

**Die Entwicklung rezeptiver Felder
und
neuronaler Karten im visuellen Kortex**

Dissertation
zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultäten
der Georg – August – Universität zu Göttingen

vorgelegt von

Norbert Michael Mayer
aus
Bobingen

Göttingen 2000

D 7

Referent: Prof. Dr. T. Geisel

Koreferent: Prof. Dr. R. Kree

Tag der mündlichen Prüfung: 1. November 2000

Inhaltsverzeichnis

1	Einleitung	7
2	Grundlagen zur Formation rezeptiver Felder	11
2.1	Neuron als Grundbaustein des Gehirns	11
2.2	Die Methode der reversen Korrelation	13
2.3	Anatomie und Physiologie des Sehens	14
2.4	Die Plastizität kortikaler rezeptiver Felder	17
2.5	Statistische Eigenschaften von natürlichen Bildern	18
2.6	Komponenten natürlicher Bilder	19
3	Formation rezeptiver Felder in Abhängigkeit der visuellen Erfahrung	23
3.1	Befunde von Vergleichsuntersuchungen von BD und normal entwickelten Tieren	23
3.2	Modellvoraussetzungen: Der retinale Filter	26
3.3	Stimulusensembles	27
3.4	Der Adaptionprozess	29
3.5	Rekonstruktion des rezeptiven Feldes	29
3.6	Numerische Resultate	32
3.7	Orientierungsselektivität	33

3.8	Wellenlängenverteilungen	34
3.9	Ähnlichkeit zu Gaborfiltern	35
3.10	Diskussion	39
3.11	Zusammenfassung	40
4	Neuronale Karten im visuellen Kortex	41
4.1	Biologische Fakten	41
4.2	Modelle zur Strukturbildung im visuellen Kortex	46
5	Kartenalgorithmen mit natürlichen Bildern als Inputs	49
5.1	Die Eigenschaften des SOM-Modells	49
5.2	Methoden der Extraktion von Merkmalskarten	51
5.3	Numerische Resultate zu selbstorganisierenden Karten	53
5.4	Der Zusammenhang zwischen lokaler Phase und Retinotopie	55
5.5	Eine Karte aus adaptiven Unterräumen	56
5.6	Numerische Resultate zum Modell mit adaptiven Unterräumen	57
5.7	Ein Vergleich der Modellresultate mit biologischen Fakten	64
5.8	Metriken und adaptive Featureräume	66
5.9	Die topologischen Eigenschaften des generischen rezeptiven Feldes in SOM und ASSOM	66
5.10	Zusammenfassung	69
6	Projektion hochdimensionaler Karten in den Featureraum	71
6.1	Grundlegende Überlegungen	71
6.2	Kodierung der Stimuli im Featureraum	72
6.3	Eine Modifikation des elastischen Netzes	73
6.4	Entsprechendes zum Kohonenalgorithmus	74

6.5	Symmetrieüberlegungen zur Metrik	76
6.6	Diskussion	83
6.7	Zusammenfassung	83
7	Ein einfaches Modell mit nichteuklidischem Featureraum	85
7.1	Eigenschaften des α -Modells	85
7.2	Der Fall $\alpha = 0$	87
7.3	Numerische Resultate zum α -Modell	87
7.4	Der Prozess der Kartenbildung	89
7.5	Vergleich mit Featurekarten hochdimensionaler Modelle	90
7.6	Diskussion	92
7.7	Zusammenfassung	93
8	Zusammenfassung und Diskussion	95
A	Rekonstruktion des rezeptiven Feldes aus der neuronalen Aktivität	99
B	Eichung der Orientierungsselektivität	101
C	Praktikable Herleitung der Metrik im Featureraum	105

Kapitel 1

Einleitung

Das menschlichen Gehirn enthält ungefähr 10^{10} Nervenzellen. Sie kommunizieren über 10^{13} synaptische Verbindungen und bilden so ein höchst komplexes System, das für unsere geistigen Fähigkeiten verantwortlich ist.

Offensichtlich liegt einer solchen Struktur eine grandiose Organisation zugrunde, die bis heute weitgehend unverstanden ist. Der Versuch, die dabei wirkenden Mechanismen zu verstehen, ist Gegenstand der Neurowissenschaften. Dabei geht der vollen Funktionsfähigkeit des Gehirns ein raffinierter Entwicklungs- und Selbstorganisationsprozess voraus. So sind bei den höheren Lebewesen nicht vom ersten Tag nach der Geburt alle kognitiven Fähigkeiten verfügbar, sondern diese entwickeln sich erst im Laufe der ersten Lebensmonate. Eine Vielzahl von biologischen Experimenten belegen das koordinierte Fortschreiten der kognitiven und mentalen Entwicklungsprozesse. Die Sehschärfe nimmt beispielsweise in den ersten Lebenswochen erheblich zu [Norcia & Tyler, 1985].

Ein besonders interessanter Aspekt ist dabei die Frage, inwieweit unsere mentalen Fähigkeiten durch Erfahrung erworben werden und inwieweit sie angeboren sind. Dabei ist die Hirnrinde, der Kortex, aus verschiedenen Gründen ein Gebiet von besonderem Interesse. Obwohl sich die Anforderungen an die Informationsverarbeitung im Bereich des Sehens, des Hörens oder des Tastens stark unterscheiden, ist die histologische Struktur des gesamten Kortex in allen Arealen relativ homogen. Wenn allgemeine Lernprozesse in bestimmten Regionen des Gehirns tatsächlich eine Rolle spielen, dann ist es am aussichtsreichsten dort danach zu suchen, wo ein homogenes Substrat verschiedenartige Reize verarbeitet. Ist im Kortex ein solcher allgemeiner Lernvorgang tatsächlich vorhanden, kann im Umkehrschluss angenommen werden, dass die Umleitung visueller Reize auf den auditorischen Kortex zu ähnlichen Strukturbildungsprozessen führen sollte, wie sie normalerweise im visuellen Kortex gefunden werden. Tatsächlich ist in neuester Zeit ein solches Experiment erfolgreich durchgeführt worden [Sharma et al., 2000]. Kortikale Strukturbildungsprozesse sind also aussichtsreiche Gebiete für den Nachweis eines stimulusgetriebenen Selbstorganisations-

prozesses.

Im Rahmen dieser Arbeit ist die Entwicklung des Sehsystems Gegenstand der Untersuchung. Das Sehen ist das Musterbeispiel neuronaler Verarbeitung, mit dem sich die kognitiven Neurowissenschaften bevorzugt auseinandersetzen, um letztlich unter anderem auf die Frage nach der Erfahrungsabhängigkeit der neuronalen Entwicklungsprozesse einzugehen und sie im Detail zu beantworten.

Das Sehen ist bei den höheren Vertebraten mit Abstand der Sinn, der die meisten kortikalen Ressourcen verbraucht. Ungefähr 30 % der Kortexoberfläche und diverse Kerne des Zwischenhirns befassen sich mit der Verarbeitung der Sehreize. Darüberhinaus sind die visuellen kortikalen Areale physiologisch gut erforscht, so dass eine Vielzahl von Details des Sehsystems bekannt ist.

Um diese Entwicklungsprozesse und speziell die des visuellen Kortex zu verstehen, hat sich eine eigene Forschungsrichtung im Rahmen der visuellen Neurowissenschaften etabliert, die sich mit der Entwicklung von Merkmalskarten befasst. Dabei werden in vielen Fällen verschieden Varianten selbstorganisierender Karten verwendet. Bereits das erste dieser Modelle [von der Malsburg, 1973] vermochte wichtige Eigenschaften des adulten visuellen Kortex mit Prinzipien der Selbstorganisation zu erklären. In der Folge sind viele andere Modelle hinzugekommen, die versuchen die Phänomene im visuellen Kortex mit selbstorganisierenden Kartenmodellen zu erklären (zB. [Erwin et al., 1995], [Brockmann et al., 1997b], [Wolf & Geisel 1998]). Von experimenteller Seite steht ebenfalls eine Vielzahl von Ergebnissen zu Verfügung, insbesondere solche, die mit der *optical imaging* Methode in Verbindung stehen (u.a. [Bonhoeffer & Grinvald, 1991], [Bosking et al., 1997], [Löwel, 1998]).

Eine andere, neuere Richtung versucht die Eigenschaften der rezeptiven Felder im Kortex mit einer nach bestimmten Kriterien optimalen Verarbeitung von visueller Information zu identifizieren. Hier werden Komponenten in natürlichen Bildern (Fotos), also den Stimulierungsarten untersucht, die unserer visuellen Erfahrung in etwa entsprechen. Vertreter dieser Richtung sind [Field, 1994], [Ruderman, 1997], [Olshausen & Field, 1996].

Diese beiden Bereiche, selbstorganisierende Merkmalskarten und Komponenten natürlicher Bilder, fließen in diese Arbeit ein. Mit ihrer Hilfe sollen einige experimentelle Befunde mit konzeptionell möglichst einfachen Modellen in Einklang gebracht werden.

Um die Frage nach der Erfahrungsabhängigkeit zu präzisieren, sollen drei Szenarien für Entwicklungsprozesse unterschieden werden. So können diese

1. genetisch determiniert,
2. durch eine interne Dynamik, die verschiedene Verarbeitungsstufen umfassen könnte, wesentlich moduliert sein oder
3. getrieben durch äußere Stimulierung als erfahrungsabhängig gelten.

Der vorliegenden Arbeit liegt das dritte Szenario als Arbeitshypothese zugrunde. Die Erfahrungsabhängigkeit zweier Aspekte der Strukturbildung im Kortex – nämlich der Entwicklung rezeptiver Felder und der Entwicklung von Merkmalskarten – soll mit weiterentwickelten Varianten der oben erwähnten Modelle getestet werden.

Die Arbeit gliedert sich in drei Teile, die in sich geschlossen sind. Zusammen stellen sie die Entwicklung in meiner persönlichen Auseinandersetzung mit den Themen rezeptive Felder, kortikale Merkmalskarten und Modelle der Selbstorganisation dar.

Nach einer Einführung in die biologischen Grundlagen wird im ersten Teil dieser Arbeit untersucht, zu welchem Grad die Entwicklung einzelner Neurone des visuellen Kortex von der visuellen Erfahrung abhängt. Die Stimulierung erfolgt im Modell durch natürliche Bilder (Fotos). Die Ergebnisse aus diesen Untersuchungen werden mit entsprechenden biologischen Resultaten in Beziehung gesetzt. Es sind eine Reihe von Experimenten an Tieren durchgeführt worden, die die Entwicklung des neuronalen Verhaltens unter verschiedenartig modifizierten Umweltbedingungen untersuchen. Diese werden zum Vergleich mit den Ergebnissen dieses Modells herangezogen. Anders als in früheren Ansätzen wird ein retinaler Entwicklungsvorgang als Erklärungsmuster für einige Phänomene bei der Entwicklung rezeptiver Felder in das Modell eingebaut.

Der zweite Teil wird mit einem Überblick über die Fakten zu biologischen Merkmalskarten eingeleitet. Im weiteren Verlauf werden Modelle zur topographischen Strukturbildung im visuellen Kortex untersucht. Dabei werden wieder natürliche Bilder als Stimulus verwendet. Es stellt sich jedoch heraus, dass die zunächst verwendeten selbstorganisierenden Karten in diesem Fall nicht mit den biologisch gefundenen Resultaten in Einklang zu bringen sind. So kann die retinotopie Anordnung der rezeptiven Felder nicht erhalten werden, was im Widerspruch zu biologischen Befunden steht. Als Grund dafür wird ein einfacher metrischer Zusammenhang zwischen zwei Stimuluseigenschaften festgestellt. Als Alternative wird ein Modell verwendet, das lokal lineare adaptive Unterräume (lokale PCA) zur Stimulusrepräsentation verwendet. Dieses Modell vermag die Retinotopie zu erhalten und ergibt zwanglos die Antworteigenschaften kortikaler *complex cells*.

Schließlich wird im dritten Teil eine kanonische Abbildung einer bestimmten Modellklasse, zu der auch das zuvor erarbeitete Modell gehört, auf äquivalente Modelle im niedrigdimensionalen Reizraum abgebildet. Dabei soll auf eine nichteuklidische Riemannsche Metrik zurückgegriffen werden, deren grundlegende Eigenschaften sich aus den Symmetrien des Modells ergeben. Der explizite differentialgeometrische Ansatz ist ein originärer Bestandteil dieser Arbeit und wurde bisher noch nicht auf *feature*-Modelle angewendet. Die Arbeit schließt mit einer Diskussion und Zusammenfassung ab.

Kapitel 2

Grundlagen zur Formation rezeptiver Felder

Der erste Teil dieser Arbeit beschränkt sich auf die Entwicklung rezeptiver Felder von Neuronen im primären visuellen Kortex. Zunächst sollen die biologischen Fakten zu diesem Thema dargestellt werden. Das rezeptive Feld beschreibt die Menge der Reize, die eine signifikant erhöhte Aktivität der Neuronen verursachen. Das Konzept des rezeptiven Feldes bezieht sich auf die in der klassischen Elektrophysiologie typische Versuchsanordnung, bei der von einem meist anästhesierten Tier die elektrische Aktivität mittels einzelner Elektroden abgeleitet wird. Meist beschränkt sich diese Untersuchungsmethode auf die primären Gebiete der Großhirnrinde. Durch systematische Platzierung der Elektroden und durch die verfügbaren bildgebenden Verfahren konnten zunehmend auch die Lagerrelationen von Neuronen mit ähnlicher Reizantwort aufgeklärt werden. Die experimentellen Fakten zu den topographischen Eigenschaften kortikaler Karten werden in Kapitel 4 nachgeliefert, um den zweiten Teil dieser Arbeit einzuleiten.

2.1 Neuron als Grundbaustein des Gehirns

Der funktionelle Grundbaustein des zerebralen Gewebes ist die Nervenzelle. Eine Nervenzelle besteht aus einem Dendritenbaum, einem Zellkörper und einem Axon (siehe Abb. 2.1). Synapsen verbinden Axone und die Dendriten anderer Nervenzellen, und bewältigen die transzelluläre Informationsübermittlung.

Das Neuron verarbeitet die eingehende Information und kodiert sie seinerseits in Aktionspotentiale (*spikes*). Diese können als diskrete, zeitliche Signale angesehen werden. Vereinfachend soll die Rate dieser Aktionspotentiale als ein analoges Signal des Neurons interpretiert werden.

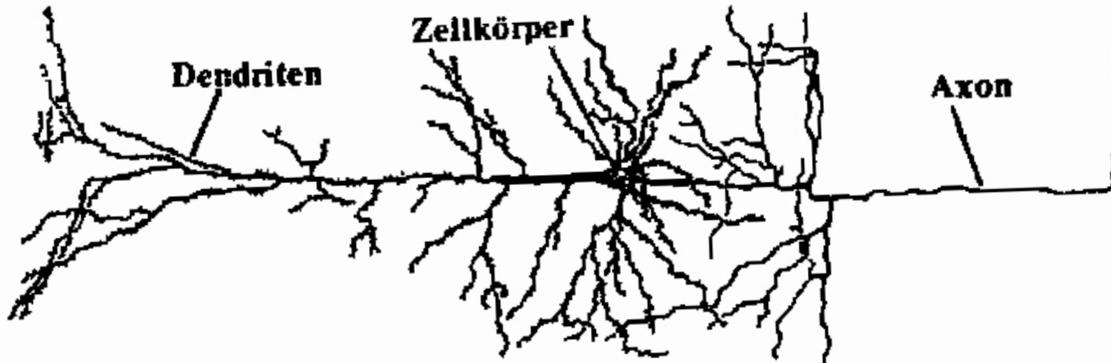


Abbildung 2.1: Ein kortikales Neuron (hier eine Pyramidenzelle) nach [Cajal, 1955]. Der Signalweg läuft von links nach rechts.

Ein Neuron lässt sich als eine Verarbeitungseinheit vorstellen, die aus einem Satz von Eingangswerten $\mathbf{V} \in \mathcal{S}$ ein Ausgangssignal R erzeugt. Eine einfache Näherung, die auf [Mc Culloch & Pitts, 1943] zurückgeht, ist

$$R = S\left(\sum V_{\xi} W_{\xi}\right). \quad (2.1)$$

Dabei ist S eine im Allgemeinen nichtlineare monoton steigende (beispielsweise sigmoide) Übertragungsfunktion, und \mathbf{W} ein Vektor, der die Gesamtheit der zu den Eingängen gehörenden synaptischen Gewichte kodiert. Innerhalb dieses Modells lässt sich dann das rezeptive Feld durch eine Bedingung für R als die Menge der Stimuli charakterisieren, die das Neuron erregen. \mathbf{W} repräsentiert die Stärke der synaptischen Verbindung zum betrachteten Neuron.

Sind die Einträge in \mathbf{W} positiv, spricht man von exzitatorischen, sind sie negativ von inhibitorischen Verbindungen. Offensichtlich wird die Antwort des Neurons zu einer gegebenen Signalstärke $\sum (V_{\xi})^2$ dann maximal sein, wenn Signal und rezeptives Feld parallel

$$V_{\xi} \propto W_{\xi} \quad (2.2)$$

sind. Der Vektor \mathbf{W} beinhaltet also hier ein Abbild des Stimulus, der im Neuron die stärkste Antwort hervorruft.

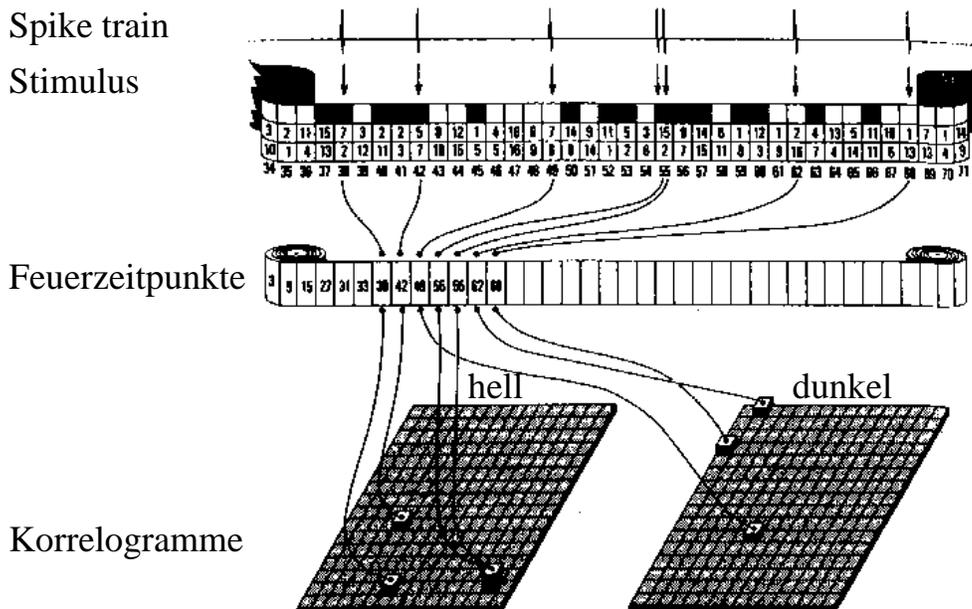


Abbildung 2.2: Reverse Korrelation zur Bestimmung eines rezeptiven Feldes (bearbeitet aus [Jones & Palmer, 1987a]). In diesem Experiment wurden zufällig positionierte helle und dunkle Pixel auf einem neutralen Hintergrund präsentiert. Falls das Neuron zum Feuern angeregt worden ist, trägt der Stimulus zum Korrelogramm bei.

2.2 Die Methode der reversen Korrelation

Seit den zwanziger Jahren des zwanzigsten Jahrhunderts kann die elektrische Aktivität einzelner Nervenzellen gemessen werden. Mit Hilfe von Elektroden, die entweder in der unmittelbaren Umgebung eines Neurons angebracht werden oder die Zellmembran durchdringen, lässt sich der Zustand des biologischen Neurons untersuchen. Insbesondere kann festgestellt werden, ob in dem Neuron bzw. der Neuronengruppe ein Aktionspotential ausgelöst wurde. Setzt man die Aktivität des Neurons mit den Umweltbedingungen des Versuchstieres in Beziehung, so lassen sich experimentell Rückschlüsse auf die rezeptiven Feldeigenschaften des Neurons treffen.

Die unvoreingenommenste Technik ein biologisches rezeptives Feld \mathbf{W} zu messen, ist die reverse Korrelation (*reverse correlation*). Die hier gegebene Darstellung folgt dem Experiment, das von [Jones & Palmer, 1987a] durchgeführt wurde. Die entscheidende Frage ist nun, welche Stimulierung am besten geeignet ist, um aus der Antwort des Neurons den rezeptiven Feldvektor zu erschließen.

Angenommen, ein Neuron erweist sich in Voruntersuchungen für einen bestimmten Ausschnitt des Sehfeldes sensitiv, dann wird dieses Neuron einer stochastischen Sequenz von Pixelmustern $Y_{r,t} \in \mathcal{T}$ ausgesetzt, die im sensitiven Bereich des Neurons zentriert sind. Im

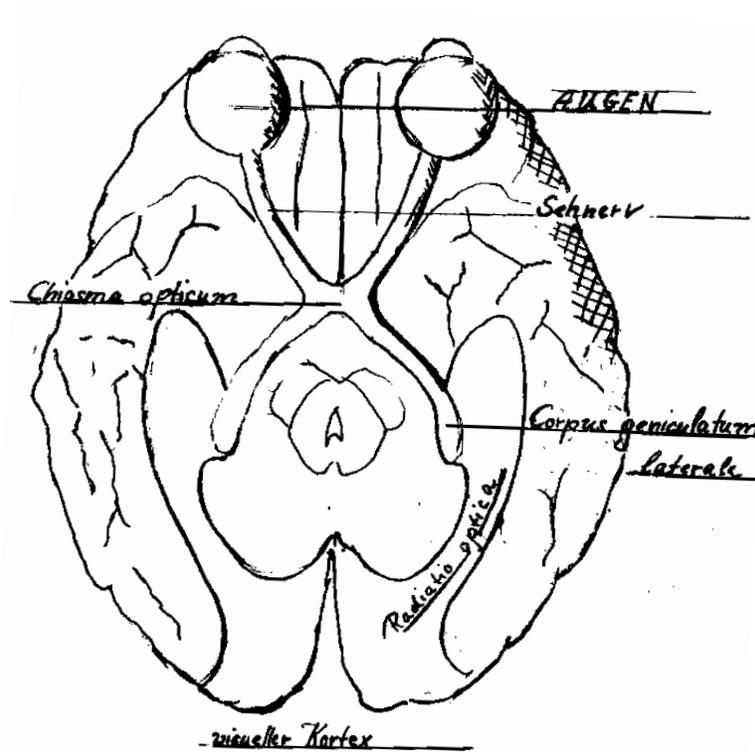


Abbildung 2.3: Die Anatomie des Sehweges. (Gezeichnet von N.M. nach [Hubel, 1990])

Anhang A wird im Rahmen einer hier hinreichenden Vereinfachung gezeigt, dass

$$E[Y_{r,t}R]_{\mathcal{T}} \propto F(S(\cdot), W_r) \quad (2.3)$$

ist, wobei F eine durch das Antwortverhalten $S(\cdot)$ des Neurons gegebene Funktion ist. Somit lässt sich aus dem Antwortverhalten R die Gestalt des rezeptiven Feldes W_r rekonstruieren.

2.3 Anatomie und Physiologie des Sehens

Die Hauptverarbeitungslinie des Sehens beginnt in der Netzhaut und führt über den Sehnerv und die Sehwegkreuzung (Chiasma opticum) zum seitlichen Kniehöcker (Corpus geniculatum laterale, CGL). Von dort geht es weiter über die Sehstrahlung (Radiatio optica) zur Sehrinde (visueller Kortex). Diesen anatomischen Gegebenheiten sollen nun physiologische Tatsachen zugeordnet werden.

Das Bild des Sehfeldes wird von dem okularen Linsensystem auf der Netzhaut (Retina) abgebildet. Hier befinden sich verschiedenartige Rezeptorzellen, die die visuellen Reize in

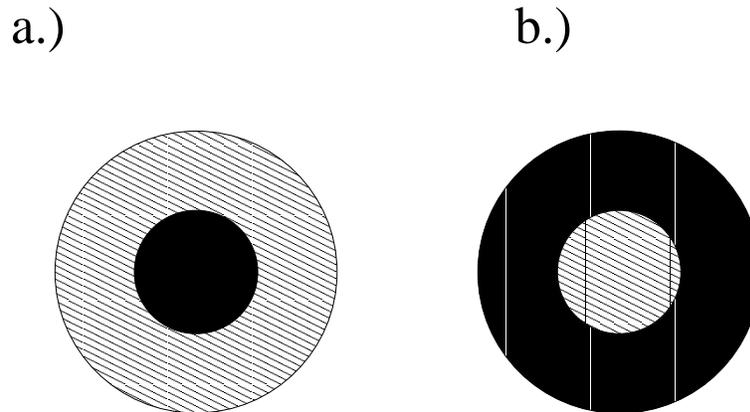


Abbildung 2.4: Rezeptive Felder a) *ON-center* und b) *OFF-center* in retinalen Ganglienzellen dargestellt in Koordinaten des visuellen Feldes. Hell schraffiert sind Bereiche, in denen das Antwortverhalten der Zelle mit hellem Licht, schwarz entsprechend Bereiche in denen das Antwortverhalten mit dunklem Licht korreliert ist.

Membranpotentialen kodieren. Dieses Signal wird schließlich in retinalen Ganglienzellen gemittelt. Bei den retinalen Ganglienzellen herrschen zwei rezeptive Feldformen vor: *ON-center* und *OFF-center* Zellen. Der erste Typ führt zum Auslösen von Aktionspotentialen der retinalen Ganglienzellen bei einer hellen Kontur auf dunklem Hintergrund, der zweite Typ löst ein Antworten der Zelle bei einer dunklen Kontur auf einem hellen Hintergrund aus. Neben diesem Unterscheidungsmerkmal lassen sich retinale Ganglienzellen auch durch andere physiologische Unterschiede in eine Vielzahl von Klassen einteilen. So wird das Antwortverhalten der Y-Zellen als eine lineare Summation über die lokalen Reizungen angegeben, während andere Zelltypen nichtlinear antworten. Der exakte Effekt der retinalen Verarbeitung von Reizen lässt sich also nicht in einer einfachen Regel wiedergeben und wird in den folgenden Modellen als ein lokaler Bandpassfilter wiedergegeben.

Die überwiegende Mehrheit retinaler Reizleitungen mündet im seitlichen Kniehöcker. Anatomisch ist der CGL als Teil des Thalamus in sechs Schichten unterteilt, die die retinalen Afferenzen sowohl nach ihrer Herkunft als auch nach ihrer Funktion trennen. Im CGL sind die Afferenzen von beiden Augen noch streng getrennt, so dass jedes Auge auf jeweils drei Schichten projiziert. Für jedes Auge zeichnet sich eine Schicht durch große Zellkörper aus und wird deshalb magnozelluläre Schicht genannt. In ihr terminieren die Afferenzen aller Ganglienzelltypen. In den beiden anderen Schichten, die aufgrund von kleineren Zellkörpern parvozelluläre Schichten genannt werden, terminieren lediglich die Y-Zellen. Die Form der rezeptiven Felder im CGL ist der der retinalen rezeptiven Felder sehr ähnlich.

Der am Hinterhauptslappen befindliche visuelle Kortex ist als Teil des Isokortex histologisch grob in sechs Schichten zu unterteilen. Der visuelle Kortex gliedert sich wiederum in verschiedene Areale, von denen eines das Areal V1 ist. In ihm zeigt sich ein sogar mit bloßem Auge erkennbarer dunkler Streifen, weshalb dieses Gebiet auch Area striata ge-

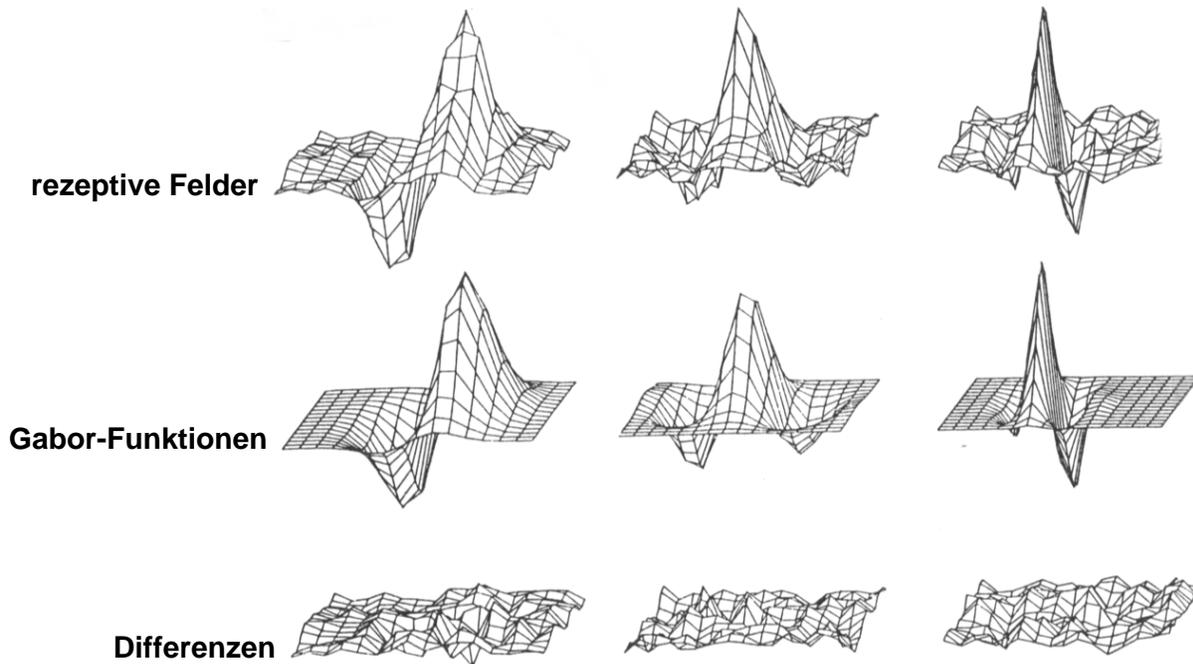


Abbildung 2.5: Rezeptive Felder kortikaler *simple cells*, wie sie mit Hilfe von *reverse correlation* Methoden gefunden werden (oben); die ähnlichsten Gabor-Filter (in der Mitte) und die Differenz aus rezeptivem Feld und Gaborfilter (unten) (aus [Jones & Palmer, 1987a]).

nannt wird. Diesem Teil des visuellen Kortex wird das Hauptinteresse dieser Arbeit gelten.

Im visuellen Kortex V1 können histologisch mehrere verschiedene Neuronentypen unterschieden werden. Experimentelle Messungen sind vor allem an Pyramidenzellen durchgeführt worden. Bei den Pyramidenzellen unterscheidet man zwischen komplexen Zellen (*complex cells*) und simplen Zellen (*simple cells*). Zunächst sollen nur die *simple cells* von Interesse sein. Mit Hilfe von reversen Korrelationstechniken konnten die mit dem Feuern der Zelle korrelierten Reizeigenschaften ohne weitergehende Annahmen studiert werden. Die Experimente führten bei *simple cells* zu rezeptiven Feldformen, die die Gestalt eines zweidimensionalen Gaborfilters haben [Jones & Palmer, 1987c] (siehe Abb. 2.5). Ein rezeptives Feld einer kortikalen *simple cell* setzt sich aus zwei bis drei parallelen Subfeldern zusammen, die alternierend über Afferenzen von genikulaten *on-center* und *off-center* Zellen verfügen.

Schon Hubel und Wiesel [Hubel & Wiesel, 1959] zeigten, dass die kortikalen Neurone im Gegensatz zu den darunter liegenden Verarbeitungsstufen, dem CGL und der Retina, selektiv auf Reize einer bestimmten Orientierung reagieren. Unter Benutzung eines rezeptiven Feldes, das einem Gaborfilter entspricht, kann dieses Phänomen leicht erklärt werden (siehe Abb. 2.6). Das Skalarprodukt eines Gaborfilters und einer Kante quer zu seiner Orientierung verschwindet im Wesentlichen, während das Skalarprodukt zwischen einer Kante längs

der Orientierung und einem Gaborfilter stark fluktuiert. Die Funktion dieser Neurone ist folglich die eines orientierungsselektiven Kantendetektors. Da auch in der technischen Bildverarbeitung zu diesem Zweck solche Filter verwendet werden, lag es nahe, diese auch im Kortex zu postulieren. Hubel und Wiesel nahmen deshalb eine entsprechende Verschaltung zwischen den kortikalen Neuronen und ihren genikulaten Afferenzen an.

Neben diesen Afferenzen aus dem CGL sind die kortikalen Neurone auch lateral verschaltet und haben darüber hinaus noch Input von höheren kortikalen Arealen. Diese Verbindungen sind teils inhibitorischer, teils exzitatorischer Natur. Die lateralen Verschaltungen lassen sich in drei Gruppen unterteilen [Kisvárdy et al., 1997]. Die erste Gruppe verbindet unspezifisch exzitatorische Neurone auf kurzen Abständen, die zweite Gruppe ist inhibitorisch und ebenfalls unspezifisch. Die dritte Gruppe verbindet selektiv Neurone mit einer ähnlichen Orientierungspräferenz exzitatorisch.

2.4 Die Plastizität kortikaler rezeptiver Felder

Das visuelle System ist bei der Geburt nicht vollständig strukturiert, sondern entwickelt sich als ein dynamischer Prozess. Wie schon in der Einleitung angedeutet wurde, ist die interessanteste Frage, inwieweit diese Vorgänge genetisch bedingt sind oder ob sie von der Umgebung beeinflusst werden. Es lassen sich verschiedene biologische Indizien angeben, die helfen können, diese Frage zu beantworten. Sie betreffen insbesondere die ersten Wochen nach der Geburt, die kritische Phase. In diesem Zeitraum werden eine Reihe von Vorgängen durch visuelle Erfahrungen beeinflusst. Neurologische Defekte, die mit Störungen der Stimulusstatistik einhergehen, wie beispielsweise durch Schielen (Strabismus), bleiben in dieser Phase reversibel. Sie sind aber im späteren Leben festgelegt. Auch Veränderungen in der kortikalen Strukturbildung durch solche Einflüsse sind beobachtet worden [Löwel, 1998] (siehe auch [Sengpiel et al., 1999]).

Solche Effekte weisen darauf hin, dass Lernprozesse im Kortex stattfinden. Eine weit hin akzeptierte Auffassung, wie solche Lernprozesse ablaufen, wurde von Donald Hebb [Hebb, 1949] formuliert. Im Hebb'schen Sinne erhöhen zwei Neurone die synaptische Stärke ihrer Verbindung, wenn ihre Aktivität positiv korreliert ist. Ein solcher Vorgang lässt sich durch die Lernregel

$$\frac{\partial \mathbf{W}}{\partial t} = \gamma R \mathbf{V}_t \quad (2.4)$$

erzielen, wobei γ eine charakteristische Zeitkonstante ist. Diese Lernregel hat sich in vielen Modellen als erfolgreich erwiesen und zeichnet sich zudem durch ihre Einfachheit aus.

Damit die funktionelle Variabilität der kortikalen Neurone als Ergebnis eines solchen Lernprozesses erklärt werden kann, ist es zudem notwendig eine laterale Wechselwirkung zwischen kortikalen Neuronen zu postulieren, die inhibitorisch sein soll. Diese dient dazu die

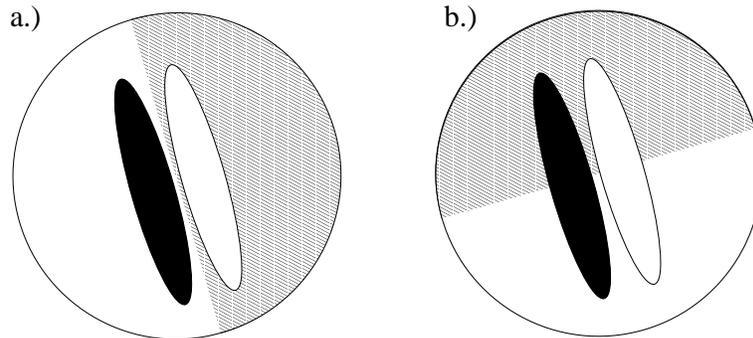


Abbildung 2.6: Angenommen das rezeptive Feld eines Neurons im visuellen Feld hätte grob die Gestalt eines Gaborfilters, so ergeben sich orientierungsselektive Antworten auf Kanten im visuellen Feld. Hat eine Kante die im Fall a skizzierte Orientierung so wird dies zu einer starken Antwort des Neurons führen, im Fall b zu keiner Antwort.

Aktivität auf bestimmte Neuronengruppen zu konzentrieren und auf diese Weise eine Spezialisierung der Neuronen auf bestimmte Reize zu ermöglichen. Die tatsächlich vorhandenen inhibitorischen lateralen Verbindungen zwischen kortikalen Neuronen mögen als Indiz für die Plausibilität dieser Modellkomponente gewertet werden.

Der Extremfall ist, dass das den Stimulus repräsentierende Neuron die anderen Neuronen so stark unterdrückt, dass nur dieses eine Neuron aktiv sein kann. Man spricht dann von einem *winner takes all*-Ansatz. Dieses Prinzip wird in der vorliegenden Arbeit vorwiegend verwendet werden. Alternativ dazu werden im nächsten Abschnitt die spärlichen Kodierung und später (siehe Kapitel 4) das Elastische Netz als abgeschwächte Varianten vorgestellt.

2.5 Statistische Eigenschaften von natürlichen Bildern

Das visuelle System von Lebewesen ist der Stimulierung durch natürliche Bilder ausgesetzt. Bei der Modellierung wird ein gewisser Realismus dadurch erreicht, daß die Stimuli aus Fotografien natürlicher Szenarien erzeugt werden. In der Praxis werden vorzugsweise Fotos von Wiesen und Wäldern mit Landschaftsdetails verwendet.

Natürliche Bilder zeichnen sich somit gegenüber Zufallsbildern durch besondere statistische Eigenschaften aus. Analysen großer Bildmengen [Ruderman, 1997], [Field, 1994] haben insbesondere die beiden folgenden Resultate erhärtet.

1. Natürliche Bilder haben auf verschiedenen Längenskalen charakteristische Funktionen. Die Fouriertransformierte der Autokorrelationsfunktion zeigt im Mittel folgende

Eigenschaft:

$$\tilde{c}(\mathbf{k}) = \frac{c_0}{|\mathbf{k}|^{2+\eta}},$$

mit einem kleinen $|\eta|$, dessen Wert normalerweise unter 0.3 liegt. \mathbf{k} ist die Wellenzahl des zweidimensionalen Spektrums eines Bildes. Die Korrelation fällt im Fourierraum also algebraisch ab und ist gegenüber Skalierungsänderungen invariant.

2. Es existieren Kanten zwischen Bereichen relativ homogener Einfärbung. Diese induzieren Korrelationen höherer Ordnung, welche einen Zusammenhang zwischen Orientierung und Ort aufweisen.

Bilder die lediglich aus Kanten bestehen, verfügen im Prinzip ebenfalls über ein $|\mathbf{k}|^{-2}$ -Spektrum. Allerdings mitteln sich diese bei zufälliger Anordnung weitgehend aus.

Es ist in der Natur kein ungewöhnlicher Vorgang, dass sich biologische Systeme unter dem Druck der Evolution sehr spezifisch auf die Optimierung hinsichtlich typischer Eigenschaften der tatsächlich vorhandenen Signale einrichten. Dabei ist zunächst unerheblich, ob die Verarbeitungsmechanismen selbst genetisch überliefert sind oder nur der Lernvorgang ein genetisch gegebener Prozess ist, der im Lebewesen im Laufe der Entwicklung zu einer optimalen Verarbeitung führt.

Deswegen ist es nicht sehr überraschend, dass Eigenschaften der kortikalen Verarbeitung, die hier an rezeptiven Feldern untersucht werden sollen, sich auch in Modellalgorithmen finden lassen, die eine Verarbeitung der natürlichen Bilder nach bestimmten Kriterien optimieren. Im nächsten Abschnitt wird ein kurzer Überblick über verschiedene informationstheoretischen Ideen gegeben, die sich als biologisch relevant erwiesen haben.

2.6 Komponenten natürlicher Bilder

In den letzten Jahren haben sich einige Resultate auf dem Gebiet der statistischen Analyse natürlicher Bilder für die kortikale Entwicklung als bedeutsam erwiesen. Hier soll kurz auf zwei die Entropie der Darstellung maximierende, lineare Kodierungen eingegangen werden.

