 - f_b)H_2^*(f_c, f_1 + f_2 - f_c) \cdot C_6^{\mathbf{u}_1}(f_a, f_1 - f_a, f_b, f_2 - f_b, -f_c) \cdot df_a df_b df_c. \quad (5.139)$$

Ist der Eingangsprozeß wie oben weiß und gaußsch, so vereinfacht sich das Integral wiederum, wenn die dezidierten Stellen $f_c = f_a$ und $f_b = f_a - f_1$ berücksichtigt werden, zu:

$$C_3^{\text{u2}}(f_1, f_2) = \gamma_2^3 \int H_2(f_a, f_1 - f_a) H_2(f_a - f_1, f_1 + f_2 - f_a) H_2(-f_a, f_a - f_1 - f_2) \cdot df_a. \quad (5.140)$$

Ein wesentliches Ergebnis dieser Herleitung ist, daß es im Allgemeinen mit einem homogenen quadratischen System nicht möglich ist, aus einem WGN-Prozeß ein Bisppektrum mit beliebigen Eigenschaften zu formen. Dies hat sich experimentell bereits früh in mehreren erfolglosen Versuchen bestätigt und kann mit diesem Ergebnis klar gemacht werden.

Zusammenfassend kann man sagen, daß die Ansätze mit Spektren und Systemen höherer Ordnung aufgrund ihres approximativen Ansatzes gewissen Grenzen unterliegen. Ein wesentlicher weiterer Punkt, der Beachtung verdient, ist, daß die einfachen Zusammenhänge zwischen statistischen Abhängigkeiten und System-Übertragungsfunktionen aus der linearen Systemtheorie nicht direkt auf statistische Abhängigkeiten und Systeme *höherer* Ordnung übertragbar sind. Eine Hauptursache hierfür ist, daß statistische Abhängigkeiten höherer Ordnung, ausgedrückt durch Kumulanten, nicht direkt auf multiplikative „Pixelinteraktionen“, welche ein Volterra-System höherer Ordnung erzeugt, abbildbar sind. Nichtsdestotrotz bietet die Frequenzbereichsanalyse der statistischen Abhängigkeiten höherer Ordnung eine elegante Möglichkeit, multiplikative Interaktionen der Frequenzen in Form von Phasenkopplungen zu *detektieren*. Der grundsätzliche Effekt der UND-Verknüpfung von Frequenzkomponenten ist bereits mit dem Bisppektrum erfaßbar, bzw. mit einem quadratischen System modellierbar. Wie in diesem Kapitel ausführlich gezeigt wurde, reichen diese Ordnungen für die qualitative Modellierung von bekannten kortikalen Phänomenen bereits aus.

6 Diskussion

Was wir wissen, ist ein Tropfen, was wir nicht wissen, ein Ozean.

Sir Isaac Newton

In dieser Arbeit konnte ich zeigen, wie das Ineinandergreifen verschiedener Ansätze zum Verständnis der visuellen Informationsverarbeitung beitragen kann. Die in Kapitel 3 eingeführte Signal-, bzw. Zustandsraumdarstellung ermöglicht prinzipiell das volle Erfassen von statistischen Abhängigkeiten in natürlichen Signalquellen und eine damit verbundene optimale Signalverarbeitung. Effekte wie Selektivität, Invarianz, Verstärkungsregelung und extra-klassische Effekte sind hiermit im Prinzip leicht zu verstehen und zu modellieren. Das immanente Problem dieses Ansatzes besteht jedoch in der Dimensionalität der Signale und Systeme, weshalb die Betrachtung immer auf sehr niedrigdimensionale Fälle reduziert bleiben muß. Auf Bildstatistiken bezogen bedeutet dies, daß lediglich zwei- oder dreidimensionale Verbundstatistiken betrachtet werden können und diese damit nicht wesentlich mehr Information beinhalten als die Erkenntnis, daß die Korrelation zwischen benachbarten Bildpunkten abnimmt. Nichtsdestoweniger können durch die Einführung einer verallgemeinerten Beschreibung die grundlegenden Verarbeitungsprinzipien neuronaler Systeme erfaßt werden. Dies wurde erstmalig von WEGMANN und ZETZSCHE als *multivariate Wavelet-Statistik* eingeführt (Wegmann und Zetzsche, 1990b) und in Kapitel 3 sowie (Zetzsche und Nuding, 2005b) weiter ausgearbeitet.

Das Lernen in neuronalen Netzen hingegen erlaubt die Evaluierung von Optimalitätskriterien. Dies motivierte die mehrstufige neuronale Architektur aus Kapitel 4. Vorarbeiten untersuchten die Transformation statistischer Abhängigkeiten in zweischichtigen Systemen, die an die Statistik natürlicher Bilder angepaßt wurden (Zetzsche und Röhrbein, 2001; Nuding, 2002). In jüngerer Zeit wurden vermehrt mehrschichtige Feed-Forward Netze verwendet, um beispielsweise invariante Objekterkennung zu ermöglichen (Riesenhuber und Poggio, 1999), oder basale Eigenschaften von kortikalen Zellen zu lernen (Hyvärinen und Hoyer, 2001). Eine hervorragende Eigenschaft des in dieser Arbeit vorgestellten und implementierten Systems ist, daß die Gewichte *nicht* durch Lernen eines gewünschten Antwortverhaltens bestimmt wurden, sondern aus der bloßen Optimierung der Statistik natürlicher Bilder. Zudem ist erwähnenswert, daß es sich um eine recht einfache Architektur mit wenigen Einschränkungen handelt. So ist die Klasse der linearen Transformationen als gültige Lösung enthalten, die bei einer entsprechenden Wahl der Koeffizienten zustande kommt. Dies ist bei ähnlichen Arbeiten auf diesem Gebiet

