

Contour Integration Models Predicting Human Behavior

Nadja Schinkel-Bielefeld

July 2007

Contour Integration Models Predicting Human Behavior

Vom Fachbereich für Physik und Elektrotechnik
der Universität Bremen

zur Erlangung des akademischen Grades eines
Doktor der Naturwissenschaften (Dr. rer. nat.)
genehmigte Dissertation

von
Nadja Schinkel-Bielefeld, M.Sc.
aus Bremen

1. Gutachter: Prof. Dr. rer. nat. Klaus Pawelzik
2. Gutachter: Prof. Dr. rer. nat. Günter Meinhardt

Eingereicht am: 17. Juli 2007

Datum des Kolloquiums: 15. August 2007

Abstract

Contour integration is believed to be a fundamental process in object recognition and image segmentation. However, its neuronal mechanisms are still not well understood. Psychophysical experiments showed that humans are remarkably efficient in integrating contours even if these are jittered or partially occluded. Therefore the brain requires a reliable algorithm for extracting contours from stimuli. Several recent publications demonstrated that the brain often uses optimal strategies to integrate sensory information. Hence in this thesis I want to tackle the question which contour integration model describes human contour integration best.

Mathematically, contour ensembles can be characterized by a conditional link probability density between oriented edge elements, termed an association field. This association field can be used to generate contours or vice versa to extract a contour from a stimulus. While in most neuronal network models all inputs to a neuron are summed up, in such a probabilistically motivated neural network for contour integration the afferent input due to the visual stimuli and the lateral input from horizontal network interactions are multiplied.

Long-range horizontal interactions in primary visual cortex link orientation columns with similar preferred orientations and are often assumed to be the neuronal substrate for the association field. Experimental findings in monkeys suggest isotropic long-range horizontal connections, spreading symmetrically into all directions from an orientation column. In contrast, probabilistic models require unidirectional lateral interactions, linking orientation columns in only one direction, in order to get optimal contour detection performance.

Using stimuli generated from given association fields, our numerical simulations show that contour detection performance for both, probabilistic-multiplicative as well as additive models reaches human performance. Hence detection performance alone is insufficient to rule out either model class. However, psychophysical experiments with humans reveal that contour detection errors are not made randomly, but are highly correlated among different subjects. Thus a model describing contour integration in the brain should not only explain human contour detection performance, but should also reproduce these systematic errors made by humans. Comparison between misdetections of humans and mispredictions of the models on a trial-by-trial basis was used to evaluate different model dynamics and association fields. This suggests that unidirectional multiplicatively coupled horizontal interactions are required in order to explain human behavior. Furthermore, cortical magnification factors have to be taken into account and a fixed association field geometry for all stimuli is preferable

instead of using for each contour the association field employed for the generation of this contour.

Contents

List of figures	v
Notations and abbreviations	xiii
Acknowledgements	xvii
1 Introduction	1
1.1 Outline of the dissertation	5
2 Biological background	7
2.1 Single neurons	7
2.2 Neuronal networks	12
2.2.1 Firing-rate models	12
2.2.2 Multiplicative gain modulation	14
2.3 The visual system	15
2.3.1 The retina and LGN	16
2.3.2 The organization of the primary visual cortex	18
2.3.3 Non-linear response properties within the primary visual cortex	22
2.3.4 Visual pathways and higher cortical areas	23
3 Contour integration - existing models and experimental findings	25
3.1 Psychophysical studies	26
3.1.1 Contour integration in the visual periphery	28
3.1.2 Subjective contours	29

Contents

3.2	Imaging studies	30
3.3	Computational models	31
3.3.1	Neuronal models	31
3.3.2	Probabilistic and ideal observer models	36
4	Theory: model classes and their dynamics	41
4.1	Bayesian contour integration	41
4.1.1	Contour creation	41
4.1.2	Contour integration	42
4.1.3	Bayesian model for specific contour position	46
4.2	Contour detection and read out mechanisms	46
4.2.1	Neuronal context	46
4.2.2	Estimating the saliency	48
4.2.3	Criteria for contour detection	48
4.2.4	Iterative models	49
4.3	Synaptic input	51
4.3.1	The afferent input	51
4.3.2	Lateral input	55
4.3.3	Relations between different forms and couplings of inputs	59
4.3.4	Summary	59
5	Analysis of model dynamics	61
5.1	Stability of the iterative models and effect of normalization	61
5.1.1	Illusions supporting the background	66
5.1.2	Star-Illusions	70
5.2	Fading activity	74
5.3	Adjustment of afferent and lateral input strength	75
6	Robustness against noise and orientation jitter	79
6.1	Stimuli and simulations	79
6.2	Performance and chance level	82

6.3	Robustness against synaptic noise	83
6.4	Robustness against orientation jitter	84
7	Evaluation of model classes	89
7.1	Human psychophysics	89
7.1.1	Generation of stimuli	89
7.1.2	The contour integration task	92
7.1.3	Results: human performance	92
7.1.4	Correlations among human subjects	95
7.1.5	Method: correlation analysis of human decisions and model predictions	96
7.1.6	The model as nullhypothesis	99
7.1.7	Testing the ROC-methods on surrogate data	101
7.1.8	Correlations among human observers	102
7.1.9	Benchmarks	105
7.1.10	Simulations and model performance	106
7.1.11	Comparison of different estimators and readout mechanisms	109
7.2	Results of model evaluation	110
7.2.1	Eccentricity dependence	115
7.2.2	Angular dependence of the association field	117
7.2.3	Symmetry of association field	120
7.2.4	The range of the association field	123
7.3	Discussion	135
8	Conclusion	141
A	Randomly emerging contours in a hexagonal grid	149
B	Average link saliency between background elements	153
	Bibliography	155

Contents ---

List of Figures

2.1	Equivalent electrical circuit of a one-compartment model	9
2.2	Membrane potential and state of ion channels during an action potential	11
2.3	Center - surround receptive fields	16
2.4	Primary visual pathway in humans	17
2.5	Orientation selective simple cells	18
2.6	Pinwheel	20
2.7	Illustration of a hypercolumn in V1	21
2.8	Orientation map and lateral connections in the tree shrew	21
2.9	Retinotopic map in V1 of a monkey	22
3.1	Examples for different grouping tasks	28
3.2	Kanizsa triangle	30
3.3	Neural components of the contour integration model by Li	32
4.1	Illustration of the circular coordinates r , α , and β	44
4.2	Overview over different model dynamics, salience estimators, contour detection criteria and representations of performance	47
4.3	Comparison of different step widths	50
4.4	Afferent input	52
4.5	Different implementations of orientation columns	53
4.6	Eccentricity dependence of the afferent input	54
4.7	Determination of curvature and alignment	56
4.8	Contours generated with different association fields	57

List of Figures

4.9	Possible implementations of unidirectional and bidirectional association fields	58
5.1	Simplified stimulus	62
5.2	Activity in the fixed point	63
5.3	Contour detection and illusions in the normalized additive model .	66
5.4	Stimulus with illusions supporting the background	67
5.5	Saliency ratios as predicted by the additive model without normalization	68
5.6	Saliency ratios as predicted by the additive model with normalization	68
5.7	Saliency ratios as predicted by the additive model with normalization and wide afferent input	69
5.8	Saliency as predicted by the normalized multiplicative model for two simple stimuli	71
5.9	Star-like edge configuration where summation and maximum estimator fail to find the contour	72
5.10	Possible infinitely long contours within an ambiguous stimulus . .	73
5.11	Star-like edge configuration where the summation estimator finds illusions	74
5.12	One-dimensional stimulus containing a finite contour surrounded by perpendicular background elements.	74
5.13	Illustration of the need for fine-tuning	77
6.1	Example stimulus: hexagonal grid	80
6.2	Performance for different contour detection criteria and association field used for the simulation	83
6.3	Contour detection performance for different levels of synaptic noise	84
6.4	Contour detection performance for various orientation jitters . . .	85
6.5	Overlap of afferent input tuning function and the lateral input . .	86
7.1	Example stimuli used for the psychophysical experiments	90
7.2	Psychophysical contour detection experiment	91
7.3	Human contour detection performance for different SOAs, element distances and orientation jitter	92
7.4	Human contour detection performance for different jitters, element distances, and SOA	93

7.5	Eccentricity dependence of human contour detection performance	94
7.6	Eccentricity dependence of human contour detection performance for tasks with high and low difficulty	94
7.7	Correlation analysis with changed null hypothesis	100
7.8	Test of ROC-methods on surrogate data	103
7.9	Correlations between responses of different human observers . . .	104
7.10	Model performance in dependence of I_l/I_a	107
7.11	Comparison of different read out mechanisms and saliency estimators	111
7.12	Distribution of inter-element distance between nearest neighboring contour elements	112
7.13	Performance of the model with unidirectional, multiplicative couplings, and ϵ -independent I_{aff} . The AF has ideal angular dependence and range ρ_{tot} taken from the nearest neighbor distribution over all stimuli	112
7.14	Correlations of the model with unidirectional, multiplicative coupling, ϵ -independent I_{aff} , and an AF with ideal angular dependence and range ρ_{tot}	113
7.15	Model performance for the model with unidirectional, additive couplings and ϵ -independent I_{aff} . The AF has ideal angular dependence and range ρ_{tot}	114
7.16	Correlations of the model with unidirectional, additive coupling, ϵ -independent I_{aff} , and an AF with ideal angular dependence and range ρ_{tot}	115
7.17	Model performance for the model with unidirectional, multiplicative couplings and ϵ -dependent I_{aff} . The AF has ideal angular dependence and range ρ_{tot}	116
7.18	Correlations of the model with unidirectional, multiplicative coupling, ϵ -dependent I_{aff} , and an AF with ideal angular dependence and a range ρ_{tot}	117
7.19	Performance for the model with unidirectional, multiplicative couplings and ϵ -dependent I_{aff} . The AF has fixed angular dependence $\sigma_\alpha = \sigma_\beta = 18 \text{ deg}$ and a range taken from the nearest neighbor distance distribution averaged over all stimuli	118
7.20	Correlations of the model with unidirectional, multiplicative coupling, ϵ -dependent I_{aff} , and an AF with fixed angular dependence $\sigma_\alpha = \sigma_\beta = 18 \text{ deg}$ and range ρ_{tot}	119

List of Figures

7.21	Performance for the model with bidirectional, multiplicative couplings and ϵ -dependent I_{aff} . The AF has fixed width and range ρ_{tot}	121
7.22	Correlations of the model with bidirectional, multiplicative coupling, eccentricity-dependent afferent input, and an AF with fixed width $\sigma_{\alpha} = \sigma_{\beta} = 18 \text{ deg}$ and a range obtained from the nearest neighbor distribution summed over all stimuli	122
7.23	Distribution of distance between nearest and next-to-nearest neighboring contour elements	123
7.24	Next-to-nearest neighbor interactions and short contours or contours with kinks	124
7.25	Performance for the model with bidirectional, multiplicative couplings and ϵ -dependent I_{aff} . The AF has fixed width and range $\rho_{2,\text{tot}}$	126
7.26	Correlations of the model with bidirectional, multiplicative coupling, eccentricity-dependent afferent input and an AF with fixed width $\sigma_{\alpha} = \sigma_{\beta} = 18 \text{ deg}$ and a range obtained from the next to nearest neighbor distribution averaged over all stimuli	127
7.27	Model performance for the model with unidirectional, multiplicative couplings and ϵ -dependent I_{aff} . The AF has fixed width and range $\rho_{2,\text{tot}}$	128
7.28	Correlations of the model with unidirectional, multiplicative coupling, eccentricity-dependent afferent input and an AF with fixed width $\sigma_{\alpha} = \sigma_{\beta} = 18 \text{ deg}$ and a range obtained from the distribution over next to nearest neighbor distances averaged over all stimuli	129
7.29	Model performance for the model with unidirectional, multiplicative couplings and ϵ -dependent I_{aff} . The AF has fixed width and range ρ_{L}	130
7.30	Correlations of the model with unidirectional, multiplicative coupling, ϵ -dependent afferent input, and an AF with fixed width $\sigma_{\alpha} = \sigma_{\beta} = 18 \text{ deg}$ and nearest neighbor connections for each individual element distance	131
7.31	Model performance for the model with bidirectional, multiplicative couplings and ϵ dependent I_{aff} . The AF has fixed width and range ρ_{L}	132

7.32	Correlations of the model with bidirectional, multiplicative coupling, eccentricity-dependent afferent input, and an AF with fixed width $\sigma_\alpha = \sigma_\beta = 18$ deg and connections between nearest neighbors for each element distance	133
7.33	Performance for the model with $I = I_a I_{\text{aff}} + I_l I_{\text{lat}} + I_m I_{\text{aff}} \cdot I_{\text{lat}}$ and ϵ dependent I_{aff} . The AF is unidirectional and has fixed width and range ρ_{tot}	134
7.34	Correlations of the model with $I = I_a I_{\text{aff}} + I_l I_{\text{lat}} + I_m I_{\text{aff}} \cdot I_{\text{lat}}$	135
A.1	Example for randomly arising contours in the background	150

Notations and abbreviations

Notations

Here we list the notations for the variables which were used during this thesis. Variables introduced in chapter 2 and 3 do not occur again in later sections of this thesis and therefore are omitted in this list.

$A(x, \phi)$	activity of a neuronal population with receptive field at position x and preferred orientation ϕ
A_b	activity of a neuronal unit representing a background element
A_c	activity of a neuronal unit representing a contour element
$A_{i,b}$	activity of a neuronal unit representing an illusory edge element at the position of a background element
$A_{i,c}$	activity of a neuronal unit representing an illusory edge element at the position of a contour element
$A_s(x, \phi)$	activity of a neuronal population with receptive field at position x and preferred orientation ϕ in the stationary state
α	viewing angle from source edge to destination edge
α'	viewing angle from destination edge to source edge
β	orientation difference between two edge elements
C	contrast of an edge element
C_c	contrast of contour elements
C_b	contrast of background elements
ϵ	mean eccentricity of the contour elements within one stimulus
η_θ	alignment jitter in degree
η_μ	synaptic noise on the afferent input
η_μ^s	static synaptic noise on the afferent input
η_μ^d	dynamic synaptic noise on the afferent input
$F(r)$	distance dependence of the association field
γ_α	alignment parameter
γ_β	curvature parameter
$g(\cdot)$	gain function

List of Figures

i	edge element with position x_i and orientation θ_i
$I(x, \phi)$	total neuronal input of an orientation column with preferred orientation ϕ and receptive field at position x
$I_0(\cdot)$	modified Bessel function of the first kind of order 0
I_a	strength of the afferent input
I_{aff}	afferent input
$I_{\text{aff}}^{\text{wide}}$	wide afferent input
I_{lat}	lateral input
I_l	strength of the lateral input
I_m	strength of total multiplicatively coupled neuronal input $I_{\text{aff}} \cdot I_{\text{lat}}$
κ	concentration parameter
k	sum over total neuronal input for all positions and orientations
K	number of preferred orientations
K_a	number of preferred orientations in a model where preferred orientations are adapted such that it matches the orientation of the stimulus
L	number of contour elements
LH	left hemifield
L_{max}	number of maximum saliencies used to determine the location of the contour
λ	wave length of Gabor patch
$M(\cdot)$	Von Mises function
M	number of rows within a hexagonal stimuli
$\mu(x, \phi)$	evidence for an edge at position x with orientation ϕ
N	total number of edge elements within a stimulus
N_l	number of elements within one row of a hexagonal stimulus
N_s	number of stimuli from which the performance is determined
n	normalization factor
$n_{A,B}$	number of correct responses of observer A, B
n_{id}	number of identical responses of two observers
$n_{\text{id}}^{\text{ex}}$	number of expected identical responses of two observers assuming randomly distributed errors
n_m	number of matching trials between generic model and observer
$n_{\text{id}}^{\text{ob}}$	observed number of identical responses of two observers
n_{ob}	number of observers
n_{ori}	number of preferred orientations
n_t	number of trials in a contour detection experiment
n_{iter}	number of iteration steps
ΔP	difference between model performance and human performance in terms of human standard deviation
\mathbf{P}	matrix containing the link probabilities $p_{i,j}$
p_A, p_B	performance of observer A or B, respectively
$p(r, \alpha, \beta)$	association field in circular coordinates

P_{element}	relative number of trials where the maximum activity of the model was at one of the contour elements
$P_{\text{hemifield}}$	relative number of trials where the observer detected the contour on the correct hemifield of the screen
$p_{\text{hemi}}^{\text{chance}}$	chance level of P_{hemi}
$P_{L_{\text{max}}}$	performance as predicted by the model using the L_{max} elements of maximum saliency to determine the location of the contour
$p_{L_{\text{max}}}^{\text{chance}}$	chance level of $P_{L_{\text{max}}}$
P_{rand}	probability for randomly arising contours in the background
ϕ_G	phase of Gabor patch
ϕ	preferred orientation of an orientation column
p^{curv}	association field that links perfectly aligned edge elements, allowing an orientation difference of up to $\pi/2$
p^δ	association field that connects only perfectly aligned collinear edge elements
$p^{\text{ex}}(n_{\text{id}}^{\text{ex}})$	discrete probability distribution over identical responses $n_{\text{id}}^{\text{ex}}$ expected when assuming that errors are distributed randomly
$p^{\text{ob}}(n_{\text{id}}^{\text{ob}})$	discrete probability distribution over identical responses $n_{\text{id}}^{\text{ob}}$ observed in an experiment
p^{f}	false alarm rate
$p_{i,j}$	probability that a contour through edge i next passes edge j
p^{t}	hit rate
p^{wide}	wide association field that links edge elements with $\alpha = \beta = 0$ and with $\alpha = \beta = \pi/4$
RH	right hemifield
r_0	average inter element distance between nearest neighbors
ρ	association field
ρ_{tot}	distribution over nearest neighboring contour elements averaged over all jitters σ_j and element distances r_0
$\rho_{2,\text{tot}}$	distribution over nearest and next-to-nearest neighboring contour elements averaged over all jitters σ_j and element distances r_0
ρ_L	distribution over nearest neighbors for contours of L elements averaged over all jitters σ_j
σ_{aff}	width of the afferent input
σ_α	alignment-dependent width of the association field
σ_β	curvature-dependent width of the association field
$\sigma_{A,B}$	sequence of contour detection decisions of observer A, B
σ_G	width of Gabor patch
$\langle s_{i,j} \rangle$	average link saliency between the two elements i and j
σ_j	curvature and alignment jitter of a contour generated from an association field with $\sigma_\alpha = \sigma_\beta = \sigma_j$
s_b	saliency of a background element

List of Figures

$s_{i,b}$	saliency of an edge at the same position as a background element but with perpendicular orientation
s_c	saliency of a contour element
$s_{i,c}$	saliency of an edge at the same position as a contour element but with perpendicular orientation
s_{LH}	saliency of the left hemifield
s_{RH}	saliency of the right hemifield
$s^{\max}(\mathbf{x})$	saliency of position \mathbf{x} estimated with the maximum estimator
s_b^{\max}	saliency at the position of a background element as estimated with the maximum estimator
s_c^{\max}	saliency at the position of a contour element as estimated with the maximum estimator
$s^{\text{sum}}(\mathbf{x})$	saliency of position \mathbf{x} estimated with the summation estimator
s_b^{sum}	saliency at the position of a background element as estimated with the summation estimator
s_c^{sum}	saliency at the position of a contour element as estimated with the summation estimator
t	iteration time
Δt	iteration time step
τ	decay constant
θ	orientation of an edge element
$u_\theta(\phi)$	afferent input to a neuron of preferred orientation ϕ caused by an edge of orientation θ
\mathbf{x}	position of an edge element or of the receptive field of a neuron, respectively

Abbreviations

AF	association field
2AFC	two-alternative forced choice
fMRI	functional magnetic resonance imaging
IT	inferior temporal cortex
LGN	lateral geniculate nucleus
MST	medial superior temporal area
MT	middle temporal area
ROC	receiver operator characteristic
SOA	stimulus onset asynchrony
V1	primary visual cortex
V2	secondary visual cortex

Acknowledgements

First of all I want to express my deep gratitude to Dr. Udo Ernst for everything he has done for me. Without his invaluable guidance and his whole-hearted support this thesis had not been possible. I am very grateful to Prof. Dr. Klaus Pawelzik for his continued support, for long and fruitful discussions, for inspiration and for his seemingly never ending motivation, and for creating a productive working atmosphere in the institute. For a collegial atmosphere I also want to thank our secretary Mrs. Janßen, without whose energy and talent for organization many Christmas parties, "Kohlfahrten" and other group activities had not taken place. Furthermore she saved us from the burden of many administrative works.

My sincere thanks to Klaus Bowe, Andreas Beuthner, Dieter Gauck and especially to our voluntary system administrators Udo Ernst and David Rotermund who kept our computers and the computer cluster running and even found space for my hundreds of Gigabytes of simulation data.

I wish to thank Simon Neitzel, Dr. Sunita Mandon and Prof. Dr. Andreas Kreiter for experimental data and a good collaboration. Many thanks also to Dr. Birgit Mathes who brought to live the "path finder meetings" and thus provided an ideal opportunity for the exchange between the research on contour integration in psychology, biology and physics.

I want to say thank you to my office mate Erich Schulzke, not only for being a walking matlab tutorial but also for long and pleasant discussions about neuroscience and teaching physics, as well as the sense of life as a PhD student.

Thanks also to our whole working group, i.e. Dr. Roland Rothenstein, Prof. Dr. Helmut Schwegler, Dr. Bailu Si, Dr. Matthias Bethge, Hedinn Steingrimsón, Axel Etzold, David Rotermund, Dr. Erich Schulzke, Dr. Christian Eurich, Markus Riegel, Onno Böhler, Ulrich Cherdron, David Engelskirchen, Daniel Bartz, Felix Patzelt and Dr. Udo Ernst and the group of Prof. Dr. Stefan Bornholdt among other things for the nice company during lunch breaks and thanks to the Mensa for providing a food selection, which is still enjoyable after four years of daily visits.

I thank the DFG and especially PIP for financial support on the visit of several conferences.

List of Figures

I am thankful to Susanne Abeld and Anneliese Niehoff for creating "PlanM", a mentoring program for woman researchers of the natural sciences, and the many interesting seminars they organized during this program. I also warmly thank the other mentees participating in this program, namely Andra Thiel, Helena Filipsson, Annette Kolb, Inessa Seifert, Uta Berger, Britta Lüder, Dagmar Kieke, Maike Isermann, Michaela Wilhelm and Kerstin Kirchner, for advice and encouragement which by far exceeded the end of the program. Many thanks to Prof. Dr. Adele Diederich for advise concerning strategic career planning and hospitality at Purdue university.

My special thanks goes to Dr. Ruth Rosenholtz and the Perceptual Science group in the Department for Brain and Cognitive Sciences at the Massachusetts Institute of Technology for their kind hospitality during several weeks of writing up my thesis. Ruth's help as well as the support of Dr. Michael Herrmann, Prof. Dr. Theo Geisel, Prof. Dr. Kerstin Schill and Dr. Christoph Zetsche during my postdoc applications and fund raising are greatly appreciated.

Last and most importantly I am grateful to my family and especially my husband Tom Bielefeld for their support. I thank Tom not only for listening to every single rehearsal talk I gave during my PhD on two different continents, but also for encouragement and the fact that he never stopped believing in me.

Chapter 1

Introduction

”He cannot see the wood for the trees!” This commonly used expression describes a situation where someone does not grasp the big picture because he concentrates too much on the details and is not able to put them together into a larger context.

When it comes to visual perception humans are remarkably good in grouping different elements of a visual scene into meaningful objects. We effortlessly recognize a squirrel in the trees even though it might be partly occluded by leaves and smaller branches and we can only intermittently see its outline. To recognize this visual input as one object is a difficult task for an image segmentation algorithm. In fact there is no object recognition or image segmentation algorithm that could compete with human visual perception of natural images. Even for non occluded, clearly visible scenes this is astonishing when considering that many cells in the retina and in the lower visual areas of the brain appear to have only localized receptive fields and are sensitive either to a certain color, or to a small edge element of a certain orientation, or to a dark dot surrounded by light. Then in a second step this local information seems to be combined in order to obtain contours or objects. So there is really a danger not to ”see the contour for the edge elements” as the brain has to group different elements in order to perceive the contours or objects within a visual scene.

Gestalt rules which describe visual perception have first been proposed by Wertheimer (133) and Koffka (64). The rule which is most important for contour integration is the law of good continuation which states that collinearly aligned elements are grouped together. The strength of this grouping effect can be seen in psychophysical experiments, where a contour consisting of collinearly aligned edge elements in a background of randomly oriented edge elements pops out. This means it requires no attention or visual search to detect this contour and the time necessary to detect this contour does not increase with the number of background elements. One of the most prominent psychophysical studies on contour integration is the study of Field and colleagues (27). There subjects had

to detect a contour consisting of aligned Gabor patches along a curved path in a background of randomly oriented edge elements. In a forced choice procedure subjects were able to identify the path when the angles between adjacent edge elements did not exceed 60 deg. Since then there have been many psychophysical studies on contour integration on humans and monkeys using similar paradigms (for a review see (67) and (47)). These dealt not only with the effect of orientation jitter on human contour integration but also analyzed i.e. the effects of closure (68) and eccentricity of the contour (45; 92; 79), as well as the speed of contour integration (83; 6; 44; 9). These show that the processing time depends on the curvature of the contour and ranges between 100 ms and 300 ms (6; 44). More recent investigations reveal that even a stimulus presentation time of only 30-60 ms is sufficient to perceive a contour (83). This poses the question which mechanisms are capable of explaining the high speed and efficiency of contour integration in humans and monkeys.

Field and colleagues (27) proposed an association field that links feature detectors with proximally located receptive fields and approximately collinearly aligned preferred orientations. It is often assumed that the neuronal substrate for the association field are the long-range horizontal interactions in V1, as these connect cells with similar preferred orientations (36; 37; 104). Furthermore in cat these excitatory horizontal connections link primarily cells whose receptive fields are approximately displaced collinearly to their preferred orientation (114; 118).

There are several ways of building a contour integration model: Many models, which can be regarded as 'bottom up approaches' start with the known neuronal anatomy of V1 and sometimes also higher visual areas and construct a model which solves the task of contour integration using these neuronal components (i.e. (77; 97; 137; 131)). Generally models which are based on empirically observed structures of the brain can have the advantage, that they sometimes do not only explain the phenomenon for which the model originally was built, but also additional effects. The opposite strategy would be a 'top-down approach'. Models employing this approach search for the most effective strategy to solve the given task. Based on that, they can draw conclusions of what components and mechanisms are needed to build such a model. Depending on whether the model aims at a technical application or at explaining brain functions in a second step one can search for an approximation of these strategies which can be implemented with neuronal components. While it is generally assumed that a biological plausible model should be constructed from known neuronal structures for which there is empirical evidence, top-down approaches have the advantage that they identify the mechanisms and components which are required for solving a certain task effectively. This could give hints on necessary neuronal components or neuronal properties which so far have been missed in empirical studies.

Most models dealing with contour integration in the brain use 'bottom-up' ap-

proaches and are based on neurophysiological and neuroanatomical data. Already one of the earlier contour integration models following this approach suggested a nonlinear coupling of different synaptic inputs, which was termed the 'bipole property' (41). The main idea of this mechanism is that a neuronal column, which does not receive direct afferent input from a visual stimulus, becomes only activated when simultaneously receiving lateral input from neighboring columns on both sides. This bipole property is also part of a later model developed in Grossberg's lab (106; 42), which focuses more on the laminar organization of the visual cortex and interlaminar cortical circuits. It also includes area V2. In contrast to this multi-layered network, several other models based on physiological findings and known neuronal components concentrate on contour integration in V1 (77; 137; 131) and show that local processing in V1 is already sufficient to extract contours even without feedback from higher cortical areas. In the models from Li (77) and Yen and Finkel (137) contours are represented by oscillatory neurons which synchronize on fast time scales. The main difference between the two models is that Li assumes inhibition for cells whose receptive fields are displaced orthogonally to their preferred orientation, while Yen and Finkel assume excitation. Ursino and La Cara (131) especially investigated the cooperation between feed-forward and feedback connections in their model and suggest that the interplay of these two connections helps to avoid too long processing times when extracting contours from noise.

Only few models follow a top-down approach. One example for this is the model of Williams and Thornber (136), who regard a contour as the trajectory of a particle moving according to brownian motion. They derive a probabilistic model for contour completion and show that illusions like the shape perceived in the center of Koffka crosses can be explained by their model.

Many models can reproduce human behavior in psychophysical experiments qualitatively, but usually they lack a quantitative comparison between psychophysical experiments and modeling and I know of no work which uses further benchmarks than just the averaged performance for such a quantitative comparison. In this thesis we close this gap by comparing quantitatively the average performance of several models to the contour detection performance of human observers using the same stimuli for simulations and experiments. In addition we develop a second method to evaluate different models by analyzing correlations between human responses on a trial-by-trial basis and requiring models to reproduce the systematic behavior revealed by this analysis.

The main question we want to tackle here is: Which contour integration model describes human behavior best? In particular we want to know the advantages and disadvantages of a nonlinear coupling of different inputs to a neuron as suggested by probabilistic models compared to a purely linear coupling of inputs as employed in many neuronal network models. Furthermore we want to test

the influence of the geometry, range, and symmetry of the association field on the model output. Experimental findings in monkeys suggest isotropic connections between columns of similar orientation preference, spreading symmetrically into all directions from an orientation column. In contrast, probabilistic models require unidirectional interactions, linking orientation columns in only one direction, in order to get optimal contour detection performance. Hence we will analyze the advantages and disadvantages of symmetric and unidirectional interactions and answer the question, which symmetry of interaction structure is more suitable to reproduce contour integration in the brain.

From a probabilistic point of view the association field can be regarded as an edge link probability density which gives the probability for each pair of edge elements to be linked into a contour. Hence an association field can be used to generate mathematically well-defined contour ensembles. This attempt of an exact definition of what humans regard as a contour enables a mathematically exact formulation of the problem, which is missing in most studies trying to explain of contour integration in the brain. This mathematical formulation is the basis for finding a probabilistic contour integration model which makes use of the information contained in the generating association field in order to infer contours from a given stimulus and hence leads to an optimal performance. Given the speed and efficiency of human contour detection an efficient mechanism for contour integration is required. Hence it seems promising to start with a probabilistic and most effective model and add biological constraints in order to approximate human behavior.

For the psychophysical experiments we consequently used stimuli generated from an association field. Therefore using the generating association field to extract contours from the stimulus results in maximum performance. However, the brain is most probably not optimized for these artificial stimuli, but rather for visual scenes important for everyday life. It has been shown, that models employing natural image statistics to obtain a probability for linking edge elements perform similarly well as humans (33). In this thesis we show that using the generating association field for contour detection leads to a very high performance but does not explain the correlations between responses of human observers.

In search for a contour integration model capable of explaining these correlations, we proceeded by including different biological constraints. I.e. the area of cortical surface representing a given area in visual space decreases with eccentricity, leading to decreased contour detection performance for contours in the periphery.

In order to answer the question, which contour integration model explains human contour detection best, we test different ways to couple the inputs of a neuron, different geometries, ranges and symmetries of association fields and different functions of the afferent input.

In a first step we want to identify special properties, advantages and disadvantages of the different model choices. Therefore we used very simplified stimuli which also allow for analytic calculations.

Often illusions or 'mistakes' of visual perception reveal more about the underlying mechanisms than tasks which can be solved correctly. Hence a model explaining contour integration in the brain should be sensitive to the same illusions as humans or monkeys. One example for this are Kanizsa figures where a contour is filled in between two inducing figures and observers seem to perceive a clear boundary where there is no physical boundary present. This illusion can be reproduced by models employing the bipole property (41). Analyzing simplified stimuli helped us in identifying overestimation of the salience of radially conjoining edge elements as an illusion to which some model choices are sensitive while others are not. This might help to rule out particular models when comparing these model predictions to experiments using stimuli tailored to these illusions.

After having singled out several properties of the analyzed model dynamics and input functions, we compare the outcome of numerical simulations using the same stimuli as employed in human psychophysics to human behavior. This way we can evaluate how closely the model predictions resemble human behavior. As several of the analyzed models reach human performance, we search for further methods to evaluate these models. We show that responses of different human observers are not made randomly but are highly correlated among different subjects. Therefore we do not only take into account the average performance but also analyze the correlations between model predictions and human decisions for each individual trial. With this we can formulate further benchmarks and rule out more of the analyzed models. This model evaluation suggests that non-linear coupling of different neuronal inputs and a unidirectional association field should be involved in contour integration in the brain.

1.1 Outline of the dissertation

The dissertation contains a part about previous work which is organized as follows: After the introduction in chapter 1 we describe the biological background which is relevant for contour integration in chapter 2, introducing the visual system and mechanisms for multiplicative gain modulation. In chapter 3 we review experimental findings, especially in psychophysics and existing contour integration models. Here we consider biologically inspired models as well as probabilistic models. A representative model for each of these two approaches will be described in some detail.

The part of this thesis that deals with my own research starts with chapter 4 explaining in detail the three models which will be investigated in this thesis:

Chapter 1. Introduction

a probabilistic exact model, a probabilistically motivated iterative model with multiplicative couplings and a standard neuronal network model with additive couplings. In chapter 5 we discuss advantages and disadvantages of the different model dynamics on the example of very simple stimuli which allow analytic computations. In chapter 6 the influence of noise and orientation jitter on a probabilistic model is investigated. In chapter 7 the psychophysical experiments are presented and several benchmarks for evaluating the models are discussed. Following we will show simulation results for different model choices and analyze their predictive power with respect to human behavior. In chapter 8 we summarize the achievement of this study and discuss open questions.

Chapter 2

Biological background

In this chapter we will introduce the biological background that is necessary to understand neuronal contour integration models. Neurons are cells which receive information, process it and transmit it to other neurons. This is done via action potentials. Connected to large networks neurons can communicate information over long distances. In section 2.1 we will describe the physiology of a neuron and the information transmission between neurons. In section 2.2 we will extend this to networks of multiple neurons. The brain has a hierarchical structures with several areas which are concerned with vision. In section 2.3 the visual system and the most important regions for processing visual input will be introduced. Furthermore multiplicative gain modulation will be explained in section 2.2.2. This is a biological plausible mechanism for implementing multiplication, which we need for one of the models introduced in chapter 4.

2.1 Single neurons

A neuron consists of a cell body, also called soma, several branches of dendrites, which receive signals from other neurons, and an elongated axon which sends signals to other neurons. At its end the axon is branched and each branch ends in a presynaptic terminal involved in the release of chemical neurotransmitters which carry the signal to the dendrites of an adjacent neuron.

Neurons contain molecules and ions which can carry charges and usually the negative charges prevail inside the neuron. The cell body of a neuron is covered by a cell membrane which is essentially impermeable to most charged molecules. The negatively charged ions inside the neuron repel each other and consequently distribute at the inner membrane surface. These charges attract an equal density of extracellular, positively charged ions which gather at the outside surface of the membrane. In this way the separating cell membrane functions as a capacitor.

Chapter 2. Biological background

Within this membrane there are different types of ion channels which are typically selective to a specific type of ions which can permeate the membrane through these channels. The number and type of the embedded ion channels determines the resulting membrane conductance. Furthermore ion pumps in the membrane spend energy to maintain a certain difference of ion concentration between the interior and exterior of the cell. By convention the potential outside a neuron is set to zero. The resting potential inside a neuron is usually around -65 mV. This negative membrane potential attracts positive ions and repels negative ions. At the same time the different concentrations of ions inside and outside the cell, which are maintained by ion pumps lead to diffusion, i.e. Na^+ and Ca^+ ions tend to diffuse inside the neuron while K^+ ions are driven out of the cell. An equilibrium potential is reached when the diffusive processes due to different ion concentrations and the ion current due to electric forces are balanced. The equilibrium potential E for an ion channel, that is selective to a single ion type, can be calculated from the ion concentration inside and outside the cell, denoted by `[inside]` and `[outside]`, with the help of the Nernst equation

$$E = \frac{k_B T}{qz} \ln \left(\frac{[\text{outside}]}{[\text{inside}]} \right) \quad (2.1)$$

Here zq is the charge of the ion where q denotes the charge of a proton. $k_B T$ is the thermal energy of an ion at a temperature T and k_B is the Boltzman constant.

The membrane current I_m of a neuron is the sum of all currents flowing through any of its ion channels. For a better comparison between neurons of different sizes one usually uses i_m , the membrane current per unit area. We label each type of ion channel by an index i . Then the conductance per unit area due to this channels is g_i which is also called specific conductance. The specific membrane current depends on the conductance g_i for each type of channel and on its driving force, which is given by the difference between the membrane potential V and the equilibrium potential E_i of that channel: $i_m = \sum_i g_i (V - E_i)$. The membrane conductances g_i can change over time. For example there are voltage dependent channels whose conductance changes with the membrane potential or channels whose conductance varies with the concentration of specific interior messengers like it is the case for Ca^+ -dependent channels. However, there are also approximately constant components of the membrane current, like for instance the currents carried by ion pumps, which are grouped together into the so called leakage current. This is given by $\bar{g}_L (V - E_L)$ where the bar over g_L indicates, that the conductance is approximately constant.

In single compartment models the membrane potential is assumed to be the same in all parts of the neuron and is accordingly described by a single variable V . One can account for the spatial geometry of a neuron and the varying properties of different parts of the neuron by a multi-compartment model. An introduction into multi-compartment models can be found in (19), chapter 6.

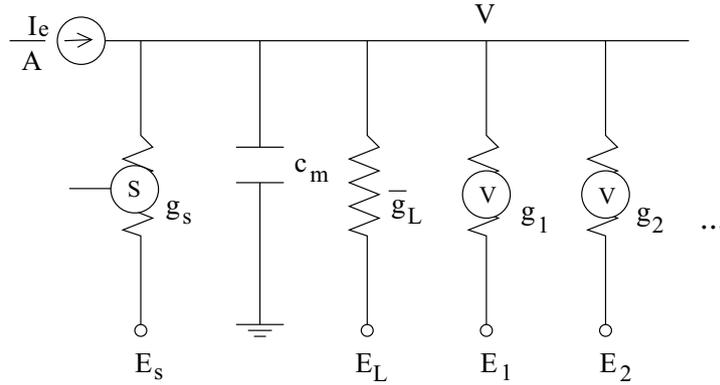


Figure 2.1: Equivalent electric circuit for a one-compartment neuron with surface A , a synapse and a current I_e injected by an electrode. Only a single synaptic conductance g_s is indicated but in principle there can be several types of synapses. g_1 and g_2 are voltage dependent conductances.

In single compartment models the membrane potential changes with the total current entering the neuron, comprising all currents due to membrane or synaptic conductances as well as any current I_e injected to the cell by an electrode. Hence the basic equation underlying all single-compartment models is

$$c_m \frac{dV}{dt} = i_m + \frac{I_e}{A} \quad (2.2)$$

Here c_m is the specific membrane capacitance, meaning the capacitance per unit area, and A is the total surface area of the neuron. The structure of the model is equivalent to an electric circuit consisting of a capacitor and various resistors that correspond to the different membrane conductances (Fig. 2.1).

The interesting dynamics of the membrane potential including the generation of action potentials comes primarily from the nonlinearities of the membrane conductances. Ion channels stochastically fluctuate between open and closed states. Due to the large number of ion channels for a specific ion, the fraction of open channels is to a good approximation the probability P_i for a single channel to be in an open state. By multiplying the density of channels and the conductance of an open channel we obtain the maximal conductance \bar{g}_i per unit area. $g_i = P_i \bar{g}_i$ gives then the actual specific conductance.

A formalism to describe voltage-dependent conductances was developed by Hodgkin and Huxley (50). We distinguish between persistent conductances like the delayed rectifier K^+ conductance and transient conductances like the fast Na^+ conductance. Persistent conductances have a swinging gate which is sensitive to voltage. The channel is activated, if the gate is open and deactivated, if the gate is closed. The opening probability increases with the depolarization of the

neuron. Because of the delayed rectifier K^+ conductance the neuron repolarizes after the action potential. This conductance contains $k = 4$ identical subunits, which must all be open for the channel to open. Hence $P_K = n^k$ where $0 < n < 1$ is the so called gating variable that describes the probability that a particular subunit gate is open. The probability that a subunit gate opens depends on the voltage and is given by the rate $\alpha_n(V)$. The voltage dependent closing rate of a subunit gate is $\beta_n(V)$. The probability that a subunit is open increases when a closed subunit opens and decreases when an open subunit closes. Hence $\frac{dn}{dt} = \alpha_n(V)(1 - n) - \beta_n(V)n$. This gating equation can be rewritten as

$$\tau_n(V) \frac{dn}{dt} = n_\infty(V) - n \quad (2.3)$$

This shows that the limiting value $n_\infty = \frac{\alpha_n(V)}{\alpha_n(V) + \beta_n(V)}$ is exponentially approached with the time constant $\tau_n = (\alpha_n(V) + \beta_n(V))^{-1}$. $\alpha_n(V)$ and $\beta_n(V)$ are obtained from fitting experimental data. α_n increases with V and β_n decreases with V .

Transient conductances are controlled by two gating mechanisms with opposite voltage dependence. The first gate works like in the case described above. For a fast Na^+ conductance we have an activation variable m and the channel consists of $k = 3$ subunits. In addition the inactivation variable h with $0 < h < 1$ is the probability that the second gate is not blocked. The probability that both gates are open and hence the ion channel can conduct is given by $P_{Na} = m^3 h$. This is the open probability for a transient conductance. The Hodgkin Huxley model (50) describes the generation of action potentials by assuming Eq. 2.2, the gating equation Eq. 2.3 and a membrane current consisting of a leakage current, a fast Na^+ current and delayed-rectified K^+ current $i_m = \bar{g}_L(V - E_L) + \bar{g}_{Na} m^3 h (V - E_{Na}) + \bar{g}_K n^4 (V - E_K)$. Fig. 2.2 shows the membrane potential and the state of the ion channels during an action potential as described by the Hodgkin Huxley model. The Na^+ concentration is higher outside the cell. When the cell reaches a certain depolarization, the Na^+ channels open and Na^+ flows inside the cell. This causes further depolarization and with some delay this leads to the opening of the K^+ channels. As the K^+ concentration is higher inside the cell, the ions flowing out of the cell slow down the depolarization due to the Na^+ ions. At a membrane potential of about 35mV the Na^+ channels are inactivated and the cell is repolarized to the resting potential due to the K^+ ions which leave the cell. Finally the K^+ channels are deactivated and all gating variables approach their initial conditions.

The Connor Stevens model (13; 14) differs from the Hodgkin Huxley model by faster kinetics due to slightly different parameters and an additional A-current from a transient K^+ conductance. It has the specific membrane current $i_m = \bar{g}_L(V - E_L) + \bar{g}_{Na} m^3 h (V - E_{Na}) + \bar{g}_K n^4 (V - E_K) + \bar{g}_A a^3 b (V - E_A)$. The main effect of the A-current is a continuous rise of the firing rate as a function of the

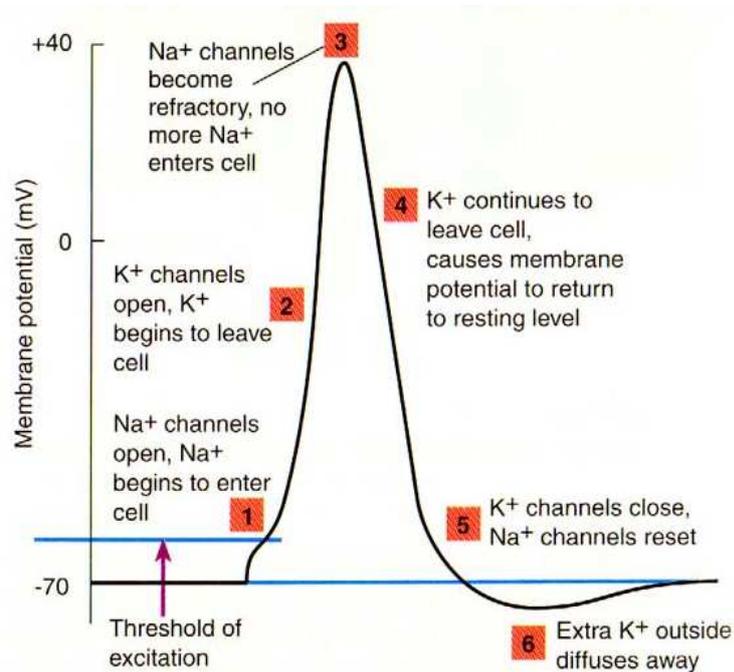


Figure 2.2: Membrane potential and state of ion channels during the firing of an action potential (modified from (1)).

electrode current injected to the neuron, above some threshold θ_E . Without the A-current the firing rate would jump to a non-zero value at θ_E instead of rising continuously.

Synaptic conductances

Neurons can communicate with each other via chemical synapses. The neuron sending the signal is called the presynaptic neuron and the postsynaptic cell receives the signal. An action potential in the presynaptic neuron propagates along the axon and activates voltage dependent Ca^{2-} channels in the presynaptic terminal causing a rise in the Ca^{2-} ion concentration. Due to the higher Ca^{2-} concentration vesicles that contain neurotransmitter molecules fuse with the cell membrane and release the transmitter into the synaptic cleft between the two neurons. The transmitter molecules diffuse through the cleft and bind to the postsynaptic neuron, which causes the ion channels to open and, subsequently, the conductance of the postsynaptic neuron to change.

2.2 Neuronal networks

In 2.1 we described how an isolated neuron can generate an action potential and how an action potential can be transferred from one neuron to another via chemical synapses. Now we want to look at the interaction of many neurons within neural networks. There are three types of interactions: Feedforward interactions go from earlier areas in the brain, concerned with earlier processing stages to higher brain region, responsible for later processing stages. Recurrent interactions are interactions within one area and between neurons at the same information processing stage. Feedback or top-down interactions go from higher areas back to regions from an earlier processing stage. As the models presented in this thesis only contain feedforward and recurrent connections, I will omit the top-down connections in this section about neuronal networks. An introduction can be found in chapter 10 of (19).

2.2.1 Firing-rate models

In principle one could build networks from the Hodgkin-Huxley neurons described in 2.1. However, depending on the size and complexity of the network this is a serious computational challenge. It is much easier to study networks that are constructed of neuronal populations whose output is a firing rate rather than a sequence of action potentials. For firing-rate models it is also possible to investigate certain aspects of the network dynamics analytically which we could not do for spiking networks. A drawback of firing rate models is that these models assume that the spikes are uncorrelated and hence cannot account for aspects of precise spike timing and synchronous firing.

Let us assume a neuron with N_u synaptic inputs labeled by $\mathbf{b} = 1, 2, \dots, N_u$ with input firing rates denoted by \mathbf{u}_b . How does the neuron's firing rate I_s depend on the presynaptic firing rates? An action potential at input \mathbf{b} generates the synaptic current $w_b K_s(t)$ in the soma of the postsynaptic neuron. Here w_b is the synaptic weight, which is positive for excitatory synapses and negative for inhibitory synapses. $K_s(t)$ is the synaptic kernel and describes the time course of the synaptic current which is evoked by a presynaptic spike that arrives at time 0. For simplicity we assume the same K_s for each synapse of a neuron and normalize $\sum_{t>0} K_s(t) = 1$.

A presynaptic spike train at input \mathbf{b} with action potentials arriving at the times t_i can be described by a sequence of δ - functions $\rho_b(\tau) = \sum_i \delta(\tau - t_i)$, also called the neural response function. The effects of action potentials at a single synapse \mathbf{b} sum linearly. Hence the synaptic current from neuron \mathbf{b} in

response to this presynaptic spike train is given by

$$I_b(t) = w_b \int_{-\infty}^t d\tau K_s(t - \tau) \rho_b(\tau)$$

The total synaptic current is then obtained by summing over different inputs. In order to finally obtain I_s in dependence of the firing rate we replace the neural response functions by the input firing rates u_b . Then we get

$$I_s = \sum_{b=1}^{N_u} w_b \int_{-\infty}^t d\tau K_s(t - \tau) u_b(t)$$

When assuming the frequently used exponential kernel $K_s(t) = \exp(-t/\tau_s)/\tau_s$ and taking the derivative of I_s with respect to t we get the differential equation

$$\tau_s \frac{dI_s}{dt} = -I_s + \mathbf{w} \cdot \mathbf{u} \quad (2.4)$$

where the input rates u_b and the weights w_b have been combined in the vectors \mathbf{u} and \mathbf{w} , respectively.

How do we now get the postsynaptic firing rate v as a function of I_s ? Here we can make use of the activation function $F(I_s)$ as computed from a specific neuron model or obtained from an experiment in which a cell was subjected to different test currents. Let us assume $v = F(I_s(t))$ where the activation function $F(I_s)$ is typically a sigmoidal function which is bounded from above and thus stabilizes recurrent networks against excessively high firing rates. An alternative choice would be $F(I_s) = [I_s - \gamma]_+$ where γ is a threshold and $[\]_+$ denotes a half-wave rectification. The activation function together with Eq. 2.4 realizes a firing-rate model.