Wir betrachten die Abbildung

$$\mathbf{a} = f(\mathbf{b}) \tag{2.5}$$

des Stimulusraumes \mathcal{S} auf den Raum der möglichen neuronalen Aktivitäten \mathcal{A} , wobei die $\mathbf{b} \in \mathcal{S}$ die zu kodierenden Daten darstellen und $\mathbf{a} \in \mathcal{A}$ den Kode, also die Aktivität eines

Neuronensembles. Der Kode soll vereinfachend als linear angenommen werden. Nun lässt sich die Gleichung

$$\mathbf{a} = \hat{\mathbf{C}} \mathbf{b} \quad (2.6)$$

schreiben, wobei Spalten der Matrix $\hat{\mathbf{C}}$ als Basisvektoren des Kodes bezeichnet und als normiert definiert werden. Um aus dem neuronalen Kode den Stimulus wieder rekonstruieren zu können, bedarf es der Umkehrabbildung $\hat{\mathbf{D}} = \hat{\mathbf{C}}^{-1}$:

$$\mathbf{b} = \hat{\mathbf{D}} \mathbf{a} \quad (2.7)$$

Um eine geeignete Kodierung zu erhalten, ist es naheliegend die Korrelationsmatrix der Bilder zu betrachten:

$$R_{ij} = \mathbf{E} [(b_i - \mathbf{E}[b_i]_{\mathcal{S}})(b_j - \mathbf{E}[b_j]_{\mathcal{S}})]_{\mathcal{S}}, \quad (2.8)$$

wobei b_i die Einträge von \mathbf{b} und $\mathbf{E}[b_i]_{\mathcal{S}}$ deren Mittelwert über \mathcal{S} darstellen. Da $\hat{\mathbf{R}}$ symmetrisch ist, ist eine Diagonaldarstellung mit reellen Eigenwerten möglich

$$\hat{\mathbf{R}} = \hat{\mathbf{C}} \text{Diag}(\lambda_{jj}) \hat{\mathbf{C}}^{-1}. \quad (2.9)$$

Dabei besteht $\hat{\mathbf{R}}$ aus den R_{ij} . Nun kann die Transformationsmatrix als Kodierungsmatrix $\hat{\mathbf{C}}$ verwendet werden. Die Eigenvektoren stellen dann die Komponenten der natürlichen Bilder dar, wie sie auch bei Hauptachsentransformationen (*principal component analysis*, PCA) gefunden werden. $\hat{\mathbf{C}}$ ist in diesem Fall orthonormal, so dass $\hat{\mathbf{D}} = \hat{\mathbf{C}}^T$ ist. Die transformierte Darstellung $\mathbf{a} = \hat{\mathbf{C}}\mathbf{b}$ zeigt nun die Korrelationsmatrix

$$\mathbf{E}[a_i a_j]_{\mathcal{S}} = \delta_{ij} \lambda_{jj}. \quad (2.10)$$

Sie ist somit unkorreliert. Offensichtlich lässt sich auch jeder Spalte \mathbf{C}_j von $\hat{\mathbf{C}}$ ein λ_{jj} zuordnen, welches die Varianz der Komponente widerspiegelt. Ein weiterer Aspekt wird deutlich, wenn angenommen wird, dass eine Komponente a_j ausfällt. Wählt man die Komponente, welche die kleinste Varianz $\sigma_j^2 = \lambda_{jj}$ besitzt, so lässt sich zeigen, dass die obige Darstellung optimal im Sinne des quadratischen Restfehlers ist (siehe [Deco & Obradovic, 1996]). Zahlreiche biologisch motivierte Modelle mit lernenden Neuronen [Oja, 1982, Sanger, 1989] produzieren ein Antwortverhalten, das einer Hauptachsentransformation auf dem Stimulus entspricht.

Allerdings führen die Basisvektoren dieser linearen Kodierung im Falle visueller Stimulierung nicht zu rezeptiven Feldern, wie sie im visuellen Kortex gefunden werden (siehe Abb. 2.7). Anders als dort sind die Hauptkomponenten nicht lokal und entsprechen in ihrem Aussehen nicht Gaborfiltern.

Ein neuerer, ebenfalls linearer Ansatz geht auf die der Beobachtung zurück [Field, 1994], dass in der Retina und im Kortex das Stimulusensemble durch die Vielzahl der Neuronen

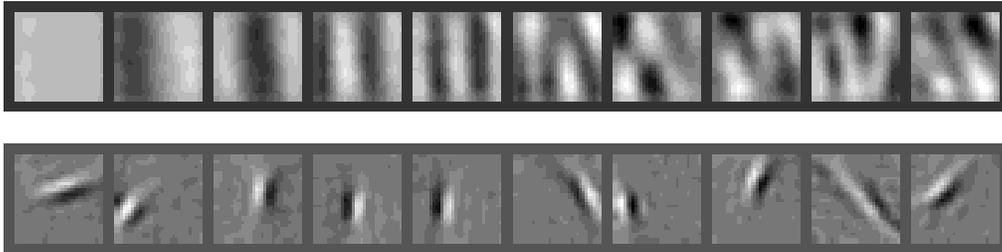


Abbildung 2.7: Rezeptive Feldformen, wie sie durch die Anwendung verschiedener informationstheoretischer Prinzipien auf natürliche Bilder zustandekommen. Im Fall der PCA (oben) sind die zehn Hauptkomponenten dargestellt, deren Eigenwerte am größten sind. Im Fall der spärliche Kodierung (unten) nach [Olshausen & Field, 1996] sind zehn repräsentative Basisvektoren ausgewählt worden. Hell bedeutet jeweils Exzitation durch helle Bildabschnitte, dunkel entsprechend durch dunkle Abschnitte.

repräsentiert ist. Der Stimulusraum ist also mit einem übervollständigen Satz an Basisvektoren repräsentiert. Zu einem gegebenen Stimulus sind aber nur wenige Neuronen gleichzeitig aktiv. Man spricht dann von einer spärlicher Kodierung (*sparse coding*). Als Maß der Spärlichkeit (*sparseness*) einer Verteilung kann zum Beispiel die Kurtosis, also die vierte Kumulante der Aktivitätsverteilung verwendet werden. Die Bestimmungsgleichung lautet hier zu einem einzelnen präsentierten Stimulus

$$K = \frac{\mathbb{E} [a_j^4 - \mathbb{E} [a_j]_{\mathcal{N}}^4]_{\mathcal{J}}}{\mathbb{E} [(\mathbf{a}_j - \mathbb{E} [\mathbf{a}_j]_{\mathcal{J}})^2]_{\mathcal{J}}^2} - 3, \quad (2.11)$$

wobei \mathcal{J} für die Gesamtzahl der Übertragungskanäle j steht. Im Falle einer gaußverteilten Aktivität unter einem gegebenen Stimulus ist die Kurtosis gleich Null.

Ein zweiter Ansatz ist, dass die neuronale Verarbeitung darauf abzielt, die Transinformation

$$I = \sum_x \sum_y p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right) \quad (2.12)$$

zwischen dem Antwortverhalten zweier Neurone zu minimieren, wobei x, y die Aktivität der beiden betrachteten Neurone und $p(\cdot)$ die Wahrscheinlichkeitsverteilung zu einem gegebenen Stimulusset darstellt.

Der Stimulus soll so kodiert werden, dass jeder Kanal des Ausgangs unabhängig von den anderen antwortet. Solche Verfahren werden als Quellentrennung bezeichnet. Geht man

von einer linearen Kodierung aus, so spricht man von einer *independent component analysis* (ICA).

Es sind verschiedene Verfahren zur ICA bekannt [Hyvärinen & Oja, 1997, Cardoso & Laheld, 1996, Bell & Sejnowski, 1995]. Alle führen zu quasi identischen Ergebnissen bei natürlichen Bildern und ähneln den Resultaten der spärlichen Kodierung. Allerdings lässt sich ein Stimulusensemble im Allgemeinen nicht in unabhängige Quellen zerlegen, so dass die Ergebnisse dieser Algorithmen im Falle von natürlichen Bildern weit davon entfernt sind, wechselseitig unabhängige Kanäle zu liefern.

Eine der vierten Kumulante (vgl. Gl (2.11)) sehr ähnlich Funktion ist

$$K_2 = \mathbf{E} [a_j^4 - \mathbf{E} [\mathbf{a}_j]_{\mathcal{J}}^4]_{\mathcal{J}} - 3\mathbf{E} [(\mathbf{a}_j - \mathbf{E} [a_j]_{\mathcal{J}})^2]_{\mathcal{J}}^2. \quad (2.13)$$

Sie wird in der Literatur Kurtosis 2 genannt und entspricht der ersten Formulierung im Falle einer normierten Varianz.

Es lässt sich zeigen [Blais et al., 1998], dass falls eine Quellentrennung vollständig möglich ist, auch die Kurtosis 2 der einzelnen Kanäle ihr Maximum erreicht, sobald die unabhängige Darstellung des Stimulusets gefunden wurde. Damit besteht auch formal ein Zusammenhang zwischen spärlicher Kodierung und ICA.

Kapitel 3

Formation rezeptiver Felder in Abhängigkeit der visuellen Erfahrung

Bis hierher wurde referiert, dass die rezeptiven Feldformen, wie sie im visuellen Kortex gefunden werden, nach verschiedenen informationstheoretischen Kriterien als optimal für die Verarbeitung von natürlichen Bildern angesehen werden können. Das allein ist aber noch kein Hinweis darauf, dass es sich im Kortex um einen stimulusgetriebenen Selbstorganisationsprozess handelt. Ebenso gut könnten, als Ergebnis der Evolution, gerade die Gene überlebt haben, die solche kortikalen rezeptiven Felder hervorbringen. So ist auch die Entwicklung der menschlichen Hand eher als genetisch vorgegeben anzusehen, denn als Resultat eines Selbstorganisationsprozesses auf der Basis des Umgangs mit Stöcken und Steinen.

Um die Rolle der visuellen Erfahrung aus der kortikalen Strukturbildung herauszufiltern, müssen sich Modelle, die die Auswirkung natürlicher Bilder bzw. die Auswirkungen ihres Fehlens betreffen, an entsprechenden Experimenten am biologischen Vorbild orientieren. Als natürliche Grundlage sind daher Vergleichsuntersuchungen an binokulär deprivierten (BD) Tieren und solchen, denen eine normale Entwicklung ermöglicht wurde, zu sehen.

3.1 Befunde von Vergleichsuntersuchungen von BD und normal entwickelten Tieren

Unter binokulärer Deprivation wird in der Praxis die Aufzucht in permanenten Dunkelheit oder die Vernäherung der Augenlider verstanden. Das bevorzugte Versuchstier ist die Katze [Smith et al., 1980, Raczkowsky et al., 1988, Ohzawa & Freeman, 1988, Crair et al., 1998].

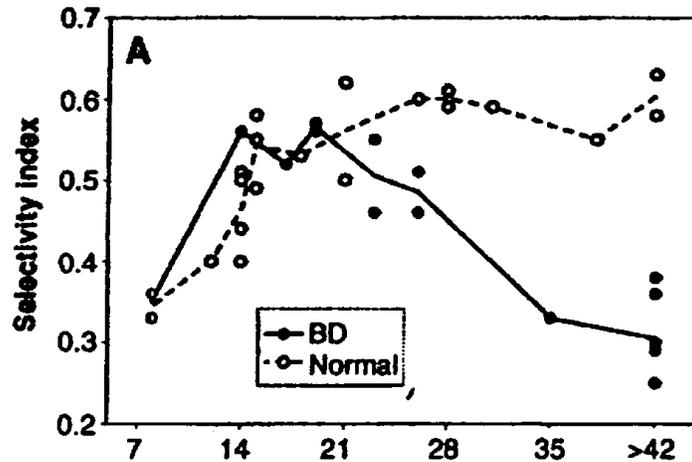


Abbildung 3.1: Entwicklung der Orientierungsselektivitäten in normalsichtigen und binokulär deprivierten Katzen aus [Crair et al., 1998].

Es zeigt sich, dass die Orientierungsselektivität sowohl bei binokulär deprivierten, als auch bei normalsichtigen Tieren zunächst anwächst. Erst später entwickeln sich die Orientierungsselektivitäten in diesen beiden Gruppen unterschiedlich: Bei den binokulär deprivierten Tieren fallen die mittleren Orientierungsselektivitäten wieder ab und gehen gegen Ende der kritischen Phase gegen Null. Bei den normalsichtigen Katzen konvergiert die Orientierungsselektivität gegen ein Lebenszeitmaximum. Einzelableitungen im Kortex [Ohzawa & Freeman, 1988] zeigen, dass als Folge binokulärer Deprivation Neurone entweder komplett insensitiv gegenüber visueller Stimulierung oder unzuverlässig werden oder dass die verbleibenden Neuronen ein stark gestörtes Orientierungs- und Frequenztoning zeigen.

Offensichtlich haben die rezeptiven Felder binokulär deprivierter Tiere andere rezeptive Feldeigenschaften als normalsichtige. Im Gegensatz zu einem rezeptiven Feld einer normalsichtigen Katze (Abb. 3.3 A) wird deshalb [Ohzawa & Freeman, 1988] für BD Katzen ein rezeptives Feld angenommen, wie es in Abb. 3.3 B skizziert ist.

Auf der anderen Seite zeigen sich bei Untersuchungen von Tieren vor dem normalen Öffnen der Augen (z.B. bei Katzen [Hubel & Wiesel, 1963, Albus & Wolf, 1984]) oder sogar in den embryonalen Entwicklungsstufen (bei Affen [Wiesel & Hubel, 1974]) Orientierungsselektivitäten und -karten im visuellen Kortex. Ein biologisch relevantes Modell sollte daher sowohl die anfängliche vorhandene Orientierungsselektivität aufgrund des unspezifischen Inputs, als auch den späteren Abfall bei der Beibehaltung des unspezifischen Inputs im Unterschied zu einem weiteren Ansteigen im Fall von natürlichen Bildstatistiken erklären können.

Die erste Aufgabe wurde von Miller [Miller, 1992, Miller, 1994] gelöst. Ihm gelang es, Ori-

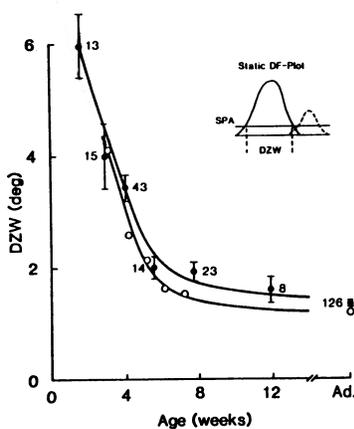


Abbildung 3.2: Die Entwicklung der Wellenlänge der retinalen und kortikalen rezeptiven Felder aus [Braastad & Heggelund, 1985]: Hier dargestellt ist die Weite des exzitatorischen Bereichs bei *ON-center* bzw. des inhibitorischen Bereichs bei *OFF-center* Zellen in der Netzhaut (Kreise). Diese Bereiche werden hier *dominant discharge field zone (DZ)* genannt. Demgegenübergestellt sind die entsprechenden *DZ*-Weiten der (orientierten) kortikalen Neurone (Punkte). Die Zahlen geben die Anzahl der untersuchten Neurone wieder.

entierungselektivitäten aus unspezifischem Input durch bestimmte lokale Korrelationen darzustellen. (Weil das im Folgenden dargestellte Modell dieses Ergebnis als Spezialfall enthält, soll auf eine weitere Erläuterung an dieser Stelle verzichtet werden.)

Eine wichtige Fragestellung ist nun, wieso einerseits die Erfahrung natürlicher Bilder wichtig für das Aufrechterhalten der Orientierungselektivität kortikaler Neurone ist, wenn sie andererseits für deren Entstehung nicht unbedingt notwendig sind.

Für ein Modell, das zum Verständnis dieses Sachverhaltes beitragen könnte, wurde ein weiterer ein retinaler Entwicklungsprozess ins Kalkül gezogen. Aus Untersuchungen zum retinalen Bandpassfilter (Katze [Rusoff & Dubin, 1977], Mensch [Norcia & Tyler, 1985]) geht hervor, dass sich die Wellenlänge des retinalen Bandpassfilters verändert. Dies wurde zumindest für die Katze direkt in der Retina [Rusoff & Dubin, 1977], als auch indirekt beim Menschen [Norcia & Tyler, 1985] festgestellt. In beiden Fällen ergab sich eine Reduzierung der Filterwellenlänge etwa um den Faktor sechs. Diese retinalen Veränderungen werden in Übereinstimmung mit Ableitungen in der Retina [Raczowsky et al., 1988] als von der Erfahrung unabhängig angenommen.

Zudem soll die anatomische Begrenzung der Enervierung des Gesichtsfeldes durch ein einzelnes kortikales Neuron in das Modell einbezogen werden. Ein einzelnes Neuron kann sein rezeptives Feld nur in dem Bereich des visuellen Feldes entwickeln, der durch die Ausdehnung seines Dendritenbaumes erreichbar ist. Das bedeutet, dass auch der Ausschnitt des Gesichtsfeldes für den ein einzelnes Neuron sensitiv werden kann, begrenzt ist. Die Weite

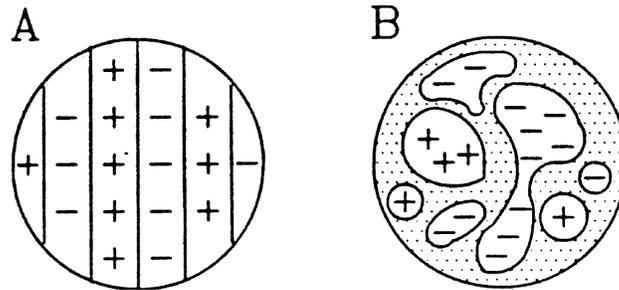


Abbildung 3.3: Aufgrund von Vergleichsuntersuchungen von Experimentatoren vermuteten rezeptiven Felder von normalsichtigen (A) und binokulär deprivierten Katzen (B) (aus [Ohzawa & Freeman, 1988]).

des Ausschnittes, den das Neuron im visuellen Feld hat, bezeichnet man als Arborweite.

Im Folgenden wird eine Gaußverteilung für die Wahrscheinlichkeitsverteilung der anatomischen Begrenzung der kortikalen Innervierung im visuellen Gesichtsfeld angenommen. Die Annahme ist nun, dass diese Weite sich nicht wesentlich ändert. Die biologischen Daten [Lübke & Albus, 1989, Kakabadze, 1985] sprechen dafür, dass ein Teil der Dendritenbäume postnatal noch im Wachstum begriffen ist. Dies trifft allerdings eher für eine Minderheit zu, während für die Mehrheit der Dendriten dieser Wachstumsprozess abgeschlossen ist. Es findet sogar eine optisch sichtbare Differenzierung in dieser Phase statt. In keinem Fall wurde jedoch eine Schrumpfung der Arborweite festgestellt, die eine ähnliche Größenordnung hätte, wie die der retinalen Filterweite. Die Annahme, daß sich das Verhältnis aus Filterwellenlänge zu Arborweite stark ändert, ist für das folgende Modell von entscheidender Bedeutung.

In den folgenden Abschnitten wird ein Modell der Formation rezeptiver Felder entwickelt, das mit Hilfe der obigen Annahmen sowohl die Entstehung der Orientierungsselektivität unter der Abwesenheit visueller Stimulierung erklären kann, als auch den späteren Abfall, falls die Deprivation andauert. Zunächst wird auf die im Modell verwendete retinale Filterfunktion eingegangen. Es schließt sich dann eine Beschreibung der unterschiedlichen Stimulierungen an. Dabei wird diskutiert, welche Stimulusstatistiken den verschiedenen Aufzuchtbedingungen zuzuordnen sind. Schließlich soll ein einfaches Modell für den kortikalen Lernprozess vorgestellt werden und seine Relevanz gegenüber dem biologischen Vorbild diskutiert werden.

3.2 Modellvoraussetzungen: Der retinale Filter

Wie bereits erläutert, lassen sich im Antwortverhalten die retinalen Ganglienzellen in zwei Gruppen unterteilen: in *on-center* und *off-center* Zellen. Das Zusammenwirken dieser Zell-

typen kann (siehe Abschn. 2.3) als linearer Bandpassfilter angesehen werden, dessen Wirkung grob gesprochen, die eines Kantendetektors ist [Marr & Hildreth, 1980].

Es wird sich im Verlauf der Untersuchungen zeigen, dass die exakte Form des Filters einen entscheidenden Einfluss auf die Statistik der rezeptiven Feldformen hat. Da die biologische Situation nicht eindeutig ist, sollen zwei variante Filterformen untersucht werden. Ein denkbarer Filter ist [Olshausen & Field, 1996]

$$f(k) = k \exp\left(-\left(k/k_0\right)^4\right). \quad (3.1)$$

Ein anderer möglicher Filter stellt die zweite Ableitung einer zweidimensionalen Gaußverteilung dar

$$f(k) = k^2 \exp\left(-\left(k/k_1\right)^2\right). \quad (3.2)$$

Beide Filter haben im Ortsraum die Form eines Sombreros. Die Filter zeigen für $k_0/k_1 = 1.62$ eine optimale Übereinstimmung bezüglich Maximum, Mittelwert und Varianz.

Der entscheidende Unterschied zwischen diesen beiden Filtern ist der Gradient in der Nähe des Ursprungs (siehe Abb. 3.4). Daraus resultiert – setzt man natürliche Bilder als Reizenensemble voraus – ein unterschiedliches Spektrum. Das der Bilder verhält sich ja im Mittel wie

$$P(k) = 1/k^{2-\eta}.$$

Da der erste Filter um den Ursprung linear ist zeigt das aus Stimulus und Filter resultierende Spektrum eine sensitive Abhängigkeit von η . Vermöge kleiner Schwankungen in η kann sich das Maximum praktisch vom Ursprung bis in die Nähe der maximal transmittierten Frequenz verschieben.

Obwohl die biologische Filterfunktion nicht exakt angegeben werden kann, lassen psychophysische Experimente eher auf die erstere Funktion schließen [Atick & Redlich, 1992].

3.3 Stimulusensembles

Die unterschiedlichen experimentellen Bedingungen sollen von unterschiedlichen Stimulustypen reflektiert werden.

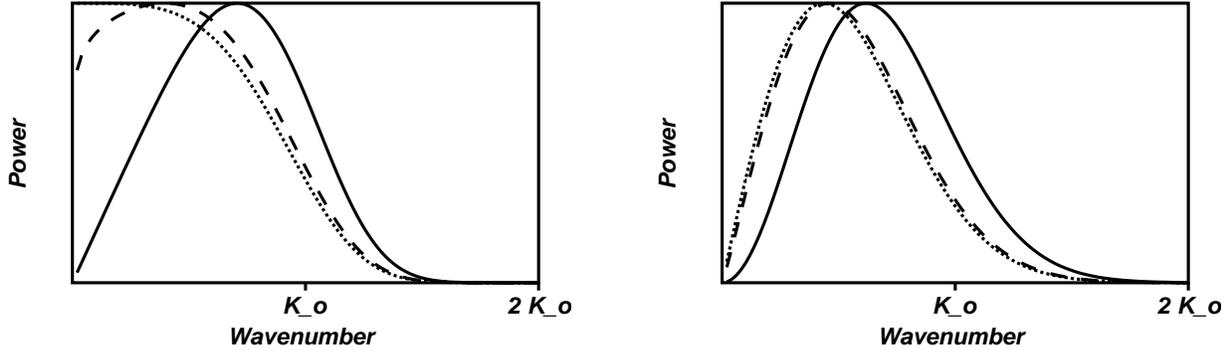


Abbildung 3.4: Illustration der Effekte des retinalen Filters auf die bevorzugte Frequenz der kortikalen rezeptiven Felder. Es wird das prognostizierte Energiespektrum (Amplitudenbetragsquadrate der Fouriertransformierten) verschiedener Stimulussensibles bei verschiedenen Filtern dargestellt. Während der Filter Gl. (3.1) (links) eine starke Abhängigkeit des resultierenden Spektrums von den Eigenschaften des Stimulusspektrums zeigt, ist Filter Gl. (3.2) (rechts) eher indifferent. Es werden das Spektrum der Filter (durchgezogene Linie), das resultierende Spektrum im Falle einer Stimulierung mit $\eta = -0.2$ (gestrichelt), und $\eta = +0.2$ (gepunktet) dargestellt.

Der erste Stimulustyp sind natürliche Bilder. Alle Bilder hatten eine Größe von 1536×1024 Pixeln und wurden jeweils gefiltert in zufällig gewählten 32×32 Pixeln großen, zufällig orientierten Ausschnitten als Inputstimuli präsentiert. Diese Ausschnitte wurden pixelweise mit einer zweidimensionalen Gaußverteilung mit $\sigma_p = 12$ Pixeln multipliziert, deren Ursprung sich in der Mitte des Ausschnittes befand.

Im Falle von binokulärer Deprivation wird eine vorherrschende spontane Aktivität retinaler Ganglienzellen postuliert, die zwar ursprünglich lokal unkorreliert ist, aber der gleichen Filterung unterliegt wie bei den anderen Bildtypen. Es wurde ein gleichverteiltes weißes Rauschen verwendet dessen Mittelwert bei Null lag.

Zusätzlich wurde die Hypothese getestet, dass das algebraische Spektrum natürlicher Bilder nicht ausreichend ist, um die Unterschiede in den Orientierungselektivitäten zu erklären. Zu diesem Zweck wurden Zufallsbilder mit algebraischem Spektrum verwendet. Zunächst wurde dabei ein Zufallsfeld $f(k_x, k_y)$ erzeugt:

$$f(k_x, k_y) = (k_x^2 + k_y^2)^{-\frac{2+\eta}{4}} \exp(2\pi i \gamma(k_x, k_y)), \quad (3.3)$$

wobei γ eine für jedes Tupel (k_x, k_y) eigens ermittelte Zufallszahl war. Dieses Feld wurde in den Ortsraum transformiert. Es wurden zwei Varianten dieser Bilder verwendet, der Fall $\eta = 0.2$ und $\eta = -0.2$.

3.4 Der Adaptionprozess

Da in diesem Kapitel allein rezeptive Feldformen betrachtet werden, sollte auch ein Modell ohne Berücksichtigung der topologische Anordnung, d.h. ohne abstandsabhängige Wechselwirkung zwischen den Neuronen, ausreichen.

Es soll nun die Adaption der kortikalen rezeptiven Felder an die Stimulusmanigfaltigkeit \mathcal{S} simuliert werden. Als Modell für den Adaptionvorgang wird ein Vektorquantisierer verwendet. Der Vektorquantisierer besteht aus einem Satz von 256 rezeptiven Feldvektoren, die jeweils aus 32×32 Pixeln bestehen und im Folgenden $\tilde{\mathbf{W}}_i$ genannt werden. Der Lernvorgang ist

$$\frac{\partial}{\partial t} \tilde{\mathbf{W}}_{i^*} = \gamma (\mathbf{V} - \tilde{\mathbf{W}}_{i^*}), \quad (3.4)$$

wobei $\mathbf{V} \in \mathcal{S}$ ist und i^* , mit

$$i^* = \operatorname{argmax}_i \frac{1}{1 + H(i)} \mathbf{V} \tilde{\mathbf{W}}_i \quad (3.5)$$

bestimmt wird, d.h. nur ein Neuron, der Sieger, darf jeweils lernen. Für eine theoretische Diskussion zu diesem Vektorquantisierungsalgorithmus siehe [Ahalt et al., 1990]. $H(i)$ wird durch die Vorgeschichte des Lernprozesses festgelegt. Es soll die Anpassung auch von solchen Neuronen ermöglichen, die zunächst selten Sieger werden. $H(i)$ wird jedesmal um eins erhöht, wenn das i -te Neuron gewinnt und fällt gleichzeitig bei jeder Iteration um den Faktor 1.001 ab. γ nimmt während der gesamten 2^{16} Iterationen von 0.1 nach 0.003 exponentiell ab. Beide, \mathbf{V} und $\tilde{\mathbf{W}}_i$, werden auf eins normalisiert. Es wurden fünf Läufe von allen möglichen Kombinationen der drei Bildtypen, der beiden Filter und von zehn verschiedenen Werten von k_0 gemacht.

3.5 Rekonstruktion des rezeptiven Feldes

Der Vektorensatz $\tilde{\mathbf{W}}_i$ am Ende des durch Gl. (3.4) definierten Lernprozesses bezeichnet nun die Gewichtung, mit der das Antwortverhalten des kortikalen Neurons auf thalamische *on-* bzw. *off-center* Afferenzen bestimmt wird. Die biologischen Experimente bezogen sich aber auf das Antwortverhalten auf visuelle Reize. Da die visuellen Testreize bei der reversen Korrelationsmethode ebenfalls der retinalen Filterung unterworfen werden, besteht zwischen der internen Repräsentation $\tilde{\mathbf{W}}_i$ und den rezeptiven Feldern, wie sie mit der Analyse festgestellt werden, ein Unterschied (siehe Abb. 3.5), der bei quantitativen Untersuchungen berücksichtigt werden muss. Hierzu wird folgende Überlegung angestellt:

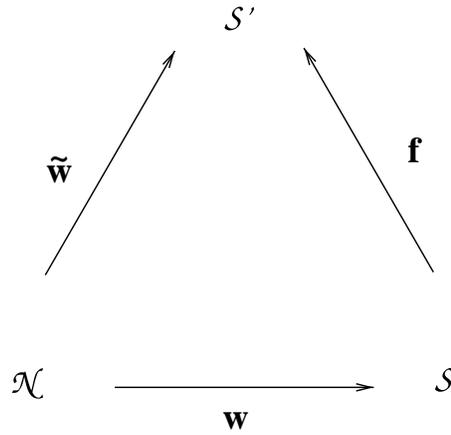


Abbildung 3.5: Eine schematische Darstellung der Stimulus- und Neuronenräume: Der Stimulus erfährt durch den retinalen Filter eine Transformation, die bei der Rekonstruktion des rezeptiven Feldes berücksichtigt werden muss. Der retinale Filter f überführt die Stimulusrepräsentation \mathcal{S} in die modifizierte Repräsentation \mathcal{S}' . Der Lernprozess der $\tilde{\mathbf{W}}_i$ findet auf dieser modifizierten Stimulusrepräsentation statt. Die reverse Korrelation misst jedoch das rezeptive Feld \mathbf{W} , das die Beziehung zwischen Neuronenmanigfaltigkeit und dem ungefilterten Stimulus angibt.

Es soll angenommen werden, dass das Gesichtsfeld des Neurons mit weißem Rauschen stimuliert wird. Das Neuron antworte linear als Funktion seiner genikulaten Afferenzen. Das rezeptive Feld definiert sich analog zum Verfahren der reversen Korrelation als

$$W_\xi = \mathbb{E} [Y_\xi R]_{\mathcal{T}}, \quad (3.6)$$

wobei $Y \in \mathcal{T}$ ein zufälliger Stimulus aus der Gesamtheit der möglichen normierten Vektoren mit der Dimensionalität des Eingangsraumes \mathcal{S} ist. ξ bezeichne die Position eines Pixels im Eingangsraum. Die Antwortfunktion ihrerseits sei

$$R = \int a_{\xi'} \tilde{W}_{\xi'} d\xi', \quad (3.7)$$

wobei a'_ξ den Stimulusvektor durch die genikulaten Afferenzen darstellt. Dies ist nichts anderes als die Stimulierung der Retina gefaltet mit dem Retinalfilter

$$a_{\xi'} = \int Y_{\xi''} f(\xi'' - \xi') d\xi''. \quad (3.8)$$

So ergibt sich für das rezeptive Feld

$$W_\xi = \int \tilde{W}_{\xi'} \mathbb{E} [a_{\xi'} Y_\xi]_{\mathcal{T}} d\xi' = \int W_{\xi'} \int f(\xi'' - \xi') \mathbb{E} [Y_{\xi''} Y_\xi]_{\mathcal{T}} d\xi'' d\xi' \quad (3.9)$$

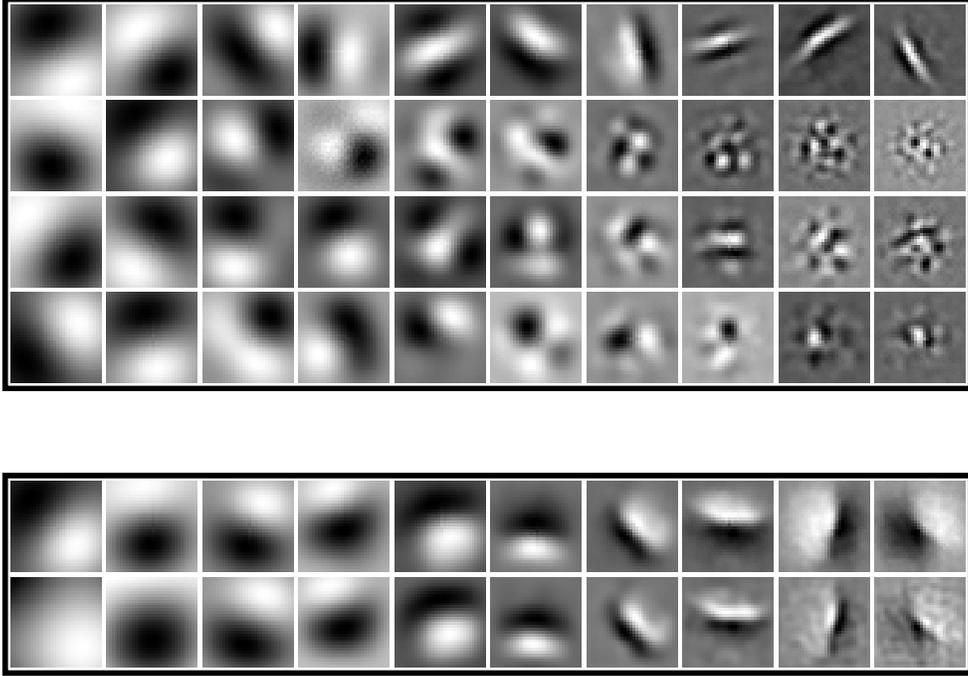


Abbildung 3.6: Oben: Die rezeptiven Felder, wie sie sich für Filter Gl. (3.2) ergeben. Die Reihen stehen für natürliche Bilder, weißes Rauschen, Zufallsbilder mit einem spektralen Exponenten von 1.8 und schließlich solche mit einem Exponenten von 2.2 (von oben nach unten). Die Spalten zeigen Ergebnisse für jeweils $k_0 = 25.0, 50.0, 75.0, 100.0, 150.0, 200.0, 300.0, 400.0, 500.0$ bzw. $600.0 \times 1/2048$ Zyklen pro Pixel (von links nach rechts), bei einem konstanten σ_p von 12 Pixeln bei einer zentrierten zweidimensionalen Gaußglocke, mit der die Stimuli multipliziert wurden. Unten: Ein Vergleich zwischen \mathbf{W} (obere Reihe) und $\tilde{\mathbf{W}}$ (untere Reihe) bei solchen rezeptiven Feldern aus Filter Gl. (3.1) und einer Stimulierung durch natürliche Bilder. Die Spalten stehen auch hier für verschiedene verwendete Filterwellenlängen. Mit dem sich vergrößernden Verhältnis aus Arborweite und spatialer Wellenlänge vergrößern sich auch die Unterschiede zwischen \mathbf{W} und $\tilde{\mathbf{W}}$. Die Inputvektordimension ist 32×32 Pixel. Die anderen Modellparameter entsprechen den in Abschnitt 3.4 gemachten Angaben.

Schließlich erhält man wegen $E[Y_{\xi''}Y_{\xi}]_{\mathcal{T}} = \delta_{\xi''-\xi}$:

$$W_{\xi} = \int \tilde{W}_{\xi'} f(\xi' - \xi) d\xi' \quad (3.10)$$

Daraus folgt, dass sich das rezeptive Feld formal als Faltung der internen Repräsentation und dem retinalen Filter ergibt.

3.6 Numerische Resultate

Die Resultate des Modells sind als Analogon zu rezeptiven Feldern kortikaler *simple cells* zu interpretieren. Da sich, wie zuvor ausgeführt, die retinale Filterwellenlänge reduziert, können verschiedene Wellenzahlen k_0 des Filters zu den verschiedenen Zeitpunkten der frühen postnatalen Entwicklung zugeordnet werden. Die resultierenden rezeptiven Felder (siehe Abb. 3.6) zeigen für die verschiedenen Stimulustypen und Wellenlängen wesentlich unterschiedliche Strukturen.

Die Simulationen zeigen, dass im Bereich großer Wellenlängen alle Stimulusarten zu einer Struktur führen die einem Gaborfilter ähnelt. So erscheint es zumindest plausibel, dass sich solange das Verhältnis zwischen Arborweite σ_p und Wellenlänge $2\pi/k_0$ bzw. $2\pi/k_1$ des Retinalfilters klein ist, auch im biologischen Vorbild rezeptive Felder ausbilden können, die eine gaborfilterartige Struktur haben. Für solche niedrigen Verhältnisse ist es also unwichtig, ob die Stimulierung, wie im Fall der natürlichen Bilder, strukturiert oder, wie im Fall der stochastische Artefakte, unstrukturiert ist.

Dieses Ergebnis korrespondiert mit entsprechenden Resultaten für die rezeptiven Felder aus anderen Modellen [Miller, 1992, Miller, 1994], die schon angesprochen wurden. Auch dort wurde lediglich eine Korrelation zweiter Ordnung verwendet, die letztlich dem hier vorgestellten Fall von weißem Rauschen und retinalem Filter entspricht. Weiterhin wird auch dort eingeräumt, dass die Qualität der Ergebnisse entscheidend von dem Verhältnis aus Arborweite und Filterwellenlänge abhängt. Schon für Änderungen von wenigen Prozenten berichtet Miller von starken Änderungen des Ergebnisses. Obwohl die Details zu diesem Aspekt dort nicht erörtert werden, ist anzunehmen, dass die Art der Veränderungen der Ergebnisse denen entspricht, die hier für größere Verhältnisse aus Arborweite und Wellenlänge im Falle des unspezifische Inputs gefunden werden.

Für größere Verhältnisse zwischen Arborweite und Wellenlänge ($\sigma_p > \lambda$, in diesem Fall für $k_0 > 170$) ergibt sich ein unterschiedliches Bild einerseits für die rezeptiven Felder aus natürlichen Bildern, und andererseits für die aus den anderen Stimulustypen gewonnenen. Im Falle der natürlichen Bilder zeigen sich auch hier rezeptive Felder mit der Ausformung eines Gaborfilters, wobei diese längs ihrer Orientierung gestreckt zu sein scheinen.

Eine andere Arbeit [Law & Cooper, 1994] kann gut mit den Ergebnissen in diesem Bereich verglichen werden. Genauso wie hier, wurde eine Vergleichsuntersuchung von natürlichen Bildern und gefiltertem weißen Rauschen vor dem Hintergrund diverser Deprivationsszenarien durchgeführt. Auch wurde eine Begrenzung der Enervierung durch eine Weite festgelegt, die etwa Wellenlänge der retinalen Filterung entsprach. Das Modell beruht auf einem Ansatz nach der Bienenstock, Cooper und Munro (BCM) Adaptionsregel [Bienenstock et al., 1982]. Auch in diesem Modell wurde ein gradueller Unterschied der Orientierungselektiven festgestellt. Dies entspricht in dem in diesem Kapitel präsentierten Modell den Verhältnissen in etwa bei $k_0 = 200$. Ein Unterschied zwischen dem hier präsen-

tierten Modell und dem BCM-inspirierten ist, dass dort als das wesentliche dynamische Element unterschiedliche Aktivitätslevel für die spontane Aktivität angenommen wurden. Der Frequenzshift der retinalen Filterung, der hier ein entscheidender Faktor ist, wird dort nicht berücksichtigt.

3.7 Orientierungsselektivität

Um den Grad der Orientierungsstärke der resultierenden rezeptiven Felder festzustellen, wurden die Vektoren \mathbf{W}_i einer Fourieranalyse unterzogen. Sie wurden skalar mit komplexen ebenen Wellen verschiedener Wellenlänge λ und Orientierung Φ multipliziert. Die Testorientierungen waren gleichverteilt auf dem Interval $[0, \pi[$. Für die Wellenlänge wurde ein geeignetes Fenster ausgewählt. Es ergibt sich dann eine Matrix komplexer Skalarprodukte $s_{\lambda, \Phi}$. Um die Orientierungsstärke zu ermitteln, bestimmt man zunächst die Antwort auf die optimale Wellenlänge

$$S_{\Phi} = \max_{\lambda}(s_{\lambda, \Phi}) \quad (3.11)$$

und mit

$$\kappa = \frac{1}{2} \log \left(\frac{\sum_{\Phi} S(\Phi) \exp(i2\Phi)}{\sum_{\Phi} S(\Phi)} \right). \quad (3.12)$$

schließlich die Orientierungsstärke

$$Z = \mathbf{E} \left[\exp(\kappa_{\xi} + \kappa_{\xi}^*) \right]_{\xi}, \quad (3.13)$$

wobei κ_{ξ}^* jeweils das komplex Konjugierte zu κ_{ξ} ist. Die finalen rezeptiven Felder (siehe Fig. 3.7) des Vektorquantisierers zeigen unterschiedliche Orientierungsstärken für weißes Rauschen und natürliche Bilder. Außerdem übersteigt die Orientierungsselektivität für natürliche Bilder bei hohen k_0 die Selektivität der beiden anderen Bildtypen, was mit den experimentellen Befunden in Einklang steht.

Bezüglich der Orientierungsselektivität ergeben sich nur relativ kleine Unterschiede zwischen den beiden verwendeten Filtertypen. Im Falle natürlicher Bilder erhält man für den ersten Filter etwas kleinere Orientierungsselektivitäten als für den zweiten. Für die anderen Stimulustypen ergeben sich lediglich im Falle des ersten Filters für große Verhältnisse aus Arborweite und Wellenlänge merkliche Unterschiede in der Orientierungsselektivität.

Auch solche rezeptiven Felder, die aus Zufallsbildern mit einer $(1/k^{2\pm\eta})$ -Statistik resultieren, zeigen im Bereich hoher Verhältnisse deutlich weniger Orientierungsselektivität als die, die aus natürlichen Bildern resultieren, Dadurch wird deutlich, dass Korrelationen

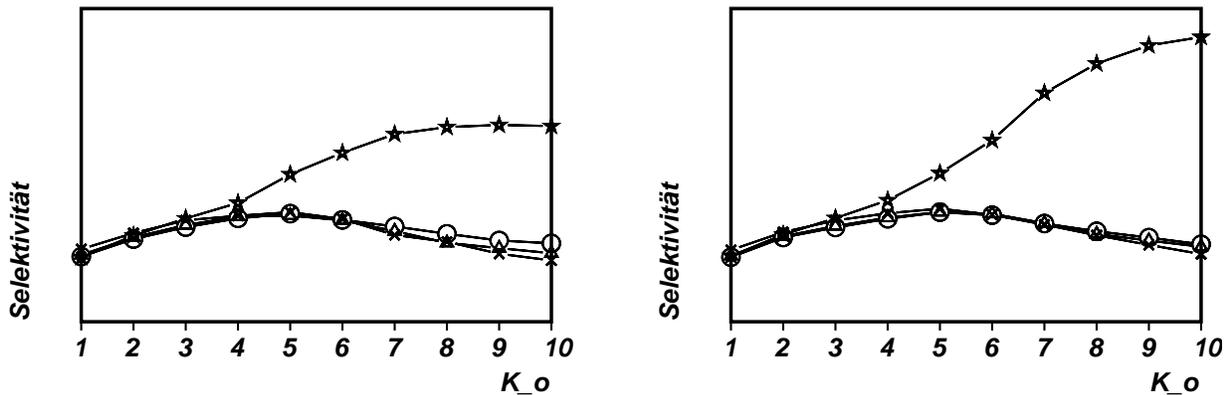


Abbildung 3.7: Orientierungsselektivitätsindex in Abhängigkeit der Arborweite im Fall von Filter Gl. (3.1) (links) und Filter Gl. (3.2) (rechts). Die Stimuli wurden von natürlichen Bildern (Sterne), weißem Rauschen (Kreuze) und Zufallsphasenbildern mit einem $(1/k^{1.8})$ Energiespektrum (Dreiecke) und schließlich von Zufallsphasenbildern mit einer $(1/k^{2.2})$ Statistik (Kreise) genommen. Die Indizierung an der Abszisse entspricht den Parameterwerten der Filterfunktion $k_0 = 25.0, 50.0, 75.0, 100.0, 150.0, 200.0, 300.0, 400.0, 500.0$ bzw. $600.0 \times 1/2048$ und k_1 jeweils $k_1 = k_0/1.62$. Die anderen Modellparameter entsprechen denen aus Abb. (3.6). Das unterschiedliche Verhalten der Filter spiegelt sich in der unterschiedlichen Orientierungsselektivität wieder.

höherer als zweiter Ordnung die Struktur der rezeptiven Feldformen beeinflussen. Es ist davon auszugehen, dass solche Korrelationen von den in natürlichen Bildern vorhandenen Kanten induziert werden. Schließlich ähneln die rezeptiven Felder dieses Modells im Falle von weißem Rauschen und hohen k_0 stark den rezeptiven Feldformen, welche für binokulär deprivierte Tiere in [Ohzawa & Freeman, 1988] vermutet wurden.