nicht der Fall, hier geht man *a priori* von bestimmten nichtlinearen Verschaltungen aus (Hyvärinen und Hoyer, 2001; Hyvärinen et al., 2005). Ich konnte zeigen, daß diese Architektur einer rein linearen Verarbeitung nicht nur bezüglich der Reduzierung der statistischen Abhängigkeiten in natürlichen Bildern überlegen ist, sondern daß auch Vorteile bei der Kodierung bestehen. Wesentlich für die beispielsweise in (Zetzsche und Krieger, 2001b; Zetzsche und Nuding, 2005b) und an anderen Stellen dieser Arbeit postulierte grundlegende Relevanz von UND-artigen Verschaltungen von Frequenzkomponenten ist auch das Auftreten von Frequenzinteraktionen, die ebenfalls in (Hyvärinen et al., 2005) gefunden wurden. Dort unterlag die Optimierung allerdings auch den oben erwähnten Einschränkungen und es wurden lediglich „Pooling-Measures“ verwendet, die von den wahren Frequenzinteraktionen abweichen, wie in Abschnitt 4.5 dargelegt. Zudem konnte ich nachweisen, daß substantielle Interaktionen auch zwischen Frequenzen unterschiedlicher *Orientierung* entstehen, eine notwendige Bedingung für eine Selektivität bezüglich der nichtredundanten Signalanteile in natürlichen Bildern.

Die Optimierung von neuronalen Netzen zur Kodierung natürlicher Bilder gewährt zwar einerseits die Möglichkeit zur Evaluierung der Architektur und des postulierten Optimierungskriteriums, andererseits kann dieser Ansatz jedoch keine Einblicke in die neuronale Verarbeitung im Sinne einfacher funktionaler Modelle bieten. Dies wird ermöglicht durch die Anwendung von polynomialen Ansätzen auf die neuronale Informationsverarbeitung. Die Analyse der statistischen Bindungen höherer Ordnung in natürlichen Bildern mittels Polyspektren wurde bereits von ZETZSCHE und Mitarbeitern mehrfach zur Klärung der Funktionsweise des visuellen Systems und der Blickbewegungssteuerung verwendet (Zetzsche et al., 1998; Krieger, 1999). Diese Untersuchungen wurden in dieser Arbeit konsequent weitergeführt, indem ich einerseits die grundlegenden Eigenschaften des Bisppektrums und den Zusammenhang mit nichtlinearen Systemen ausgearbeitet habe, und indem andererseits die bispektralen Eigenschaften natürlicher Bilder weiter quantifiziert wurden.

Die Relevanz von UND-artigen Verknüpfungen von Frequenzkomponenten für die visuelle Informationsverarbeitung kann mit Hilfe der Volterra-Wiener Systemtheorie erklärt werden. Ich konnte zeigen, daß aktuelle neurophysiologische Ergebnisse, die bislang keine konsistente Erklärung fanden, bereits mit quadratischen Volterra-Systemen leicht konstruktiv synthetisierbar sind und damit im Sinne funktionaler Modelle einfach interpretiert werden können. Die Frage nach der optimalen Signalverarbeitung für natürliche Bilder kann allerdings auch hiermit nicht vollständig zufriedenstellend gelöst werden. Ein wesentlicher Schritt zur Klärung ist jedoch der Nachweis, daß *i2D*-selektive Systeme in der Tat zur Dekorrelation höherer Ordnung notwendig sind. Dies geht einerseits aus den oben erwähnten Frequenzinteraktionen der gelernten neuronalen Architektur hervor, andererseits aber auch aus der expliziten Dekorrelation des Bisppektrums natürlicher Bilder mittels quadratischen Volterra-Systemen (Abschnitt 5.2.6).

Die Analyse physiologischer Systeme mit Volterra-Reihen ist nicht neu, sondern wurde

in vielen peripheren sensorischen Systemen bereits angewandt (Marmarelis und Marmarelis, 1978). Neurone in den höheren visuellen Arealen von Primaten haben diesbezüglich bislang nur wenig Aufmerksamkeit erfahren, was zweierlei Ursachen hat. Einerseits antworten Neurone nur schwach auf die klassischen WGN-Stimuli, was zu einer schlechten Schätzung der Kerne höherer Ordnung führt. Andererseits fehlen ohne einen konzeptuellen Rahmen die Interpretationsmöglichkeiten für die gemessenen Ergebnisse, da eine reine Reproduktion des Antwortverhaltens aufgrund der schwierigen Schätzung meist nur mangelhaft möglich ist. Durch die Weiterentwicklung verschiedener Meßverfahren konnte ich zeigen, daß eine Schätzung der Kerne unter biologisch realistischen Bedingungen insofern möglich ist, als daß relevante Strukturen wie beispielsweise Frequenzinteraktionen und *i2D*-Selektivität identifiziert werden können. Dies wird möglich durch die Frequenzbereichsanalyse, wie ich ausführlich in den Abschnitten 5.2.3 und 5.2.4 herausgearbeitet habe. Die Weiterentwicklung der Meßverfahren hin zu natürlichen Stimuli führt zudem vermutlich zu einem weniger spärlichen Antwortverhalten der kortikalen Neurone und somit zu einem besseren S/N der Messung.