An alternative firing-rate model assumes that the firing rate does not change with I_s instantaneously. Because of the membrane capacitance and resistance the sum of the postsynaptic action potentials is approximately a low-pass filtered version of the synaptic current. Hence we can write the time-dependent firing rate as a low-pass filtered version of the steady state firing rate

$$\tau_r \frac{dv}{dt} = -v + F(I_s(t)) \quad (2.5)$$

The time constant τ_r describes how rapidly the firing rate approaches its steady state value for constant I_s . This second model consist of Eq. 2.4 and Eq. 2.5. However, for $\tau_r \ll \tau_s$ Eq. 2.5 rapidly approaches $v = F(I_s(t))$ and the model equals the first model description. If vice versa $\tau_r \gg \tau_s$ the synaptic current comes quickly to its steady state $I_s = \mathbf{w} \cdot \mathbf{u}$ and it holds

$$\tau_r \frac{dv}{dt} = -v + F(\mathbf{w} \cdot \mathbf{u}) \quad (2.6)$$

Now let us now assume a feedforward network, which has the input rates \mathbf{u}_b ($b = 1, 2, \dots, N_v$) and the output rates v_a ($a = 1, 2, \dots, N_v$). This can be described by the differential equation

$$\tau_r \frac{dv}{dt} = -v + F(\mathbf{W} \cdot \mathbf{u}) \quad (2.7)$$

In contrast to Eq. 2.6 here the weights are given by a matrix \mathbf{W} whose components W_{ab} represent the mean strength of the synapses between output unit a and input unit b . Furthermore we can introduce recurrent connections into the network described above by including an additional weight matrix \mathbf{M} with components $M_{aa'}$ describing the synaptic strength between output unit a' and output unit a . The output rates can then be described by

$$\tau_r \frac{dv}{dt} = -v + F(\mathbf{W} \cdot \mathbf{u} + \mathbf{M} \cdot \mathbf{v}) \quad (2.8)$$

According to Dale's law a neuron can be either excitatory or inhibitory. An excitatory neuron can only excite postsynaptic neurons and an inhibitory neuron inhibits all its postsynaptic neurons. Hence for a given presynaptic neuron α the components of \mathbf{M} should have the same sign for all postsynaptic neurons α' . Hence we can distinguish between the output rates \mathbf{v}_E and \mathbf{v}_I of excitatory neurons and inhibitory neurons, respectively. We end up with a set of coupled differential equations for the two types of neurons.

$$\tau_E \frac{dv_E}{dt} = -v_E + F_E(\mathbf{h}_E + \mathbf{M}_{EE} \cdot \mathbf{v}_E + \mathbf{M}_{EI} \cdot \mathbf{v}_I) \quad (2.9)$$

and

$$\tau_I \frac{dv_I}{dt} = -v_I + F_I(\mathbf{h}_I + \mathbf{M}_{IE} \cdot \mathbf{v}_E + \mathbf{M}_{II} \cdot \mathbf{v}_I) \quad (2.10)$$

Here $\mathbf{h} = \mathbf{W} \cdot \mathbf{u}$ is the total feedforward input to each neuron. The components of \mathbf{M}_{EE} and \mathbf{M}_{IE} are positive or zero and analogously the elements of \mathbf{M}_{EI} and \mathbf{M}_{II} are smaller or equal to zero. This set of coupled differential equations gives the possibility to choose different time constants, activation functions and feedforward inputs for inhibitory and excitatory neurons.

2.2.2 Multiplicative gain modulation

As we will see later in more detail, multiplication of neuronally represented variables is a desirable operation for general computations. In a neurophysiological context this can be implemented by multiplicative gain modulation. The term 'multiplicative gain modulation' describes an approximately multiplicative or divisive change of the amplitude of the stimulus-response function of a neuron without modification of the response selectivity. This is equivalent to a change of

the slope of the firing rate plotted against synaptic input current without shifting the threshold.

Earlier models (11; 89) proposed that multiplicative gain modulation could be obtained by changing the input conductance by so called shunting or silent inhibition, an inhibitory input with a reversal potential close or equal to the resting potential of the cell, which will affect the membrane potential of the cell in a divisive manner. It was suggested that when the inhibitory input rate increases, the slope of the firing rate in dependence of the synaptic input current decreases, while the threshold stays approximately constant.

However, if the reversal potential of the inhibition is far from the spiking threshold, the inhibitory synapse rather acts like a current source. The cell's conductance is not changed much by the inhibitory current I_{inh} and the input current I_{syn} is shifted to $I_{\text{syn}} - I_{\text{inh}}$.

In contrast to (11; 89) Holt and Koch (51) as well as Salinas and Abbott (110) found that shunting which increases the total membrane conductance has a subtractive effect and cannot explain multiplicative gain modulation. Only if the shunting is accompanied by an increase of the input noise that the cell receives, it has an approximately multiplicative effect. This can be the case if a neuron receives significant amounts of balanced excitatory and inhibitory input (110) as was shown in the neocortical pyramidal neurons (12).

There is plenty of experimental evidence for multiplicative gain modulation in many cortical areas. For example in the primary visual areas the gain of orientation or direction tuning curves increases with attention (85; 128). Gain modulation has also been proposed to be a neural mechanism underlying coordinate transformations, which are relevant for visually guided reaching (139; 111; 110; 100; 101) and necessary for invariant object recognition (112).

2.3 The visual system

The processing of visual information in the brain happens in parallel as well as hierarchically. Visual information is split into its components like color, orientation, motion, texture, shape and depth. These features are processed in parallel in separate areas of the visual cortex specialized for these particular features. It is assumed that afterwards the information is reassembled into a single coherent percept of the visual scene in higher visual areas. In this section we will give an overview over the brain areas concerned with the processing of visual information. Here our main focus will be on the areas processing orientation, as this is the feature being crucial for the perception of contours.

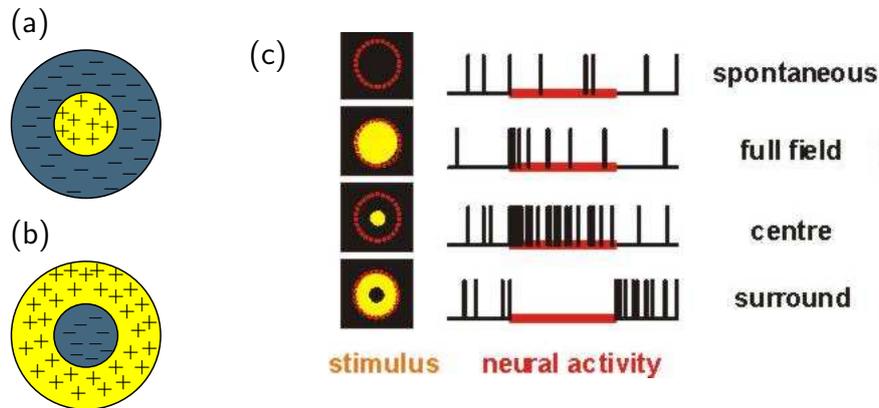


Figure 2.3: Retinal ganglion cells with on-center-off-surround receptive fields (a) respond maximally when their center is exposed to light while their surround lies in darkness. Vice versa off-center-on surround cells (b) respond maximally if only their surround is exposed to light. (c) shows the neural response of an on-center-off surround cell for different stimuli.

2.3.1 The retina and LGN

The retina

Processing of visual information starts in the retina, which is already part of the central nervous system. The retina contains two types of photoreceptors: rods and cones. While cones are important for day vision, rods are responsible for night vision and can be excited by even one single photon (43). These photoreceptors absorb light and transduce it into electrical signals which are transferred to the retinal ganglion cells via bipolar cells. Ganglion cells can have two different types of receptive fields. Cells with on-center receptive fields are maximally excited if the center region of the receptive field is stimulated by light while the region surrounding the center lies in the dark. Light in the surround has an inhibitory effect on these cells. Vice versa ganglion cells with off-center receptive fields are inhibited due to light in the center of the receptive field and excited by light in the region surrounding the center (see Fig. 2.3). The axons of the retinal ganglion cells of each eye come together at the optic disc where they are bundled into the optic nerve (see Fig. 2.4). In the optic chiasm the two optic nerves join. The fibres from the temporal hemiretina continue on the same side of the brain. However, the axons from ganglion cells from the nasal hemiretina cross to the opposite side of the brain, such that the right optic tract contains only information from the left visual hemifield, while the left optic tract carries only information from the right hemifield. The optic tracts project to the pretectum, the superior colliculus and the lateral geniculate nucleus (LGN).

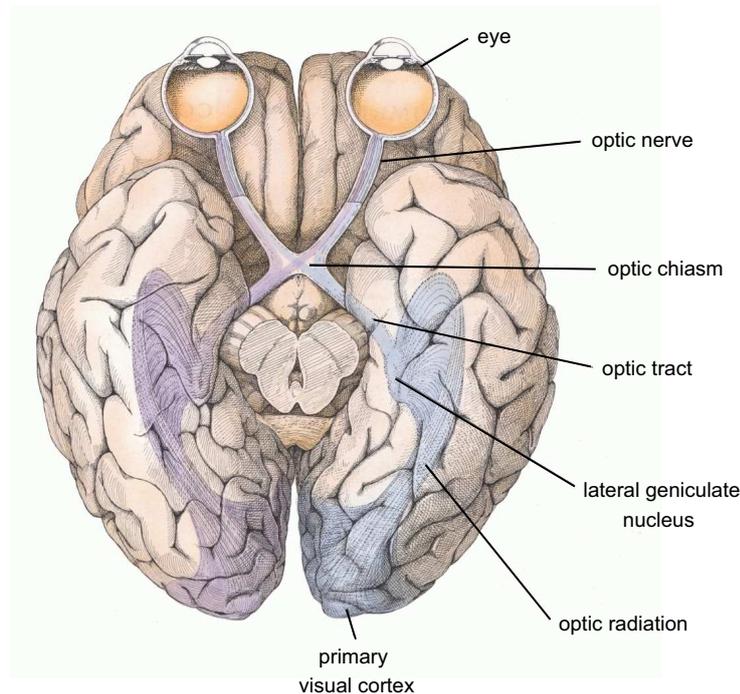


Figure 2.4: Visual pathway from the retina to the primary visual cortex in the human brain, Adapted from: (54)

The lateral geniculate nucleus

The fibres of the optic tracts terminate topographically organized in the left and right LGNs. Each LGN contains a retinotopic representation of the contralateral visual field, meaning that the right hemifield is represented in the left LGN and vice versa. In the neural map emerging from the topological structure the periphery has a proportionally much smaller representation than the fovea, which is the retinal area with the highest acuity.

The human LGN has six layers of cell bodies which each get input from only one eye. The two most ventral layers are called magnocellular layers. They are named after the relatively large cells they consist of. They mainly receive input from the M-cells of the retina. The four dorsal layers are termed parvocellular layers which get input from the retinal P-cells respectively. Like the retina the LGN has concentric receptive fields with on-center and off-surround structure or vice versa with off-center and on-surround organization. The retinal P-cells and M-cells form the origin of separate pathways, which are subdivided into a magnocellular and a parvocellular pathway. The magnocellular pathway is responsible for the detection of motion and the location of an object while the parvocellular pathway processes information about color.

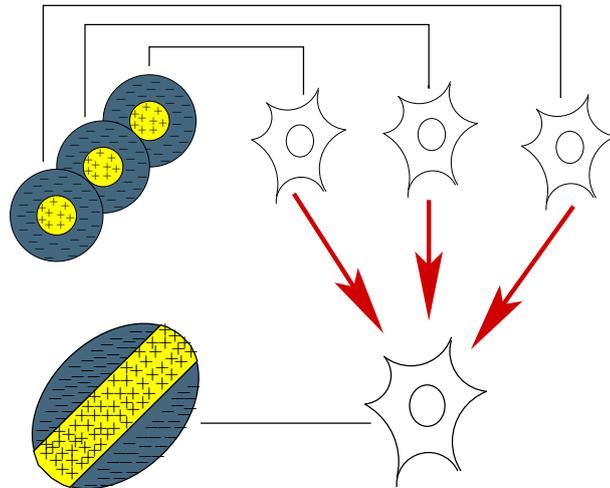


Figure 2.5: Simple scheme proposed by Hubel and Wiesel (52) showing how cells with circular receptive fields from LGN are combined to form orientation selective simple cells in V1.

2.3.2 The organization of the primary visual cortex

The primary visual cortex (V1) consists of six layers with different cell and fibre density. Each cerebral hemisphere receives only direct visual information from the contralateral visual hemifield. The fourth layer is again divided into four sublayers 4A, 4B, 4C α and 4C β . Layer 4C α belongs to the magnocellular pathway and receives mainly input from the M-cells of the LGN. The layer 4C β belongs to the parvocellular pathway. Axons of the P-cells of the LGN terminate in layer 4C β . Like the cells in the LGN the spiny stellate cells in 4C β have concentric receptive fields. In this layer one finds so-called simple cells which are pyramidal neurons with elongated rectilinear receptive fields being larger than the concentric receptive fields in the LGN and in the spiny stellate cells. These cells respond maximally to a bar of light that matches their orientation of the receptive fields (see Fig. 2.5).

For an optimal response this bar must have a specific position within the receptive field, i.e. it must lie in the excitatory region of the receptive field. If the bar is not well positioned or if its orientation deviates from the preferred orientation of the cell, the response will not be maximal because the excitatory region is not completely filled and because the bar partly extends into the inhibitory region. Cortical cells which receive input from the same location on the retina can have the same spacial shape of receptive fields but with different preferred orientations, such that all orientations are represented for each point of the retina. The receptive field properties of simple cells could be generated by a certain topographical arrangement of cells with concentric receptive fields.

They could arise from stellate cells with similar center-surround receptive fields with slightly displaced positions on the retina which are aligned such that their centers taken together represent a certain axis of orientation within the visual field.

A second class of receptive fields, named complex receptive fields, can be found for pyramidal cells in layer 2, 3, 5 and 6. Like the simple cells they have rectilinear receptive fields of a certain orientation. However, in contrast to the simple cells they have no clearly defined on and off zones. Hence the response of the cell is independent of the position of the bar. There are certain complex cells which respond to stimuli that move in a certain direction through the receptive field. Hubel and Wiesel (52) proposed that these movement-sensitive cells receive input from simple cells with the same axis of preferred orientation and receptive fields positioned along the preferred direction of movement.

Organization of V1 into blobs and columns

The neurons in V1, which receive input from the parvocellular and the magnocellular pathway, are organized into orientation columns. Neurons within an orientation column all respond optimally to the same preferred orientation and their receptive fields are located at nearly the same retinal position, even though they might differ in size. The topology of orientation columns on the surface of the cortex can be visualized by using the optical imaging technique. Full-field gratings of varying orientations are presented, and active cells are then detected using voltage-sensitive dyes and a highly sensitive camera. This technique shows that orientation columns are arranged around orientation singularities which are termed pinwheels. Around each pinwheel every single orientation from 0 to 180 degrees is represented. The axes of similar preferred orientations are organized radially around the pinwheel. Orientation maps have specific length scales which differ among species. In addition to the structure of orientation columns the striate cortex is also organized into other feature maps as for example ocular dominance columns, spatial frequency, columns sensitive to the color of a stimulus, and others. The structures of these maps are not independent of each other, but show certain statistical correlations. In analyzing the processing of visual information it is useful to introduce the concept of a hypercolumn, which contains cells sensitive to all combinations of features on a localized region in retinal space. When focusing on orientations as a feature of great importance for contour integration, this means that a hypercolumn contains orientation-sensitive cells for all preferred orientations representing edge elements at approximately the same position in the visual scene.

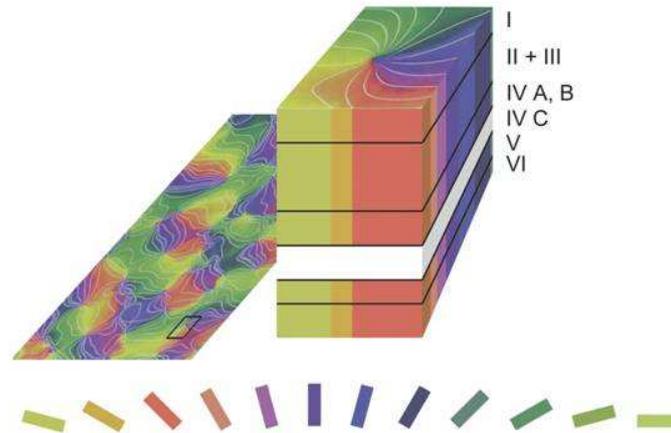


Figure 2.6: On the left an orientation preference map from visual cortex is shown. Each color represents a certain preferred orientation. On the right we see a close-up of this map. Here all preferred orientations are arranged around a pinwheel. Perpendicular to the surface of cortex, orientation preference is similar throughout all layers of V1 with the exception of layer 4C.

Horizontal connections

The three constituents of the hypercolumns, namely the orientation columns, ocular dominance columns, and blobs, extend perpendicular to the cortex surface and cross the layers of V1. Between these vertically oriented structures there are horizontal connections which link cells within one cortical layer.

Anatomical and neurophysiological studies in tree shrews (105), cats (36) and monkeys (114) reveal that long range, lateral connections which link cells with non-overlapping receptive fields, project primarily to cells of similar orientation preference. Malach et al. (82) and Bosking et al. (8) used optical imaging to determine the orientation preferences of the orientation columns across the visual cortex. They combined this method with inducing biochemical tracers to track where the horizontal axons of a particular cell project. The result is depicted in Fig. 2.8, which shows a map of the visual cortex, where the preferred orientation of the cells is color-coded. The white dots mark the cells where the tracer was injected. The orientation preference of these cells is also shown in the upper edge of the pictures. The black dots mark the terminals of the axons of those cells. It can be seen that long-range connections mainly project to orientation columns with similar preferred orientation. Bosking also showed that axons along the axis of the receptive field extend for greater distances than axons orthogonal to the orientation of the receptive field.

While long-range horizontal connections link neurons with similar preferred

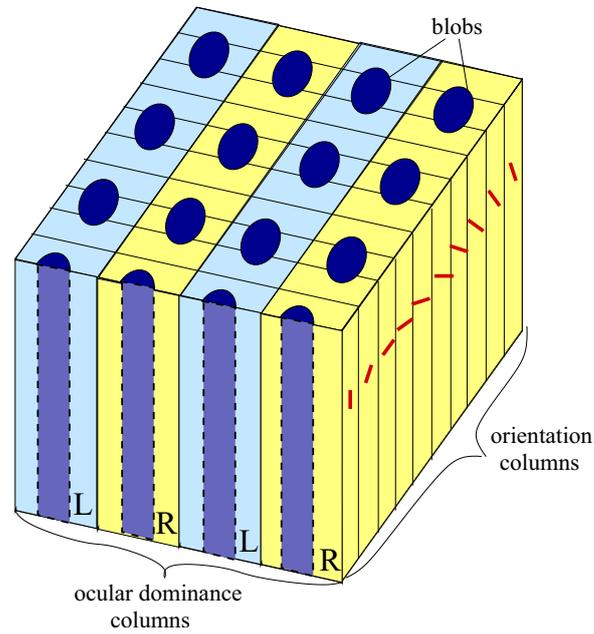


Figure 2.7: A hypercolumn is an idealized scheme of a basic cortical building block, comprising ocular dominance columns for both eyes, orientation columns, and blobs which are concerned with processing color.

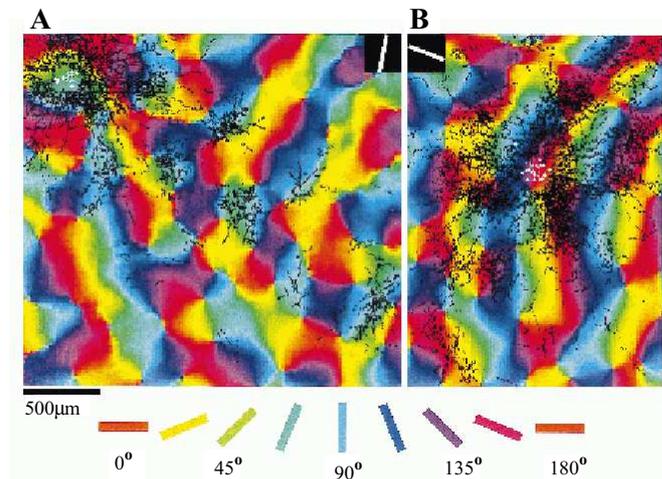


Figure 2.8: Color coded orientation map within primary visual cortex of the tree shrew. The lateral connections are made visible by induced biochemical tracers. The injection points are marked in white. The termination of the axons are shown in black. Source: (8).

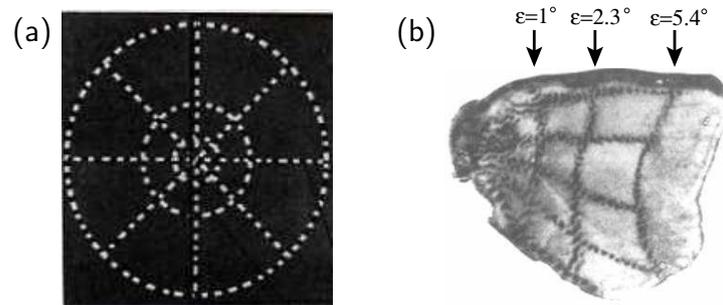


Figure 2.9: (a) Stimulus shown to a macaque monkey. (b) Radioactive trace of the activity evoked by the image shown in (a). The vertical lines correspond to concentric circles in the stimulus at an eccentricity of 1° , 2.3° and 5.4° (126).

orientations and hence are thought to mediate contour integration, short-range connections appear to link cells with different preferred orientations and have been suggested to mediate detection of border discontinuities (17; 18).

Cortical magnification factors

Visual acuity is far better in the fovea than in the periphery. The reason for this lies in the structure of the retina. The visual field is mapped topographically on the surface of the striate cortex. However, the projection has a certain size for the central visual field and is gradually compressed towards the periphery (87; 116). This compression can be measured by a cortical magnification factor (16), which is defined as the distance of cortical surface to which one degree of visual angle is projected at a given eccentricity. Its reciprocal increases linearly from about 0.15 degrees per mm at the fovea to 1.5 degrees per mm at an eccentricity of 20 degrees. The resulting retinotopic map is illustrated in Fig. 2.9. The pattern seen in Fig. 2.9 (b) was obtained by imaging a radioactive analogue of glucose in macaque monkeys (126). It shows the neurons which were active while the monkey observed the pattern shown in Fig. 2.9 (a).

2.3.3 Non-linear response properties within the primary visual cortex

Center-surround modulations of spiking neuron responses in V1

The classical receptive field of a neuron is defined as the region where the presentation or the distinction of light evokes a response, like described above. However, this is an incomplete description of the area that influences a neuron. Even

though stimuli in the surround of a classical receptive field are not sufficient for driving spiking responses, they can modulate the neuronal response evoked by a stimuli inside the classical receptive field (29; 31). These modulations are termed center/surround modulations and provide a comparison of stimuli within the receptive field and outside the receptive field. In V1 the surround can cover up to 10 times the size of the classical receptive field (73). Some neurons in the primary visual cortex have a stronger response to a single bar that is flashed in their receptive field, if this bar is embedded in a field of orthogonally oriented bars in compared to a bar surrounded by edges of the same orientation (63). The bar surrounded by orthogonal edges is perceived as being much more salient than the bar under the other condition and can be detected without paying attention. One says it "pops out" (127). Another hint for the preattentiveness of these processes is that such orientation-selective modulations by the surround can also be measured in anesthetized animals (74; 91; 115; 119).

End-stopped cells

As long as a bar of light is within the excitatory zone of the receptive field of a simple or a complex cell, the response of that cell increases with the length of that bar. If the receptive field consists of longitudinal on- and off- zones, extending the length of the bar beyond the borders of the receptive field has no effect on the response. However, so called end-inhibited or end-stopped cells have an additional inhibitory region on top and bottom of the excitatory region. Hence lengthening the stimulus bar in either direction beyond the excitatory region leads to attenuation and abolition of the cell's response. Cells termed single end-stopped cells have this inhibitory region either on top or on bottom of the excitatory region. Thus this limitation occurs only on one side. End-stopped cells are assumed to be necessary for measuring the length of a line. Furthermore they could be important for detecting corners, curvature and breaks in a line. For these signals not only the orientation but also the length of a bar is relevant.

2.3.4 Visual pathways and higher cortical areas

The information processing which is necessary to perceive an object or a person clearly goes beyond the primary visual cortex with its simple and complex cells. Information relevant for form perception is processed in the parvocellular interblob system while color information is processed in the parvocellular blob system. The parvocellular pathway starts in the parvocellular cells in the retina, and continues through LGN and primary visual cortex via V2 to V4. Similar as in V1, there are cells in V2 which are sensitive to orientation, to color, or to horizontal disparity. The orientation-selective neurons have larger receptive fields as the neurons in V1. Hence a given neuron in V1 receives input from cells in V2

or higher area cells whose receptive fields extend far beyond the receptive field of the target cell in V1 (108; 109). For example in macaque monkey feedback connections from V2 to V1 convey information from a region within the visual scene which is five to six times larger than an average receptive field in V1 (5). Hence feedback connections from V2 or higher areas are another good candidate for the neuronal substrate of center-surround modulation. However, when inactivating V2 Hupé and colleagues (55) found no effect on the center-surround modulation in V1 in the case of a bar surrounded by a static texture.

Another property of V2 is that in contrast to the cells in V1 many cells in V2 also respond to illusionary contours and not only to edge elements defined by contrast.

V4 is organized retinotopically and contains neurons selective to form and color. From there signals are projected to the inferior temporal cortex (IT), a region which is relevant for the *what* rather than the *where* of the visual scene. The inferior temporal cortex is not retinotopically organized and has much larger receptive fields than the striate cortex, with an average size of $25\text{deg} \times 25\text{deg}$. Ten percent of these cells respond to specific images like, for example, a hand or a face. Lesions of the inferior temporal cortex lead to impairment in the recognition of faces, forms, and patterns without affecting more basic functions of visual perception like acuity or the perception of color or motion.

The magnocellular pathway projects from the magnocellular layers of the LGN via the layers 4C α and 4B of the primary visual cortex to the middle temporal area (MT). Another projection goes from V2 to MT, which then projects to the posterior parietal cortex. This pathway is termed dorsal pathway and it is concerned with where in the scene an object can be seen. The area MT (V5) mostly consists of motion-sensitive cells. These are direction-selective, meaning that they respond to spots or bars moving in one direction, but not to stimuli moving into the opposite direction. Like V1, the MT is retinotopically organized and cells with similar preferred direction are forming columns. For each part of the visual field there is a set of columns with different direction preferences.

When we move through the environment we perceive the motion of the entire visual field. This global motion is called optic flow, which is processed in the medial superior temporal area (MST), a cortical area adjacent to MT which contains cells that respond to optic flow.

Chapter 3

Contour integration - existing models and experimental findings

To make sense of a visual scene our visual system has to group different parts of the scene into coherent objects. During this process of perceptual grouping the Gestalt rules play an important role (64; 133). Following these rules, elements of a visual scene are grouped according to proximity, common faith, similarity, closure, and good continuation.

Receptive fields of the early visual system are localized in space. This means that independent local inputs must be combined in order to obtain information over the image structure of extended areas of the visual scene, i.e. information about contours. This spatial integration process has been investigated in many psychophysical studies.

Contour integration plays an important role in the process of image segmentation and object recognition. Contours 'pop out' of their environment if they fulfill the Gestalt rule of good continuation. Pop out means that the contour can be recognized without effort and without directing attention to it. Also the time needed for contour detection is independent of the number of distractors. In contrast, many search tasks, like for example searching for a combination of several features (conjunction search), require attention and take more time with increasing number of distractor elements.

As already mentioned in chapter 2 neurons in the early visual system have receptive fields which are tuned to a specific orientation. These receptive fields are localized (52; 53; 26). In order to perceive extended objects of a visual scene, like for example contours, the input from these localized units needs to be combined. This spatial integration process which groups single elements into contours is termed contour integration. Contour integration and image segmentation are believed to be fundamental steps in the process of object recognition.

3.1 Psychophysical studies

The quantification of human perception goes back to (133; 64) who first formulated the Gestalt rules, namely the rule of good continuation, proximity, closure, common fate, and similarity.

Proximity

The Gestalt rule of proximity states, that elements which are close to each other are more likely to be grouped than elements with a greater distance. This is demonstrated in Fig. 3.1 (a). An observer will most likely perceive the dots as being arranged in two horizontal rows rather than being arranged in three vertical columns. The Gestalt rule of proximity was investigated psychophysically with stimuli consisting of dots (see i.e. (124)) as well as with typical contour detection stimuli consisting of oriented Gabor patches.

Good Continuation

The gestalt law of good continuation states that elements which are collinearly aligned are likely to be grouped together. For example, in Fig. 3.1 (b) an observer would rather perceive two straight lines from A to C and from B to D that are crossing in point E than two curved lines (A – E – D) and (B – E – C) which meet at point E. Similarly, distinct elements like bars or Gabor patches are grouped together if they are aligned. This Gestalt rule is fundamental to the process of contour integration. Its importance for contour integration has repeatedly shown in psychophysical experiments, i.e. (27; 98; 99; 83). Review articles on contour integration have been written by Hess (46) and Kovacs (67). One of the most influential psychophysical studies was (27). For their 'path paradigm' Field and colleagues used Gabor patches whose spatial frequency and orientation properties resembled that of visual cortical cells. Contours of various curvature consisted of ten aligned Gabor elements which were embedded in a background of identical but randomly oriented Gabors. The contour path was constructed of invisible line segments joining at an angle which is uniformly distributed between $+\alpha$ and $-\alpha$, where α is the 'path angle'. This stimulus served as the target in a two-alternative-forced-choice (2AFC)-task. A stimulus consisting of the same, but randomly rotated Gabors was used as the distractor. Subjects were required to decide which of the two presentations contained the contour. Elements were spaced at a distance at which it is unlikely that a single neuron responds to more than one element.

The experiments demonstrated that the contour detection performance decreased with increasing path angle. The linking strength between two elements

depends on their orientations. This dependence is commonly termed the 'association field'. It is often assumed that the neuronal substrate of this association field are the long-ranging horizontal connections between cortical V1 neurons (35; 103).

Closure

Elder and Zucker (22) proposed that closure is crucial for the inference of shape from a contour. Their hypothesis is based on the fact that an isolated object always has a closed boundary. However, there are varying opinions whether closure as an additional Gestalt criterion enhances salience or if the enhanced saliency of closed contours is already explained by the rule of good continuation. Kovacs (68) and Braun (9) both found a closure effect. However, as in their stimuli there was a larger variability in shape for open contours than for closed ones and they did not control for the eccentricity of the contour, this could also be the reason for an enhanced saliency for the closed contours. In contrast, Tversky and colleagues (130), who tested for the impact of these effects found no evidence for an enhancement of closed contours as compared to open contours. He proposed that the closure effect can be explained solely by good continuation and probability summation. For example, in a closed contour of eight elements there are eight different possibilities to find a contour comprising eight elements. In an open contour of the same number of elements one has only one possibility to find a contour of that length. Hence a closed contour is supposed to be more salient.

Similarity

The rule of similarity states that similar elements, i.e. elements that have the same color or the same shape, are more likely to be grouped than differently looking elements. Psychophysical studies investigating the role of color or of spatial frequency as an additional feature facilitating contour integration are, e.g. (84; 95).

Common fate

The rule of common fate predicts that humans group elements which have the same fate, i.e. which move coherently into one direction or which are flashed at the same time. This seems to be the first Gestalt rule which emerges in the developing visual system. For example in contrast to adults, infants do not make use of the Gestalt rule of good continuation in order to recognize an object as a unity (123). Instead they use the common motion of parts of objects in order to

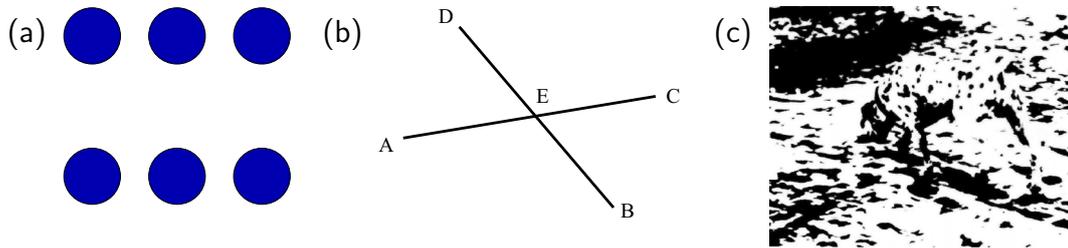


Figure 3.1: Examples for the Gestalt laws of (a) proximity and (b) good continuation; (c) Richard Gregory's picture of a Dalmatian dog, which can hardly be recognized without prior knowledge.

group them into one percept (61). Similarly older children or adults which have been blind since birth and have recovered vision after medical treatment, first recognize moving objects as unity, but have problems to tell overlapping objects apart. With visual experience they learn from the Gestalt rule of common fate, to make use of the Gestalt rule of good continuation in order to recognize objects (93). So even though this Gestalt rule is of no relevance when trying to find contours in static images, it seems to be fundamental for building an internal representation of a contour, i.e. for learning an association field.

Apart from motion (72; 49), additional cues that can facilitate contour detection are contrast and spatial scale (95), depth (48), and disparity (2). In addition to these and the Gestalt criteria, other knowledge about typical feature conjunctions in the world can also be crucial in order to interpret a visual scene. One example for an object which can hardly be recognized without prior knowledge about the world is Richard Gregory's Dalmatian dog, which is shown in Fig. 3.1(c). This picture contains a Dalmatian dog, sniffing on the floor and heading to the left, whose texture is so similar to the environment, that he can only be detected with the help of prior knowledge.

3.1.1 Contour integration in the visual periphery

Several psychophysical contour integration studies report that contour integration is better near the fovea than in the periphery (46; 45; 92). Hess and Dakin (46; 45) reported that human subjects were unable to detect contours consisting of alternating-phase Gabors in the periphery (beyond 10 degrees of visual angle eccentricity) of a display of randomly positioned and oriented edge elements. However, the subjects could see contours if they consisted of same-phase Gabor elements. Hence Hess and Dakin concluded, that there must be a fundamental difference between contour processing in central and peripheral vision. They claim, that an association field mechanism is available only for eccentricities < 10

degrees of visual angle, and that contour integration in the periphery can be modeled by a simple filter model without any subsequent cellular linking interactions. In contrast, (92) found that contour detection performance decreases gradually with eccentricity between 0 and 30 degrees visual angle. Only for the more difficult condition of alternating-phase stimuli the performance decreased slightly faster. Furthermore, subjects who suffered from the loss of their central vision could detect contours of 13 to 17 degree visual angle eccentricity in both phase conditions. Therefore (92) assume a similar mechanism for contour integration in central and peripheral vision, with similarity between edge elements of identical phase improving perception by adding a second Gestalt property as a hint to the location of a target, in this case a contour. Also (79) finds that both, phase-aligned and phase-alternated contours can be detected at an eccentricity of 13 degree visual angle. Furthermore Lovell changed smoothness and eccentricity of the contours, where a smooth contour is defined as a contour existing of few long arcs rather than having many inflections. Hence a closed contour has a greater smoothness than an open contour, given that path angle, element distance, and contour length are the same in both cases. It was shown that contour detection performance increases with the smoothness of the contour regardless of the eccentricity at which the contour was presented. However, the simple filter model proposed by Hess and collaborators shows decreasing performance for increasing smoothness and hence cannot explain the experimental findings, while an association field model could account for increasing performance for smooth contours. In contrast to the above works, there is a recent publication (70), which states that for circles or ellipses with good Gestalt properties, displayed centered within the visual field, contour integration performance for shape detection and discrimination is nearly constant between 0 and 35 degree eccentricity.

3.1.2 Subjective contours

Not all perceived contours need to be physically present. Two of the most prominent examples for subjective contours are Kanizsa figures (58) and Koffka crosses. A Kanizsa triangle is shown in Fig. 3.2 (a). Observers perceive a triangle, though there is no physical boundary line of the triangle. Similarly observers tend to perceive a white square or circle in the center of a Koffka cross (see Fig. 3.2 (b)). The perceived shape changes from a square for large arm widths of the cross to a circle for small arm widths (113). In neither case is a physical boundary of the shape present. These perceived boundaries are termed subjective contours or illusory contours. There is experimental evidence that real contours and illusory contours are processed or integrated by the same neuronal substrate in primary visual cortex (81). As a consequence a plausible contour integration model should also reproduce the perception of illusory contours.

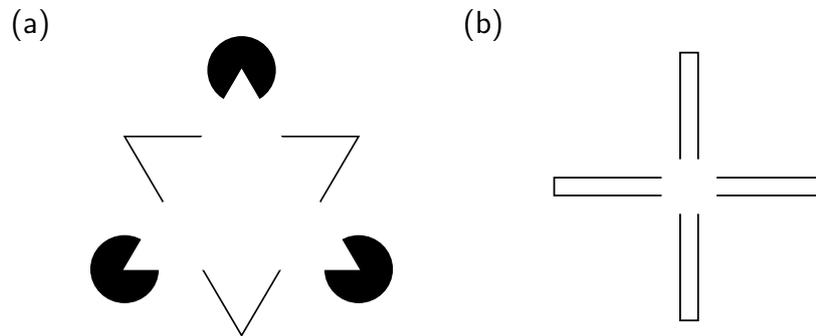


Figure 3.2: Two of the most famous contour illusions, the Kanizsa triangle (a) and the Koffka cross (b).

3.2 Imaging studies

Several studies tested the role of lower and higher visual areas in contour integration by combining functional magnetic resonance imaging (fMRI) and psychophysics (3; 66; 2; 84). fMRI lacks the high spatial resolution of intracortical recordings. However, compared to standard recording techniques it has the advantage that one can simultaneously measure the responses to the same set of stimuli in multiple visual areas.

Kourtzi and colleagues (66) tested responses across visual areas to collinear patterns, consisting of similarly oriented edge elements embedded in a background or randomly oriented edges, versus random patterns containing solely randomly oriented edge elements. Both, in humans and anesthetized monkeys, they found responses selective to collinear rather than to random patterns in earlier visual areas (V1 to V4) as well as in higher cortical areas (regions in the inferotemporal cortex). Selective responses to collinear patterns varied for different early visual areas depending on the size of receptive fields in that area, suggesting the involvement of multiple visual areas in contour integration at different spatial scales.

Despite some differences between the experiments with humans and monkeys Kourtzi obtained similar results for both species, suggesting similar mechanisms for the integration of local elements to global shapes for human and nonhuman primates.

3.3 Computational models

To construct a contour integration model one can have two different approaches. One approach would be to gather information about the response properties of the neurons, the organization of the visual system, and the connections between the neurons and construct a model on the basis of these known constituents of the visual system, and finally hope that it is as efficient in finding contours as humans. These models are bottom-up approaches.

As the brain is very effective in integrating contours and often uses strategies which are close to optimal for solving a given task, one could also think of a different strategy. One could try to build a model which processes and combines the given information efficiently in order to solve a given task in an optimal way. Such a model follows a top-down approach to explain visual perception. For stochastic problems a top-down approach requires a probabilistic algorithm. In a further step towards biological realism one can then search for neuronal substrates which can approximate the dynamics of the model.

3.3.1 Neuronal models

There are a number of models aiming to explain the mechanism underlying human contour integration. While some models use only known structures of the brain and emphasize biological plausibility others just try to reproduce human behavior or find an effective model of contour extraction for computer vision.

One influential biophysical model which uses only known elements and connections from primary visual cortex is the one from Li (77).

It assumes hypercolumns at discrete spatial locations \mathbf{i} . Each hypercolumn has $K = 12$ pairs of neurons having different preferred orientations. Each pair of neurons can represent an edge segment and consists of an excitatory neuron and an inhibitory interneuron, which are mutually connected. The excitatory cells receive visual input, and their output to higher visual areas quantifies the salience of the edge element. For a sketch of the model see Fig. 3.3.

The input a cell with preferred orientation ϕ receives from an edge segment at position \mathbf{i} with orientation θ is $I_{\mathbf{i}\theta} = \hat{I}_{\mathbf{i}\theta} \exp(-|\phi - \theta|/(\pi/8))$, where $\hat{I}_{\mathbf{i}\theta}$ is the input strength. The membrane potentials of the excitatory and inhibitory cells are denoted by $x_{\mathbf{i}\phi}$ and $y_{\mathbf{i}\phi}$, respectively. Their outputs are firing rates which are obtained from the membrane potential by the sigmoidal functions $g_x(\cdot)$ and $g_y(\cdot)$, respectively. Neuron pairs representing an oriented edge segment mutually exchange their activity $g_x(x_{\mathbf{i}\phi})$ and $g_y(y_{\mathbf{i}\phi})$. In addition the excitatory cells excite themselves with synaptic strength J_0 . An edge segment $\mathbf{j}\phi'$ at a different position can excite edge $\mathbf{i}\phi$ by directing an excitatory signal $J_{\mathbf{i}\phi,\mathbf{j}\phi'} g_x(x_{\mathbf{j}\phi'})$ to the

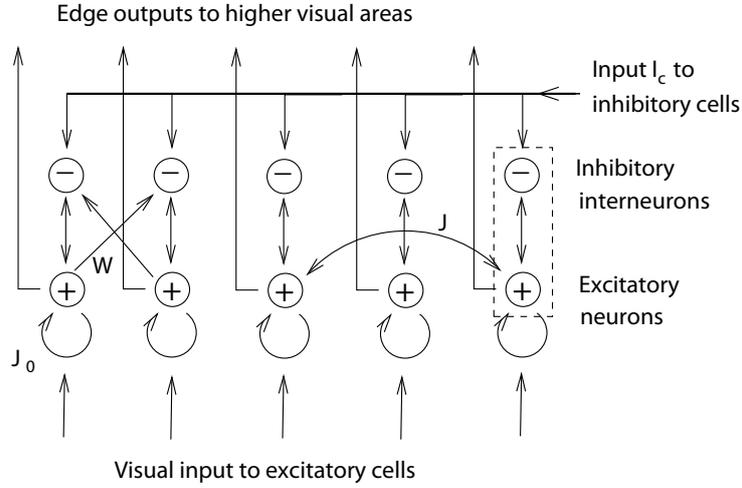


Figure 3.3: Neural components of the contour integration model by Li. J_0 denotes self-excitatory connections and J lateral excitatory connections between neurons representing edges at different locations. W is the lateral disinaptic inhibitory connection between edge elements. Figure redrawn from (77).

excitatory cell representing edge $i\phi$. Similarly edge $j\phi'$ can inhibit edge $i\phi$ by sending an excitatory signal $W_{i\phi,j\phi'}g_x(x_{j\phi'})$ to the inhibitory cell.

In total the neural dynamics is given by

$$\begin{aligned} \dot{x}_{i\phi} = & -\alpha_x x_{i\phi} - \sum_{\Delta\phi} \psi(\Delta\phi) g_y(y_{i,\phi+\Delta\phi}) + J_0 g_x(x_{i\phi}) \\ & + \sum_{j \neq i, \phi'} J_{i\phi, j\phi'} g_x(x_{j\phi'}) + I_{i\phi} + I_0 \end{aligned} \quad (3.1)$$

$$\dot{y}_{i\phi} = -\alpha_y y_{i\phi} + g_x(x_{i\phi}) + \sum_{j \neq i, \phi'} W_{i\phi, j\phi'} g_x(x_{j\phi'}) + I_c \quad (3.2)$$

where $J_{i\phi, j\phi'}$ and $W_{i\phi, j\phi'}$ are the strength of horizontal cortical connections and $1/\alpha_x = 1/\alpha_y = 1$ are the membrane time constants. I_c is the background input to inhibitory cells which is used to model top-down control signals. I_0 is the background input to the excitatory neurons which also serves to normalize the activity. $\psi(\Delta\phi)$ is a symmetric function that decreases with $|\Delta\phi|$ and provides inhibition within a hypercolumn. Furthermore each neuron gets noisy input. After a time interval of several membrane time constants the network approaches a limit cycle and the activity pattern $g_x(x_{i\phi})$ can be interpreted as a salience map.

The connection structure $J_{i\phi, j\phi'} g_x(x_{j\phi'})$ and $W_{i\phi, j\phi'} g_x(x_{j\phi'})$ is designed such that edge elements which are part of a smooth contour enhance each other while isolated edge elements should be suppressed or at least not enhanced. Excitatory

and inhibitory connections link cells with similar preferred orientations. However, the connections $J_{i\phi,j\phi'}$ mainly connect cells whose preferred orientations are aligned to the relative displacement of their receptive fields, while the connections $W_{i\phi,j\phi'}$ link cells whose preferred orientation is approximately orthogonal to the relative displacement of their receptive fields. This is consistent with experimental results (34), showing that there are some horizontal connections between neurons whose orientation preference and relative receptive field displacement are approximately orthogonal to each other. These connections $W_{i\phi,j\phi'}$ also cause competition between two alternative paths of a single contour. This interaction pattern resembles qualitatively the edge compatibility function in (140).

Contour detection is a fast process and only few membrane time constants after stimulus onset the saliency of contour elements is significantly higher than those for the background elements. On the examples of a line and a circle Li shows that the response levels of a segment which is part of the contour as compared to an isolated segment differ by the factor $(s_y - J_0 + 1)/s_y - J_0 - \sum_{i \neq j} J_{ij} + 1)$ where s_y denotes the slope of the curve at $\dot{y} = 0$ and $\sum_{i \neq j} J_{ij}$ is the sum of horizontal connection strengths between a contour element and all others. s_y usually decreases with lower input strength. Hence this model predicts much stronger contour enhancement for low input contrasts like used e.g. in experiments by Kapadia et al. (59) or Kovacs et al. (68). In a finite contour the end segments get less extinction from other segments and thus a closed contour or a longer contour is more salient than a finite shorter contour. However, this closure effect is weaker than predicted by psychophysical experiments (68) as salience between closed and open contours only differs towards the end elements of the open contour.

Each neuron pair of excitatory and inhibitory cells can be seen as a neural oscillator (78), that oscillates around the equilibrium point (\bar{x}, \bar{y}) . These oscillators are coupled and express collective behavior when embedded into a contour. When changing the notation such that $x_i - \bar{x}_i \rightarrow \tilde{x}_i$ with $\mathbf{X} = (x_1, x_2, \dots)^T$ and $\mathbf{Y} = (y_1, y_2, \dots)^T$. Then it holds for small \mathbf{X} and \mathbf{Y} :

$$\dot{\mathbf{X}} = -\mathbf{X} - \mathbf{G}'_y \mathbf{Y} + \mathbf{J}\mathbf{X} \quad (3.3)$$

$$\dot{\mathbf{Y}} = -\mathbf{Y} - \mathbf{G}'_x \mathbf{X} \quad (3.4)$$

where the elements of \mathbf{J} are $(\mathbf{J})_{ij} = J_0 g'_x(\bar{x}_j)$ for $i = j$ and $(\mathbf{J})_{ij} = J_{ij} g'_x(\bar{x}_j)$ for $i \neq j$. The diagonal matrices \mathbf{G}'_x and \mathbf{G}'_y have the entries $(\mathbf{G}'_x)_{ii} = g'_x(\bar{x}_i)$ and $(\mathbf{G}'_y)_{ii} = g'_y(\bar{y}_i)$. For contours of constant curvature \mathbf{J} has an orthogonal set of eigenvectors $\{\mathbf{X}^k\}$ with real eigenvalues λ^k and $k = 1, 2, \dots$. Ordering the eigenvectors such that $\lambda^1 \leq \lambda^2 \leq \dots$ and taking $\{\mathbf{X}^k\}$ as new basis for \mathbf{X} and \mathbf{Y} ,

results in the differential equations

$$\dot{x}^k = -x^k - g'_y(\bar{y})y^k + \lambda^k x^k \quad (3.5)$$

$$\dot{y}^k = -y^k - g'_x(\bar{x})x^k \quad (3.6)$$

with the solution

$$x^k(t) = x^k(0) \exp(-(1 - \lambda^k/2)t) \cos \omega^k t + \phi^k \quad (3.7)$$

Here $\omega^k = \sqrt{g'_y(\bar{y})g'_x(\bar{x}) - (\lambda^k)^2/4}$ is the oscillation frequency, $x^k(0)$ is the amplitude and ϕ^k the phase. As λ^1 is the highest eigenvalue, the system will be dominated by the first oscillation mode \mathbf{X}^1 . The components of the complex vector \mathbf{X}^1 determine the relative oscillation amplitudes and phases of the contour segments. For a contour of constant curvature with uniform inputs to its segments all components of \mathbf{X}^1 are the same. Hence all segments of a contour have the same amplitude and oscillate in phase. The strength of oscillation is largest for a straight line and decreases with curvature. Although Li derives these predictions on the example of a line and a circle, she expects them to carry over to arbitrary contours.

Even if the input for different contour segments is turned on at different times, synchronization will occur within that contour, while synchronization is rare between the elements of different contours.

This model is capable of enhancing contours solely by V1 neural circuits. However, when assuming that higher visual areas already know which elements belong to the contour, or expect the contour in a specific part of the visual field, Li's model can in addition make use of feedback sent back to V1. This feedback signal I_c is assumed to be sent to the inhibitory neurons. It influences the output of an edge detector as follows:

$$\delta g_x(\bar{x})/\delta I_c = -\frac{g'_y(\bar{y})g'_x(\bar{x})}{1 + g'_y(\bar{y})g'_x(\bar{x}) - I_0 g'_x(\bar{x})} \quad (3.8)$$

Here $I_c = I_{c,\text{background}} + I_{c,\text{control}}$ contains a background signal $I_{c,\text{background}}$ which is identical for all elements and modulates the overall visual alertness. $I_{c,\text{control}}$ can be either positive or negative and enhances or suppresses selected edge segments. The feedback signal in this model is sent to inhibitory cells in order to prevent a mixing of visual input from the external world and internal feedback. In this way weak or incomplete contours can be enhanced or filled in under the control feedback. However, no contour is enhanced, that does not at least partially exist in the visual input.