3.8 Wellenlängenverteilungen

Besonders gilt es zwischen den Verbindungstärken der kortikalen Neuronen zu den genikulaten Afferenzen $\tilde{\mathbf{W}}_i$ und den rezeptiven Felder \mathbf{W}_i , wie sie durch reverse Korrelation entstehen, zu unterscheiden (vgl. Abschn. 3.5). In den Abb. 3.8 und 3.9 sind deshalb jeweils die Histogramme, die aus der Verteilung der $\tilde{\mathbf{W}}_i$ hervorgehen, gestrichelt wiedergegeben, während die Verteilung der rezeptiven Felder \mathbf{W}_i mit einer durchgezogenen Linie markiert ist. Zudem muss auch zwischen den Simulationen bei verschiedenen Verhältnissen zwischen Arborweite und Wellenlänge unterschieden werden. Abb. 3.9 b) zeigt ein Beispiel für ein kleines Verhältnis (Stimulus: natürliche Bilder). Es lassen sich kaum Unterschiede zwischen den Filtertypen feststellen. Auch der Unterschied zwischen \mathbf{W}_i und $\tilde{\mathbf{W}}_i$ ist sehr gering.

Anders dagegen verhält es sich im Falle großer Wellenzahlen, also großer Verhältnisse zwischen Arborweite und Wellenlänge (Abb. 3.9 b-d). Auf Abb. 3.9 b) ist die Wellenlängen-

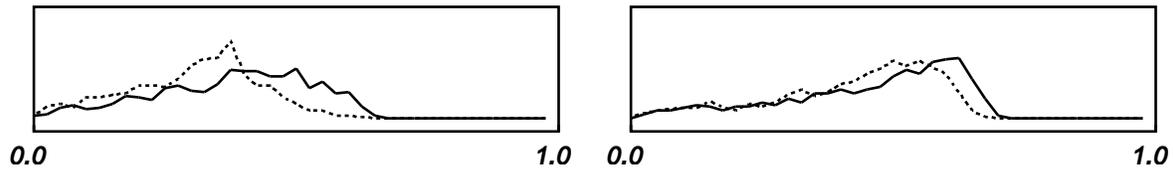


Abbildung 3.8: Die Verteilung der Orientierungsselektivitäten im Falle der Stimulierung durch natürliche Bilder bei $k_0 = 600$, $k_1 = 600/1.62 \times 1/2048$ Zyklen pro Pixel. Zu sehen sind die Ergebnisse für Filter Gl. (3.1) (links) und Filter Gl. (3.2) (rechts), sowohl im Fall der rezeptiven Felder \mathbf{W}_i (Linie), als auch der Afferenzen $\tilde{\mathbf{W}}_i$ (gestrichelt).

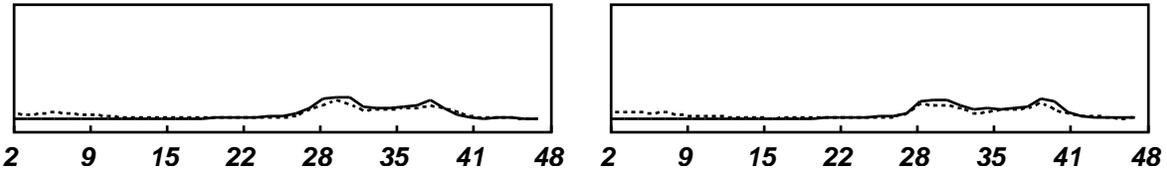
verteilungen bei einem großen Verhältnis und der Stimulierung durch natürliche Bilder zu sehen. Hier ist ein deutlicher Unterschied zwischen den beiden verwendeten Filtern zu erkennen. Während Filter Gl. (3.1) bei natürlichen Bildern eine breite Verteilung an Wellenlängen über mehrere Oktaven zeigt, ergibt Filter Gl. (3.2) eine schmale Verteilung mit einer deutlichen Spitze bei einer etwas über der Wellenlänge des Stimulus liegenden Wellenlänge. Dieser Unterschied ist im Falle der Gewichte der kortikogenikulaten Afferenzen $\tilde{\mathbf{W}}$ deutlicher als im Falle der rezeptiven Felder \mathbf{W} , wie sie sich aus der Rekonstruktion ergeben.

Die Verteilung der resultierenden rezeptiven Felder bei Filter Gl. (3.1) entspricht den Verteilungen, wie sie in physiologischen Untersuchungen im Kortex gefunden werden, eher als Filter Gl. (3.2) (vergleiche [Schiller et al., 1976] für den fovealen Bereich des Areals 17 im Rhesusaffen, und den Grenzbereich zwischen Areal 17 und 18 in der Katze [Holub & Morton-Gibson, 1981]). Auch für den Fall der Zufallsphasenbilder mit algebraischem Spektrum (siehe Abb. 3.9 c) und d) finden sich diese Unterschiede zwischen den Filtern bei einem Exponenten von 2.2 wieder. Im Falle eines Exponenten von 1.8 (siehe Abb. 3.9 d)) ist die Verteilung unabhängig vom Filtertyp spitz und enger begrenzt. Während Filter Gl. (3.2) also nicht wesentlich zwischen einer Verteilung mit einem Exponenten von 1.8 und einer mit einem Exponenten von 2.2 unterscheidet, ergeben sich deutliche Unterschiede bei Filter Gl. (3.1). Die Ergebnisse passen zu den Überlegungen, die in Abschn. 3.2 angestellt wurden.

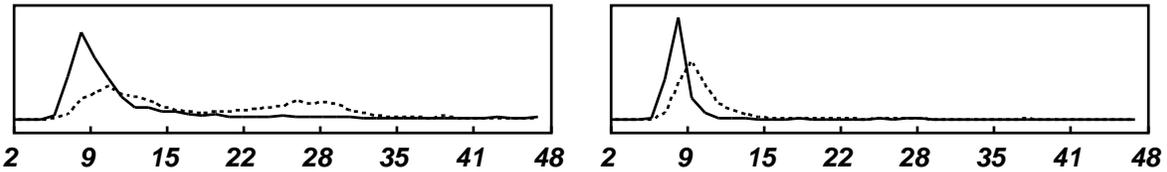
3.9 Ähnlichkeit zu Gaborfiltern

Analog zu den Arbeiten mit reversen Korrelationsmethoden an rezeptiven Feldern [Jones & Palmer, 1987a, Jones & Palmer, 1987b, Jones & Palmer, 1987c] wurde nun die Ähnlichkeit der hier erhaltenen rezeptiven Felder mit zweidimensionalen Gaborfiltern gemessen.

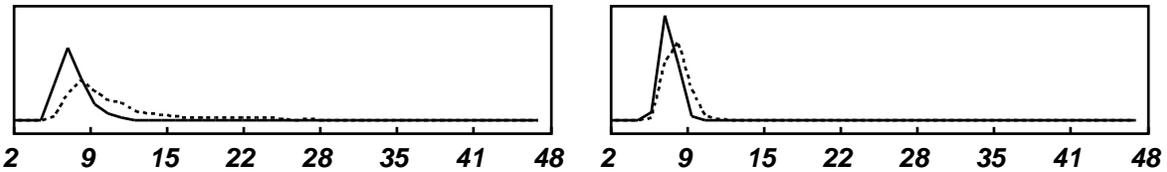
a



b



c



d

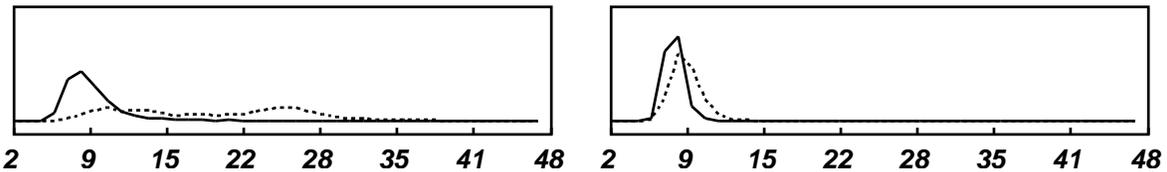


Abbildung 3.9: Histogramme für verschiedene Wellenlängenverteilungen und Filter Gl. (3.1) (links), sowie rechts Filter Gl. (3.2) (rechts). Die anderen Modellparameter entsprechen denen aus Abb. 3.6. Die durchgezogene Linie zeigt die Ergebnisse der Analyse der rezeptiven Felder \mathbf{W}_i , während die gestrichelte Linie jeweils die Analyse der entsprechenden kortikogenikulaten Afferenzen $\tilde{\mathbf{W}}_i$ darstellt. **a)** Die Verteilung der Wellenlängen bei der Stimulierung durch natürliche Bilder. Die Ordinate zeigt die Wellenlängenhistogramme in Pixeln. Die Wellenlängen der Filter betragen hier $k_0 = 75$ bzw. $(k_1 = 75/1.62) \times 1/2048$ Zyklen pro Pixel. Bei den in **b-d)** dargestellten Histogrammen handelt es sich um Filter mit $k_0 = 600$ bzw. $k_1 = 600/1.62 \times 1/2048$ Zyklen pro Pixel. **b)** Die Verteilung der Wellenlängen bei der Stimulierung durch natürliche Bilder. **c)** dito, bei Zufallsbildern mit einem algebraischen Energiespektrum mit einem Exponenten von 1.8. **d)** dito, bei Zufallsbildern mit einem algebraischen Energiespektrum mit einem Exponenten von 2.2.

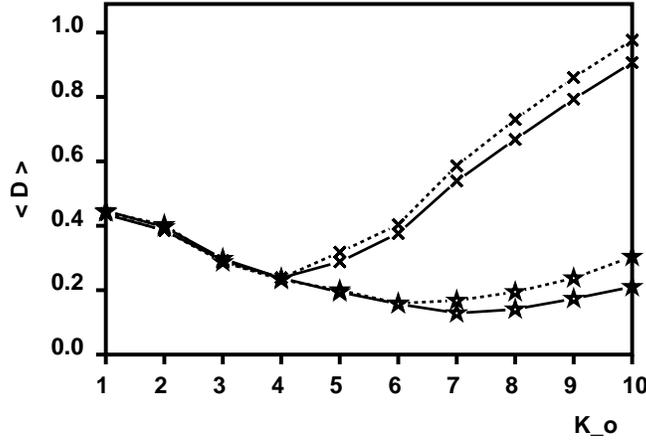


Abbildung 3.10: Mittlere Distanz der rezeptiven Felder zu einem optimalen Gaborfilter in Abhängigkeit der Wellenzahl k_0 . Die Modellparameter entsprechen denen aus Abb. 3.6 .

Ein zweidimensionaler Gaborfilter ist das Produkt

$$\mathbf{G} = G(x, y) = g_w \times g_g, \quad (3.14)$$

wobei

$$g_w(x, y) = \cos \left(\omega \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \hat{T}(\Phi_1) \cdot \begin{pmatrix} x \\ y \end{pmatrix} + \Theta \right) \quad (3.15)$$

und

$$g_g = \exp \left(- \left| \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \cdot \hat{T}(\Phi_2) \cdot \begin{pmatrix} \xi_1 - \bar{\xi}_1 \\ \xi_2 - \bar{\xi}_2 \end{pmatrix} \right|^2 \right) \quad (3.16)$$

mit

$$\hat{T}(\theta) = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}. \quad (3.17)$$

$(\Phi_1, \Phi_2, \Theta, \bar{\xi}_1, \bar{\xi}_2, \omega, b_1, b_2)$ sind freie Parameter. Dabei legen Φ_1 und Φ_2 die Orientierungspräferenz fest. Die Größen Θ, ω sollen im Folgenden lokale oder spatiale Phase bzw. Wellenlänge genannt werden. Diese Art Filter soll so optimiert werden, dass der Abstand

$$d(\mathbf{G}, \mathbf{W}_i) = \frac{|\mathbf{G} - \mathbf{W}_i|}{|\mathbf{G}| |\mathbf{W}_i|} \quad (3.18)$$

minimal wird.

Diese Minimierung wurde mit Hilfe der Simplexmethode [Nelder & Mead, 1965] durchgeführt, einem Algorithmus, der auch dazu verwendet wurde, die Ähnlichkeit zwischen biologischen rezeptiven Feldern und Gaborfiltern zu quantifizieren [Jones & Palmer, 1987c]. Es wurden acht Simplexvektoren verwendet, um die acht Parameter des Gabors zu optimieren. Die Simplexvektoren wurden mit Schätzungen, wie sie sich durch die zuvor gemachten Analysen (Orientierung, retinale Position) ergaben, initialisiert; dort, wo Schätzungen fehlten, wurden sie um Standardwerte gestreut.

Der Mittelungsfaktor σ_{SI} des Algorithmus wurde auf 0.8 der Spiegelungsfaktor γ_{SI} auf 1.5 festgesetzt. Die Optimierung wurde abgebrochen, sobald die Varianz der auf eins normierten Simplexvektoren im Durchschnitt kleiner war als 10^{-5} . Der Algorithmus erwies sich als äußerst robust.

Im Falle von kleinen k_0 liegen viele unentwickelte rezeptive Felder vor. Das führt dazu, dass die mittlere Distanz zwischen Stimulus und der nächsten Gaborfunktion dort hoch (siehe Abb. 3.10) ist. Sie fällt zunächst für ansteigendes k_0 bei beiden Stimulustypen ab, da die mittlere Zahl unentwickelter rezeptiver Felder gegen Null abfällt. Für hohe Werte von k_0 findet man wieder ein gegensätzliches Verhalten für rezeptive Felder aus natürlichen Bildern und solchen aus weißem Rauschen. Während bei weißem Rauschen die mittlere Distanz mit ansteigendem k_0 wieder ansteigt, geht sie bei den natürlichen Bildern gegen einen minimalen Wert. Dieses Verhalten ist nicht unerwartet, wenn man sich die Struktur der rezeptiven Felder selbst wieder in Erinnerung ruft (siehe Abb. 3.6). Auch dort wurde deutlich, dass der wesentliche strukturelle Unterschied zwischen den Stimulustypen der Grad der Ähnlichkeit zu Gaborfiltern ist.

Für die weiteren Untersuchungen wurden nur solche rezeptiven Felder verwendet, die einen kleineren Abstand als 0.5 zum nächsten Gaborfilter hatten. Diese Bedingung wurde je nach Wellenzahl des Filtes von mehr als 43% der rezeptiven Felder erfüllt. In dieser Gruppe wurde das Verhältnis von r_1 bzw. r_2 zu der Wellenlänge $\lambda = 2\pi/k$ gemessen. Es zeigt sich, dass für solche rezeptiven Felder diese Verhältnisse die größten Werte bei großen k_0 (siehe Abb. 3.11) annimmt. Auch die normale kortikale Entwicklung bedingt also strukturelle Veränderungen in der Gestalt der rezeptiven Felder, obwohl ihr prinzipielles Aussehen weiterhin dem von Gaborfiltern entspricht.

Die Tabelle 3.1 enthält Mittelwerte, die sich für biologischen Neuronen [Jones & Palmer, 1987c] extrahieren lassen und vergleicht diese mit den entsprechenden Werten, die sich aus dem Modell für die verschiedenen Filtertypen bei $k_0 = 600.0$ ergeben. Ein Blick auf die in Abb. 3.11 dargestellten Verhältnisse für andere Modellparameter macht deutlich, dass große Werte für $E[r_1/\lambda]$ bzw. $E[r_2/\lambda]$ nur bei großen Verhältnissen zwischen Arborweite und Wellenlänge erzielt werden können. Die in [Jones & Palmer, 1987c] gemessenen reifen rezeptiven Felder zeigen (siehe Tabelle 3.1) ebenfalls große Verhältnisse zwischen $E[r_1/\lambda]$ bzw. $E[r_2/\lambda]$. Sie tragen also die Signaturen einer späten Entwicklung, die durch den Einfluss natürlicher Bilder geprägt ist.

Filtertyp	E $[r_1/\lambda]$	E $[r_2/\lambda]$
Gl. (3.1)	0.77 ± 0.02	0.43 ± 0.01
Gl. (3.2)	0.93 ± 0.02	0.62 ± 0.01
[Jones & Palmer, 1987c]	0.82 ± 0.07	0.48 ± 0.06

Tabelle 3.1: Vergleich zwischen den rezeptiven Feldern des Modells unter den Parametern von Abb. 3.6, sowie einer Filterung von $k_0 = 600.0$ bzw. $k_1 = 600.0/1.62$. Die angegebene Standardabweichung ist der mittlere Fehler.

3.10 Diskussion

Der Anspruch der hier vorgestellten Modelle ist, allgemeine Prinzipien der postnatalen Entwicklung der kortikalen rezeptiven Felder wiederzugeben. Die hier erhaltenen Resultate entsprechen im Wesentlichen den Entwicklungswerten im biologischen Subjekt unter verschiedenen Aufzuchtbedingungen. Sowohl das initiale Ansteigen unter allen vier verwendeten Stimulierungsszenarien als auch der folgende Abfall bei der unstrukturierten Stimulierung entsprechen den biologischen Befunden. Auch die detaillierteren Untersuchungen der rezeptiven Felder im Modell im Vergleich mit den Eigenschaften, wie sie im visuellen Kortex gefunden werden, stimmen sehr gut überein. Es lassen sich sogar Kriterien festmachen, an denen aus der Form der rezeptiven Felder entschieden werden kann, ob strukturierte Stimulierung, also Stimulierung durch natürliche Bilder, an der Formation des rezeptiven Feldes beteiligt war, oder ob der Prozess in einem intermediären Stadium, in dem auch noch unstrukturierter Input zu orientierungsselektiven rezeptiven Feldern führen kann, zum Stehen kam. Der Vergleich mit biologischen Daten zeigt, daß diese Lernprozesse bis in die Phase visueller Erfahrung andauern. Darüberhinaus können verschiedene Rückschlüsse über die Ursachen dieser unterschiedlichen Entwicklung gezogen werden.

Hierbei sind die Ergebnisse der Stimulierung durch Inputs mit Zweipunktkorrelationen hilfreich, die der Statistik natürlicher Bilder entsprechen. Es zeigt sich, daß diese Statistik nicht ausreicht, um die Orientierungsselektivität in den späteren Entwicklungsstadien zu erhalten. Das Entstehen der Orientierungsselektivität ist deshalb als eine Konsequenz höherer Korrelationen in natürlichen Bildern zu verstehen.

Die Modellannahmen sind eigentlich schlicht. Es wurde versucht, die Entwicklung rezeptiver Felder als Prozess darzustellen, der einerseits durch intrakortikale laterale Inhibition und andererseits durch Adaption an den Stimulusraum bestimmt wird.

Da sich das verwendete Modell im Hinblick auf die rezeptive Feldentwicklung als hinreichend zum Verständnis der hier dargelegten Prozesse erwiesen hat, kann geschlossen werden, dass andere Eigenschaften des Kortex, wie seine Topographie für ein entwicklungstheoretisches Verständnis der rezeptiven Feldbildung nicht essentiell sind.

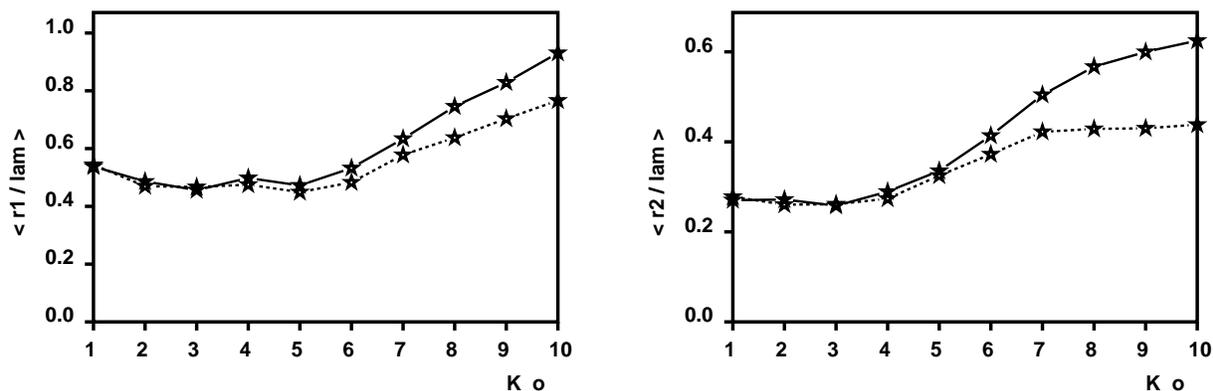


Abbildung 3.11: Die Verhältnisse zwischen r_1 (links) bzw. r_2 (rechts) zur Wellenlänge der rezeptiven Felder als Funktion von k_0 bei natürlichen Bildern und Filter Gl. (3.1) (Linie), (3.2) (Striche). Für Parameter des Modells, sowie der Gaboranalysen, siehe Abb. 3.10.

3.11 Zusammenfassung

Hier wurde ein einfaches Modell vorgestellt, das auf einem Vektorquantisierer basiert. Um die kortikale Entwicklung während der ersten Lebenswochen rekonstruieren zu können, wurde berücksichtigt, dass die Wellenlänge der retinalen Filterung während dieser Phase essentiell abnimmt. Durch eine Änderung des Verhältnisses zwischen retinaler Wellenlänge und des Bereichs der kortikalen Innervierung kann sowohl die Entstehung von Orientierungselektivität als auch deren Verschwinden im Falle unkorrelierter, spontaner Aktivität der retinalen Ganglienzellen reproduziert werden. Rezeptive Feldformen, die durch Stimulierung mit natürlichen Bildern hervorgebracht werden, sind in beiden Fällen orientierungselektiv und haben die Gestalt von zweidimensionalen Gaborfiltern. Allerdings weisen die rezeptiven Felder jeweils bestimmte Unterschiede auf. Die im reifen Kortex gemessenen rezeptiven Felder zeigen genau die Eigenschaften, die den rezeptiven Felder des Modells entsprechen, die bei großen Wellenzahlen entstanden sind. Gleichzeitig erweist sich die Zweipunktkorrelation natürlicher Bilder als nicht ausreichend für das Entstehen von Orientierungspräferenzen. Die Kanten, die sich als die signifikanten Komponenten natürlicher Bilder erweisen, ergeben dagegen die Orientierungselektivität auch für große Verhältnisse zwischen Arborweite und retinalem Bandpassfilter.

Kapitel 4

Neuronale Karten im visuellen Kortex

In diesem Kapitel wird ein kurzer Überblick über die biologischen Fakten zur Entwicklung kortikaler Merkmalskarten nachgereicht.

4.1 Biologische Fakten

Versuche, bei denen die Ableitungselektroden längs der Kortexoberfläche verschoben werden, machen klar, dass einige der rezeptiven Feldeigenschaften benachbarter Neurone im visuellen Kortex ähnlich sind [Hubel & Wiesel, 1962]. So zeigen Messungen in Abhängigkeit der kortikalen Position einen kontinuierlichen Wechsel der präferierten Orientierung (siehe Abb. 4.1). Diese Beobachtung lässt den Schluss zu, dass bestimmte Reizeigenschaften auf den visuellen Kortex zumeist kontinuierlich in dessen Oberflächenkoordinaten abgebildet werden.

Bewegt man sich auf der Kortexoberfläche, so ergibt sich ein fließender Übergang in der Repräsentation bestimmter Stimulusparameter. Damit lässt sich nicht nur einem bestimmten Neuron ein rezeptives Feld zuordnen, sondern allen Neuronen, die sich an einer bestimmten kortikalen Position befinden. Wenn vorkommende Abweichungen (*scatter*) vernachlässigt werden, lassen sich diese rezeptiven Feldeigenschaften als Funktionen der Koordinaten der Kortexoberfläche definieren. Die Abbildungen der kortikalen Oberfläche in die Merkmalsräume der Stimulusmanigfaltigkeit werden kortikale Merkmalskarten genannt. Eine mögliche bildliche Darstellung dieser Eigenschaften ist die Kodierung mit Farben oder Grauwerten auf der Kortexoberfläche.

Mit Hilfe verschiedener bildgebender Verfahren eröffnet sich die Struktur der Reizrepräsentation.

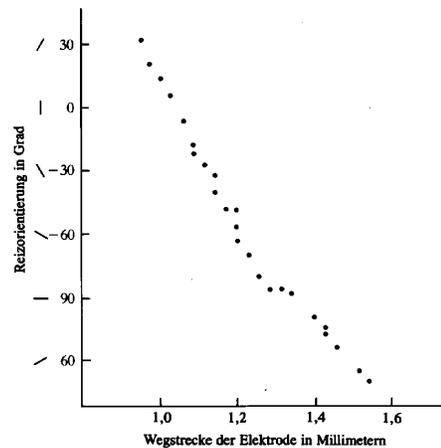


Abbildung 4.1: Die präferierte Reizorientierung bei einer sehr schrägen Penetration des Kortex. Es zeigt sich, dass sich die Orientierungspräferenz der kortikalen Neurone kontinuierlich als Funktion der Wegstrecke ändert. Aus [Hubel, 1990]

tation auf der Kortexoberfläche experimentell. Eine Methode ist die 2-Deoxyglukose (2-DG) Autoradiographie. Hier wird diese Verbindung, die mit ^{14}C markiert ist, injiziert. Sie gelangt in die Zelle, kann dort aber nicht abgebaut werden. In der Folge reichert sich der Stoff in Abhängigkeit der Zellaktivität an [Sokoloff et al., 1977]. Der dadurch entstehende Nachteil der Methode ist, dass sich an einem Tier nur eine Messung durchführen lässt.

Unter anderem kann mit Hilfe der 2-DG Methode gut gezeigt werden [Tootell et al., 1982], dass die Positionen der Neurone im Kortex die Position des rezeptiven Feldes im Wesentlichen topographisch widerspiegeln (siehe Abb. 4.2). Die Wiedergabe des visuellen Feldes auf der Kortexoberfläche behält die retinale Anordnung in etwa bei. Ein Neuron, das sich im visuellen Kortex unten links befindet wird sein rezeptives Feld oben rechts im visuellen Feld haben, respektive ein Neuron mit einem rezeptiven Feld oben rechts unten links. Dieses Phänomen nennt man Retinotopie.

Weiterhin steht von der experimentellen Seite her noch die sogenannte *optical imaging* Methode zur Verfügung. Hier wird ausgenutzt, dass sich die spektralen Eigenschaften des Hirngewebes in Abhängigkeit seiner neuronalen Aktivität schwach ändert [Grinvald et al., 1986]. Am lebenden Tier wird die Kortexoberfläche durch ein Fenster für eine Kamera zugänglich gemacht. Durch geringfügige Kontraste zwischen aktivem und inaktivem kortikalem Gewebe können diese in einer Analyse voneinander getrennt werden.

In neuester Zeit ist auch die funktionelle Kernspintomographie in der Lage die hier relevanten Strukturen darzustellen [Kim et al., 2000]. Hier liegen allerdings zur Zeit nur unsichere, wenn auch perspektivisch hochinteressante Ergebnisse vor.

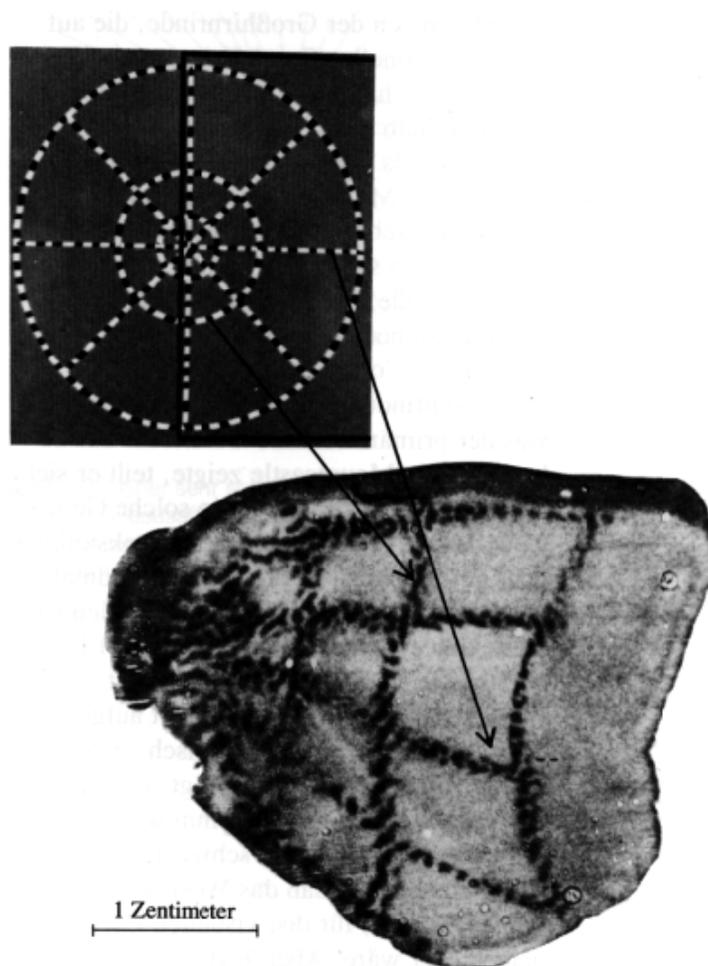


Abbildung 4.2: Wie eine Dartscheibe in V1 von Makakken in neuronale Aktivität umgesetzt wird (aus [Tootell et al., 1982]). Einem Makakken wurde 2-DC injiziert und dem Tier ein flackernder Stimulus in Form einer Dartscheibe präsentiert. Das angeregte neuronale Gewebe verbraucht mehr Zucker, wodurch sich der radioaktive, unverdauliche Agent anlagert.

Eine Orientierungskarte kann am besten mit der *optical imaging* Methode ermittelt werden. Hierbei werden statt der Zufallsmuster der reversen Korrelation triftende Streifenmuster (sog. *gratings*) verwendet. Das endgültige Orientierungsmuster wird dann mit Hilfe der vektoriellen Summation von kortikalen Aktivitätsmustern zu verschiedenen Orientierungen ermittelt (siehe Abb. 4.3). Für die Darstellung dieser Karten hat sich die Kodierung jeder einzelnen Orientierung mit einer Farbe auf dem Farbkreis als geeignet erwiesen.

Ein typisches Merkmal dieser Orientierungskarten sind Punktsingularitäten, die wegen ihres Aussehens im Farbkode *pinwheels* (Windrädchen) genannt werden.

Als letzte hier noch zu erwähnende Eigenschaftskarte existiert die Karte der präferierten

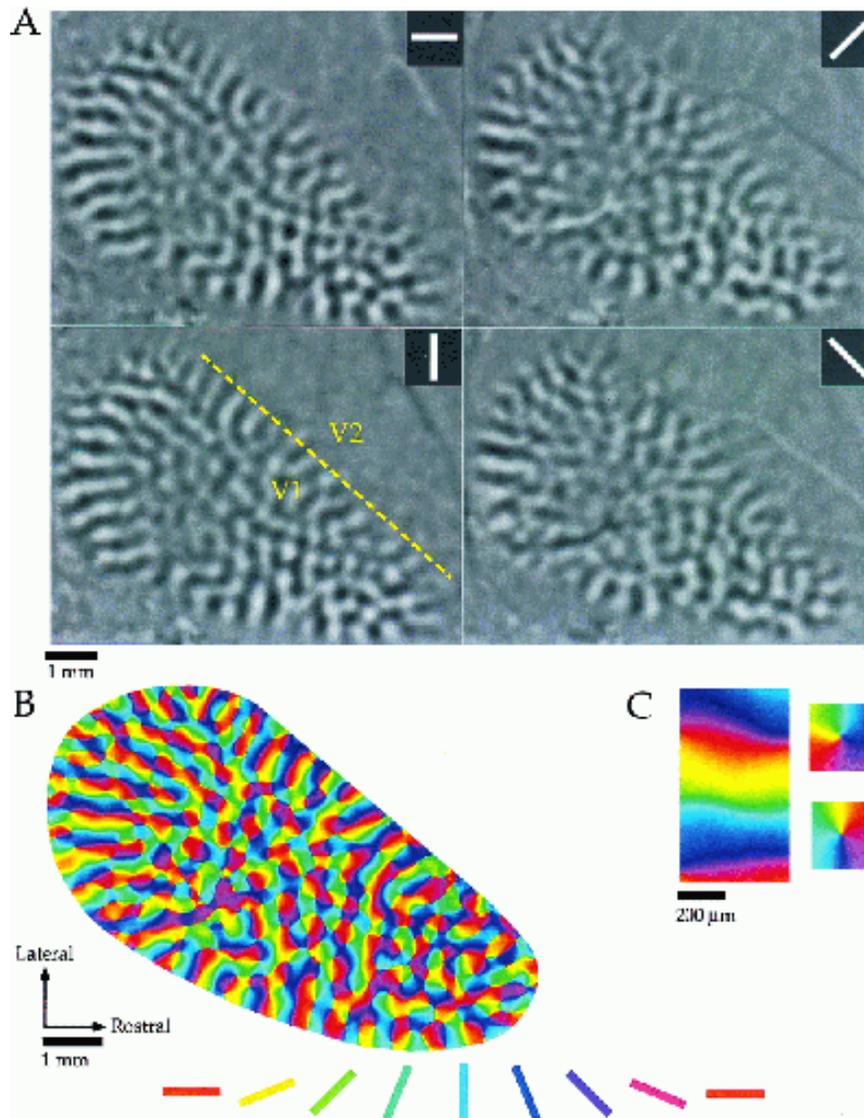


Abbildung 4.3: Beispiel für die Messung einer Orientierungskarte. Hier eines Tupaia aus [Bosking et al., 1997]. In A) sind Differenzbilder zu sehen, wie sie mit Hilfe der *optical imaging* Methode aus den vier Stimuluswinkeln 0° , 45° , 90° und 135° aus einem einzelnen Tier aus V1 gewonnen werden. Die dunklen Gebiete deuten hier auf eine höhere Aktivität hin als die helleren. B) zeigt eine Orientierungskarte, wie sie durch Vektorsummation aus A) erzeugt wird. Darunter befindet sich eine Legende der Farbkodierung. C) zeigt vergrößerte Ausschnitte aus B), einen von einem Gebiet mit relativ parallelen Orientierungsanordnungen (links) und zwei Singularitäten (*pinwheels*) (rechts).

spatialen Wellenlängen (vergleiche Abb. (4.4)). Sie ist schwächer ausgeprägt als die Orientierungskarte. In ihr treten Inseln mit kleinen spatialen Frequenzen [Tootell et al., 1981, Hübener et al., 1997] in einer See von hohen spatialen Frequenzen auf.

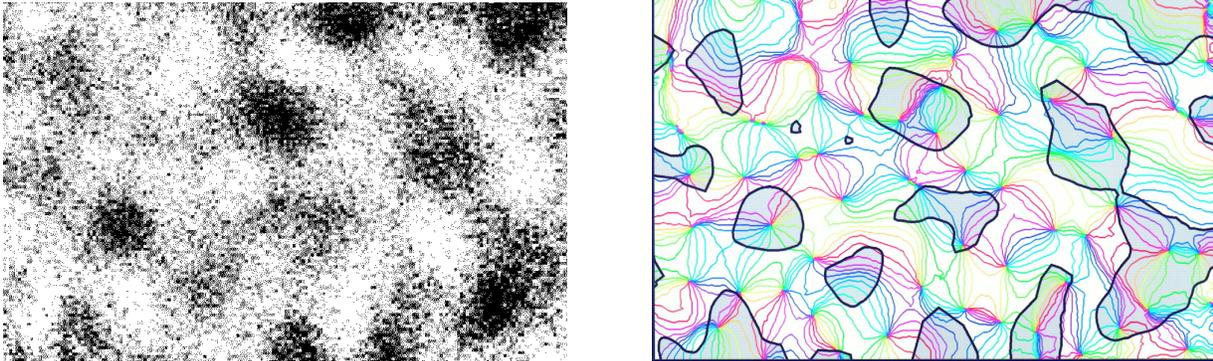


Abbildung 4.4: Karten der präferierten Wellenlänge. Eine Karte zeigt die präferierten Wellenlänge (links). Hell steht für hohe spatiale Frequenzen, dunkel entsprechend niedrige. Rechts: Die selben Daten überlagert mit Isoorientierungslinien. Aus [Hübener et al., 1997]

Im visuellen Kortex zeigt sich eine – abgesehen von Punktsingularitäten – stetige Abbildung der zweidimensionalen Oberfläche (r_1, r_2) auf den Stimulusraum mit diversen Freiheitsgraden (Orientierung, retinale Position, Okulardominanz, etc.).

Diese Nachbarschaftsabhängigkeit der rezeptiven Felder lässt sich durch eine abstandsabhängige Korrelationsfunktion messen. Im Falle der Okulardominanz und auch der präferierten Orientierung ist die Korrelationsfunktion positiv für kurze Abstände und negativ für die weiter entfernt liegenden. In Anlehnung an ihr Aussehen in zwei Dimensionen wird sie auch Mexikanerhutfunktion genannt. Diese Korrelationsfunktion spiegelt dabei die typischen Wellenlängen der Orientierungs- und Okulardominanzkarte wieder, die in beiden Fällen bei etwa 1 mm liegt. Diese abstandsabhängige Korrelation ist ein Phänomen des Hirns und findet sich auch in anderen Hirnregionen ubiquitär wieder.

Eine interessante Frage ist schließlich, in welcher Weise sich die diversen Merkmalskarten beeinflussen. In einer Arbeit [Das & Gilbert, 1997] wurde speziell die Beziehung zwischen Positionskarten und Orientierungskarten gemessen. Es wurde ein beinahe linearer Zusammenhang zwischen Orientierungs- und Positionsgredient auf der Kortexoberfläche gefunden (siehe Abb. (4.5)) also

$$\frac{\partial \Phi}{\partial |r|} \propto \frac{\partial |\xi|}{\partial |r|} \quad (4.1)$$

wobei $\Phi(r)$ die Orientierungs- bzw $\xi(r)$ die Positionspräferenz am Kortexpunkt r darstellen soll. Allerdings ist dieses Ergebnis nicht unumstritten und konnte bisher von anderen Arbeitsgruppen nicht bestätigt werden [White et al., 1999b].

Im folgenden sollen Eigenschaften der kortikalen Karten wie sie sich aus den sogenannten *optical imaging* - Methoden ergeben, untersucht werden. Speziell sollen die Auswirkungen der im vorherigen Kapitel gefundenen rezeptiven Feldformen auf die topographische

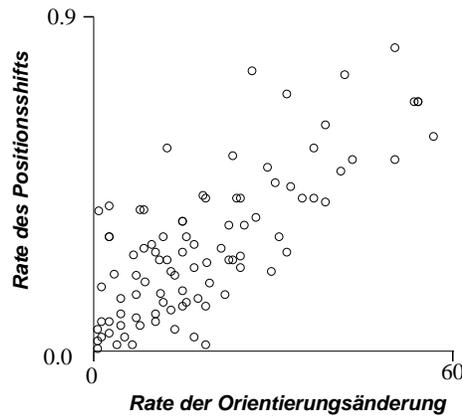


Abbildung 4.5: Die Rate der Änderung des Orientierungsshifts gegen die Rate der Positionsänderung (Nach Daten aus [Das & Gilbert, 1997]).

Struktur ihrer Anordnung geprüft werden.

4.2 Modelle zur Strukturbildung im visuellen Kortex

Kehrt man zurück zur tragenden Prämisse dieser Arbeit, dass nämlich einerseits die Stimulierung in irgendeiner Form für die rezeptive Feldformation im visuellen Kortex relevant ist und andererseits benachbarte Neurone in irgendeiner Form wechselwirken, so ist es konsequent anzunehmen, dass letztlich auch die Kartenbildung durch den Stimulus beeinflusst wird. Dem bereits präsentierten Modell muss eine solche Komponente noch hinzugefügt werden. Im biologischen Subjekt sind laterale Verbindungen erwähnt worden [Kisvárdy et al., 1997], die teils exzitatorischer, teils inhibitorischer Natur ist. Diese Verbindungen mögen bei der biologischen Begründung zur dynamischen Kartenbildung eine Rolle spielen.

Ein solches Konzept lag schon dem ersten Modell zur Kartenbildung [von der Malsburg, 1973] und praktisch allen folgenden zugrunde (siehe unter anderem [Kohonen, 1995a], [Obermayer et al., 1992], [Riesenhuber et al., 1996], [Brockmann et al., 1997a], [Wolf, 1998], [Wiemer et al., 2000a], für eine Übersicht siehe [Swindale, 1996]).

In diesen Realisierungen wird jedem Punkt der Kortexoberfläche \mathbf{r} ein rezeptives Feld $\mathbf{W}_{\mathbf{r}}$ zugeordnet. Alle Kartenmodelle, die auf dem Prinzip der stimulusgetriebenen Selbstorganisation beruhen, optimieren zwei letztlich konkurrierende Kriterien.