Ein wesentlicher Aspekt, der allen Ansätzen gemein ist, ist die Abstraktion von realen Neuronen, d.h. Kodierung durch Aktionspotentialstöße. Es bleibt offen, *wie* die entsprechenden Antwortverhalten in einer Neuronenpopulation kodiert sind. Es ist mittlerweile akzeptiert, daß einfache Feuerraten-Modelle für viele biologische Systeme nicht ausreichend realistisch sind. Ein häufiger Einwand, den man bei einer derartigen Modellierung erfährt, ist, daß wesentliche Informationsanteile durch das präzise *zeitliche* Auftreten der neuronalen Aktionspotentiale und durch dynamische Oszillationen in rekurrenten Netzwerken übertragen wird. Diese Art der neuronalen Kodierung wurde in dieser Arbeit nicht adressiert, da hier die *funktionalen* Aspekte der Informationsverarbeitung in Form von Basisselektivitäten, bzw. Antwortcharakteristika im Mittelpunkt standen. Analytische Betrachtungen der Verarbeitung in *Spiking Neural Networks* ist meistens nur unter Verwendung von starken Vereinfachungen möglich, die dann selbst wieder in Frage zu stellen sind. Aus diesem Grund wurde für diese Arbeit eine relativ abstrakte Vorgehensweise gewählt, da die behandelte Signalverarbeitung sehr komplex ist und grundlegende funktionale Eigenschaften noch nicht verstanden sind.

Spätestens seit den Pionierarbeiten von ATTNEAVE und BARLOW ist klar, daß die effiziente Verarbeitung visueller Information nicht ohne Kenntnis der strukturellen Eigenschaften unserer Umwelt möglich ist. Obwohl hier besonderer Wert auf die Eliminierung von statistischen Abhängigkeiten gelegt wird, so bedeutet dies nicht, daß das einzige Ziel einer sensorischen Verarbeitung die bestmögliche Quellenkodierung im Sinne einer *sparsamen* Repräsentation ist. BARLOW betont, daß ein derartig reduktionistischer Ansatz biologisch irrelevant und unplausibel ist (Barlow, 2001b). Darauf habe ich bereits ausführlich in Kapitel 2 hingewiesen. Die hier vorgestellten Mechanismen dienen hauptsächlich dem *Erfassen* von statistischen Abhängigkeiten und der *Transformation* dieser in eine dem Organismus zugänglichen Art und Weise.

7 Literaturverzeichnis

- Abramowitz, M. und Stegun, I. A., Herausgeber (1970). *Handbook of Mathematical Functions*. Dover Publications, Inc., New York, 9. Auflage.
- Atick, J. J. (1992). Could information theory provide an ecological theory of sensory processing? *Network: Computation in Neural Systems*, 3:231–251.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, 61:183–193.
- Baddeley, R. (1996). An efficient code in V1? *Nature*, 381:560–561.
- Baddeley, R., Abbott, L. F., Booth, M. C., Sengpiel, F., Freeman, T., Wakeman, E. A. und Rolls, E. T. (1998). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. R. Soc. London Ser. B*, 264(1389):1775–1783.
- Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. Rosenblith, Herausgeber, *Sensory Communication*, Seiten 217–234. MIT Press, Cambridge, MA.
- Barlow, H. B. (1972). Single units and sensation: a neuron doctrine for perceptual psychology? *Perception*, 1:371–394.
- Barlow, H. B. (2001a). The exploitation of regularities in the environment by the brain. *Behav. Brain Sci.*, 24(4):602–607.
- Barlow, H. B. (2001b). Redundancy reduction revisited. *Network: Computation in Neural Systems*, 12:241–253.
- Barnard, G. A. (1955). Statistical calculations of word entropies for four western languages. *IRE Trans. Inform. Theory*, IT-1:49–53.
- Bartelt, H., Lohmann, A. W. und Wirtzinger, B. (1984). Phase and amplitude recovery from bispectra. *Appl. Opt.*, 23(18):3121–3129.
- Bell, A. J. und Sejnowski, T. J. (1997). The 'independent components' of natural images are edge filters. *Vision Research*, 37(23):3327–3338.

- Berger, T. (1979). *Rate Distortion Theory*. Prentice-Hall, Englewood Cliffs, NJ.
- Berkes, P. und Wiskott, L. (2004). Slow feature analysis yields a rich repertoire of complex-cells properties. In *Proc. Early Cognitive Vision Workshop, Isle Of Skye, Scotland, May 28–June 1*.
- Blakemore, C. und Campbell, F. W. (1969). On the existence of neurones in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology*, 203:237–260.
- Blakemore, C. und Sutton, P. (1969). Size adaptation: a new aftereffect. *Science*, 166:245–247.
- Boring, E. (1942). *Sensation and Perception in the History of Experimental Psychology*. D. Appleton-Century Co., New York.
- Brillinger, D. R. (1965). An introduction to polyspectra. *Ann. Math. Statist.*, 36:1351–1374.
- Brillinger, D. R. (2002). John W. Tukey’s work on time series and spectrum analysis. *The Annals of Statistics*, 30(6):1595–1618.
- Campbell, F. W. und Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 197:551–556.
- Carter, B. E. und Henning, G. B. (1971). Detection of gratings in narrow-band visual noise. *Journal of Physiology*, 219:355–365.
- Chaitin, G. (1987). *Algorithmic Information Theory*. Cambridge Univ. Press, Cambridge.
- Comon, P. (1994). Independent Component Analysis – a new concept? *Signal Processing*, 36:287–314.
- Cosman, P. C., Gray, R. M. und Vetterli, M. (1996). Vector quantization of image subbands: a survey. *IEEE Trans. on Image Processing*, 5(2):202–225.
- Dan, Y., Atick, J. J. und Reid, R. C. (1996). Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. *Journal of Neuroscience*, 16(10):3351–3362.
- DeValois, K. K. und Tootell, R. B. H. (1983). Spatial-frequency-specific inhibition in cat striate cortex cells. *Journal of Physiology*, 336:359–376.
- DeValois, R. L. und DeValois, K. K. (1990). *Spatial Vision*. Oxford University Press, New York.
- Diamantaras, K. I. und Kung, S. Y. (1996). *Principal Component Neural Networks*. John Wiley & Sohns, Inc, New York.