A similar model was published by Yen and collaborators (137). Here hypercolumns consist of multi-compartment pyramidal cells and interneurons. Cells

within one hypercolumn are densely interconnected with negligible delay. Half of the cells within a hypercolumn project to cells of the same orientation preference in neighboring hypercolumns via long-range horizontal connections. These connections have a short axonal delay when they extend parallel to the preferred orientation axis of the cell. These co-axial connections serve to synchronize the neuronal activity of spatially aligned cells. In contrast to Li (77) Yen and Finkel also assume connections orthogonal to the preferred orientation which facilitate synchronization as there is anatomical (103; 104; 86; 80; 28) and psychophysical (27; 99) evidence consistent with these trans-axial connections. The salience of a contour is represented by the sum of activities of all synchronized elements. The network output corresponds to the most salient contour.

Pettet et al. (97) also start with orientation selective spatial filters like those found in primary visual cortex. However, they assume that each Gabor patch in the stimulus display activates only one optimally tuned neuron. Units in the second stage of the neural network get constant input from the orientation selective cells in the first stage and facilitatory input from the other units of the network. In addition there is a normalizing shunting inhibition which leads to the convergence of the network to a steady state with bounded activity. The facilitatory interaction decreases with the distance between two receptive fields, with the curvature, and with the change in curvature of a function fitted through the orientation axis of the receptive fields. The parameters of these facilitatory inputs were matched to find the best agreement between simulations and psychophysical experiments. The model performance was then estimated using signal detection theory comparing responses for noise-alone stimuli and stimuli containing the contour.

Like the three models presented so far (77; 97; 137) also a more recent model by Ursino and La Cara (131) shows that local processing in V1 without the influence of higher areas can already emphasize contours. This model has four main components, namely feed-forward input from LGN, inhibitory feed-forward input which is maximally orthogonal to the target cell and suppresses non-optimal stimuli, excitatory cortical feedback, and long-range isotropic feedback inhibition, which suppresses noise and small contours due to the competition with longer contours. The main interest of the model was to investigate the cooperation between feedforward and feedback inputs. It suggests that the cooperation of these two interactions is necessary to extract contours from noise within short processing times. The model has been tested on artificial and natural images and was able to extract contours within only 30-40 ms stimulus presentation time.

Another model based on psychophysical and electrophysiological data which uses feed-forward and feedback is the one from Neumann and Sepp (90). This mainly investigates the role of feedback from V2.

Van Rullen and colleagues (132) simulated contour integration in a spiking

neural network. The main idea of their work was to use few spikes exchanged over horizontal connections between collinearly aligned edge detectors in order to cause neuronal units representing contour elements to fire prior to units responding to the presence of background elements. This mechanism basing on the ideas of temporal rank order coding put forward by Thorpe and coworkers (125) was supposed to explain the high speed of contour integration. In contrast to Ursino and colleagues (131), Van Rullen expected feedback cortical mechanisms to converge too slowly to reach an equilibrium solution within the time required for human contour integration or object recognition tasks.

Grossberg and Mingolla suggested the so-termed 'bipole property' of cells (40; 41; 38) as a key computational principle: For contour completion a linking mechanism must be activated only when a cell is stimulated simultaneously from two sides by horizontal connections. Input from only one side should not activate the linking mechanism, independently of the input strength. In order to avoid this extrapolation of a contour from a single inducer, and only interpolate the contour in between two appropriate inducers, a non-linear computing mechanism is needed. A more recent model developed in Grossberg's group (106) which incorporates the bipole property stresses the laminar organization of the visual cortex and investigates how the interacting structures of laminae, columns and maps in V1 and V2 cooperate to perceptual grouping.

Sherwood (117) modeled a network of simple cells in layer 2 and 3 of the primary visual cortex. These cells, which are involved in contour integration were modeled with high biophysical plausibility using Izhikevich's model (56). It was shown, that for certain sets of synaptic weights the neuronal activity of the cells representing the contour was synchronized. However the frequency of action potentials was much higher than the frequency observed empirically (21). Furthermore the simulations were performed using a synchronized initial condition. It remains unclear, whether synchrony would also arise when starting with irregular neuronal behavior.

3.3.2 Probabilistic and ideal observer models

As an example for probabilistic models of contour integration I want to introduce the framework of Williams and Thornber (136). Their approach is based on the idea that the distribution of contour paths can be described by particles moving with constant velocity in one direction while subjected to Brownian motion in the perpendicular direction (88). For this model Williams and Jacobs introduced the stochastic completion field, given by the distribution of particle trajectories passing a set of certain positions in certain directions (135), and devised an algorithm to compute the stochastic completion field in a local parallel network (134). Williams and Thornber modified this model in order to compute a scale-

3.3. Computational models

invariant distribution of closed contours ('completion field') from the positions of distinct edge elements (136).

The advection, diffusion and decay of particles can be described by a Fokker-Planck equation. Solving the Fokker Planck equation yields the Green's function $G_\gamma(\vec{x}, \theta; t_1|\vec{u}, \phi; t_0)$. This describes the probability that a particle traveling at speed γ , that passes position \vec{u} at time t_0 in direction ϕ will subsequently pass position \vec{x} at time t_1 in direction θ . This Green's function is invariant to time reversal, meaning the probability of passing two edges is identical when reversing the order and directions of the edges.

$$G_\gamma(\vec{x}, \theta; t_1|\vec{u}, \phi; t_0) = G_\gamma(\vec{u}, \phi + \pi; t_1|\vec{x}, \theta + \pi; t_0)$$

Furthermore the Green's function is invariant when scaling the speed of the particle and the distance between the two edges by the same factor

$$G_\gamma(\vec{x}, \theta; t_1|\vec{u}, \phi; t_0) = G_1(\vec{x}/\gamma, \theta; t_1|\vec{u}/\gamma, \phi; t_0)$$

Given an input pattern of N edges i with positions \vec{x}_i and orientations θ_i , Williams and Thornber assume two states i and \bar{i} for each edge i . Both states have the same position $\vec{x}_i = \vec{x}_{\bar{i}}$ but opposite directions $\theta_i = \theta_{\bar{i}} + \pi$ representing the two directions in which a randomly moving particle could join the edge element.

The probability $P(j|i)$ that a particle traveling with unit speed which was observed at edge i will later pass edge j can be obtained from the Green's function by integrating over time.

$$P(j|i) = \int_0^\infty dt_1 G_1(\vec{x}_j, \theta_j; t_1|\vec{x}_i, \theta_i; 0) \quad (3.9)$$

This is a path of length one, as there are no intermediate edges visited. The probability of a path of length n between edge i and edge j visiting $n-1$ intermediate edges is given by

$$P^{(n)}(j|i) = \sum_{k_n} \dots \sum_{k_1} P(j|k_n) \dots P(k_1|i) = (\mathbf{P}^n)_{ji} \quad (3.10)$$

where \mathbf{P} is a $2N \times 2N$ matrix with $P_{ji} = P(j|i)$. Hence it is no longer necessary to look at the continuous time random process underlying a particle visiting n edges at n real valued times. Instead the analysis of the particle trajectory can be based on a discrete time random process with no memory (Markov process) and with transition probabilities specified by the matrix \mathbf{P} .

Williams and Thornber define the edge saliency c_i of element i as the relative number of closed contours passing through edge i which is given by

$$c_i = \lim_{n \rightarrow \infty} \frac{\mathbf{P}^{(n)}(i|i)}{\sum_j \mathbf{P}^{(n)}(j|j)}$$

c_i can also be expressed as

$$c_i = \frac{(\mathbf{s}\bar{\mathbf{s}}^T)_{ii}}{\mathbf{s}^T\bar{\mathbf{s}}} \quad (3.11)$$

where \mathbf{s} and $\bar{\mathbf{s}}$ are the right and left eigenvectors of \mathbf{P} with the largest real positive eigenvalue λ , meaning $\lambda\mathbf{s} = \mathbf{P}\mathbf{s}$ and $\lambda\bar{\mathbf{s}} = \mathbf{P}^T\bar{\mathbf{s}}$. With this framework one can not only compute the relative number of contours passing a certain edge element in the display, but also the probability for a contour passing an arbitrary position \vec{u} in an arbitrary direction ϕ in the plane, i.e. one can determine the stochastic completion field. With $\eta = (\vec{u}, \phi)$ this is given by

$$c_\eta = \frac{1}{\lambda\mathbf{s}^T\bar{\mathbf{s}}} \sum_i P(\eta|i)s_i \cdot \sum_j P(j|\eta)\bar{s}_j \quad (3.12)$$

In an extension of this algorithm the model input consisting of exact positions and directions is replaced by an input bias function $\mathbf{b}(\vec{x}, \phi)$, a probability density function, that gives the probability that an edge with orientation θ exists at position \vec{x} . One of the successes of this model is the explanation of the experimental finding that the perceived shape of the illusory contour in the center of a Koffka cross changes from a circle to a square with increasing widths of the arms of the cross (113). A Koffka cross is shown in Fig. 3.2 (b) in Section 3.1.2.

Another probabilistic model which uses natural image statistics was proposed by Geisler and colleagues (33). Already in 1953 Brunswick hypothesized that environmental image statistics may underly the Gestalt principles of perceptual grouping, and that using statistical correlations of the visual world would allow the brain to disambiguate natural scenes, leading to its amazing performance in object segmentation and perception. Geisler and collaborators measured the co-occurrence of all edges within a visual scene, as well as the co-occurrence of edge elements which belong to the same contour, termed the Bayesian edge co-occurrence. Using this natural image statistics data to link all edges whose probability to belong to the same contour exceeds a certain threshold, can already explain human contour detection performance.

Now we have looked at two different kinds of contour integration models in more detail. Li's model is an example for a bottom-up approach to the contour integration problem. It is constructed from known components of V1 and works with operations and on connection patterns which were found in experiments. Li analyzes how well this reproduces human contour integration behavior. In contrast, Thornber and collaborators pursued a top-down approach. They start by quantifying the task the model is supposed to solve. From this they construct a model that ideally combines all the information which can be gathered from the stimuli in order to solve the problem. Then they speculate how far this probabilistic model can be mapped on the structure of the brain.

So what are the main similarities and the fundamental differences of these two model classes? Both models start with some evidence for the edge elements

in the stimulus, which is the afferent input in the neuronal model and a probability distribution in the probabilistic model. These could easily be identified. Between the representations of single edge elements there are interactions which can be excitatory or inhibitory in the neuronal framework, while the probabilistic approach uses only excitatory connections. However, due to the normalization, little or no excitation in the Thornber model should have a similar effect as inhibition in the Li model. So even though the motivation for the interaction pattern is quite different in both models, the connections in both models are similar in the sense that they strongly link edge elements between collinearly aligned edges and weaken or do not strengthen the connections between edge elements whose centers are displaced orthogonally to their orientation. However one qualitative difference lies in their symmetry: While edge elements in the neuronal model connect symmetrically to neurons in both directions, connections in the probabilistic model are unidirectional and connect only to one direction.

In contrast to Thornber, Li uses two neurons to represent one edge element. Computationally, a single cell representing one edge element can realize the same operations, as long as it can receive and send both, excitatory and inhibitory signals. However, such a cell is not biologically plausible as it violates Dale's principle after which a neuron has either inhibitory or excitatory outputs, but not both.

So far, we can at least to some approximation map the components of the two models onto each other. However, one fundamental difference between the two approaches is how the lateral interactions are coupled to the evidence for some edge element. While these two quantities are multiplied in the probabilistic model, they are summed in Li's model as well as in many other bottom-up models (i.e. (97; 137)) with the exception of the model by Grossberg (i.e. (42)).

This difference provides the motivation to compare these two classes of models: Models with multiplicative couplings and unidirectional interactions on the one side and models with additive couplings and bidirectional interactions on the other side.

Chapter 3. Contour integration - existing models and experimental findings

Chapter 4

Theory: model classes and their dynamics

In this section the models used for this study will be introduced. We want to start with a Bayesian model for contours of finite length that inverts the contour generation process and extracts contours from stimuli generated by a Markov process. Following that we will describe a model, which is also probabilistically motivated but less restrictive than the Bayesian model. The third model is a standard neuronal network.

4.1 Bayesian contour integration

The first model we want to introduce is a Bayesian model. This is the model that "inverts" the contour generation process in order to optimally detect contours in a stimulus. Before we construct such a model, we want to discuss, how mathematically well-defined contours can be generated.

4.1.1 Contour creation

Mathematically, a contour ensemble can be characterized by a conditional link probability, a so called 'association field' ρ (88; 27). If an edge at position \mathbf{x} with orientation ϕ belongs to a contour, then $p(\mathbf{x}', \phi' | \mathbf{x}, \phi)$ gives the probability that the contour next passes through position \mathbf{x}' with orientation ϕ' . The association field can be used to generate contour ensembles with well-defined statistical properties. Vice versa, the association field can be used in order to determine the probability that a certain edge configuration belongs to a contour. Hence it can be employed to extract contours from a visual stimulus.

For the generation of contours we first choose the hemifield in which the contour will be located, assuming that the probability for each hemifield is $1/2$. In this hemifield we generate the contour configuration \vec{i}_j by a statistical process. We start by randomly choosing the first element $i_{j1} = (x_{j1}, \phi_{j1})$ of the contour \vec{i}_j with position x_{j1} within the given hemifield and orientation ϕ_{j1} . Then we draw the next contour element from the probability distribution $p(x_{j,l+1}, \phi_{j,l+1} | x_{j,l}, \phi_{j,l})$. We continue drawing successive contour elements from this probability distribution until we have the desired number L of contour elements. In this process the probability for an edge i_{jl} to belong to the contour only depends on the preceding edge element $i_{j,l-1}$. The probability for an edge configuration \vec{i}_j given that all its edge elements belong to the contour (**cont**) is given by $p(\vec{i}_j | \vec{i}_j \in \text{cont}) = \prod_{l=1}^{L-1} p(i_{j,l+1} | i_{j,l})$.

When the generation of the contour is completed we place background elements of random position and orientation in the two hemifields. In this process we want the distance distribution ρ_{CC} between contour elements, the distance distribution ρ_{BB} between background elements and the distance distribution ρ_{CB} between contour and background elements to be identical. Hence the contour is generated with a distance distribution which usually occurs in the background when randomly placing background elements and in a second step the background elements are randomly shifted until $\rho_{BB} = \rho_{CC} = \rho_{CB}$ is true in a good approximation.

4.1.2 Contour integration

Let us consider we have a stimulus with N Gabor elements whose positions and orientations are given by $\vec{i} = \{\vec{x}, \vec{\phi}\}$ and that we can split this stimulus into a right hemifield RH which is characterized by \vec{i}_{RH} and a left hemifield LH with elements at \vec{i}_{LH} , respectively, where $\vec{i} = \vec{i}_{RH} \cup \vec{i}_{LH}$. Furthermore we assume that one contour consisting of L elements is placed in either hemifield with equal probability $p(\text{cont} \in \text{RH}) = p(\text{cont} \in \text{LH})$. Now we want to know, which side contains the contour (**cont**) and which side contains solely background elements (**back**). To answer this question we use the rule of Bayes to derive the conditional probability for a contour on the right hemifield, given the stimulus configuration \vec{i}

$$p(\text{cont} \in \text{RH} | \vec{i}) = \frac{p(\vec{i} | \text{cont} \in \text{RH}) \cdot p(\text{cont} \in \text{RH})}{\sum_{S=\text{RH,LH}} p(\vec{i} | \text{cont} \in S) \cdot p(\text{cont} \in S)} \quad (4.1)$$

Let us assume there coexist N_j possible contour configurations in the stimulus. A contour configuration is one possible way to route a contour through L edge elements. We denote a contour configuration by $\vec{i}_j = (i_{j1}, \dots, i_{jL})$. Summing up the probabilities for all possible contour configurations \vec{i}_j within the right hemifield

4.1. Bayesian contour integration

RH, associating all other elements to the background, we obtain

$$p(\vec{i}|\text{cont} \in \text{RH}) = \sum_{\vec{i}_j \in \text{RH}} p(\vec{i}_{\text{RH}}/\vec{i}_j | \vec{i}_{\text{RH}}/\vec{i}_j \in \text{back}) p(\vec{i}_j | \vec{i}_j \in \text{cont}) p(\vec{i}_{\text{LH}} | \vec{i}_{\text{LH}} \in \text{back}) \quad (4.2)$$

$p(\vec{i}_{\text{RH}} | \vec{i}_{\text{RH}} \in \text{back})$ can be factorized into a probability $p(\vec{i}_j | \vec{i}_j \in \text{back})$ to observe an edge configuration $\vec{i}_j \in \vec{i}_{\text{RH}}$ which is part of the background and the probability $p(\vec{i}_{\text{RH}}/\vec{i}_j | \vec{i}_{\text{RH}}/\vec{i}_j \in \text{back})$ for the remaining edge elements $\vec{i}_{\text{RH}}/\vec{i}_j$ within this hemifield, given that they belong to the background. Hence we can rewrite Eq. 4.2 as

$$p(\vec{i}|\text{cont} \in \text{RH}) = \sum_{\vec{i}_j \in \text{RH}} \frac{p(\vec{i}_{\text{RH}} | \vec{i}_{\text{RH}} \in \text{back})}{p(\vec{i}_j | \vec{i}_j \in \text{back})} p(\vec{i}_j | \vec{i}_j \in \text{cont}) p(\vec{i}_{\text{LH}} | \vec{i}_{\text{LH}} \in \text{back}) \quad (4.3)$$

With this and the analog expression for $p(\vec{i}|\text{cont} \in \text{LH})$ Eq. 4.1 becomes

$$p(\text{cont} \in \text{RH} | \vec{i}) = \frac{\sum_{\vec{i}_j \in \text{RH}} \frac{p(\vec{i}_j | \vec{i}_j \in \text{cont})}{p(\vec{i}_j | \vec{i}_j \in \text{back})}}{\sum_{\vec{i}_j \in \text{RH}} \frac{p(\vec{i}_j | \vec{i}_j \in \text{cont})}{p(\vec{i}_j | \vec{i}_j \in \text{back})} + \sum_{\vec{i}_j \in \text{LH}} \frac{p(\vec{i}_j | \vec{i}_j \in \text{cont})}{p(\vec{i}_j | \vec{i}_j \in \text{back})}} \quad (4.4)$$

Eq. 4.4 and the analog expression for $p(\text{cont} \in \text{LH} | \vec{i})$ are normalized such that $\sum_{S=\text{RH,LH}} p(\text{cont} \in S | \vec{i}) = 1$. Hence Eq. 4.4 should equally hold when replacing the probabilities $p(\vec{i}_j | \vec{i}_j \in \text{cont})$ and $p(\vec{i}_j | \vec{i}_j \in \text{back})$ by the likelihoods $q(\vec{i}_j | \vec{i}_j \in \text{cont})$ and $q(\vec{i}_j | \vec{i}_j \in \text{back})$. Here $q(\vec{i}_j | \vec{i}_j \in \text{cont})$ is the likelihood that we obtain the configuration \vec{i}_j by drawing the contour elements i_{j1}, \dots, i_{jL} from an association field. $q(\vec{i}_j | \vec{i}_j \in \text{back})$ is the likelihood to place the background elements $i_{j1} \dots i_{jL}$.

The likelihood $q(\vec{i}_j | \vec{i}_j \in \text{cont})$ can be expressed by the edge link probabilities given by the association field. $q(\vec{i}_j | \vec{i}_j \in \text{cont}) = \prod_{l=1}^{L-1} p(x_{jl+1}, \phi_{jl+1} | x_{jl}, \phi_{jl})$. Here the link probabilities depend on the positions and orientations of the edge elements. However, when performing a coordinate transformation we can write the link probability in terms of the distance $r_{jl,jl+1} = \sqrt{(x_{jl} - x_{jl+1})^2}$ between edge i_{jl} and edge i_{jl+1} , the viewing angle $\alpha_{jl,jl+1} = \arctan 2 \left(\frac{x_{jl+1,2} - x_{jl,2}}{x_{jl+1,1} - x_{jl,1}} \right) - \phi_{jl}$ under which edge i_{jl+1} is seen from edge i_{jl} and the difference $\beta_{jl,jl+1} = \phi_{jl+1} - \phi_{jl}$ between the orientations of edge i_{jl+1} and edge i_{jl} (see Fig. 4.1). Here $x_{jl,1}$ and $x_{jl,2}$ denote the two components of the position vector of edge i_{jl} . These coordinates have the advantage that the link probability between edge elements factorizes into a distance-dependent part and an angle-dependent part $p(r_{jl,jl+1}, \alpha_{jl,jl+1}, \beta_{jl,jl+1}) = p^{\text{dist}}(r_{jl,jl+1}) \cdot p^{\text{ang}}(\alpha_{jl,jl+1}, \beta_{jl,jl+1})$.

Hence the likelihood $q(\vec{i}_j | \vec{i}_j \in \text{cont})$ can be split into the likelihood $q_{\text{C}}^{\text{dist}}(\vec{r}_j)$ to draw contour elements with the inter-element distances $\vec{r}_j =$

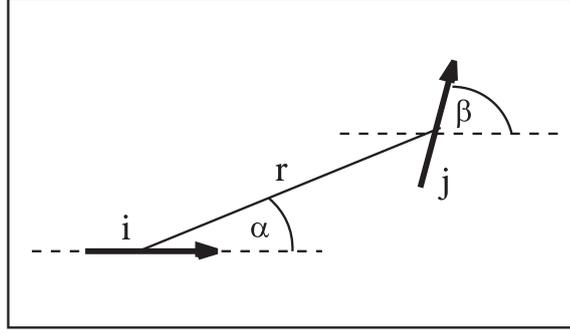


Figure 4.1: Illustrated are the inter element distance r , the viewing angle α under which the destination edge is seen from the source edge, and the orientation difference β between the two edges.

$r_{j1,j2}, \dots, r_{jL-1,jL}$ and the likelihood $q_C^{\text{ang}}(\alpha_{j1,j2}, \dots, \alpha_{jL-1,jL}, \beta_{j1,j2}, \dots, \beta_{jL-1,jL}) = \prod_{l=1}^{L-1} p^{\text{ang}}(\alpha_{j,l,j,l+1}, \beta_{j,l,j,l+1})$ that these contour elements have the orientation differences $\beta_{j1,j2}, \dots, \beta_{jL-1,jL}$ and the viewing angles $\alpha_{j1,j2}, \dots, \alpha_{jL-1,jL}$. It holds

$$q(\vec{i}_j | \vec{i}_j \in \text{cont}) = q_C^{\text{dist}}(\vec{r}_j) \cdot \prod_{l=1}^{L-1} p(\alpha_{j,l,j,l+1}, \beta_{j,l,j,l+1})$$

In the background each orientation occurs with the same probability and independently of the orientations of the neighboring edges. Hence the likelihood for any orientation configuration in the background is $q_B^{\text{ang}}(\alpha_{j1,j2}, \dots, \alpha_{jL-1,jL}, \beta_{j1,j2}, \dots, \beta_{jL-1,jL}) = \text{const} \forall j$ and the likelihood $q(\vec{i}_j | \text{back})$ basically corresponds to the likelihood $q_B^{\text{dist}}(\vec{r}_j)$ to place background elements with element distances \vec{r}_j .

With these likelihoods Eq. 4.4 becomes

$$p(\text{cont} \in \text{RH} | \vec{i}) = \frac{\sum_{\vec{i}_j \in \text{RH}} \frac{q_C^{\text{dist}}(\vec{r}_j)}{q_B^{\text{dist}}(\vec{r}_j)} \prod_{l=1}^L p(\alpha_{j,l,j,l+1}, \beta_{j,l,j,l+1})}{\sum_{S=\text{RH,LH}} \sum_{\vec{i}_j \in S} \frac{q_C^{\text{dist}}(\vec{r}_j)}{q_B^{\text{dist}}(\vec{r}_j)} \prod_{l=1}^L p(\alpha_{j,l,j,l+1}, \beta_{j,l,j,l+1})} \quad (4.5)$$

The likelihoods of a distance configuration \vec{r}_j can be expressed as the likelihoods of each single position x_{jL} . It yields $q_C^{\text{dist}}(\vec{x}_j) = \prod_{l=1}^L q_B^{\text{dist}}(x_{jL})$ and $q_B^{\text{dist}}(\vec{x}_j) = \prod_{l=1}^L q_B^{\text{dist}}(x_{jL})$, respectively.

$q_C^{\text{dist}}(x_{jL})$ and $q_B^{\text{dist}}(x_{jL})$ depend on the probability distributions over the distance to the neighboring elements of x_{jL} . As we generated the contours such that position and orientation of each element only depend on the preceding element, we only take into account nearest neighbors here. With $d_{mnjL} = \sqrt{x_{mn}^2 - x_{jL}^2}$ being the distance between x_{jL} and its nearest neighboring elements x_{mn} , the

4.1. Bayesian contour integration

probability to find a background element at position x_{jl} is

$$q_B^{\text{dist}}(x_{jl}) = \prod_{mn; i_{mn} \in \text{back}} \sqrt{\rho_{BB}(d_{mnjl})} \cdot \prod_{mn; i_{mn} \in \text{cont}} \sqrt{\rho_{BC}(d_{mnjl})}$$

where ρ_{BB} is the distance distribution between nearest neighboring background elements and ρ_{BC} the distance distribution between nearest neighboring contour and background elements. Accordingly

$$q_C^{\text{dist}}(x_{jl}) = \prod_{mn; i_{mn} \in \text{back}} \sqrt{\rho_{CB}(d_{mnjl})} \cdot \prod_{mn; i_{mn} \in \text{cont}} \sqrt{\rho_{CC}(d_{mnjl})}$$

where ρ_{CC} is the distance distribution between nearest neighboring contour elements. However, all stimuli subjected to this model throughout this thesis are designed such that $\rho_{BB} = \rho_{CC} = \rho_{CB} = \rho_{BC}$. It follows $p_C^{\text{dist}} = p_B^{\text{dist}}$ and hence the distance dependence cancels when inserting this in Eq. 4.5. Then we obtain

$$p(\text{cont} \in \text{RH} | \vec{i}) = \frac{\sum_{\vec{i}_j \in \text{RH}} \prod_{l=1}^L p(\alpha_{j,l,j,l+1}, \beta_{j,l,j,l+1})}{\sum_{S=\text{RH,LH}} \sum_{\vec{i}_j \in S} \prod_{l=1}^L p(\alpha_{j,l,j,l+1}, \beta_{j,l,j,l+1})} \quad (4.6)$$

However, for stimuli where the distance statistics differ for contour and background elements, this could lead to the integration of contour elements of random orientation based on element distance alone. Assume for example that the distance distribution ρ_{BB} between background elements is shifted to larger element distances compared to ρ_{CC} . Then an association field which employs the correct distance distribution ρ_{CC} on average predicts a higher link probability between contour elements than between background elements and hence this prior knowledge about this distance distribution can be sufficient to find the contour even if the orientations of the contour elements are not aligned.

So far we assumed that the exact orientation of each edge element is known. Now we consider the more general case in which we know only a probability distribution over n_{ori} possible orientations $\phi_{j,l,k}$ with $k = 1, \dots, n_{\text{ori}}$ for the edge i_{jl} . Let us denote the probability that edge i_{jl} has orientation $\phi_{j,l,k}$ by u_{jlk} . Comprising the orientation probabilities u_{jlk} for all edge elements within the right hemifield to vector \vec{u}_{RH} and analogously the probabilities for the orientations of elements on the left hemifield to \vec{u}_{LH} we can include this orientation uncertainty into Eq. 4.6 rewriting it as matrix multiplication. Therefore we define the matrix \mathbf{P}_{RH} with the elements $P_{\text{RH},mn} = p(\alpha_{m,n}, \beta_{m,n}) \forall i_m, i_n \in \text{RH}$. Analogously \mathbf{P}_{LH} contains the link probabilities between all edge elements on the left hemifield. Then it holds

$$p(\text{cont} \in \text{RH} | \vec{i}) = \frac{\sqrt{\vec{u}_{\text{RH}}^T} \mathbf{Q}_{\text{RH}}^{L-1} \sqrt{\vec{u}_{\text{RH}}}}{\sqrt{\vec{u}_{\text{RH}}^T} \mathbf{Q}_{\text{RH}}^{L-1} \sqrt{\vec{u}_{\text{RH}}} + \sqrt{\vec{u}_{\text{LH}}^T} \mathbf{Q}_{\text{LH}}^{L-1} \sqrt{\vec{u}_{\text{LH}}}} \quad (4.7)$$

Here the orientation probability is included into the matrix $\mathbf{Q}_{\text{RH}} = \sqrt{\vec{u}_{\text{RH}}} \mathbf{P}_{\text{RH}} \sqrt{\vec{u}_{\text{RH}}^T}$ and $\mathbf{Q}_{\text{LH}} = \sqrt{\vec{u}_{\text{LH}}} \mathbf{P}_{\text{LH}} \sqrt{\vec{u}_{\text{LH}}^T}$ and \mathbf{Q}^l denotes the l -th power of the matrix \mathbf{Q} .

4.1.3 Bayesian model for specific contour position

So far we derived a Bayesian model which is optimal for the task of detecting a contour on either hemifield of the screen. However, this is an artificial task imposed by an experimentalist on a human observer. In everyday life we usually want to know the exact location of an object rather than just the side of our visual field where the object is located. Thus it seems more plausible to assume that subjects first determine the location of the most salient object and in a second step decide which hemifield this location belongs to.

This observation motivated to compute for each edge element the probability that it belongs to the contour. This probability also serves as a measure of the salience for this edge element. The probability $P_L^l(i)$ that edge i is the l -th element of a contour of length L is given by

$$P_L^l(i) = \frac{\left(\sqrt{\bar{\mathbf{u}}^T} \mathbf{Q}^{L-l} \mathbf{Q}^{l-1} \sqrt{\bar{\mathbf{u}}}\right)_i}{\sum_j \left(\sqrt{\bar{\mathbf{u}}^T} \mathbf{Q}^{L-l} \mathbf{Q}^{l-1} \sqrt{\bar{\mathbf{u}}}\right)_j} \quad (4.8)$$

Here $\mathbf{Q} = \sqrt{\bar{\mathbf{u}}^T} \mathbf{P} \sqrt{\bar{\mathbf{u}}}$, where the matrix \mathbf{P} contains link probabilities between all edge elements within the stimulus, regardless on which hemifield they are located. The probability $P_L(i)$ that edge i belongs to the contour is then obtained by summing over the $P_L^l(i)$.

$$P_L(i) = \sum_{l=1}^L \frac{P_L^l(i)}{\sum_i P_L^l(i)} \quad (4.9)$$

From a probabilistic point of view this model is not ideal as it ignores the prior knowledge that no contour passes the border between the two hemifields. However, this model is still optimal under the assumption that the visual system is not specialized to the task of finding an object in either hemifield and rather searches for the position of an object. Hence we will use this model rather than the one defined by Eq. 4.7. Also the term 'Bayesian model' will refer to the model given by Eq. 4.9 and Eq. 4.8.

4.2 Contour detection and read out mechanisms

4.2.1 Neuronal context

This purely probabilistic algorithm can be mapped onto a neuronal context. There the evidence μ for an oriented edge corresponds to the afferent input to a neuron with preferred orientation ϕ . The link probabilities resemble the efficacy of horizontal connections between orientation columns. $P_L(x, \phi)$ can be

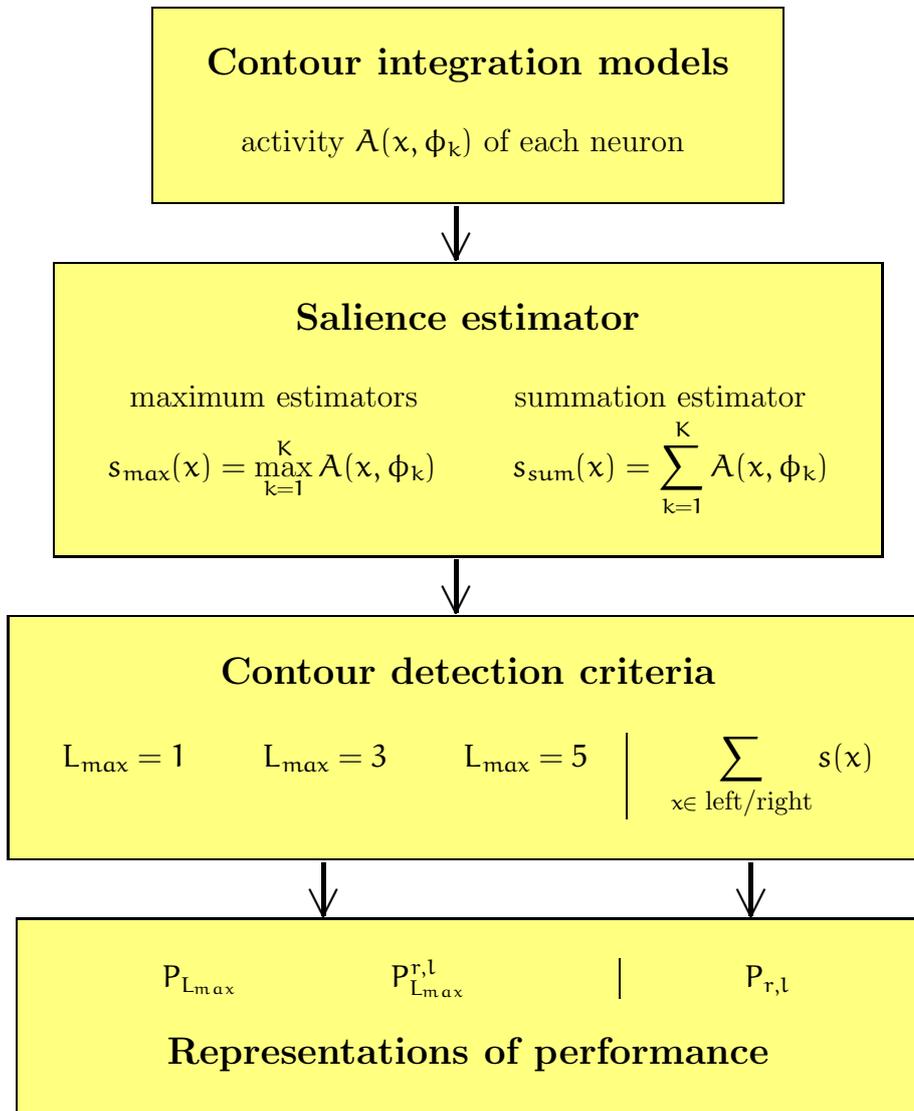


Figure 4.2: From model to performance: Overview over different model dynamics, saliency estimators, contour detection criteria, and representations of performance.

identified with the activity $A(x, \phi)$ of a neuron with preferred orientation ϕ and receptive field located at position x . This neuronal activity can be interpreted as the firing-rate of the neuron.

4.2.2 Estimating the saliency

There are many possibilities to define the saliency of a certain location x , of which we will discuss two generic choices here. We can either sum the activity $A(x, \phi_k)$ over all possible orientations in a hypercolumn

$$s_{\text{sum}}(x) = \sum_{k=1}^K A(x, \phi_k) \quad (\text{summation estimator}) \quad (4.10)$$

or we can take the maximum

$$s_{\text{max}}(x) = \max_k(A(x, \phi_k)) \quad (\text{maximum estimator}) \quad (4.11)$$

4.2.3 Criteria for contour detection

Having determined the saliency of each edge element the model has to make a decision on which hemifield of the screen the contour is located. There are again several possibilities: One can evaluate the saliencies of the complete hemifields $s_{\text{RH}} = \sum_{i \in \text{RH}} s(i)$ or $s_{\text{LH}} = \sum_{i \in \text{LH}} s(i)$, respectively and choose the more salient hemifield, which would be the ideal strategy for the given task. Or one could search for the most likely position of a contour and choose the corresponding hemifield. This corresponds to detecting the contour at the position x with the highest saliency s_{max} or s_{sum} , respectively. However, when taking only into account the maximum saliency, contours sometimes are mispredicted despite having a very high saliency, because of a single background element which exceeds the saliency of the contour element. Especially when counting a contour as detected only when the maximum saliency is located at one of the contour elements, many 'mistakes' of the model arise due to a 'background element' i_b adjacent to the contour, which is collinearly aligned to the last contour element. In this case i_b often is the most salient edge element in the display as the multiplicative model predicts a particularly high saliency for the ends of the contours. To avoid this we used as criterion for contour detection that more than half of the $n = 1, 3$ or 5 highest saliencies should be on the hemifield containing the contour.

Restrictions

This model assumes open contours of a given length L . In real life observers effortlessly detect contours without prior knowledge about their length. Hence

4.2. Contour detection and read out mechanisms

this model is too restrictive and therefore not plausible. Nevertheless it can serve as a benchmark: provided that the brain uses no other information than is also available to the model, human contour detection performance should not exceed the performance of the model. Hence if we find a case where humans significantly outperform the model, we know that humans must take into account additional information.

As the performance of human observers increases with stimulus presentation time one could assume that contour integration is an iterative process. However this model is only iterative when calculating the saliency s_{RH} or s_{LH} for a complete hemifield, not when calculating the saliency for each edge element i .

4.2.4 Iterative models

The dynamics of these two models is given by

$$\tau \dot{A}(x, \phi) = -A(x, \phi) + n(t)g(I(x, \phi)) \quad (4.12)$$

where $n(x, \phi) = (\sum_{x, \phi} g(I(x, \phi)))^{-1}$ is a normalization factor, τ is the decay constant, and $g(I) = I$ for $I > 0$, and 0 otherwise is the gain function. The total synaptic input is given by

$$I(x, \phi) = I_a \cdot I_{aff}(x, \phi) * I_l \cdot I_{lat}(x, \phi) \quad (4.13)$$

where $*$ means a multiplication for the multiplicative model and a summation for the additive model. For the additive model this is a simplified linear Wilson Cowan model. To solve the differential equation 4.12 numerically, the Euler method is employed. The precision of this iterative method depends on the integration time step δt as illustrated in Fig. 4.3. There $\Delta s_{max}(\Delta t_1, \Delta t_2) = |s_{max}(\Delta t_2) - s_{max}(\Delta t_1)|/s_{max}(\Delta t_1)$ is used as a measure for the quality of the approximation of the differential equation. While for the additive model a precision in the one percent range is achieved already with $\Delta t = 0.032$. For the multiplicative model a step width of about $\Delta t = 0.004$ is needed for this precision.

The multiplicative model approximates the probabilistic model for closed contours or infinitely long contours. This can be seen when rewriting Eq. 4.12 as a difference equation. Then we get

$$A(x, \phi, t + dt) = A(x, \phi, t)(1 - \Delta t/\tau) + n(t)I(x, \phi, t)$$

where $A(x, \phi, t + dt)$ is the activity of the neuronal unit with preferred orientation ϕ and receptive field at position x at iteration time $t + \Delta t$. Accordingly $A(x, \phi, t)$ is the activity at time t . When the time Δt for each iteration is set to τ , exactly

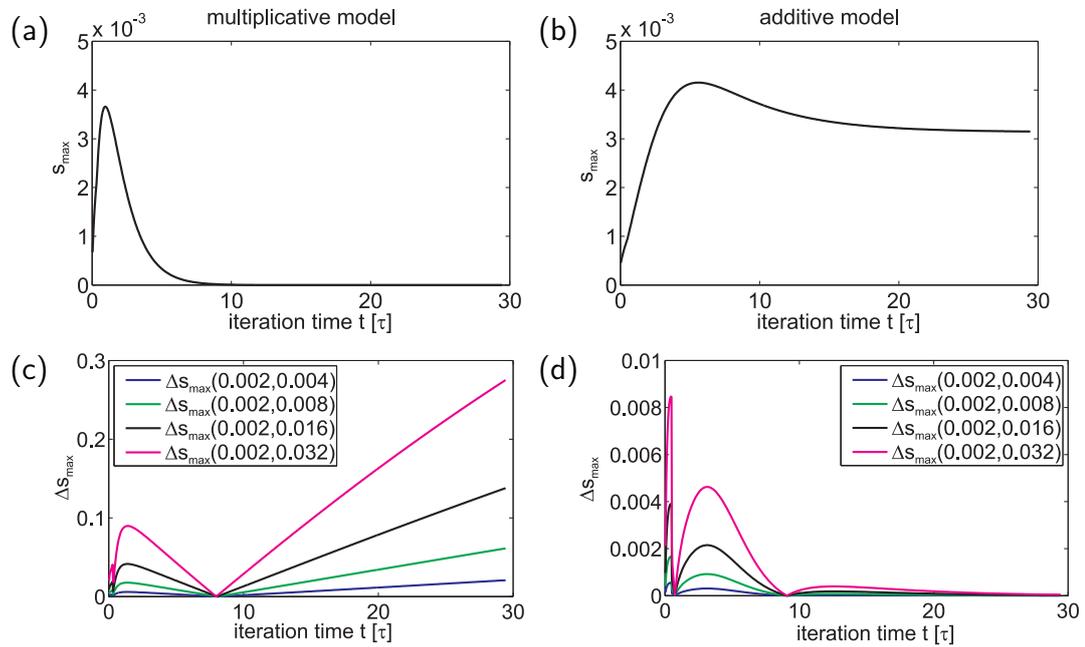


Figure 4.3: Maximum salience s_{\max} in a hypercolumn representing a contour element predicted by the multiplicative (a) and the additive (b) model for an example stimulus, computed with a step width of 0.002; (c) + (d) Δs_{\max} for different Δt for multiplicative and additive model, respectively.

one further edge element is integrated in each iteration time step. Then we obtain in the z -th iteration

$$A(\mathbf{i}_z, \mathbf{t} + z \cdot \Delta \mathbf{t}) = \frac{\sum_{\mathbf{i}_{z-1}} \dots \sum_{\mathbf{i}_1} I_{\text{aff}}(\mathbf{i}_z) p(\mathbf{i}_z | \mathbf{i}_{z-1}) \dots I_{\text{aff}}(\mathbf{i}_2) p(\mathbf{i}_2 | \mathbf{i}_1) I_{\text{aff}}(\mathbf{i}_1)}{\sum_{\mathbf{i}_z} \text{numerator}} \quad (4.14)$$

where $\mathbf{i}_z = (\mathbf{x}_z, \phi_z)$. This is the probability $p_z^z(\mathbf{i}_z)$ that an edge at position \mathbf{x}_z with orientation ϕ_z is the last element of a contour of length z . Hence in the stationary state $A(\mathbf{x}, \phi, \mathbf{t})$ gives the probability that edge element (\mathbf{x}, ϕ) belongs to a closed or infinitely long contour.

4.3 Synaptic input

The contour integration models described in this chapter have two kinds of synaptic input: afferent input from lower visual areas due to the visual stimuli and lateral input from the horizontal interaction between neuronal edge detectors. In the simulations we generally use Von Mises functions

$$M(z, \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp(\kappa \cos(z - \mu)) \quad (4.15)$$

to characterize lateral and afferent input distributions. This is a continuous distribution defined on the range $z \in [0, 2\pi)$ with mean μ and a concentration parameter $\kappa > 0$. $I_0(\kappa)$ is the modified Bessel function of the first kind of order 0. The Von Mises function is the circular analog in \mathbb{S}^1 of a Gaussian distribution in \mathbb{R} .

4.3.1 The afferent input

The afferent input

$$I_{\text{aff}}(\phi, \theta) = M(2\phi, 2\theta, 1/\sigma_{\text{aff}}^2) \quad (4.16)$$

is centered around the preferred orientation ϕ of the neuronal orientation column with a width of σ_{aff} . Here the angles are multiplied by two in order to get the same afferent input for θ and $\theta + \pi$ as a bar or Gabor patch has only an orientation. Thus no distinction between these two directions is possible.

Discretization of orientation columns

So far we assumed that the possible orientations ϕ_k are the same for each hypercolumn. This means that especially for small σ_{aff} one needs a large number K of orientations for each hypercolumn in order to avoid sampling problems, assuring that each Gabor patch within the stimulus results in an afferent input

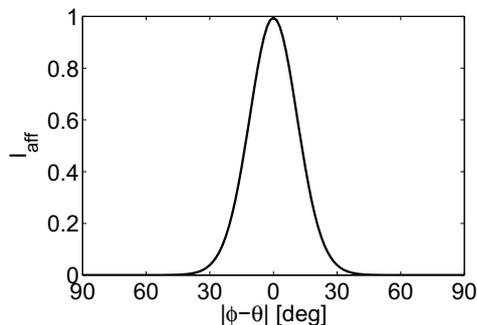


Figure 4.4: Afferent input with $\sigma_{\text{aff}} = \pi/8$.

to at least one orientation column. For this reason the performance decreases with decreasing K due to the lack of information about the orientation of certain Gabor elements. The impact of decreasing K is shown in Fig. 4.5 (a) and (b).

Large K result in long running times for the simulations, which makes it extremely difficult to search certain parameter ranges in order to find the optimal working regime for a model. However, this seems to be inevitable for the additive model where afferent and lateral input need to be balanced in order to compute a plausible performance. One possibility to avoid sampling problems as well as long computing times is to adapt the orientations of each hypercolumn such that one orientation matches the orientation of the Gabor patch at the position represented by this hypercolumn, i.e. $\phi_k = \text{mod}(\theta + k \cdot 2\pi, 2\pi)$ where $k = 1 \dots K_a$. Here the performance increases for smaller K_a , especially for $K_a \leq 12$ (see Fig. 4.5 (c) and (d)). Both, the additive and the multiplicative model show the best performance for $K_a = 1$ as there is no uncertainty about the orientation of an edge. In this case the afferent input is just a constant. For the multiplicative model this constant factor cancels due to the normalization, such that the differential equation describing this model simplifies to

$$\tau \dot{A} = -A + \frac{I_{\text{lat}}}{\sum I_{\text{lat}}} \quad (4.17)$$

With the additive model optimal performances are achieved when the contributions of afferent and lateral input are balanced. Therefore the ratio of I_l/I_a needs to be adjusted because for $I_l I_{\text{lat}} \ll I_a I_{\text{aff}}$ or $I_a I_{\text{aff}} \gg I_l I_{\text{lat}}$ the performance decreases. However, for $I_a I_{\text{aff}} \ll I_l I_{\text{lat}}$ the afferent input can be neglected and Eq. 4.12 approximates Eq. 4.17. Hence for $K = 1$ the multiplicative model is equivalent to an additive model with large lateral input and performs worse than the additive model with balanced inputs.

Of the two options just discussed, the biologically more plausible implementation of orientation columns is certainly a model with neuronal columns with

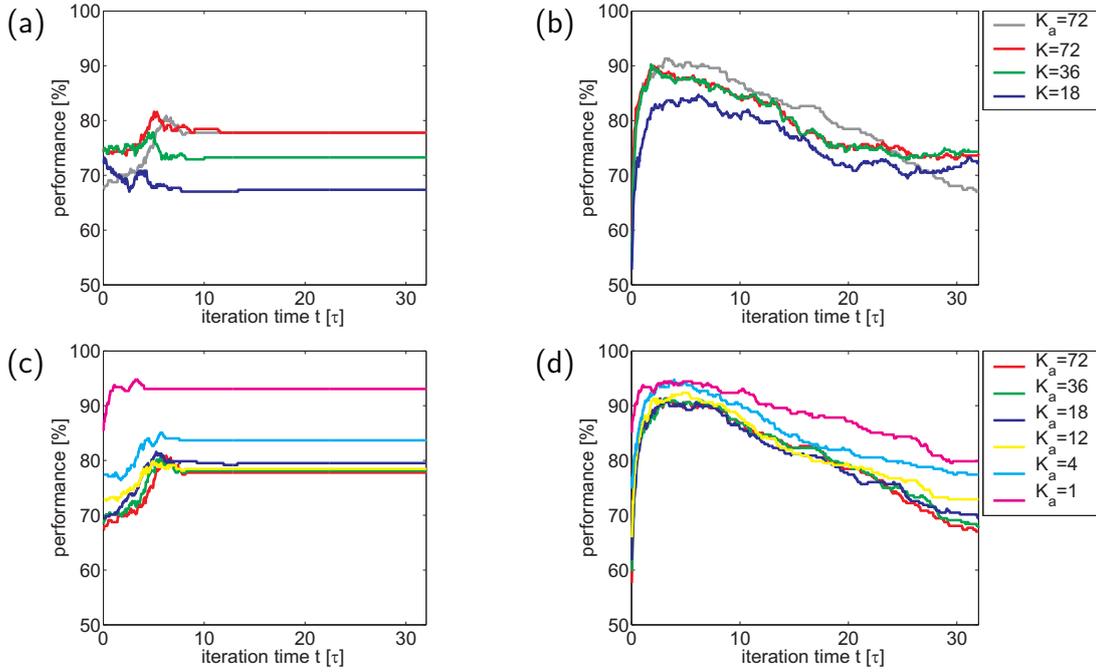


Figure 4.5: The colored lines show the performance of (a) the additive model and (b) the multiplicative model for different numbers K of equidistributed preferred orientations ϕ_k . The ϕ_k are the same for each hypercolumn independently of the stimulus presented. For comparison the grey line shows the model performance for $K_a = 72$ preferred orientations when assuming that the preferred orientations ϕ_k are adapted such that one preferred orientation ϕ_k of each hypercolumn exactly matches the orientation θ of the Gabor element presented in the receptive field of this hypercolumn. The lower diagrams depict the performance for (c) the additive model and (d) the multiplicative model for different numbers K_a of preferred orientations employing these adapted orientation preferences. All performances shown are for stimuli of jitter $\sigma_\alpha = \sigma_\beta = 13.5$ degree and element distance of 3.5 degree visual angle. Simulations were performed with a unidirectional association field of the same width as the generating association field and saliences estimated with $L_{\max} = 3$.