1. Die Karte soll den Stimulusraum möglichst gleichmässig und vollständig repräsentieren. Die Karte soll also – näherungsweise – den Stimulusraum ausfüllen oder anders

formuliert, soll sich die Karte so in den Stimulusraum einfallen, dass jedem Punkt des Stimulusraumes ein korrespondierender Punkt auf der Karte möglichst nahe kommt. Es gibt in neuester Zeit auch einige Untersuchungen, die zu dem Ergebnis kommen, dass dies in biologischen Karten gut erfüllt ist [Swindale et al., 2000].

2. Benachbarte Punkte auf der Karte sollen sich im Stimulusraum möglichst nahe kommen. Auf der Karte soll also in einer zunächst nicht näher bestimmten Weise die Glattheit optimiert werden, was letztlich die topographischen Eigenschaften der Karte hervorbringt.

Die relative Stärke dieser beiden Kriterien zueinander bestimmt die Entwicklung der Karte. In einem bestimmten Parameterbereich wird eine entsprechende Dynamik zu einem Auffalten der Karte im Stimulusraum führen.

An dieser Stelle sollen zwei konkrete Modelle vorgestellt werden, deren Vertreter und Varianten für den Rest der Arbeit eine große Rolle spielen werden. Das erste Modell ist die von Kohonen entwickelte selbstorganisierende Karte (*self-organizing map*, SOM) [Kohonen, 1995a]. Das zweite das elastische Netz, das ursprünglich zur Lösung des *traveling salesman*-Problems entwickelt wurde [Durbin & Willshaw, 1987]. Beide Modelle zeichnen sich durch ihre konzeptionelle Einfachheit aus und wurden u.a. in [Wolf, 1998] für ein Modell mit effektiv zweidimensionalem Stimulusraum ausführlich untersucht.

In den hier verwendeten Modellen werden die Neuronen auf einem quadratischen Gitter angeordnet und mit \mathbf{r} indiziert. Die Lerndynamik erfolgt in diskreten Zeitschritten.

Dem Modell aus dem letzten Kapitel liegt die selbstorganisierende Karte von der Konzeption her näher. Zu diesem Modell muss lediglich eine weitere Komponente hinzugefügt werden. Wieder wird in jeder Iteration das dem Stimulus nächstliegende Neuron bestimmt, also das mit dem größten Überlapp

$$\mathbf{r}^* = \operatorname{argmax}_{\mathbf{r}} [\mathbf{V}\mathbf{W}_{\mathbf{r}}]. \quad (4.2)$$

Statt wie zuvor von einem zufällig vernetzten kortikalen Gewebe auszugehen, soll nun berücksichtigt werden, dass sich Neuronen in *Abhängigkeit ihres Abstandes* beeinflussen können. Dieser Tatsache wird mit der Verwendung einer gaußförmigen Nachbarschaftsfunktion Rechnung getragen

$$H(\mathbf{r}, \mathbf{r}^*) = \exp(-(\mathbf{r} - \mathbf{r}^*)^2/2/\sigma_K^2) \quad (4.3)$$

Diese Nachbarschaftsfunktion fließt in die Lerndynamik mit ein

$$\frac{\partial \mathbf{W}}{\partial t} = \gamma H(\mathbf{r}, \mathbf{r}^*)[\mathbf{V} - \mathbf{W}], \quad (4.4)$$

wobei γ die Lernrate ist. Das elastische Netz lässt sich durch die Gleichung

$$\frac{\partial \mathbf{W}}{\partial t} = \gamma [\mathbf{V} - \mathbf{W}_r] \mathbf{e}(\mathbf{r}, \mathcal{S}, \mathbf{W}(\cdot)) + \eta \sum_{|\mathbf{r}_2 - \mathbf{r}|=1} \mathbf{W}_{\mathbf{r}_2} - \mathbf{W}_r \quad (4.5)$$

definieren, wobei

$$\mathbf{e}(\mathbf{r}, \mathcal{S}, \mathbf{W}(\cdot)) = \frac{\exp(-(\mathbf{V} - \mathbf{W}_r)^2/2/\sigma_{EL}^2)}{\sum_{\mathbf{r}^* \in \mathcal{N}} \exp(-(\mathbf{V} - \mathbf{W}_{\mathbf{r}^*})^2/2/\sigma_{EL}^2)} \quad (4.6)$$

und \mathcal{N} die Neuronenmenge ist.

Diese Gleichung ist so konstruiert, dass sie im Kontinuumsfall zu

$$\frac{\partial \mathbf{W}}{\partial t} = \gamma [\mathbf{V} - \mathbf{W}] \mathbf{e}(\mathbf{r}, \mathcal{S}, \mathbf{W}(\cdot)) + \eta \Delta \mathbf{W}, \quad (4.7)$$

wird, was sich für analytische Zwecke eignet, wobei Δ den Laplaceoperator symbolisiert. Die Größen γ , sowie σ_K bzw. η und σ_{EL} sind freie Parameter der Modelle.

Erscheinen beide Modelle formal und auch von der konzeptionellen Grundidee ähnlich, so sind sie in zweierlei Hinsicht verschieden. Während die selbstorganisierende Karte das Konzept des Vektorquantisierers durch eine Nachbarschaftsfunktion erweitert und vor allem durch seine Einfachheit überzeugt, hat die Adaptionregel des elastischen Netzes die Eigenschaft, dass die Dynamik als Gradientenabstieg der Kostenfunktion

$$E = \sigma_{EL}^2 \mathbb{E} \left[\log \left(\sum_{\mathcal{N}} \exp \left(-\frac{(\mathbf{V} - \mathbf{W}_r)^2}{2 \sigma_{EL}^2} \right) d\mathbf{r} \right) \right] + \frac{\eta}{2} \sum_{\mathcal{N}} \sum_{|\mathbf{r}_2 - \mathbf{r}|=1} (\mathbf{W}_{\mathbf{r}_2} - \mathbf{W}_r)^2 \quad (4.8)$$

darstellbar ist. Für weitere Betrachtungen wird sich diese Eigenschaft als vorteilhaft erweisen.

Kapitel 5

Kartenalgorithmen mit natürlichen Bildern als Inputs

Hier soll untersucht werden, inwieweit die im letzten Kapitel vorgestellten Algorithmen, speziell die selbstorganisierende Karten, unter der Stimulierung durch natürliche Bilder analog zu der Vorgehensweise im Kapitel 3 verwendbar sind. Die Resultate dieser Modelle werden mit den im vorangegangenen Kapitel beschriebenen biologischen Befunden verglichen.

Es wird sich schnell zeigen, dass einige Modifikationen notwendig sind, um biologisch relevante Ergebnisse zu erhalten. Bei der Erweiterung der Modelle wurde eher versucht, die konzeptionelle Einfachheit des Modells zu erhalten, als physiologische Details zu reproduzieren.

5.1 Die Eigenschaften des SOM-Modells

Zunächst sollen selbstorganisierende Kartenalgorithmen diskutiert werden, in denen ähnlich wie im vorigen Modell Derivate von natürlichen Bildern verwendet werden. Für die hier diskutierten Karten werden Arrays aus 128×128 Neuronen benutzt. Jedes Neuron hat einen Eingangsraum von 24×24 Pixeln. Es werden wiederum vorgefilterte natürliche Bilder verwendet. Aus diesen wird in jeder Iteration ein zufälliger Ausschnitt in der Größe des Eingangsraumes verwendet. Diese Ausschnitte werden mit Gaußverteilungen einer Weite von σ_p von 3.2 Pixeln multipliziert, die diesmal aber nicht zentriert, sondern an einer zufälligen Position im Eingangsraum innerhalb eines Randes von 8 Pixeln positioniert sind.

Die verwendeten natürlichen Bilder haben eine Dimension von 516×516 Pixeln. (Sie sind der von [Olshausen & Field, 1996] verwendeten Datenbank entnommen worden.) Sie sind

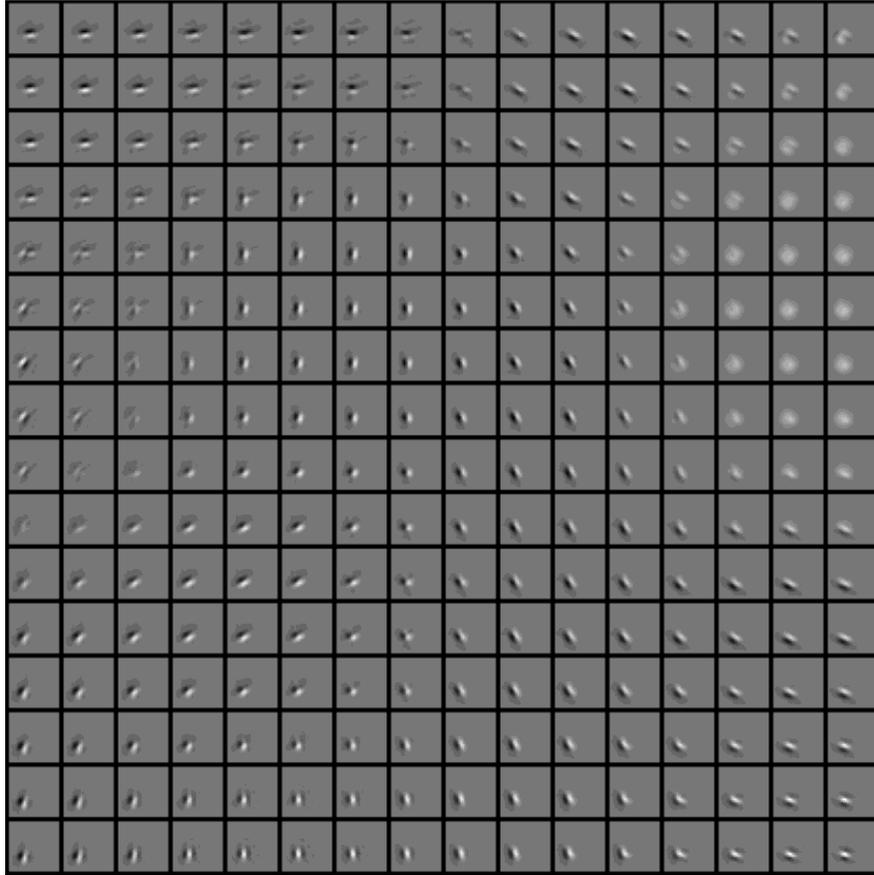


Abbildung 5.1: Ein Ausschnitt von 16×16 aus einer Karte mit insgesamt 128×128 Neuronen. Der Eingangsraum eines Neurons besteht aus 24×24 Pixeln. Die Karte wurde retinotop mit Gaussverteilungen initialisiert, die mit weißem Rauschen multipliziert wurden. Als Stimulus wurden zufällige Bildausschnitte aus mit Filter Gl. (3.1) und $k_0 = 0.38 \times 2\pi / \text{Pixel}$ gefilterten Bildern verwendet, was einer Wellenlänge von etwa 3.65 Pixeln entspricht. Diese wurden ihrerseits mit einer Arborfunktion mit $\sigma_p = 3.2$ Pixel multipliziert, die innerhalb von Rändern von 8 Pixeln zufällig positioniert wurden. Das Modell hatte eine Lernrate, die γ während der gesamten 2^{20} Iterationen exponentiell von 0.1 auf 0.005 abfiel. Die Nachbarschaftsweite σ_K des Modells betrug 2.0 Neurone.

mit dem Filter nach Gl. (3.1) bei $k_0 = 200$ Zyklen / Bild gefiltert worden. Die maximale Transmission dieses Filters hat dann eine Wellenzahl von rund 0.27 Zyklen / Pixel, was einer Wellenlänge von etwa 3.65 Pixeln entspricht. Das Verhältnis aus Arborweite und spatialer Wellenlänge liegt daher etwas unter eins. Nach den Überlegungen aus dem vorigen Kapitel handelt es sich hier um ein intermediäres oder sogar frühes Stadium der Entwicklung und nicht um den adulten Reifezustand.

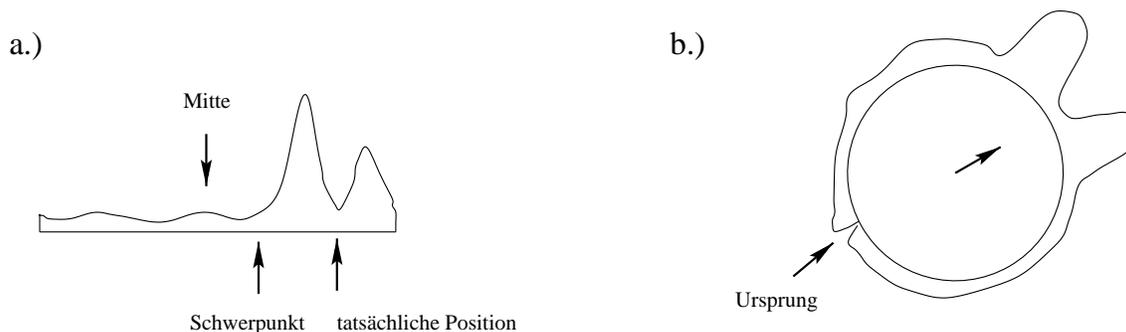


Abbildung 5.2: Die Vorteile einer Positionsbestimmung des rezeptiven Feldes in einem periodischen Raum im Vergleich zu einem abgeschlossenen (schematisch). Während im Falle eines endlichen Raumes der Schwerpunkt durch ein gleichmäßiges Rauschen zur Mitte des Raumes verschoben ist, kann durch die im Abschnitt 5.2 vorgeschlagenen Modifizierung der mit dem optischen Zentrum koinzidierende Punkt bestimmt werden.

In diesem Bereich ist darüberhinaus

$$\tilde{\mathbf{W}} \approx \mathbf{W} \quad (5.1)$$

und daher soll im Rahmen dieses Kapitels auf die Berücksichtigung des Unterschiedes zwischen kortikogenikulaten Verbindungstärken und dem rezeptiven Feld (siehe Abschn. 3.5) verzichtet werden.

Die Karten wurden mit retinotop angeordneten Gaußverteilungen initialisiert, die mit weißem Rauschen multipliziert wurden. Die Lernrate γ fiel exponentiell ab. Die Nachbarschaftsweite σ_K wurde konstant gehalten.

5.2 Methoden der Extraktion von Merkmalskarten

Es lassen sich aus den erhaltenen Simulationsergebnissen die Eigenschaftskarten, die zu den mit *optical imaging* Methoden gefundenen äquivalent sind, extrahieren, insbesondere die Karte der präferierten Orientierung und die der präferierten lokalen Wellenlänge der Gaborfilter. Die Methoden sind dabei ähnlich wie in Kapitel 3. Insbesondere kann die Orientierung nach Gl. (3.13) bestimmt werden. Anders als vorher ist die Analyse der resultierenden Positionsverteilungen im visuellen Feld von erheblichem Interesse.

Die einfachste Methode ist die Bestimmung des Schwerpunktes

$$\xi_i^{sp} = \int_{\Xi} \xi_i |\mathbf{W}_{\xi_1, \xi_2}^2| d\xi, \quad (5.2)$$

wobei ξ_i wiederum eine der beiden Koordinaten des visuellen Felds und \mathbf{W} den zu untersuchenden Vektor des rezeptiven Felds darstellt.

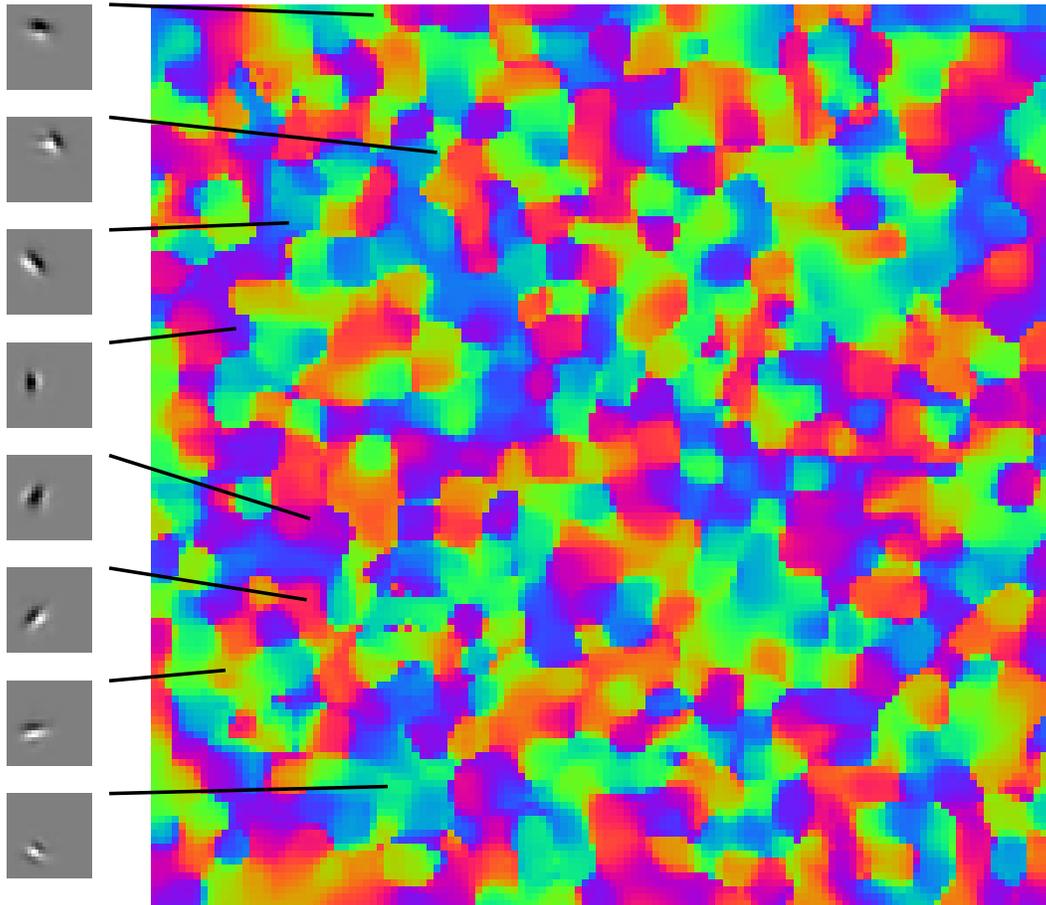


Abbildung 5.3: Eine Orientierungskarte, wie sie nach Gl. (3.13) aus den Daten von Abb. 5.1 gewonnen werden kann.

Mit dieser Methode findet allerdings eine systematische Verschiebung des ermittelten Positionswertes in Richtung des Mittelpunktes des Eingangsraumes statt (siehe Abb. 5.2). Der Grund dafür ist ein ubiquitäres Rauschen, dessen Stärke an allen Punkten des Eingangsraumes im Mittel konstant ist, dessen Schwerpunkt also ungefähr in der Mitte des Eingangsraumes lokalisiert ist. $\mathbf{W}_{\xi_1, \xi_2}^2$ ist also die Summe aus dem rezeptiven Feld und diesem Rauschen. Der Schwerpunkt wird dann zur gewichteten Summe aus diesen beiden Termen. Damit verschiebt sich der berechnete Schwerpunkt systematisch in die Mitte des Eingangsraumes Ξ . Das kann vermieden werden, indem man stattdessen die Position als periodische Variable des Eingangsraumes Ξ definiert. Auf diese Weise läßt sich der Punkt, der optisch dem Zentrum des rezeptiven Feldes entspricht genauer als die zuvor beschriebene Methode bestimmen, siehe Abb. 5.2.

Es ist hilfreich die komplexe Größe

$$\rho_i = \log \int_{\Xi} \exp(i\xi_i/\Xi_x) \mathbf{W}^2(\xi_1, \xi_2) d\xi \quad (5.3)$$

einzuführen, wobei Ξ_x für die Grösse des Eingangsraumes in einer Richtung steht. ρ_i hat dann für die Position des rezeptiven Feldes eine ähnliche Bedeutung wie die Variable κ aus Gl. (3.13) bei der Orientierungsanalyse. Der Imaginärteil kodiert den so berechneten Schwerpunkt, der sich mit

$$\xi_i^{pr} = \text{Im} [\rho_i] \quad (5.4)$$

extrahieren läßt. Der Realteil kann als Maß für die Signifikanz des ermittelten Schwerpunktes interpretiert werden.

Außerdem können die Größen Position, Orientierung, lokale Wellenlänge durch eine Gaboranalyse, wie sie in Kapitel 3 vorgestellt wurde, ermittelt werden.

5.3 Numerische Resultate zu selbstorganisierenden Karten

Die rezeptiven Felder der einzelnen Neuronen (siehe Abb. 5.1) ähneln in ihrem Aussehen denen, die mit Hilfe des Vektorquantisierers (Kapitel 3) aus dem nichttopographischen Modell gewonnen wurden. Sie zeigen also ebenfalls ein Aussehen, das dem von biologischen *simple cells* entspricht.

Vergleicht man rezeptive Felder von Neuronen, die zueinander benachbart liegen, bemerkt man, dass diese einander ähnlich sind. Geht man auf einem Pfad über die Oberfläche \mathbf{r} des Modells so ändern sich die Eigenschaften der rezeptiven Felder kontinuierlich. Das Modell besitzt in dieser Hinsicht die Eigenschaften, die im visuellen Kortex gefunden werden (vgl. Abschn. 4.1) und wie sie auch andere Modelle mit reduzierten Stimuluseigenschaften aufweisen [Swindale, 1996]. Das Modell erscheint daher zunächst als plausibles Pendant sowohl zu anderen Kartenmodellen als auch zu den Modellen, in denen natürlichen Bildern verwendet werden (vgl. [Mayer et al., 1998b]).

Allerdings deutet die Analyse der Karteneigenschaften auf Unstimmigkeiten hin. Die extrahierten Orientierungskarten (siehe Abb. 5.3) sehen dem biologischen Gegenstück leidlich ähnlich. Es fällt allerdings eine etwas fleckige Musterung auf. Die Karten der präferierten Wellenzahlen (siehe Abb. 5.4) zeigen isolierte Gebiete mit kleineren Wellenzahlen, wie sie auch in biologischen Karten [Hübener et al., 1997] gefunden werden. In vielen Fällen weisen Gebiete mit einer schwachen Orientierungsselektivität auch eine großen spatiale Wellenlänge auf. In anderen Fällen können niedrige Orientierungsselektivitäten auch durch Abweichungen vom Aussehen eines Gaborfilters verursacht sein. In Abb. 5.5 sind die Gebiete

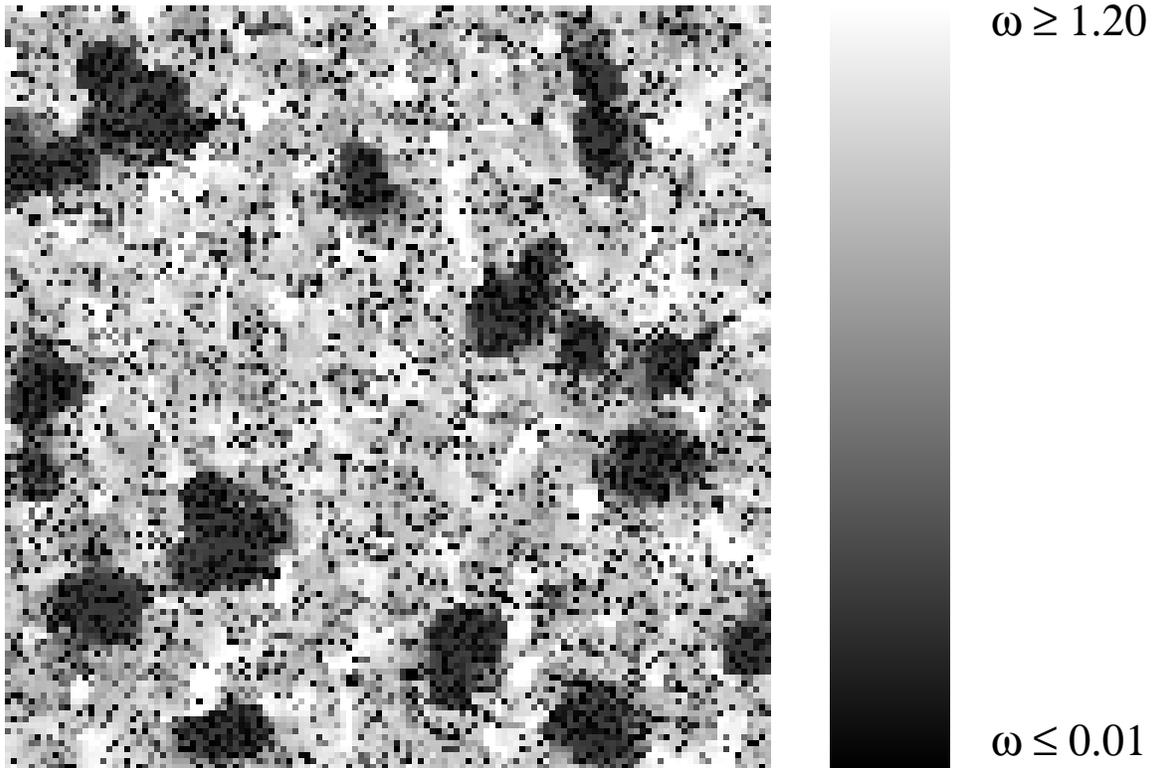


Abbildung 5.4: SOM: Die extrahierte Karte der präferierten Wellenzahlen. Es handelt sich um die selben Daten wie Abb. 5.7). Die Wellenzahlen wurde mit der in Kapitel 3 beschriebenen Gaboranalyse ermittelt.

geringer Orientierungsselektivität im Vergleich zu einer kombinierte Karte der präferierten Wellenlängen und des Abstandes zum nächstgelegenen Gaborfilter dargestellt. Einerseits erkennt man ein netzartiges Muster geringer Orientierungsselektivität, das die Orientierungsdomänen voneinander abgrenzt. Dort wird die geringe Orientierungsselektivität durch größere Distanzen zum nächsten Gaborfilter bewirkt. Das andere – eher fleckige – Muster entsteht durch die Bereiche großer spatialer Wellenlängen. In biologischen Karten treten solche linienartigen Gebilde nicht auf. Vielmehr bleiben Stellen geringer Orientierungsselektivität auf einzelne Punkte beschränkt. Aufgrund dieser qualitativen Unterschiede zeigt sich das klassische SOM-Modell als ungeeignet, wenn mit angenähert realistischen Reizen gearbeitet wird.

Dies wird durch eine Analyse der extrahierten Positionskarten unterstützt. Die Karten der Position des rezeptiven Feldes im Stimulusraum (siehe Abb. (5.7)) haben ein unstrukturiertes Aussehen. Sie sind kaum retinotop und deshalb nicht mit dem biologischen Vorbild zu vereinbaren. Der Grund dafür ist ein metrisches Problem zwischen der Repräsentation der spatialen Phase und der Position im visuellen Feld. Diese Zusammenhänge bilden den Ausgangspunkt zu einer Erweiterung des Modells.

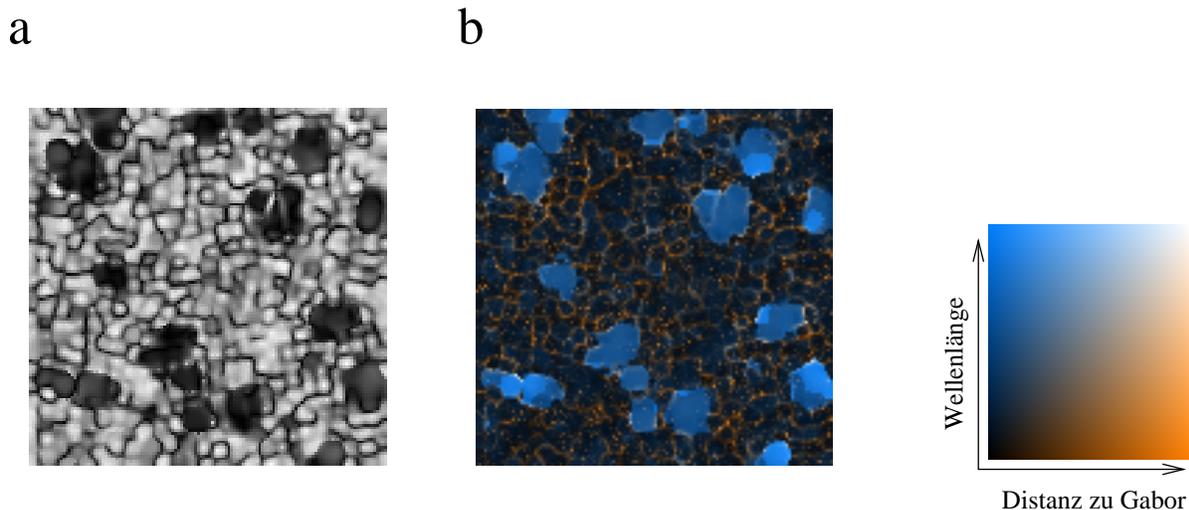


Abbildung 5.5: Es sind Orientierungspräferenzen und die Größen, die die Ursache für geringe Orientierungspräferenzen sind, dargestellt. **a)** Die aus der Abb. 5.1 gewonnenen Orientierungsstärken. Die dunklen Bereiche stehen für niedrige Orientierungsstärken, die hellen für große. **b)** Eine kombinierte Karte, die mit gelben Farben Stellen kodiert, die eine starke Abweichung von dem nach der Optimierung nächstgelegenen Gaborfilter haben; hellblau für solche Stellen, die eine niedrige spatiale Frequenz besitzen. Die verschiedenen Kombinationen werden mit korrespondierenden Farbkombinationen kodiert. Solche Stellen, die entweder eine niedrige lokale Frequenz aufweisen oder eine große Abweichung von dem ähnlichsten Gaborfilter aufweisen, befinden sich meist an Stellen geringer Orientierungspräferenz.

5.4 Der Zusammenhang zwischen lokaler Phase und Retinotopie

Eine wesentliche Größe für die Strukturbildung in neuronalen Karten ist der euklidische Abstand zwischen zwei hochdimensionalen Vektoren, in diesem Fall zwischen dem Stimulus \mathbf{V} und den rezeptiven Feldern \mathbf{W} . Das führt dazu, dass zwei rezeptive Felder mit entgegengesetzter Phase, die sich an derselben Position befinden einen doppelt so größeren Abstand im hochdimensionalen Raum besitzen wie zwei dislozierte rezeptive Felder (siehe Abb. 5.6). Da selbstorganisierende Karten topographieerhaltend sind, also dazu tendieren, Distanzen in ihrer Repräsentation zu erhalten, kann die retinotopie Anordnung von rezeptiven Feldern in der Karte nicht erhalten bleiben. Zur Lösung dieses Problems wurde eine Variante der selbstorganisierenden Karte verwendet, die selbstorganisierende Abbildung adaptiver Unterräume oder *adaptive subspace self-organizing map* (ASSOM) [Kohonen, 1995a].

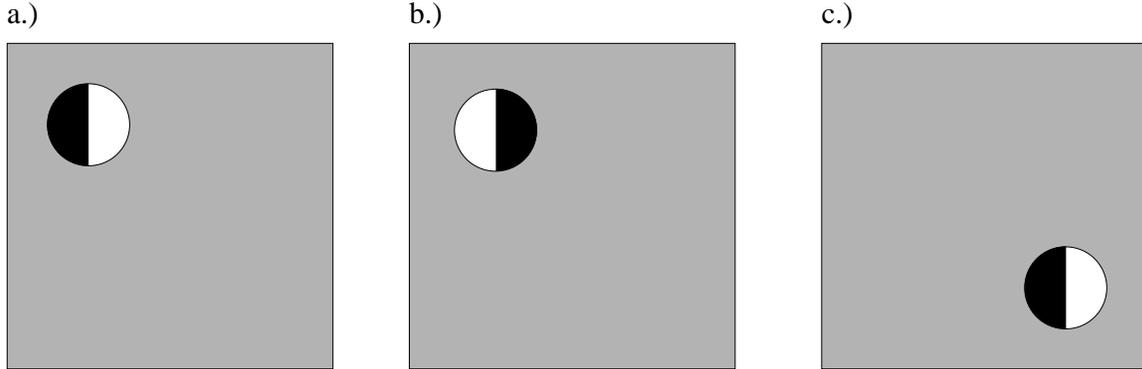


Abbildung 5.6: Rezeptive Felder (schematisch, in Koordinaten des visuellen Feldes). Schwarze und weiße Regionen korrespondieren, wie vorher mit Subfeldern aus genikulaten Afferenzen zu *ON-center* und *OFF-center* Zellen respektive. **a** und **c** haben unterschiedliche Positionen, während ihre Phase koinzidiert. Bei **a** und **b** stimmen die Positionen überein, während die Phase verschoben ist. Betrachtet man die rezeptiven Felder als Vektoren, so ist der euklidische Abstand zwischen den beiden rezeptiven Feldern **a** und **b** größer als zwischen **a** and **c**.

5.5 Eine Karte aus adaptiven Unteräumen

Die *adaptive subspace SOM* [Kohonen, 1995a] ist eine Variante der normalen selbstorganisierenden Karte. Die Darstellung ihrer Anwendung auf die Gegebenheiten im visuellen Kortex folgt den Darstellungen in [Mayer et al., 1998a, Mayer et al., 2000a]. Jedes Neuron des $N_x \times N_x$ großen Arrays verfügt statt über einen rezeptiven Feldvektor \mathbf{W}_r über L Basisvektoren $\mathbf{W}_{r,l} = (\mathbf{W}_l)_r$, deren Abstand zu einem Stimulus \mathbf{V}

$$D^2(\mathbf{V}, (\mathbf{W})_r) = \min_{a_l} \left(\left| \mathbf{V} - \sum_L a_l \mathbf{W}_{r,l} \right| \right)^2 \quad (5.5)$$

definiert ist. Die Einzelvektoren $\mathbf{W}_{r,l}$ eines Ortes r sollen zueinander orthonormiert sein, so dass

$$a_{l,\min} = \mathbf{V} \cdot \mathbf{W}_{r,l}. \quad (5.6)$$

sein muss. Das kann durch folgende Überlegung leicht bewiesen werden: Man kann durch ein geeignetes Verfahren die $\mathbf{W}_{r,l}$ mit anderen Vektoren $\mathbf{K}_{L+1} \dots \mathbf{K}_\Xi$ so ergänzen, dass die Vektormenge $\mathbf{W}_{r,l}, \mathbf{K}_l$ eine Orthonormalbasis des Stimulusraumes darstellt, die die Dimension Ξ des Eingangsraumes aufweist. Der Abstandsvektor zwischen Stimulus \mathbf{V} und einer beliebigen Superposition der Basisvektoren $(\mathbf{W})_r$ ist

$$(\beta_1, \dots, \beta_L, \mathbf{V} \cdot \mathbf{K}_{L+1}, \dots, \mathbf{V} \cdot \mathbf{K}_\Xi), \quad (5.7)$$

wobei

$$\beta_l = (\mathbf{V} - a_l \mathbf{W}_{l,r}) \cdot \mathbf{W}_{l,r}, \quad (5.8)$$

was offensichtlich wegen $\mathbf{W}\mathbf{W} = 1$ für

$$a_l = \mathbf{V} \cdot \mathbf{W}_{l,r} \quad (5.9)$$

die β_l verschwinden läßt und so zum kürzest möglichen Abstand führt.

Anstatt den *winner* als das Neuron zu definieren, dessen Abstand in der in Gl. (5.5) gegebenen Form minimal ist, ist es unter den gegebenen Nebenbedingungen gleichwertig die maximale Aktivierung zu verwenden. Danach ist der *winner* das Neuron an der Stelle

$$\mathbf{r}^* = \operatorname{argmax}_{\mathbf{r}} \left(\sum_L a_{l,r} \right) \quad (5.10)$$

mit der Antwortfunktion

$$a_{l,r} = \mathbf{V} \cdot \mathbf{W}_{l,r}. \quad (5.11)$$

Der Adaptionprozess erfolgt in Analogie zur normalen selbstorganisierenden Karte dann durch

$$\frac{\partial \mathbf{W}_{l,r}}{\partial t} = \gamma \exp\left(-\frac{(\mathbf{r} - \mathbf{r}^*)^2}{2\sigma_K^2}\right) (a_{l,r} \mathbf{V} - \mathbf{W}_{l,r}). \quad (5.12)$$

Nach der Präsentation eines Stimulus werden die $\mathbf{W}_{l,r}$ in jeder einzelnen Einheit l mit Hilfe des Gram-Schmidt-Verfahrens orthogonalisiert. Dieses Modell hat sich insofern als erfolgreich erwiesen, als dass sich diese beiden Basisvektoren immer derart anordnen, dass sich alle lokalen Phasen durch Superpositionen von den jeweiligen Basisvektoren dargestellt werden können. Auf diese Weise wird die lokale Phase in der Karte nicht topographisch repräsentiert.

5.6 Numerische Resultate zum Modell mit adaptiven Unterräumen

Um die Vergleichbarkeit zwischen selbstorganisierenden Karten und den entsprechenden Karten von adaptiven Unterräumen zu gewährleisten, wurden gleiche Parameterwerte für σ_K und γ gewählt. In den in diesem Kapitel beschriebenen Modellen entsprechen deshalb die Werte dieser Größen den unter Abb. 5.1 gegebenen. Auch die Initialisierung dieses Modells wurde in gleicher Weise durchgeführt. Es werden hier nur Karten des ASSOM diskutiert, die einen zweidimensionalen Unterraum L aufweisen.

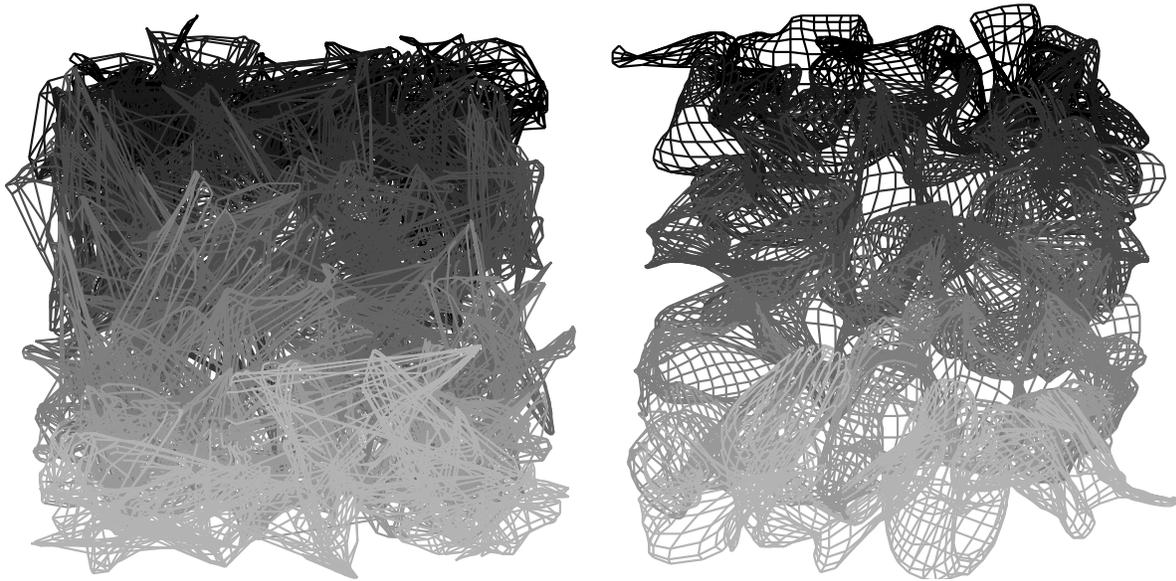


Abbildung 5.7: Die Einfaltung des Netzes der rezeptiven Felder im visuellen Feld. Links: SOM, Rechts: ASSOM. Die ASSOM wurde mit $L = 2$ Dimensionen betrieben. Sonst entsprechen die Parametereinstellungen denen der SOM von Abb. 5.1.

Die beiden Basisvektoren entwickeln rezeptive Feldformen, die in ihrem Aussehen denen des Vektorquantisierers und denen der selbstorganisierenden Karte entsprechen. Vergleicht man die beiden Basisvektoren an einem Maschenpunkt \mathbf{r} des Modells, so zeigen sich im Allgemeinen rezeptive Felder, die in ihrer Position im visuellen Feld und in ihrer präferierten Orientierung koinzidieren, aber die eine um etwa 90° verschobene lokale Phase aufweisen. In Abb. 5.9 sind die Histogramme der Winkeldifferenzen und Positionsunterschiede jeweils zwischen den beiden Basisvektoren dargestellt. Es ist eine deutliche Bevorzugung der gleichen präferierten Orientierung und der Position im visuellen Feld erkennbar.

Wie an einem typischen rezeptiven Feld des Modells erkennbar wird (siehe Abb. 5.14), ist im generischen Fall die lokale Phase die Größe, die am stärksten zwischen beiden Basisvektoren variiert. Auf diese Weise gelingt es also in der Tat, die Repräsentation der lokalen Phase auf der Kortexoberfläche zu eliminieren. Damit werden Störungen in der Retinotopie minimiert, was sich an entsprechend unterschiedlichen Entwicklungen der lokalen Repräsentationen von selbstorganisierender Karte und ASSOM zeigt.