- Dowling, J. E. (1987). *The Retina: An Approachable Part of the Brain*. Harvard University Press, Cambridge, MA.
- Emerson, R. C., Citron, M. C., Vaughn, W. J. und Klein, S. A. (1987). Nonlinear directionally selective subunits in complex cells of cat striate cortex. *Journal of Neurophysiology*, 58(1):33–65.
- Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A*, 4(12):2379–2394.
- Field, D. J. (1999). Wavelets, vision and the statistics of natural scenes. *Philosophical Transactions of the Royal Society: Mathematical, Physical and Engineering Sciences*, Seiten 2527–2542.
- Fiori, S. (2003). Overview of independent component analysis with an application to synthetic aperture radar (sar) imagery processing. *Neural Networks*, 16:453–467.
- Fishman, R. S. (1973). Kepler’s discovery of the retinal image. *Arch Ophthalmol.*, 89(1):59–61.
- Foldiak, P. (1995). Sparse coding in the primate cortex. In M. A. Arbib, Herausgeber, *The Handbook of Brain Theory and Neural Networks*, Seiten 895–989. MIT Press, Cambridge, MA.
- Franz, M. O. und Schölkopf, B. (2004). Implicit wiener series for higher-order image analysis. In *Proc. Neural Information Processing Systems (NIPS 2004)*.
- Fréchet, M. (1910). Sur les fonctionelles continues. *Annales Scientifiques de L’ecole Normale Supérieure*, 27:193–216.
- Gabor, D. (1946). Theory of communication. *J. IEE*, 93:429–457.
- Gallant, J. L., Connor, C. E., Rakshit, S., Lewis, J. W. und Van Essen, D. C. (1996). Neural responses to polar, hyperbolic, and Cartesian gratings in area V4 of the macaque monkey. *Journal of Neurophysiology*, 76:2718–2739.
- Gantmacher, F. R. (1960). *The Theory of Matrices*. Chelsea, New York.
- Gardner-Medwin, A. R. und Barlow, H. B. (2001). The limits of counting accuracy in distributed neural representations. *Neural Computation*, 13(3):477–504.
- Goethe, J. W. (1953). *Farbenlehre*. Kohlhammer, Stuttgart. Erstausgabe 1810.
- Gottschalk, A., Sexton, M. G. und Roschke, G. (2004). Multiplicative neural noise can favor an independent components representation of sensory input. *Network: Computation in Neural Systems*, 15(4):291–311.
- Graham, N. und Nachmias, J. (1971). Detection of grating patterns containing two

- spatial frequencies: A comparison of single channel multichannel models. *Vision Research*, 11:251–259.
- van Hateren, J. H. und van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. Lond.*, 265:359–366.
- Hauske, G. (1994). *Systemtheorie der visuellen Wahrnehmung*. B. G. Teubner, Stuttgart.
- Hauske, G. (1997). A self organizing map approach to image quality. *BioSystems*, 40:93–102.
- Hauske, G. und Zetzsche, C. (1990). Die Bedeutung des analytischen Signals in Bildanalyse und Bildcodierung. *Frequenz*, 44:68–73.
- Haykin, S. (1978). *Communication Systems*. John Wiley & Sohns, Inc., New York.
- Hegd , J. und van Essen, D. C. (2003). Strategies of shape representation in macaque visual area v2. *Visual Neuroscience*, 20:313–328.
- Helmholtz, H. (1850). Messungen  ber den Zeitlichen Verlauf der Zuckung animalischer Muskeln und die Fortpflanzungsgeschwindigkeit der Reizung in den Nerven. *Archiv Anatomie, Physiologie und wiss. Medicin*, Seiten 276–381.
- Helmholtz, H. v. (1910). *Handbuch der physiologischen Optik*, Band 3. Verlag von Leopold Voss, Hamburg und Leipzig, 3 Auflage.
- Hinich, M. J. (1982). Testing for Gaussianity and linearity of a stationary time series. *J. Time Series Analysis*, 3(3):169–176.
- Holambe, R. S., Ray, A. K. und Basu, T. K. (1996). Signal phase recovery using the bispectrum. *Signal Processing*, 55:321–337.
- Hoyer, P. O. und Hyv rinen, A. (2002). A multi-layer sparse coding network learns contour coding from natural images. *Vision Research*, 42(12):1593–1605.
- Hubel, D. H. und Wiesel, T. N. (1965). Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology*, 28:229–289.
- Hubel, D. H. und Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195:215–243.
- Hubel, D. H. und Wiesel, T. N. (1977). Functional architecture of macaque monkey visual cortex. *Proceedings of the Royal Society of London B*, 198(1–59).
- Hunter, I. W. und Korenberg, M. J. (1986). The identification of nonlinear biological systems: Wiener and Hammerstein cascade models. *Biological Cybernetics*, 55:135–144.