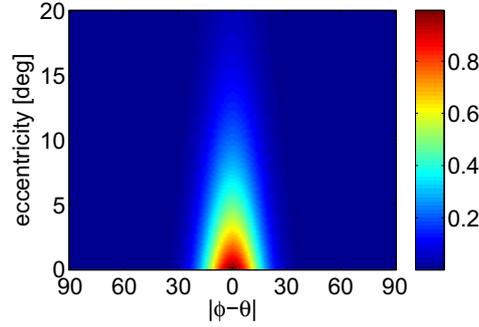


Figure 4.6: Eccentricity dependence of the afferent input as described by Eq. 4.18. The width of the afferent input is $\sigma_{\text{aff}} = \pi/8$ and the wavelength $\lambda = 0.36$ degree visual angle.

fixed preferred orientations which are independent of the stimulus and where K is sufficiently high to avoid sampling problems. The models with adapted preferred orientations have its limitations for $K_a = 1$. However, the performance for $K_a \geq 18$ seems not to depend on the exact K_a any more and in this range of K_a it approximates the performance of the stimulus-independent model with large K far better, than the stimulus independent model with small K .

Modulation by stimulus eccentricity

So far our afferent input is independent of the eccentricity of a stimulus. However, as already discussed in 2.3.2, the size of the area of cortical surface onto which an object in the visual field is mapped decreases with eccentricity. With the decrease of this area of cortical surface the visual acuity decreases (15; 20) and the contrast threshold increases (107). In a recent study Foley and colleagues (30) measured contrast thresholds of Gabor patterns as a function of eccentricity and obtained the eccentricity parameter α , which is the factor by which the sensitivity of a receptive field decreases with eccentricity ϵ for a given wavelength λ of the Gabor pattern. They modeled sensitivity of the receptive field as well as excitation of the corresponding cells to be proportional to $\alpha^{\epsilon/\lambda}$ where ϵ is the eccentricity of the center of the receptive field. We used this factor to model the decrease of afferent input with eccentricity.

$$I_{\text{aff}}(\theta, \phi, \epsilon) = I_{\text{aff}}(\theta, \phi) \alpha^{\epsilon/\lambda} \quad (4.18)$$

Here the eccentricity parameter was $\alpha = 0.947$. The decrease of afferent input with eccentricity as described by Eq. 4.18 is illustrated in Fig. 4.6.

4.3.2 Lateral input

The lateral input is given by

$$I_{\text{lat}}(x_l, \phi_l, x_1 \dots x_N, \phi_1 \dots \phi_N) = \sum_{i,j}^N p(x_l, \phi_l | x_i, \phi_j) A(x_i, \phi_j) \quad (4.19)$$

and depends crucially on the connection strength $p(x_l, \phi_l | x_i, \phi_j)$ between neuronal orientation columns which is specified by the association field.

Association fields for contour integration

The models are supposed to judge the same edge configurations as being salient as humans do. Hence the association field should be based on the Gestalt rules. The Gestalt criterion of proximity states that edge elements at small distances are more likely to be perceived as belonging to the same object than distant edge elements. Consequently, there should be a distance-dependent factor $F(r)$ in the association field.

According to the Gestalt law of good continuation collinearly aligned edge elements should have a high probability to be linked together. Connection strength should then decrease for higher curvature or alignment jitter.

A contour has no direction. Hence the probability $p(x_j, \phi_j | x_i, \phi_i)$ that a contour which passes an element at position x_i with orientation ϕ_i next passes an edge element at location x_j with orientation ϕ_j should be the same as the probability $p(x_i, \phi_i + \pi | x_j, \phi_j + \pi)$ to pass these two contour elements in reversed order and in the opposite direction. An association field which fulfills these requirements is

$$p(r, \alpha, \beta) = F(r) \cdot \left[M(\gamma_\alpha, 0, 1/\sigma_\alpha^2) M(\gamma_\beta, 0, 1/\sigma_\beta^2) + M(\gamma_\alpha, \pi, 1/\sigma_\alpha^2) M(\gamma_\beta, \pi, 1/\sigma_\beta^2) \right] \quad (4.20)$$

where $F(r)$ describes the dependence on the inter-element distance r . $\gamma_\alpha = \beta/2 - \alpha$ specifies the alignment and $\gamma_\beta = \beta/2$ the curvature of two elements, where α is the viewing angle from the source edge to the destination edge and β is the difference of the two edge orientations (see Fig. 4.7 (a)). Assuming the source edge has location $x_i = (x_{i,1}, x_{i,2})$ and orientation ϕ_i and the destination edge is located at $x_j = (x_{j,1}, x_{j,2})$ and has orientation ϕ_j , we can express r as $r(x_i, x_j) = \sqrt{(x_i - x_j)^2}$, α as $\alpha(x_i, x_j, \phi_i) = \arctan\left(\frac{x_{j,1} - x_{i,1}}{x_{j,2} - x_{i,2}}\right) - \phi_i$ and β as $\beta(\phi_i, \phi_j) = \phi_i - \phi_j$. With this we can rewrite Eq. 4.20 in terms of element positions and orientation. Then the transition probability from edge i to edge j

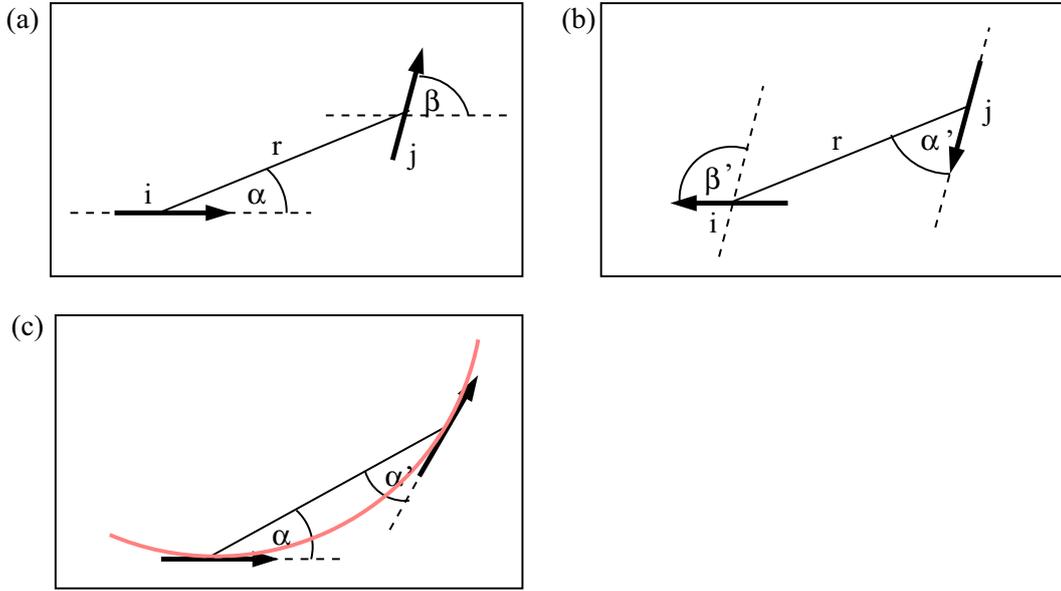


Figure 4.7: (a) Curvature and alignment for a contour passing from edge i to edge j can be expressed in terms of the viewing angle α from source edge i to destination edge j and $\beta = \phi_i - \phi_j$. (b) Same as (a) for a contour passing from j to i . (c) Two edges are perfectly aligned, if they are tangents of a circle passing through the edges, i.e. $\alpha = \alpha'$. Figure modified from (94)

is given by

$$\begin{aligned}
 p(x_j, \phi_j | x_i, \phi_i) &= F(x_i - x_j) \cdot & (4.21) \\
 &\left[M\left(\frac{\phi_i + \phi_j}{2} - \arctan\left(\frac{x_{j,1} - x_{i,1}}{x_{j,2} - x_{i,2}}\right), 0, 1/\sigma_\alpha^2\right) M(\phi_i - \phi_j, 0, 1/\sigma_\beta^2) + \right. \\
 &\left. M\left(\frac{\phi_i + \phi_j}{2} - \arctan\left(\frac{x_{j,1} - x_{i,1}}{x_{j,2} - x_{i,2}}\right), \pi, 1/\sigma_\alpha^2\right) M(\phi_i - \phi_j, \pi, 1/\sigma_\beta^2) \right]
 \end{aligned}$$

The two notations are equivalent and throughout this thesis we will switch between them where appropriate.

We get the same link probability between edge i and j as in the opposite direction from j to i , $p(r, \alpha, \beta) = p(r, \alpha', \beta')$ (see Fig. 4.7 (b) for the definition of α' and β'). Two edges are aligned, if they are tangents of a circle passing through these edges, i.e. $\alpha = \alpha'$ (see Fig. 4.7(c))(94).

Symmetry of the association field

Eq. 4.20 describes a unidirectional association field. In contrast to this choice, many models for contour integration in striate cortex use a bidirectional association field where connections to both directions in retinal space have equal

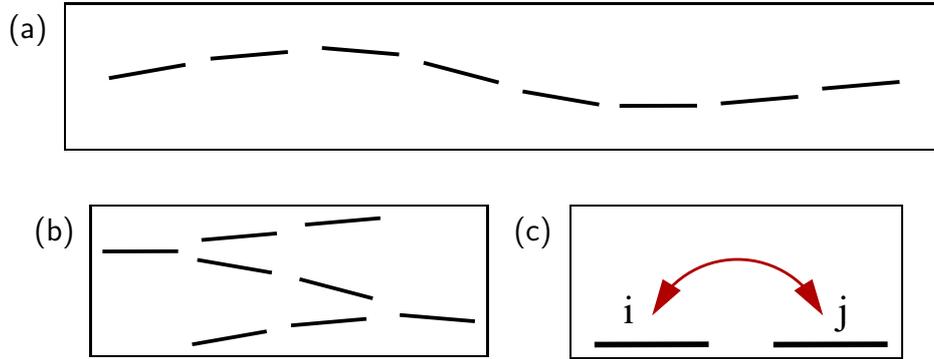


Figure 4.8: The contour in (a) can be generated by a unidirectional or by a bidirectional association field. Contours in (b) and (c) are equally likely to be generated by a bidirectional association field, but cannot occur when using a unidirectional association field.

strength (e.g. (77; 137)). We use association fields for two purposes: First, we generate mathematically well-defined contours by successively drawing the contour elements from an association field in a Markov process. This means that the probability of a contour element at position x_i with orientation ϕ_i depends only on the position and orientation of the preceding contour element (x_{i-1}, ϕ_{i-1}) . Second, we invert the process of contour generation to extract a contour from a given stimulus. Therefore we use the association field to determine the probability of each edge element to belong to the contour. This probability corresponds then to the saliency of the contour.

To generate contours we use a unidirectional association field. This gives long and, depending on the width σ_β of the association field, relatively straight contours like the one shown in Fig. 4.8 (a). However, when generating contours using a bidirectional association field, a zigzagged contour like in Fig. 4.8 (b) is produced with equal probability. In fact there are more possibilities to produce contours with kinks than to generate straight contours. Furthermore, if we generate a contour of $L \gg 2$ elements, we can get the edge configuration shown in Fig. 4.8 (c). Starting with element i it is very likely to draw element j as the next contour element. Then we have a high probability to go back to element i and again a high probability to go to element j next. By going back and forth we can generate infinitely long contours which consist only of two different elements and which humans would perceive only as a contour of length two. Because humans perceive long straight contours as much more salient than contours with an abrupt change in direction or short contours, we always used the unidirectional association field for contour generation. For a probabilistic model one had to use the same association field for contour generation and contour

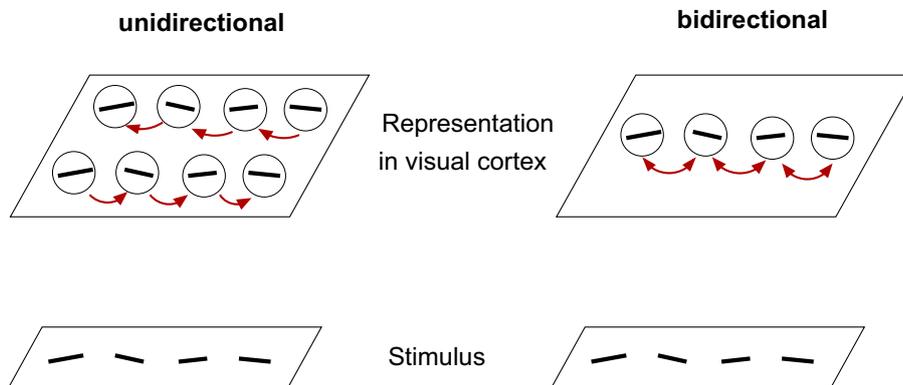


Figure 4.9: Possible implementations of unidirectional and bidirectional association fields: While for bidirectional interactions one representation of the contour in the visual cortex is sufficient one would assume two separate representations when using a unidirectional association field. In one representation edge detectors are linked only in one direction, in the other representation interaction go solely in the opposite direction.

integration. As there is no experimental evidence for unidirectional connections between neuronal orientation columns, we also tried to use bidirectional association fields for contour integration. A possible implementation of unidirectional and bidirectional association fields is depicted in Fig. 4.9. As already mentioned we require the probability to pass a contour in one direction to be the same as for passing the contour elements in reverse order in opposite direction. For bidirectional association fields a single representation of each edge element would be sufficient to fulfill this symmetry requirement as interactions between two edge elements could go in both directions. In contrast, two separate representations of each edge element are necessary when using a unidirectional association field. In one of the representations for the contour elements activity could propagate into one direction, while in the other it propagates in the opposite direction.

Association field geometries

A similar conflict between a probabilistic approach and the biophysically most plausible approach exists for the question whether we need different association fields for contours with different statistical properties. A mathematically ideal model would require to use the same association field that was applied for contour generation also for contour detection. This would mean that we need different association fields for contours with different curvature, alignment, or element distances. From a biological point of view it is rather unlikely that the brain changes its functional connectivity structure for contours of varying statistical properties. And even if it were able to adapt its connection structure, it first needed to know

the characteristics of the stimulus before it could choose a suitable connection structure. Hence we will investigate both possibilities: contour integration with a fixed width for curvature and alignment jitter as well as contour integration with an association field using the same σ_α and σ_β as the contour generating process. Where simulations were directly compared to human behavior, we always used an association field of the same range for all contours independently of their inter-element distance.

4.3.3 Relations between different forms and couplings of inputs

So far we posed the question which of the models, the multiplicative or the additive one would be better. However, these two are not the only models one can think of and the model best suitable to describe human behavior could also be a mixture of these two models. As we will see in 5.2 it can be advantageous for purely multiplicative models with unidirectional association field to introduce an offset c_l for the lateral input. In this way $I_{lat} > 0$ for all edge elements and hence the activity A will not drop to zero due to edge detectors which receive or spread zero lateral input. Biologically such an offset can be interpreted as spontaneous activity. However, introducing this offset is nothing else as mixing multiplication and addition of the two inputs. A multiplicative coupled input $I = I_{aff} \cdot \tilde{I}_{lat}$ with $\tilde{I}_{lat} = I_{lat} + c_l$ is equivalent to $I = I_{lat} \cdot I_{aff} + c_l \cdot I_{aff}$. Similarly one could think of spontaneous activity for the afferent input $\tilde{I}_{aff} = I_{aff} + c_a$. Then $I = \tilde{I}_{aff} \cdot \tilde{I}_{lat} = c_l \cdot I_{aff} + c_a \cdot I_{aff} + c_l c_a + I_{aff} \cdot I_{lat}$. Hence if we use the most general input $I = I_a I_{aff} + I_l I_{lat} + I_m I_{aff} I_{lat} + I_0$, all offsets for lateral and afferent input are already included. This is true provided the offset is added to each neuron and not e.g. only to the lateral input for the nearest neighbors. In principle we can also skip the last term. As I_0 is added to each neuron and the decision of the model depends only on the relative values and not on the absolute values, this does not interfere with the decision for the neuron with the maximum activity, at least not if there is no noise included in the model.

4.3.4 Summary

In this chapter we introduced a Bayesian model which from a mathematically point of view is optimal for contour integration. From this we derived an iterative model which can evaluate the afferent input from the visual stimuli against the lateral input from horizontal interactions in various ways. In the following chapters we will investigate the two generic cases of multiplicatively bound afferent and lateral input and a total neuronal input obtained by the summation of all inputs. Throughout this thesis we will discuss advantages and disadvantages

Chapter 4. Theory: model classes and their dynamics ---

of these two cases and compare them to psychophysical experiments. Similarly we analytically investigate special properties of the two different estimators we introduced in this chapter. In chapter 7 we will test the different possible choices for the afferent input and different symmetries and geometries of the association field. By analyzing the model performance compared to the human performance and by analyzing the correlations between the model and human observers we will evaluate different model and association field choices and search for the model which is closest to human behavior.

Chapter 5

Analysis of model dynamics

Large simulations can take into account many different influences on a dynamical system, but in the end it can be complicated to identify what effect is caused by which mechanism. Sometimes one can learn more about a complex system by simplifying it such that one can investigate basic principles in great detail. Also simplification often allows the application of analytic methods which are difficult to use otherwise. However, one should be aware that for simplified systems one might miss many phenomena.

Having introduced the different models and salience estimators in chapter 4 we now want to investigate some of the basic dynamic properties of these models subjected to simple stimuli. Later we compare the different models to psychophysical experiments, where more complex stimuli are used. In particular, this chapter will tackle questions like the effect of normalization and the parameter range for a stable fixed point. Furthermore we will see that under certain circumstances the representation of an edge element which is not present in the stimulus is predicted by the additive model to have non-zero salience, as e.g. for star-like edge configurations which sometimes are predicted to be more salient than a contour. In addition we investigate what happens if the balance between afferent and lateral input is changed.

5.1 Stability of the iterative models and effect of normalization

We investigate the simplified stimulus depicted in Fig. 5.1 assuming periodic boundary conditions. Accordingly, the lower two edge elements form an infinitely long contour. The upper two edges are considered as background elements. We also assume that background elements and contour elements are of different contrast C_b and C_c , respectively. In the simplified model we have neurons with only

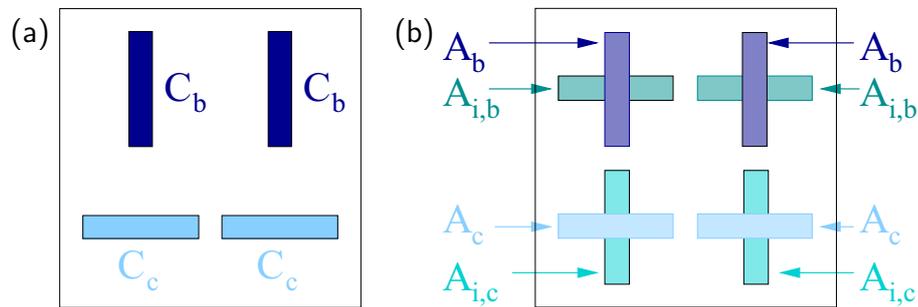


Figure 5.1: (a) Simplified stimulus with periodic boundary conditions containing a horizontal contour. Contour and background elements have different contrast C_c and C_b , respectively. (b) Neuronal representation of the stimulus situation depicted in (a). Each hypercolumn has neurons with two different preferred orientations. A_c is the activity of the neuron representing the contour element, $A_{i,c}$ the activity for the element at the same position but perpendicular to the contour element. A_b and $A_{i,b}$ are the corresponding activities in the hypercolumns representing the background.

two different preferred orientations: horizontal and vertical. The association field connects only perfectly aligned, neighboring edges with $p_{i,j} = I_l$.

In our model the saliency of an element corresponds to the activity of the neuron representing this edge element. As the stimulus contains an infinitely long contour the model is supposed to predict the highest saliency for the contour elements.

This means, that

- (a) the activity A_c of the contour elements should be higher than the activity A_b of the background elements and
- (b) the activities A_c and A_b of the neurons representing the elements which are actually present should exceed the activities $A_{i,c}$ and $A_{i,b}$ of the corresponding neurons with preferred orientations perpendicular to the bars presented in their receptive field.

It can happen that the model predicts the highest saliency for an edge element which is not present in the stimulus. This can be an edge at the same position x_i as an edge i which is depicted in the stimulus, but with an orientation $\theta \neq \theta_i$ or it can be an edge element at a position where no edge is present in the stimulus. In these cases we speak of illusions and accordingly we call elements which are not physically present in the stimulus but which activate an orientation column representing this edge element, 'illusory elements'. The fact that a model finds illusory elements is not necessarily bad, as humans also perceive certain illusions,

5.1. Stability of the iterative models and effect of normalization

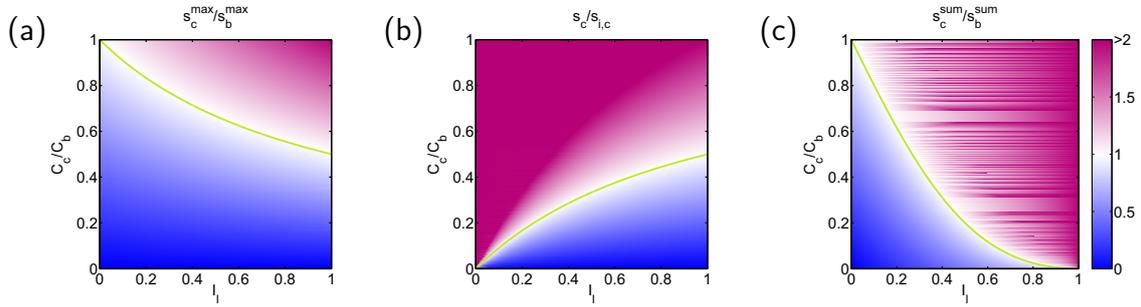


Figure 5.2: The three pictures show the salience ratios discussed in the text. The green line denotes a salience ratio as specified by the subfigure title equal to 1. In the purple area above the green line the model can detect (a) contour elements in contrast to background elements predicted by the maximum estimator (b) contour elements opposed to illusory elements (c) contour elements in contrast to background elements as predicted by the summation estimator.

i.e. the filling in of the borders of the Kanizsa triangle (see Fig. 3.2) where there are no physical boundaries present. However, the salience of illusory contours is generally lower than that of a contour which is physically present. We want to investigate which illusions the different models are sensitive to. In comparison with psychophysics these results can give hints which models are plausible.

Additive model without normalization

One potential difficulty for the dynamical systems given by Eq. 4.12 is that, depending on the employed parameter range and on the stimuli subjected to this systems, they do not necessarily have a stable stationary point, if we omit the global divisive normalization by setting the normalization factor $\mathfrak{n} = 1$. So here we want to determine the stable regime for both models without normalization. Let us start with the additive model and carry out a linear stability analysis around the fixed point. We find a stable fixed point for $0 < I_1 < 1$.

We calculate the actual stationary state A_s by solving for $\dot{A} = 0$ and build the ratio A_c/A_b

$$\frac{A_c}{A_b} = \frac{C_c}{C_b}(I_1 + 1) \quad (5.1)$$

This relation is plotted in Fig. 5.2. When using the maximum estimator, the activity corresponds directly to the saliency s^{\max} and if $A_c/A_b = s_c^{\max}/s_b^{\max} > 1$, the contour is detected correctly. $A_c/A_b = 1$ is denoted by the green line in Fig. 5.2 (a) and in the purple region above this line the contour position is detected by the maximum estimator.

Furthermore we want to know under which circumstances the model finds illusory elements. For this stimulus the neurons with preferred orientation perpendicular to the background elements get no input at all. Hence $A_{i,b} = 0$ for all I_l and all C_c/C_b . This means that the model finds no illusory elements at the positions of the background elements. However, this is not so simple for the positions occupied by the contour elements. Neurons representing elements perpendicular to the contour elements get lateral input from the cells representing background elements. Depending on the strength I_l of the lateral input and the contrast C_c and C_b it can indeed happen, that the model finds illusory elements instead of contour elements. This means it predicts a higher activity for neurons representing elements at the position of the contour elements which are not present in the stimulus than it determines for cells representing contour elements. This shows the ratio

$$\frac{A_c}{A_{i,c}} = \frac{C_c(1 + I_l)}{C_b I_l} \quad (5.2)$$

which is illustrated in Fig. 5.2 (b). The green line shows where real elements and illusory contour elements are equally salient. Here we see that illusions occur for high overall coupling strength and high contrast C_b , while in the purple region above the green line the model does not exhibit any illusions. Furthermore, if we compare with Fig. 5.2 (a), it also finds the contour in this region. Hence this is the region where for this particular stimulus the model does what it is supposed to do. This changes with different stimuli.

Next we want to investigate the behavior of the summation estimator. As the activity of neurons representing illusory elements in this case strengthen the salience of the position of the contour in contrast to the background, we would expect that the summation estimator finds the contour in a larger parameter range. To check this we calculate the ratio

$$\frac{s_c^{\text{sum}}}{s_b^{\text{sum}}} = \frac{A_c + A_{i,c}}{A_b + A_{i,b}} = \frac{C_c}{C_b}(1 + I_l) + I_l = \frac{s_c^{\text{max}}}{s_b^{\text{max}}} + I_l \quad (5.3)$$

and plot it in Fig. 5.2 (c). As expected we find a larger area where the model detects the contour. This estimator only finds a position, without distinguishing between the contour element and the corresponding illusory element. Comparing Fig. 5.2 (a) and (c) we see that the summation estimator predicts a higher salience for the contour elements than for the background elements in a much larger parameter range than the maximum estimator does. Also $\frac{s_c^{\text{sum}}}{s_b^{\text{sum}}}$ exceeds $\frac{s_c^{\text{max}}}{s_b^{\text{max}}}$ especially for high I_l . This is consistent with finding illusory elements supporting the position of the contour especially for a large strength I_l of the lateral input.

5.1. Stability of the iterative models and effect of normalization

with normalization	without normalization
$\frac{s_c^{max}}{s_b^{max}} = \frac{C_c}{C_b} \left(1 + \frac{I_l}{k}\right)$	$\frac{s_c^{max}}{s_b^{max}} = \frac{C_c}{C_b} (1 + I_l)$
$\frac{s_c}{s_{i,c}} = \frac{C_c}{C_b} \left(1 + \frac{k}{I_l}\right)$	$\frac{s_c}{s_{i,c}} = \frac{C_c}{C_b} \left(1 + \frac{1}{I_l}\right)$
$\frac{s_c^{sum}}{s_b^{sum}} = \frac{C_c}{C_b} \left(\frac{I_l}{k} + 1\right) + \frac{I_l}{k}$	$\frac{s_c^{sum}}{s_b^{sum}} = \frac{C_c}{C_b} (I_l + 1) + I_l$

Table 5.1: Saliency ratios for the additive model with and without normalization.

Additive model with normalization

Having investigated the additive model without normalization, we are now interested in the changes which occur if we reintroduce the normalization $\mathbf{n}(\mathbf{t}) = \left(\sum_{\mathbf{x}, \phi} I(\mathbf{x}, \phi)\right)^{-1}$ in Eq. 4.12. \mathbf{n} prevents the neuronal activity to grow to infinity and it is chosen such that $\sum \mathbf{A}(\mathbf{i}, \mathbf{t}) = 1 \ \forall \mathbf{t}$. Because of this stabilization we find for all values of I_l a stable stationary point for this stimulus. In order to study the behavior of the model in this range, we compute the same three ratios as already done in 5.1 for the model without normalization.

These ratios are presented in Table 5.1, which gives an overview over the saliency ratios for the models with and without normalization. Here $k = \sum_{\mathbf{x}, \phi} I(\mathbf{x}, \phi)$ where $I(\mathbf{x}, \phi)$ is the total neuronal input as defined in Eq. 4.13. k is a constant consisting of two parts: $\sum_{\mathbf{x}, \phi} I_{aff}(\mathbf{x}, \phi)$ is constant for every stimulus which does not change with time and $\sum_{\mathbf{x}, \phi} I_{lat}$ is constant for this edge configuration as here $\sum_{\mathbf{x}, \phi} \mathbf{P} \cdot \mathbf{A}(\mathbf{x}, \phi) = \sum_{\mathbf{x}, \phi} \mathbf{A}(\mathbf{x}, \phi) = 1$. Here the matrix $\mathbf{P} = \{p_{i,j}\}$ denotes the link probability between different edge elements. If we also normalize the afferent input to $\sum_{\mathbf{x}, \phi} I_{aff}(\mathbf{x}, \phi) = N$, where N is the total number of edge elements within the stimulus, the saliency ratios depend only on I_l/I_a and on C_c/C_b . The three ratios are shown in Fig. 5.3. Here the green lines indicate saliency ratios equal to one. The dashed black line shows the same for the additive model without normalization to allow for a better comparison. For the x-axis of the dashed black line we assumed $I_l/I_a = I_l$. However, when taking the limit $I_l/I_a \rightarrow \infty$ we find for all three saliency ratios the same values as for the model without normalization for $I_l \rightarrow 1$. So the behavior at the borders of the parameter range leading to a stable fixed point is the same. Hence the main effect of the normalization is to stabilize the system on a much broader range than without normalization without changing the qualitative model behavior which is possible in this range.

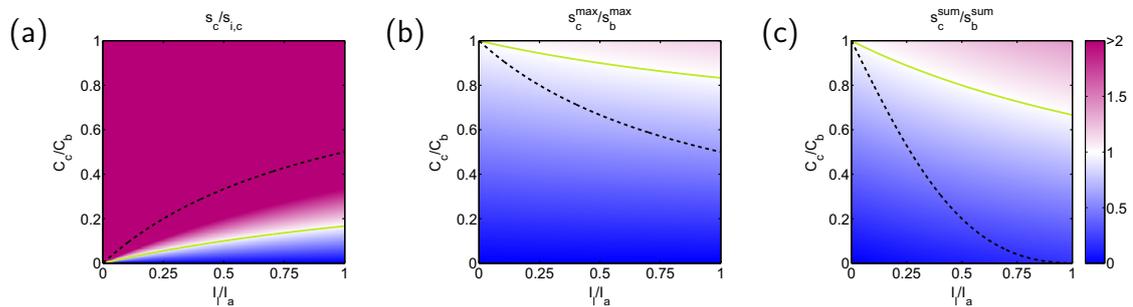


Figure 5.3: The three pictures illustrate the investigated salience ratios as specified by the subfigure titles for the additive model with normalization. The green line denotes a salience ratio equal to 1. In the purple area above the green line the model can detect (a) contour elements in contrast to background elements predicted by the maximum estimator, (b) contour elements opposed to illusory elements and (c) contour elements in contrast to background elements with the summation estimator.

5.1.1 Illusions supporting the background

For the stimulus investigated so far the activity of neurons representing illusory elements help the summation estimator to detect the contour. However, we can think of different stimuli where illusory elements have the opposite effect and are in rivalry to the contour. One such stimulus is depicted in Fig. 5.4 (a).

Again we assume periodic boundary conditions, two preferred orientations in each hypercolumn and interaction of perfectly aligned neighboring edge elements only. The contour elements have contrast C_c and form an infinitely long contour along the diagonal due to the periodic boundary conditions. The crucial difference to the stimulus in Fig. 5.1 is that here the illusions strengthen the salience at the positions occupied by background elements. Each neuronal unit corresponding to an illusory element at the position of a background element gets lateral input from two neighboring units, like the orientation columns corresponding to contour elements get lateral input from two neighboring columns. Like before we determine the saliency ratios for the additive model with and without normalization. They are presented in Table 5.2.

The ratios for the normalized model are plotted in Fig. 5.6. Even though the background elements are more salient than the corresponding illusory elements for the full parameter range (see Fig. 5.6 (a)), the parameter range where the summation estimator finds the contour is reduced compared to the parameter range, in which the maximum estimator finds the contour. The reason for this is that the neuronal units corresponding to illusory background elements get lateral input from two neighboring edge detectors. As before, the activation of

— 5.1. Stability of the iterative models and effect of normalization

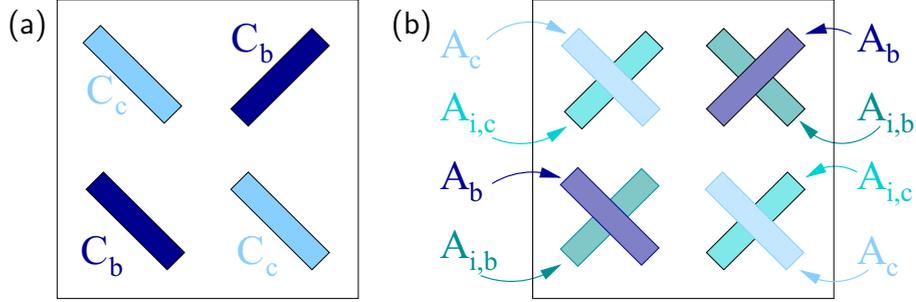


Figure 5.4: (a) Simple stimulus where illusory elements support the background. Periodic boundary conditions are assumed. Contour and background elements have contrast C_c and C_b , respectively. (b) Neuronal representation of the stimulus. Hypercolumns contain two preferred orientations for each position. A_c and A_b denote the activity of the neurons representing the contour or background elements respectively. $A_{i,c}$ and $A_{i,b}$ are the activities for the neurons representing illusory elements not present in the stimulus.

with normalization	without normalization
$\frac{s_c^{max}}{s_b^{max}} = \frac{C_c}{C_b} \frac{k+I_l}{I_l}$	$\frac{s_c^{max}}{s_b^{max}} = \frac{C_c}{C_b} \frac{1+I_l}{I_l}$
$\frac{s_c}{s_{i,c}} = \frac{I_l}{k}$	$\frac{s_c}{s_{i,c}} = I_l$
$\frac{s_c^{sum}}{s_b^{sum}} = \frac{C_c}{C_b}$	$\frac{s_c^{sum}}{s_b^{sum}} = \frac{C_c}{C_b}$

Table 5.2: Saliency ratios for the additive model with and without normalization.

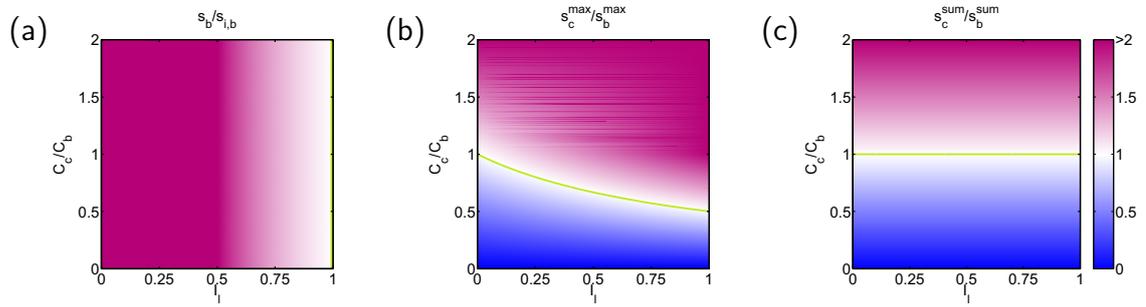


Figure 5.5: Saliency ratios in the stationary state of the additive model without normalization. The green line corresponds to a saliency ratio equal to one.

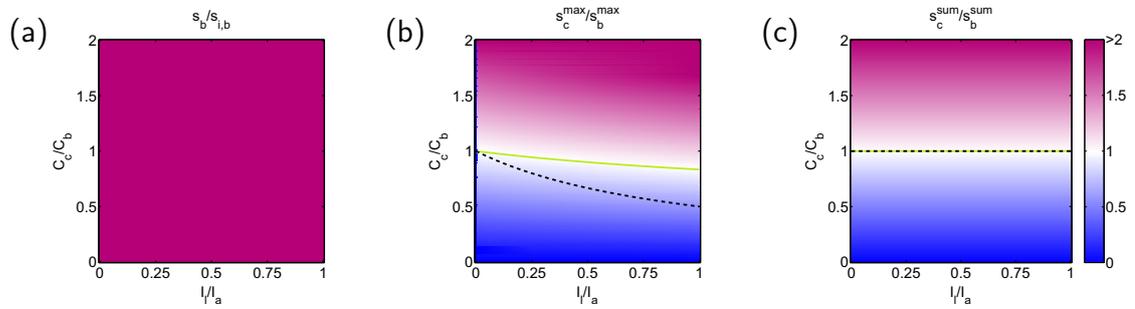


Figure 5.6: Saliency ratios in the stationary state for the additive model with normalization. The green line corresponds to a saliency ratio equal to one. The black dashed line shows this ratio for the additive model without normalization.

edge detectors representing elements which are absent in the stimulus is stronger for higher I_1 . Normalization scales the saliency ratios from the stable I_1 -regime $0 \leq I_1 \leq 1$ to the full possible range of I_1 , namely $0 \leq I_1 \leq \infty$.

Wider tuning curves for input functions

In the simplified model investigated so far, each edge element can get lateral input from at most two neighboring elements. However, in a plausible model we have much more than just two possible orientations and neither the afferent input nor the association field is δ -shaped. Hence we would expect more complex effects. However, for a complex model with n_{ori} preferred orientations, afferent input tuning width σ_{aff} , and association field curvature and alignment tuning $\sigma_{\alpha,\beta}$ it would be difficult to calculate the stationary point analytically. Hence in order to investigate the influence of more possible orientations and wider input functions we restrict ourselves $n_{\text{ori}} = 4$. We use the same stimulus as in Fig. 5.4,

5.1. Stability of the iterative models and effect of normalization

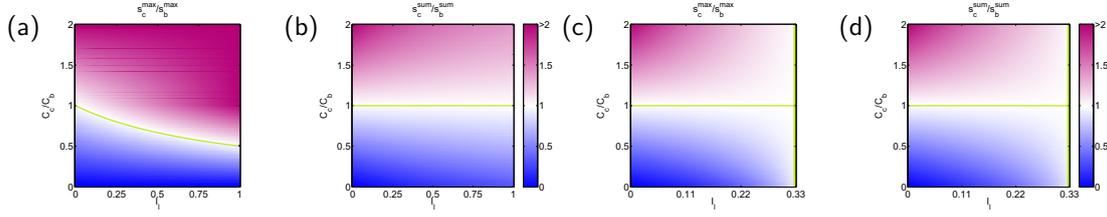


Figure 5.7: Saliency ratios in the stationary state for the additive model with (a+b) $I_{\text{aff}}^{\text{wide}}$ and δ -shaped association field. (c+d) The same for wide afferent input and wide association field. The green line corresponds to an activity ratio equal to one.

but modify the afferent input to

$$I_{\text{aff}}^{\text{wide}}(\phi, \theta) = \begin{cases} C/2 & \text{for } |\theta - \phi| = 0 \\ C/4 & \text{for } |\theta - \phi| = \pi/4 \\ 0 & \text{else} \end{cases}$$

where C is the contrast of the edge element. A wider afferent input gives input to more neurons representing edge elements which are not present in the stimulus. However, this is also the case for cortical neurons which also have a finite tuning widths and do not only respond to stimuli in their preferred orientation. Just using a wider afferent input and allowing for more possible orientations leaves the parameter range in which we find a stable attractor unchanged. There are more possibilities for illusions as there are more illusory elements now. However, as long as only neurons representing perfectly aligned elements interact with each other, we have basically four mutually independent systems of edge detectors, which have no impact on the other systems: Neurons with a horizontal preferred orientation interact only with other neurons of horizontal preferred orientation. Equally we have a system of vertical edges, one of right oblique edges and one of left oblique edges. As those neurons, which get maximum afferent input are the same, which also would be activated when assuming a δ -shaped tuning curve, $s_c^{\text{max}}/s_b^{\text{max}}$ stays unchanged (see Fig. 5.7 (a)).

The summation estimator still detects the contour in the same cases as before. However, due to the additional illusory elements the gradient $\frac{d(s_c^{\text{sum}}/s_b^{\text{sum}})}{d(C_c/C_b)}$ near $s_c^{\text{sum}}/s_b^{\text{sum}} = 1$ is much lower than with a sharply tuned afferent input, especially for large I_l . Hence noise is expected to have a much bigger impact on making contour detection more difficult.

Choosing a broader tuned association field couples the previously separated subsystems. In our simplified example such an association field could be

$$p^{\text{curv}}(i, j) = \begin{cases} 1/2 & \text{for } \alpha = \alpha' = 0 \\ 0 & \text{for } |\phi_1 - \phi_2| = \pi/4, 3\pi/4 \\ 1/4 & \text{for } \alpha = \alpha' \text{ and } |\phi_1 - \phi_2| = \pi/2 \end{cases}$$

where the angles α and α' are defined as the viewing angles under which the destination edge is seen from the originating edge, and vice versa (see Fig. 4.7). ϕ_i is the preferred orientation of neuron i . This association field demands perfect alignment but allows higher curvature between edge elements that are to be linked.

This modified association field narrows the parameter range where we find a stable fixed point for the system. Now this is only the case for $0 \leq I_1 \leq 1/3$. As there is now a link probability greater than zero between contour and background elements, the parameter range in which the maximum estimator finds the contour decreases. For both maximum and summation estimator we find a small gradient $\frac{d(s_c^{sum}/s_b^{sum})}{d(C_c/C_b)}$ near $s_c^{sum}/s_b^{sum} = 1$ which in combination with noise makes it very difficult to detect the contour in this parameter range (see Fig. 5.7 (c) and (d)).

In general, we observe that the impact of elements which are not present in the stimulus increases with n_{ori} and with broader tuning curves for afferent input and association field. The width of the association field has a bigger impact than the width of the afferent input, as the association field determines whether the differential equations for our model are mutually coupled or can be separated into isolated subsystems. The association field used in this example might seem inappropriate, as it links edge detectors with $\beta = \pi/2$ while in the human visual system we actually find inhibition for perpendicular edge elements. However, in this simplified model there is no possibility to allow a smaller non-zero curvature jitter. And even though the analyzed association field has little similarity with the association field that might be implemented in the human visual system, it shows the general effect that illusions become more influential for wider association fields. This finding should also be valid for a biologically plausible association field.

The multiplicative model

In contrast to the additive model, the multiplicative model only has a stable stationary point when it is normalized. However, the normalized model detects the contour for both stimuli. It judges all elements of the infinitely long contour equally salient while it predicts zero saliency for the background elements (see Fig. 5.8). Furthermore the saliency of all non-present elements i vanishes, as their total input I is proportional to $I_{aff}(x_i, \phi_i) = 0$.

5.1.2 Star-Illusions

Having seen in detail how illusions emerge in different model classes, we can now construct a stimulus trying to maximize the effect that can be caused by such illusions. Such a stimulus is depicted in Fig. 5.9 (a). The longest contour within

— 5.1. Stability of the iterative models and effect of normalization

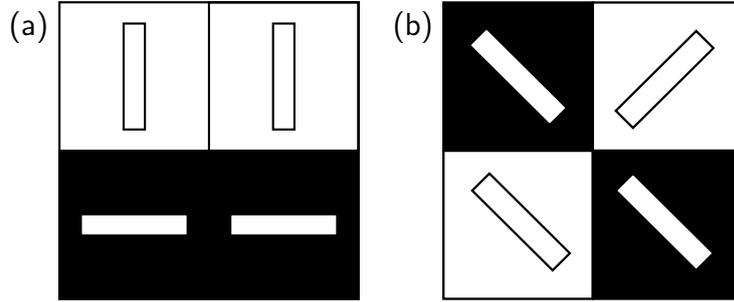


Figure 5.8: (a) and (b) show the saliency as predicted by the normalized multiplicative model for the two stimuli investigated before for the additive model. In the above diagram the white bars with black borders represent the stimulus and the color of the background demonstrates the saliency of that edge element. Black corresponds to maximum saliency and white represents zero saliency. Here both estimators predict identical saliences s^{\max} and s^{sum} .

this stimulus (contour A) has three perfectly aligned elements and should thus be the most salient contour within the stimulus. However, the background is arranged such that only one background element is directed towards the center of a neighboring contour element. Hence few illusory elements supporting contour A are expected. In addition, there are several contours of length two. One of these shorter contours includes an edge element $i = (x_i, \theta_i)$ that has 7 out of 8 neighboring elements oriented such that they are directed towards the center x_i of edge i . We assume neurons of $n_{\text{ori}} = 4$ possible preferred orientations ϕ_j with $j = 1 \dots n_{\text{ori}}$ at each position x . The afferent input is chosen as $I_{\text{aff}}(\phi_i, \theta_j) = I_a$ for $\theta_j = \phi_i$, and 0 otherwise. The association field is delta-shaped with $p^\delta(\alpha, \beta) = 1$ for $\alpha = \beta = 0$ and $p^\delta(\alpha, \beta) = 0$ otherwise. Here α is the viewing angle under which the target edge i_2 is seen from the source edge i_1 , β is the difference of preferred orientations $\phi_2 - \phi_1$ and the range of $p^\delta(\alpha, \beta)$ is limited to nearest neighbors. For an illustration of α and β see Fig. 4.7.

Fig. 5.9 (b) and (c) show the saliences s^{sum} and s^{\max} as predicted by the additive model. As desired from a good contour integration model the maximum estimator finds the highest saliency for the three elements that belong to contour A. In contrast the summation estimator predicts the highest saliency for edge i , the center of a star-like edge configuration. The multiplicative model with p^δ shows a saliency greater zero only for the three contour elements. In the multiplicative model $s^{\max} = s^{\text{sum}}$ because only the neurons representing the orientations $\vec{\phi}$ that are present in the stimulus receive a total input $I \neq 0$.

Why does the summation estimator applied to the additive model find this high saliency for edge i , despite it is not part of the longest and hence most

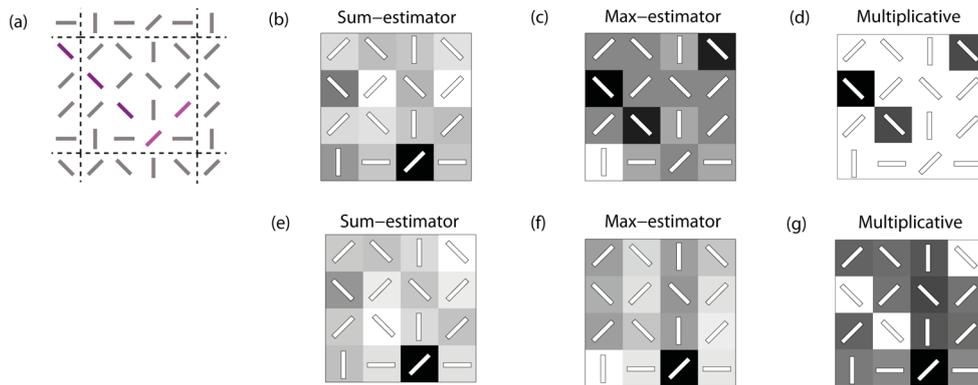


Figure 5.9: (a) A stimulus with periodic boundary conditions is depicted within the dashed square. It contains one contour with three elements (purple) and several contours of length two. One of these contours of length two is highlighted (pink), as it contains an element i with most neighboring elements directed towards the location of i , which makes it very salient for the additive model. (b) and (c) show the salience for each edge element as predicted by the additive model with couplings solely between perfectly aligned collinear edge detectors. (d) shows the same for the multiplicative model. As the predictions with summation and maximum estimator are the same only one diagram is shown. (e) and (f) show the corresponding salience as predicted by the additive model using a wider association field that also couples elements with 45° alignment jitter. (g) shows the salience predictions of the multiplicative model with wide association field. In each diagram black corresponds to the highest predicted salience and white corresponds to zero salience.

5.1. Stability of the iterative models and effect of normalization

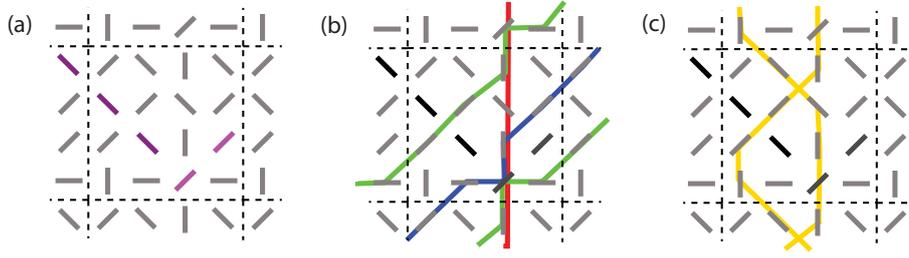


Figure 5.10: (a) Stimulus with periodic boundary conditions. Two finite perfectly aligned contours are highlighted. (b) All closed contours consisting of four elements are highlighted. (c) One closed contour consisting of 8 elements is highlighted as an example. Combinations of the contours highlighted in (b) will give further closed contours consisting of more than four elements.

salient contour? Only one neighboring element is perfectly aligned with edge i . However, there are n_{ori} neurons representing edge elements at position x_i with different orientations. And even though only one of the represented orientations is present in the stimulus, the other $n_{\text{ori}} - 1$ neurons get lateral input from neighboring edges that are aligned to the non-present edge elements represented by these neurons. In the multiplicative model this has no effect as $I_{\text{lat}}(x_i, \phi_j)$ is multiplied by $I_{\text{aff}}(\phi_i, \theta_j)$ which is zero for the non-present edges. However, in the additive model I_{lat} always contributes to the total input of a neuron independent of its afferent input. And as the summation estimator sums the salience of all non-present edge elements, it predicts a salience $s^{\text{sum}}(x_i)$ that is twice as high as the salience of any other position in this stimulus. The maximum estimator counts only the activity of the most active neuron at position x_i , and this activity is still lower than the highest activity at the contour elements.

In summary, for a δ -shaped association field the additive model with summation estimator is the only model which is sensitive to star-like edge configurations instead of the contour. However, if a wider association field is used, the situation changes. Now let us assume the association field $p^{\text{wide}}(\alpha, \beta) = 1/2$ for $\alpha = \beta = 0$ and $p^{\text{wide}}(\alpha, \beta) = 1/4$ for $\alpha = \beta = \pi/4$. With p^{wide} both models and both estimators predict the highest salience for edge element i instead for the contour A . However, the reason for this is different. While the additive model indeed predicts high saliences for star-like edge configurations, the multiplicative model computes a high salience for edge i . The reason for this is that i is part of many strongly jittered, but **closed** contours, while contour A is a **finite** contour (see Fig. 5.10).