Um den Erfolg der Verwendung der adaptiven Unterräume zu visualisieren, wurde die Einfaltung der neuronalen Karte in das visuelle Feld dargestellt (Abb. 5.7). Wenn auch die retinotopie Initialisierung und die im Vergleich zum Gesamtausmaß der selbstorganisierenden Karte geringe Nachbarschaftsweite σ_K dazu geführt hat, dass die retinotopie Anordnung nicht gänzlich zerstört wurde, so zeigen sich unstrukturierte Verschiebungen. Auch die ASSOM zeigt deutliche Verzerrungen, deren Wellenlänge in etwa mit der Größe der Orientierungskolumnen übereinstimmt. Allerdings bleibt das Netz der Neuronen als

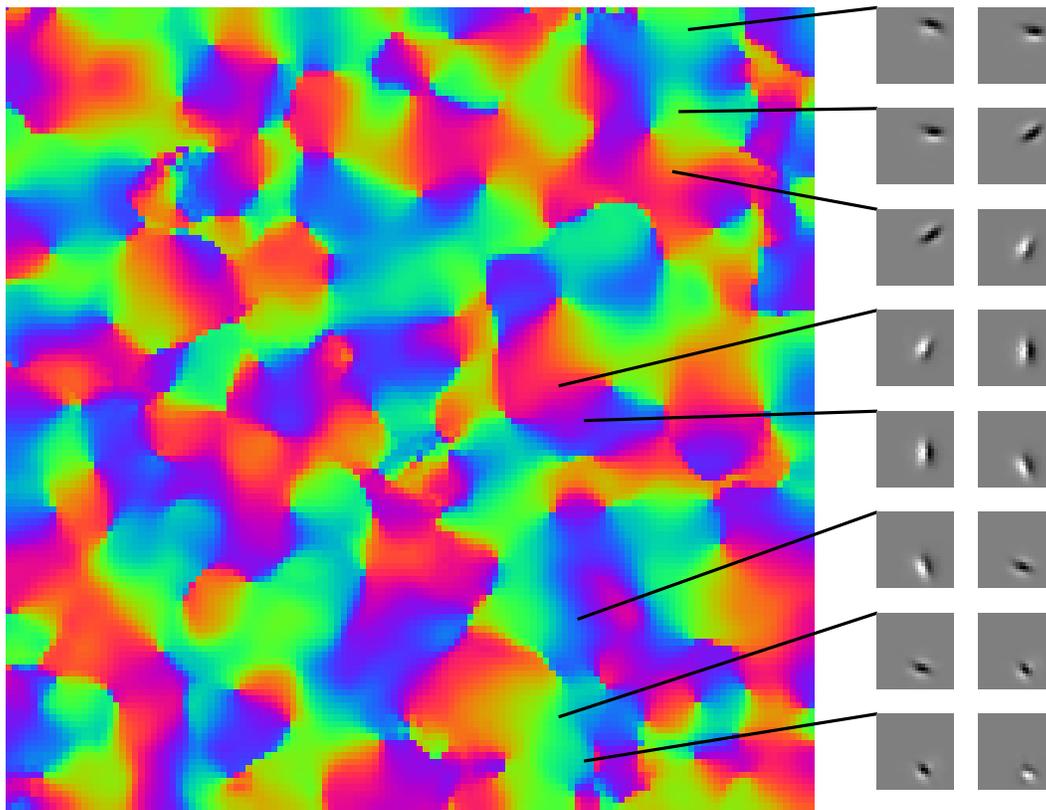


Abbildung 5.8: Die Orientierungskarte jeweils der ersten Basisvektoren der adaptiven Unterraummodells (ASSOM) und Beispiele für receptive Felder, nach der Orientierung geordnet. Es wurden die gleichen Daten wie in Abb. 5.7 verwendet.

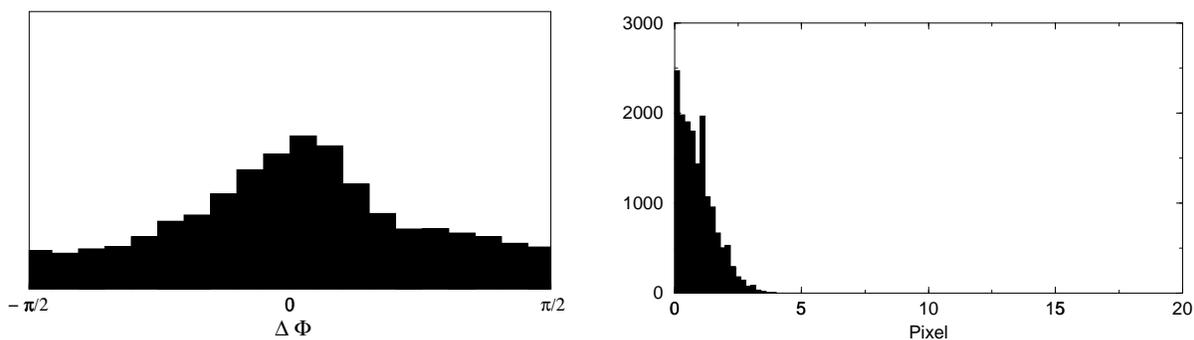


Abbildung 5.9: Histogramme der Differenzen jeweils der beiden Basisvektoren. Die Differenz der präferierten Orientierung (links) und die Differenz der Position im visuellen Feld (rechts, in horizontaler Richtung, identische Daten wie Abb. 5.7).

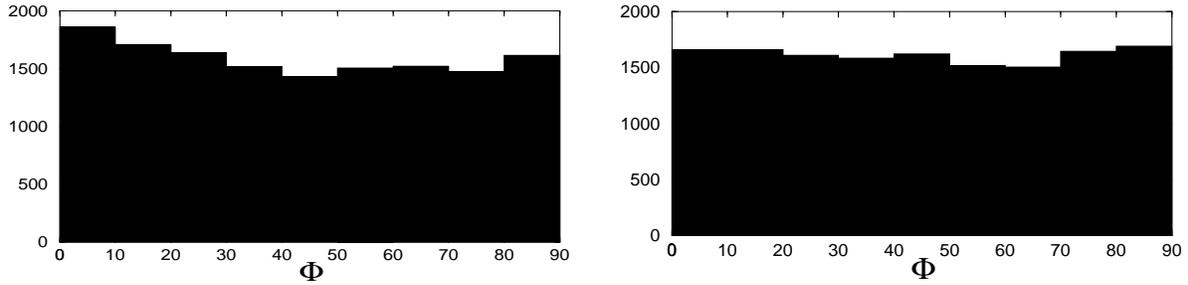


Abbildung 5.10: Histogramme typischer Schnittwinkel beim selbstorganisierenden Kartenalgorithmus (links), und dem ASSOM-Modell (rechts). Während sich im SOM-Modell eine leichte Präferenz für senkrechte Schnittwinkel ergeben, ist die Schnittwinkelverteilung beim Modell adaptiver Unterräume gleichverteilt.

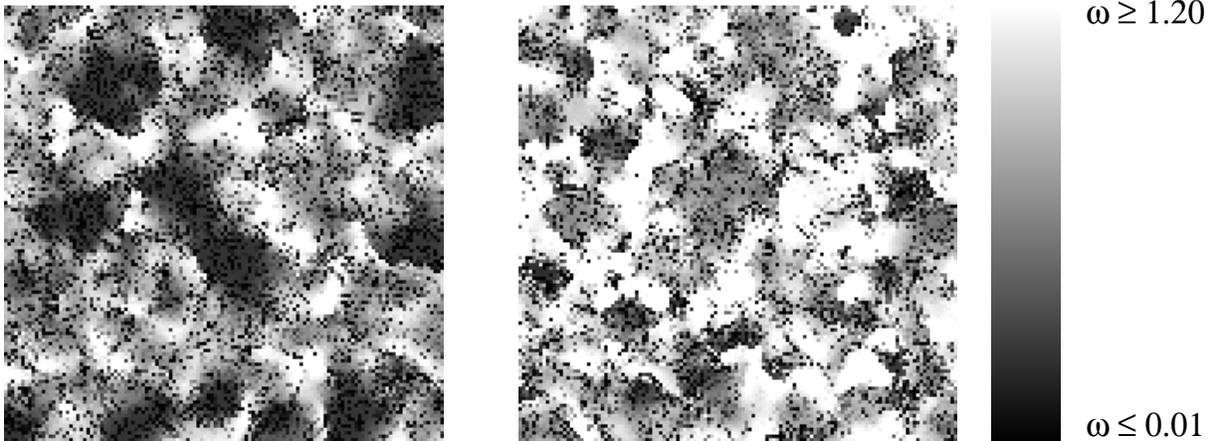


Abbildung 5.11: ASSOM: Extrahierte Karten der präferierten Wellenzahlen. Links, jeweils die ersten Basisvektoren, rechts die zweiten. (es handelt sich wiederum um die gleichen Daten wie Abb. 5.1). Diese wurde mit der in Kapitel 3 beschriebenen Gaboranalyse ermittelt.

solches erkennbar; die Strukturen sind hier geordneter als bei der selbstorganisierenden Karte. Die Störungen, die aber auch hier massiv sind, können als Hinweis darauf gewertet werden, dass es bei der ASSOM eine Kopplung zwischen der Orientierungskarte und der Positionskarte existiert. Diese Vermutung befindet sich auch in Übereinstimmung mit von [Dimitrov & Cowan, 1984] untersuchten Orts-Richtungs-Korrelationen in natürlichen Bildern.

Um diesen Zusammenhang zu beleuchten, wurde das Spektrum der rezeptiven Feldpositionen relativ zur retinotopen Karte

$$\tilde{\xi}_r = \xi_r - \xi_{\text{retinotop},r} \quad (5.13)$$

ermittelt. Dabei wurde für beide kartesischen Koordinaten eine Wellenlänge von rund 20

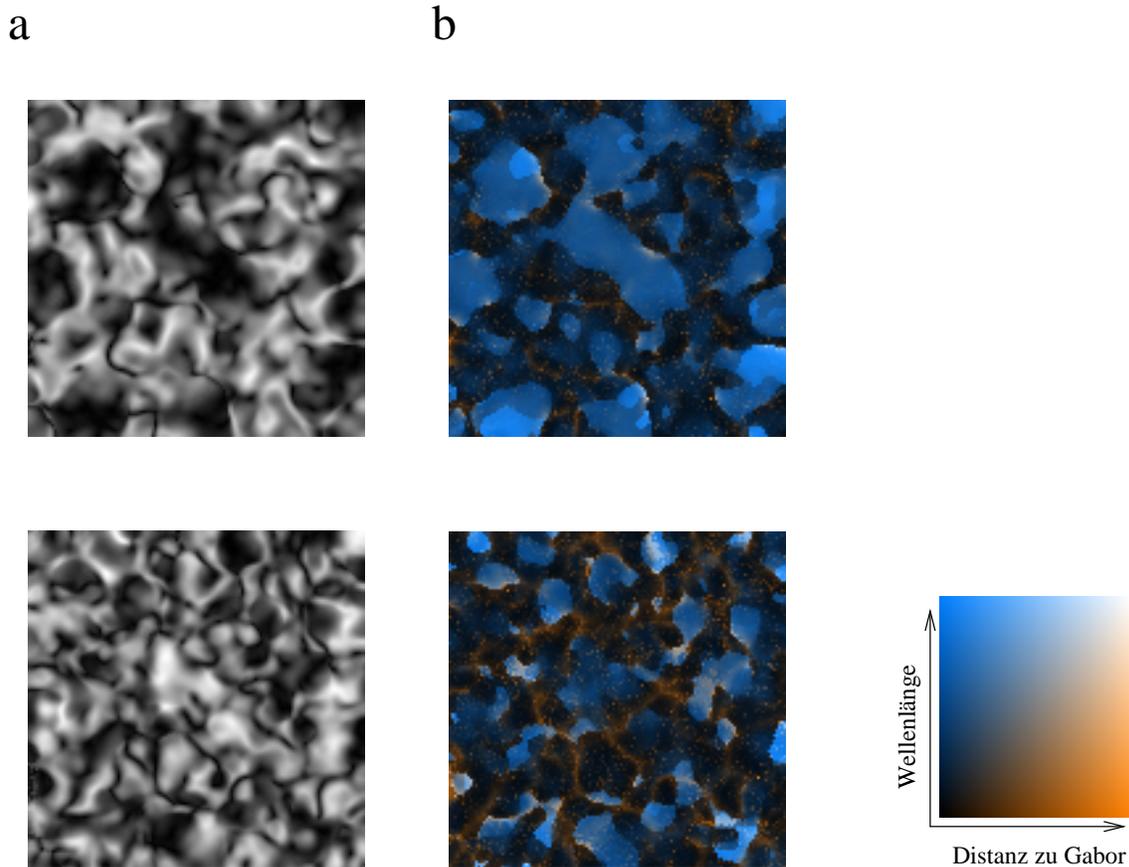


Abbildung 5.12: Die Ursachen von schwachen Orientierungsstärken bei der Karte adaptiver Unterräume sind denen der selbstorganisierenden Karte ähnlich. Hier die zu Abb. 5.5 analoge Abbildung zur ASSOM. Oben, jeweils die ersten Basisvektoren, unten die zweiten. Unter **a** ist die Karte der Orientierungsstärke wiedergegeben. Hell repräsentiert sind Gebiete mit einer starken Orientierungspräferenz, dunkel Gebiete mit einer schwachen. Unter **b** sind für dieselbe Karte die Messwerte für die lokale Wellenlänge und die Abweichung vom nächstgelegenen Gaborfilter angegeben. Ganz rechts ist dargestellt, welche Farbcodes für die verschiedenen Kombinationen von spatialer Wellenlänge und Abweichung bei der Gaboranalyse verwendet wurden.

Neuronen festgestellt.

Die Orientierungskarte der ASSOM zeigt ebenfalls eine Wellenlänge ungefähr 20 Neuronen, was deutlich grösser ist als die Wellenlänge der Orientierungsdomänen in der selbstorganisierenden Karte mit ungefähr 12 Neuronen unter den gleichen Simulationsbedingungen (siehe Abb. 5.8). Die Zahl der *pinwheels* pro Wellenlängenquadrat liegen bei der ASSOM etwa bei 3.3, bei der selbstorganisierenden Karte etwa bei 2.7.

Ein Blick auf die Karte der lokalen Wellenlängen (siehe Abb. 5.11) zeigt, dass die schon im selbstorganisierenden Kartenmodell prävalenten Inseln mit einer großen spatialen Wel-

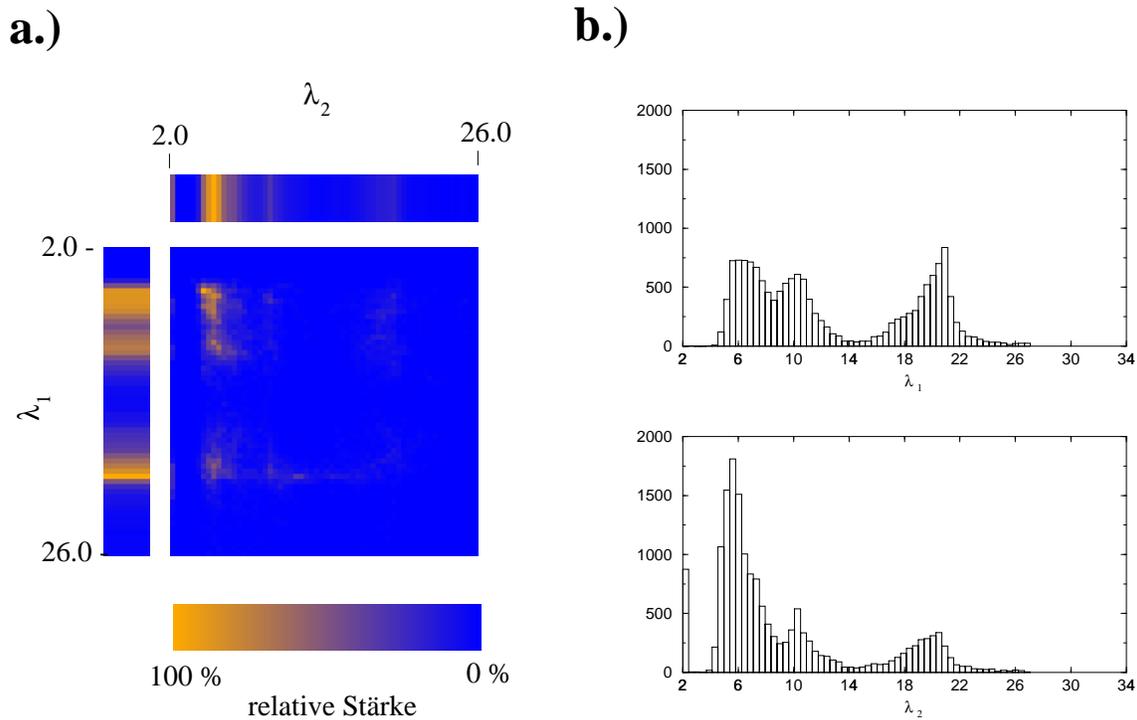


Abbildung 5.13: Histogramme der präferierten Wellenlängen (b) und deren Kombination bzgl. der beiden Basisvektoren (a). Die Häufigkeit des Auftretens einer Kombination ist mit dem Farbübergang von Dunkelblau (selten) zu Gelb (häufig) kodiert. Es ist zu erkennen, dass der zweite Basisvektor eher kürzere lokale Wellenlängen repräsentiert. (identische Daten wie Abb. 5.7).

Wellenlänge hier nur im ersten Basisvektor auftreten. Der zweite Basisvektor hat in vielen Fällen an diesen Stellen eine niedrigere Wellenlänge (vgl. Abb. 5.13). An diesen Stellen variiert also eher die lokale Wellenlänge, als die lokale Phase.

Die rezeptiven Felder der ASSOM werden durch zwei orthogonale Vektoren in jeder Einheit produziert. Diesen entsprechen ähnliche Positionen und Orientierungen, aber eine 90° Phasenverschiebung führen (siehe Fig. 5.14 a)). Stimuli an einer ähnlichen Position und mit einer ähnlichen Orientierung aber einer beliebigen Phase können also durch eine Einheit der ASSOM repräsentiert werden.

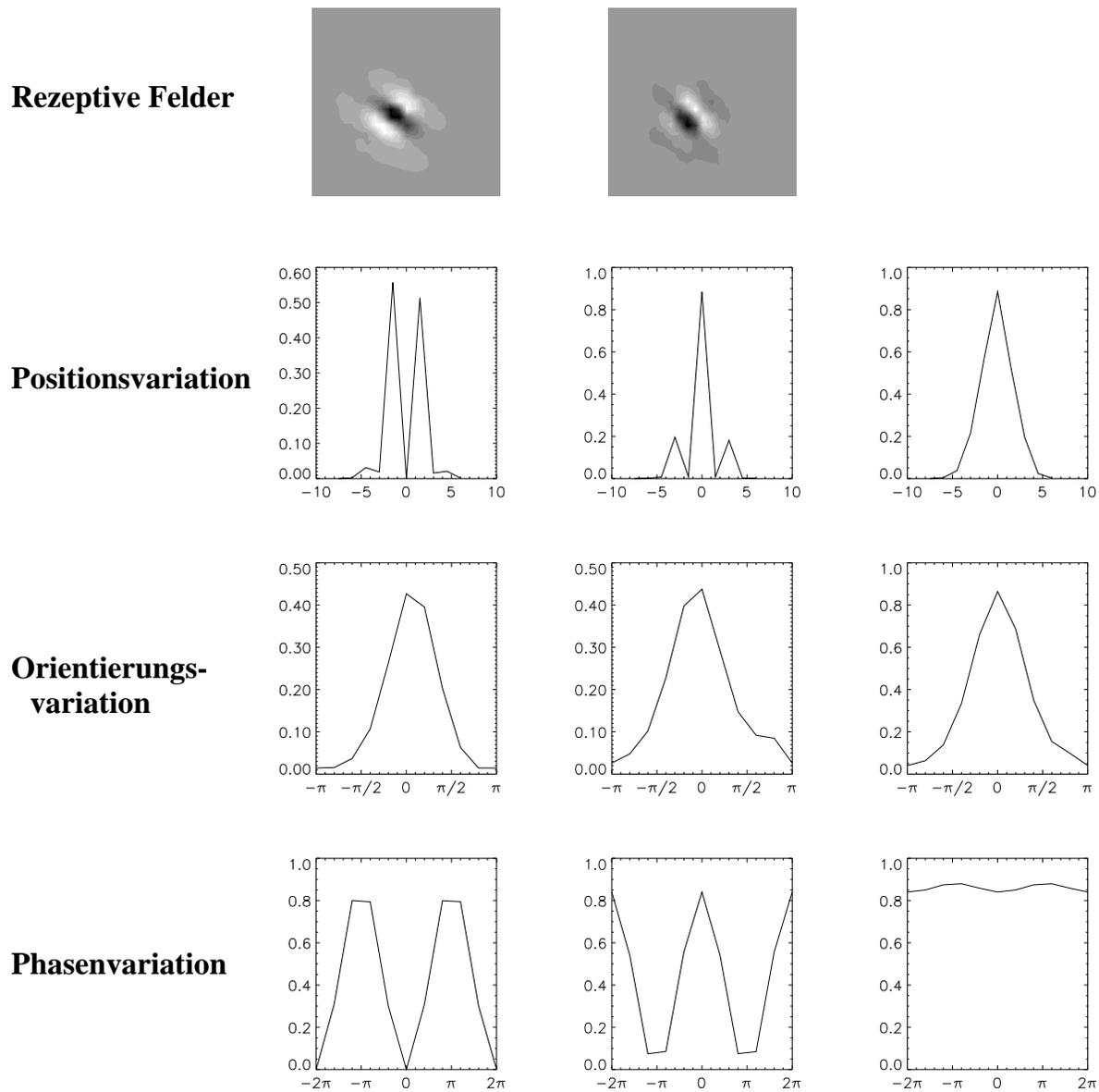


Abbildung 5.14: Ein Beispiel für ein rezeptives Feld des ASSOM-Modells $(\mathbf{W})_r$ (aus den Daten Abb. 5.7). Die beiden Basisvektoren zeigen rezeptive Felder von einem gaborartigen Aussehen. Die präferierte Orientierung und Position im visuellen Feld sind beinahe identisch, die spatiale Phase zeigt einen Unterschied von etwa 90° . Darunter die Antwort a_l der beiden rezeptiven Felder l bzw. ihrer Quadratsumme $\sum_l a_l^2$ auf einen Teststimulus (Gaborfilter) der in der Position variiert wurde. Darunter die Antwort auf eine Orientierungsvariation und auf eine Variation der lokalen Phase eine Phasenvariation.

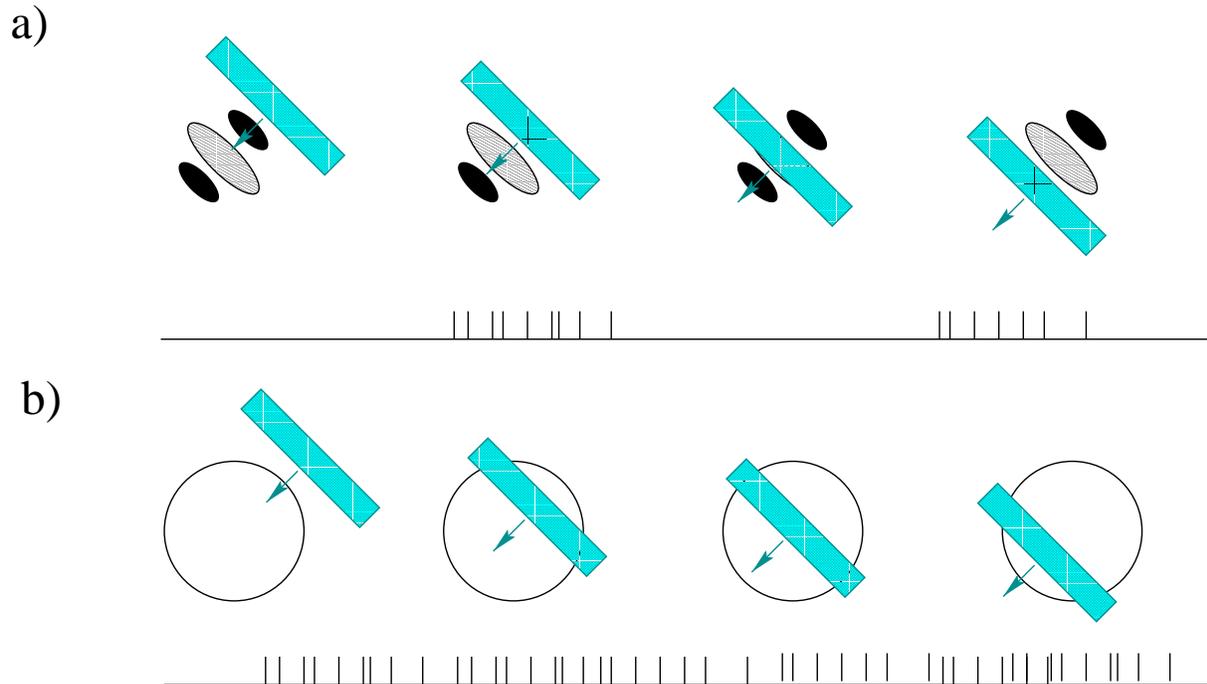


Abbildung 5.15: Neuronales Antwortverhalten: a.) *simple cells* und b.) *complex cells*. Unten ist jeweils eine denkbare Folge von Aktionspotentialen der Zellen gezeichnet. Hell schraffiert positive synaptische Gewichte, schwarz negative synaptische Gewichte. Während *simple cells* sehr selektiv auf die Position des Reizbalkens reagieren, antwortet eine *complex cell* während des ganzen Durchgangs gleichmäßig.

5.7 Ein Vergleich der Modellresultate mit biologischen Fakten

Das Verhalten der einzelnen Einheiten der ASSOM ähnelt dem Antwortverhalten kortikaler *complex cells* (siehe Abb. 5.15). Dieser Zelltyp antwortet auf Stimulierung selektiv bezüglich Orientierung und Position im visuellen Feld aber unabhängig von der Phase des Stimulus (siehe Abb. 5.14). Auch im Modell antwortet ein Punkt des Kortex auf einen Stimulus in Abhängigkeit seiner Orientierung und Position aber phasenunabhängig. Dabei ist festzustellen, dass einzelne Stellen des Kortex fast konstant auf alle Phaseneinstellungen antworten. Die rezeptiven Feldeigenschaften der beiden Basisvektoren stimmen dabei sowohl mit theoretischen Überlegungen überein, wie sie für das Zustandekommen rezeptiver Feldeigenschaften für komplexe Zellen schon früher gefordert wurden [Heeger, 1991, Yuille, 1989], als auch mit den physiologischen Gegebenheiten, wie sie tatsächlich in biologischen rezeptiven Feldern von einander benachbarten *simple cells* gefunden werden [Freeman et al., 1997].

Während nämlich die Eigenschaften rezeptiver Felder wie Orientierung, Position und Oku-

lardominanz, wie eingangs dieses Kapitels erwähnt, stetig auf die Kortexoberfläche abgebildet werden, gilt dies nicht für die lokale Phase. Zunächst wurde allerdings angenommen, dass sie sich in regelmäßigen Sprüngen von 90° anordnen würden [Pollen & Ronner, 1981, Liu et al., 1991]. Dagegen geht man heute davon aus, dass benachbarte Neuronen beliebige Verschiebungen der spatialen Phase aufweisen [Freeman et al., 1997]. Die Karten bildet also diese Eigenschaft nicht stetig ab.

Wie funktioniert nun eine komplexe Zelle? Die Einheiten des hier vorgestellten Modells sind hierarchisch aufgebaut. Zwei Antworten von Neuronen mit orthogonalen Phasen, die also von kortikalen *simple cells* stammen können, efferieren auf eine Verrechnungseinheit, die man als eine kortikale *complex cell* interpretieren kann. Damit befindet sich dieses Modell in Übereinstimmung mit theoretischen Überlegungen und Modellen, die in der Vergangenheit vorgeschlagen wurden [Yuille, 1989, Heeger, 1991, Chance et al., 1999]. Allerdings scheinen im biologischen Substrat diese Hierarchien nicht zu existieren [Alonso & Martinez, 1998]. Die *complex cells* erhalten eher direkt Zugang über genikulate Afferenzen bzw. deren Zwischenstationen als über *simple cells*. Daher scheint die Vermutung angebracht, dass sich *complex cells* und *simple cells* während der Lernphase zwar beeinflussen, sich möglicherweise gegenseitig optimieren, dass aber eine reife kortikale *complex cell* keine afferenten *simple cells* für ihr typisches Antwortverhalten benötigt, sondern die obengenannten Verrechnungsvorgänge intern bewältigt. Eine komplementäre Erklärungsmöglichkeit bieten Berechnungen auf dem Dendritenbaum [Mel & Archie, 2000]. Die Unterstützung des Lernens würde sich dann durch Hebbsches Lernen in einer strukturierten afferenten Verschaltung niederschlagen.

Typische Schnittwinkel zwischen Phasenkarte und Orientierungskarten werden beim AS-SOM nicht gemessen (siehe Abb. 5.10). Bei der selbstorganisierenden Karte gibt es eine schwache Präferenz für senkrechte Schnittwinkel. Bei biologischen Karten liegen zwei widersprüchliche Ergebnisse vor. Während eine Arbeit die Existenz typischer Schnittwinkel negiert [Kim et al., 1999], berichtet eine zweite von einer schwachen statistischen Präferenz für senkrechte Schnittwinkel zwischen Orientierungsgradienten und Wellenlängengradienten [Hübener et al., 1997].

Bei den hier vorgenommenen Simulationen hat die Entwicklung der Karten den stabilen Endzustand nicht mit Sicherheit erreicht. Stattdessen wurde ein intermediärer Zustand numerisch durch die exponentielle Abnahme der Lernrate γ quasi während des Lernprozesses eingefroren. Der Grund dafür liegt darin, dass der numerische Aufwand bei der Berechnung dieser hochdimensionalen Modelle erheblich ist. In den nachfolgenden Kapiteln soll daher das Modell reduziert werden, um die stabilen Zustände tatsächlich zu erhalten.

5.8 Metriken und adaptive Featureräume

Metrische Überlegungen zu einem Zusammenhang zwischen spatialer Phase und Retinotopie haben zu den richtigen rezeptiven Feldformen von *complex cells* im visuellen Kortex geführt. Anders als in Kapitel 3 standen hier metrische Überlegungen explizit im Vordergrund und haben in der Konsequenz zu einem alternativen Konzept eines Modells selbstorganisierender Karten beigetragen.

In der Zwischenzeit ist auch von anderer Seite [Hyvärinen & Hoyer, 2000] ein ähnliches Modell entwickelt worden, das ebenfalls auf adaptiven Unterräumen aufbaut, aber statt eines *winner-takes-all*-Prinzips auf einer *independent component analysis* basiert zu der in ähnlicher Weise wie hier beschrieben topographische Elemente hinzukommen. Wie in dem hier vorgestellten Ansatz ist das Ziel dieses Modells den Anspruch des Antwortverhalten kortikaler komplexer Zellen korrekt wiederzugeben.

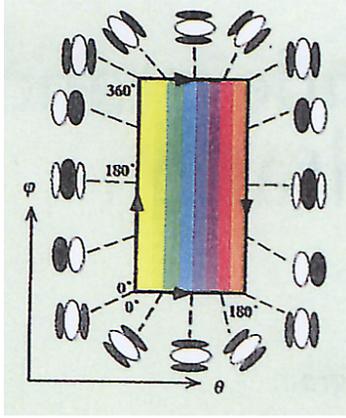
Die Wellenlänge des ASSOM-Modells ist fast doppelt so groß wie die des unter gleichen Bedingungen entstandenen einfachen Kartenmodells. Die Ursache dafür könnte darin liegen, dass die bei der selbstorganisierenden Karte dominierende Eigenschaft, nämlich die spatiale Phase, beim ASSOM nicht mehr in ihren verschiedenen Kombinationen mit den anderen Stimuluseigenschaften repräsentiert wird. Das effektive Volumen des Stimulusraumes verringert sich durch die Reduktion der spatialen Phase, während sich die effektiven Abstände zwischen zwei benachbarten Orientierungen nicht ändern.

Ein Blick auf die Einfaltung der ASSOM-Modellneuronen in das visuelle Feld gibt einen Hinweis darauf welche Eigenschaften besonders von den einfacheren Featureansätzen abweichen. Ist die Retinotopie der ASSOM auch im Groben erhalten (siehe Abb. 5.7), so lassen sich doch lokal etwa in der Wellenlänge der Orientierungskolumnen, erhebliche Störungen feststellen. Diese wurden auch schon in anderen hochdimensionalen Modellen gefunden. Schon weil die Wellenlängen dieser Störungen mit der Wellenlänge der Orientierungspräferenzen übereinstimmen, ist es naheliegend, anzunehmen, dass diese Störungen auf einer metrischen Kopplung zwischen der Position und Orientierung beruhen. Um eine Untersuchung der essentiellen Eigenschaften dieser Kopplung durchzuführen soll im folgenden Kapitel ein vereinfachtes Modell untersucht werden, das auf einem reduzierten Stimulusraum aufbaut, jedoch im Wesentlichen mit den gerade präsentierten Modellen übereinstimmt.

5.9 Die topologischen Eigenschaften des generischen rezeptiven Feldes in SOM und ASSOM

Die generischen rezeptiven Felder der Karten haben sowohl im Falle der selbstorganisierenden Karte als auch bei der ASSOM näherungsweise die Form zweidimensionaler Ga-

a)



b)

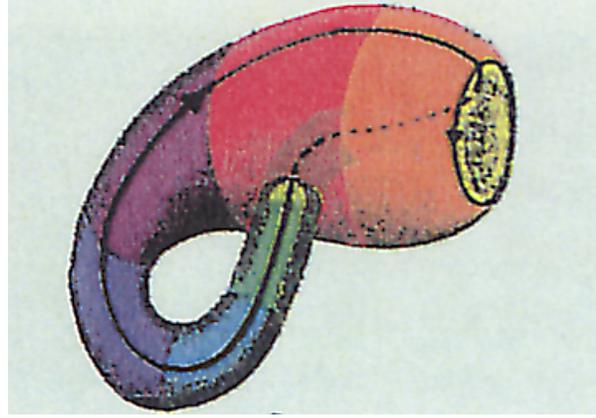


Abbildung 5.16: Kleinsche Flasche. Die kombinierte Variierung der präferierten Orientierung und Φ und der spatialen Phase Θ bei Gaborfiltern führt zu einer geschlossenen Manigfaltigkeit, die die Eigenschaften einer Kleinschen Flasche hat (aus [Tanaka, 1995]).

borfilter. Um die Vorgehensweise zu vereinfachen, soll nun angenommen werden, daß auch der generische Stimulus im Wesentlichen ein Gaborfilter ist. Stimulus und rezeptive Felder lassen sich also formal durch (vgl. Gl. (3.14)):

$$\mathbf{G} = G(\xi_1, \xi_2) = g_w \times g_g \quad (5.14)$$

mit

$$g_w = \cos \left(\omega \times \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cdot \hat{T}(\Phi_1) \cdot \left(\begin{pmatrix} \xi_1 - \bar{\xi}_1 \\ \xi_2 - \bar{\xi}_2 \end{pmatrix} + \Theta \right) \right) \quad (5.15)$$

und

$$g_g = \exp \left(- \left| \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \cdot \hat{T}(\Phi_2) \cdot \begin{pmatrix} \xi_1 - \bar{\xi}_1 \\ \xi_2 - \bar{\xi}_2 \end{pmatrix} \right|^2 \right) \quad (5.16)$$

beschreiben. Dabei soll die Orientierung Φ durch eine koinzidierende Ausrichtung Φ_1 und Φ_2 festgelegt sein. Zudem werden für Stimulus und Karte konstante Werte für b_1 und b_2 angenommen, die beide gleich groß sein sollen. Weiterhin soll von einer Variation der spatialen Frequenz und der Position des Stimulus im visuellen Feld abgesehen werden.

Daraus ergeben sich folgende Fragen. Welche topologischen Eigenschaften hat der Stimulusraum dieser vereinfachten Beschreibung? Wie unterscheiden sich die topologischen Eigenschaften des Stimulusraums von SOM und ASSOM?

Im Falle der selbstorganisierenden Karte ist der um die Positionsvariation reduzierte Stimulusraum eine Manigfaltigkeit, die von den beiden periodischen Variablen lokale Phase Θ und Orientierung Φ abhängt, wobei die spatiale Phase 2π periodisch und die Orientierung π periodisch ist. In Abb. 5.16 ist dargestellt, wie sich die beiden Variablen Θ und Φ auf dieser Fläche darstellen lassen. Versucht man nun die Ränder in der Weise aneinanderzuheften, dass eine an allen Punkten stetige Abbildung entsteht, so müssen der eine vertikale Rand an der Horizontalen gespiegelt an den linken geheftet werden. Der obere horizontale Rand kann dagegen ohne Spiegelung an den unteren horizontalen geheftet werden. Eine solche Manigfaltigkeit ist homöomorph zu einer Kleinschen Flasche. Der Bezug zu Eigenschaften rezeptiver Felder wurde zuerst von S. Tanaka [Tanaka, 1995] dargestellt. Die erwähnten neuen Ergebnisse zur Repräsentation der räumlichen Phase lassen die an sich interessanten Betrachtungen von Tanaka jedoch als irrelevant erscheinen. Bei der Karte linearer adaptiver Unterräume reduziert sich die Darstellung von Phase und Orientierung des Stimulusraumes durch den Wegfall der spatialen Phase nämlich auf einen eindimensionalen π -periodischen. Alle Phasen sind in einer einzelnen Einheit des Modells vereinigt. Die weiteren topologischen und metrischen Auswirkungen linearer adaptiver Unterräume können formal mit Hilfe eines Abstandsmaßes untersucht werden, das den Eigenschaften linearer adaptiver Unterräume entspricht. Ein plausibler Ansatz ist der in Gl. (5.5) gegebene:

$$D^2(\mathbf{V}, (\mathbf{W})_r) = \min_{\alpha_l} \left(\left| \mathbf{V} - \sum_L \alpha_l \mathbf{W}_{r,l} \right| \right)^2 \quad (5.17)$$

Zwei Annahmen soll nun gemacht werden: Erstens sollen sich sowohl die Stimuli als auch die rezeptiven Felder auf die oben genannten Gaborfilter reduzieren lassen. Zweitens soll die lokale Phase zwischen den beiden rezeptiven Feldvektoren \mathbf{W}_1 und \mathbf{W}_2 einer Einheit um etwa 90° differieren. Die anderen Eigenschaften des Gaborfilters sollen in beiden Basisvektoren gleich sein. Das führt dazu, dass durch die Variation der α_l im Wesentlichen die spatiale Phase variiert wird. Damit wird das Abstandsmaß äquivalent zu

$$D^2(\mathbf{V}, (\mathbf{W})_r) = \min_{\Theta} (\mathbf{G}(\mathbf{v}) - \mathbf{G}(\mathbf{w}, \Theta))^2, \quad (5.18)$$

wobei die Eigenschaftsvektoren \mathbf{v}, \mathbf{w} lokale Phase, Orientierung und Position wiedergeben sollen. Weiterhin definieren wir

$$\Theta_{\min} = \operatorname{argmin}_{\Theta} (\mathbf{G}(\mathbf{v}) - \mathbf{G}(\mathbf{w}, \Theta))^2. \quad (5.19)$$

Da Spiegelungen eines Gaborfilters längs seiner Orientierung zu einem Gaborfilter führen, der die inverse spatiale Phase aufweist, ändert sich bei gleichzeitiger Spiegelung beider rezeptiver Felder der Abstand nach Gl. (5.19) nicht. Lediglich Θ_{\min} transformiert sich zu $-\Theta_{\min}$. Treten alle spatialen Phasen gleichmäßig in der Stimulierung auf, so kann argumentiert werden, daß auch der effektive Stimulus diese Symmetrie aufweist. Weiterhin

ist ein Gaborfilter spiegelsymmetrisch gegenüber Siegelungen quer zu seiner präferierten Orientierung. Diese Symmetrie führt nicht zu einer Veränderung Θ_{\min} .

Diese beiden Eigenschaften der Metrik bezüglich des Stimulusraumes für das Modell adaptiver Unterräume sollen nun festgehalten werden:

1. Die Metrik ist spiegelsymmetrisch bezüglich Spiegelungen der Positionskordinaten längs und quer zur präferierten Orientierung.
2. Die Metrik ist π - periodisch bezüglich der präferierten Orientierung.

Im nächsten Kapitel werden diese Eigenschaften benutzt.

In der Darstellung der Analysemethoden in Abschnitt 3.7 wurde der Begriff Orientierungsstärke als Messgröße eingeführt, die ein Maß für die Signifikanz einer ermittelten Orientierung darstellt. In den bisher vorgestellten Modellen ließ sich die Variation der Orientierungsstärke schon auf mindestens zwei Ursachen zurückführen (vgl. Abb. 5.5 und 5.12). Einerseits erwiesen sich Bereiche mit weniger großen spatialen Wellenlängen als weniger orientierungsselektiv. Es gab andererseits aber auch solche Stellen, die Gaborfiltern unähnlich waren. Im Folgenden soll auf die einzelnen Ursachen der Orientierungsselektivität nicht mehr eingegangen werden, d.h. die Stärke der Orientierungsselektivität wird lediglich als skalarer Modellparameter einbezogen, siehe Abb. 5.17.

Die Defizite des vereinfachten Modells können durch eine kompliziertere metrische Struktur ausgeglichen werden. Bevor im Kapitel 7 auf die Karteneigenschaften eingegangen wird, werden nun die Bedingungen für eine dem Kartenbildungsprozess zugrundeliegende Metrik abgeleitet.

5.10 Zusammenfassung

Die Erweiterung des Modells aus dem 3. Kapitel auf die klassische selbstorganisierende Karte führt zu topographischen Problemen. So wurde aus metrischen Gründen die normale hochdimensionale selbstorganisierende Karte durch ein topographisches Modell das lineare Unterräume adaptiert (ASSOM) ersetzt. Dieses Modell vermag sowohl die retinotopie Anordnung der kortikalen rezeptiven Felder weitgehend zu erhalten als auch des Antwortverhalten komplexer Zellen im visuellen Kortex zu reproduzieren.