- Hurri, J. (1997). *Independent Component Analysis of Image Data*. Diplomarbeit, Helsinki University of Technology Department of Computer Science and Engineering.
- Hyvärinen, A. (1999). Fast and robust fixed-point algorithm for Independent Component Analysis. *IEEE Transactions on Neural Networks*, 10(3):626–634.
- Hyvärinen, A., Gutmann, M. und Hoyer, P. O. (2005). Statistical model of natural stimuli predicts edge-like pooling of spatial frequency channels in V2. *BMC Neuroscience*, 6(12):1–12.
- Hyvärinen, A. und Hoyer, P. O. (2001). A two-layer sparse coding model learns simple and complex cell receptive fields and topography from natural images. *Vision Research*, 41(18):2413–2423.
- Hyvärinen, A., Karhunen, J. und Oja, E. (2001). *Independent Component Analysis*. John Wiley & Sohns, Inc., New York.
- Hyvärinen, A. und Oja, E. (2000). Independent component analysis: Algorithms and applications. *Neural Networks*, 13:411–430.
- Ito, M. und Komatsu, H. (2004). Representation of angles embedded within contour stimuli in area v2 of macaque monkeys. *Journal of Neuroscience*, 24(13):3313–3324.
- Jacobson, H. (1951). The informal capacity of the human eye. *Science*, 113:292–293.
- Jayant, N. S. und Noll, P. (1984). *Digital Coding of Waveforms: Principles and Applications to Speech and Video*. Signal Processing Series. Prentice-Hall, Englewood Cliffs, NJ.
- Jones, J. und Palmer, L. (1987a). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58:1233–1258.
- Jones, J. P. und Palmer, L. A. (1987b). The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58(6):1187–1211.
- Julesz, B. und Schumer, R. A. (1981). Early visual perception. *Annual Review of Psychology*, 32:575–627.
- Kammeyer, K. D. (1996). *Nachrichtenübertragung*. B. G. Teubner, Stuttgart.
- Kaplan, E. und Shapley, R. M. (1982). X and Y cells in the lateral geniculate nucleus of macaque monkeys. *Journal of Physiology*, 330:125–143.
- Kato, H., Bishop, P. O. und Orban, G. A. (1978). Hypercomplex and simple/complex cell classifications in cat striate cortex. *Journal of Neurophysiology*, 41:1071–1095.
- Kay, S. M. (1988). *Modern Spectral Estimation*. Prentice-Hall, Englewood Cliffs, NJ.

- Kepler, J. (1604). *Ad Vitellionem Paralipomena*. C. Marnius and Heirs of J. Aubrius, Frankfurt.
- Kersten, D. (1987). Predictability and redundancy of natural images. *J. Opt. Soc. Am. A*, 4(12):2395–2400.
- Kim, K. I., Franz, M. O. und Schölkopf, B. (2003). Kernel hebbian algorithm for iterative kernel principal component analysis. Technischer Bericht TR-109, Max-Planck-Institut für biologische Kybernetik.
- Kim, K. I. und Powers, E. J. (1988). A digital method of modeling quadratically nonlinear systems with a general random input. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 36(11):1758–1769.
- Klein, S. A. und Tyler, C. W. (1986). Phase discrimination of compound gratings: generalized autocorrelation analysis. *J. Opt. Soc. Am. A*, 3(868–879).
- Korenberg, M. J. und Hunter, I. W. (1986). The identification of nonlinear biological systems: LNL cascade models. *Biological Cybernetics*, 55:125–134.
- Koukoulas, P. und Kalouptsidis, N. (1995). Nonlinear system identification using gaussian inputs. *IEEE Trans. Signal Process.*, 43(8):1831–1841.
- Krieger, G. (1999). *Nichtlineare Informationsverarbeitung in biologischen und technischen Sehsystemen: eine Analyse mit Volterra-Reihen und Statistiken höherer Ordnung*. Doktorarbeit, Lehrstuhl für Nachrichtentechnik, Technische Universität München.
- Krieger, G. und Zetsche, C. (1996). Nonlinear image operators for the evaluation of local intrinsic dimensionality. *IEEE Transactions on Image Processing*, 5(6):1026–1041.
- Krieger, G., Zetsche, C. und Barth, E. (1995). Nonlinear image operators for the detection of local intrinsic dimensionality. In *Proc. IEEE Workshop Nonlinear Signal and Image Processing*, Seiten 182–185.
- Krieger, G., Zetsche, C. und Barth, E. (1997). Higher-order statistics of natural images and their exploitation by operators selective to intrinsic dimensionality. In *Proc. IEEE Signal Processing Workshop on Higher-Order Statistics*, Seiten 147–151.
- Krüger, N. und Felsberg, M. (2003). A continuous formulation of intrinsic dimension. In *British Machine Vision Conference*.
- Kuffler, S. W. (1953). Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 16:37–68.
- Laughlin, S. B. (1981). A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch.*, 36C:910–912.