As the multiplicative model computes the probability for each element to belong to an infinitely long contour, the activity of the finite contour A increases during the first iterations but then decreases again. To show that the multiplica-

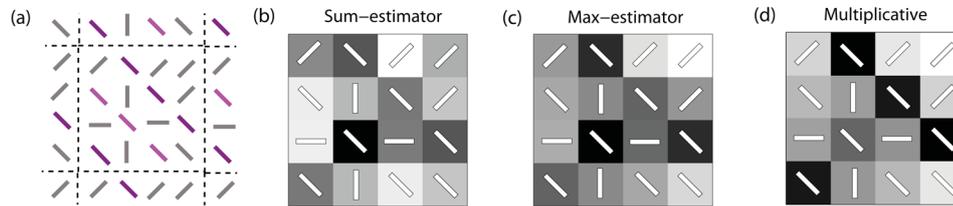


Figure 5.11: (a) Within the dashed square a stimulus with periodic boundary conditions is depicted. It contains one infinitely long contour (dark purple) and a finite contour of three elements (pink). (b)+(c) show the salience for each edge element as predicted by the additive model and (d) shows the corresponding salience as predicted by the multiplicative model, which is identical for both estimators. In each diagram black corresponds to the highest predicted salience and white corresponds to zero. For all diagrams $p^{\text{wide}}(\alpha, \beta)$ was used.

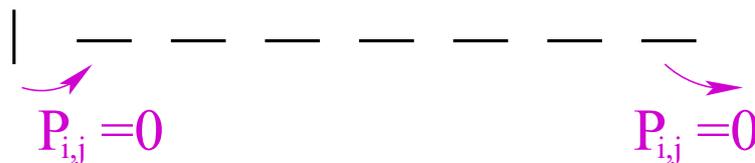


Figure 5.12: One-dimensional stimulus containing a finite contour surrounded by perpendicular background elements.

tive model is sensitive to closed contours rather than star-like edge configurations a second stimulus was analyzed (see Fig. 5.11). This stimulus contains a perfectly aligned infinitely long contour and the center edge of the star-like edge configuration is only part of strongly jittered closed contours. As before, the additive model with p^{wide} judges the center of the star-like edge configuration as most salient. However, the multiplicative model obtains the highest salience for the least jittered closed contour, as a good contour integration model is supposed to do.

5.2 Fading activity

Having seen situations in which the additive model with summation estimator has some disadvantages in the last section, we now want to look at a situation in which the multiplicative model has some shortcomings. Consider a one-dimensional stimulus containing a finite contour of L elements. If this contour is surrounded

5.3. Adjustment of afferent and lateral input strength

by edge elements perpendicular to the contour elements, like depicted in Fig. 5.12, there is no lateral connection between the neuron activated by the last contour element and the neuron activated by to the neighboring background element. During the first $L-1$ iteration steps the activity of the contour elements, particularly the activity of the first and the last element, will increase in the both the additive and the multiplicative model. However, at the L th iteration the activity will be multiplied by 0 in the multiplicative model and thus drop to zero. Originally this model was designed to detect closed or infinitely long contours and in the stationary state this model gives for each element the probability to belong to a closed contour. Hence very jittered closed contours are judged more salient than straight finite curves.

In practice this model nevertheless works very well for the complex stimuli we use for the psychophysical experiments, as there is almost always a way to close the contour via a more or less jittered and curved path in the background. However, what this example demonstrates, is that when searching for finite contours, the model performs much better transiently before reaching the stationary state.

One possibility to preserve the high activity which builds up during the first $L-1$ iterations could be to use an association field that is greater zero for all angles α and β of two adjacent edge elements. This can be achieved by a small offset to the association field, which is mathematically similar to introducing spontaneous activity. Another solution would be to combine the multiplicative model with the additive model, which would not have this decrease in activity at the L th iteration. As already discussed in 4.3.3 such a combination of linear and nonlinear model dynamics would already include the effect, that an offset for the association field in the multiplicative model has. It remains to be shown, to what extend such a combined model dynamics will be less sensitive to the problem of fading activity from which a purely multiplicative model suffers. Also it is not clear to what extent the problem of illusions for star like edge geometries, a disadvantage of purely linear model dynamics, can be cured by this combined model dynamics.

A different method to fix this shortcoming of the multiplicative model is to use a bidirectional association field. However, when assuming horizontal interactions in the range of nearest neighbors this leads to a very high salience for two perfectly aligned edge elements in comparison to longer, but slightly jittered contours.

5.3 Adjustment of afferent and lateral input strength

In the additive model the total neuronal input is $I = I_a \cdot I_{aff} + I_l \cdot I_{lat}$. Therefore the influence of the afferent and the lateral input and hence the model performance

depend on the balance of I_a and I_l . There are three qualitatively different working regimes which can be considered: $I_a \gg I_l$, $I_a \ll I_l$ and $I_a \approx I_l$.

If $I_a \gg I_l$, the lateral input is negligible and we obtain $A \approx I_{\text{aff}}$. So s_{sum} is the same in each hypercolumn, which represents a visible edge of a given contrast. Equally s_{max} is identical for all hypercolumns representing visible edges with the same contrast. Hence the saliency depends only on the contrast of an edge element and the contour will be detected at the element with the highest contrast regardless of its alignment with neighboring edges. In this case, no real contour integration takes place.

This effect is demonstrated in Fig. 5.13. We designed a stimulus, where the contour is very easily distinguishable from the background. For this purpose we assumed periodic boundary conditions and an infinitely long, oblique contour. Furthermore, we chose a very regular background, that does not even contain contours of length 2 (see Fig. 5.13 (a)). And as expected in the working regime $I_l \approx I_a$ the additive model clearly finds a high saliency s^{max} for the contour elements and a much lower saliency for the background elements (see Fig. 5.13 (c)). However, for $I_a \gg I_l$ the difference between the predicted s^{max} for contour and background elements is much smaller (see Fig. 5.13 (d)).

If in contrast $I_a \ll I_l$, the additive model also will not detect the contour. If the overall afferent input is negligible compared to the lateral input, we lose the information about the orientation of the edges. As we have neurons with all preferred orientations at each edge location x , the activity will depend only on the position of an edge and its distance to the neighboring edge elements, but not on their alignment and curvature. Hence the regular distance distribution between contour elements, background elements and contour and background elements, like in the stimuli used for the psychophysical experiments will lead to an equidistributed activity. This is also true for the simple stimulus depicted in Fig. 5.13 (a). The saliency for this stimulus also increasingly resembles an equidistribution the more the lateral input exceeds the afferent input (see Fig. 5.13 (b)).

The only working regime, in which the additive model performs well is $I_a \approx I_l$, where both inputs have a similar influence. This subtle balance of the two input strengths is only necessary for the additive model. In the multiplicative model the total neuronal input I depends only on $I_m = I_a \cdot I_l$. Thus no adjustment of I_a and I_l is needed.

Having analyzed the dynamical properties of both model classes, revealing both, potential pitfalls as well as putative advantages in detecting contours, we are now well prepared to investigate which of the models provides the most plausible account of human behavior in psychophysical experiments.

5.3. Adjustment of afferent and lateral input strength

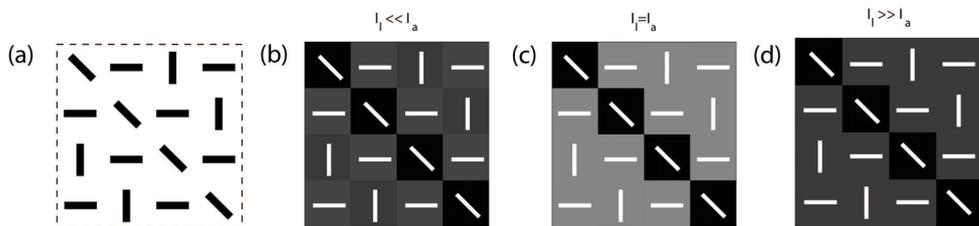


Figure 5.13: Plot (a) shows a stimulus with an infinitely long contour embedded in a very regular background. Plot (b) - (d) show the saliency as predicted by the additive model with different balance of afferent and lateral input strength. Here black corresponds to the maximum saliency and white represents zero saliency. In plot (c) afferent and lateral input are optimally tuned and the model judges the contour clearly more salient than the background. If either the overall afferent input strongly exceeds the lateral input (diagram (b)) or vice versa the lateral input is much stronger than the afferent input (diagram (d)) the predicted saliency approaches more and more an equidistribution.

Chapter 5. Analysis of model dynamics

Chapter 6

Robustness against noise and orientation jitter

One of the important questions of this thesis is, which contour integration model can explain the surprising speed and high performance of human observers in psychophysical experiments. In contrast to the simple stimuli used for the theoretical investigations in chapter 5, these experiments employed complex displays of Gabor patches similar to the ones first introduced by Field and colleagues (27). However, before processing stimuli employed during the psychophysical experiments, we study stimuli based on a hexagonal grid, which due to their symmetry still give the opportunity to supplement numeric simulations with analytic computations. Of the three model classes introduced in chapter 4 we select the model from which we expect the highest contour detection performance for contours of finite length and analyze the impact of orientation jitter and noise.

6.1 Stimuli and simulations

As we are interested in the robustness of the model against orientation jitter, we want to exclude any influence of the distance between edge elements on the model's performance. Hence we place Gabor patches on a hexagonal grid with periodic boundary conditions. This topology has the advantage that all elements have exactly the same distance r to their neighbors. We place a straight contour consisting of $L = 9$ elements horizontally, right oblique, or left oblique on the hexagonal grid which contains in total $N = 324$ edge elements. Now we can vary the alignment of the edges without changing the contour's curvature by rotating each contour element by $\pm\eta_\theta(2\pi/K)$, with $K = 72$ and an orientation jitter η_θ varying between 0 and 6. The orientation of each background element was randomly chosen from all possible orientations $\theta_i = i \cdot 2\pi/K$ where $i = 1 \dots K$.

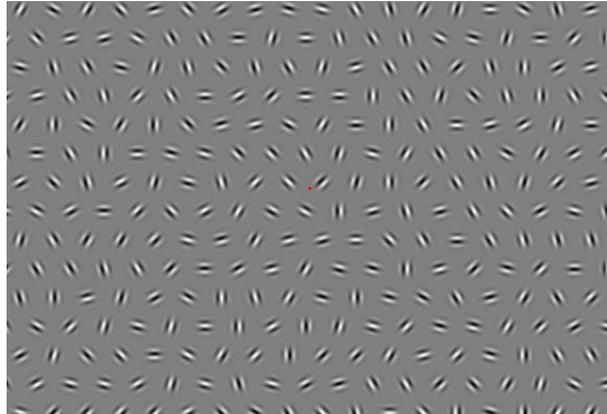


Figure 6.1: Example stimulus: Hexagonal grid with a straight contour without orientation jitter ($\eta_\theta = 0$) consisting of 9 elements.

An example of these stimuli is depicted in Fig. 6.1. These stimuli were only used for modeling studies investigating the robustness against orientation jitter. In order to investigate the dependence of the contour detection performance on the curvature of the contour, these stimuli are not suitable as for this hexagonal geometry only contours of zero curvature or contours with an angle of a multiple of 60° between two edge elements are possible. These highly curved contours are hardly recognized by humans, even if they are perfectly aligned.

These very regular stimuli have the second drawback that edge positions line up in three orientations: horizontal, right oblique and left oblique. Hence there is a high probability that contour-like structures along these lines arise by chance in the randomly generated background. In order to be able to include horizontal, right oblique or left oblique perfectly aligned contours in the stimuli we have to discretize the possible orientations for the edge elements into $K = 3K_0$ orientations with $K_0 \in \mathbb{N}$, such that they contain the three orientations necessary to construct these contours.

Hence, given that these three orientations are possible, the smaller K the higher is the probability to get contours in the background, which occur by chance when choosing random orientations for the background elements. Let us derive the probability for such randomly occurring contours in the background.

The probability that L randomly chosen edge elements located in either a horizontal, a right oblique, or a left oblique line are perfectly aligned is $(2/K)^L$ and the probability $P(L|N_l, K)$ that in a line of N_l elements with K possible directions, there exists at least one contour of L aligned elements is

$$P(L|N_l, K) = \left(\frac{2}{K}\right)^L \left(1 + (N_l - L) \frac{K-2}{K}\right) \quad (6.1)$$

for $N_l \leq 2L$ and

$$P(L|N_l, K) = \left(\frac{2}{K}\right)^L + \frac{K-2}{K} \left(\frac{2}{K}\right)^L \left[N_l - L - \sum_{Z=L}^{N_l-L-1} P(L|Z, K) \right] \quad (6.2)$$

for $N_l > 2L$. The derivation of Eq. 6.1 and 6.2 can be found in Appendix A The probability $P_{\text{rand}}(L|M, N_l, K)$ to find at least one such contour in a display of M lines of N_l elements each is then

$$P_{\text{rand}}(L|M, N_l, K) = \sum_{k=1}^M \binom{M}{k} P(L|N_l, K)^k (1 - P(L|N_l, K))^{M-k} \quad (6.3)$$

So for example in a hexagonal grid with $N_l = M = 18$, like the one used for the simulations in this section, the probability to find by chance a horizontal contour of length $L = 9$ is $P_{\text{rand}} = 1.7 \cdot 10^{-12}$. When assuming only half the possible directions $K = 36$ the probability increases by more than two orders of magnitude to $P_{\text{rand}} = 8.6 \cdot 10^{-10}$. The probability to find a contour of length $L = 4$ is $P_{\text{rand}} = 1.6 \cdot 10^{-4}$ for $K = 72$ and $P_{\text{rand}} = 0.0024$ for $K = 36$.

If we would assume a less regular geometry of edge locations x_i such that no L elements are positioned in a straight row and only curved contours are possible, it is not possible to build a perfectly aligned contour, when only few different orientations θ_i are possible. Many different θ_i are needed and hence for irregular stimuli the probability for randomly arising contours of small alignment jitter is far lower than for a hexagonal geometry.

If the stimulus includes a huge number of different orientations θ_i , but the contour integration model only has few orientation preferences ϕ_i and employs a narrow input tuning, we can encounter sampling problems when no orientation preference matches the orientation of a particular edge element. These problems can be avoided either by choosing sufficiently many orientation preferences ϕ_j together with an afferent input tuning curve which is wide enough such that each edge causes afferent input to at least one orientation column. Alternatively we could adapt the preferred orientation at each position x_j such that one orientation preference $\phi_{k,j}$ matches θ_i . In the case of the hexagonal grid stimulus we avoid sampling problems as all θ_i occurring in the stimuli are also preferred directions of the model neurons.

We modeled the contour detection performance for the stimuli described in section 6.1 using the ideal model (see 4.1.2) using the maximum estimator to determine the saliency s_{max} (see 4.2.2). We choose the distance dependence of the association field as

$$F(r) = \begin{cases} 1 & \text{for } r = r_0 \\ 0 & \text{else} \end{cases} \quad (6.4)$$

which means that there were only couplings of a neuron to its nearest neighbors.

6.2 Performance and chance level

In psychophysical contour detection experiments where the contour has to be detected on either hemifield of the screen, the chance level is $P_{\text{hemi}}^{\text{chance}} = 50\%$. For varying degrees of difficulty the performance P_{hemi} ranges from $P_{\text{hemi}}^{\text{chance}} \leq P_{\text{hemi}} \leq 100\%$. However, even if the observer gives a correct response we cannot know, whether he really detected the contour, or something different which is by chance located in the same hemifield as the contour itself. It is also possible that the observer just made a lucky guess without recognizing anything in particular. When running a simulation, we first get a saliency distribution for all elements from which we can decide whether the highest saliency was predicted for a contour or for a background element. When using the contour detection criterion that more than half of the L_{max} elements with the highest saliency must be contour elements in order to predict the location of the contour correctly, we get a chance level $P_{L_{\text{max}}}^{\text{chance}} \ll P_{\text{hemi}}^{\text{chance}}$. The probability $P_{L_{\text{max}}}(X = k)$ that k of these L_{max} elements are contour elements is given by the hypergeometric distribution, a probability distribution describing selection from a finite population without replacement. Hence it yields

$$p_{L_{\text{max}}}(X = k) = \frac{\binom{L}{k} \binom{N-L}{L_{\text{max}}-k}}{\binom{N}{L_{\text{max}}}} \quad (6.5)$$

Here L is the number of contour elements and N is the number of total edge elements within the stimulus. The chance level is then $P_{L_{\text{max}}}^{\text{chance}} = p_{L_{\text{max}}}(X > L_{\text{max}}/2)$ with a standard deviation of

$$\sigma^{\text{chance}} = \sqrt{\frac{P_{L_{\text{max}}}^{\text{chance}} \cdot (1 - P_{L_{\text{max}}}^{\text{chance}})}{N_s}}$$

where N_s is the number of stimuli over which the performance is averaged. Here we used $N_s = 300$ and get the chance level $P_1^{\text{chance}} = 2.8 \pm 0.9\%$, $P_3^{\text{chance}} = 0.2 \pm 0.26\%$ and $P_5^{\text{chance}} = 0.014 \pm 0.069\%$, in 1- σ -precision. Hence we have a much larger regime for performance values than we have when using P_{hemi} . The drawback is that these performance values cannot be compared to human performances in a 2AFC-task directly.

However, assuming that there is an equal number of stimuli containing the contour on either side and that furthermore all randomly arising edge configurations in the background which happen to be similarly or more salient than the contour are also equally likely on both hemifields, the relation between the two different performance measures is simply

$$P_{\text{hemi}} = \frac{1 - P_{L_{\text{max}}}}{2} + P_{L_{\text{max}}}$$

6.3. Robustness against synaptic noise

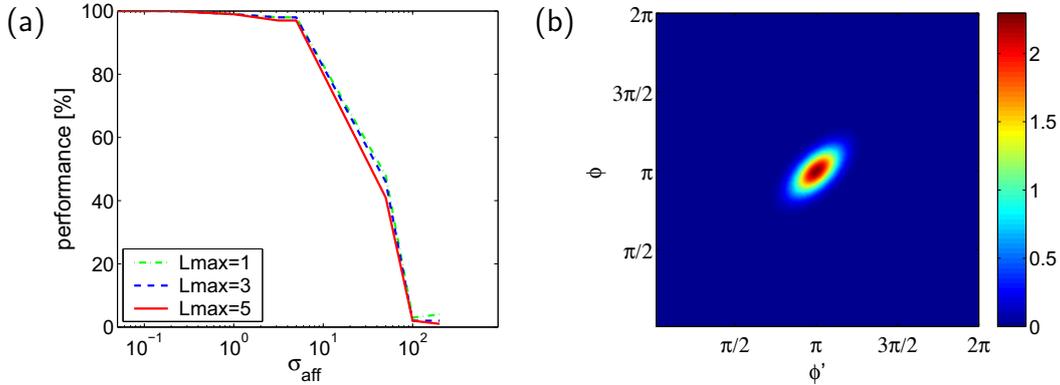


Figure 6.2: (a) Performance for different L_{max} : Performance decreases with larger σ_{aff} but is approximately the same for all L_{max} . (b) Association field with $\sigma_{\alpha} = 2\sigma_{\beta} = \pi/12$ for two horizontal neighboring edge elements. This association field was used for all simulations within this chapter.

So a comparison to psychophysical performances is possible if there is no bias for contours on either side of the display.

Performances for $L_{\text{max}} = 1, 3$ and 5 are approximately equal (see Fig. 6.2 (a)). For this simulation we used a stimulus containing a contour with $\eta_{\theta} = 0$. However, also for higher orientation jitter all L_{max} lead to a very similar performance. Thus we show in the following diagrams in this chapter only the performance for $L_{\text{max}} = 5$. The simulations are done with an association field having the width $\sigma_{\alpha} = 2\sigma_{\beta} = \pi/12$. The association field is illustrated in Fig. 6.2 (b). There the connection strength between two neurons with preferred orientation ϕ and ϕ' corresponding to two horizontal neighboring edge elements is shown.

6.3 Robustness against synaptic noise

In the brain we have random fluctuations in the membrane potential leading to synaptic background activity. Consequently, we added synaptic noise to the afferent input I_{aff} . This noise is uniformly distributed between 0 and $\eta_{\mu} \cdot \max_{\phi}(I_{\text{aff}})$. It can be either static noise η_{μ}^s which is constant for each iteration of the model or dynamic noise η_{μ}^d varying with time. We investigated dynamic and static noise with $\eta_{\mu} = 0.001$ and $\eta_{\mu} = 0.05$ for the orientation jitter $\eta_{\theta} = 0$ (see Fig. 6.3). For both noise levels the contour detection performance was better for dynamic noise than for static noise, as the dynamic noise in contrast to the static noise averages out over time. With increasing width σ_{aff} of the afferent input the information about the orientation ϕ of an edge becomes less evident and the performance decreases. Here it starts decreasing at the critical width of $\sigma_{\text{aff}} \approx 10$ for a noise

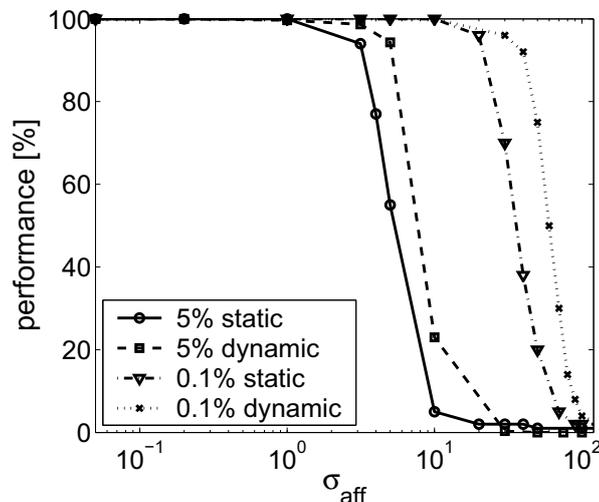


Figure 6.3: Contour detection performance in dependence of the afferent tuning width for different levels of synaptic noise. Static noise has a stronger impact on the performance than dynamic noise.

level of $\eta_\mu = 0.001$ and $\sigma_{\text{aff}} \approx 4$ for $\eta_\mu = 0.05$, respectively, where the noise level exceeds the amplitude of the tuning curve of the afferent input. When the width of the afferent input increases further until a distinction between a noisy Von Mises-function and a noisy equidistribution is no longer possible, the performance drops to chance level. When applying no noise at all the decrease of performance is determined solely by the computing precision, which in our case is 10^{-16} .

6.4 Robustness against orientation jitter

Another type of noise that we investigated was noise on the stimulus, i.e. the orientation jitter. We varied η_θ from 0 to 6 and plotted the model performance for different η_θ against σ_{aff} (see Fig. 6.4). As a wider σ_{aff} means less specific information about the orientations θ_i of the edge elements, we expect that the performance monotonically decreases with increasing σ_{aff} . For small jitters up to $\eta_\theta = 2$ performance indeed starts at 100% and monotonically decreases to chance level. However, for $\eta_\theta \geq 3$ the performance first increases with σ_{aff} until it reaches a plateau starting between $\sigma_{\text{aff}} = 0.1$ and $\sigma_{\text{aff}} = 1.0$. Finally it decreases to chance level because of synaptic noise. The critical value of σ_{aff} where the performance starts decreasing, and where it finally reaches chance level seem to be independent of η_θ .

At first glance this result seems to be counterintuitive. Why does the perfor-

6.4. Robustness against orientation jitter

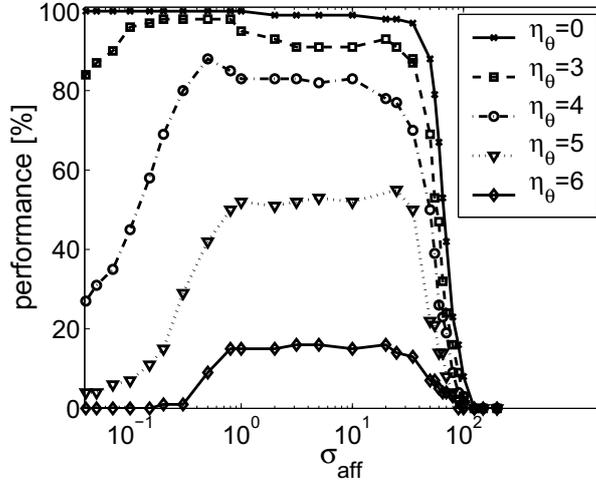


Figure 6.4: Contour detection performance as predicted by the ideal model in dependence of orientation jitter η_θ and tuning width σ_{aff} of the afferent input. The performances for $\eta_\theta = 1$ and $\eta_\theta = 2$ are omitted here, because they are nearly indistinguishable from the performance for $\eta_\theta = 0$.

mance increase for less specific information about θ ? To answer this question let us consider two horizontally neighboring edges i and i' . In the worst case i is jittered by $+\eta_\theta$ and i' by $-\eta_\theta$ or vice versa as depicted in Fig. 6.5 (a). In order to be grouped together as a contour the mean link probability between the two hypercolumns corresponding to edge i and i' should exceed the average link probability $\langle s_l(i_b, i'_b) \rangle$ between the hypercolumns of two neighboring background elements i_b and i'_b . This leads to the condition

$$o(x'_i, x') = \sum_{k'=1}^K I_{\text{aff}}(x'_i, \Phi_{k'}) \cdot I_{\text{lat}}(x'_i, \Phi_{k'}, x_i, \vec{\Phi}) > \langle s_l(i_b, i'_b) \rangle \quad (6.6)$$

where $\langle s_l(i_b, i'_b) \rangle$ for continuous orientations is given by

$$\langle s_l(i_b, i'_b) \rangle = \frac{I_a \cdot F(r)}{4\pi^2} \quad (6.7)$$

Here I_a is the overall strength of the afferent input and $F(r)$ is the distance dependence of the association field (see 6.4). A detailed calculation of Eq. 6.7 can be found in appendix B. The link salience between element i and i' increases with the overlap of $\vec{I}_{\text{aff}}(x'_i)$ and $\vec{I}_{\text{lat}}(x'_i, x_i, \vec{\Phi})$ which is illustrated in Fig. 6.5. This overlap increases with wider σ_{aff} and hence the contour detection performance increases. For a broader σ_{aff} , horizontal interactions which were inactive before can be recruited in order to strengthen the effective link between the edges i and i' . The same effect occurs when instead of σ_{aff} the association field is

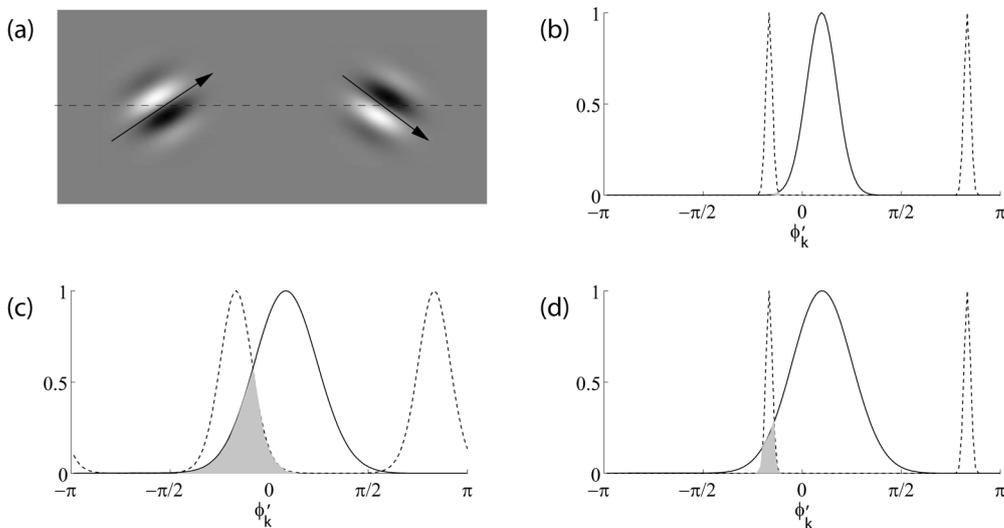


Figure 6.5: (a) Two horizontally neighboring edge elements i and i' with orientation jitter $+\eta_\theta$ and $-\eta_\theta$, respectively. (b) - (d) show the afferent input distribution $I_{\text{aff}}(x'_i, \phi'_k)$ at destination element i' (dashed line), together with the lateral input $I_{\text{lat}}(x'_i, \phi'_k, x_i, \vec{\phi})$ (solid line) in different situations, namely (b) with small σ_{aff} and small $\sigma_\alpha = 2\sigma_\beta$, (c) with large σ_{aff} and small $\sigma_\alpha = 2\sigma_\beta$, and (d) with narrow σ_{aff} and wide $\sigma_\alpha = 2\sigma_\beta$. The shaded areas depict the overlap of the two curves symbolizing the relative link probability of i' and i .

broadened by increasing σ_α or σ_β . However, for very wide tuning widths for either the afferent input or the association field, contour integration does not benefit from enlarged σ 's as the probability for the detection of strongly jittered pseudo-contours in the background also increases with the widths of afferent input and association field.

We saw in this chapter that contour detection performance is approximately independent of the width of the afferent input for a wide range of σ_{aff} . The performance plateau we found is shifted to smaller σ_{aff} 's compared to tuning widths measured experimentally. For example Lampl and colleagues (71) found intracellular recorded tuning widths in simple cells of cat primary visual cortex that varied between 12 deg and 34 deg, depending on the spatial frequency. However, the used association field was chosen in order to get a high performance of the model and is also smaller than experimentally obtained association fields (60). For broader σ_α and σ_β we would also expect a broader plateau which might include experimentally found σ_{aff} .

Even though the dependence of the performance on the afferent input here was only investigated for the probabilistic model for contours of finite length, we would expect a qualitatively similar result for the iterative model that couples afferent and lateral input multiplicatively. All investigations about the link

6.4. Robustness against orientation jitter

salience should yield there in exactly the same way.

Chapter 6. Robustness against noise and orientation jitter _____

Chapter 7

Evaluation of model classes

7.1 Human psychophysics

In order to compare different models we need a benchmark. And as we desire to explain the human brain, rather than finding an algorithm for contour integration that works perfect for technical image processing applications, this benchmark needs to be the human behavior. Hence we compare the results of different contour integration models to psychophysical experiments where human subjects have to detect a contour consisting of curvilinear aligned Gabor elements embedded in a background of randomly oriented Gabor elements. This is a paradigm that is often used in psychophysics and there are several contour detection experiments of this kind that can be found in the literature (i.e. (27; 98; 68; 99; 6; 44; 7; 83)). The experiments described in this chapter were conducted by Simon Neitzel, Sunita Mandon and Andreas Kreiter. What distinguishes this empirical work from the publications just mentioned and makes it particularly useful for our work, is that the stimuli were generated by an association field and thus the probabilistic model described in 4.1.2 is the optimal model for detecting the contours hidden in these stimuli.

7.1.1 Generation of stimuli

Stimuli were composed of randomly oriented Gabor patches as background and $L = 4$ to 10 collinearly aligned Gabors as contour. Element distances (1.2 - 3.5 degree of visual angle) varied with the number of contour elements L such that the length of the contour in degrees of visual angle was kept constant. Gabor patches were used because they closely resemble the shape of experimentally measured

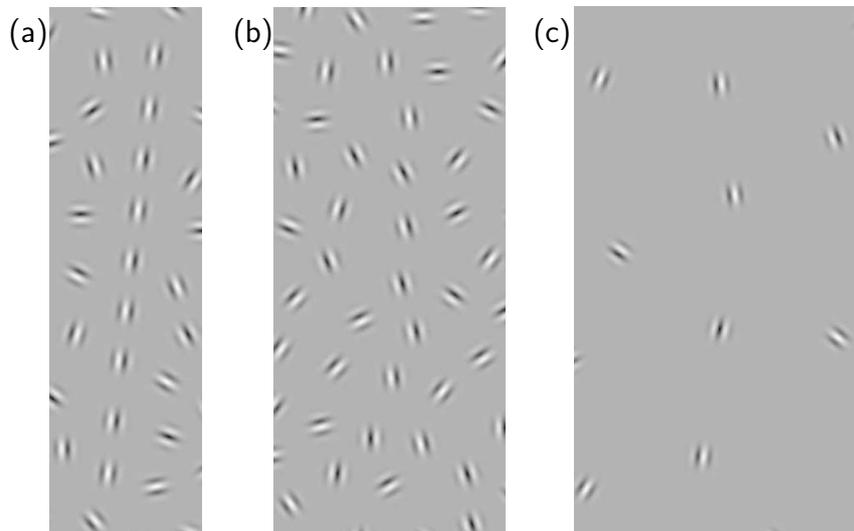


Figure 7.1: Section of the stimuli used for the psychophysical contour detection experiment: (a) contour consisting of 9 elements with zero orientation jitter, (b) contour with 10 elements and the largest orientation jitter used (22.5°), (c) contour consisting of four elements with a jitter of 9° .

receptive fields (57). Here they were defined by

$$G(x, y, \theta) = \cos\left(\frac{2\pi x \cos \theta + 2\pi y \sin \theta}{\lambda} + \phi_G\right) \exp\left(-\frac{x^2 + y^2}{2\sigma_G^2}\right) \quad (7.1)$$

where (x, y) denote the position of the edge element and θ its orientation. λ is the wavelength and σ_G the width of the Gabor element. The phase $\phi_G \in [0, \pi]$ was chosen randomly for each individual Gabor patch.

The element distances were chosen such that even for the smallest employed element distance of 1.2 degree of visual angle it can be assumed that long-range horizontal connections in V1 link only nearest neighboring elements, but not next-to-nearest neighbors. Long-range horizontal interactions span up to 5 – 8mm in cats (129; 36), but only 1.5 – 2.1mm in primates (104; 62; 4; 120). When mapping element distances in the stimulus onto the cortical surface, one has to take into account the eccentricity at which the elements were shown. The highest eccentricity in our stimuli, averaged over the elements of a contour, is 12.9 degree of visual angle. Using data on cortical magnification in human visual cortex (121) we find that for this eccentricity and the smallest element spacing the distance to next-to nearest neighbors corresponds to at least 3.6mm in the cortex, which is not linked by long-range horizontal interactions in primates. For an average eccentricity of 7.6 degree visual angle, this even corresponds to more than 6.1mm.

The contour was created by starting with a random element and successively drawing the following $L - 1$ elements from an association field $p(r, \alpha, \beta)$ (see Eq.

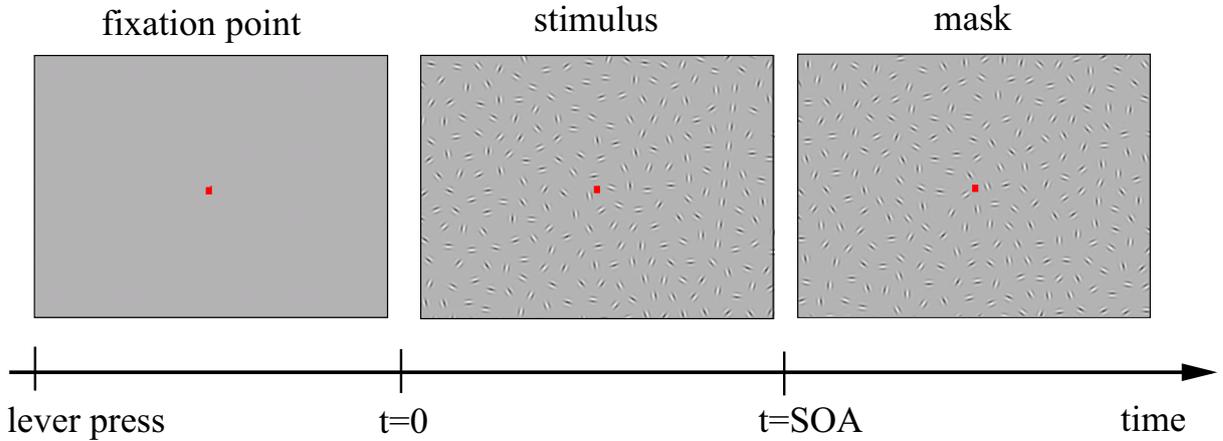


Figure 7.2: Psychophysical contour detection experiment: First a fixation point is shown. Then the stimulus is shown and with a certain SOA a second stimulus serving as backward mask is shown.

4.20). The contour was placed either on the right or on the left hemifield of the display. Then background elements of random orientation were added, such that the distance distribution between contour elements, between background elements and between contour and background elements were the same. To complicate contour integration we also introduced curvature jitter and alignment jitter. This is specified by the width σ_α and σ_β of the generating association field. However, as all contours used in the psychophysical experiment have the same curvature and alignment jitter, we summarize the two parameters into a combined curvature and alignment jitter $\sigma_j = \sigma_\alpha = \sigma_\beta$ of the stimulus. To distinguish between contour jitter and the possibly different width of an association field employed by a contour integration model, σ_j will only be used to describe the jitter of the stimulus, while from now on solely σ_α and σ_β will be used to denote properties of an association field employed to detect contours. The curvature and alignment jitter of the contours for this experiment varied between $\sigma_j = 0\text{deg}$ and $\sigma_j = 22,5\text{deg}$. Three example contours for different σ_j and different average element distances r_0 are shown in Fig. 7.1.

The stimuli were displayed on a 21 inch CRT (cathode ray tube) monitor with a resolution of 1152×864 . The refresh rate was 100 Hz. The screen was placed approximately 80cm in front of the subjects eyes, such that the display spanned $26.6^\circ \times 20.0^\circ$ visual angle.

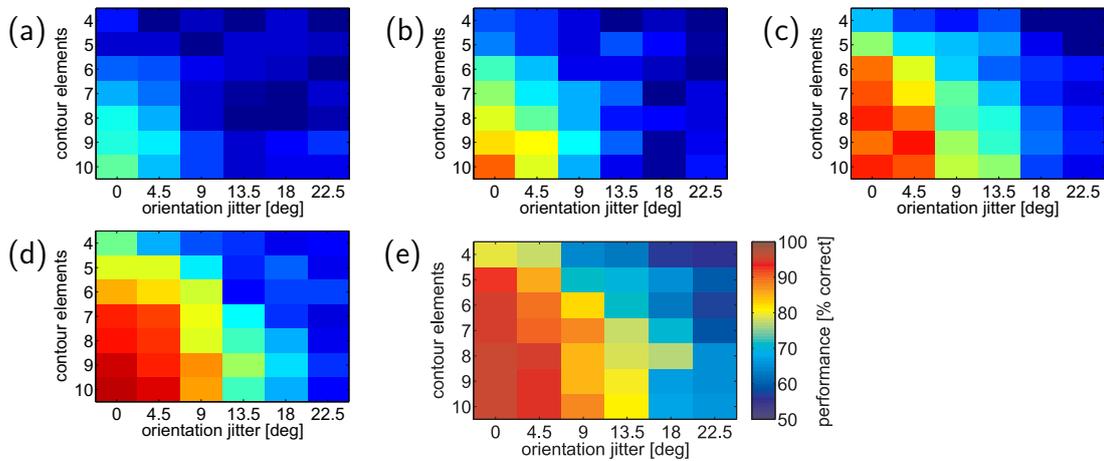


Figure 7.3: The diagrams (a) - (e) show the human contour detection performance in percent correct, averaged over all observers for 20ms, 30ms, 60ms, 100ms, and 200ms SOA. The color coding is the same in each diagram

7.1.2 The contour integration task

The experiment was performed by five human subjects with normal or corrected to normal vision. Subjects were trained to do a two-alternative forced choice (2AFC) contour detection task. For each combination of the seven element distances and the six orientation jitters a set of 48 stimuli was used. In the beginning of each trial a fixation point appeared in the middle of the screen and the subject had 3 sec time to start the trial by pressing a lever. 1 sec after the lever press the contour stimulus was presented. Next a backwards mask was shown. The stimulus onset asynchrony (SOA) between target and mask stimulus varied between 20 - 200 ms. The backward mask consisted of Gabor elements at the same position as in the original stimulus. However, all Gabors were randomly rotated, such that the contour was no longer visible. Finally subjects had to report on which hemifield of the screen they saw the contour by pressing a left or right button. Chance level was 50%. Stimuli of different jitter and element distances were shown interleaved and subjects saw each stimulus only once. The order of the stimuli was different for each observer.

7.1.3 Results: human performance

In the experiments the contour detection performance of the observers in dependence of orientation jitter, average element distance r_0 , which in our case reciprocally corresponds to the number of contour elements L , and stimulus presentation time was measured. A summary of all experimental results is shown in Fig. 7.3. Performance generally increases with SOA, lower jitter, and number of contour

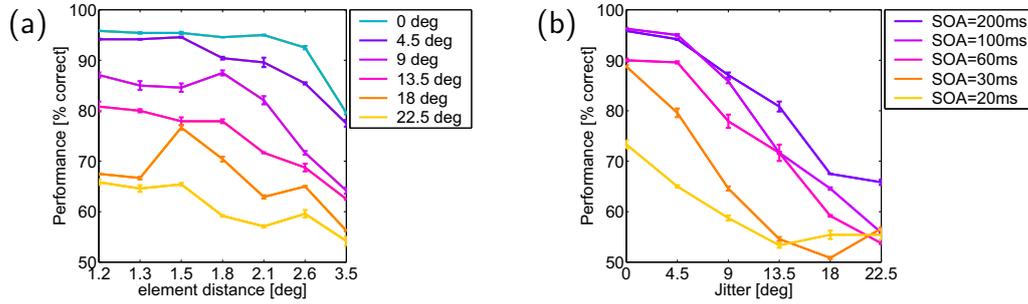


Figure 7.4: (a) Human contour detection performance with 200 ms SOA for different curvature and alignment jitter σ_j and different element distances r_0 , corresponding to different L . Performance decreases with higher σ_j and with larger r_0 . (b) Human contour detection performance for different SOAs. Only the data for $L = 10$ is shown. Performance increases with longer stimulus presentation times. In both plots performance is averaged over five observers and the chance level is 50%. Error bars in both plots denote the variance of the performance.

elements. Fig. 7.4 (a) shows how the contour detection performance for 200 ms SOA decreases with increasing orientation jitter and element distances. Qualitatively this has been found by many psychophysical studies (27; 68; 75; 96; 9; 83). However, quantitatively the decrease of performance in dependence of orientation jitter can differ enormously between studies. Field and colleagues found a human contour detection performance near chance level for an orientation jitter near 30° , and similarly in a contour detection study with highly trained macaque monkeys a contour detection rate near chance level for 36° was measured (83). Contrary, in a psychophysical study with amblyopic monkeys a performance of 75% could still be reached for orientation jitters between 40° and 60° (69). The experimental data used in this thesis shows a much faster decrease in performance, i.e. the performance for 22.5° orientation jitter, $L=10$ and $SOA=200$ ms is $65\% \pm 7\%$. However, here orientation jitter is not defined in the same way as in the studies just mentioned. One major difference to the study with amblyopic monkeys is that in our study the distance statistics between contour elements and background elements is the same. This was not the case in their stimuli and could serve as an additional cue for the position of the contour (9).

Fig. 7.4 (b) shows that the performance decreases for shorter stimulus presentation times. However, while in this study the performance for straight contours of $L = 10$ is still $89\% \pm 5\%$ for $SOA=30$ ms, and a similarly rapid contour detection was found in a recent study with macaque monkeys (83), earlier works (6; 44) stated that contour integration is a rather slow process, requiring 100 ms processing time for straight contours and 250-300 ms for higher curvature.

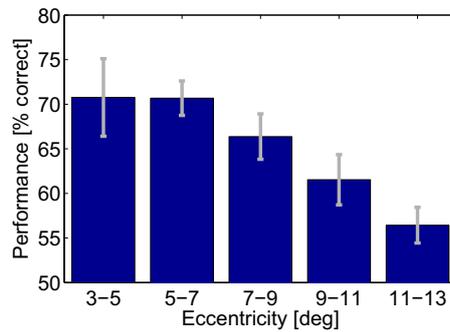


Figure 7.5: Human contour detection performance in percent correct versus mean eccentricity of the contour elements in degrees of visual angle. The performance was averaged over all jitters, element distances, and SOAs. The error bars denote one standard deviation.

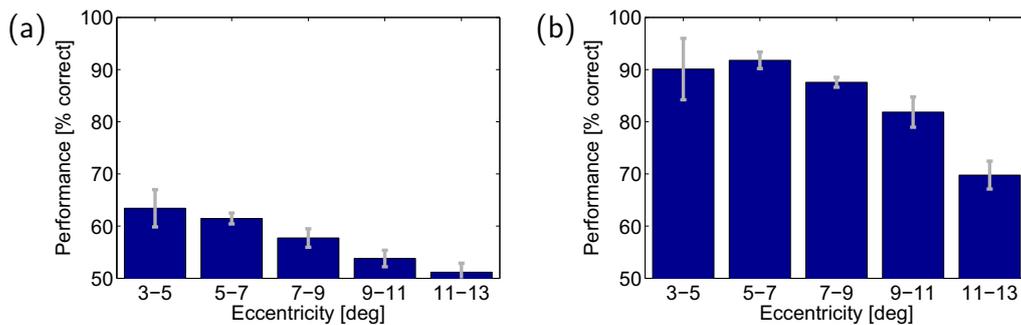


Figure 7.6: Human contour detection performance versus mean eccentricity of the contour elements in degrees of visual angle. The performance was averaged over all combinations of jitter, element distances, and SOA, which lead to a performance of (a) more than 75%, and (b) less than 75% when averaging over all e . The errorbars denote one standard deviation.

Human contour detection in the periphery

Many psychophysical studies have shown that human contour detection performance is worse in the periphery than in central vision. However, there is an ongoing debate whether the performance decreases gradually with eccentricity (79; 92) or whether there are fundamental differences between contour integration in the fovea and in the periphery (eccentricities > 10 degree visual angle), suggesting a different mechanism for contour integration (45; 46).

In this psychophysical study the mean eccentricity ϵ of all contour elements within one stimulus varied between 2.9 and 12.9 degree visual angle. As it was not the aim of this study to investigate contour integration in the periphery, the stimuli do not contain contours at large enough eccentricities to draw any conclu-

sion about the existence of a different mechanism for eccentric contour detection. Also the eccentricity of every single element can differ by up to five degree from the mean eccentricity ϵ as the contours span up to 10 degrees of visual angle. However, when plotting the performance against ϵ we do see a significant decrease of the performance with eccentricity (see Fig. 7.5), which would support the theses of gradually decreasing performance. Nugent (92) reported a faster decrease of performance for the more difficult condition, where stimuli consisted of phase-alternating Gabor patches, as compared to phase aligned stimuli. The study used in this thesis uses only random phase Gabor elements, but the difficulty varies with average element distance r_0 , stimulus presentation time (SOA) and curvature and alignment jitter σ_j . In order to investigate whether we find a stronger effect for the more difficult tasks, we distinguished between combinations of r_0 , σ_j , and SOA-values that lead to an eccentricity-averaged performance higher, respectively lower than 75%. The corresponding histograms are shown in Fig. 7.6. However, here we find a stronger absolute effect for easier stimuli.

7.1.4 Correlations among human subjects

Performance is one important quantity to characterize human contour integration. However, performance only measures how many mistakes an average observer makes regardless of the nature of these misdetections. On the one hand there are errors which occur randomly. The probability for such a random error should depend on the difficulty of the task, i.e. the element distance, the orientation jitter and the presentation time, but not on the particular stimulus geometry. On the other hand, it can happen that the 'contour' that was implemented in the stimulus on purpose is not the most salient or contour-like object within the stimulus, but that by chance some edge configuration in the background is more salient than the actual 'contour'. In this case the detection of the most salient edge configuration is counted as a mistake, even though the subject did indeed find the most salient or contour-like object within the display. Hence one can hardly call this a misdetection. For such stimuli we would expect most observers to give the same 'wrong' answer.

In the case where a 'wrong' response from an observer is caused by a longer, less jittered contour in the background, the Bayesian model should also fail. Hence we could avoid counting trials as mistakes even though the longest and least jittered contour was detected, by defining a wrong response as a response that differs from the prediction of the Bayesian model. Even though this example suggests that the Bayesian model predictions seem to be a more plausible basis for evaluating human responses than the task, posed by the experimentalist, it is still unclear if or to what extent the Bayesian model resembles the model underlying human observers.

In fact the first question is whether there is such an underlying strategy for all observers and how much responses of individual subjects differ from each other. Hence it will be interesting to see, to what extent responses of different observers are correlated on the basis of individual trials. Thus a correlation analysis can reveal to what extent the underlying contour integration mechanism is the same for each observer.

7.1.5 Method: correlation analysis of human decisions and model predictions

Discriminating systematic and random decisions

In this section I will explain the method for correlation analysis involving ROC analysis (receiver operator characteristic analysis) for the example of the 2AFC-contour detection experiments described in 7.1 (24). Here we assume we have n_t trials of the same difficulty, meaning with constant σ_j and L . We denote a correct detection of the contour with 1 and a misdetection with 0. For two observers **A** and **B** we then get two sequences σ_A and σ_B of ones and zeros containing their contour detection hits and misses for the n_t trials. Now we want to know whether observers **A** and **B** have the same underlying contour integration strategy. If this is the case, and the common strategy is not just to randomly guess a contour position, the responses σ_A and σ_B of the two observers should be correlated and there should be a high number n_{id} of identical decisions of observers **A** and **B**. If both observers had the same underlying model without any noise, then we would expect $n_{id} = n_t$. However, $n_{id} \approx n_t$ does not necessarily mean a high correlation between σ_A and σ_B . E.g. for $p_A = p_B = 1$ it yields $n_{id} = n_t$ independent of the underlying models.