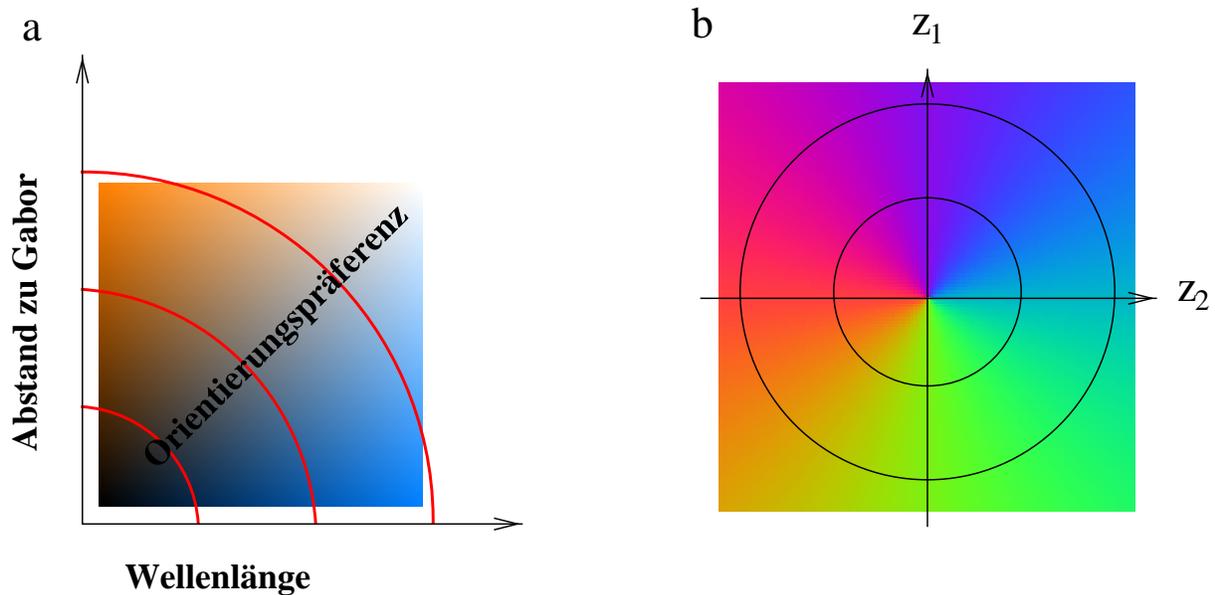


Abbildung 5.17: Die Vereinfachung des Modells durch die Einführung eines skalaren Orientierungsparameters. **a)** Die Orientierungsselektivität hängt in den Modellen, die natürliche Bilder als Stimulierung verwenden, von verschiedenen Ursachen ab. In den in diesem Kapitel präsentierten Modellen wurden mindestens zwei Ursachen beobachtet (vgl. Abb. 5.5 und 5.12): erstens Die Variation der spatialen Wellenlänge (in diesem Schema horizontal) und zweitens die Abweichung vom Aussehen eines Gaborfilters (vertikal). Die Kombination von Abweichung vom nächsten Gaborfilter und Variation der Wellenlänge führt zu unterschiedlichen Orientierungsstärken, wobei die ungefähren Konturen der Orientierungsstärke rot eingezeichnet sind. **b)** Der kombinierte Raum der Orientierungsstärke und der Orientierung soll im Folgenden als homöomorph zu \mathbb{C} bzw. \mathbb{R}^2 gedacht werden, wobei die Transformationsgleichung $(S\cos(2\Phi), S\sin(2\Phi)) \rightarrow (z_1, z_2)$ ist.

Kapitel 6

Projektion hochdimensionaler Karten in den Featureraum

Am Ende des Kapitels 5 wurden topologische Überlegungen dazu angestellt, wie eine in beiden Richtungen stetige Abbildung eines hochdimensionalen Reizensembles in bestimmte Unterräume des Eigenschaftsraumes aussehen könnte.

Nun sollen, analog zu den topologischen die metrischen Eigenschaften untersucht werden. Es soll also die Frage beantwortet werden, wie die Abstände bei einer Projektion in den Eigenschaftsraum zu transformieren sind. Es stellt sich heraus, dass die Kreisscheibe eine mit Einschränkungen akzeptable Darstellung der Orientierungseigenschaften ergibt und dass außerdem ein zweidimensionaler Raum für die Darstellung der möglichen Positionen notwendig ist. Es lässt sich dann unter Annahme einiger plausibler Symmetrien, ein infinitesimales Abstandsmaß finden, welches für eine große Klasse von Stimulusräumen geeignet ist.

6.1 Grundlegende Überlegungen

Bis hierher wurden Untersuchungen in einer Modellklasse vorgenommen, die im Allgemeinen als hochdimensionale Modelle bezeichnet werden. Die Vertreter der dazu komplementären Klasse werden Featuremodelle genannt. Zunächst soll der wesentliche Unterschied dargestellt werden.

Wie bereits diskutiert, werden bei hochdimensionalen Ansätzen die Stimulusvektoren \mathbf{V} als Aktivitätsverteilung über viele parallele Kanäle ξ aus genikulaten Afferenzen V_ξ repräsentiert, wobei jeder Kanal eine Dimension des Vektors darstellt. Konkret resultiert ein Stimulus von 16×16 Pixeln in einen 256 dimensionalen Vektor. Die Aktivierung $a_{\mathbf{r}}$ eines

kortikalen Neurons an der Stelle \mathbf{r} wird durch den Überlapp zwischen Stimulus \mathbf{V} und dem rezeptivem Feldvektor $\mathbf{W}_{\mathbf{r}}$

$$a_{\mathbf{r}} = (\mathbf{V} \cdot \mathbf{W}_{\mathbf{r}}). \quad (6.1)$$

definiert, beziehungsweise, im Falle des Modells adaptiver Unterräume, durch

$$a_{\mathbf{r}} = \sum_L (\mathbf{V} \cdot \mathbf{W}_{l,\mathbf{r}}). \quad (6.2)$$

Das Ziel der neuronalen Optimierung ist es, diesen Überlapp zu maximieren. Wie schon gezeigt, ist dieser Überlapp im normierten Fall dem Abstand $D(\mathbf{V}, (\mathbf{W})_{\mathbf{r}})$ entgegengesetzt proportional. Dabei ist die Stimulusmanigfaltigkeit \mathcal{S} formal Ξ -dimensional, wobei Ξ die Zahl der Kanäle bezeichnet.

Die Aktivitäten treten in diesen Kanälen keineswegs unabhängig voneinander auf. Vielmehr sind die Aktivitäten stark korreliert, wodurch biologische Neuronen im Falle einer normalen Entwicklung zumeist gaborfilterartige rezeptive Felder entwickeln. Die effektive Dimension dieser Strukturen liegt dabei deutlich unter der Anzahl der verwendeten Kanäle. Mit anderen Worten, die Stimulusmanigfaltigkeit \mathcal{S} ist lediglich ein Unterraum des Raums aller theoretischen Aktivitätsverteilungen der Kanäle, der eine deutlich niedrigere Dimension hat als die Zahl der Pixel. Featurekartenmodelle benutzen diese Tatsache und repräsentieren auch intern Stimuli und rezeptive Felder im niedrigdimensionalen Eigenschaftsraum s .

Die Struktur der relevanten Dimensionen, die s bestimmen, kann aus den Eigenschaften der Stimulussets im hochdimensionalen Raum berechnet werden. Zunächst sollen allgemeine Eigenschaften aus den Symmetrien der Stimulusmenge extrahiert werden.

6.2 Kodierung der Stimuli im Featureraum

Für die folgende Darstellung reicht zur Beschreibung von Stimulus und rezeptiven Feldeigenschaften eine Beschränkung auf vierdimensionale Vektoren aus. Zwei Parameter entfallen auf die Orientierung und zwei auf die Position des Stimulus bzw. des rezeptiven Felds.

Für die Position des Stimulus können die beiden Positionskoordinaten $(\bar{\xi}_1, \bar{\xi}_2)$ verwendet werden. Bei der Orientierungsstärke Z und der Orientierung Φ , die π periodisch ist, eignet sich die Kodierung

$$z_1 = Z \cos(2\Phi) \quad (6.3)$$

$$z_2 = Z \sin(2\Phi). \quad (6.4)$$

Diese Darstellung erscheint deshalb sinnvoll, weil sie stetig die π -periodische Größe der Orientierung auf den 2π -periodischen Richtungsvektor (z_1, z_2) abzubilden vermag.

6.3 Eine Modifikation des elastischen Netzes

Bevor die metrischen Eigenschaften des Featureraumes detailliert untersucht werden, sollen in diesem und im folgenden Abschnitt die Adaption der hochdimensionalen Kartenmodelle auf den niedrigdimensionalen Fall vorgenommen werden. Zunächst wird das elastische Netz betrachtet. Dessen Kostenfunktion (vgl. Gl. (4.8))

$$E = \sigma_{EL}^2 E \left[\log \left(\sum_{\mathcal{N}} \exp \left(-\frac{(\mathbf{V} - \mathbf{W}_{\mathbf{r}})^2}{2\sigma_{EL}^2} \right) d\mathbf{r} \right) \right]_{\mathcal{S}} + \frac{\eta}{2} \sum_{\mathcal{N}} \sum_{|\mathbf{r}_2 - \mathbf{r}|=1} (\mathbf{W}_{\mathbf{r}_2} - \mathbf{W}_{\mathbf{r}})^2 \quad (6.5)$$

enthält die Vektoren der rezeptiven Felder in zwei Termen. In beiden Fällen handelt es sich um quadratische euklidische Abstände. Im ersten Fall um den quadratischen Abstand zwischen dem Stimulus \mathbf{V} und dem rezeptiven Feld $\mathbf{W}_{\mathbf{r}}$

$$(\mathbf{V} - \mathbf{W}_{\mathbf{r}})^2 = D^2(\mathbf{V}, \mathbf{W}_{\mathbf{r}}). \quad (6.6)$$

Im zweiten Fall um den Abstand aus benachbarten rezeptiven Feldern

$$\sum_{|\mathbf{r}_2 - \mathbf{r}|=1} (\mathbf{W}_{\mathbf{r}_2} - \mathbf{W}_{\mathbf{r}})^2. \quad (6.7)$$

Nimmt man an, dass sich sowohl alle hochdimensionalen Stimuli \mathbf{V} als auch die rezeptiven Felder $\mathbf{W}_{\mathbf{r}}$ formal durch die niedrigdimensionalen Eigenschaftsvektoren \mathbf{v} und $\mathbf{w}_{\mathbf{r}}$ mit $\mathbf{V} = G(\mathbf{v})$ und $\mathbf{W} = G(\mathbf{w})$ ausdrücken lassen, so sollte schon die Kenntnis des Abstandsmaßes

$$D(\mathbf{V}, \mathbf{W}_{\mathbf{r}}) = D(G(\mathbf{v}), G(\mathbf{w})) \quad (6.8)$$

ausreichen, um das Minimum der Kostenfunktion zu bestimmen. Es ist gleichwertig, die Distanzfunktion auf den Featureraum herunterzuziehen, das heißt

$$D(G(\mathbf{v}), G(\mathbf{w})) = d(\mathbf{v}, \mathbf{w}). \quad (6.9)$$

Damit wird die Kostenfunktion im elastischen Netz zu

$$E = \sigma_{EL}^2 \mathbb{E} \left[\log \left(\int_{\mathcal{N}} \exp \left(-\frac{d^2(\mathbf{v}, \mathbf{w}_r)}{2 \sigma_{EL}^2} \right) d\mathbf{r} \right) \right]_{\mathcal{S}} + \frac{\eta}{2} \sum_{\mathcal{N}} \sum_{|\mathbf{r}_2 - \mathbf{r}|=1} d^2(\mathbf{w}_r, \mathbf{w}_r) . \quad (6.10)$$

Ist die Distanz klein, so genügt es, die quadratischen Terme des Abstandsmaßes, das heißt die ersten nichtverschwindenden Terme einer Polynomreihenentwicklung zu betrachten. Da eine Metrik bei identischen Positionen verschwindet und obendrein eine gerade Funktion der beiden Positionsvektoren ist, sind die ersten Entwicklungsterme, die nicht verschwinden mindestens zweiter Ordnung. Ist tatsächlich schon die zweite Ordnung relevant, handelt es sich um eine Riemannsche Metrik. Im Falle kleiner Abstände ist also

$$d^2(\mathbf{w}, \mathbf{w} + \Delta \mathbf{w}) \approx \Delta \mathbf{w} g(\mathbf{w}) \Delta \mathbf{w}, \quad (6.11)$$

wobei (g) eine quadratische Matrix ist, die im Grenzfall infinitesimaler Abstände einen metrischen Tensor darstellt. In den numerischen Simulationen soll angenommen werden, dass die Abstände zwischen Stimulus und rezeptivem Feld ebenfalls gut durch ein quadratisches Abstandsmaß genähert werden können.

Eine Transformation des Stimulusraumes führt immer zu einer entsprechenden kovarianten Transformation der Metrik. Gesetzt den Fall, der stabile Endzustand, welchen die Karte unter \mathcal{S} schließlich einnimmt, ließe sich auch auf die Mannigfaltigkeit \mathcal{S}' abbilden, so lässt sich ein Tupel (\mathcal{S}', D') finden, das zum identischen Endzustand des Algorithmus wie unter (\mathcal{S}, D) führt. Dabei kann sich jedoch die Dynamik des Gradientenabstieges in Abhängigkeit der Darstellung der Mannigfaltigkeit deutlich unterscheiden.

6.4 Entsprechendes zum Kohonenalgorithmus

Da der Kohonenalgorithmus nicht über eine Kostenfunktion verfügt, die zu minimieren wäre, kann der Endzustand nicht über ein Minimum definiert werden. Es bleibt, die kovariante Dynamik gegenüber der Darstellungstransformationen des Featureerraumes zu berechnen. Es ist naheliegend, von einer Darstellung der Dynamik

$$\partial_t \mathbf{W}_r = F(\mathbf{W}, \mathcal{S}) \quad (6.12)$$

auszugehen und durch Koordinatentransformation $\mathbf{W}(\mathbf{w}) \rightarrow \mathbf{w}$ zur Dynamik

$$\partial_t \mathbf{w}_r = \frac{\partial \mathbf{w}_r}{\partial \mathbf{W}_r} F(\mathbf{W}(\mathbf{w}), \mathcal{S}) = f(\mathbf{w}) \quad (6.13)$$

überzugehen. Dieser Weg ist jedoch aus verschiedenen Gründen meist praktisch nicht anwendbar (jedoch siehe [Thomas & Cowan, 2000]). Ist das Distanzmaß im Featureerraum bekannt (vgl. Abschnitt 6.5), so kann man einen anderen Weg beschreiten. Die Lernregel im

hochdimensionalen Kohonenmodell (vgl. Gl. (4.3)) lässt sich als

$$\partial_t \mathbf{W} = \gamma (\mathbf{V} - \mathbf{W}_r) \exp(-(\mathbf{r} - \mathbf{r}^*)^2 / 2\sigma_K^2) \quad (6.14)$$

schreiben, wobei

$$\mathbf{r}^* = \operatorname{argmin}_r [(\mathbf{V} - \mathbf{W}_r)^2]. \quad (6.15)$$

Gl. (6.14) kann durch

$$\partial_t \mathbf{W} = \gamma \mathbf{E} D(\mathbf{V}, \mathbf{W}_r) \exp(-\frac{(\mathbf{r} - \mathbf{r}^*)^2}{2\sigma_K^2}) \quad (6.16)$$

ausgedrückt werden. Hierbei stellt

$$\mathbf{E} = \frac{\mathbf{V} - \mathbf{W}}{|\mathbf{V} - \mathbf{W}|} \quad (6.17)$$

den normierten Richtungsvektor in Richtung der stärksten Abnahme der Distanz im hochdimensionalen Raum dar. Diese sollte mit der Richtung der stärksten Abnahme der Distanz im Featureerraum übereinstimmen. So kann die korrespondierende Dynamik im Featureerraum

$$\partial_t \mathbf{w}_r = \gamma \mathbf{e} d(\mathbf{v}, \mathbf{w}_r) \exp(-(\mathbf{r} - \mathbf{r}^*)^2 / 2\sigma_K^2) \quad (6.18)$$

geschrieben werden. \mathbf{e} stellt wie \mathbf{E} einen Einheitsvektor in Richtung der stärksten Abnahme, jetzt allerdings von $d(\mathbf{v}, \mathbf{w}_r)$, dar. \mathbf{e} lässt sich dann aus

$$\tilde{\mathbf{e}} = -(g)^{-1} \cdot \frac{\partial d(\mathbf{v}, \mathbf{w}_r)}{\partial \mathbf{w}_r}, \quad (6.19)$$

mit $(g)^{-1}$ als dem inversen metrischen Tensor, zu

$$\mathbf{e} = \frac{\tilde{\mathbf{e}}}{|\tilde{\mathbf{e}}|} \quad (6.20)$$

berechnen. Die Gl. (6.19) leitet sich aus der Minimierung der Kostenfunktion (siehe [Amari 2000] für ein identisches Vorgehen im Falle des natürlichen Gradientenabstiegs, *natural gradient descent*)

$$E = \frac{\partial d(\mathbf{v}, \mathbf{w}_r)}{\partial \mathbf{w}_r} \cdot \tilde{\mathbf{e}} + \lambda (\tilde{\mathbf{e}}^T (g) \tilde{\mathbf{e}} - N) \quad (6.21)$$

ab, wobei in diesem Falle λ einen Lagrangemultiplikator darstellt, der die Normierung von $\tilde{\mathbf{e}}$ auf den Betrag N sicherstellt.

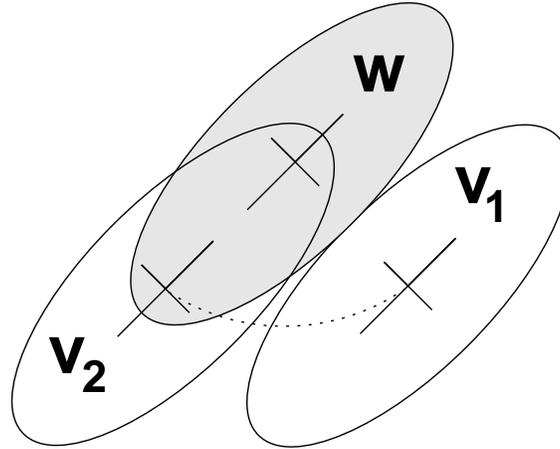


Abbildung 6.1: Zwei Stimuli und ein rezeptives Feld in einem hochdimensionalen Modell (schematisch). Obwohl beide Stimuli die gleiche Distanz in retinalen Koordinaten aufweisen ist der pixelweise Überlapp von v_1 mit w kleiner als der zwischen v_2 und w . Der Grund dafür liegt darin, dass einer der Stimuli zum rezeptiven Feld längs seiner Orientierung disloziert ist, während der andere quer dazu verschoben ist. Die Kreuze inmitten der Stimuli respektive des rezeptiven Feldes geben ihre Symmetrieachsen wieder.

6.5 Symmetrieüberlegungen zur Metrik

Es soll nun das Ziel sein, zu verschiedenen hochdimensionalen Stimulusmanigfaltigkeiten die jeweilige Metrik der Featureerräume zu berechnen. Im Stimulusraum hat die infinitesimale Abstandsgleichung für zwei vierdimensionale Eigenschaftsvektoren $\mathbf{u}_1, \mathbf{u}_2$, die zueinander den Abstand $d\mathbf{u} = (d\xi, dz) = (d\xi_1, d\xi_2, dz_1, dz_2)$ haben, grundsätzlich folgendes Aussehen:

$$d(\mathbf{u}_1, \mathbf{u}_2) = g_{\xi\xi}d\xi^2 + 2g_{\xi z}d\xi dz + g_{zz}dz^2. \quad (6.22)$$

(Statt der Bezeichnungen $(\bar{\xi}_1, \bar{\xi}_2)$ sollen in den verbleibenden Kapiteln die Bezeichnungen (ξ_1, ξ_2) für die Schwerpunkte, also die Positionen der rezeptiven Felder bzw. der Stimuli im visuellen Feld verwendet werden.) Die Zerlegung soll der Übersichtlichkeit dienen. Der gesamte metrische Tensor setzt sich dann zusammen als

$$(g) = \begin{pmatrix} g_{\xi\xi} & g_{\xi z} \\ g_{\xi z} & g_{zz} \end{pmatrix}. \quad (6.23)$$

Zunächst soll versucht werden, durch Einschränkung der Modellklasse einige generelle Aussagen über die Eigenschaften des metrischen Tensors zu machen. Diese Einschränkungen sind Symmetrieannahmen, wie sie in den gängigen hochdimensionalen Modellansätzen vorkommen und die im visuellen Kortex plausibel erscheinen. Alle bisherigen Featureansätze

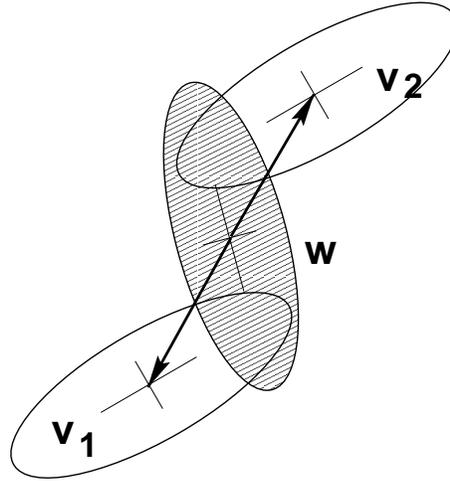


Abbildung 6.2: Invarianz gegen Positionsinversionen. Da die Orientierung von \mathbf{v}_1 und \mathbf{v}_2 gleich sind, der Betrag der Dislozierung übereinstimmt, und außerdem Stimulus und rezeptives Feld π -symmetrische Strukturen sind, muss die Abstandsfunktion d für $d(\mathbf{v}_1, \mathbf{w})$ gleich $d(\mathbf{v}_2, \mathbf{w})$ sein.

nehmen implizit ein euklidisches Abstandsmaß an, also

$$g_{\xi\xi} = g_{zz} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad (6.24)$$

$$g_{\xi z} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (6.25)$$

Dieses vermag aber bestimmte Eigenschaften des Stimulusraumes, wie sie in hochdimensionalen Karten auftreten, nicht zu reproduzieren. Am anschaulichsten lassen sich die dabei auftretenden Phänomene am Beispiel einer hochdimensionalen selbstorganisierenden Karte mit rezeptiven Feldern verstehen, die eine zigarrenartige Form haben. Die Richtung der Zigarre indiziert die Orientierung. Solche Stimulusräume waren schon Gegenstand von Untersuchungen zu hochdimensionalen Karten [Obermayer & Blasdel, 1997]. Man stelle sich nun ein Stimuluspaar $\mathbf{v}_1, \mathbf{v}_2$ vor, das dem rezeptiven Feld ähnelt (siehe Abb. 6.1). Der Unterschied besteht darin, daß \mathbf{v}_1 quer zur gemeinsamen Orientierung und \mathbf{v}_2 längs der Orientierung disloziert ist. Der Abstand in retinalen Koordinaten des visuellen Feldes sei in beiden Fällen gleich. Das bedeutet, dass im Falle eines euklidischen Abstandsmaßes im Feature Raum der Abstand in beiden Fällen gleich sein muss, denn $d\xi^2$ ist in beiden Fällen gleich und dz^2 jeweils Null. Dagegen wird der Abstand zwischen zwei hochdimensionalen Vektoren durch den Überlapp bestimmt. Dieser ist bei dem Vektorpaar \mathbf{v}_2 und \mathbf{w} wesentlich größer als im Fall \mathbf{v}_1 und \mathbf{w} .

Um ein Abstandsmaß im Feature Raum zu erhalten, das in der Lage ist, hochdimensionale Überlapps durch Featureabstände auszudrücken, muss letztlich über die Aktivitätsverteilung integriert werden.

Zunächst soll der Frage nachgegangen werden, in welcher Weise sich der allgemeine metrische Tensor einschränken lässt. Dazu sollen folgende Symmetrieüberlegungen angestellt werden.

1. Betrachtet wird zuerst die Invarianz der Energiefunktion gegenüber Translationen im visuellen Feld. Zumindest in den hier vorgestellten Modellen ist diese erfüllt. Biologisch mag das fragwürdig erscheinen, denn bekannterweise ist die Repräsentation des visuellen Feldes in der Retina, wie auch im visuellen Kortex stark verzerrt. Jedoch spielen sich diese Verzerrungen auf größeren Längenskalen als die Orientierungskolumnen ab. Deshalb soll im Rahmen dieses Modells – wie auch der meisten anderen [Obermayer et al., 1992, Obermayer et al., 1990, Wolf, 1998, Wolf & Geisel 1998] – auf die Berücksichtigung von Inhomogenitäten verzichtet werden. So mag das Modell als Simulation im fovealen Bereich interpretiert werden, wo die Effekte der Verzerrung am schwächsten sind. Daraus folgt, dass

$$d(\boldsymbol{\xi}_1, \boldsymbol{\xi}_2) = d(\boldsymbol{\xi}_1 - \boldsymbol{\xi}_2) \quad (6.26)$$

lediglich eine Funktion der Distanz der beiden Positionskoordinaten sein darf. Aufgrund der infinitesimalen Formulierung der Metrik darf (g) auch keine explizite Funktion von $d\boldsymbol{\xi}$ enthalten. Es treten also weder die Positionskoordinaten $\boldsymbol{\xi}$ noch ihre Verrückung $d\boldsymbol{\xi}$ im metrischen Tensor auf.

2. Der Abstand d soll invariant gegenüber der Inversion der retinalen Positionsdifferenz sein (vgl. Abb. 6.2), d. h.

$$d(d\boldsymbol{\xi}) = d(-d\boldsymbol{\xi}). \quad (6.27)$$

Daraus folgt, dass die gemischten Terme $g_{\boldsymbol{\xi}\mathbf{z}}d\boldsymbol{\xi} d\mathbf{z}$ verschwinden müssen, da die Metrik dann keine ungerade Funktion in $d\boldsymbol{\xi}$ sein darf. Daher ergibt sich

$$g_{\boldsymbol{\xi}\mathbf{z}} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \quad (6.28)$$

3. Weiterhin wird Invarianz gegen Drehungen des Gesamtsystems gefordert. Die Metrik sollte also mindestens für $\Phi_a \equiv \Phi_b$ invariant gegenüber der Koordinatentransformation

$$(\boldsymbol{\xi}, \mathbf{z}) \cdot \begin{pmatrix} \hat{\mathbf{T}}_{\Phi_a} & 0 \\ 0 & \hat{\mathbf{T}}_{2\Phi_b} \end{pmatrix} \rightarrow (\boldsymbol{\xi}, \mathbf{z}) \quad (6.29)$$

sein, wobei \hat{T}_{Φ} zweidimensionale Rotationen bezeichnen.

Zunächst erfüllt die Metrik

$$\tilde{d}(\mathbf{d}\mathbf{u}) = \mathbf{d}\mathbf{u} \tilde{g} \mathbf{d}\mathbf{u} \quad (6.30)$$

mit

$$\tilde{g} = \begin{pmatrix} \nu & 0 & 0 & 0 \\ 0 & \nu & 0 & 0 \\ 0 & 0 & \delta & 0 \\ 0 & 0 & 0 & \delta \end{pmatrix} \quad (6.31)$$

die genannten Symmetrien. Es handelt sich hierbei um die allgemeine ungekoppelte Metrik, die diagonal sein muss. Aufgrund der letzten geforderten Symmetrie darf sie nur die beiden freien Funktionen δ und ν abhängen. Dieses Abstandsmaß erfüllt allerdings noch eine weitere Symmetrie. Beispielsweise ließen sich hier die rezeptiven Felder \mathbf{w}_1 und \mathbf{w}_2 (in Abb. 6.1) ohne Änderung des Abstandes ineinander überführen. Es lassen sich also auch Transformationen nach Gl. (6.29) durchführen, bei denen $\Phi_a \neq \Phi_b$ ist. Um der Situation im Hochdimensionalen gerecht zu werden, führe ich nun Kopplungen zwischen $d\mathbf{z}$ bzw. $d\xi$ mit \mathbf{z} ein. Aufgrund der dritten Symmetrie, die streng nur für $\Phi_a = \Phi_b$ gelten soll, ist nur eine Kopplung der Art

$$c_1 = f(|\mathbf{z}|, |d\xi|, 2\chi - 2\Phi) \quad (6.32)$$

zwischen $d\xi$ und \mathbf{z} möglich. Dabei ist

$$d\xi = |d\xi| \begin{pmatrix} \cos(\chi) \\ \sin(\chi) \end{pmatrix} \quad (6.33)$$

und wie definiert

$$\mathbf{z} = |\mathbf{z}| \begin{pmatrix} \cos(2\Phi) \\ \sin(2\Phi) \end{pmatrix}. \quad (6.34)$$

Es muss nun berücksichtigt werden, dass es sich bei der Kopplung von $d\xi$ und \mathbf{z} um eine Kopplung eines 2π -periodischen und eines π periodischen Vektors handelt. Das macht die Hilfskonstruktion

$$d\tilde{\xi} = |d\xi|^2 \begin{pmatrix} \cos(2\chi) \\ \sin(2\chi) \end{pmatrix} = \begin{pmatrix} d\xi_1^2 - d\xi_2^2 \\ 2d\xi_1 d\xi_2 \end{pmatrix} \quad (6.35)$$

notwendig. Da

$$2\chi - 2\Phi = \arccos \left(\frac{d\tilde{\xi}}{|d\tilde{\xi}|^2} \cdot \frac{\mathbf{z}}{|\mathbf{z}|} \right) \quad (6.36)$$

ist, lässt sich Gl. (6.32) zu

$$c_1 = f \left(|\mathbf{z}|, |d\xi|, \arccos \left(\frac{d\tilde{\xi}}{|d\tilde{\xi}|^2} \cdot \frac{\mathbf{z}}{|\mathbf{z}|} \right) \right) = \tilde{f} \left(|\mathbf{z}|, |d\xi|, d\tilde{\xi} \cdot \mathbf{z} \right) \quad (6.37)$$

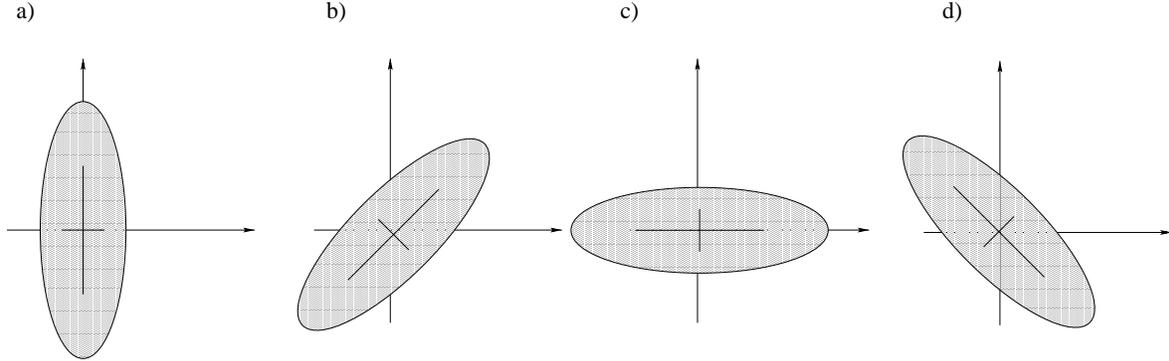


Abbildung 6.3: Die Festlegung der Kodierung von \mathbf{z} , die bei Gl. (6.44) vorausgesetzt ist. Die Pfeile bezeichnen die Koordinatenachsen x_1 und x_2 des visuellen Feldes. Die Darstellungen der Orientierungen im Featureerraum muss dann für $\mathbf{z} = (z_1, z_2)$ lauten a) $(|\mathbf{z}|, 0)$, b) $(0, |\mathbf{z}|)$, c) $(-\mathbf{z}, 0)$, d) $(0, -|\mathbf{z}|)$.

vereinfachen. Da der durch den metrischen Tensor dargestellte infinitesimale Abstand ausschließlich quadratische Formen von $d\xi$ und $d\mathbf{z}$ enthalten darf, verbleibt als einzig möglicher Ausdruck für c_1 der quadratische Entwicklungsterm der Polynomentwicklung in $d\xi$ von Gl. (6.37):

$$c_1 = \mu(|\mathbf{z}|) \begin{pmatrix} d\xi_1^2 - d\xi_2^2 \\ 2d\xi_1 d\xi_2 \end{pmatrix} \cdot \hat{\mathbf{T}}(\Phi') \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}^T, \quad (6.38)$$

wobei $\hat{\mathbf{T}}(\Phi')$ eine zunächst nicht näher bestimmte Drehmatrix einer Drehung um den Winkel θ darstellen kann.

Offensichtlich lässt sich diese Kopplung auf den vereinfachten Ausdruck

$$c_1 = \mu(|\mathbf{z}|) \begin{pmatrix} d\xi_1^2 - d\xi_2^2 \\ 2d\xi_1 d\xi_2 \end{pmatrix} \cdot \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}^T \quad (6.39)$$

reduzieren, wenn man die Orientierung entsprechend

$$\hat{\mathbf{T}}(-\Phi') \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}^T \rightarrow \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}^T \quad (6.40)$$

undefiniert. Wie einige einfache Überlegungen zeigen, ist das der Fall (siehe Abb. 6.3), wenn die beiden Orientierungen, bei denen die beiden Symmetrieachsen auf den Hauptachsen des Koordinatensystems der Positionskodierung liegen, mit $(z_1 = |\mathbf{z}|, z_2 = 0)$ bzw. $(z_1 = -|\mathbf{z}|, z_2 = 0)$ kodiert werden. Von einer solchen Kodierung soll im Folgenden ausgegangen werden.

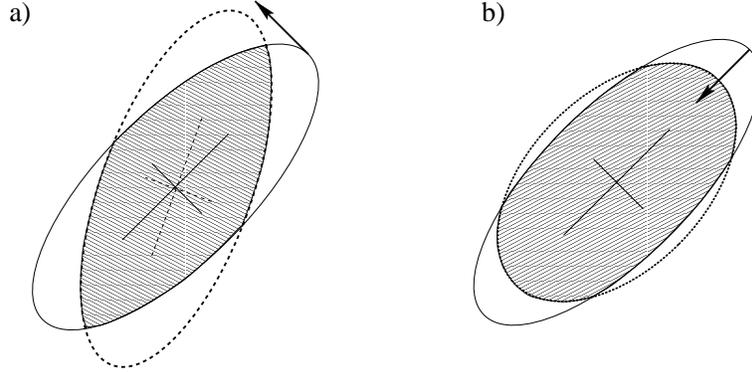


Abbildung 6.4: Die Kopplung zwischen \mathbf{z} und $d\mathbf{z}$ ist anschaulich der metrische Unterschied zwischen einer Drehung und einer Stauchung. Die Pfeile symbolisieren jeweils die Richtung des infinitesimalen Orientierungsunterschiedes $d\mathbf{z}$ in ihrer π -periodischen Darstellung. Im Fall a) ist $d\mathbf{z} \cdot \mathbf{z} = 0$ im Fall b) $-|\mathbf{z}||d\mathbf{z}|$. Dieser Unterschied korrespondiert mit den im allgemeinen Fall unterschiedlichen Überlappungen in der hochdimensionalen Darstellung.

Schließlich ist noch eine Kopplung zwischen \mathbf{z} und $d\mathbf{z}$ möglich. Anschaulich handelt es sich hierbei um die Darstellung von Unterschieden zwischen einer Drehung und einer Stauchung bei der Projektion in den Feature Raum (siehe Abb. 6.4). Diese sollten wegen der Rotationsinvarianz eine Funktion des Skalarprodukts

$$c_2 = f(\mathbf{z} \cdot d\mathbf{z}) \quad (6.41)$$

sein. Aufgrund analoger Argumente im Fall der Kopplung zwischen Orientierung und Position erhält man schließlich

$$c_2 = \zeta(|\mathbf{z}|)(dz_1^2 z_1^2 + dz_2^2 z_2^2 + dz_1 dz_2 z_1 z_2). \quad (6.42)$$

Insgesamt hat dann die Metrik die Form

$$d = \tilde{d} + c_1 + c_2 = d\mathbf{u}(g) d\mathbf{u}^T \quad (6.43)$$

mit

$$(g) = \begin{pmatrix} g_{\xi\xi} & g_{\xi\mathbf{z}} \\ g_{\mathbf{z}\xi} & g_{\mathbf{z}\mathbf{z}} \end{pmatrix} = \begin{pmatrix} \nu + \mu z_1 & \mu z_2 & 0 & 0 \\ \mu z_2 & \nu - \mu z_1 & 0 & 0 \\ 0 & 0 & \delta + \epsilon z_1 z_1 & \epsilon z_1 z_2 \\ 0 & 0 & \epsilon z_1 z_2 & \delta + \epsilon z_2 z_2 \end{pmatrix} \quad (6.44)$$

als allgemeinste Darstellung unter den gegebenen Symmetrien. Auch über die Eigenschaften der vier freien Parameter lassen sich Aussagen machen. Die Funktionen $\nu, \delta, \mu, \epsilon$ dürfen nur von $|\mathbf{z}|$ abhängen. Eine Abhängigkeit von ξ ist wegen der Translationsinvarianz nicht erlaubt. Abhängigkeiten von $d\xi$ und $d\mathbf{z}$ sind aufgrund des infinitesimalen Kalküls des Tensors

verboten. Schließlich verletzen Abhängigkeiten von \mathbf{z} die Rotationsinvarianz des Gesamtsystems. Da die Metrik positiv definit sein muss, ist es notwendig, dass ihre Eigenwerte positiv sind. Da die Matrix ihrer Konstruktion nach eine Metrik definiert, die gegenüber Orientierungsvariationen der oben beschriebene Art invariant ist, genügt es, den Spezialfall $z_1 = |\mathbf{z}|, z_2 = 0$ zu betrachten. Hier ist (g) schon diagonal und es ergeben sich

$$\nu - \mu|\mathbf{z}| > 0 \quad (6.45)$$

$$\nu + \mu|\mathbf{z}| > 0 \quad (6.46)$$

$$\delta + \epsilon|\mathbf{z}|^2 > 0 \quad (6.47)$$

$$\delta > 0 \quad (6.48)$$

als Nebenbedingungen für den metrischen Tensor.

Im Anhang C wird für einen Spezialfall, nämlich elongierten Gaußverteilungen der konkrete metrische Tensor berechnet. Für die später folgende Simulation wurde der Einfachheit halber, ein abstraktes Featureset verwendet, welches die die Symmetrien (1)-(3) aufweist und außerdem über einen nichtverschwindenden Kopplungsterm verfügt.

Wie oben schon dargelegt, enthält die obige Darstellung noch einen Freiheitsgrad, die Definition von $|\mathbf{z}|$. Das kann dazu führen, dass äquivalente Systeme unterschiedlich erscheinen. Zudem kann eine sinnvolle Definition von $|\mathbf{z}|$ eine der vier freien Funktionen eliminieren helfen. In Anhang B wird gezeigt, dass die Transformation

$$\tilde{z}_1 \leftarrow z_1 f(|\mathbf{z}|^2) \quad (6.49)$$

$$\tilde{z}_2 \leftarrow z_2 f(|\mathbf{z}|^2) \quad (6.50)$$

mit

$$f = C \exp \left(\int_0^{|\mathbf{z}|} -\frac{1}{2|\mathbf{z}|^2} + \frac{1}{2} \sqrt{\frac{1}{|\mathbf{z}|^4} + \frac{\epsilon}{\delta} \frac{1}{|\mathbf{z}|^2}} d|\mathbf{z}| \right) \quad (6.51)$$

zu einer Eliminierung von $\epsilon(|\mathbf{z}|)$ führt, so dass sich alle relevanten Metriken auf die Äquivalenzklassen

$$(\tilde{g}) = \begin{pmatrix} \tilde{\nu} + \tilde{\mu}\tilde{z}_1 & \tilde{\mu}\tilde{z}_2 & & \\ \tilde{\mu}\tilde{z}_2 & \tilde{\nu} - \tilde{\mu}\tilde{z}_1 & & \\ & & \tilde{\delta} & \\ & & & \tilde{\delta} \end{pmatrix} \quad (6.52)$$

reduzieren lassen. In diesem Anhang wird außerdem gezeigt, warum der Versuch, statt $\epsilon(|\mathbf{z}|)$ eine der andern drei Funktionen (ν, μ, δ) durch eine Transformation von $|\mathbf{z}|$ zu eliminieren, scheitern muss.

6.6 Diskussion

Solange von hochdimensionalen Modellen die Rede war, schien es, dass sich die Analyse der verschiedenen Phänomene auf eine Beschreibung von Ergebnissen der numerischen Berechnung zu beschränken hatten.

Die Situation hat sich nun geändert. Statt mit einem Stimulusraum von mehreren hundert Dimensionen hat man es nun mit einem vierdimensionalen zu tun. Erstens kann man jetzt die Auswirkungen bestimmter Eigenschaften der hochdimensionalen Modelle direkt im Eigenschaftsraum testen. Das verringert den numerischen Aufwand der Simulation. Zweitens sind nun formale, analytische Untersuchungen möglich. Drittens kann man das Resultat dieses Kapitels mit dem Katalog bestehender Eigenschaftsmodelle, die auf ähnlichen Prinzipien, beruhen vergleichen und diese bestehenden Modelle in den nun größeren Raum von metrischen Möglichkeiten einordnen.

6.7 Zusammenfassung

In diesem Kapitel wird eine Klasse von Metriken im Eigenschaftsraum vorgestellt, die als die nach den Einschränkungen durch verschiedene Symmetrieüberlegungen allgemeinste anzusehen ist. Sie verfügt über drei verschiedene freie Parameterfunktionen.

Kapitel 7

Ein einfaches Modell mit nichteuklidischem Featureraum

Nun soll an einigen Repräsentanten der im vorhergehenden Kapitel diskutierten Metriken untersucht werden, inwiefern unter Berücksichtigung realistischer Symmetrieeigenschaften im stationären Zustand Strukturänderungen der Merkmalskarte auftreten können. Besonderes Augenmerk soll dabei auf die Orientierungskarte und ihre strukturellen Eigenschaften gerichtet werden, weil hierzu eine Vielzahl von experimentellen, modellbasierten und theoretischen Untersuchungen vorliegt.