- Leonov, V. P. und Shiryaev, A. N. (1959). On a method of calculation of semi-invariants. *Theory Probab. Appl.*, 4(3):319–328.
- Lettvin, J. Y., Maturana, H. R., Pitts, W. H. und McCulloch, W. S. (1961). Two remarks on the visual system of the frog. In W. A. Rosenblith, Herausgeber, *Sensory Communications*, Seiten 757–776. John Wiley & Sons, New York.
- Li, M. und Vitanyi, P. (1993). *An Introduction to Kolmogorov Complexity and its Applications*. Springer, New York, Heidelberg, Berlin.
- Lindberg, D. C. (1976). *Theories of Vision from Al-Kindi to Kepler*. University of Chicago Press.
- Livingstone, M. und Hubel, D. H. (1988). Segregation of form, color, movement and depth: anatomy, physiology and perception. *Science*, 240:740–750.
- Livingstone, M. S., Pack, C. C. und Born, R. T. (2001). Two-dimensional substructure of MT receptive fields. *Neuron*, 30:781–793.
- Mach, E. (1922). *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen*. Fischer, 9. Auflage. Erstaussgabe 1886.
- Marcellin, M. W., Gomish, M. J., Bilgin, A. und Boliek, M. P. (2000). An overview of JPEG-2000. In *Proc. IEEE Data Compression Conf.*, Seiten 523–541.
- Mariotte, E. (1681). *Traité de la nature des couleurs*. Paris.
- Marko, H. (1994). *Systemtheorie – Methoden und Anwendungen für ein- und mehrdimensionale Systeme*. Springer Verlag, Berlin, Heidelberg, New York.
- Marmarelis, P. Z. und Marmarelis, V. Z. (1978). *Analysis of Physiological Systems*. Plenum Press, New York, NY.
- Marčelja, S. (1980). Mathematical description of the responses of simple cortical cells. *J. Opt. Soc. Am.*, 70:1297–1300.
- Mathews, V. J. und Sicuranza, G. L. (2000). *Polynomial Signal Processing*. Telecommunications and Signal Processing. John Wiley & Sohns, Inc., New York.
- Mendel, J. M. (1991). Tutorial on higher-order statistics (spectra) in signal processing and system theory: theoretical results and some applications. *Proceedings of the IEEE*, 79(3):278–305.
- Nelder, J. A. und Mead, R. (1965). A simplex method for function minimization. *Comput. J.*, 7:308–313.
- Nikias, C. L. und Petropulu, A. P. (1993). *Higher-Order Spectral Analysis: A Nonlinear Signal Processing Framework*. Prentice-Hall, Englewood Cliffs, NJ.

- Nuding, U. (2002). *Transformation statistischer Abhängigkeiten durch nichtlineare Filterung*. Diplomarbeit, Lehrstuhl für Nachrichtentechnik, Technische Universität München.
- Nuding, U. und Zetsche, C. (2003). Nonlinear two-layer wavelet networks for feature extraction and coding. In *Proc. IEEE Workshop on Nonlinear Signal and Image Processing*. Grado, Italy.
- Nuding, U. und Zetsche, C. (2004). Measurement of nonlinear 2nd-order kernels with polyspectra. In U. Ilg, H. H. Bülthoff und H. A. Mallot, Herausgeber, *Dynamic Perception: Workshop of the GI Section Computer Vision*, Seiten 189–194. IOS Press.
- Nuding, U. und Zetsche, C. (2006). Learning the selectivity of V2 and V4 neurons using nonlinear multi-layer wavelet networks. *Biosystems*. Im Druck.
- Nuding, U., Zetsche, C., Schill, K. und Hauske, G. (2004). Measurement of nonlinear 2nd-order kernels using Gaussian and natural inputs. In *Thirty-Eighth Asilomar Conference on Signals, Systems and Computers*. IEEE, Pacific Grove, CA.
- Oja, E. (1982). A simplified neuron model as a principal component analyzer. *J. Math. Biology*, 15:267–273.
- O’Kusky, J. und Colonnier, M. (1982). A laminar analysis of the number of neurons, glia and synapses in the visual cortex (area 17) of adult macaque monkeys. *J. Comp. Neurol.*, 210:278–290.
- Olshausen, B. und Field, D. J. (1997). Sparse coding with an overcomplete basis set: A strategy employed by V1? *Vision Research*, 37(23):3311–3325.
- Olshausen, B. A. und Field, D. J. (1996). Natural image statistics and efficient coding. *Network: Computation in Neural Systems*, 7(2):333–339.
- Pack, C. C., Born, R. T. und Livingstone, M. S. (2003a). Two-dimensional substructure of stereo and motion interactions in macaque visual cortex. *Neuron*, 37:525–535.
- Pack, C. C., Livingstone, M. S., Duffy, K. R. und Born, R. T. (2003b). End-stopping and the aperture problem: Two-dimensional motion signals in macaque v1. *Neuron*, 39:671–680.
- Pantle, A. und Sekuler, R. (1968). Size detecting mechanisms in human vision. *Science*, 162:1146–1148.
- Papoulis, A. (1991). *Probability, random variables, and stochastic processes*. McGraw-Hill, New-York.
- Pearson, K. (1892). *The Grammar of Science*. Walter Scott, London.