In general the observed number n_{id}^{ob} of identical responses of observer **A** and **B** is given by

$$n_{id}^{ob} = \sum_{i=1}^{n_t} (\sigma_A(i) \cdot \sigma_B(i) + (1 - \sigma_A(i))(1 - \sigma_B(i))) \quad (7.2)$$

In order to appraise the number of identical responses of the two observers we determine the probability $p(n_{id}^{ex}|p_A, p_B)$ of expected identical responses n_{id}^{ex} given their performances p_A and p_B under the assumption that errors are not made systematically and thus are randomly distributed over all trials. If the two observers have the observed performances $p_A = \sum_{i=1}^{n_t} \sigma_A(i)/n_t$ and $p_B = \sum_{i=1}^{n_t} \sigma_B(i)/n_t$, respectively, the probability to find n_{id}^{ex} identical responses is given by

$$p(n_{id}^{ex}|p_A, p_B) = \frac{r_{n_B}(a)r_{n_t-n_B}(n_A - a)}{r_{n_t}(n_A)} \quad (7.3)$$

Here $n_{id}^{ex} = a + n_t - n_B - (n_A - a)$ where $n_A = p_A \cdot n_t$ and $n_B = p_B \cdot n_t$ is the number of correct responses of observer A and B, respectively. $r_{n_B}(a) = \binom{n_B}{a}$ denotes the number of possibilities to distribute a correct responses onto the subset of n_B trials in which observer B answered correctly. $r_{n_t - n_B}(n_A - a) = \binom{n_t - n_B}{n_A - a}$ stands for the number of possibilities to distribute $n_A - a$ correct responses of observer A onto the $n_t - n_B$ trials in which observer B failed to detect the contour. $r_{n_t}(n_A) = \binom{n_t}{n_A}$ gives the number of possible distributions of the n_A correct responses of observer A onto the n_t trials. With this we obtain

$$p(n_{id}^{ex} | p_A, p_B) = \frac{n_A! n_B! (n_t - n_A)! (n_t - n_A)!}{n_t! a! (n_A - a)! (n_B - a)! (n_t - n_A - n_B - a)!} \quad (7.4)$$

where $a = (n_{id}^{ex} - n_t + n_A + n_B)/2$. This expression holds for $n_B < a < n_A + n_B - n_t$. For $a > n_B$ it is not possible to distribute a correct responses of observer A onto only n_B correct responses of observer B. Hence for $a > n_B$ it holds $r_{n_B}(a) = 0$ and accordingly $p(n_{id} | n_A, n_B) = 0$. Similarly for $n_A - a$ it holds $p(n_{id} | n_A, n_B) = 0$, as $n_A - a$ correct responses cannot be distributed onto $n_t - n_B$ responses.

We assume that human observer use a similar contour integration strategy and that therefore the discrete distribution of observed identical responses $p^{ob}(n_{id}^{ob})$ is shifted to higher values than the discrete distribution of expected identical responses $p^{ex}(n_{id}^{ex})$. Consequently we can formulate the hypothesis $p^{ob} \neq p^{ex}$. If n_{id}^{ob} exceeds a threshold z we assume that the hypothesis is true. If this guess is correct we call this decision a "true positive" otherwise it is a "false alarm". In signal detection theory the quantity n_{id} used to verify or falsify the hypothesis $p^{ob} \neq p^{ex}$ is called a test. The probability for a true positive $p^t(z) = p(n_{id}^{ex} > z | p^{ob} \neq p^{ex})$ is termed the hit rate or the power of the test. $p^f(z) = p(n_{id}^{ex} > z | p^{ob} = p^{ex})$ is the false alarm rate or the size of the test. The following table shows the probabilities for the test to give a correct answer.

underlying model	probability	
	correct	incorrect
$p^{ob} \neq p^{ex}$	$p^t(z)$	$1 - p^t(z)$
$p^{ob} = p^{ex}$	$1 - p^f(z)$	$p^f(z)$

The performance of the test depends crucially on the value of the threshold z . Ideally we want the power of the test to be close to 1 and the size of the test to be near 0. However, generally it is not possible to optimize both, power and size of a test. Instead one can maximize the probability p_c to obtain a correct

test result. This is given by $p_c(z) = p(p^{ex}(n_{id}^{ex}) = p^{ob}(n_{id}^{ob}))p^t(z) + p(p^{ex}(n_{id}^{ex}) \neq p^{ob}(n_{id}^{ob}))(1 - p^t(z))$. Assuming that

$$p(p^{ex}(n_{id}^{ex}) = p^{ob}(n_{id}^{ob})) = p(p^{ex}(n_{id}^{ex}) \neq p^{ob}(n_{id}^{ob})) = 0.5 \quad (7.5)$$

the maximum probability p_c^{max} for correct test results is

$$p_c^{max}(p_0^f) = 0.5(p^t(p_0^f) + 1 - p_0^f) \quad \text{where} \quad \left. \frac{dp^t}{dp^f} \right|_{p_0^f} = 1 \quad (7.6)$$

Another possibility where we do not need the assumption 7.5 is the ROC-curve. Here the threshold z is varied continuously, and the hit rate p^t is plotted against the false alarm rate p^f parameterized by the decision criterion z . This ROC-curve demonstrates how the size and the power of the test depend on z . For the hypothesis that observers make systematic errors rather than random errors the probability for a true positive is

$$p^t(z) = \sum_{n_{id}^{ob}=z}^{\infty} p^{ob}(n_{id}^{ob}) \quad (7.7)$$

and the probability for a false positive is

$$p^f(z) = \sum_{n_{id}^{ex}=z}^{\infty} p^{ex}(n_{id}^{ex}) \quad (7.8)$$

Equivalently we can write

$$p^t(p^f) = \sum_{n_{id}=z: \sum_{n_{id}=z}^{\infty} p^{ex}(n_{id})=p^f}^{\infty} p^{ob}(n_{id}) \quad (7.9)$$

Receiver operator characteristic for trials of varying difficulty

In the psychophysical experiments we have stimuli of different alignment and curvature jitter σ_j and of different element distances r_0 . For each set of stimuli with different parameters r_0 and σ_j observers show a different performance. Let us assume that we have the stimuli sets $m = 1, \dots, M$. The trials from all sets were viewed by each observers in a different order. This way a possible influence of the previous trial cannot have the same effect on the responses of all observers. For each combination of observers **A** and **B** we have a different distribution $p_{A,B,m}^{ex}(n_{id}) = p(n_{id}|n_A, n_B)$ of expected identical responses n_{id} under the assumption that the errors are distributed randomly. A second 'probability distribution' $p_{A,B,m}^{ob}(n_{id}) = \delta_{n_{id}, n_{id}^{ob}(\sigma_A, \sigma_B, m)}$ is given by the

actually observed number $n_{id}^{ob}(\sigma_A, \sigma_B, m)$ of identical responses of observers A and B. We now define $z_{A,B,m} = \sum_{n_{id}=n_{id}^{ob}(\sigma_A, \sigma_B, m)+1}^{n_t} p_{A,B,m}^{ex}(n_{id})$ and $Z_{A,B,m} = \sum_{n_{id}=n_{id}^{ob}(\sigma_A, \sigma_B, m)}^{n_t} p_{A,B,m}^{ex}(n_{id})$. Between those two magnitudes we interpolate linearly and obtain

$$p_{A,B,m}^t(p^f) = \begin{cases} 0 & \text{for } p^f \leq z_{A,B,m} \\ 1 & \text{for } p^f \geq Z_{A,B,m} \\ \frac{p^f - z_{A,B,m}}{Z_{A,B,m} - z_{A,B,m}} & \text{otherwise} \end{cases} \quad (7.10)$$

This interpolation is necessary to obtain a continuous function from the discrete distributions $p^{ex}(n_{id}^{ex})$ and $p^{ob}(n_{id}^{ob})$. From $p_{A,B,m}^t(p^f)$ we can compute the mean receiver operator characteristics for all possible pairs of observers and all sets of stimuli of various difficulty

$$\langle p^t(p^f) \rangle_{A,B,m} = \frac{1}{M n_{ob}(n_{ob} - 1)} \sum_{A=1}^{n_{ob}} \sum_{B \neq A}^{n_{ob}} \sum_{m=1}^M p_{A,B,m}^t(p^f) \quad (7.11)$$

Here n_{ob} is the number of observers.

7.1.6 The model as nullhypothesis

So far we counted a decision as correct, if the contour was detected on the hemifield where we placed the contour, and as wrong otherwise. However, maybe other criteria for classifying human or model decisions as hits or misses are more sensible. Let us assume, we placed a contour consisting of four elements with a strong orientation jitter in the left hemifield. Then it could happen just by chance that in the right hemifield four elements in the background form a contour with stronger alignment and less curvature than the contour we inserted on purpose. In this case not only humans, but also the best possible contour integration model would predict the contour on the right hemifield. So how can we say that this is wrong? Furthermore what we are searching for is not the best model in the sense that it is most efficient in finding contours, but the model that describes human behavior best. Hence we would like the model to reproduce the 'generic human decision' for a specific stimulus. Assuming an infinite number of human observers this could be obtained from the response that the majority of human observers give. Let us assume for a moment that we would know the model that exactly reproduces this mean human behavior. For a given stimuli sequence this model produces a sequence of left or right responses (L/R) which we can take as null hypothesis for the generic model (see Fig. 7.7). Now we evaluate the responses of the observers according to this null hypothesis: Identical decisions of an observer and the model taken as null hypothesis for the generic model are counted as correct responses of the observers and different decisions as mistakes

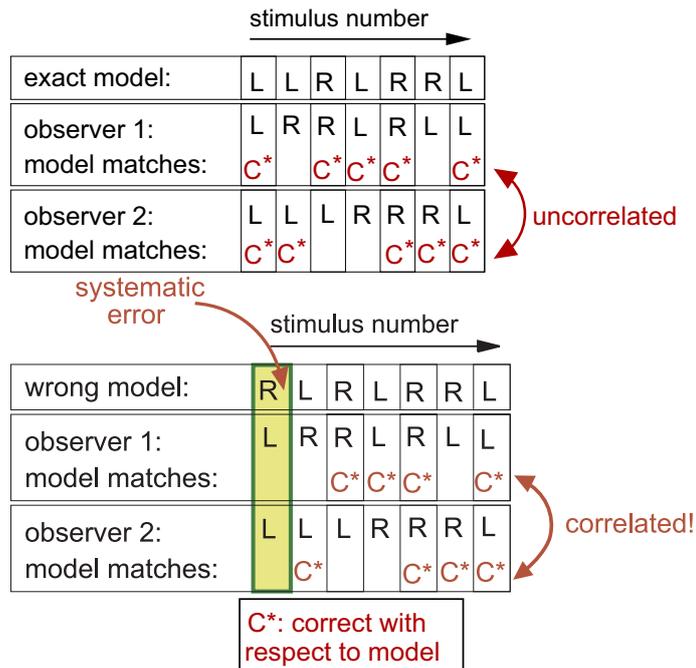


Figure 7.7: Correlation analysis with changed null hypothesis: In the upper table we assume we knew the exact model underlying human behavior. Its decisions for different trials are denoted by R (right) and L (left). Then observer 1 and observer 2 would give similar responses as the exact model but differ for some trials due to noise. When an observer gives the same response as the exact model we regard this as a correct response C*. Wrong responses are due to noise and hence not correlated. In the lower picture we assume that the model which we take as null hypothesis for the human behavior is not the generic model describing human behavior but differs from that model for the first trial. Hence we expect significantly more wrong responses for this trial (marked in green) and thus a correlated behavior of the two observers.

of the observers. Now there will always be deviations of the human response to the generic model as human behavior is influenced by other factors which may be regarded as noise. However, as errors because of noise are made randomly, there should be no correlations of these deviations from the null hypothesis for different subjects. If in contrast the model does not describe the generic human behavior, but makes a systematic error, then the deviations of different observer decisions from the model predictions should be correlated. This means when taking a certain model as null hypothesis the correlation between the decisions of different human observers with respect to this null hypothesis can serve as a measure for the quality of this model. The less correlation we have, the closer does this model resemble the generic human behavior.

Throughout this thesis we will refer to these correlations of the deviations of human decisions from the model prediction as "ROC2"-correlations. The previously described correlations between model predictions and human responses will be termed "ROC1"-correlations.

7.1.7 Testing the ROC-methods on surrogate data

In order to better understand the ROC1 and ROC2 method we tested it on surrogate data. Therefore we generated a string σ_{generic} containing the decisions of a surrogate generic model for n_t trials. Correct responses are denoted by 1 and wrong decisions by 0. These were drawn from a random probability distribution, assuming the probability for the generic model to detect a contour is p_{generic} . Then we generated the responses σ_{obs} for $n_{\text{ob}} = 5$ surrogate observers assuming their responses match the decisions of the generic model with the probability p_m , and that their deviations from the generic model were distributed randomly. Then we computed the ROC1-correlation between these n_{ob} observers. For an example of $p_{\text{generic}} = 0.8$ and $p_m = 0.8$ this is shown in Fig. 7.8 (a). Next we wanted to know, the curve progression of the ROC1-correlation between the responses of a model and these observers in dependence of the match of this model to the generic model. Therefore we generated a random sequence σ_{rand} of hits (1) and misses (0), assuming the same probability for correct decisions as in the generic model. Then we can combine the first n_m components of σ_{generic} and the last $n_t - n_m$ entries of σ_{rand} in order to create the responses $\sigma_m = \{\sigma_{\text{generic},1}, \dots, \sigma_{\text{generic},n_m}, \sigma_{\text{rand},n_m+1}, \dots, \sigma_{\text{rand},n_t}\}$ of a model whose match with the generic model increases with $n_m = 0, \dots, n_t$. In Fig. 7.8 (a) we see that with increasing n_m the ROC1 correlation between σ_m and σ_{obs} rises above the correlations among the n_{ob} observers. This suggests that the random deviations of the observer responses from the generic model have a bigger impact on the correlation between the observers responses than on the correlations between observers and model. Consequently we also expect the correlations among humans and a plausible contour integration model to exceed the correlations

among humans as human behavior is also affected by noise and hence should randomly deviate from the generic model which describes human contour integration. However, due to the limited number of observers, it is not possible to obtain the decisions of the generic model and the probability for human responses different from these decisions. Hence we cannot determine to which extent the correlations between a plausible contour integration model and humans should exceed the human correlations.

In addition Fig. 7.8 (a) shows that the ROC2-correlation indeed decreases to 0.5 for $n_m \rightarrow n_t$. This decrease to 0.5 can also be observed, if the observers have different performances, or if they are subjected to different amounts of noise, corresponding to different p_m for each observer, as long as p_m is the same for each trial (diagrams not shown). However, if the n_t trials can be divided into two subsets with different amounts of noise, this is not the case any more. This is shown in Fig. 7.8 (b). There we used again a performance of the generic model of $p_{\text{generic}} = 0.8$, but split the n_t trials in two subsets of equal size, assuming $p_m = 0.7$ for one subset and $p_m = 0.9$ for the other subset. Trials from both subsets were arranged in random order. Here the ROC2-correlation stays far above 0.5 even for $n_m = n_t$. Nevertheless it is monotonously decreasing, indicating that a lower ROC2-correlation indicates an improvement of the tested model.

In human contour detection experiments an example for such a bias could be the preference of a subject for a particular hemifield. For example voting for the right side whenever one did not perceive the real location of the contour, leads to a higher performance for contours on the right hemifields than for contours on the other side. In the psychophysical study used for our correlation analysis, a difference in performance for the two hemifields of up to 14% was observed. As this is presumable not the only bias the observers had, it is not possible to give a realistic lower bound of the ROC2-correlation. However, despite the fact that the ROC2-correlation cannot serve as an absolute measure for the predictive power of a model, the monotonous decrease of the ROC2-correlation with n_m indicates, that a lower ROC2-correlation still corresponds to a better description of human behavior by the tested model. Thus the ROC2-correlation still provides an additional measure for the evaluation of our contour integration models.

7.1.8 Correlations among human observers

Now we apply the ROC-correlation analysis on the data from the psychophysical contour integration experiment. Fig. 7.9 (a) shows the ROC-curve for human decisions for a stimulus presentation time of 200 ms. To have a measure for the error and the significance of this result we obtained surrogate data by randomly shuffling the order of human responses within each subset of 48 trials. This

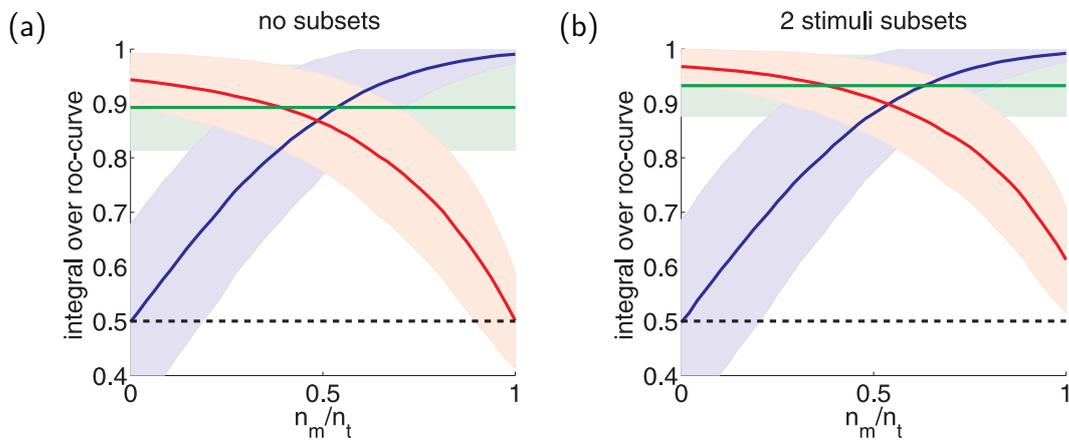


Figure 7.8: ROC-correlations in dependence of the match n_m/n_t between the analyzed model and the generic model underlying the observer responses: (a) shows ROC1-correlation between model and observers (blue line), the ROC2-correlation (red line) and the correlation between different observers (green line) averaged over 1000 data sets of $n_t = 48$ trials each. The shaded areas around the curves denote one standard deviation. Here we assumed, that the probability for deviations of the observer responses from the responses of the generic model is equal for each trial. (b) shows the same assuming two subsets of stimuli with different probability $(1 - p_m)$ for mismatches between the decisions of the generic model and the observer's responses. Here the ROC2 curve does not reach the benchmark of 0.5 (black dashed line).

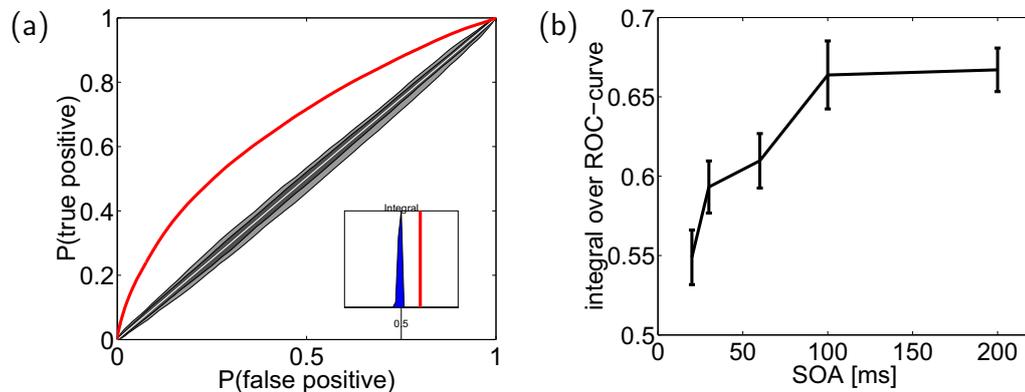


Figure 7.9: (a) ROC-curve (red line) representing the correlations between human observers for 200 ms SOA. The diagonal shows the ROC curve for random decisions. The three grey shaded areas surrounding the diagonal illustrate three standard deviations for random decisions. They were computed from 100 surrogate ROC curves which were obtained by randomly shuffling the decisions of each human observer in each trial sequence before computing the ROC-curve. The inset shows the integral over the randomly shuffled curves (blue) and over the ROC curve of the human observers (red). (b) Correlation between different human observers for different stimulus presentation times. The error bars denote one standard deviation.

uncorrelated data set yields ROC-curves close to the diagonal. The grey shaded area in Fig. 7.9 (a) shows 3 standard deviations of these curves obtained from the shuffled data. This gives a measure which shows whether the ROC-curve significantly differs from ROC-curves obtained from uncorrelated data sets of finite size. The red curve representing the correlations of the human observers is clearly above the grey shaded area which means that human decisions are indeed significantly correlated. The degree of correlation is represented by the area under the ROC-curve, whose integral is a very convenient measure for the correlations. The integral over the ROC-curve for 200 ms SOA is 0.67. The area under the diagonal is 0.5 ± 0.01 in $1-\sigma$ -precision. Fig. 7.9 (b) shows the integral over the ROC-curve for different stimulus presentation times. Not surprisingly the correlation among human observers increases with higher SOA, meaning their decisions become less random the longer they see the stimulus.

For the experiments two different sets of stimuli were used. Only for 200 ms SOA all five observers saw the same set of stimuli, allowing here to average the correlations of 10 different observer combinations. In contrast, for 100 ms SOA one set of stimuli was employed for the first two observers, and the remaining three subjects did the experiment with the second set of stimuli. This means that only four comparisons between observers having seen the same stimuli are possible. Experiments with all other SOA times were performed with four ob-

servers searching contours in the same stimuli. These limitations in the existing data cause higher uncertainties in the correlations derived for certain SOAs.

For the ROC-analysis we distinguished stimuli sets of different curvature and alignment jitter and of different element distance. However, the performance also varies with eccentricity of the contour. Hence in principal one should also distinguish between different eccentricities and take their performance into account. However, as it was not the original intention of the psychophysical study to analyze the effect of eccentricity, the stimuli were not generated such that there is a given number of contours for certain ranges of eccentricity. Still one could simply sort all stimuli for a given σ_j and r_0 by their eccentricity and divide them into several classes of eccentricities. However, a stimuli set for one combination of curvature and alignment jitter and element distance only contains 48 stimuli. When reducing the number of stimuli even further, the performance obtained becomes increasingly inaccurate. Also the number of expected identical responses of two observers and thus the whole ROC-analysis suffers from a huge amount of noise. For this reason the given data does not allow to determine to what extent the correlation between human observers is caused by the eccentricity- dependence of the performance and to what extent it might be caused by any other influences which we did not take into account so far.

7.1.9 Benchmarks

The benchmark for our models has to be the human behavior. So the first thing to do is to compare the contour detection performance of any model with the performance of humans in psychophysical contour detection experiments. Human contour detection performance increases with longer SOA (compare Fig. 7.9). As a plausible contour integration model should also explain the highest observed human performance, we compare the model performance with the human performance at a contour processing time of 200 ms. Usually one would rule out a model that exceeds the human performance by far more than 3σ . However, human contour integration is subjected to different kinds of noise, but in our models there is no noise included. Noise would certainly decrease the contour detection performance (compare chapter 6 for the Bayesian model). Furthermore it is possible that human contour detection performance improves even further when increasing the stimulus presentation time above 200 ms and we do not know which iteration time corresponds to which stimulus presentation time. Hence we require a plausible model to reach or exceed the average human performance. This, however, should not only be the case for the human performance averaged over all element distances, curvature and alignment jitters, and eccentricities, but also for the performance of each combination of these parameters. Furthermore we saw in section 7.1.3 that human performance decreases with eccentricity, with higher curvature and alignment jitter, and with higher element distance. Hence

we want to find the dependencies on ϵ , σ_j , and r_0 also in the model performance. A performance that for example exceeds the human performance for each eccentricity, but does not decrease with eccentricity, is not a plausible model.

Performance is an important and the most commonly used benchmark for contour integration models. However, we will see that there are many models which exceed human performance. Hence performance alone is not sufficient to decide which model explains human behavior best. Given that human contour detection behavior is highly correlated, it is logical to request in addition that the predictions of a plausible contour integration model are correlated to the human responses. As human behavior in contrast to the model is subject to noise, these correlations should be at least as high as the correlations among humans. We found different correlations among humans for different SOA. The correlations between model and humans for different SOA should each reach or exceed the corresponding human-human correlation. We do not know the correlation between our five human subjects and the underlying model. For large n_{ob} the decisions of the generic model could be obtained for each trial from the responses given by the majority of observers, and correlations between the responses of the generic model and single observers could be analyzed. However, in order to perform this investigation with sufficient statistics we would need many more subjects and many more experiments. Hence we do not know how much bigger than the human-human correlation the correlation between the observer and the model should be. However, no matter what amount of noise the human behavior is subjected to, the ROC2-analysis gives us the possibility to check whether our model systematically differs from the model underlying human behavior. So the ROC2-analysis should be as low as possible. Ideally it should reach 0.5 which would mean that the model makes no systematic errors, when considering the generic human behavior as null hypothesis. For a given SOA, all those conditions need to be fulfilled by the model at the same iteration time.

Not surprisingly none of the investigated models fulfills all those requirements, especially that of ROC2-correlations of 0.5 which would be only given for the best possible model that exactly describes the model underlying human behavior and unbiased observers. Nevertheless these criteria help us to rule out several model and to decide, which of the analyzed models explains human behavior best.

7.1.10 Simulations and model performance

Numerical complexity

All simulations presented in this chapter were performed with an afferent input width of $\sigma_{aff} = \pi/8$ which lies well within the range of experimental results for tuning curves of orientation selective cells in V1 (i.e. (71; 102)). As discussed in

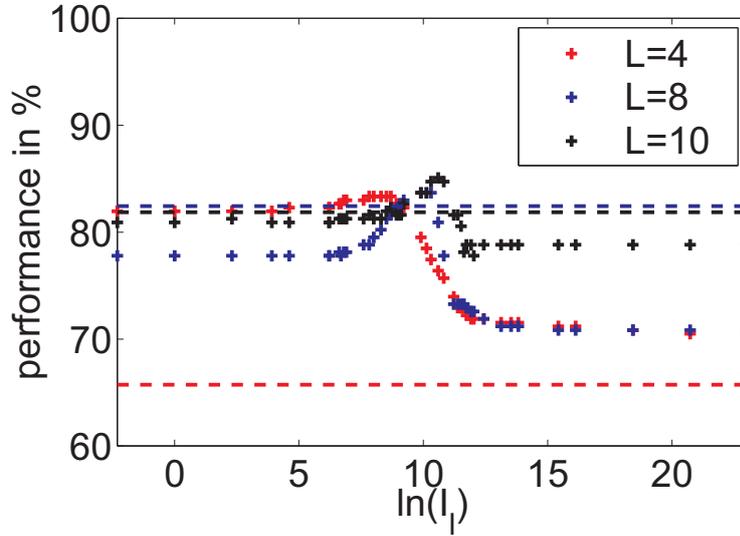


Figure 7.10: The crosses show the maximum performance of the additive model reached during iteration time in dependence of I_l for constant I_a . The performance was averaged over all orientation jitter. The dashed lines illustrate the corresponding human performance after 200 ms SOA.

4.3.1 we use $n_{\text{ori}} = 18$ preferred orientations which is a good compromise between reducing computation time and artifacts from low n_{ori} . For sufficient accuracy in solving the differential equation 4.12, simulations were performed with iteration time steps of $\Delta t = 0.002\tau$. However, to cut down computation time, the correlation analysis was done only for every 16th time step. Depending on the inter element distance r_0 and the range of the association field it takes between 1 and 48 hours on a Pentium 4 processor at 3,2 GHz to compute the performance of the additive and multiplicative model for just one stimulus subset containing 48 contours of given r_0 and σ_j . In order to perform a complete correlation analysis this procedure needs to be repeated for all 84 subsets of the two complete sets of stimuli which were used in the psychophysical experiments.

Parameter scaling

As already described in section 5.3 the additive model relies on the adjustment of I_l and I_a such that afferent and lateral input have a balanced influence. The ratio I_l/I_a where the model performs best changes with the stimulus as well as with the properties of afferent input and association field.

Due to normalizing the total activity $\sum_{x,\phi} A(x, \phi) = 1$ it yields $A \propto 1/N$ where N is the number of edge elements in the display. Hence $I_{\text{lat}} \propto 1/N$ while I_{aff} is independent of N , which means that we had to adjust I_l/I_a for different

N in order to find the best performance for each N . Stimuli with $L = 10$ which contain on average $\langle N_{10} \rangle = 342$ elements while stimuli with $L = 4$ consist on average of only $\langle N_4 \rangle = 55$ Gabors. Assuming that horizontal interactions link only nearest neighbors or next to nearest neighbors, respectively, this means that the maximum performance for $L = 10$ is shifted to $\langle N_{10} \rangle / \langle N_4 \rangle = 6.2$ times higher values of I_l compared to the maximum performance for $L = 4$. If we choose an association field of constant range independently of r_0 , we also have to take into account, that the number of cells from which a cell gets input increases with smaller r_0 such that I_l should decrease again. For an association field with constant range this is shown in Fig. 7.10. Apart from the range of the association field also the width $\sigma_\alpha = \sigma_\beta$ of the association field has an influence on the average lateral input and hence on the optimal I_l/I_a . Also we expect a lower average I_{aff} when eccentricity dependence is taken into account than without eccentricity dependence.

When searching for a model of optimal contour detection performance we had to adjust I_l/I_a for each stimulus and each association field. However, from a biological point of view it is rather unlikely that the brain changes its synaptic weight for each stimulus. Hence we search for a ratio I_l/I_a which leads to reasonable results for all stimuli. If possible we try to adjust I_l/I_a such that performance for each L reaches or exceeds the corresponding human performance.

In principle we should not only optimize the performance for each L , but also ROC1 and ROC2 correlations. Therefore we would have to perform simulations for the two complete sets of stimuli for several values of I_l while keeping I_a constant. However, because of the enormous additional computational effort this would mean, we only optimized the performance. From a few examples where we computed the maximum performance in dependence of I_l for $L = 4$, $L = 8$, and $L = 10$ we know how the maximum performance is shifted with higher L . For the remaining input and association field parameters we first look at the maximum performance for $L = 4$ as this is determined comparatively quickly and try to infer a suitable I_l . If we are in doubt we also compute the performance for $L = 8$. This solution allows us to do the adjustment of I_l and I_a with reasonable computation time. However, with this method it is still possible that the results for the additive model could be improved when adjusting I_l/I_a on the basis of the full data and taking into account both, performance and correlation. But as one requirement for a plausible model is that its performance reaches or exceeds human performance and the additive model fulfills this requirement only for a limited range of I_l/I_a , especially for the more plausible choices of afferent input and association field, the balance of afferent and lateral input obtained when taking into account correlation and performance cannot differ much from our choice of I_l/I_a .

We did not include noise in the simulations as this would require averaging

over even more trials and hence increase the computational effort even further.

7.1.11 Comparison of different estimators and readout mechanisms

In 4.2.2 we introduced two different ways of defining the saliency, either by summing or by taking the maximum of the activities within one hypercolumn. For very simple stimuli these two salience estimators were already discussed in chapter 5. In this special case we saw no difference between s^{\max} and s^{sum} for the non-linear model, but for the additive model the summation estimator was much more sensitive to star-illusions than the maximum-estimator (see 5.1.2). However, lacking psychophysical experiments revealing how sensitive humans are to star-illusions, we cannot rule out either estimator. As a consequence we did all simulations for both estimators.

As one example, performance and correlations for the model without ϵ -dependence and with fixed-range unidirectional association field of the same angular dependence as the generating association fields is depicted in Fig. 7.11. There is little difference between the two estimators for the multiplicative model. However, for the model with linear coupling between different inputs we see a much higher performance for the maximum estimator. Also the correlations between model and human behavior are higher for s^{\max} than for s^{sum} and the ROC2-correlations, where a low correlation is desirable, are lower for the maximum estimator. Even though we show only one example here, we observed similar results for all simulations we performed. This suggests that the human brain rather uses the maximum estimator than the summation estimator and hence we would conclude, that humans are not as sensitive to star illusions as the linear model with summation estimator.

After obtaining a saliency for each element we have to make a decision, which hemifield contains the contour. In 4.2.3 we introduced several possible contour detection criteria. Either one builds the sums $s_{\text{RH}} = \sum_{i \in \text{RH}} s(i)$ and $s_{\text{LH}} = \sum_{i \in \text{LH}} s(i)$ over all saliences within the right hemifield (RH) and the left hemifield (LH), respectively. Then the contour is detected in the right hemifield if $s_{\text{RH}} > s_{\text{LH}}$ and vice versa in the left hemifield if $s_{\text{LH}} > s_{\text{RH}}$. Another possible contour detection criterion is to detect the contour on the hemifield that contains more than half of the L_{\max} highest saliences $s(i)$.

Performance and correlations for these different contour detection criteria are also shown in Fig. 7.11. For the non-linear model there are little differences between the contour detection criteria. However, already here we obtain better results, i.e. the highest performance and ROC1-correlation and the lowest ROC2-correlation, for higher L_{\max} . Summing over the complete hemifield should mathematically be the optimal strategy to detect contours in probabilistic models

and we find that performance and correlations for this contour detection criterion are similar to the results for the highest L_{\max} . For the additive model the differences between performance and correlations obtained with different contour detection criteria increase. Here it becomes even clearer that within the tested L_{\max} the performance and the predictive power of the model increase with L_{\max} . However, in contrast to the multiplicative model, here summing over hemifields is the worst contour detection criterion of all. The reason for this could be that the differences in the computed salience for contour and background elements is much smaller for the additive model than for the multiplicative model. So if the salience of the hemifield which does not contain the contour is increased by some smaller or strongly jittered contours which arise by chance in the background, in the additive model this salience is more likely to exceed the salience of the hemifield containing the contour than in the multiplicative model. Even though we just show one example here, the result that the maximum estimator with $L_{\max} = 5$ seems to be the best of the tested estimators and contour detection criteria is consistent with the results of all other simulations. Hence in the following we will only show performances or correlations which were obtained from s^{\max} with $L_{\max} = 5$.

7.2 Results of model evaluation

In the following we will analyze different choices for model dynamics, association fields and afferent input functions. For this exploration of different models we performed many simulations of which the essential results will be illustrated and discussed here in more detail. An overview over the results of all simulations is given in table 7.1.

We started by investigating both, the additive and the multiplicative model as described by Eq. 4.12 with an afferent input given by Eq. 4.16. This afferent input depends solely on the preferred orientation of an orientation column and the orientation of the Gabor patch displayed in the receptive field of this neuronal population. As a plausible contour integration model has to explain the high human performance, we first use for each stimulus an association field with the same angular dependence as the association field employed to generate this stimulus. This way we expect to obtain an optimal performance. For the spatial extension of the association field we employed the distribution r_{tot} over the distances between nearest neighboring contour elements of all stimuli (see Fig. 7.12).

As we see in Fig. 7.13 (a) the averaged simulation performance clearly exceeds the human performance over a certain range of iteration time and hence meets the first requirement for a plausible model formulated in 7.1.9. To analyze how well human performance P_{human} for each combination of jitter and element distance is

7.2. Results of model evaluation

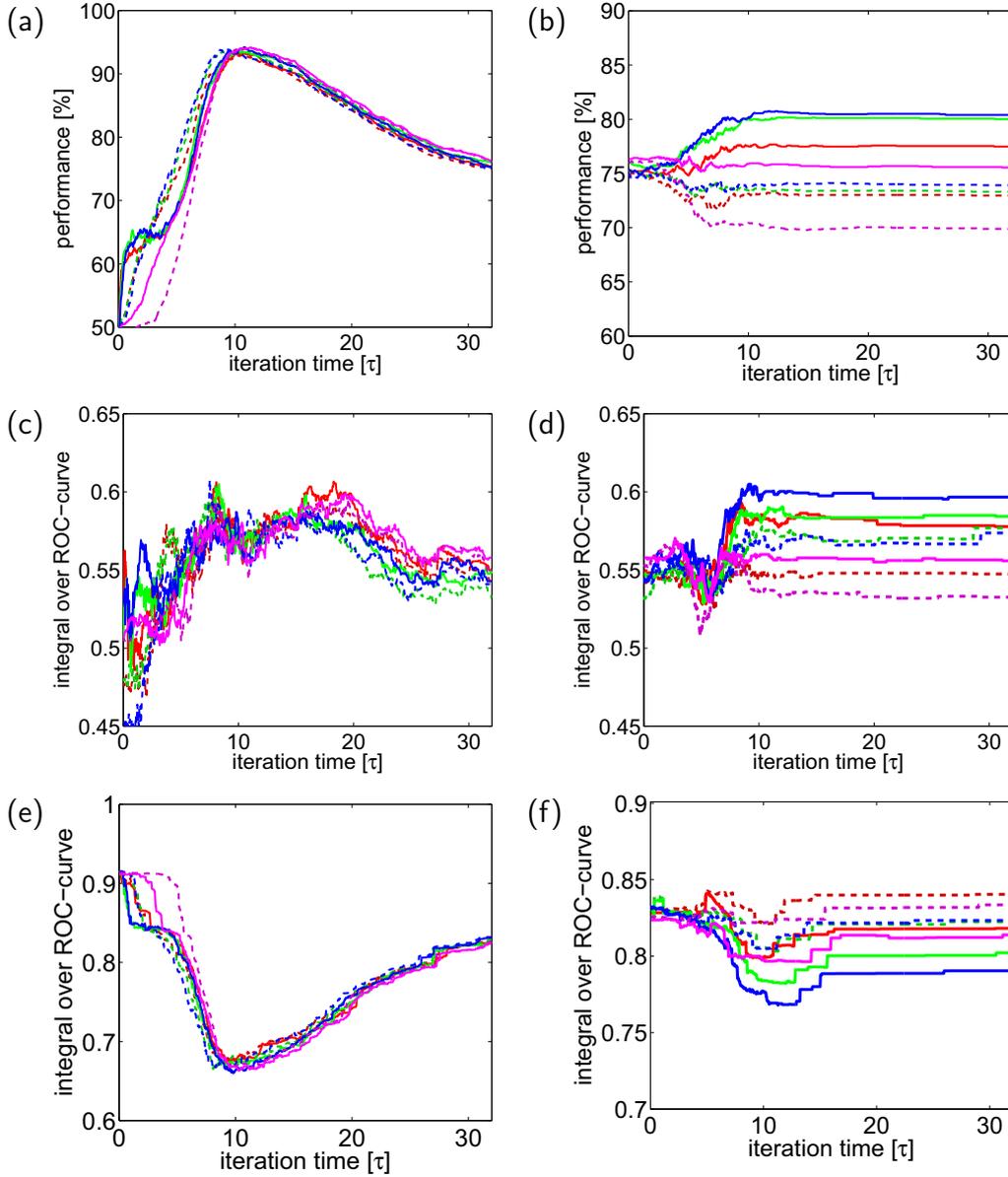


Figure 7.11: The first column depicts for the multiplicative model (a) the performance, (c) correlation to human behavior and (e) correlation between deviation of human responses from model prediction. (b), (d) and (f) show the same for the additive model. The solid lines show these magnitudes as derived from s^{\max} , the slightly darker dashed lines show them as obtained from s^{sum} . The different colors correspond to different contour detection criteria, namely detecting the contour on the hemifield where more than half of the $L_{\max} = 1$ (red line), $L_{\max} = 3$ (green line) or $L_{\max} = 5$ (blue line) elements of maximum salience can be found. For the pink/purple line the contour was detected on the basis of the saliences for the total hemifields s_{LH} and s_{RH} , respectively.

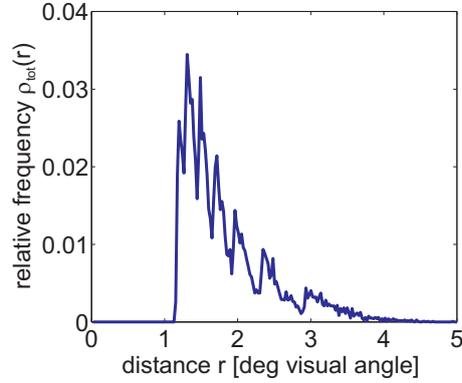


Figure 7.12: Distribution ρ_{tot} of inter-element distances between nearest neighboring contour elements for the complete set of stimuli used for the experiment.

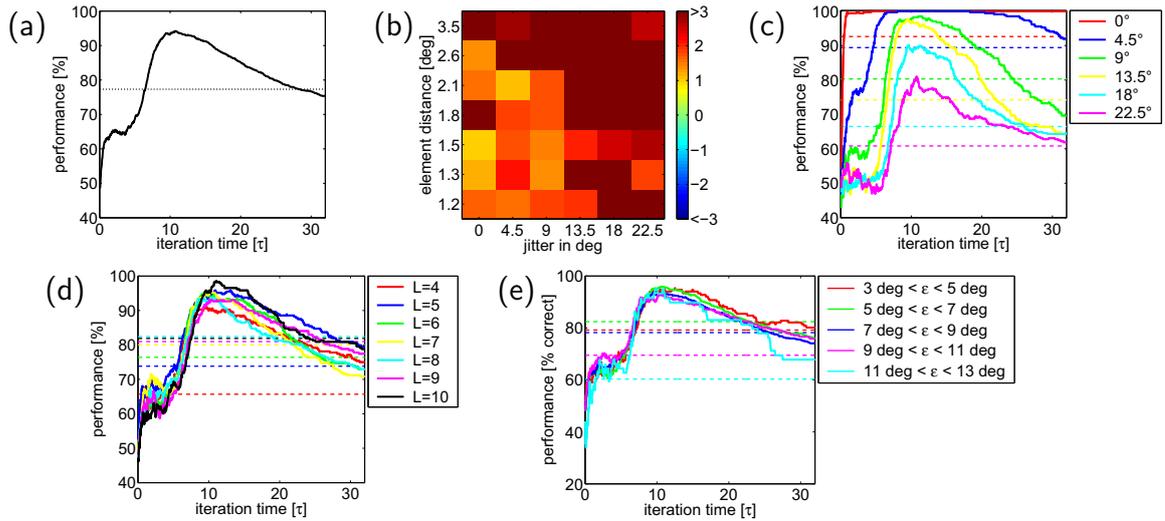


Figure 7.13: Performance of the model with unidirectional, multiplicative couplings and ϵ -independent afferent input. The AF has ideal angular dependence and range ρ_{tot} taken from the nearest neighbor distribution summed over all stimuli: (a) Performance averaged over all jitter and element distances. The solid line shows the model performance, the dotted line the human performance for 200 ms SOA. (b) ΔP for all combinations of σ_j and r_0 . (c) Performance for different jitters averaged over all element distances and eccentricities. (d) Performance for different element distances averaged over jitters and eccentricities. (e) Performance in dependence of mean eccentricity of the elements of a contour averaged over jitters and element distances. For (c-e) solid lines denote model performance while dashed lines of the same color illustrate the corresponding human performance.

7.2. Results of model evaluation

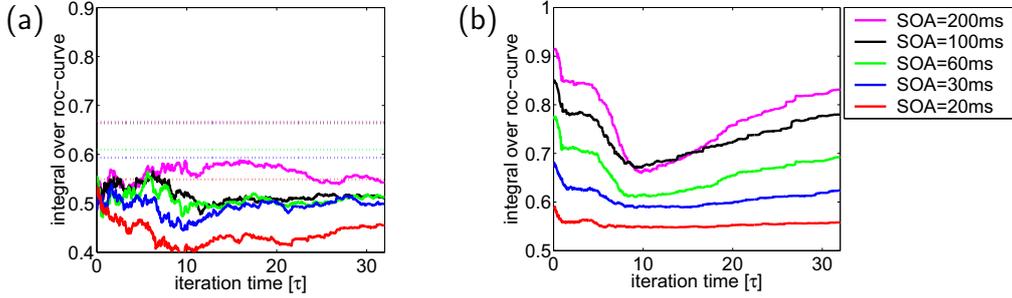


Figure 7.14: Correlations of the model with unidirectional, multiplicative coupling, ϵ -independent I_{aff} , and an AF with ideal angular dependence and range ρ_{tot} : (a) Solid lines show the correlations between model predictions and human responses color coded for different SOA (see legend in (b)). Dotted lines show the correlations between human responses of different subjects for the stimulus presentation times shown in the same color code. (b) Correlations between deviations from human responses from model predictions for different SOA (see legend).

reached by the model performance P_{model} we compute $\Delta P = \frac{P_{\text{model}} - P_{\text{human}}}{\sigma_{\text{human}}}$ where σ_{human} is the standard deviation of human performance. Ideally ΔP should exceed -3 for all combinations of r_0 and curvature and alignment jitter σ_j . For the analyzed model $\Delta P > 0$ for all r_0 and σ_j with especially high values for strong jitters and large element distances (Fig. 7.13 (b)). Due to the ceiling effect human performance for the easier tasks cannot be exceeded by far. The model performance increases with decreasing curvature and alignment jitter until it reaches 100% and cannot increase any more (Fig. 7.13 (c)). This reproduces the trend of the jitter dependence of the human contour detection performance. Like human performance the model performance decreases also for larger element distances, which in our case is equivalent to having fewer contour elements L (see Fig. 7.13 (d)). In 7.1.3 we saw that human performance decreases with eccentricity. Hence our model should show the same eccentricity dependence. However, as neither afferent input nor the association field is eccentricity-dependent, the performance is independent of eccentricity (see Fig. 7.13 (e)). As the model does not show any decrease of performance in the periphery it is not surprising, that the correlations between these model's predictions and human responses do not reach the correlations among the responses of human observers (see Fig. 7.14). Nevertheless there is comparatively little correlation between deviations of the human responses from the model prediction, indicating that the model already partly explains human behavior even though it neither reproduces the eccentricity dependence of the human performance nor does it explain the correlations between human observers.

Let us now look at the additive model with the same association field and

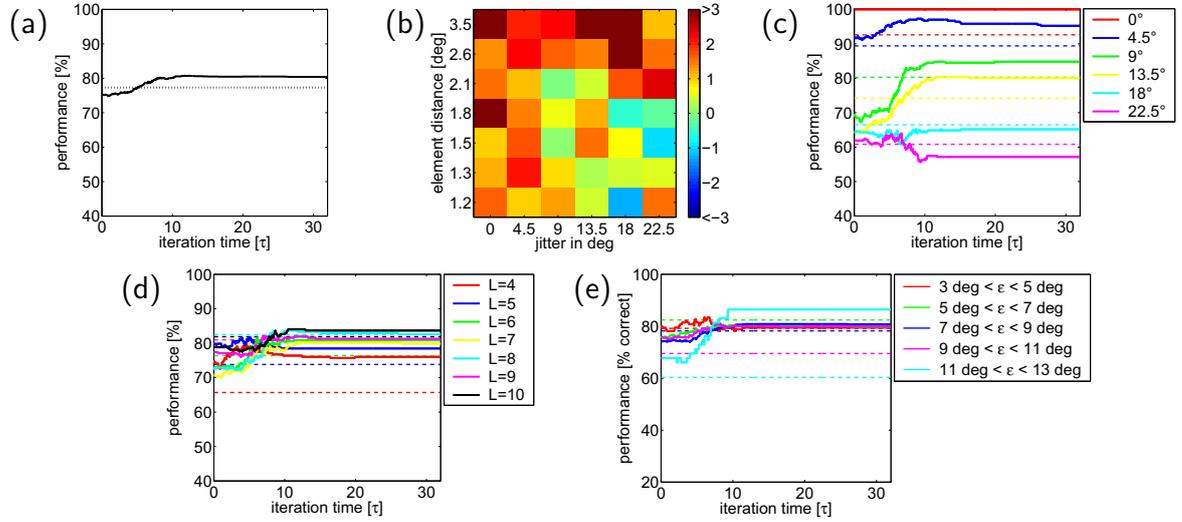


Figure 7.15: Performance of the model with unidirectional, additive couplings, and eccentricity-independent afferent input. The AF has ideal angular dependence and a range ρ_{tot} taken to be identical to the nearest neighbor distribution ρ_{tot} summed over all stimuli: (a) Performance averaged over all jitters and element distances. The solid line shows the model performance, the dotted line the human performance for 200 ms SOA. (b) ΔP for each combination of σ_j and r_0 . (c) Performance for different jitters averaged over all element distances and eccentricities. (d) Performance for different element distances averaged over jitters and eccentricities. (e) Performance in dependence of mean eccentricity of the elements of a contour averaged over jitters and element distances. For (c-e) solid lines denote model performance while dashed lines of the same color illustrate the corresponding human performance.

afferent input as the multiplicative model just shown.

The performance of the model with additive couplings is lower than that of the model with nonlinear dynamics. However, it still exceeds the average human performance and is in fact much closer to the average human performance (Fig. 7.15 (a)). Also $\Delta P > -3$ for all σ_j and r_0 combinations. In fact there are few combinations for which $\Delta P < 0$ (Fig. 7.15 (b)). When looking at the dependence of the performance on jitter, element distance, and eccentricity we see the same trends as in the multiplicative models: decreasing performance with higher element distance and alignment, but no change of performance with higher eccentricity (Fig. 7.15 (c-e)).