7.1 Eigenschaften des α -Modells

Es geht in diesem Abschnitt darum, an einem Beispiel die Konsequenzen der metrischen Kopplungen zwischen der Richtung der Positionsverschiebung und Orientierungskarte zu untersuchen. Als ein einfacher denkbarer Ansatz soll als Abstandsmaß zwischen den zwei Punkten $\mathbf{u}^a = (\xi_1^a, \xi_2^a, z_1^a, z_2^a)$ und $\mathbf{u}^b = (\xi_1^b, \xi_2^b, z_1^b, z_2^b)$ des Stimulusraumes

$$d^2(\mathbf{u}^a, \mathbf{u}^b) = (|\mathbf{u}^a - \mathbf{u}^b|)^2 + \alpha C_1, \quad (7.1)$$

gewählt werden, wobei

$$C_1 = \frac{z_1^a + z_1^b}{2} \left((\xi_1^a - \xi_1^b)^2 - (\xi_2^a - \xi_2^b)^2 \right) + 2 \frac{z_2^a + z_2^b}{2} (\xi_2^a - \xi_2^b) (\xi_1^a - \xi_1^b) \quad (7.2)$$

die metrische Kopplung zwischen Orientierungs und Positionskarte definiert. Im Grenzfall konvergierender Punkte \mathbf{u}^a und \mathbf{u}^b geht die Abstandsfunktion in

$$d^2 = d\mathbf{u} (g) d\mathbf{u} \quad (7.3)$$

mit einem metrischen Tensor

$$(g) = \begin{pmatrix} 1 + \alpha z_1 & \alpha z_2 & & \\ \alpha z_2 & 1 - \alpha z_1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \quad (7.4)$$

über. Im Kontext der Überlegungen des letzten Kapitels ergibt sich $\nu = \delta = 1$, $\mu = \alpha$, $\epsilon = 0$. Diese Metrik ist durch die Nebenbedingungen nach den Gln. (6.45)- (6.48) bei der Orientierungsstärken $\max(|\mathbf{z}|) < 1/\alpha$ positiv definit.

Die Abweichung vom Euklidischen Fall geht über den Parameter α ein, der es erlaubt, zwischen einer stark gekrümmten und der euklidischen Metrik zu interpolieren.

Der Lernalgorithmus soll das im vorigen Kapitel modifizierte elastische Netzmodell (vgl. Gl. (6.10)) sein. In jeder Iteration wird dem Modell ein Stimulus $\mathbf{v} = (\mathbf{z}, \boldsymbol{\xi})$ präsentiert. Die Orientierungskomponente $\mathbf{z} = (z_1, z_2)$ des Stimulus ist dabei auf einer Kreisscheibe gewählt, wobei die Repräsentation von $|\mathbf{z}|$ gleichmäßig auf dem Intervall $[0, 2]$ liegt. Die Schwerpunktkoordinaten im visuellen Feld $\boldsymbol{\xi} = (\xi_1, \xi_2)$ werden zufällig aus einem Intervallquadrat mit der Kantenlänge S_x ermittelt.

Für die Positionskoordinaten $\boldsymbol{\xi}$ im visuellen Feld und für die kortikalen Koordinaten \mathbf{r} werden periodische Randbedingungen angenommen.

Initialisiert wird die Karte nicht orientierungsspezifisch und retinotop, wobei auf der Orientierungskarte an jedem Punkt eine Zufallsorientierung mit einer Amplitude der Orientierungsselektivität von 0.01 appliziert wird. Die Adaption erfolgt iterativ durch die durch den Gradienten der Kostenfunktion gegebene Lernregel

$$\frac{\partial \mathbf{w}_{\mathbf{r}}}{\partial t} = -\gamma \frac{\partial E}{\partial \mathbf{w}_{\mathbf{r}}}. \quad (7.5)$$

In den Untersuchungen wird ein Gitter von $N_x \times N_x = 64 \times 64$ Neuronen verwendet, was einer Größe eines visuellen Feldes von $S_x \times S_x = 10 \times 10$ Einheiten gleichkommt. Es werden sowohl bei den Positionskoordinaten des visuellen Feldes ($\boldsymbol{\xi}$), als auch bei denen des Kortex (\mathbf{r}) periodische Randbedingungen verwendet, wobei die Periodizität im Stimulusraum S_x ist. Die Gesamtheit der Neuronen \mathcal{N} , sieht man sie als Kontinuum, formt damit einen Torus.

Das Minimum der Kostenfunktion wird durch den Gradientenabstieg ermittelt, wobei statt des Erwartungswertes über den Stimulusraum in jeder Iteration ein zufällig ermittelter Vertreter verwendet wird. Die Lernrate γ in jeder Iteration ist konstant 0.01

7.2 Der Fall $\alpha = 0$

Für $\alpha = 0$ erhält man die euklidische Metrik $d(\mathbf{u}^a, \mathbf{u}^b) = \|\mathbf{u}^a - \mathbf{u}^b\|^2$, die sowohl für die selbstorganisierende Karte als auch im Fall des elastischen Netzes schon früher eingehend analytisch untersucht worden ist [Wolf, 1998]. Insbesondere ist ein Phasenübergang von einer Phase, bei der alle Neuronen orientierungsunspezifisch sind, zu einer orientierungspezifischen Phase für $\sigma_{EL} > \sigma_{EL}^*$ beobachtet und analytisch untersucht worden, wobei

$$\sigma_{EL}^* = \sqrt{\frac{\text{var}[\mathbf{z}]}{2} - \eta^* + \eta^* \log\left(\frac{2\eta^*}{\text{var}[\mathbf{z}]}\right)} \quad (7.6)$$

Die Wellenzahl der dann entstehenden Orientierungskarte ist

$$|\mathbf{k}| = \frac{1}{\sigma_{EL}} \sqrt{\log\left(\frac{\text{var}[\mathbf{z}(\mathbf{r})]}{2\eta^*}\right)} \quad (7.7)$$

wobei $\text{var}[\mathbf{z}] = \mathbf{E}[\mathbf{z}^2]_s$ die Varianz der Orientierungskordinaten des Stimulus darstellt und $\eta = \frac{S_x^2}{N_x} \eta^*$. Für das elastische Netz wird in [Wolf, 1998] analytisch gezeigt worden, dass in der Nähe des oben beschriebenen Phasenübergangs keine stabilen Orientierungssingularitäten existieren. Numerisch ist in dieser Arbeit gefunden worden, dass auch im selbstorganisierenden Kartenalgorithmus für dieses Abstandsmaß keinen stabilen Orientierungssingularitäten erhalten bleiben.

7.3 Numerische Resultate zum α -Modell

In dem hier verwendeten Modell wird die Entwicklung – insbesondere die der Orientierungssingularitäten – auf großen Zeitskalen untersucht. Es zeigt sich, dass auch über sehr lange Zeiträume hinweg in Abhängigkeit der Kopplungsstärke *pinwheels* erhalten bleiben. Die hier diskutierten Karten werden über einen Zeitraum von bis zu 240×2^{20} Iterationen untersucht.

Bei Simulationen, die sich weiter vom Phasenübergang entfernt befinden, werden zwar ebenfalls über lange Zeitskalen Orientierungssingularitäten beobachtet, allerdings entwickeln sich keine regelmäßigen Muster, wie sie in Abb. 7.2 beobachtet werden.

Es stellt sich heraus, dass sich für verschiedene Kopplungsstärken α der für den Fall $\alpha = 0$ ermittelte Phasenübergang etwas verschiebt (siehe Abb. 7.1). Der wesentliche Grund dafür ist wohl die Veränderung in der Varianz des Stimulussets durch den hinzukommenden Stimulusparameter.

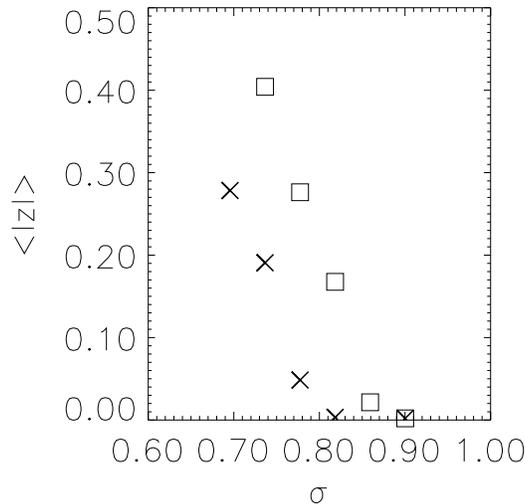


Abbildung 7.1: Die mittleren Orientierungsstärken der Karten nach 32×2^{20} Iterationen gegen die Stärke von σ_{EL} im Fall $\alpha = 0$ (Kreuze) und $\alpha = 0.5$ (Quadrate). Die Lernrate γ wird bei 0.01, sowie η bei 0.025 konstant gehalten. Stimuliert wird das Modell mit Reizen deren Orientierungspräferenz gleichmäßig zwischen 0 und 2.0 verteilt ist.

Weiterhin ergeben sich in der Nähe des Phasenübergangs auf langen Zeitskalen für alle untersuchten α -Werte in einem Bereich von 0.0 bis 0.5 regelmäßige Muster, die bei einem α unter 0.3 *pinwheel*-frei sind und bei Werten darüber *pinwheels* enthalten (siehe Abb. 7.2). Die stabilen *pinwheel*-Dichten entsprechen in etwa den *pinwheel*-Dichten, wie sie im visuellen Kortex gefunden werden.

Es ist anzumerken, dass der Kohonenalgorithmus (vgl. Abschn. 6.4) sich in der α -Dynamik etwas vom elastischen Netz unterscheidet. Während für kleine α beide Algorithmen in ähnlicher Weise Streifenmuster produzieren, kommen geordnete *pinwheel*-Gitter bei größerem α nur transient vor. D.h. die ungeordnete Pinwheelkonfiguration durchläuft eine Phase, die ähnlich der des elastischen Netz ist, um dann wieder in ein streifenähnlicher Muster zu fallen. Dieser Effekt kann in Analogie zu den analytischen Untersuchungen in [Der & Herrmann, 1994] damit erklärt werden, dass der Zustand des Netzes durch stimulusbedingte Fluktuationen aus dem Minimum eines effektiven Ginzburg-Landau-Potenzials getrieben und erst durch Terme 3. Ordnung weit weg von der instabilen Fixpunktlage aufgefangen wird. Aufgrund der Symmetrien des Systems resultieren dominierende Terme 3. Ordnung in Streifenmustern [Wolf, 1998].

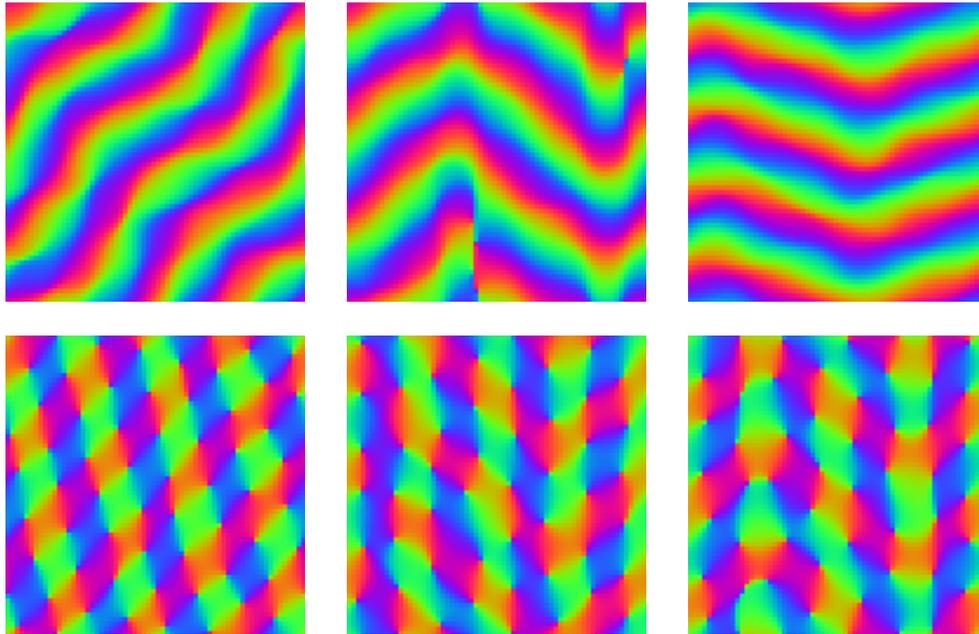


Abbildung 7.2: Die finalen Karten nach 120×2^{20} Iterationen. Die Parametereinstellungen von links nach rechts: erste Zeile: $\alpha = 0.0, \sigma_{EL} = 0.78$; $\alpha = 0.1, \sigma_{EL} = 0.75$; $\alpha = 0.2, \sigma_{EL} = 0.77$; zweite Zeile: $\alpha = 0.3, \sigma_{EL} = 0.78$; $\alpha = 0.4, \sigma_{EL} = 0.80$; $\alpha = 0.5, \sigma_{EL} = 0.82$. Die leichte Verschiebung von σ_{EL} ist in der im Text erwähnten Verschiebung des Phasenübergangs in Anhängigkeit der Kopplungsstärke begründet. Für die anderen Modellparameter siehe Abb. 7.1.

7.4 Der Prozess der Kartenbildung

Die Entwicklung der Karten lässt sich in zwei Phasen unterteilen: das anfängliche Entstehen von Orientierungspräferenzen und einen darauf folgenden globalen Strukturbildungsprozess. Während das anfängliche Ansteigen der Orientierungspräferenzen unter den hier für das Modell benannten Bedingungen meist deutlich unter 2^{20} Iterationen stattfindet, ist die Dauer der globalen Strukturbildung durchaus nicht sehr weit (etwa $1/8 - 1/1$) unter der Größenordnung der Dauer der Simulation von 120×2^{20} Iterationen. Diese Werte stimmen in etwa mit denen in [Wolf, 1998] angegebenen Werten für den dort untersuchten Spezialfall von $\alpha = 0$ überein.

Es wird versucht heuristische Größen zu finden, die den Fortschritt während des Strukturbildungsprozesses quantitativ widerspiegeln sollen. Dabei erweist sich die Gaußsche Krümmung der Karte als hilfreich.

Da es sich bei den Simulationen um Karten mit periodischen Randbedingungen handelt, formt die Karte eine periodisch Untermanigfaltigkeit im Stimulusraum, genauer handelt es

sich um einen Torus. Nach dem globalen Gauß - Bonnet - Theorem gilt

$$\int_{\mathcal{N}} \kappa_{\mathbf{r}} d\mathbf{r} = 2\pi\chi. \quad (7.8)$$

In unserem Fall gilt also, da die Euler - Poincaré Charakteristik χ eines Torus gleich 0 ist

$$E[\kappa_{\mathbf{r}}]_{\mathcal{N}} = 0. \quad (7.9)$$

Hier wird die Varianz der Gaußschen Krümmung $\text{var}(\kappa_{\mathbf{r}})$ untersucht. Es zeigt sich (siehe Abb. 7.3), dass diese sowohl im Falle einer schwachen Kopplung, also bei der Konvergenz der Karte gegen den *pinwheel*-freien Fall, als auch im Falle einer starken Kopplung kontinuierlich abnahm. Dieses Phänomen wird bei allen Simulationen des α - Modells beobachtet.

7.5 Vergleich mit Featurekarten hochdimensionaler Modelle

In diesem Abschnitt sollen nun die Orientierungskarten, wie sie nach langen Zeitskalen aus dem α - Modell entstehen, mit originär hochdimensionalen Ansätzen verglichen werden.

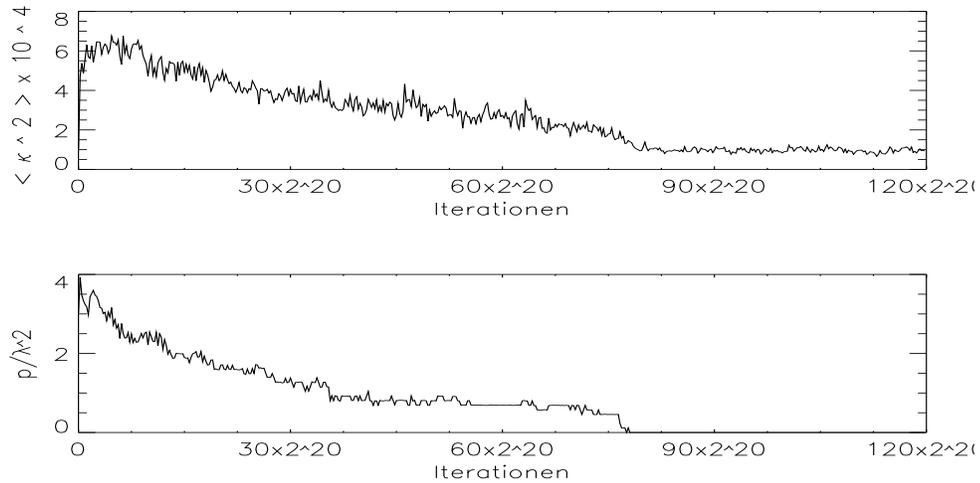
Statt der ASSOM, die numerisch aufwendig ist, von der auch nicht bekannt ist, wo sich der Phaenübergang zwischen der unselektiven und der orientierungsselektiven Phase befindet, soll ein anderes, einfacheres Modell verwendet werden. Es wird eine hochdimensionale Karte verwendet, deren Stimulierung durch zweidimensionale elongierte Gaussverteilungen

$$G_{\text{gs}}(\xi_1, \xi_2) = \exp \left(- \left| \begin{pmatrix} 1/2/\sigma_{2,gs}^2 \\ 1/2/\sigma_{1,gs}^2 \end{pmatrix} \cdot \hat{u}(\Phi_{\text{gs}}) \cdot \begin{pmatrix} (\xi_1 - \bar{\xi}_1) \\ (\xi_2 - \bar{\xi}_2) \end{pmatrix} \right|^2 \right) \quad (7.10)$$

erfolgt. In solchen Karten ist der Phasenübergang zwischen der Orientierungsunspezifischen und der orientierungsselektiven Phase schon früher erforscht worden [Riesenhuber et al., 1996]. Es werden Simulationen von einem Gitter mit $N_x \times N_x = 64 \times 64$ Neuronen durchgeführt. Die Größe des Eingangsraumes Ξ beträgt 24×24 Pixel. Die Stimulation erfolgte mit elongierten Gaussverteilungen wobei diese konstante Halbmesser $\sigma_{1,gs}$ und $\sigma_{2,gs}$ aufweisen. Die Weite der Nachbarschaftsfunktion σ_K wird so gewählt, dass es sich in der Nähe des von [Riesenhuber et al., 1996] ermittelten Phasenübergangs befinden.

Es zeigt sich, dass die sich ausbildende Struktur (siehe Abb. 7.5) in beiden Fällen ähnlich ist.

a)



b)

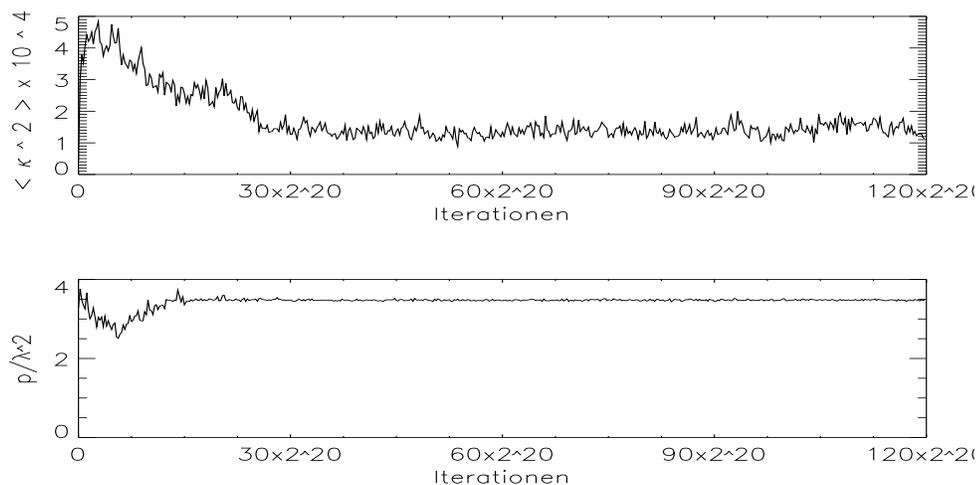


Abbildung 7.3: Beispiele für zwei Entwicklungsprozesse. a) Eine Simulation mit $\alpha = 0.2$ und $\sigma_{EL} = 0.77$, restliche Parameter siehe Abb. 7.2), mit dem *pinwheel*-freie Karten erzielt werden und b) $\alpha = 0.3$ und $\sigma_{EL} = 0.79$, der in einer *pinwheel*-reichen, kristallinen, Karte endet. Gezeigt wird die Anzahl der *pinwheels* pro Wellenlängenquadrat der Orientierungskolumnen (jeweils oben) und die Varianz der Gaußschen Krümmung $\text{var}(\kappa_{\mathbf{r}})$. Obwohl die beiden Karten in unterschiedliche Strukturen konvergieren nehmen in beiden Fällen die ermittelten Werte für $\text{var}(\kappa_{\mathbf{r}})$ kontinuierlich ab.

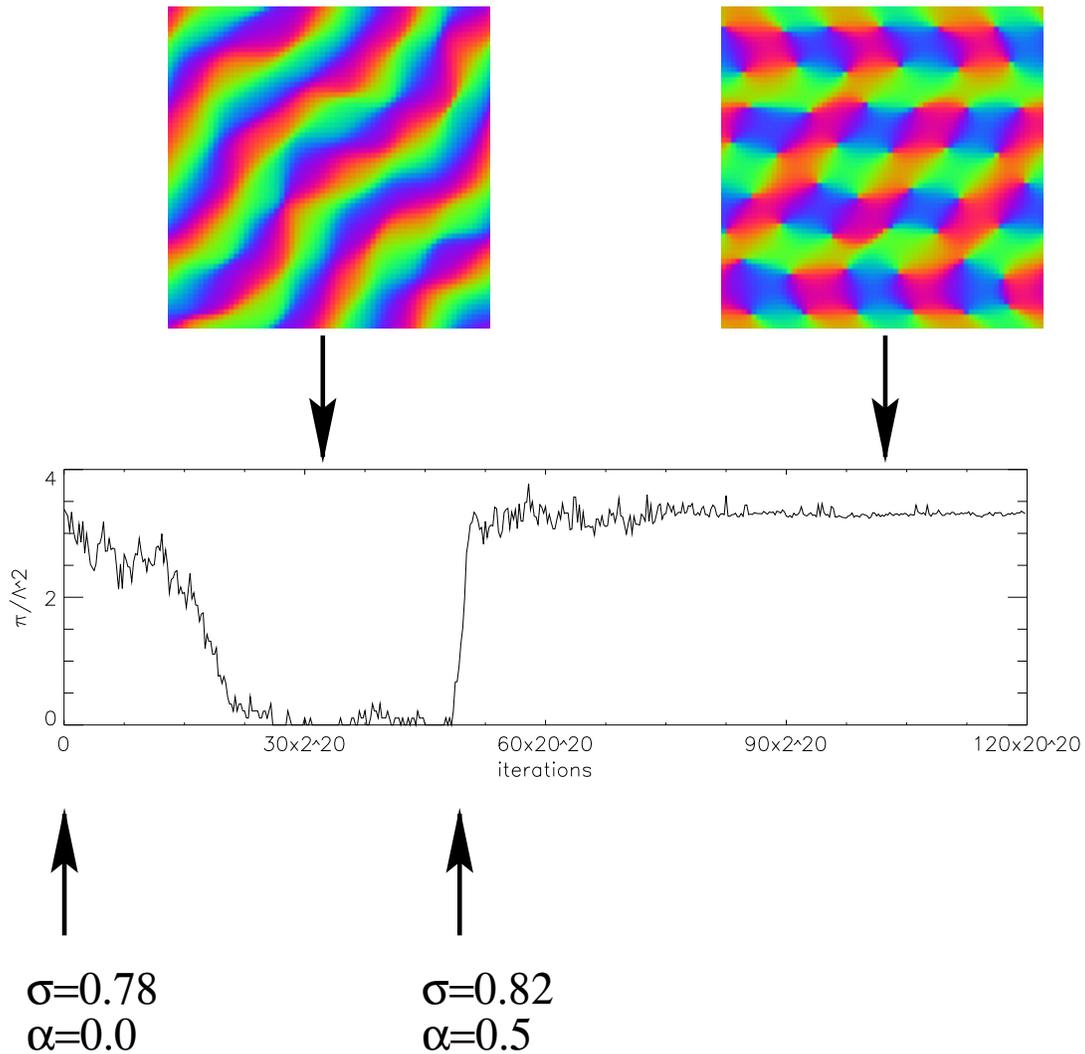


Abbildung 7.4: Die Simulationen zeigen sich stabil gegen die Initialisierung. Die Karte wird retinotop mit schwachen, stochastischen Orientierungspräferenzen initialisiert. Die Karte wird einem Lernprozess von $\alpha = 0.0$ und $\sigma_{EL} = 0.86$ 48×2^20 Iterationen ausgesetzt. In diesem Zeitraum verschwanden praktisch alle *pinwheels*. Nach dieser ersten Phase wird die Simulation unter geänderten Konditionen fortgeführt. Der Parameter der Metrik α beträgt 0.5 bei $\sigma_{EL} = 0.96$

7.6 Diskussion

Für das elastische Netz wird in [Wolf, 1998] analytisch gezeigt, dass in der Nähe des oben beschriebenen Phasenübergangs für den Fall der hier als der Fall $\alpha = 0$ bezeichnet wird keine stabilen Orientierungssingularitäten existieren. Numerisch wird in dieser Arbeit auch im selbstorganisierenden Kartenalgorithmus für dieses Abstandsmaß keinen stabilen Orientierungssingularitäten gefunden. Aus diesen Befunden wird dort geschlossen, dass die große

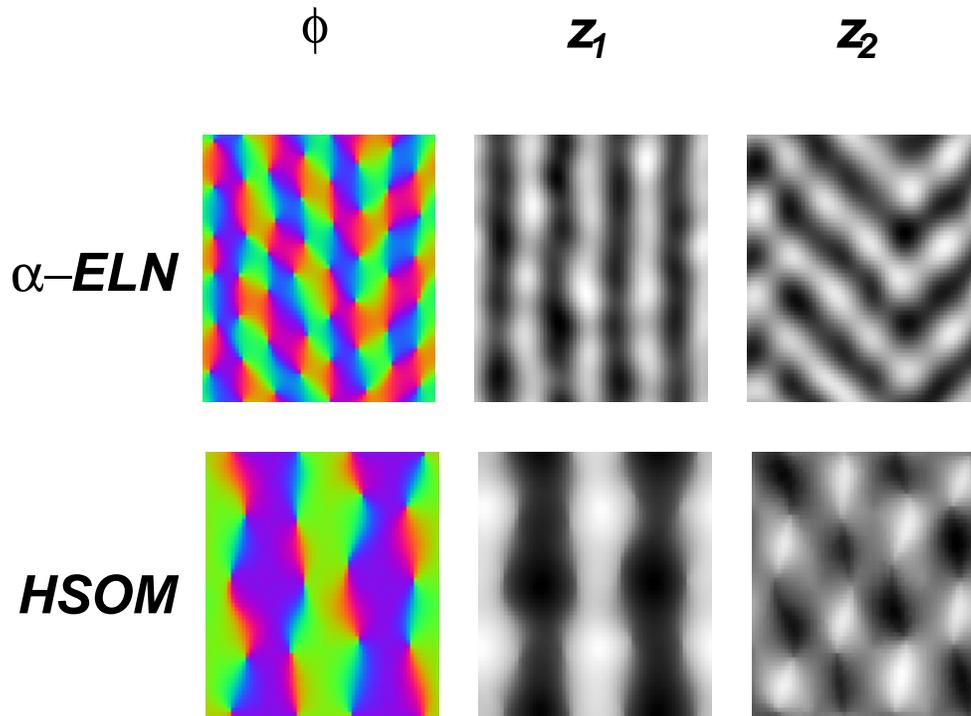


Abbildung 7.5: Ein Vergleich zwischen Karten wie nach langen Zeitskalen aus hochdimensionalen selbstorganisierenden Karten entstehen können mit den finalen Karten des α -Modells. Der Repräsentant des α -Modells hat eine Kopplung α von 0.4 und ein σ_{EL} von 0.80. Die damit verglichene Repräsentantin einer hochdimensionalen selbstorganisierenden Karte wird mit elongierten Gaußverteilungen stimuliert. Deren größerer Halbmesser $\sigma_{1,gs}$ ist konstant bei 7.46 und der kleinere $\sigma_{2,gs}$ bei 2.15 lag. Die Weite der Nachbarschaftsfunktion σ_K ist in diesem Lauf auf 5.0 konstant gehalten worden; die Lernrate γ beträgt 0.05. Es sind 2^{24} Iterationen ausgeführt worden.

Zahl von *pinwheels* im adulten Kortex nicht durch afferente Gegebenheiten erklärt werden kann, sondern ein Phänomen sind, das durch orientierungsspezifische, langreichweitige, laterale Wechselwirkungen im visuellen Kortex hervorgerufen wird. In diesem Kapitel konnte gezeigt werden, dass das Auftreten von Orientierungssingularitäten unter Umständen eine Konsequenz bestimmter Eigenarten des Stimulusraumes sein können, und dass diese Möglichkeit bei zukünftigen Überlegungen in Betracht gezogen werden muss.

7.7 Zusammenfassung

In diesem Kapitel werden einige Vertreter der im letzten Kapitel ermittelten möglichen Metriken im Hinblick auf die beiden Modellsysteme, selbstorganisierende Karte und elasti-

sches Netz numerisch untersucht. Es wird ein Parameter eingeführt, der in den Metriken die Stärke der Kopplung der Positionskarte mit der Orientierungskarte repräsentiert. Die Simulationen zeigen, dass die Strukturbildung durchaus von der Diversität dieser Metriken affiziert wird. Insbesondere zeigen sich diese Einflüsse in der Zahl der über lange Zeiträume stabilen Orientierungssingularitäten.

Kapitel 8

Zusammenfassung und Diskussion

Die Erfahrungsabhängigkeit der Entwicklung des visuellen Kortex und damit in Verbindung stehende strukturelle Selbstorganisationsprozesse werden in dieser Arbeit mit verschiedenen methodischen Ansätzen modelliert.

Die Arbeit teilt sich in drei Themenbereiche. Der erste Teil befasst sich mit einem nichttopographischen Modell, welches die Antworteigenschaften kortikaler rezeptiver Felder simuliert. Im zweiten Teil wird diesem Modell eine topographische Komponente hinzugefügt. Da allerdings die biologischen Befunde von kortikalen Merkmalskarten so nicht reproduziert werden können, wird dieses Modell zu einer Karte mit adaptiven Unterräumen erweitert. Der dritten Teil dieser Arbeit beschreibt die metrischen Eigenschaften hochdimensionaler Modelle mit bestimmten Symmetrien in einem vierdimensionalen Eigenschaftsraum. Es kann eine Klasse von relevanten Metriken extrahiert werden, die in der Lage ist, Eigenschaften hochdimensionaler Inputmanigfaltigkeiten zu beschreiben. Eine einfache Variante aus dieser Klasse wird verwendet, um die Kartenbildung in diesem Fall zu untersuchen. Dabei interessiert, inwiefern die notwendige Verallgemeinerung der Euklidischen Metrik auf den nichteuklidischen Fall die Strukturbildung in diesen Modellen beeinflusst. Es soll nun detaillierter auf die einzelnen Teile eingegangen werden.

Im ersten Teil wird gezeigt, wie die kortikale Entwicklung durch Veränderungen bei der retinalen Verarbeitung beeinflusst wird. Während einer kritischen Phase zeigt sich, dass die normale Erfahrungswelt für eine normale Entwicklung des visuellen Kortex benötigt wird. Als Steuerungsparameter dieses Entwicklungsvorgangs wird die retinale Filterwellenlänge identifiziert, die sich während der hier relevanten Entwicklungsphasen erheblich verändert. Die rezeptiven Felder des Modells entsprechen in hohem Maße denen im biologischen Kortex in den betrachteten Entwicklungsstadien.

Insgesamt erweist sich dieses Modell als effektiv in der Erklärung verschiedener, bis dato unverständener Phänomene, die mit Experimenten unter visueller Deprivation in Zu-

sammenhang stehen, wobei sich die Reduzierung der retinalen Filterwellenlänge, die in bisherigen Modellen unberücksichtigt geblieben ist, als essentiell erweist.

Am Anfang der Entwicklung weist der retinale Filter eine hohe Wellenlänge auf. In dieser Phase verläuft die kortikale Entwicklung laut biologischen Befunden eher unabhängig von der visuellen Erfahrung. Dies kann im Modell reproduziert werden. Unterhalb eines kritischen Wertes dieser Wellenlänge ist die Art der Erfahrung essentiell für die kortikale Strukturbildung, was hier anhand der Entwicklung rezeptiver Felder gezeigt werden konnte.

Als Maß für das Entwicklungsdefizit wurde experimentell die Stärke der Orientierungsselektivität kortikaler Neuronen verwendet. Das dazu analog am Modell abgeleitete Maß der Orientierungsstärke zeigt einen synchronen Verlauf unter den verschiedenen Stimulusbedingungen. Damit ist die Annahme bestätigt, dass die retinale Filterwellenlänge in den frühen postnatalen Entwicklungsstufen die Funktion eines Kontrollparameters in der visuellen Entwicklung einnimmt. Der Sinn dieser Veränderung könnte darin liegen, dass das visuelle System im Laufe dieses Prozesses an das Umfeld angepasst wird. Als solches Umfeld im weiteren Sinne kann auch die Physiognomie des Tieres angesehen werden, auf die sich das visuelle System erst einstellen muss.

Im zweiten Teil werden natürliche Bilder als Inputs für selbstorganisierenden Karten benutzt. Ein Teil der Eigenschaften dieses Modells deckt sich durchaus mit experimentellen Befunden. Es stellt sich jedoch heraus, dass die zunächst verwendeten selbstorganisierenden Karten in diesem Fall nicht mit den biologisch gefundenen Resultaten in Einklang zu bringen sind. So kann die retinotopie Anordnung der rezeptiven Felder nicht stabil bleiben. Als Grund dafür wird ein einfacher metrischer Mismatch zwischen zwei Stimuluseigenschaften – nämlich der lokalen Phase und der retinotopen Position – festgestellt. Alternativ wird ein Modell verwendet, das lokal lineare adaptive Unterräume (lokale PCA) zur Stimulusrepräsentation verwendet. In diesem Modell bleibt die grobe Retinotopie erhalten. Die einzelnen Einheiten zeigen ohne, dass dies explizit in die Modellannahme einging, ein Antwortverhalten, das unabhängig von der spatialen Phase des Stimulus ist. Dies entspricht dem Antwortverhalten von *complex cells* im visuellen Kortex. Auf einer kleinen Größenskala zeigen sich Abweichungen von der retinotopen Karte. Die Wellenlänge dieser Abweichungen entspricht in etwa der Wellenlänge der Orientierungskarte. Diese Tatsache kann als Hinweis darauf interpretiert werden, dass zwischen der Positions- und der Orientierungskarte eine Kopplung besteht.

Im dritten Teil der Arbeit wird eine Metrik in einem vierdimensionalen Eigenschaftsraum eingeführt, die einer bestimmten Eigenschaft einer Klasse von hochdimensionalen Karten Rechnung trägt, nämlich, daß – wie oben gesagt – Positionskarte und Orientierungskarte metrisch aneinander gekoppelt sind. Unter Berücksichtigung von Symmetrien, die in den meisten hochdimensionalen Modellen exakt und im visuellen Kortex näherungsweise erfüllt sind, wird eine Metrik in der Form eines metrischen Tensors als allgemeinste mögliche Darstellung dieses Abstandes eingeführt. Durch Normierung eines Freiheitsgrades werden die vier ursprünglichen Parameter auf nur noch drei skalare freie Parameterfunktionen

reduziert

Daraufhin wird ein einfacher Repräsentant solcher Metriken vorgestellt, welcher mit Hilfe eines Parameters die Stärke der Kopplung einzustellen vermag. Es zeigt sich nun, daß sich Unterschiede in der topologischen Struktur der Karten ergeben. Insbesondere ändert sich die Dichte der schließlich verbleibenden Ortientierungssingularitäten (*pinwheels*). Im Falle einer starken Kopplung zwischen Positions und Orientierungskarte erhält man eine finale Pinweeldichte, die nahe bei der im biologischen Kortex gefundenen liegt.

Es wurde bei höheren Säugetieren mehrfach die Frage diskutiert, inwieweit der visuelle Kortex in seiner Struktur, sowohl der Ausformung seiner rezeptiven Felder, als auch in der Musterbildung, ein Produkt der Erfahrung ist, inwieweit das Ergebnis einer genetisch bedingten Vorlage. In dieser Arbeit werden unter der Prämisse, daß die kortikale Entwicklung stark erfahrungsabhängig ist, Phänomene reproduziert, wie sie tatsächlich im visuellen Kortex gefunden werden.

Insbesondere trifft dies auf die Untersuchungen zur Entwicklung rezeptiver Felder zu. Hier können tatsächlich verschiedene Details, wie der initiale Anstieg der Orientierungsselektivität, wie auch der spätere Abfall im Falle binokulärer Deprivation rekonstruiert werden. Dies kann in der Konsequenz so interpretiert werden, dass die Vorgänge, die die Entwicklung der rezeptiven Felder betreffen, tatsächlich stark erfahrungsabhängig sind und daß die zugrundeliegenden Prozesse im Rahmen des Modells gut verstanden sind.

Bei dem selbstorganisierenden Kartenmodell adaptiver Unterräume kommt ein anderer Aspekt hinzu. Hier wird der Stimulusraum nicht nur repräsentiert, sondern auch von einer Eigenschaft des Stimulusraumes, der lokalen Phase, abstrahiert. Der visuelle Kortex sieht von einer Eigenschaft ab und ermittelt stattdessen die gegenüber dieser Eigenschaft invarianten Größen Orientierung und Position im visuellen Feld. Auf einer abstrakteren Ebene ist ein solches Verhalten in zunehmenden Maß in höheren Arealen zu beobachten, in denen beispielsweise Gesichter unabhängig von der Perspektive erkannt werden, und diese Neuronen genau dann feuern, wenn ein Gesicht im visuellen Feld erscheint [Bruce et al., 1980]. Oder man denke an intermediäre visuelle Areale in denen Neurone exakt dann feuern, wenn ein Schraubenzieher im visuellen Feld erscheint [Kobatake & Tanaka, 1994]. Obwohl hier in der Entwicklung andere Abstraktionskriterien verwendet werden, ist das Prinzip dort ähnlich. Statt der Verwendung linearer Superponierbarkeit, können in diesem Fall andere Kriterien verwendet werden, wie beispielsweise zeitlich Koinzidenz (zB. [Wiskott, 1998],[Wiemer et al., 2000b]).

Die im letzten Teil der Arbeit praktizierte Anwendung der Differentialgeometrie ist ein gänzlich neuer Aspekt, nicht nur im Bereich der biologischen Modelle des visuellen Kortex, sondern generell in der Theorie der selbstorganisierenden Karten. Zwar wurde in einer technischen Anwendung ein hyperbolischer Selbstorganisierender Kartenalgorithmus [Ritter et al., 1999] verwendet. In diesem war allerdings nicht die Stimulusmanigfaltigkeit explizit gekrümmt, sondern die Manigfaltigkeit der Neuronen. Die Qualitäten dieses hyper-

bolischen Raumes lassen es lohnend erscheinen, weitergehende Überlegungen zu biologisch relevanten Kartenalgorithmen mit allgemeinen Geometrien sowohl im Input- und im Outputraum anzustellen.