- Petropulu, A. P. (1998). Higher-order spectral analysis. In V. K. Madisetti und D. B. Williams, Herausgeber, *Digital Signal Processing (DSP) Handbook*, Kapitel 76. CRC Press, Boca Raton, FL.
- Petropulu, A. P. und Pozidis, H. (1998). Phase reconstruction from bispectrum slices. *IEEE Trans. Signal Processing*, 46(2):527–530.
- Pollen, D. A. und Ronner, S. F. (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science*, 212(4501):1409–1411.
- Press, William, H., Teukolsky, Saul, A., Vetterling, William, T. und Flannery, Brian, P. (1992). *Numerical Recipes in C*. Cambridge University Press, 2. Auflage.
- Purkinje, J. E. (1825). *Neuere Beiträge zur Kenntnis des Sehens in subjectiver Hinsicht*, Kapitel Beobachtungen und Versuche zur Physiologie der Sinne. Bd II. Reimer, Berlin.
- Råde, L. und Westergren, B. (1997). *Springers mathematische Formeln: Taschenbuch für Ingenieure, Naturwissenschaftler, Wirtschaftswissenschaftler*. Springer, Berlin, Heidelberg, New York, 2. korr. und erw. Aufl. Auflage.
- Rao, T. S. und Gabr, M. M. (1984). An introduction to bispectral analysis and bilinear time series models. *Lecture Notes in Statistics*, 24.
- Riesenhuber, M. und Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2(11):1019–1025.
- Röhrbein, F. und Zetzsche, C. (1998). Exploitation of natural image statistics by biological vision systems: $1/f^2$ power spectra and self-similar bandpass decompositions. In *Proc. CVPR-98*, Seiten 86–91. IEEE Computer Society, IEEE, Los Alamitos, CA.
- Rosenblatt, M. und Van Ness, J. W. (1965). Estimation of the bispectrum. *Annals of Mathematical Statistics*, 36:1120–1136.
- Ruderman, D. L. (1997). Origins of scaling in natural images. *Vision Research*, 37(23):3385–3398.
- Rugh, W. J. (1981). *Nonlinear System Theory. The Volterra-Wiener Approach*. John Hopkins University Press, Baltimore, Maryland.
- Sabra, A. I. (1981). *Theories of Light, from Descartes to Newton*. Cambridge University Press.
- Saito, H., Tanaka, K., Fukada, Y. und Oyamada, H. (1988). Analysis of discontinuity in visual contours in area 19 of the cat. *Journal of Neuroscience*, 8:1131–1143.
- Schetzen, M. (1980). *The Volterra and Wiener Theories of Nonlinear Systems*. John Wiley & Sohns, Inc., New York.

- Schölkopf, B. und Smola, A. (2002). *Learning with Kernels*. MIT Press, Cambridge, MA.
- Schölkopf, B., Smola, A. und Müller, K. (1998). Nonlinear component analysis as a kernel eigenvalue problem. *Neural Computation*, 10(5):1299–1319.
- Schwartz, O. und Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature Neuroscience*, 4(8):819–825.
- Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423.
- Shannon, C. E. (1951). Prediction and entropy of printed english. *The Bell System Technical Journal*.
- Shannon, C. E. und Weaver, W. (1976). *Mathematische Grundlagen der Informationstheorie*. R. Oldenbourg Verlag GmbH, München.
- Shapley, R. und Enroth-Cugell, C. (1984). Visual adaptation and retinal gain controls. *Prog. Retinal Research*, 3:263–346.
- Shevelev, I. A., Lazareva, N. A., Sharaev, G. A., Novikova, R. V. und Tikhomirov, A. S. (1998). Selective and invariant sensitivity to crosses and corners in cat striate neurons. *Neuroscience*, 84(3):713–721.
- Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J. und Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature*, 378:492–496.
- Skottun, B., DeValois, R., Grosf, D., Movshon, J., Albrecht, D. und Bonds, A. B. (1991). Classifying simple and complex cells on the basis of response modulation. *Vision Research*, 31(7/8):1079–1086.
- Stanley, G. B., Li, F. F. und Dan, Y. (1999). Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *Journal of Neuroscience*, 19(18):8036–8042.
- Sziklai, G. (1956). Some studies in the speed of visual perception. *IRE Trans. Inform. Theory*, IT-2:125–128.
- Szulborski, R. G. und Palmer, L. A. (1990). The two-dimensional spatial structure of nonlinear subunits in the receptive fields of complex cells. *Vision Research*, 30(2):249–254.
- Theunissen, F. E., David, S. V., Singh, N. C., Hsu, A., Vinje, W. E. und Gallant, J. L. (2001). Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. *Network: Computation in Neural Systems*, 12:289–316.