The correlation of the model with the human responses seems to be a little higher for the linear model than for the model with non-linear dynamics. However, the correlation between human observers is not reached by the model for any of the investigated SOA values. This suggest that neither the additive nor

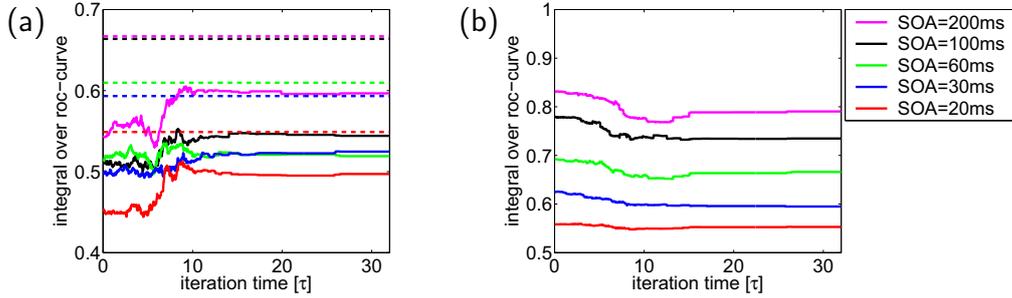


Figure 7.16: Correlations of the model with unidirectional, additive coupling, eccentricity-independent afferent input, and an AF with ideal angular dependence and range ρ_{tot} : (a) Solid lines show the correlations between model predictions and human responses color coded for different SOA (see legend in (b)). Dotted lines show the correlations between human responses of different subjects for the stimulus presentation times using the same color code. (b) Correlations between deviations from human responses from model predictions for different SOA (see legend).

the multiplicative model with these properties of afferent and lateral input is capable of explaining human behavior. However, when comparing the correlation between deviations of the human responses from the model predictions, we see much lower correlations for the multiplicative model, which implies that the multiplicative model at least for these parameters explains human behavior better than the additive model. In both models we have the least correlation between deviation of human responses from model predictions for the shortest SOA. The reason for this is that for short SOA the correlation between human responses also decreases and there is less systematic behavior which needs to be explained by a contour integration model. As the models can not explain the performance decrease with eccentricity, this suggests to insert eccentricity dependence into the model.

Similar as in the presented example the additive model compared to the multiplicative model is worse in performance and ROC2 correlation for most tested model choices. For this reason we will focus on the multiplicative model in this discussion and come back to the additive model after having finished the exploration of eccentricity scaling and association fields for the multiplicative model. However, all results for the additive model are summarized in table 7.1.

7.2.1 Eccentricity dependence

In order to account for the eccentricity dependence of the human performance we used eccentricity modulated afferent input (see Eq. 4.18). As afferent input

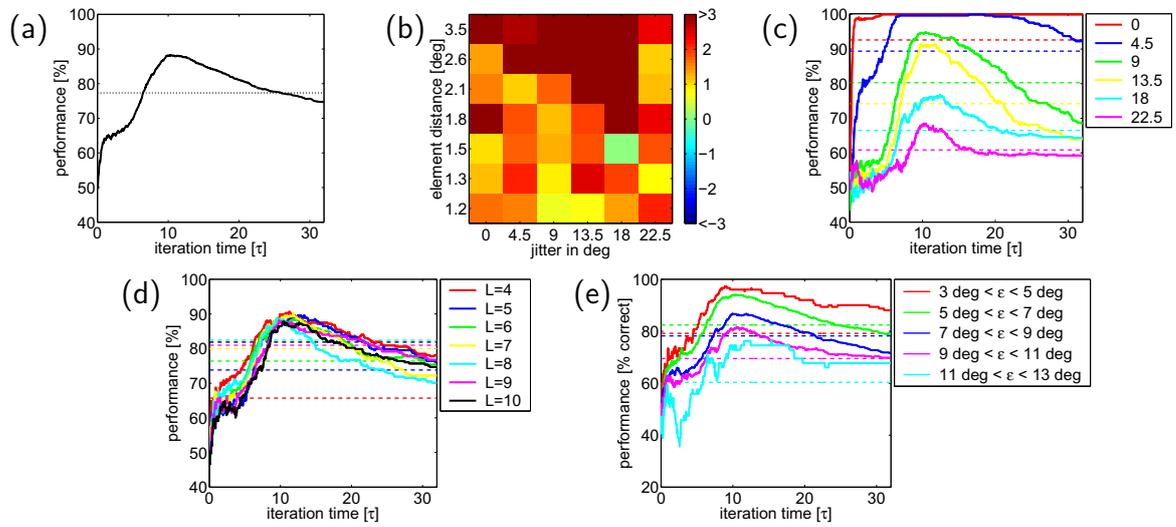


Figure 7.17: Model performance for the model with unidirectional, multiplicative couplings and eccentricity-dependent afferent input. The AF has ideal angular dependence and range ρ_{tot} : (a) Performance averaged over all jitter and element distances. The solid line shows the model performance, the dotted line the human performance for 200 ms SOA. (b) ΔP for different combinations of r_0 and σ_j ; (c) Performance averaged over all element distances and eccentricities. (d) Performance for different element distances averaged over jitters and eccentricities. (e) Performance in dependence of mean eccentricity of the elements of a contour averaged over jitters and element distances. For (c-e) solid lines denote model performance while dashed lines of the same color illustrate the corresponding human performance.

in the periphery is lower than close to the fovea the performance for eccentric contours and hence also the averaged performance is decreased. Fortunately, the performance of the multiplicative model is still far above the average human performance (Fig. 7.17 (a)). Model performance for very easy tasks is slightly below human performance, but still $\Delta P > -1$ for all σ_j and r_0 (Fig. 7.17 (b)).

As before the performance decreases with higher jitter σ_j (Fig. 7.17 (c)). However, it seems now that the performance decreases with smaller element distances. This is exactly the opposite of what the human behavior shows. This problem arises not only for this parameter combination, and one possible explanation which has to do with the range of the association field will be discussed in subsection 7.2.4 where the influence of the range of the association field on the model performance and correlation is analyzed.

Now the performance decreases with eccentricity. The maximum performance of the model still exceeds the performance of human observers for each analyzed eccentricity (see Fig. 7.17 (e)).

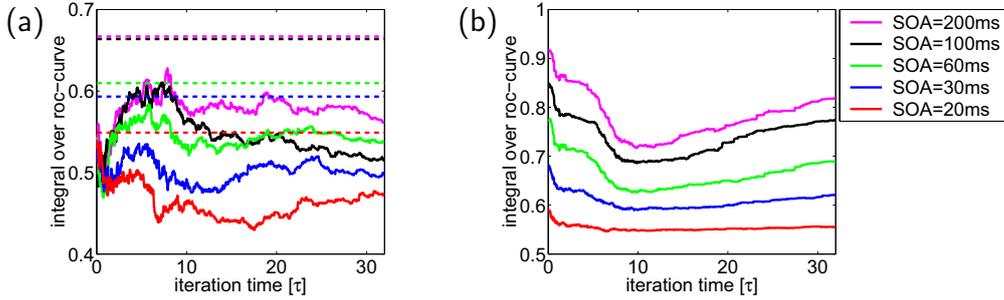


Figure 7.18: Correlations of the model with unidirectional, multiplicative coupling, eccentricity-dependent afferent input, and an AF with ideal angular dependence and a range ρ_{tot} obtained from the nearest neighbor distribution summed over all stimuli: (a) Solid lines show the correlations between model predictions and human responses color coded for different SOA (see legend in (b)). Dotted lines show the correlations between human responses of different subjects for the stimulus presentation times using the same color code. (b) Correlations between deviations from human responses from model predictions for different SOA (see legend).

The correlation is now higher than for the multiplicative model without dependence on ϵ . However, the correlation among human observers is still not reached by the correlations between the multiplicative model and humans. Hence it is surprising that the ROC2-correlations are higher, which suggests that this model explains less of the human behavior than the multiplicative model without ϵ dependence.

7.2.2 Angular dependence of the association field

For the probabilistic model we would expect an optimal performance when using an association field with the same angular dependence as the generating association field. However, this would mean using a different association field for different stimuli. Even if this was possible, without prior knowledge about the contour it is not possible to decide which is the most suitable association field for the shown stimulus. From a biological point of view we would expect that there is a fixed association field implemented in the brain, which might be adapted to the statistics of natural scenes, but which is independent of the stimulus the observer is viewing. Hence we performed simulations employing a fixed association field for which we chose $\sigma_\alpha = \sigma_\beta = 18 \text{ deg}$. This association field resembles the spatial distribution of surround interactions around a neuron's receptive field in V1 (60). As shown in Fig. 7.20 (a) here the average performance decreases compared to simulations using an association field with optimal angular dependence. However, it still exceeds human performance. When looking at the performances for

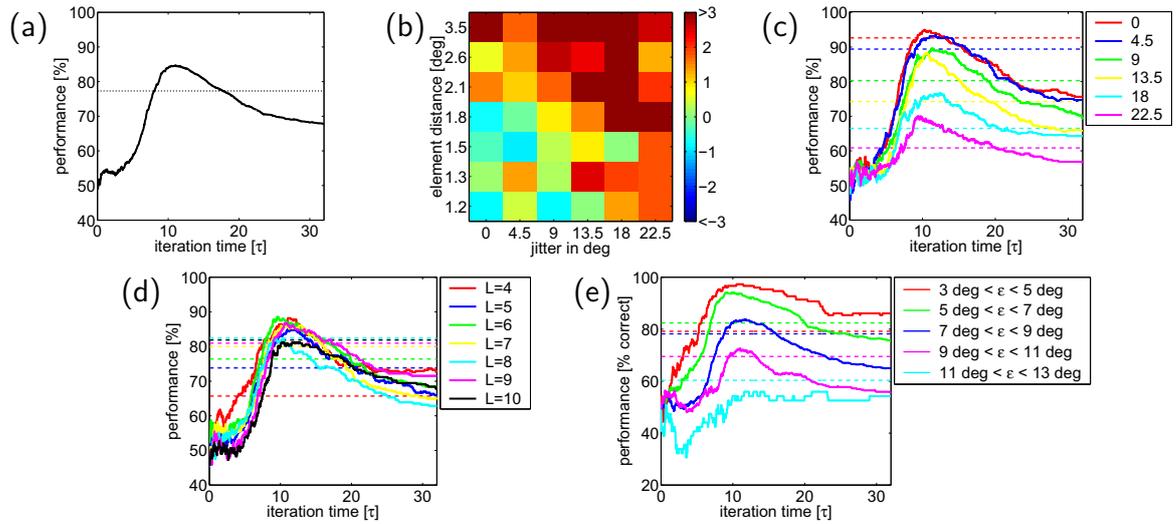


Figure 7.19: Performance of the model with unidirectional, multiplicative couplings and eccentricity-dependent afferent input. The AF has fixed angular dependence $\sigma_\alpha = \sigma_\beta = 18$ deg and range ρ_{tot} : (a) Performance averaged over all jitters and element distances. The solid line shows the model performance, the dotted line the human performance for 200 ms SOA. (b) ΔP for different combinations of r_0 and σ_j (c) Performance for different jitters averaged over all element distances and eccentricities. (d) Performance for different element distances averaged over jitters and eccentricities. (e) Performance in dependence of mean eccentricity of the elements of a contour averaged over jitters and element distances. For (c-e) solid lines denote model performance while dashed lines of the same color illustrate the corresponding human performance.

7.2. Results of model evaluation

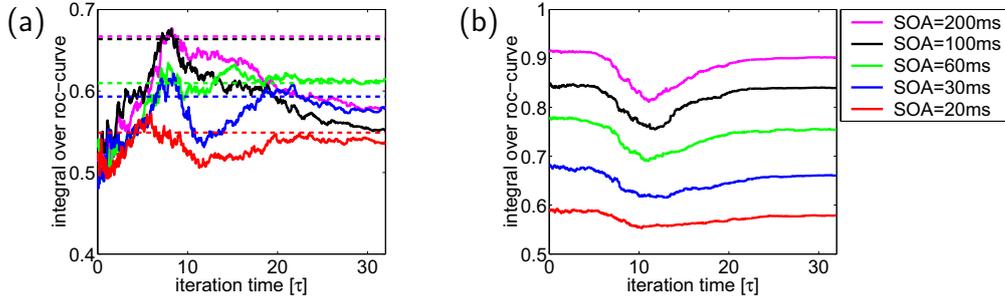


Figure 7.20: Correlations of the model with unidirectional, multiplicative coupling, eccentricity-dependent afferent input, and an AF with ideal angular dependence and range ρ_{tot} : (a) Solid lines show the correlations between model predictions and human responses color coded for different SOA (see legend in (b)). Dotted lines show the correlations between human responses of different subjects for the same color coded stimulus presentation times. (b) Correlations between deviations from human responses from model predictions for different SOA (see legend).

each jitter and element distance separately, they lie within or above the 3σ -range around human performance (Fig. 7.20 (b)). Not surprisingly we find particularly high performances for jitters close to the width $\sigma_\alpha = \sigma_\beta = 18$ of the association field for contour integration. As required the performance decreases with jitter and eccentricity (Fig. 7.19 (c+e)). As in Fig. 7.17 the performance does not decrease with wider element distances (Fig. 7.19 (d)).

When looking at the correlation we see, that with the change from the generating association field to a fixed association field with properties similar to those found in experiments, the multiplicative model with ϵ -dependent I_{aff} can reproduce the correlations among human subjects for all SOAs at least for a certain iteration time. The ROC-curve for SOA=200 ms reaches its maximum at an iteration time where the model performance is close to the human performance after 200 ms stimulus presentation time. Then performance increases further and the ROC1-curve decreases again.

We assume that contour integration is an iterative process. Hence we would expect the iteration time t for which the correlation between model predictions and human responses for a certain SOA surpass the human-human correlation or reach a maximum should be proportional to the stimulus presentation time. However, in Fig. 7.20 (a) this shift to longer iteration times for higher SOAs cannot be observed. Also the minimum of the ROC2-correlation is reached at about the same iteration time for different SOA (7.20 (b)). Furthermore the minimum of the ROC2-curves comes at later iteration times as the maximum of the ROC1-correlations. Hence neither ROC1 nor ROC2 correlations can be used

to map iteration time to stimulus presentation time.

Here ROC2-correlations are comparatively high. So even though this model finally fulfills the benchmark for ROC1-correlations we seem not to have found a model which explains human behavior satisfactorily.

When using only the AF with fixed width but no ϵ -dependence, the ROC1-correlation again decreases and cannot explain human behavior (corresponding graphs not shown, see table 7.1 for benchmark values).

7.2.3 Symmetry of association field

It is often assumed that long-range horizontal interactions in V1 serve as the neuronal substrate for the association field (36). So far we assumed unidirectional lateral interactions as required by probabilistic models. These link orientation columns in only one direction and should lead to an optimal performance. However, experimental findings in monkeys rather suggest isotropic connections, spreading symmetrically into all directions from an orientation column. This inspired us to also investigate contour integration with a bidirectional association field. Here we see that for the multiplicative model the average model performance does not reach human performance (Fig. 7.21 (a)). However, in contrast to the unidirectional model the performance only slightly decreases after reaching its maximum. When looking at the performance for each σ_j and r_0 we see that especially for easy tasks the human performance is not reached and not all performance values lie within the 3σ -range of the human performance (Fig. 7.21 (b)).

A possible explanation for the low performance is that if by chance there are two collinearly aligned edge elements, the bidirectional model will judge them as being more salient than a much longer, but slightly jittered contour. Consequently, it finds much more contours in the background as the unidirectional model. Also it evaluates zigzagged contours as being equally salient as straight contours (see Fig. 4.8). This also increases the probability to find illusory contours in the background.

When looking at the performance for each element distance averaged over all jitters (Fig. 7.21 (c)), we see that in contrast to human performance the model computes the highest performance for the largest element distances and vice versa. This was already the case for the unidirectional association field. However, here it is easier to understand than in the unidirectional case.

The probability to find two collinearly aligned edge elements in the background increases with the number of background elements, which is highest for the smallest inter-element distances. Hence we would expect a bigger impact of the detection of these short contours in the background for smaller inter-element

7.2. Results of model evaluation

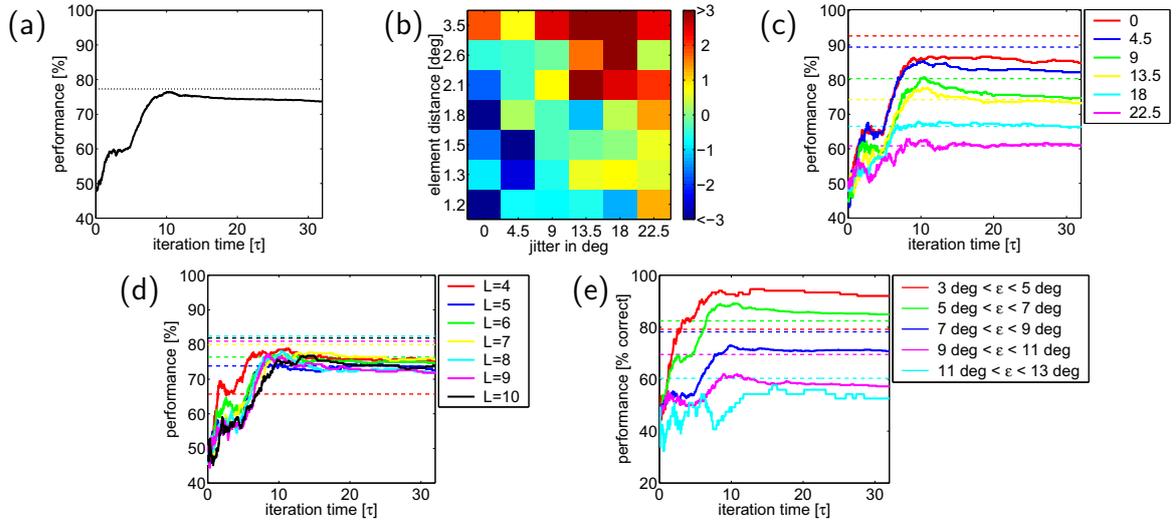


Figure 7.21: Performance of the model with bidirectional, multiplicative couplings and eccentricity-dependent afferent input. The AF has fixed width $\sigma_\alpha = \sigma_\beta = 18$ deg and range ρ_{tot} : (a) Performance averaged over all jitters and element distances. The solid line shows the model performance, the dotted line the human performance for 200 ms SOA. (b) ΔP for different combinations of r_0 and σ_j (c) Performance for different jitters averaged over all element distances and eccentricities. (d) Performance for different element distances averaged over jitters and eccentricities. (e) Performance in dependence of mean eccentricity of the elements of a contour averaged over jitters and element distances. For (c-e) solid lines denote model performance while dashed lines of the same color illustrate the corresponding human performance.

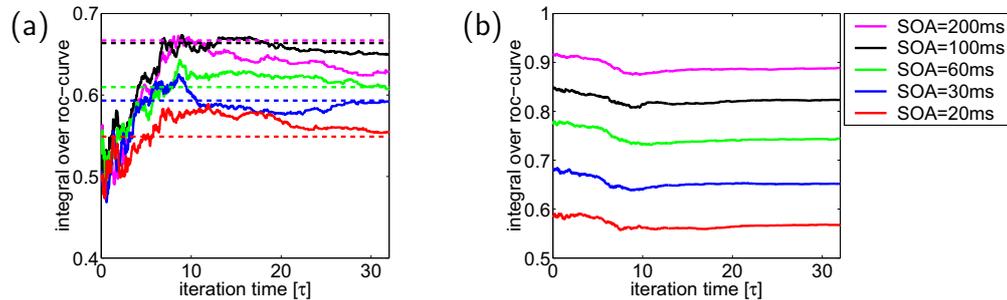


Figure 7.22: Correlations of the model with bidirectional, multiplicative coupling, eccentricity-dependent afferent, input and an AF with fixed width $\sigma_\alpha = \sigma_\beta = 18$ deg and a range obtained from the nearest neighbor distribution summed over all stimuli: (a) Solid lines show the correlations between model predictions and human responses color-coded for different SOA (see legend in (b)). Dotted lines show the correlations between human responses of different subjects for the stimulus presentation times using the same color code. (b) Correlations between deviations from human responses from model predictions for different SOA (see legend).

distances. This could then account for the fact, that we find higher performances for larger element distances. Also for unidirectional models the probability for random contours in the background increases for small element distances. However, unidirectional models usually judge longer contours much more salient than short contours. Hence for $L = 10$ a short contour in the background should not change the decision of the model, and a contour of 10 elements in the background is very unlikely.

Despite these shortcomings with respect to the performance, this model reaches the correlations among human observers for each of the tested stimulus presentation times. The maximum correlations are similarly high as for the unidirectional model. However, in contrast to the unidirectional model the bidirectional model keeps a high level of correlations over a large time interval. The reason for the nearly constant performance and the high level of correlation over a large time is, that once the model detected a contour within the first few iterations it can go back and forth between the contour elements. The unidirectional model cannot do this and hence its neuronal activity for cells representing contour elements decreases when iterating further than needed to link all elements of a contour (compare 5.12).

As this bidirectional model does not reach a high enough performance, it might be worth trying to use the bidirectional association field with optimal angular dependence. As expected this increases the performance. However, for this model the correlation decreases and hence the correlation between different sub-

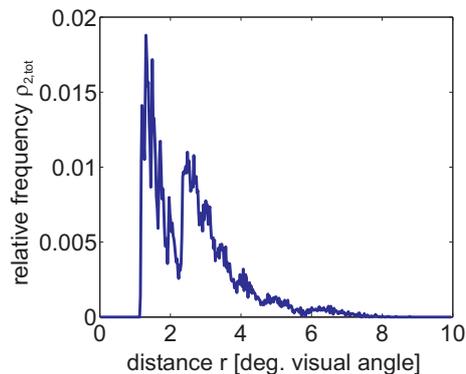


Figure 7.23: Distribution $\rho_{2,\text{tot}}$ of inter-element distances between nearest and next-to-nearest neighboring contour elements. Shown is the distribution of element distances for the complete set of stimuli used for the experiment.

jects is not reached (see table 7.1). So this bidirectional model cannot reproduce human behavior.

7.2.4 The range of the association field

So far we used the distance distribution ρ_{tot} over nearest neighboring contour elements to describe the range of the association field. This means that for the larger element distances (2.1, 2.7 and 3.5 degree visual angle) the association field connects only nearest neighboring edge elements. However, a problem of the bidirectional association field is that it finds contours by alternately linking two adjacent aligned edge elements. Possibly the bidirectional model improves with longer range of the horizontal connections. When assuming connections to the next-to-nearest neighboring elements a contour of length $L = 2$ and zero jitter should not become equally salient to a longer, but slightly jittered contour. At least $L = 3$ adjacent aligned elements are needed for the described alternately linking of few elements that leads to a high salience of short, straight and perfectly aligned contours (see first row of the table in Fig. 7.24). As randomly arising contours of $L = 3$ are a lot less likely than contours of $L = 2$ they should have less influence on the performance. Even more importantly, the salience of contours containing abrupt changes of direction should decrease compared to the salience of straight contours. For contours with only moderate changes in directions all nearest neighboring elements as well as all next-to-nearest neighboring elements are linked by connections of the according range (see second row of the table in Fig. 7.24). Now we "fold" this contour by flipping edge elements l and m to the opposite site of element k but leaving the orientations of each element unchanged. A model using bidirectional nearest neighbor interactions would compute exactly the same saliency for the folded contour as for the unfolded contour. However,

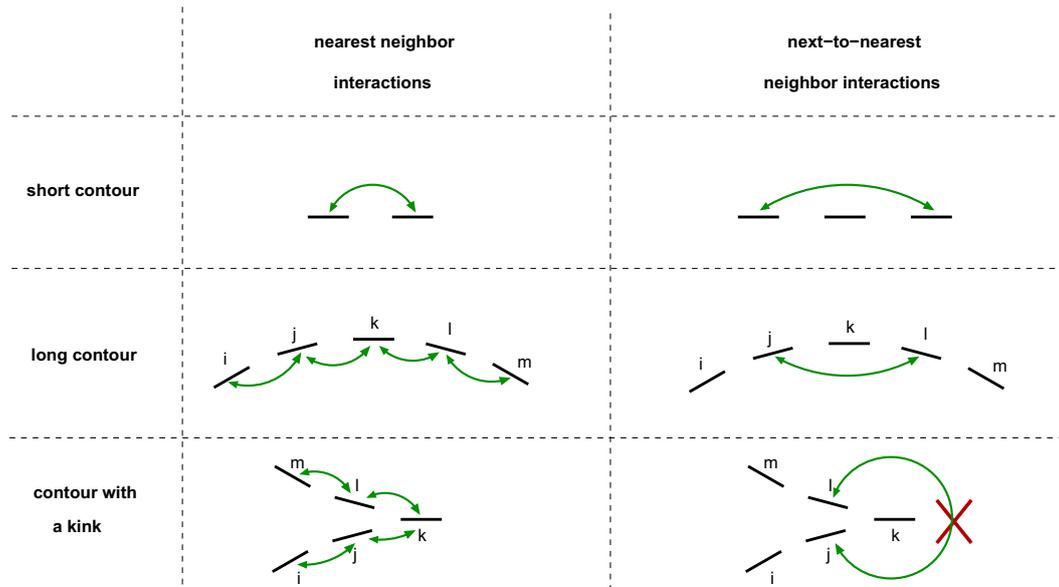


Figure 7.24: Here the effect of a longer range of bidirectional connections is shown for the examples of a short contour (first row), a longer contour without abrupt changes in direction (second row), and a longer contour with a kink (third row): While bidirectional interactions between nearest neighbors allow activity to be recursively exchanged between the neuronal representations of two adjacent collinearly aligned edge elements, at least three such elements are necessary for this effect when assuming interactions between next-to-nearest neighbors (first row). For longer contours without kinks depending on the range of the association field all next-to-nearest neighbors or all next-to-nearest neighbors are connected (second row, for lucidity of the figure next-to-nearest neighbor connections between edges i and k as well as between k and m are omitted). When using edge elements of the same orientations to generate a contour with a kink, by flipping over edge element l and m to the opposite side of element k , we obtain the same nearest neighbor connections as before. In contrast, next-to-nearest neighbor connections do not link edge detectors representing edge j and l any more because of the kink in the contour.

when analyzing next-to-nearest neighbor connections we realize that in contrast to the straight contour edge j and l are not linked any longer for the contour with kink. Hence the salience as computed using next-to-nearest neighbor interactions is lower for the contour with abrupt change in direction than for the straight contour. This suggests, that a longer range of the association field could enhance the salience of longer, approximately straight contours compared to short contours or contours with kinks, like observed in human perception.

In order to test whether a longer range of the horizontal connections indeed improves performance and correlations to humans, we extend the range of the association field such that each element is at least connected to its next-to-nearest neighbors. As we want the same association field range for each stimulus, we take the distance distribution $\rho_{2,\text{tot}}$ over next-to-nearest contour elements of all stimuli. However, this comes at the cost of having connections spanning 5 times the inter-element distance for the stimuli with the smallest element distance. Contrary to our hope that this model would show increased performance as it should detect fewer short contours in the background and fewer contours containing kinks, its averaged performance is decreased due to the longer association field range (Fig. 7.25 (a)). It shows especially low performance for small element distances (Fig. 7.25 (b+c)). As already mentioned for the smallest element distance an association field of range $\rho_{2,\text{tot}}$ connects elements with a distance of up to $5r_0$. This suggests that the connections extending beyond next-to-nearest neighbors make the performance worse. It seems highly unlikely that within human perception two collinearly aligned elements at such a distance would be grouped together if the intermediate edges are not nearly aligned to these two edges. Hence it is not surprising that for small element distances human performance is not reached by the model. However, when comparing only the performance for $L = 4$ and $L = 5$ for which connections do not reach beyond next-to-nearest neighbor elements, we see an increase of performance compared to the same model employing an association field with range ρ_{tot} . This suggests that next-to-nearest neighbor connections indeed facilitate contour detection for a bidirectional model.

Despite the bad average performance, the correlation between model predictions and human responses is equally good as for the model of smaller range. However, the ROC2-correlation is much higher here than in the model with the shorter range, which means that a model with an association field of shorter range explains human behavior far better. This finding together with the decreased performance for small r_0 indicates that the chosen range $\rho_{2,\text{tot}}$ of the association field is too large.

For comparison we also tested the long-range association field for unidirectional connections. Here the decrease in average performance was much more drastic than for the bidirectional model (Fig. 7.27 (a)). In contrast to the bidirectional model, here also the performance for small element distances decreases

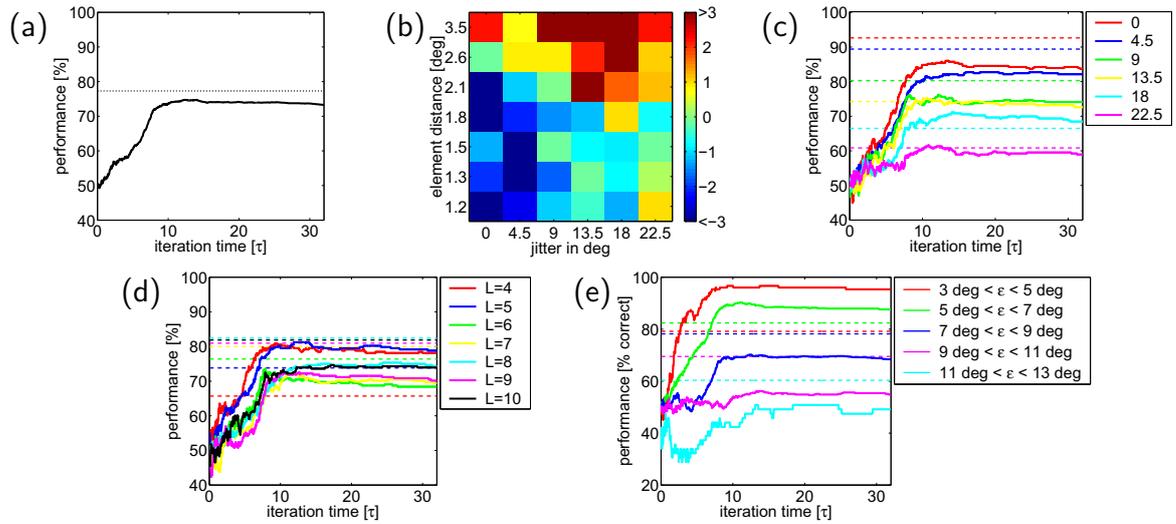


Figure 7.25: Performance of the model with bidirectional, multiplicative couplings and eccentricity-dependent afferent input. The AF has fixed width $\sigma_\alpha = \sigma_\beta = 18$ deg and range $\rho_{2,\text{tot}}$: (a) Performance averaged over all jitters and element distances. The solid line shows the model performance, the dotted line the human performance for 200 ms SOA. (b) ΔP for different combinations of r_0 and σ_j (c) Performance for different jitters averaged over all element distances and eccentricities. (d) Performance for different element distances averaged over jitters and eccentricities. (e) Performance in dependence of mean eccentricity of the elements of a contour averaged over jitters and element distances. For (c-e) solid lines denote model performance while dashed lines of the same color illustrate the corresponding human performance.

7.2. Results of model evaluation

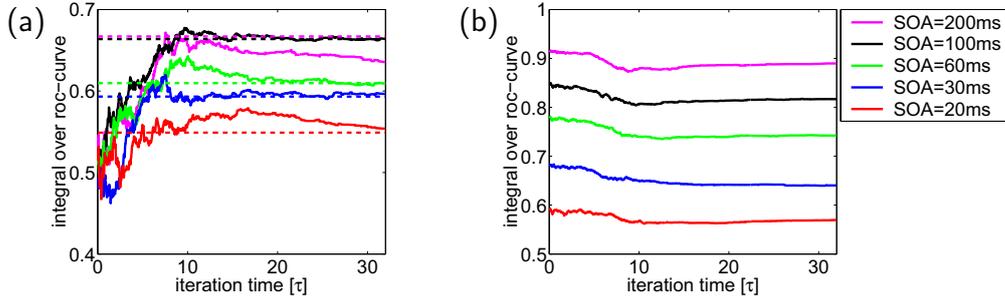


Figure 7.26: Correlations of the model with bidirectional, multiplicative coupling, eccentricity-dependent afferent input, and an AF with fixed width $\sigma_\alpha = \sigma_\beta = 18\text{deg}$ and a range obtained from the next to nearest neighbor distribution averaged over all stimuli: (a) Solid lines show the correlations between model predictions and human responses color coded for different SOA (see legend in (b)). Dotted lines show the correlations between human responses of different subjects for the same color coded stimulus presentation times. (b) Correlations between deviations from human responses from model predictions for different SOA (see legend).

(Fig. 7.27 (b+c)). Also performance for small jitters and performance for eccentric contours does not reach human performance (Fig. 7.27 (d+e)). When looking at the r_0 -dependence of the performance we see that again the highest performance is reached for the largest r_0 (Fig. 7.27 (c)). Here and in the bidirectional model with range $\rho_{2,\text{tot}}$ the difference between performances for $L = 4$ and $L = 10$ is larger than in the models using an association field with range ρ_{tot} .

Despite the low performance, correlations among human observers are reached by the ROC1-correlation for all SOAs. However, ROC2-correlations are clearly higher than in the corresponding model with an association field of range ρ_{tot} . Hence performance and ROC2-correlations give a strong hint, that the unidirectional connections assumed in this model extend too far.

Using different association field ranges for different stimuli seems not to be biologically plausible as horizontal connections do not change with the stimulus the subject is viewing. However, despite this biological implausibility we want to investigate the principal effects of connections between nearest neighbors. This would also be the range one would use for a probabilistic model. Therefore we perform simulations using an association field which extends exactly to the nearest neighboring units for each element distance. Average performance as well as performances for all combinations of jitter, element distance, and eccentricity exceed human performance (Fig. 7.29). And in contrast to the models that used the same association field for all stimuli, in this case the performance decreases with wider element distances like the human performance does. So it seems that

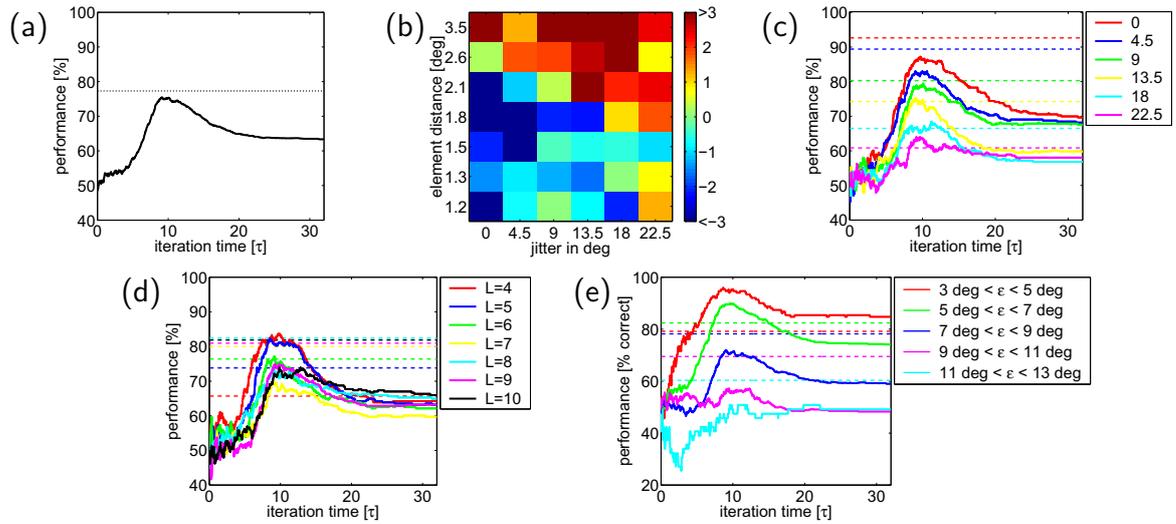


Figure 7.27: Performance of the model with unidirectional, multiplicative couplings, and eccentricity-dependent afferent input. The AF has fixed width $\sigma_\alpha = \sigma_\beta = 18$ deg and range $\rho_{2,\text{tot}}$: (a) Performance averaged over all jitters and element distances. The solid line shows the model performance, the dotted line the human performance for 200 ms SOA. (b) ΔP for different combinations of r_0 and σ_j (c) Performance for different jitters averaged over all element distances and eccentricities. (d) Performance for different element distances averaged over jitters and eccentricities. (e) Performance in dependence of mean eccentricity of the elements of a contour averaged over jitters and element distances. For (c-e) solid lines denote model performance while dashed lines of the same color illustrate the corresponding human performance.

7.2. Results of model evaluation

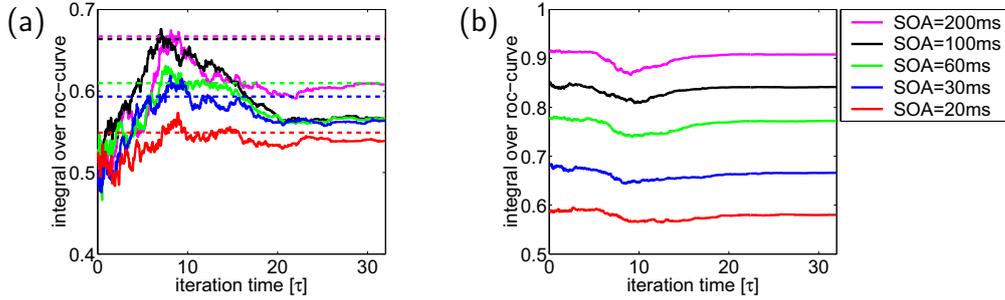


Figure 7.28: Correlations of the model with unidirectional, multiplicative coupling, eccentricity-dependent afferent input, and an AF with $\sigma_\alpha = \sigma_\beta = 18$ deg and a range obtained from the distance distribution between next-to-nearest neighbors averaged over all stimuli: (a) Solid lines show the correlations between model predictions and human responses color coded for different SOA (see legend in (b)). Dotted lines show the correlations between human responses of different subjects for the same color coded stimulus presentation times. (b) Correlations between deviations from human responses from model predictions for different SOA (see legend).

the reason for the unexpected increase of performance with element distance when using the same association for all element distances, was the fact that for smaller element distances the same range of association field links more elements together than for large element distances and interactions beyond next-to-nearest neighbors decrease the performance.

The correlation analysis shows that for all SOAs except the 200 ms condition the correlation between human responses is reached by the correlation between humans and models. And also the correlation for 200 ms SOA is within the $2\text{-}\sigma$ -range of the correlation among humans. From all models which reach the human-human correlation this is the one which has the lowest ROC2-correlation for the longer SOAs. Unfortunately, the minimum of the ROC2-curves is reached at a later iteration time when the ROC1-correlation already dropped again far below the correlation among humans. This means we still did not find a model that fulfills all three requirements, namely a high enough performance, a high enough ROC1 correlation, and a low ROC2-correlation, at the same time.

Given that performance for unidirectional association fields increases when going from an association field range of $\rho_{2,\text{tot}}$ over ρ_{tot} to ρ_L suggests that the ρ_{tot} distribution already contains connections which extend too far. Hence it would be interesting to use an association field with shorter horizontal connections. However, at the moment we assume a hypercolumn for each edge element in the stimulus, which means that for the longest element distance hypercolumns representing adjacent edge elements are not linked any more when using shorter

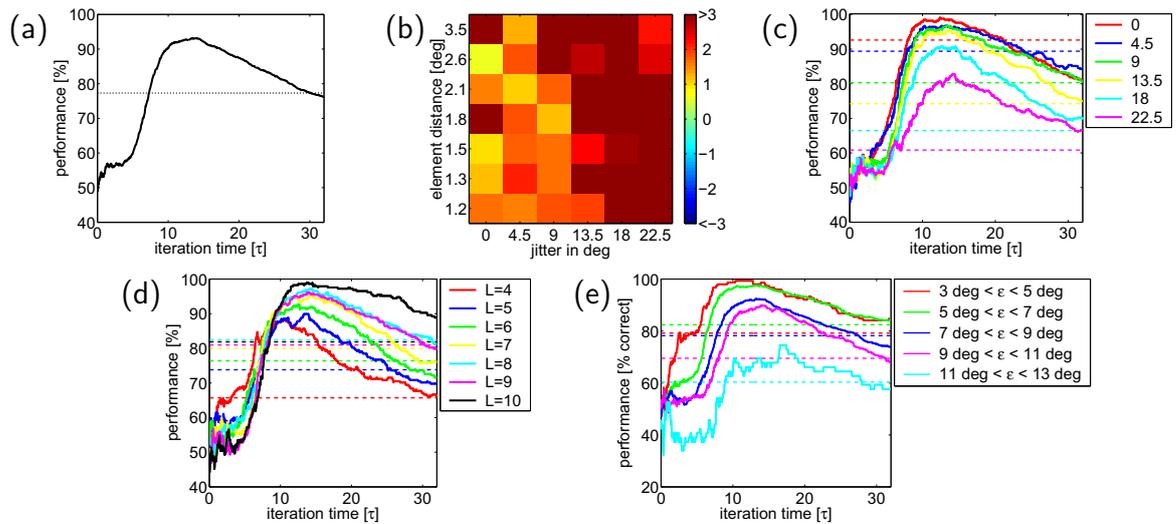


Figure 7.29: Performance of the model with unidirectional, multiplicative couplings and eccentricity-dependent afferent input. The AF has fixed width $\sigma_\alpha = \sigma_\beta = 18$ deg and range ρ_L to the nearest neighbor for each element distance: (a) Performance averaged over all jitters and element distances. The solid line shows the model performance, the dotted line the human performance for 200 ms SOA. (b) ΔP for different combinations of r_0 and σ_j (c) Performance for different jitters averaged over all element distances and eccentricities. (d) Performance for different element distances averaged over jitters and eccentricities. (e) Performance in dependence of mean eccentricity of the elements of a contour averaged over jitters and element distances. For (c-e) solid lines denote model performance while dashed lines of the same color illustrate the corresponding human performance.

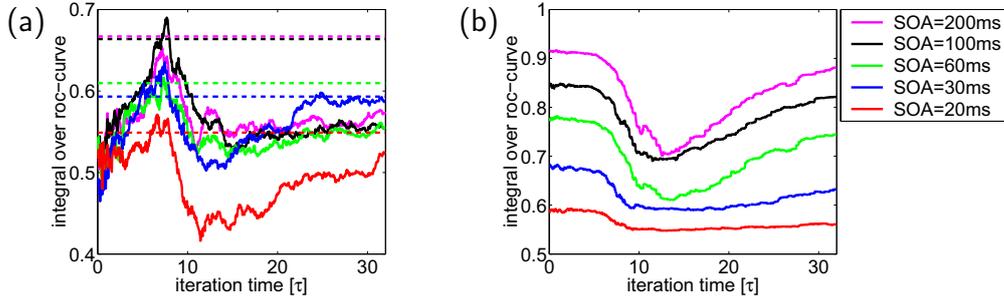


Figure 7.30: Correlations of the model with unidirectional, multiplicative coupling, eccentricity-dependent afferent input, and an AF with $\sigma_\alpha = \sigma_\beta = 18$ deg and nearest neighbor connections for each element distance: (a) Solid lines show the correlations between model predictions and human responses color coded for different SOA (see legend in (b)). Dotted lines show the correlations between human responses of different subjects for the same color coded stimulus presentation times. (b) Correlations between deviations from human responses from model predictions for different SOA (see legend).

range connections. Hence one had to insert additional hypercolumns between the already existing hypercolumns representing the edge elements of the stimulus. This would be a natural extension of our model as in visual cortex there are hypercolumns with receptive fields at each position in the visual field. This way hypercolumns representing neighboring edge elements in the used stimuli get not linked directly in one iteration step but with further iteration steps they get linked via the intermediate hypercolumns. This mechanism could also explain the filling in effect effect like for example observed for Kanizsa figures. However, this also means that the computational effort for the simulation will be significantly increased.

Additive model

As can be seen in table 7.1 (corresponding diagrams not shown) the additive model has generally a lower performance, lower ROC1-correlations, and higher ROC2-correlations than the corresponding multiplicative model. The only additive model which nearly reaches correlations among human observers uses an ϵ -modulated afferent input and a fixed, bidirectional association field of range ρ_{tot} . However, for all tested additive models with eccentricity-modulated I_{aff} and fixed association field geometry human performance is not reached within the 3σ -range for all jitters and element distances. Due to compromises in the determination of I_L/I_a it is possible that this improves to some extent for a more precise adjustment of afferent and lateral input. However, these results are still a strong hint that some nonlinear interaction is required for a contour integration

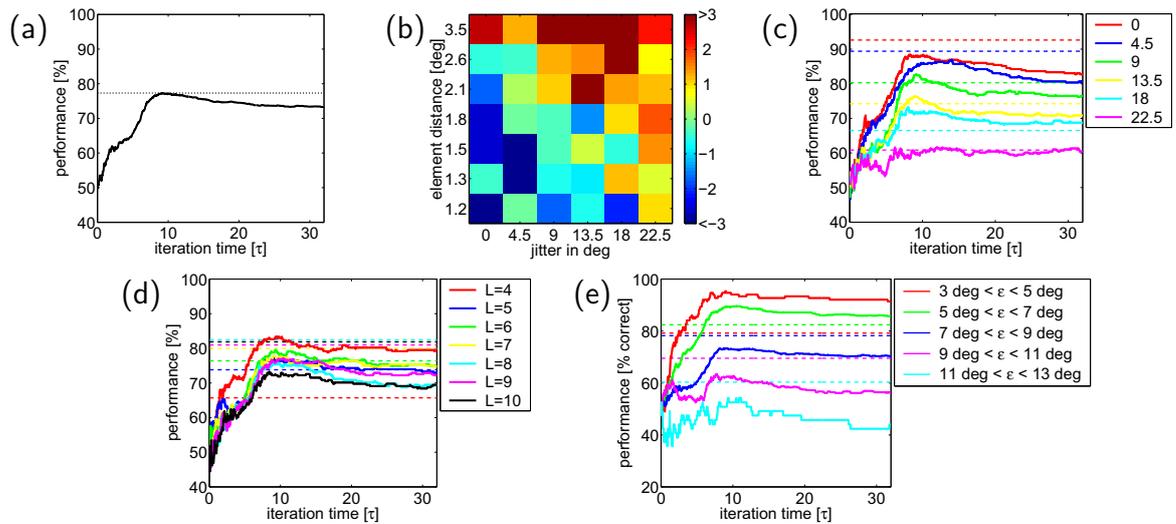


Figure 7.31: Performance of the model with bidirectional, multiplicative couplings and eccentricity-dependent afferent input. The AF has fixed width $\sigma_\alpha = \sigma_\beta = 18\text{deg}$ and range ρ_L to nearest neighboring elements for each individual element distance: (a) Performance averaged over all jitters and element distances. The solid line shows the model performance, the dotted line the human performance for 200 ms SOA. (b) ΔP for different combinations of r_0 and σ_j (c) Performance for different jitters averaged over all element distances and eccentricities. (d) Performance for different element distances averaged over jitters and eccentricities. (e) Performance in dependence of mean eccentricity of the elements of a contour averaged over jitters and element distances. For (c-e) solid lines denote model performance while dashed lines of the same color illustrate the corresponding human performance.

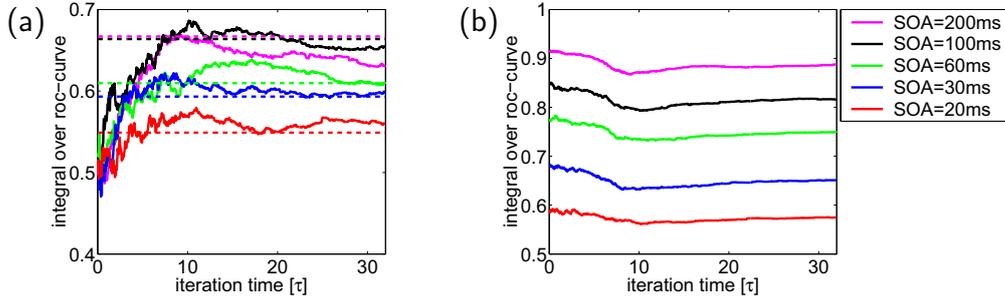


Figure 7.32: Correlations of the model with bidirectional, multiplicative coupling, eccentricity-dependent afferent input, and an AF with $\sigma_\alpha = \sigma_\beta = 18$ deg and couplings between nearest neighbors for each element distance: (a) Solid lines show the correlations between model predictions and human responses color coded for different SOA (see legend in (b)). Dotted lines show the correlations between human responses of different subjects for the same color coded stimulus presentation times. (b) Correlations between deviations from human responses from model predictions for different SOA (see legend).

model which is capable of reproducing human behavior. This does not need to be a purely multiplicative model, but can also be a combination of multiplicative and additive models.