Anhang A

Rekonstruktion des rezeptiven Feldes aus der neuronalen Aktivität

Angenommen, ein Neuron erweist sich in Voruntersuchungen als für einen bestimmten Ausschnitt des Sehfeldes sensitiv, dann wird dieses Neuron einer Sequenz von Pixelmustern $y_{r,t}$ ausgesetzt, die im sensitiven Bereich des Neurons zentriert sind. Im Folgenden wird sich zeigen, daß gerade solche Stimulusensembles geeignet sind, die für alle geraden $n > 0$ folgende Eigenschaft besitzen:

$$E [y_{r_1,t_1} y_{r_2,t_2} \cdots y_{r_n,t_n}]_{\mathcal{S}} = \mu^n \delta_{r_1,r_2,\dots,r_n} \delta_{t_1,t_2,\dots,t_n} \quad (\text{A.1})$$

die $E [y_{r_1,t_1} y_{r_2,t_2} \cdots y_{r_n,t_n}]_{\mathcal{S}}$ mit ungeraden n sollen dagegen verschwinden. Eine solches Stimulusensemble lässt sich leicht erzeugen, indem einzelne Pixel entweder den Wert eins oder minus eins zugeordnet bekommen. Es ist naheliegend anzunehmen, daß die Korrelation

$$E [y_{r,t} \cdot R_t]_{\mathcal{S}} \quad (\text{A.2})$$

die Information über W_r enthält. Es wird nun in mehreren Serien die bedingte Wahrscheinlichkeit einer Spikenantwort R gemessen. Die Aktivierung einer Nervenzelle sei durch Gl. (2.1) bestimmt. Die Nichtlinearität der Aktivierungsfunktion läßt sich in einer Taylorreihe entwickeln:

$$S(x) = \sum_n a_n x^n \quad (\text{A.3})$$

Als Response eines Neurons auf den Stimulus ergibt sich dann:

$$R_t = \sum_n a_n \left(\int y_{r,t} W_r dr \right)^n \quad (\text{A.4})$$

Setzt man Gl. (A.4) in (A.2) ein, ergibt sich

$$\mathbb{E} \left[y_{r,t} \left(\sum_n a_n \left(\int y_{r',t} W_r dr' \right)^n \right) \right] \quad (\text{A.5})$$

$$= \sum_n a_n \left(\int \cdots \int (\mathbb{E} [y_{r,t} y_{r_1,t} \cdots y_{r_n,t}]_{\mathcal{S}} W_{r_1} \cdots w_{r_n}) dr_1 \cdots dr_n \right) \quad (\text{A.6})$$

$$= \sum_{(n+1 \bmod 2)=0} a_n \mu^{n+1} W_r^n \quad (\text{A.7})$$

$$= \mu \frac{S(\mu W_r) - S(-\mu W_r)}{2}. \quad (\text{A.8})$$

Damit erhält man

$$\mathbb{E} [y_{r,t} R_t]_{\mathcal{S}} = F(S(\cdot), W_r) = \mu \frac{S(\mu W_r) - S(-\mu W_r)}{2}. \quad (\text{A.9})$$

Anhang B

Eichung der Orientierungsselektivität

Dieser Abschnitt der Arbeit zeigt, wie aus der aus den Symmetrien ermittelten allgemeinen Metrik Gl. (6.44)

$$\begin{pmatrix} \nu + \mu z_1 & \mu z_2 & 0 & 0 \\ \mu z_2 & \nu - \mu z_1 & 0 & 0 \\ 0 & 0 & \delta + \epsilon z_1 z_1 & \epsilon z_1 z_2 \\ 0 & 0 & \epsilon z_1 z_2 & \delta + \epsilon z_2 z_2 \end{pmatrix} \quad (\text{B.1})$$

durch Redefinition der Orientierungsstärke $|\mathbf{z}|$ der Parameter ϵ zu Null gemacht werden kann. Deshalb betrachten wir die Koordinatentransformation

$$\tilde{z}_1 \leftarrow z_1 f(|\mathbf{z}|^2) \quad (\text{B.2})$$

$$\tilde{z}_2 \leftarrow z_2 f(|\mathbf{z}|^2). \quad (\text{B.3})$$

die zu einer Metrik mit dem Tensor

$$\begin{pmatrix} \tilde{\nu} + \tilde{\mu} \tilde{z}_1 & \tilde{\mu} z_2 & 0 & 0 \\ \tilde{\mu} z_2 & \tilde{\nu} - \tilde{\mu} z_1 & 0 & 0 \\ 0 & 0 & \tilde{\delta} & 0 \\ 0 & 0 & 0 & \tilde{\delta} \end{pmatrix} \quad (\text{B.4})$$

führt Für die folgenden Ausführungen ist das Abstandsmaß bezüglich $d\xi$ unerheblich, deshalb werden die zu $d\xi$ gehörenden Terme zunächst nicht berücksichtigt. Da das Abstandsmaß d sich bei der Transformation nicht ändern soll, wird der Ansatz

$$d = \delta d\mathbf{z}^2 + \epsilon (d\mathbf{z}^2 \mathbf{z})^2 = \tilde{\delta} d\tilde{\mathbf{z}}^2 \quad (\text{B.5})$$

versucht. Aus den Gln. (B.2) und (B.3) ergibt sich

$$d\tilde{z}_1 = (f + 2z_1^2 f') dz_1 + 2z_1 z_2 f' dz_2 \quad (\text{B.6})$$

$$d\tilde{z}_2 = (f + 2z_2^2 f') dz_2 + 2z_1 z_2 f' dz_1, \quad (\text{B.7})$$

was in Gl. (B.5) eingesetzt werden kann. Man erhält

$$\tilde{\delta} (d\mathbf{z}f + 2\mathbf{z}f' (d\mathbf{z} \mathbf{z}))^2 = \tilde{\delta} d\mathbf{z}^2 f^2 + 4\tilde{\delta} (d\mathbf{z} \mathbf{z})^2 (ff' + \mathbf{z}^2 (f')^2). \quad (\text{B.8})$$

Damit wird Gl. (B.5) zu

$$\delta d\mathbf{z}^2 + \epsilon (d\mathbf{z}^2 \mathbf{z})^2 = \tilde{\delta} d\mathbf{z}^2 f^2 + 4\tilde{\delta} (d\mathbf{z} \mathbf{z})^2 (ff' + \mathbf{z}^2 (f')^2). \quad (\text{B.9})$$

Durch den Koeffizientenvergleich von $d\mathbf{z}^2$ und $d\mathbf{z} \mathbf{z}$ bekommt man nun

$$\delta = \tilde{\delta} f^2 \quad (\text{B.10})$$

$$\epsilon = 4\tilde{\delta} (ff' + \mathbf{z}^2 (f')^2) \quad (\text{B.11})$$

und damit für den Quotienten

$$\frac{\epsilon}{\delta} = 4 \frac{f'}{f} + 4\mathbf{z}^2 \left(\frac{f'}{f} \right)^2. \quad (\text{B.12})$$

Diese quadratische Gleichung hat zwei Lösungen

$$\frac{f'}{f} = -\frac{1}{2|\mathbf{z}|^2} \pm \frac{1}{2} \sqrt{\frac{1}{|\mathbf{z}|^4} + \frac{\epsilon}{\delta} \frac{1}{|\mathbf{z}|^2}}, \quad (\text{B.13})$$

die Integration liefert

$$\log(f) = \int_0^{|\mathbf{z}|} -\frac{1}{2|\mathbf{z}|^2} \pm \frac{1}{2} \sqrt{\frac{1}{|\mathbf{z}|^4} + \frac{\epsilon}{\delta} \frac{1}{|\mathbf{z}|^2}} d|\mathbf{z}| + C. \quad (\text{B.14})$$

Demnach ist die Transformation durch

$$f = C' \exp \left(\int_0^{|\mathbf{z}|} -\frac{1}{2|\mathbf{z}|^2} \pm \frac{1}{2} \sqrt{\frac{1}{|\mathbf{z}|^4} + \frac{\epsilon}{\delta} \frac{1}{|\mathbf{z}|^2}} d|\mathbf{z}| \right) \quad (\text{B.15})$$

gegeben, wobei $C' > 0$. Offensichtlich ist $f > 0$. Es muss nun noch überprüft werden, in welchen Fällen diese Transformation erlaubt ist. Damit die beiden Definitionen von \mathbf{z} zueinander homöomorph sind, ist es notwendig, dass die Transformation unter Gl. (6.50) bijektiv die beiden Definitionen der metrischen Darstellung aufeinander abbildet. Die Funktionen

$$z_1 f(|\mathbf{z}|^2) \text{ und } z_2 f(|\mathbf{z}|^2) \quad (\text{B.16})$$

müssen also monoton steigend bezüglich z_1 beziehungsweise z_2 sein. Daraus folgt für z_1 , dass

$$\frac{\partial z_1 f(|\mathbf{z}|^2)}{\partial z_1} = f + 2z_1^2 f' \stackrel{!}{>} 0 \quad (\text{B.17})$$

sein muss. Setzt man Gl. (B.13) ein, so bekommt man

$$f \left[1 + 2z_1^2 \left(-\frac{1}{2|\mathbf{z}|^2} \pm \frac{1}{2} \sqrt{\frac{1}{|\mathbf{z}|^4} + \frac{\epsilon}{\delta} \frac{1}{|\mathbf{z}|^2}} \right) \right] \stackrel{!}{>} 0 \quad (\text{B.18})$$

Offensichtlich erfüllt die positive Variante dieser Gleichung eher die Bedingung. Berücksichtigt man zusätzlich noch die metrische Nebenbedingung von Gl. (6.47), so lässt sich schließen, dass

$$\frac{\delta}{\epsilon} > -\frac{1}{|\mathbf{z}|^2}. \quad (\text{B.19})$$

Weil nun $f > 0$, erhalten wir

$$f \left[1 + 2z_1^2 \left(-\frac{1}{2|\mathbf{z}|^2} + \frac{1}{2} \sqrt{\frac{1}{|\mathbf{z}|^4} + \frac{\epsilon}{\delta} \frac{1}{|\mathbf{z}|^2}} \right) \right] > \quad (\text{B.20})$$

$$f \left[1 + 2z_1^2 \left(-\frac{1}{2|\mathbf{z}|^2} + \frac{1}{2} \sqrt{\frac{1}{|\mathbf{z}|^4} - \frac{1}{|\mathbf{z}|^4}} \right) \right] = \quad (\text{B.21})$$

$$f \left[1 - \frac{z_1^2}{|\mathbf{z}|^2} \right] \geq 0. \quad (\text{B.22})$$

Damit ist die Bedingung in jedem Fall erfüllt. Es ist jedoch unmöglich, eine andere der vier freien Funktionen der Metrik zu eliminieren. Die beiden Funktionen ν und δ müssen aufgrund der Nebenbedingungen der Gln. (6.45) und (6.46) bzw. (6.47) erhalten bleiben. Der Versuch μ zu eliminieren scheitert, da aus

$$d = \nu(d\xi_1^2 + d\xi_2^2) + \mu [(d\xi_1^2 - d\xi_2^2)z_1 + 2d\xi_1 d\xi_2 z_2] = \tilde{\nu}(d\tilde{\xi}_1^2 + d\tilde{\xi}_2^2) \quad (\text{B.23})$$

bei einem analogen Vorgehen wie zwischen den Gln. (B.8) und (B.11) sofort $f \equiv 0$ folgt. Somit transformieren die Funktionen unter der oben genannten Bedingung:

$$\tilde{\epsilon} = 0 \tag{B.24}$$

$$\tilde{\delta} = \frac{\delta(f^{-1}(|\tilde{\mathbf{z}}|))}{f^2} \tag{B.25}$$

$$\tilde{\nu} = \nu(f^{-1}(|\tilde{\mathbf{z}}|)) \tag{B.26}$$

$$\tilde{\mu} = \frac{\mu(f^{-1}(|\tilde{\mathbf{z}}|))}{f}. \tag{B.27}$$

Anhang C

Praktikable Herleitung der Metrik im Featureraum

Im Folgenden soll eine Methode entwickelt werden, wie die in Kapitel 6 eingeführten Metriken, also letztlich die Funktionen μ , δ , ν und ϵ , relativ einfach aus einem hochdimensionalen Stimulusraum konstruiert werden können.

Die direkte Bestimmung einer konkreten Metrik ist oft mühselig. Allerdings lassen sich die Symmetrien auch hier als Hilfe anwenden. Es genügt, die entsprechenden Integrale für zwei Orientierungen zu berechnen. Hier wird, der Einfachheit halber, die senkrechte und die waagrechte Orientierung verwendet. $G(\xi_1, \xi_2)(|z|)$ sei ein hochdimensionales rezeptives Feld, dessen Orientierung parallel zur ξ_2 -Achse liegt. Der Schwerpunkt liegt auf dem Ursprung.

Zunächst ergeben die Integrale

$$a = g_{11}|_{z_1=-|z|, z_2=0} = - \int \left(\frac{\partial G}{\partial \xi_1} \right)^2 d\xi \quad (\text{C.1})$$

und

$$b = g_{11}|_{z_1=|z|, z_2=0} = - \int \left(\frac{\partial G}{\partial \xi_2} \right)^2 d\xi \quad (\text{C.2})$$

aus denen erhält man die beiden Funktion ν und μ durch

$$\nu = \frac{a + b}{2} \quad (\text{C.3})$$

$$\mu = \frac{b - a}{2|z|} \quad (\text{C.4})$$

Bei den Funktionen δ und ϵ soll analog vorgegangen werden. Es ist dabei sinnvoll, zunächst die Darstellung in Polarkoordinaten $(|z|, \phi)$ zu wählen. Man berechnet zunächst an der Stelle $\phi = 0$

$$\begin{aligned} c &= \int \left(\frac{\partial}{\partial \phi} G(\hat{\mathbf{T}}(\phi) \cdot \boldsymbol{\xi}^T) \right)^2 d\boldsymbol{\xi} \\ &= \int \left| \left(\frac{\partial G}{\partial \xi_1}, \frac{\partial G}{\partial \xi_2} \right)^T \cdot (-\xi_2, \xi_1) \right|^2 d\boldsymbol{\xi} \end{aligned} \quad (\text{C.5})$$

und außerdem

$$e = \int \left(\frac{\partial G}{\partial |z|} \right)^2 d\boldsymbol{\xi}. \quad (\text{C.6})$$

Diese beiden Einträge sind genau die Diagonalkomponenten eines metrischen Tensors der Darstellung in Polarkoordinaten

$$d = \begin{pmatrix} d|z| \\ d\phi \end{pmatrix} \cdot \begin{pmatrix} e & 0 \\ 0 & c \end{pmatrix} \begin{pmatrix} d|z| \\ d\phi \end{pmatrix}. \quad (\text{C.7})$$

Da $z_1 = |z|$ ist, gilt $dz_1 = d|z|$ und $dz_2 = 2|z|d\phi$. Damit wird der metrische Tensor an dieser Stelle zu

$$\begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix} \begin{pmatrix} e & 0 \\ 0 & \frac{c}{4|z|^2} \end{pmatrix} \begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix}. \quad (\text{C.8})$$

Da andererseits gilt

$$d = \begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix} \begin{pmatrix} \delta + \epsilon|z|^2 & \\ & \delta \end{pmatrix} \begin{pmatrix} dz_1 \\ dz_2 \end{pmatrix}, \quad (\text{C.9})$$

erhält man

$$\delta = \frac{c}{4|z|^2} \quad (\text{C.10})$$

$$\epsilon = \frac{e - \delta}{|z|^2}. \quad (\text{C.11})$$

Als kleine Fingerübung soll nun eine konkrete Metrik berechnet werden. Ein einfaches Beispiel ist die Metrik elongierter Gaußblobs. Ein elongierter Gaußblob ist die Funktion

$$G(\mathbf{p}, \sigma_1, \sigma_2, \Theta)(\mathbf{x}) = \mathbf{A} \exp(-(\boldsymbol{\xi} - \boldsymbol{\xi}_0) \hat{\Omega} (\boldsymbol{\xi} - \boldsymbol{\xi}_0)) \quad (\text{C.12})$$

mit

$$\hat{\Omega} = \hat{\mathbf{T}}(\Theta) \cdot \text{Diag} \left(\frac{\sigma_1^{-2}}{2}, \frac{\sigma_2^{-2}}{2} \right) \cdot \hat{\mathbf{T}}(-\Theta) \text{ und } \hat{\mathbf{T}}(\Theta) = \begin{pmatrix} \cos(\Theta) & -\sin(\Theta) \\ \sin(\Theta) & \cos(\Theta) \end{pmatrix}. \quad (\text{C.13})$$

Das Volumen der Gaußverteilung soll auf eins normiert sein. Zudem wird $A = 1$ definiert. In der numerischen Simulation können diese Größen durch die Wahl der Stimulussets bestimmt werden. Es ergibt sich unter Benutzung von $\int \exp(-a^2 x^2) d\mathbf{x} = \sqrt{\pi}/a$

$$\int G_g^2 d\boldsymbol{\xi} = \pi \sigma_1 \sigma_2 := 1. \quad (\text{C.14})$$

Das hochdimensionale Abstandsmaß ist der euklidische Abstand. Als Orientierung des Gaußblobs sei Θ definiert. Der Schwerpunkt liegt bei $\boldsymbol{\xi}_0$. Für die obige Berechnung wird er nach Null verschoben. Dieses Stimulusensemble ist spiegelsymmetrisch längs und quer der präferierten Orientierung. Es erfüllt also die in Kapitel 6 vorausgesetzten Symmetrien. Es bleibt eine geeignete Definition der Stärke der Orientierungspräferenz zu definieren. Es wird im Folgenden

$$|z| = (\sigma_2 - \sigma_1) \quad (\text{C.15})$$

als Definition von $|z|$ verwendet werden.

Man erhält

$$a = - \int \left(\frac{\partial G}{\partial \xi_1} \right)^2 d\boldsymbol{\xi} = \pi \frac{\sigma_1}{\sigma_2} \quad (\text{C.16})$$

sowie

$$b = - \int \left(\frac{\partial G}{\partial \xi_2} \right)^2 d\boldsymbol{\xi} = \pi \frac{\sigma_2}{\sigma_1} \quad (\text{C.17})$$

und

$$c = \int \left| \left(\frac{\partial G}{\partial x_1}, \frac{\partial G}{\partial x_2} \right)^T \cdot (-x_2, x_1) \right|^2 \frac{\pi |z|^2 (4 + \pi |z|^2)}{4} \, d\mathbf{x} = \quad (\text{C.18})$$

und schließlich

$$e = \int \left(\frac{\partial G}{\partial |z|} \right)^2 \, d\mathbf{x} = \frac{\pi}{4 + \pi |z|^2}. \quad (\text{C.19})$$

Damit ergibt sich:

$$\nu = \frac{\pi^2}{2} |z|^2 + \pi \quad (\text{C.20})$$

$$\mu = \frac{\pi^2}{2} \sqrt{|z|^2 + \frac{4}{\pi}} \quad (\text{C.21})$$

$$\delta = \frac{\pi(4 + \pi |z|^2)}{8} \quad (\text{C.22})$$

$$\epsilon = -\frac{\pi(8\pi + \pi^2 |z|^2)}{32 + 8\pi |z|^2}. \quad (\text{C.23})$$

Literaturverzeichnis

- [Ahalt et al., 1990] Ahalt SC, Krishnamurthy AK, Chen P, Melton DE (1990). Competitive Learning Algorithms for Vector Quantization. *Neural Networks***3**,277-290.
- [Alonso & Martinez, 1998] J-M. Alonso, LM. Martinez (1998). Functional connectivity between simple cells and complex cells in cat striate cortex. *Nat. Neurosci.* **1**(5),395-403.
- [Atick & Redlich, 1992] Atick JJ, Redlich AN (1992). What does the Retina Know about Natural Scenes? *Neural Computation* **4**,196-210.
- [Albus & Wolf,1984] K. Albus, W. Wolf (1984). Early post-natal development of neuronal function in the kitten's visual cortex; a laminar analysis. *J. Physiol.* **348**:153-185.
- [Amari 2000] S. Amari (2000). Natural Gradient Works Efficiently in Learning. *Neural Computation* **10**, 251-276.
- [Bell & Sejnowski, 1995] A. J. Bell, T. Sejnowski (1995). An information maximization approach to blind separation and blind deconvolution. *Neural Computation* **7**, 1129-1159.
- [Bell & Sejnowski, 1997] A. J. Bell, T. Sejnowski (1997). Edges are the independent components of natural images. *Advances in Neural Information Processing Systems* Vol. **9**, 831-837.
- [Bienenstock et al., 1982] EL. Bienenstock, LN. Cooper, PW. Munro (1982). Theory for the Development of Neuron Selectivity: Orientation Specificity and Binocular Interaction in Visual Cortex. *J. Neurosci.* **2**, 32-48.
- [Blais et al., 1998] BS. Blais, N. Intrator, H. Shouval LN. Cooper, (1998). Receptive Field Formation in Natural Scene Environments: A Comparison of Single-Cell Learning Rules. *Neural Computation* **10**, 1797 - 1813.
- [Blasdel & Salama, 1986] G. G. Blasdel, G. Salama (1986). Voltage-sensitive dyes reveal a modular organization in monkey striate cortex. *Nature* **321**, 579-585.

- [Bonhoeffer & Grinvald, 1991] T. Bonhoeffer, A. Grinvald (1991). Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature* **353**, 429-430.
- [Bosking et al., 1997] WH. Bosking, Y. Zhang, B. Schofield, D. Fitzpatrick (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex, *J Neurosci.* **17**(6),2112-27.
- [Braastad & Heggelund, 1985] B. O. Braastad, P. Heggelund (1985). Development of spatial receptive-field organization and orientation selectivity in kitten striate cortex. *J. Neurophysiol.* **53**:(5), 1158-1179.
- [Brockmann et al., 1997a] D. Brockmann, H.-U. Bauer, M. Riesenhuber, T. Geisel (1997). SOM-model for the development of oriented receptive fields and orientation maps from non-oriented ON-center OFF-center inputs. ICANN97.
- [Brodman, 1909] K. Brodmann (1909), Vergleichende Lokalisationslehre der Großhirnrinde in ihren Prinzipien dargestellt aufgrund ihres Zellenbaues. J. A. Barth, Leipzig.
- [Brockmann et al., 1997b] D. Brockmann, M. Riesenhuber, H.-U. Bauer, T. Geisel (1997). Conditions for the joint emergence of orientation and ocular dominance in a high-dimensional self-organizing map. J. Bower, editor, *Proceedings to the CNS*97*.
- [Bruce et al., 1980] C. Bruce, R. Desimone, C. G. Gross (1980). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* **46**, 369-384.
- [Cajal, 1955] R. y Cajal (1955) Histologie du Système Nerveux II. C.I.S.C, Madrid.
- [Cardoso & Laheld, 1996] JF. Cardoso, B. Laheld (1996) Equivariant adaptive source separation, IEEE Trans. on Signal Proc.
- [Chance et al., 1999] Complex cells as cortically amplified simple cells (1999). FS. Chance, SB. Nelson, LF Abbott. *Nat Neurosci* **2**(3),277-82.
- [Chapman & Stryker, 1993] B. Chapman, M. P. Stryker (1993). Development of orientation selectivity in ferret visual cortex and effects of deprivation. *J. Neurosci.* **13**, 5251-5262.
- [Crair et al., 1998] M. C. Crair et al. (1998). The role of visual experience in the development of columns in cat visual cortex. *Science* **279**, 566-569.
- [Das & Gilbert, 1997] A. Das, C. Gilbert. (1997) Distortions of visuotopic map match orientation singularities in primary visual cortex. *Nature* **387**,594-598.
- [Daugman, 1984] J. Daugman (1984). Two-Dimensional spectral analysis of cortical receptive field profiles. *Vision research* **24**(9),847-856.

- [Deco & Obradovic, 1996] G. Deco, D. Obradovic. An Information-Theoretic approach to Neural Computing, Springer, New York 1996.
- [Der & Herrmann, 1994] R. Der, J.M. Herrmann (1994). Critical Phenomena in self-organizing feature maps: Ginzburg Landau approach. *Phys Rev E* **49**:5 5840-5848.
- [Dimitrov & Cowan, 1984] Dimitrov A, Cowan JD (1998). Spatial decorrelation in orientation -selective cortical cells. *Neural Computation* **10**, 1779-1795.
- [Durbin & Willshaw, 1987] R. Durbin, D. Willshaw (1987). An Analogue Approach to the Traveling salesman Problem Using an Elsatric Net Algorithm. *Nature* **326**,689-691.
- [Erwin & Miller, 1997] E. Erwin , K. D. Miller (1997). A subregion-correspondence model of binocular simple cells. *Soc. Neurosci. Abs.* **23**,227.10.
- [Erwin et al., 1995] E. Erwin, K. Obermayer, K. Schulten (1995), Models of Orientation and Ocular Dominance Columns in the Visual Cortex: A Critical Comparison. *Neural Computation* **7**,425-468.
- [Field, 1994] D. J. Field (1994), What is the goal of sensory coding ? *Neural Computation* **6** , 559-601.
- [Freeman et al., 1997] R. D. Freeman, et al. (1997), Clustering of response properties of neurons in the visual cortex, *Soc. Neurosci. Abs.* 23:227.1.
- [Grinvald et al., 1986] A. Grinvald,E. Lieke,RD. Frostig,CD. Gilbert,TN. Wiesel (1986) Functional architecture of cortex revealed by optical imaging of intrinsic signals. *Nature* **324**,361-4.
- [Gupta et al., 2000] A. Gupta, Y. Wang, H. Markram (2000) Organizing Principles for a Diversity of GABAergic Interneurons and Synapses in the Neocortex, *Science* **287**, 273-278.
- [Hateren & van der Schaaf, 1998] J.H. van Hateren, A. van der Schaaf (1998), Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. Royal Soc. Lond. B* **265**, 359-366.
- [Hateren Database] J. H. van Hateren, Natural stimuli collection. WWW: <http://hlab.phys.rug.nl/archive.html>.
- [Hebb, 1949] D. O. Hebb (1949) The Organization of Behaviour. Wiley, New York.
- [Heeger, 1991] DJ. Heeger (1991), Normalization of cell responses in cat striate cortex, *Visual Neuroscience* **9**:181-197,
- [Herrmann et al., 2000] J. M. Herrmann, N. Mayer and T. Geisel (2000), Pinwheels in a Curved Feature Map Model. *Proc. of the 4th ICCNS*, 2000.

- [Hodgkin & Huxley, 1952] AL. Hodgkin AF. Huxley (1952) A quantitative description of membrane current and its application to conduction and excitation in nerve, *J. Physiol.* **117**, 500-544.
- [Holub & Morton-Gibson, 1981] R. A. Holub, M. Morton-Gibson M. (1981), Response of visual cortical neurons of the cat to moving sinusoidal gratings: Response-contrast functions and spatiotemporal interactions. *J. Neurophysiol.* **46**:6, 1244-1259.
- [Hubel, 1990] Hubel D (1990), Auge und Gehirn: Neurobiologie des Sehens, Spektrum Verlag.
- [Hubel & Wiesel, 1959] D. H. Hubel, T. N. Wiesel (1959), Receptive Fields of single neurons in the cats striate cortex. *J. Physiol.* **160**, 106-154.
- [Hubel & Wiesel, 1963] D. H. Hubel, T. N. Wiesel (1963), Receptive fields of cells in striate cortex of very young, visually unexperienced kittens. *J. Neurophysiol.* **26**, 994-1002.
- [Hubel & Wiesel, 1962] D. H. Hubel, T. N. Wiesel (1962), Receptive Fields, Binocular Interaction and Functional Architecture in the Cat's Visual Cortex *J. Physiol.***160**:106-154.
- [Hübener et al., 1997] M. Hübener, D. Shoham, A. Grinvald, T. Bonhoeffer (1997) Relationships among three columnar systems in cat area 17. *J. Neurosci.* **17** 9270-9284.
- [Hyvärinen & Hoyer, 2000] A. Hyvärinen, P. Hoyer (1999), Emergence of phase and shift invariant features by by decomposition of natural images into independent feature subspaces *Neural Computation* **12**(7)1705-1720
- [Hyvärinen & Oja, 1997] A. Hyvärinen, E. Oja (1997), A fast fixed point algorithm for independent component analysis. *Neural Computation* **9**(7),1483-1492.
- [Jones & Palmer, 1987a] J. Jones, L. Palmer (1987), The Two-Dimensional Spatial Structure of Simple Receptive Fields in Cat Striate Cortex, *Journal of Neurophysiology* **58**(6), 1187-1211.
- [Jones & Palmer, 1987b] J. Jones, L. Palmer (1987), The Two-Dimensional Spectral Structure of simple Receptive Fields in Cat Striate Cortex, fields in cat striate cortex, *Journal of Neurophysiology* **58**(6), 1212-32.
- [Jones & Palmer, 1987c] J. Jones, L. Palmer (1987), An evaluation of the two-dimensional Gabor-filter model of simple receptive fields in cat striate cortex, *Journal of Neurophysiology* **58**(6), 1233-58.
- [Kakabadze, 1985] IM. Kakabadze (1985), (russisch) A comparative electron microscope study of neurons in the cat visual cortex during the postnatal ontogenesis, *Zitologija*, **10**, 1123-28.

- [Kim et al., 2000] DS. Kim, TQ. Duong, SG. Kim (2000) High-resolution mapping of iso-orientation columns of fMRI *nature neurosci.* **3**(2):164-169.
- [Kim et al., 1999] DS. Kim, Y. Matsuda, K. Ohki, A. Ajima, S. Tanaka (1999) Geometrical and topological relationships between multiple functional maps in cat, *Neuroreport* **10**,2515-2522.
- [Kisvárday et al., 1997] ZF. Kisvárday, E. Tóth, M. Rausch, U. Eysel (1997), Orientation-specific Relationship between Populations of Excitatory and Inhibitory Lateral Connections in the Visual Cortex of the Cat, *Cerebral Cortex*, **7**: 605-618.
- [Kobatake & Tanaka, 1994] E. Kobatake, K. Tanaka (1994). Neuronal selectivities to complex object features of ventral visual pathway of the macaque cerebral cortex. *J. Neurophysiol.* **71**:856-867.
- [Kohonen, 1995a] T. Kohonen, The self-organizing map, Springer, 1995.
- [Kohonen, 1995b] T. Kohonen (1995). The Adaptive-Subspace SOM (ASSOM) and its Use for the Implementation of Invariant Feature Detection, Proc. Int. Conf. on Artificial Neural Networks, Vol. I, pp. 3-10, EC2.
- [Kohonen et al., 1997] T. Kohonen, S. Kaski, H. Lappalainen, J. Salojaervi, The adaptive-subspace self-organizing map (ASSOM) (1997). Proceedings of WSOM'97, Workshop on Self-Organizing Maps, Espoo, Finland, June 4-6, pp. 191-196, Helsinki University of Technology, Neural Networks Research Centre.
- [Law & Cooper, 1994] Law CC, Cooper LN (1994) Formation of receptive fields in realistic visual environments according to the Bienenstock, Cooper, and Munro (BCM) theory. *Proc. Natl. Acad. Sci.* **91**: 7797-7801.
- [Liu et al., 1991] Liu Z, Gaska JP, Jacobson LD, Pollen DA (1991). Interneural Interaction between Members of Quadrature Pairs in the Cat's Visual Cortex. *Vision Res.* **32**:7 1193-1198.
- [Löwel, 1998] S. Löwel. (1998). The layout of orientation and ocular dominance domains in area 17 of strabismic cats. *Eur. J. Neurosci.* **10**,8, 2629-43.
- [Lübke & Albus, 1989] Lübke J, Albus K (1989). The postnatal development of layer VI pyramidal neurons in the cats striate cortex, as visualized by intracellular Lucifer yellow injections in aldehyde-fixed tissue, *Dev. Brain. Res.* **45**:29-38.
- [Marr & Hildreth, 1980] D. Marr, E. Hildreth. (1980) Theory of edge detection, Proc. R. Soc. Lond. B **207**, 187-217
- [Mayer et al., 1998a] N. Mayer, J. M. Herrmann and T. Geisel (1998) A cortical interpretation of ASSOMs. *Proc. Int. Conf. Art. Neural Networks (ICANN)* Vol. **2**, 961-966.

- [Mayer et al., 1998b] N. Mayer, J. M. Herrmann and T. Geisel (1998). Orientation Maps from Natural Images. n N. Elsner, R. Wehner, editors, Proceedings of the 26th Göttingen Neurobiology Conference , volume 1, page 759. Thieme, 1998.
- [Mayer et al., 1999] N. Mayer, J. M. Herrmann and T. Geisel (1999). Receptive field formation in BD cats depends on higher order statistics. In N. Elsner, U. Eysel, editors, Proceedings of the 27th Göttingen Neurobiology Conference , page 905. Thieme, 1999.
- [Mayer et al., 2000a] N. Mayer, J. M. Herrmann and T. Geisel (2000). Retinotopy and spatial phase in topographic maps. *Neurocomputing* 2000.
- [Mayer et al., 2000b] N. Mayer, J. M. Herrmann and T. Geisel (2000). Structure Formation in Visual Cortex Based on a Curved Feature Space. *Proc. of IJCNN* 2000.
- [Mayer et al., 2000c] N. Mayer, J. M. Herrmann and T. Geisel (2000). Shaping of Receptive Fields in Visual Cortex During Retinal Maturation. (submitted)
- [Mayer et al., 2000d] N. Mayer, J. M. Herrmann and T. Geisel (2000). The Impact of Receptive Field Shape on Cortical Map Formation. Proc. of the ICONIP (eingeladener Beitrag).
- [Mc Culloch & Pitts, 1943] W. S. Mc Culloch, W. H. Pitts (1943). A logical calculus of the Ideas Imminent in Neural Nets, *Bulletin of Mathematical Biophysics*, **5** 115-133.
- [Mel & Archie, 2000] Mel BW., Archie K (2000). A model for intradendritic computation of binocular disparity. *Nature Neurosci***3**:1 54 - 63.
- [Miller, 1992] K. D. Miller (1992). Development of orientation columns via competition between ON- and OFF-center inputs. *NeuroRep.* **3**, 73-79.
- [Miller, 1994] K. D. Miller (1994). A model for the development of simple-cell receptive fields and the arrangement of orientation columns through activity-dependent competition between On- and Off-center inputs. *J. Neurosci.* **14**, 409-441.
- [Nelder & Mead, 1965] Nelder JA, Mead RA (1965). A simplex method for function minimization. *Comput. J.* **7**: 308-313.
- [Noether, 1918] E. Noether (1918). Invariante Variationsprobleme. *Nachr. d. Kgl. Ges. d. Wiss. (Math. phys. Klasse)* Göttingen 235.
- [Norcia & Tyler, 1985] A. M. Norcia, C. W. Tyler (1985), Spatial frequency sweep VEP: Visual acuity during the first year of life. *Vision Res.* **25**:10, 1399-1408.
- [Obermayer et al., 1992] K. Obermayer, G. G. Blasdel, K. Schulten (1992), Statistical-mechanical analysis of self-organization and pattern formation during the development of visual maps. *Phys. Rev. A* **45**, 7568-7569.

- [Obermayer & Blasdel, 1993] K. Obermayer, G. G. Blasdel (1993). *J. Neurosci.* **13** 4114–4129.
- [Obermayer & Blasdel, 1997] Obermayer K, Blasdel GG (1997). Singularities in primate orientation maps. *Neural Computation* 9, 555-575.
- [Obermayer et al., 1990] K. Obermayer, H. Ritter, K. Schulten (1990). A principle for the formation of the spatial structure of cortical feature maps. *Proc. Nat. Acad. Sci. USA* **87**, 8345-8349.
- [Ohzawa & Freeman, 1988] I. Ohzawa, R. D. Freeman (1988), Binocularly deprived cats: Binocular tests of cortical cells show regular patterns of interaction. *J. Neurosci.* **8**:7, 2507-2516.
- [Oja, 1982] E. Oja (1982) A simplified Neuron Model as a Principal Component Analyser. *Journal of Mathematical Biology*, **15**, 267-73.
- [Olshausen & Field, 1996] B. Olshausen, D. J. Field (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **78**, 315-333.
- [Pollen & Ronner, 1981] D. Pollen, S. Ronner (1981). Phase relationships between adjacent simple cells in the visual cortex. *Science* **212**:1409-1411.
- [Raczowsky et al., 1988] D. Raczowski, D. J. Uhlrich, S. M. Sherman (1988). Morphology of Retinogeniculate X and Y Axon Arbors in Cats Raised With Binocular Lid Suture. *J. Neurophysiol.* **60**:6 2152-2167.
- [Riesenhuber et al., 1998] M. Riesenhuber, H.-U. Bauer, D. Brockmann, and T. Geisel (1998). Breaking rotational symmetry in a self-organizing map model for orientation map development. *Neural Computation*, **10**:717-730.
- [Riesenhuber et al., 1996] M. Riesenhuber, H.-U. Bauer, T. Geisel (1996). Analyzing phase transitions in high-dimensional self-organizing maps, *Biol. Cyb.* **75**, 397-407.
- [Ritter et al., 1992] H. Ritter, Th. Martinetz, K. Schulten, *Neuronale Netze*, Addison Wesley, Bonn (1992).
- [Ritter et al., 1999] H. Ritter (1999). Self-organizing Maps in non-Euclidean Spaces. *Kohonen Maps*, pp:97-108, ed.:E. Oja and S. Kaski ,
- [Ruderman, 1997] D. Ruderman (1997). Origins of scaling in natural images. *Vision Res.* **37**:23, 3385-3398.
- [Ruderman & Bialek, 1992] D. Ruderman, W. Bialek (1992). Statistics in Natural Images: Scaling in the woods, *Phys. Rev. Lett.* **73**, 814-817.

- [Rusoff & Dubin, 1977] A. C. Rusoff, M. W. Dubin (1977). Development of receptive-field properties retinal ganglion cells in kittens. *J. Neurophysiol.* **40**:5, 1188-1198.
- [Sanger, 1989] T. Sanger (1989), Optimal Unsupervised Learning in a Single-Layer Feed-forward Neural Network. *Neural Networks* **2**, 459-473.
- [Sengpiel et al., 1999] Sengpiel F, Stawinski P, Bonhoeffer T (1999). Influence of experience on orientation maps in cat visual cortex. *Nature Neurosci.* **2**, 727-732.
- [Schiller et al., 1976] P. H. Schiller, B. L. Finlay, S. F. Volman (1976). Quantitative studies of single-cell properties in monkey striate cortex. I. spatiotemporal organization of receptive fields. *J. Neurophysiol.* **39**:6, 1288-1319.
- [Sharma et al., 2000] J. Sharma, A. Angelucci, S. Mriganka (2000). Induction of visual orientation modules in auditory cortex. *Nature* **404** 841-847.
- [Sirosh & Miikkulainen, 1997] J. Sirosh, R. Miikkulainen (1997). Topographic receptive fields and patterned lateral interaction in a self-organizing model of the primary visual cortex. *Neural Comput.* **9**(3):577-94.
- [Smith et al., 1980] D. C. Smith et al. (1980). Visual acuity following binocular deprivation in the cat. *Brain Research*, **183** 1-11.
- [Sokoloff et al., 1977] L. Sokoloff, M. Reivich, C. Kennedy, M. H. Des Rosiers, C. S. Patlak, K. D. Pettigrew, O. Sakurada, M. Shinohara (1977). The [14C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. *J Neurochem.* **28**(5):897-916.
- [Swindale, 1996] N. V. Swindale (1996). The development of topography in the visual cortex: A review of models. *Network* **7**, 161-247.
- [Swindale et al., 2000] N. V. Swindale, D. Shoham, A. Grinvald, T. Bonhoeffer, M. Hübener (2000). Optimizing coverage in the cortex. *Nature Neuroscience* **3**:8 750-751.
- [Tanaka, 1995] S. Tanaka (1995). Topology-based approach to the understanding of visual information representation. RIKEN Review No. 9 (April 1995): Focused on Brain and Information Science, 39-40.
- [Thomas & Cowan, 2000] Thomas PJ, Cowan JD, Spin model for orientation map via reduction of Hebb's rule. To appear in *Neurocomputing*.
- [Tootell et al., 1982] RBH. Tootell, MS. Silverman, E. Switkes, RL. de Valois (1982). Deoxyglucose Analysis of Retinotopic Organization in Primate Striate Cortex. *Science* **218** 902-4.

- [Tootell et al., 1981] RBH. Tootell, MS. Silverman, E. Switkes, RL. de Valois (1981). Spatial frequency columns in primaty visual cortex *Science* **214** 813-815.
- [von der Malsburg, 1973] Ch. von der Malsburg (1973). Self-organization of orientation sensitive cells in striate cortex. *Kybernetik*, **14** 49-54.
- [Wässle und Boycott, 1991] H. Wässle, B. B. Boycott (1991). Functional architecture of the mammalian retina. *Physiol. Rev.* **71**, 447-480.
- [White et al., 1999a] L. E. White, W. H. Bosking, S. M. Williams, D. Fitzpatrick (1999). Maps of Central Visual Space in Ferret V1 and V2 Lack Matching Inputs from the two eyes. *J. Neurosci* **19**(16):7089-7099.
- [White et al., 1999b] L. E. White, W. H. Bosking, D. Fitzpatrick (1999). Visuotopic discontinuity at the V1/V2 border without disruption of the map of orientation preference in ferret visual cortex. In *Soc. Neurosci. Abs.* **25I**, 313.2
- [Wiemer et al., 2000a] J. Wiemer, T. Burwick, W. von Seelen (2000), Self-organizing maps for visual feature representation based on natural binocular stimuli, *Biol. Cybern.* **82**, 97-110
- [Wiemer et al., 2000b] J. Wiemer, F. Spengler, F. Joubin, P. Stagg, S. Wacquant, Learning cortical topography from a spatiotemporal stimuli, *Biol. Cybern.* **82**, 173-187.
- [Wiesel & Hubel, 1974] T. N. Wiesel, D. N. Hubel (1974), Ordered arrangement of orientation columns in monkeys lacking visual experience. *J. Comp. Neur.* **158**, 307-318.
- [Wiskott, 1998] Learning Invariance Manifolds (1998), *Neurocomputing*, **26/27**(1):925-932.
- [Wolf et al., 1994a] F. Wolf, H.-U. Bauer, T. Geisel (1994), Formation of field discontinuities and islands in visual cortical maps. *Biol. Cyb.* **70**, 525-531.
- [Wolf & Geisel 1998] F. Wolf, T. Geisel (1998) Spontaneous pinwheel annihilation during visual development. (paper and electronic supplement) *Nature* 395, 73-76.
- [Wolf et al., 1994b] F. Wolf, K. Pawelzik, and T. Geisel (1994), Linear and nonlinear evolution in maps of orientation preference . F. H. Eeckman and J. Bower, editors, *Computation and Neural Systems*, Boston, Kluwer Academic Publishers
- [Wolf, 1998] F. Wolf (1998) Strukturbildung in der Entwicklung des visuellen Kortex, Dissertation, Universität Frankfurt am Main.
- [Yuille, 1989] AL. Yuille, DM. Kammen, SD. Cohen (1989), Quadrature and the Development of Orientation Selective Cortical Cells by Hebb Rules, *Biol. Cybern.* **61**:183-194.

Danksagung

Ich danke Herrn Prof. Geisel für die Aufnahme an das Institut, sowie für die guten Arbeitsbedingungen, die mit der Betreuung dieses Vorhabens einhergingen. Michael Herrmann danke ich für die freundschaftliche Unterstützung und engagierte Hilfe, sowie beinahe väterlichen Beistand beim Zustandekommen dieser Arbeit. Hans-Ulrich Bauer sei dafür gedankt, dass er mich nach Göttingen in diese Arbeitsgruppe geholt hat. Meiner Schwester Michaela gilt mein besonderer Dank für ihr Gegenlesen, für das sie viel Zeit geopfert hat. Weiterhin danke ich der Arbeitsgruppe, Steffen Arnich, Rita Bartels, Dmitri Bibitchkov, Wolf-Dieter Brandt, Dirk Brockmann, Markus Diesmann, Silke Dodel, Ragnar Fleischmann, Lars Hufnagel, Matthias Kaschube, Roland Ketzmerick, Tsampikos Kottos, Björn Naundorf, Alexander Ossipov, Mark-Felix Otto, Felix Petri, Holger Schanz, Michael Schnabel, Dennis Springguth, Frank Steinbach, Tom Tezlaff, Marc Timme, Corinna Trautsch, Matthias Weiß und Fred Wolf, für die angenehme Atmosphäre und Diskussionsbereitschaft.

Lebenslauf

Name: Norbert Michael Mayer

- 30.8.1971 geboren in Bobingen
- 1978-1982 Grundschule Gablingen (Bayern)
- 1982-1991 Gymnasium Bischof-Neumann-Schule in Königstein / Taunus
- 1991-1996 Physikstudium in Frankfurt am Main
- 1996-2000 Promotion am Max-Planck-Institut für Strömungsforschung