- Tick, L. J. (1961). The estimation of transfer functions of quadratic systems. *Technometrics*, 3:563–567.
- Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer-Verlag, Berlin, Heidelberg, New York.
- Vetterli, M. und Kovacevic, J. (1995). *Wavelets and Subband Coding*. Prentice-Hall, Englewood Cliffs, NJ.
- Vinje, W. E. und Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287(5456):1273–1276.
- Volterra, V. (1887). Sopra le funzioni che dipendono da altre funzioni. *Rend. Regia Accademia dei Lincei, 2o Sem.*, Seiten 97–105, 141–146, 153–158.
- Walker, G. A., Ohzawa, I. und Freeman, R. D. (2000). Suppression outside the classical cortical receptive field. *Visual Neuroscience*, 17(3):369–379.
- Wegmann, B. und Zetzsche, C. (1990a). Statistical dependence between orientation filter outputs used in a human vision based image code. In M. Kunt, Herausgeber, *Visual communication and image processing*, Band 1360 von *Proc. SPIE*, Seiten 909–923. Bellingham, WA.
- Wegmann, B. und Zetzsche, C. (1990b). Visual system based polar quantization of local amplitude and local phase of orientation filter outputs. In B. Rogowitz, Herausgeber, *Human Vision and Electronic Imaging: Models, Methods, and Applications, Proc. SPIE*, Band 1249, Seiten 306–317.
- Wegmann, B. und Zetzsche, C. (1996). Feature-specific vector quantization of images. *IEEE Transactions on Image Processing*, 5(2):274–288.
- Wiener, N. (1958). *Nonlinear Problems in Random Theory*. MIT Press, Cambridge, MA.
- Xu, W. und Hauske, G. (1995). Perceptually relevant error classification in the context of image coding. In *Proc. IEE Image Processing and its Applications*, Nummer 410, Seiten 589–593.
- Yellott, J. I. (1993). Implications of triple correlation uniqueness for texture statistics and the Julesz conjecture. *J. Opt. Soc. Am. A*, 10(5):777–793.
- Young, T. (1801). On the mechanism of the eye. *Philos Trans Royal Soc*, 92:23–88.
- Yu, Y. und Lee, T. S. (2003). Dynamical mechanisms underlying contrast gain control in single neurons. *Physical Review E*, 68(1).
- Zetzsche, C. (1990). Sparse coding: the link between low level vision and associative memory. In R. Eckmiller, G. Hartmann und G. Hauske, Herausgeber, *Parallel Processing in Neural Systems and Computers*, Seiten 273–276. Elsevier, Amsterdam.

- Zetzsche, C. (2002). The visual system is „blind“ to almost all possible images. In *4. Workshop Dynamic Perception*. Bochum, Germany.
- Zetzsche, C. und Barth, E. (1990a). Fundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Research*, 30(7):1111–1117.
- Zetzsche, C. und Barth, E. (1990b). Image surface predicates and the neural encoding of two-dimensional signal variations. In B. Rogowitz, Herausgeber, *Human Vision and Electronic Imaging: Models, Methods, and Applications*, Band SPIE 1249, Seiten 160–177.
- Zetzsche, C., Barth, E. und Rentschler, I. (1993a). Cortical end-stopped cells and the unambiguous detection of object motion. In N. Elsner und M. Heisenberg, Herausgeber, *Gene - Brain - Behavior. Proc. 21st Göttingen Neurobiology Conf.*, Seite 882. Thieme, Stuttgart, New York.
- Zetzsche, C., Barth, E. und Wegmann, B. (1993b). The importance of intrinsically two-dimensional image features in biological vision and picture coding. In A. Watson, Herausgeber, *Digital Images and Human Vision*, Seiten 109–138. MIT Press, Cambridge, MA.
- Zetzsche, C., Barth, E. und Wegmann, B. (1993c). Nonlinear aspects of primary vision: entropy reduction beyond decorrelation (invited address). In J. Morreale, Herausgeber, *SID International Symposium - Digest of Technical Papers*, Band XXIV, Seiten 933–936. Soc. Inform. Display, Playa del Ray, CA.
- Zetzsche, C. und Krieger, G. (1999). Nonlinear neurons and higher-order statistics: new approaches to human vision and electronic image processing. In B. Rogowitz und T. Pappas, Herausgeber, *Human Vision and Electronic Image Processing*, Band 3644 von *Proc. SPIE*. SPIE, SPIE, Bellingham, WA.
- Zetzsche, C. und Krieger, G. (2001a). Intrinsic dimensionality: Nonlinear image operators and higher-order statistics. In S. K. Mitra und G. L. Sicuranza, Herausgeber, *Nonlinear Image Processing, Communications, Networking and Multimedia*, Kapitel 14, Seiten 403–448. Academic Press, San Diego, CA.
- Zetzsche, C. und Krieger, G. (2001b). Nonlinear mechanisms and higher-order statistics in biological vision and electronic image processing: review and perspectives. *Journal of Electronic Imaging*, 10(1):56–99.
- Zetzsche, C., Krieger, G. und Mayer, G. (2001). Nonlinear AND interactions between frequency components and the selective processing of intrinsically two-dimensional signals by cortical neurons. In B. E. Rogowitz und T. N. Pappas, Herausgeber, *Human Vision and Electronic Imaging*, Band 4299 von *Proc. SPIE*, Seiten 36–68. SPIE, Bellingham, WA.

- Zetzsche, C., Krieger, G. und Wegmann, B. (1999). The atoms of vision: Cartesian or polar? *J. Opt. Soc. Am. A*, 16(7):1554–1565.
- Zetzsche, C. und Nuding, U. (2005a). Natural scene statistics and nonlinear neural interactions between frequency-selective mechanisms. *Biosystems*, 79(1–3):143–149.
- Zetzsche, C. und Nuding, U. (2005b). Nonlinear and higher-order approaches to the encoding of natural scenes. *Network: Computation in Neural Systems*, 16(2/3):191–221.
- Zetzsche, C. und Röhrbein, F. (2001). Nonlinear and extra-classical receptive field properties and the statistics of natural scenes. *Network: Computation in Neural Systems*, 12:331–350.
- Zetzsche, C., Schill, K., Deubel, H., Krieger, G., Umkehrer, E. und Beinlich, S. (1998). Investigation of a sensorimotor system for saccadic scene analysis: an integrated approach. In R. Pfeifer, B. Blumenberg, J.-A. Meyer und S. Wilson, Herausgeber, *Proc. 5th Intl. Conf. Soc. Adaptive Behavior*, Band 5, Seiten 120–126. MIT Press, Cambridge, MA.