Combining linear and nonlinear model dynamics

We saw in chapter 5 that the multiplicative model has the problem of fading activity for finite contours or contours containing gaps. This is no problem for the additive model. In contrast, the additive model is quite sensitive to star-illusions. Therefore it is a good idea to combine the two dynamics in order to make use of the advantages of both models and to overcome their shortcomings. Here we tested an ϵ -dependent model with combined dynamics, assuming $I = I_a I_{\text{aff}} + I_l I_{\text{lat}} + I_m I_{\text{aff}} \cdot I_{\text{lat}}$, using a unidirectional association field with $\sigma_\alpha = \sigma_\beta = 18$ deg and distance distribution ρ_{tot} over all nearest neighbor distances between contour elements. We chose $I_m = I_l \cdot I_a$ as this corresponds to a balanced influence of linear and non-linear model dynamics and like for the purely additive model we adjusted the strength of afferent and lateral input by searching for the I_a/I_l that leads to the maximum performance by exploring the I_l/I_a dependence for $N = 4$ and $N = 8$ and interpolating for the remaining N . For the chosen relation of I_l and I_m performances and ROC1-correlation are slightly reduced compared to the purely multiplicative model. Due to the range of the association field we still find the wrong dependence for the performance on the element distance r_0 . The ROC2-correlation is higher than in the purely multiplicative model. Similarly

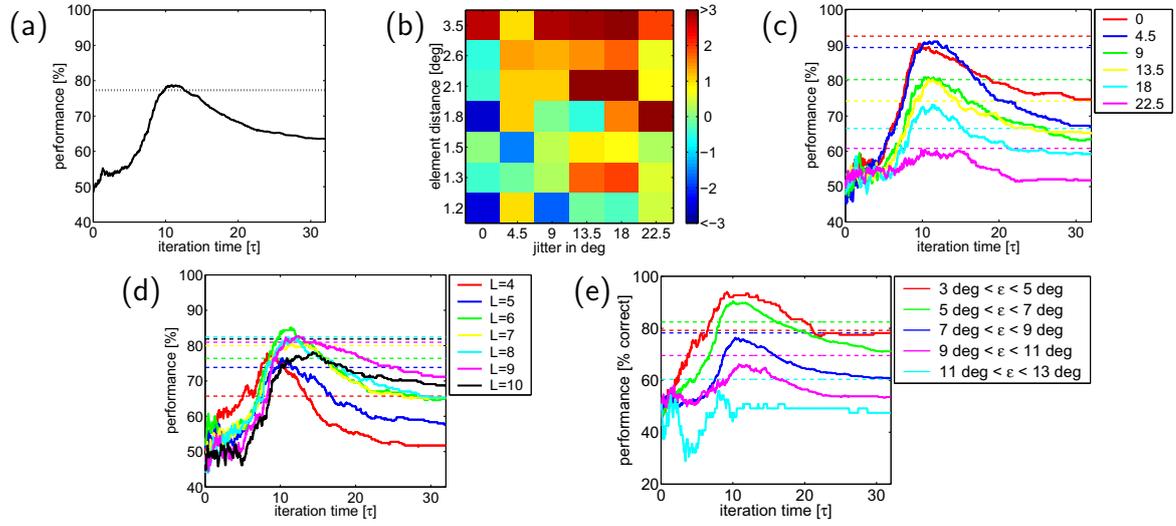


Figure 7.33: Performance of the model with $I = I_a I_{aff} + I_l I_{lat} + I_m I_{aff} \cdot I_{lat}$ and eccentricity-dependent afferent input. The AF is unidirectional and has fixed width $\sigma_\alpha = \sigma_\beta = 18\text{deg}$ and range ρ_{tot} : (a) Performance averaged over all jitters and element distances. The solid line shows the model performance, the dotted line the human performance for 200 ms SOA. (b) ΔP for different combinations of r_0 and σ_j (c) Performance for different jitters averaged over all element distances and eccentricities. (d) Performance for different element distances averaged over jitters and eccentricities. (e) Performance in dependence of mean eccentricity of the elements of a contour averaged over jitters and element distances. For (c-e) solid lines denote model performance while dashed lines of the same color illustrate the corresponding human performance.

the results for the bidirectional model are slightly worse than the results of the corresponding purely multiplicative model (see table 7.1). Obviously the model does not improve by simply averaging additive and multiplicative dynamics with $I_l \cdot I_a = I_m$. As the purely multiplicative model outperforms both, the additive model and a model with balanced influence of additive and multiplicative input coupling, it seems that the contribution of the multiplicative dynamics is more important than the contribution of the additive dynamics. However, in order to really answer the question, whether a mixing of the two dynamics can improve the model, one needs to explore further possible relations of the values I_l , I_a and I_m . Here we expect large I_m to be promising, as the multiplicative model so far leads to the best results and even a small input contribution of $I_a I_{aff} + I_l I_{lat}$ could be sufficient to fix the problem of fading activity for finite contours or contours with gaps.

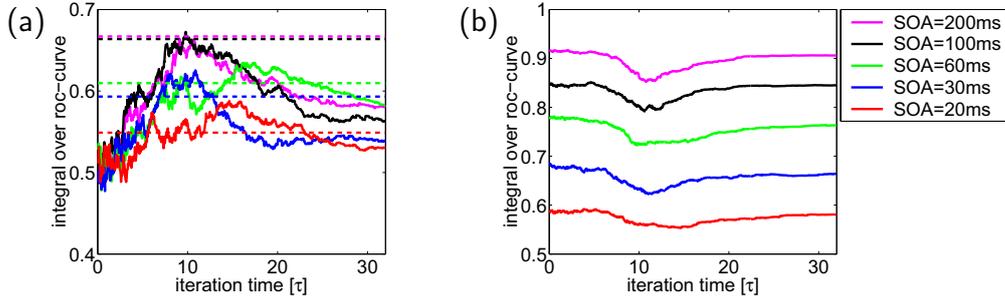


Figure 7.34: Correlations of the model with $I = I_a I_{\text{aff}} + I_l I_{\text{lat}} + I_m I_{\text{aff}} \cdot I_{\text{lat}}$, eccentricity-dependent afferent input, and a unidirectional AF with $\sigma_\alpha = \sigma_\beta = 18\text{deg}$ and range ρ_{tot} : (a) Solid lines show the correlations between model predictions and human responses color coded for different SOA (see legend in (b)). Dotted lines show the correlations between human responses of different subjects for the same color coded stimulus presentation times. (b) Correlations between deviations from human responses from model predictions for different SOA (see legend).

7.3 Discussion

In this chapter we realized, that there are many different influences which need to be taken into account in order to build a plausible contour integration model and there are several free parameters which need to be investigated. We explored different choices for model dynamics and input functions. For assessing the validity of a model we formulated several benchmarks based on human behavior. We saw that performance already can rule out some model choices, but especially if we do not know how much the noiseless model performance can be reduced due to a realistic amount of noise, performance alone is not sufficient to decide which of the analyzed models describes human behavior best. For example, models with ideal angular dependence as well as models with a fixed association field geometry of $\sigma_\alpha = \sigma_\beta$ reach human performance. But when looking at the correlations between model prediction and the responses of human observers, the model with fixed association field clearly explains human behavior better than a model employing multiple optimal association fields. Hence correlation analysis turns out to be a powerful tool for evaluating possible model choices. We saw that models using a multiplicative coupling between afferent and lateral input can explain human behavior better than those who sum all inputs to a neuron. Due to constraints of computer power we could only approximately determine the optimal balance between afferent and lateral input in the additive model and hence cannot say whether the model may improve further for better adjusted I_a/I_l . The most general choice of an input would be to combine the sum and the product of different inputs to a neuron. We tested one such model assuming balanced influ-

ences of additively and multiplicatively coupled inputs. This, however, did not improve correlations compared to the purely multiplicative model, but it remains to be tested whether a different balance of the additive and multiplicative term could improve the match between simulation and reality.

We showed that the correlations among human observers could only be reached when modulating the afferent input by eccentricity. However, a different way to include this dependence would be to scale the range of horizontal connections with eccentricity. Different cortical magnification factors depending on eccentricity have been found in experiments. We did not test yet, whether this improves the model further. However, the modulation of afferent input seems already sufficient to reproduce the correlation among human responses.

It would also be interesting to test to what extent the correlations among human observers are caused by this eccentricity dependence of contour integration performance. However, to test this we would need additional psychophysical data which has stimuli subsets for different contour eccentricity. With sufficiently many trials in each subset we could then distinguish between different eccentricity classes in the ROC analysis and eliminate the effect of the different performances for different eccentricities on the correlation analysis as we did for different element distances and different jitters. Furthermore it would be helpful to have more than 48 trials for each stimuli subset in order to get a better statistics for the ROC analysis which allows us to evaluate the correlations for each subset and not only the correlation for the complete set of data. For the correlation analysis it would also be sensible to use the same set of stimuli for each human observer in order to increase the number of pairs of subjects between which response correlations can be analyzed.

We saw that an association field with a fixed geometry for all stimuli leads to better results than using the generating association field for each stimulus. Also unidirectional association fields lead to better results as bidirectional ones because in models with bidirectional interactions the salience of short contours or contours with kinks is overestimated. This effect is stronger for nearest neighbor connections than for next-to-nearest neighbor connections. However, when assuming that the range of the association field does not change for different stimuli, we cannot adapt the association field such that it connects nearest and next-to-nearest neighboring elements regardless of r_0 . However, as especially for unidirectional models performance improves when going from an association field range of $\rho_{2,tot}$ over ρ_{tot} to ρ_L , we expect that the connections of range ρ_{tot} are already too long. In our study the range was an ad hoc assumption based on the fact that a probabilistic model performs best when taking the same range for the association field used to detect contours as for the generating association field. It needs to be explored with which range of the association field human contour detection can be explained best. As already mentioned in 7.1.1 the element

distances used in these stimuli are already larger than the long-range horizontal interactions found in V1. This is an additional reason to test shorter horizontal interaction than we did in this study. However, recent work suggests that higher cortical areas as V1 are involved in contour integration (76). In those areas long-range horizontal connections are assumed to link edge detectors for spatially further separated edge elements, and our model could equally well be implemented in those areas.

It is in the nature of exploratory studies with numerically demanding simulations that not for all parameters the optimal choice can be determined. However, we did introduce an effective method for evaluating models and ruled out several choices of input functions.

model dynamics	afferent input	AF geometry	AF symmetry	$F(r)$	$\max_t(\bar{P}_{\text{model}})$	$\#(\text{dP}(\sigma_j, r_0) < -3)$	$\max_t(\text{ROC1})$	$\min_t(\text{ROC2})$	
linear dynamics	$I_{\text{aff}}(\phi)$	optimal	uni	ρ_{tot}	80,7	0	0,60	0,77	
				$\rho_{2,\text{tot}}$	82,4	0	0,61	0,74	
			bi	ρ_{tot}	86,5	0	0,62	0,72	
				$\rho_{2,\text{tot}}$	90,9	0	0,58	0,69	
			fixed	uni	ρ_{tot}	78,4	0	0,63	0,82
					$\rho_{2,\text{tot}}$	82,5	0	0,59	0,78
		bi	ρ_{tot}	85,1	0	0,62	0,74		
			$\rho_{2,\text{tot}}$	90,1	0	0,59	0,69		
		$I_{\text{aff}}(\phi, \epsilon)$	optimal	uni	ρ_{tot}	74,6	2	0,60	0,82
					$\rho_{2,\text{tot}}$	71,2	1	0,58	0,82
					ρ_L	84,6	0	0,54	0,74
				bi	ρ_{tot}	79,6	0	0,63	0,79
	$\rho_{2,\text{tot}}$				80,1	0	0,61	0,79	
	ρ_L				79,6	0	0,60	0,81	
	fixed		uni	ρ_{tot}	68,6	9	0,62	0,89	
				$\rho_{2,\text{tot}}$	64,0	12	0,61	0,91	
				ρ_L	76,6	3	0,57	0,88	
			bi	ρ_{tot}	73,6	4	0,66	0,89	
				$\rho_{2,\text{tot}}$	73,6	5	0,65	0,89	
				ρ_L	73,5	7	0,63	0,88	
	combined dynamics	$I_{\text{aff}}(\phi, \epsilon)$	fixed	uni	ρ_{tot}	78,8	0	0,66	0,85
				bi	ρ_{tot}	75,1	3	0,65	0,89

model dynamics	afferent input	AF geometry	AF symmetry	$F(\mathbf{r})$	$\max_t(\bar{P}_{\text{model}})$	$\#(\text{dP}(\sigma_j, \tau_0) < -3)$	$\max_t(\text{ROC1})$	$\min_t(\text{ROC2})$	
nonlinear dynamics	$I_{\text{aff}}(\phi)$	optimal	uni	ρ_{tot}	94,2	0	0,59	0,66	
				$\rho_{2,\text{tot}}$	91,1	0	0,60	0,69	
			bi	ρ_{tot}	87,4	0	0,62	0,71	
				$\rho_{2,\text{tot}}$	91,1	0	0,61	0,69	
			fixed	uni	ρ_{tot}	94,3	0	0,61	0,68
					$\rho_{2,\text{tot}}$	91,5	0	0,61	0,69
		bi		ρ_{tot}	87,2	0	0,62	0,72	
				$\rho_{2,\text{tot}}$	91,4	0	0,63	0,68	
		$I_{\text{aff}}(\phi, \epsilon)$	optimal	uni	ρ_{tot}	88,2	0	0,63	0,72
					$\rho_{2,\text{tot}}$	82,8	0	0,63	0,74
					ρ_L	93,8	0	0,62	0,67
				bi	ρ_{tot}	81,7	0	0,62	0,76
	$\rho_{2,\text{tot}}$				81,9	0	0,62	0,76	
	ρ_L				82,2	0	0,63	0,77	
	fixed		uni	ρ_{tot}	84,8	0	0,68	0,81	
				$\rho_{2,\text{tot}}$	75,5	5	0,67	0,87	
				ρ_L	93,2	0	0,65	0,70	
			bi	ρ_{tot}	76,4	3	0,67	0,87	
				$\rho_{2,\text{tot}}$	74,8	5	0,67	0,87	
				ρ_L	77,3	3	0,67	0,87	

Table 7.1: This table gives an overview over different model dynamics, association field ranges, symmetries and geometries and afferent input properties. The most important model results are summarized here. \bar{P}_{model} denotes the average model performance. For ΔP and the ROC-correlations given in this table we used the human data for SOA=200 ms.

Chapter 8

Conclusion

The aim of this study was to learn more about contour integration in the brain, an important process in human and animal visual perception. For this purpose human psychophysical contour detection experiments were combined with extensive numerical simulations.

Many contour detection experiments use a similar paradigm as the psychophysical study which underlies this thesis. For the stimuli employed in this thesis, the distance distributions between contour elements, between background elements and between contour and background elements were carefully controlled to be identical in a good approximation. In contrast to this, several other studies (i.e. (48; 69; 27; 68)) used stimuli, where this was not the case and prior knowledge about the distance distribution of the contour could facilitate the detection of the contour (9). Another distinguishing feature of the stimuli employed in this thesis is that they contain contours generated from association fields. To my knowledge, no other study created contours in this manner. This mathematical precise contour generation makes it possible to formulate a Bayesian algorithm which from a mathematical point of view is the best contour integration model for this kind of stimuli and the given task. In recent years there were several publications demonstrating an optimal integration of sensory information (i.e. (65; 23)), and studies using probabilistically motivated models for contour integration (136). Stochastically generated stimuli have the advantage that we can derive a probabilistic model which in a way inverts the generation process and leads to the best possible contour detection performance for these stimuli and the given task. This gives us the possibility to see how much the human performance differs from the performance of an ideal observer model.

As the brain is assumed to be optimized to tasks and stimuli important in everyday life and not to artificial stimuli used in psychophysical experiments, it is not surprising that the Bayesian model leads to a high contour detection performance, but not to model predictions which were closest to human decisions.

However, as this ideal model bears the possibility to include prior information about the geometry of contours, it might be interesting to use this model for image processing applications where a high performance but not necessarily a resemblance to human perception is desired. In case a visual scene contains several contours with different geometrical properties, but only contours of a certain geometry are of interest, the prior knowledge about the geometry of the relevant contours can help to find them opposed to other less interesting contours.

Even though human behavior cannot be explained by an ideal observer model, there are some similarities between human performance and the ideal observer performance. Human performance is very high and its increase with stimulus presentation resembles the performance increase of the ideal observer model with longer iteration times, hinting to an iterative process also in the brain. Thus it seems possible that the brain employs similar strategies to those suggested by a probabilistic model. For this reason we first investigated the characteristic properties of probabilistic and of additive neural network models and the general differences between the model predictions. Therefore we analyzed basic properties of the models for very simple stimuli. This revealed that the additive model finds a very high salience for star-like edge configurations. Especially the additive model with summation estimator is very sensitive to these star-illusions, which motivates to perform psychophysical experiments with stimuli containing star-like edge configurations as well as contours in order to test to what extent human observers are sensitive to these star-illusions. However, our simulations with more complex stimuli showed, that the additive model with summation estimator leads to a comparatively low performance and also the correlations between model and human behavior are lowest for this model estimator combination. We did not include star-like edge configurations on purpose into the stimuli. However, edge configurations which could lead to stronger or weaker star-illusions can arise in the background just by chance. Hence there are almost certainly some stimuli containing star-like edge configurations and it is very well possible that these cause the bad performance of the summation estimator for the additive model. It remains to be tested, how many of these stimuli really contain star-like edge configurations and how much the average performance predicted for those stimuli by the additive model with summation estimator is decreased compared to the performance for the rest of the stimuli.

The multiplicative model is not sensitive to this kind of star illusions. However, in contrast to the additive model the multiplicative model has the problem that for finite contours the neuronal activity of the cells representing contour elements first increases but decreases after a certain iteration time. Depending on the association field and the elements surrounding the contour neuronal activity can even decrease to zero. This can be prevented by introducing spontaneous activity for the lateral input. It has a similar effect to replace the purely multiplicative coupling between afferent and lateral input by an input

$I = I_a I_{\text{aff}} + I_l I_{\text{lat}} + I_m I_{\text{aff}} \cdot I_{\text{lat}}$ that consists of a linear and a non-linear part. The multiplicative model further has the disadvantage that a global divisive normalization of the neuronal activity is needed in order to prevent that the activity drops to zero or rises beyond all bounds. For the additive model one can find a stable regime also without normalization. Another difficulty for the additive model is the need of a subtle balance of afferent and lateral input. This complicates simulations for complex stimuli as it is an enormous computational effort to determine this balance and due to constraints in computer power we could only approximately adjust this parameter.

In order to obtain a model which closer resembles human contour detection, we included several biological constraints in our model. This leads to a variety of model choices, starting with the probabilistic model, which uses unidirectional, multiplicative interactions with differing association field geometry and range for different contour ensembles, and no modulation of any input with eccentricity, and ending with a model relying on bidirectional additively interactions, an association field of fixed range and geometry for all stimuli, and an eccentricity-modulated afferent input.

In order to evaluate all these model choices, we did extensive numerical simulations on the same stimuli used in psychophysics. The model predictions were then compared to human decisions. For this quantitative comparison of model and human behavior we simulated for each model more than 4000 stimuli with widely varying curvature and alignment jitter, element distance, and mean eccentricity of contour elements. We showed that performance can rule out some models, but also realized that many models reach human performance and hence a comparison of performance is not sufficient to decide which of the analyzed models explains human behavior best.

A correlation analysis between the responses of different human observers showed, that human errors are not made randomly but are highly correlated among different subjects. This result can be used to better evaluate contour integration models, as a model capable of explaining human contour integration should reproduce not only human performance but also these systematic errors made by human observers. Hence the correlations between model predictions and human responses should reach the correlations among different subjects. Finally we found with the ROC2-correlation analysis a method to analyze to what extent our model reproduces the generic behavior of humans and how far we are away from the contour integration mechanism underlying human behavior. We showed, that with such a quantitative comparison of model and experiment, it is possible to draw predictions about connection structures and coupling dynamics from purely behavioral data. This kind of correlation analysis could be a powerful tool for the comparison of numerical simulations and behavioral data wherever complex stimuli are used which contain random components. The predictions

resulting from this study will be presented in the following paragraphs.

The quantitative comparison of model predictions and human responses showed that an eccentricity modulation of the input is necessary in order to explain the decrease of performance with the eccentricity of the contour. We only modulated the afferent input, which was sufficient to obtain a model capable of reproducing performance as well as ROC1-correlations. However, in many animals we also observe a modulation of the range of horizontal connections with eccentricity. It remains to be tested, whether an implementation of this increasing range of long-range horizontal interactions in the periphery improves the model further.

In some situations the human brain works close to optimal (65; 23). If the brain was optimized for detecting our contours, it would make sense to use the generating association field for contour integration. However, our correlation analysis showed that a single association field geometry leads to a better reproduction of human behavior than using for each stimulus the association field employed for stimulus creation. This is not surprising as the brain is not necessarily optimized for artificial tasks used in psychophysical experiments but rather for natural images. Hence when trying to find the optimal association field it would be interesting to analyze the statistics of natural images. Here the more interesting edge co-occurrence is not the one between all edge elements within the stimulus but the one between edge elements belonging to the same object. In fact one needs humans to mark contours of objects in order to determine this edge co-occurrence. This procedure was performed by Geisler and colleagues (33) who termed the resulting statistics 'Bayesian edge co-occurrence'. With this data we could test the predictive power of the model when using an association field obtained from natural image statistics. Geisler and colleagues did this only for a very simplified contour integration model which binds all edge elements if their link probability exceeds a certain threshold. Hence it would be interesting to perform simulations with the multiplicative model analyzed in this thesis and an association field obtained from Geisler's data.

In our simulations it was not possible to reproduce human performance with bidirectional association fields because models with this connection structure overestimate the salience of very short contours or contours with kinks. Also the well established model of Li (77) which also uses bidirectional interactions cannot reproduce human performance for the stimuli set we used here (25). This need for unidirectional interactions seems to be a disagreement to the biological finding that long range horizontal interactions which are believed to be the neuronal substrate for the association field spread isotropically in all directions. However, this asymmetry does not necessarily need to be caused by the anatomical structure. It would also be sufficient if there was some symmetry breaking by a refractory period of a neuron starting an asymmetrical process like the

propagation of a wave of activity.

A contour per se has no direction and passing the edge elements of a contour in one direction should be equivalent to passing them in reverse order in the opposite direction. The Gestalt rule underlying contour integration is the rule of good continuation. Psychophysical studies show that while adults use the Gestalt rule of good continuation in order to perceive an object as unity, children younger than one year cannot use this rule for object perception (123). The dominating principle for perceiving an object as unity in 12 week old infants is the Gestalt rule of common fate realized in the form of common motion of different object parts (61). Also adults who have been deprived from visual experience due to blindness from birth and have gained vision after medical treatment, can recognize moving objects in a visual scene as a unity before they can perceive stationary objects as a unity (93). This suggest that the rule of good continuation is learned from the visual experience of moving objects, and motion is a directed process. Hence it seems plausible to learn a unidirectional association field from moving scenes.

An influence we did not include in our models so far are top-down interactions from higher cortical areas. The receptive field size in humans and monkeys increases with higher level of the processing hierarchy from V1 to V4 (138; 122). At any given eccentricity receptive fields in V2 have more than twice the size than receptive fields in V1 (32; 10). Hence a cell in V2 gets stronger excitation from a longer contour matching its receptive field than from a shorter contour. Accordingly top-down interactions will enhance longer contours more than shorter contours and improve the performance of models with bidirectional horizontal interactions. However, for a model which uses only feed-forward and recurrent connections it is rather difficult to reach human performance with a single bidirectional association field.

In physiological experiments one finds facilitation for neurons representing collinearly aligned elements and inhibition between receptive fields whose centers are displaced perpendicular to their preferred orientation (60; 114; 118). Due to this inhibition the salience of contours with kinks is reduced compared to contours without kinks. The association fields employed in this study have only excitatory connections as we originally interpreted those as link probabilities which cannot become negative. Due to the overall normalization of neuronal activity a zero-excitation has a similar effect as an inhibition. However, it would be possible that an inhibition for parallel displaced edge detectors improves the performance of models using bidirectional association fields, especially for models containing additive interactions.

We assumed that contour integration is an iterative process. However, we observed neither that the maximum correlation between model and human observers shifted to longer iteration times for longer SOA, nor that the minimum ROC2-correlation shifted to longer iteration times for longer SOA, as we would

expect for an iterative model. Maximum correlations between model and humans increased with longer SOA. Hence it might be interesting to see, what performance and correlations can be reached for SOAs larger than 200 ms. Furthermore it could be that we then observe clearer maxima and the expected shift to larger SOAs.

Another way of obtaining more information about the temporal dynamics of contour integration could be to use dynamic stimuli where positions and orientations of Gabors vary with time instead of the static stimuli used so far.

If larger correlations between human subjects occur for longer SOA this would lead to even higher demands on the model, as the correlations between model and human observers should at least reach correlations among humans. An important issue which could not be solved satisfactorily in this study is the question about the origin of the correlations among human observers. At least partly they arise due to the decreased performance for contours in the periphery. For curvature and alignment jitter as well as element spacing we defined different subsets of data and took into account the different performance for each subset in the ROC-analysis. However, we could not distinguish between different eccentricities as there was too little data in each subset if we had partitioned the data further into several eccentricity classes. In order to determine to what extent the correlations are caused by eccentricity and to what extent other influences are responsible for the correlations, it would be necessary to obtain more psychophysical data using stimuli subsets of different eccentricities of contour elements with sufficiently many trials for each eccentricity.

An influence which was neglected in the comparison of simulations and human psychophysics because of the large additional computational effort this would mean, was the effect of noise on performance and correlations between models and humans. Generally we would expect noise to reduce the performance of the model. This effect of noise is particularly relevant wherever the contour is detected on the basis of small differences between saliences. As the differences in salience for contour and background elements are generally smaller in the additive model than in the multiplicative model, we would expect a bigger impact of noise on the additive model. This could result in the exclusion of additional models, which do not reach human performance when subjected to noise.

Even though the nature of the systematic errors of humans could not be completely identified, this study provided a quantitative comparison between model predictions and human decisions which does not stay on the level of average performances but takes into account correlations between human responses and model predictions on a trial-by-trial basis as an additional benchmark. It analyzed basic properties of different model dynamics and coupling structures and evaluated many possible model choices employing stimuli of a large range of eccentricity, element distance and curvature and alignment jitter. This allowed

us to rule out several possible model choices and suggested that multiplicative unidirectional lateral interactions are involved in contour integration in the brain.

Chapter 8. Conclusion

Appendix A

Randomly emerging contours in a hexagonal grid

In chapter 6 we discussed that the probability of randomly emerging contours in the background increases with the symmetry of the background positions. Here we want to determine the probability for randomly arising contours in a hexagonal grid, like the stimuli used for the simulations in 6.1 analytically.

As stimuli are not perceived with periodic boundary conditions we use non-periodic boundary conditions here. Given a horizontal row of N_1 background elements and K possible orientations for each background element, what is the probability $P(L|N_1, K)$ to find a randomly arising contour consisting of L elements? Under these conditions a contour of length L arises if L adjacent elements of this row are horizontally oriented. The probability for an element to have a horizontal orientation is $p = 2/K$. Vice versa $q = 1 - p$ is the probability for an orientation other than horizontal. We distinguish three different cases: If $N_1 = L$ we simply get $P(L|N_1, K) = p^L$. For $L < N_1 \leq 2L$ we have $N_1 - L + 1$ possibilities to position a contour of L elements. An example with $L = 3$ and $N_1 = 6$ is shown in Fig. A.1. The first possible position for the contour is a contour starting with the first element. All elements following behind the contour may have arbitrary orientations. The probability for such a situation is p^L . The second possible position of the contour is a contour starting with the second element in the row. However, to avoid double counting of the case where the first $L + 1$ elements form a contour, we require that in this case the first elements must not be horizontal. All elements following the contour can again have arbitrary orientations. The probability for this is qp^L . Similarly we require for all following starting positions x of the contour that the element at position $x - 1$ has not a horizontal orientation. At the positions $1 \dots (x - 2)$ we allow arbitrary orientations. For $N_1 \leq 2L$ there are not sufficiently many elements in front of the contour, such that these could form a second contour of length L . Hence these cases also have

Appendix A. Randomly emerging contours in a hexagonal grid _____

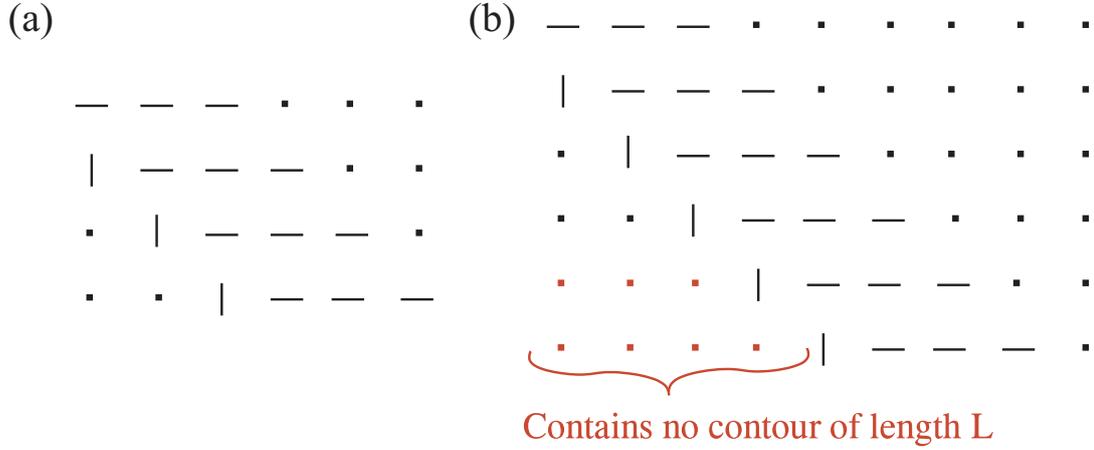


Figure A.1: Here horizontal bars stand for horizontal edge elements, vertical bars denote all orientations except horizontal and dots mean arbitrary orientations. We determine the probability for each possible contour position requiring some non-horizontal elements prior to the contour in order to avoid double counting of some edge configurations. (a) Example for $N_l \leq 2L$. (b) Example for $N_l > 2L$.

the probability qp^L . Summing over all these contour starting positions we get $P(L|N_l, K) = p^L(1 + (N_l - L)q)$.

The last possibility is that we have $N_l > 2L$ elements in a row. The first $L + 1$ starting positions for the contour have the same probability as in the case with $N_l \leq 2L$, as we just add elements of arbitrary orientation behind the contour. However, when the starting position of the contour is $x > L + 1$ and we have again a non-horizontal element at position $x - 1$ we still could have a contour of length L before the contour starts. This would lead to double counting of certain edge configurations. Hence for all elements at positions larger than $x - 1$ all combinations of orientations are allowed except for configurations which contain a contour of length L . It follows that the probability to find the first $x - 2$ elements in such an orientation configuration is $1 - P(L|x - 2, K)$. This term must again be multiplied by qp^L , and if we sum over all possible starting positions of the contour we get the recursive expression

$$P(L|N_l, K) = p^L + qp^L \cdot \left[N_l - L - \sum_{Z=L}^{N_l-L-1} P(L|Z, K) \right] \quad (\text{A.1})$$

Now we have the probability to obtain a randomly arising contour in one row. If we furthermore assume that the display has M such rows, the probability to

find at least one such contour is given by

$$P_{\text{rand}}(L|M, N_l, K) = \sum_{k=1}^M \binom{M}{k} P(L|N_l, K)^k (1 - P(L|N_l, K))^{M-k} \quad (\text{A.2})$$

according to the binomial distribution. Similarly this can be determined for diagonal contours. In that case we just have to take into account that the rows consist of different numbers of elements.

Appendix A. Randomly emerging contours in a hexagonal grid _____

Appendix B

Average link saliency between background elements

We want to determine the average link saliency $\langle s_{i,j} \rangle$ between two randomly oriented background elements i and j . In order to extract a contour from a stimulus, the link saliency between two contour elements as computed by the model should significantly exceed $\langle s_{i,j} \rangle$. In order to calculate $\langle s_{i,j} \rangle$ we assume a continuous distribution of possible orientations and use mean field theory.

The main idea of mean field theory is to replace all interactions, which in a specific realization of a stimulus target to one body, in this case edge i , by an averaged or effective interaction. In this way we can reduce a multi-body problem to an effective one-body problem.

We start with the link saliency between two elements i and j , which is given by

$$s_{i,j} = \int d\phi_i \int d\phi_j \mathbf{u}_{\theta_j}(\phi_j) P(\phi_j|\phi_i) \mathbf{u}_{\theta_i}(\phi_i)$$

where $\mathbf{u}_{\theta_j}(\phi_j)$ is the afferent input for a neuron that has the preferred orientation θ_j . Here θ_j is the orientation of the Gabor patch. In order to evaluate the average link saliency we have to include the probabilistic distribution over edge orientations $\rho(\theta)$ and to integrate over all stimulus orientations

$$\langle s_{i,j} \rangle = \int d\theta_i \int d\theta_j \int d\phi_i \int d\phi_j \rho(\theta_i) \rho(\theta_j) \mathbf{u}_{\theta_j}(\phi_j) P(\phi_j|\phi_i) \mathbf{u}_{\theta_i}(\phi_i)$$

Now we have to distinguish between contour elements and background elements. For background elements the orientations of the stimuli are equally distributed, thus $\rho(\theta) = \frac{1}{2\pi}$. With this we get

$$\langle s_{i,j} \rangle^{\text{back}} = \frac{1}{(2\pi)^2} \int d\phi_i \int d\phi_j \int d\theta_j \mathbf{u}_{\theta_j}(\phi_j) P(\phi_j|\phi_i) \int d\theta_i \mathbf{u}_{\theta_i}(\phi_i)$$

Appendix B. Average link saliency between background elements —

We can solve the θ -integrals by substituting $z = 2\theta - 2\phi$.

$$\begin{aligned} & \int_0^{2\pi} d\theta \frac{I_a}{2\pi I_0(\kappa)} \exp(\kappa \cos(2\theta - 2\phi)) \\ &= \int_{-2\phi}^{4\pi-2\phi} dz \frac{I_a}{2\pi I_0(\kappa)} \exp(\kappa \cos(z)) \end{aligned}$$

and as the Van Mises-function is periodic in 2π and normalized to 1,

$$= 2 \int_0^{2\pi} dz \frac{I_a}{2\pi I_0(\kappa)} \exp(\kappa \cos(z)) = 2I_a$$

So we find for the averaged background link saliency

$$\langle s_{i,j} \rangle^{\text{back}} = \frac{I_a^2}{\pi^2} \int d\phi_i \int d\phi_j P(\phi_j|\phi_i)$$

The association field is given by

$$\begin{aligned} P(\mathbf{r}, \alpha, \beta) &= \frac{F(\mathbf{r})}{2 \cdot (2\pi)^2 I_0(\kappa_\alpha) I_0(\kappa_\beta)} [\exp(\kappa_\alpha \cos(\beta/2 - \alpha)) \exp(\kappa_\beta \cos(\beta/2)) \\ &+ \exp(\kappa_\alpha \cos(\beta/2 - \alpha - \pi)) \exp(\kappa_\beta \cos(\beta/2 - \pi))] \end{aligned}$$

For simplicity we assume $\kappa_\alpha = \kappa_\beta = \kappa_{\text{aff}} = \kappa$. As we have only horizontal connections between the elements we have $\alpha = \phi_i$ and $\beta = \frac{\phi_i - \phi_j}{2}$. With this coordinate transformation we obtain the association field

$$\begin{aligned} P(\phi_j|\phi_i) = P(\mathbf{r}, \alpha, \beta) &= \frac{F(\mathbf{r})}{2 \cdot (2\pi)^2 I_0^2(\kappa)} \left[\exp\left(\kappa \left(\cos\left(\frac{\phi_i + \phi_j}{2}\right) + \cos\left(\frac{\phi_i - \phi_j}{2}\right)\right)\right) \right. \\ &+ \left. \exp\left(\kappa \left(\cos\left(\frac{\phi_i + \phi_j}{2} - \pi\right) + \cos\left(\frac{\phi_i - \phi_j}{2} - \pi\right)\right)\right) \right] \\ &= \frac{F(\mathbf{r})}{8\pi^2 I_0^2(\kappa)} \left[\exp\left(2\kappa \cos\left(\frac{\phi_i}{2}\right) \cos\left(\frac{\phi_j}{2}\right)\right) \right. \\ &+ \left. \exp\left(2\kappa \cos\left(\frac{\phi_j}{2} - \pi\right) \cos\left(\frac{\phi_i}{2}\right)\right) \right] \end{aligned}$$

Finally the average background link saliency reads

$$\begin{aligned} \langle s_{i,j} \rangle^{\text{back}} &= \frac{I_a \cdot F(\mathbf{r})}{8\pi^4 I_0^2(\kappa)} \left[\exp\left(2\kappa \cos\left(\frac{\phi_i}{2}\right) \cos\left(\frac{\phi_j}{2}\right)\right) \right. \\ &+ \left. \exp\left(2\kappa \cos\left(\frac{\phi_j}{2} - \pi\right) \cos\left(\frac{\phi_i}{2}\right)\right) \right] \end{aligned}$$

Substituting $\phi_1/2 = x$ and performing the integrations with

$$\int_0^\pi \exp(z \cos(x)) dx = 2\pi I_0(z)$$

(see i.e. (39), p.336) we obtain

$$\langle s_{i,j} \rangle^{\text{back}} = \frac{I_a \cdot F(r)}{8\pi^3 I_0^2(\kappa)} \int_0^{2\pi} [I_0(2\kappa \cos(\phi_j/2)) + I_0(2\kappa \cos(\phi_j/2 - \pi))] dx$$

By substituting $u = x - \pi$ in the second integral, we see that the two integrals are the same because $\cos(u) = \cos(-u)$.

$$\langle s_{i,j} \rangle^{\text{back}} = \frac{I_a \cdot F(r)}{4\pi^3 I_0^2(\kappa)} \int_0^{2\pi} I_0(2\kappa \cos(\phi_j/2)) dx$$

We can again substitute $\phi_j/2$. With

$$\int_0^{\pi/2} I_0(2\kappa \cos(x)) dx = \int_0^{\pi/2} I_0(2\kappa \sin(x)) dx = \pi/2 J_0^2(\kappa)$$

((39) p.657) and

$$I_0(z) = J_0(iz)$$

we find

$$\int_0^\pi I_0(2\kappa \cos(x)) dx = \pi I_0^2(\kappa)$$

as $\int_0^{\pi/2} I_0(2\kappa \sin(x)) dx = \int_0^{\pi/2} I_0(2\kappa \cos(x - \pi/2)) dx = \int_{\pi/2}^\pi I_0(2\kappa \cos(y)) dy$. This allows us to compute the average background link saliency as

$$\langle s_{i,j} \rangle^{\text{back}} = \frac{I_a \cdot F(r)}{4\pi^2}$$

Appendix B. Average link saliency between background elements —

Bibliography

- [1] http://gargoyle.arcadia.edu/psychology/blustein/neuro/lecture_notes/week_1/week_2/action_potential.jpg. 11/20/2006.
- [2] C. F. Altmann, H. H. Bühlhoff, and Z. Kourtzi. Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology*, 13:342–349, 2003.
- [3] C. F. Altmann, A. Deubelius, and Z. Kourtzi. Shape saliency modulates contextual processing in the human lateral occipital complex. *Journal of Cognitive Neuroscience*, 16:794–804, 2004.
- [4] Y. Amir, M. Harel, and R. Malach. Cortical hierarchy reflected in the organization of intrinsic connections in macaque monkey visual cortex. *Journal of Comparative Neurology*, 334(1).
- [5] A. Angelucci, J. B. Levitt, J. M. Hupé, E. J. S. Walton, J. Bullier, and J. S. Lund. Anatomical circuits for local and global integration of visual information: intrinsic and feedback connections in macaque visual cortical area V1 (Abstract). *European Journal of Neuroscience*, 12:285, 2000.
- [6] W. H. A. Beaudot and K. T. Mullen. Processing time of contour integration: the role of colour, contrast, and curvature. *Perception*, 30:833–853, 2001.
- [7] W. H. A. Beaudot and K. T. Mullen. How long range is contour integration in human colour vision? *Visual Neuroscience*, 20:51–64, 2003.
- [8] W. H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *Journal of Neuroscience*, 17(6):2112–2127, 1997.
- [9] J. Braun. On the detection of salient contours. *Spatial Vision*, 12:211–225, 1999.
- [10] A. Burkhalter and D. C. Van Essen. Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey. *Journal of Neuroscience*, 6:2327–2351, 1986.

Bibliography

- [11] M. Carandini and D. J. Heeger. Summation and division by neurons in primate visual cortex. *Science*, 264:1333–1335, 1994.
- [12] F. Chance, L. F. Abbott, and A. D. Reyes. Gain modulation from background synaptic input. *Neuron*, 35:773–782, 2002.
- [13] J. A. Connor and C. F. Stevens. Prediction of repetitive firing behaviour from voltage clamp data on an isolated neuron soma. *Journal of Physiology*, 213:31–53, 1971.
- [14] J.A. Connor, D. Walter, and R. McKown. Neural repetitive firing: modifications of the Hodgkin-Huxley axon suggested by experimental results from crustacean axons. *Biophysical Journal*, 18:81–102, 1977.
- [15] A. Cowey and E. T. Rolls. Human cortical magnification factor and its relation to visual acuity. *Experimental Brain Research*, 21(5):447–454, 1974.
- [16] P. M. Daniel and D. Whitteridge. The representation of the visual field on the cerebral cortex in monkeys. *Journal of Physiology*, 59:203–221, 1961.
- [17] A. Das and C. D. Gilbert. Distortions of visuotopic map match orientation singularities in primary visual cortex. *Nature*, 387:594–598, 1997.
- [18] A. Das and C. D. Gilbert. Topography of contextual modulations mediated by short-range interactions in primary visual cortex. *Nature*, 399:655–661, 1999.
- [19] P. Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modelling of Neural Systems*. The MIT Press, 2001.
- [20] R. O. Duncan and G. M. Boyton. Cortical magnification within human primary visual cortex correlates with acuity threshold. *Neuron*, 38(4):659–671, 2003.
- [21] R. G. P. Eckhorn, A. Gail, A. Bruns, A. Gabriel, and B. Al-Shaikhli. Neural mechanisms of visual associative processing. *Acta Neurobiologica Experimentalis*, 64:239–252, 2004.
- [22] J. H. Elder and S. W. Zucker. The effect of contour closure on the rapid discrimination of two-dimensional shapes. *Vision Research*, 33:981–991, 1993.
- [23] M. O. Ernst and M. S. Banks. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, 415(6870):429–433, 2002.
- [24] U. A. Ernst. *personal communication*.

- [25] U. A. Ernst. *personal communication*.
- [26] D. Ferster and K. D. Miller. Neural mechanisms of orientation selectivity in the visual cortex. *Annual Review of Neuroscience*, 23:441–471, 2000.
- [27] D. J. Field, A. Hayes, and R. F. Hess. Contour integration by the human visual system: evidence for a local association field. *Vision Research*, 33:173–193, 1993.
- [28] D. Fitzpatrick. The functional organization of local circuits in visual cortex: insights from the study of tree shrew striate cortex. *Cerebral Cortex*, 6:329–341, 1996.
- [29] D. Fitzpatrick. Seeing beyond the receptive field in primary visual cortex. *Current Opinion Neurobiology*, 10:438–443, 2000.
- [30] J. M. Foley, S. Varadharajan, C. C. Koh, and M. C. Q. Farias. Detection of gabor patterns of different sizes, shapes, phases and eccentricities. *Vision Research*, 47:85–107, 2007.
- [31] Y. Fregnac and V. Bringuier. Spatio-temporal dynamics of synaptic integration on cat visual cortical receptive fields. In A. Aertsen and V. Braitenberg, editors, *Brain Theory - Biological and Computational Principles*, pages 143–199. Elsevier, 1996.
- [32] R. Gattass, C. G. Gross, and J. H. Sandell. Visual topography of V2 in the macaque monkey. *Journal of Comparative Neurology*, 201:519–539, 1981.
- [33] W. S. Geisler, J. S. Perry, B. J. Super, and D. P. Gallogly. Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, 41:711–724, 2001.
- [34] C. D. Gilbert and T. N. Wiesel. Clustered intrinsic connections in cat visual cortex. *Journal of Neuroscience*, 3.
- [35] C. D. Gilbert and T. N. Wiesel. Morphology and intracortical connections of functionally characterised neurons in the cat visual cortex. *Nature*, 280:120–125, 1979.
- [36] C. D. Gilbert and T. N. Wiesel. Columnar specificity of intrinsic horizontal and cortico-cortical connections in cat visual cortex. *Journal of Neuroscience*, 9:2432–2442, 1989.
- [37] C. D. Gilbert and T. N. Wiesel. Receptive field dynamics in adult primary visual cortex. *Nature*, 356:150–152, 1992.

Bibliography

- [38] A. Gove, E. Mingolla, and S. Grossberg. Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience*, 12:1027–1052, 1995.
- [39] I. S. Gradsteyn and I. M. Ryzhik. *Table of Integrals, Series and Products*. Academic Press, San Diego, 1980.
- [40] S. Grossberg. Cortical dynamics of three-dimensional form, color, and brightness perception: I. monocular theory. *Perception and Psychophysics*, 41:87–116, 1987.
- [41] S. Grossberg and E. Mingolla. Neural dynamics of perceptual grouping: textures, boundaries, and emergent segmentations. *Perception and Psychophysics*, 38:141–171, 1985.
- [42] S. Grossberg, E. Mingolla, and W. D. Ross. Visual brain and visual perception: how does the cortex do perceptual grouping? *Trends in Neuroscience*, 20:106–111, 1997.
- [43] S. Hecht, S. Schlaer, and M. H. Pirenne. Energy, quanta and vision. *Journal of General Physiology*, 25:819–840, 1942.
- [44] R. F. Hess, W. H. A. Beaudot, and K. Mullen. Dynamics of contour integration. *Vision Research*, 41:1023–1037, 2001.
- [45] R. F. Hess and S. C. Dakin. Absence of contour linking in peripheral vision. *Nature*, 390:602–604, 1997.
- [46] R. F. Hess and S. C. Dakin. Contour integration in the peripheral field. *Vision Research*, 39:947–959, 1999.
- [47] R. F. Hess and D. Field. Integration of contours: new insights. *Trends of Cognitive Science*, 3:480–486, 1999.
- [48] R. F. Hess and D. J. Field. Contour integration across depth. *Vision Research*, 35:1699–1711, 1995.
- [49] R. F. Hess, T. Ledgeway, and S. Dakin. The detection of direction-defined and speed-defined spatial contours: one mechanism or two? *Vision Research*, 43:597–606, 2003.
- [50] A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117:500–544, 1952.
- [51] G. R. Holt and C. Koch. Shunting inhibition does not have a divisive effect on firing rates. *Neural Computation*, 9:1001–1013, 1997.

- [52] D. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, 160:106–154, 1962.
- [53] D. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *Journal of Physiology*, 195:215–243, 1968.
- [54] D. H. Hubel. *Eye, Brain and Vision*. New York: Scientific American Library, 1988.
- [55] J. M. Hupé, A. James, P. Girard, and J. Bullier. Response modulation by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. *Journal of Neurophysiology*, 85:146–163, 2001.
- [56] E. M. Izhikevich. Simple model of spiking neuron. *IEEE Transaction on Neural Networks*, 14:1569–1572, 2003.
- [57] J. P. Jones and L. A. Palmer. An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiol.*, 58(6):1233–1258, 1987.
- [58] G. Kanisza. Subjective contours. *Scientific American*, 234(4):48–52, 1976.
- [59] M. K. Kapadia, M. Ito, C. D. Gilbert, and G. Westheimer. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 alert monkeys. *Neuron*, 15:843–856, 1995.
- [60] M. K. Kapadia, G. Westheimer, and C. D. Gilbert. Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology*, 84:2048–2062, 2000.
- [61] P. J. Kellmann, E. Spelke, and K. Short. Infant perception of object unity from translatory motion in depth and vertical translation. *Child Development*, 58.
- [62] G. Kenan-Vaknin, G. E. Ouaknine, N. Razon, and R. Malach. Organization of layers II-III connections in human visual cortex revealed by *in vitro* injections of biocytin. *Brain Research*, 594(2):339–342, 1992.
- [63] J. J. Knierim and D. C. Van Essen. Neuronal responses to static texture patterns in area V1 of alert macaque monkey. *Journal of Neurophysiology*, 67:961–980, 1992.
- [64] K. Koffka. *Principles of gestalt psychology*. New York: Hartcourt, 1935.

Bibliography

- [65] K.P. Körding and D. M. Wolpert. Bayesian integration in sensorimotor learning. *Nature*, 427:244–247, 2004.
- [66] Z. Kourtzi, A. S. Tolias, C. F. Altmann, M. Augath, and N. Logothetis. Integration of local features into global shapes: monkey and human fMRI studies. *Neuron*, 37:333–346, 2003.
- [67] I. Kovacs. Gestalten of today: early processing of visual contours and surfaces. *Behavioural Brain Research*, 82:1–11, 1996.
- [68] I. Kovacs and B. Julesz. A closed curve is much more than an incomplete one: Effect of closure in figure-ground segregation. *Proceedings of the National Academy of Sciences USA*, 90:7495–7497, 1993.
- [69] P. Kozma and L. Kiorpes. Contour integration in amblyopic monkeys. *Visual Neuroscience*, 20:577–588, 2003.
- [70] S.-G. Kuai and C. Yu. Constant contour integration in peripheral vision for stimuli with good gestalt properties. *Journal of Vision*, 6(12):1412–1420, 2006.
- [71] I. Lampl, J. Anderson, D. C. Gillespie, and D. Ferster. Prediction of orientation selectivity from receptive field architecture in simple cells of cat visual cortex. *Neuron*, 30:263–274, 2001.
- [72] T. Ledgeway, R. F. Hess, and W. S. Geisler. Grouping local orientation and direction signals to extract spatial contours: Empirical tests of ‘association field’ models of contour integration. *Vision Research*, 45:2511–2522, 2005.
- [73] J. B. Levitt and J. S. Lund. Spatial summation properties of macaque striate neurons. *Society of Neuroscience Abstracts*, 23:455, 1997.
- [74] C. Y. Li and W. Li. Extensive integration field beyond the classical receptive field of cat’s striate cortical neurons-classification and tuning properties. *Vision Research*, 34:2337–2344, 1994.
- [75] W. Li and C. D. Gilbert. Global contour saliency and local colinear interactions. *Journal of Neurophysiology*, 88:2846–2856, 2002.
- [76] W. Li, V. Piech, and C. D. Gilbert. Learning to link contours in primary visual cortex. *Society of Neuroscience Abstracts*, 503:4, 2006.
- [77] Z. Li. A neural model of contour integration in the primary visual cortex. *Neural Computation*, 10:903–940, 1998.
- [78] Z. Li and J. J. Hopfield. Modeling the olfactory bulb and its neural oscillatory processing. *Biological Cybernetics*, 61.

- [79] P. G. Lovell. Evaluating accounts of human contour integration using psychophysical and computational methods, PhD-Thesis, Department of Psychology, University of Stirling, Scotland, 2002.
- [80] J. Lund, D. Fitzpatrick, and A. L. Humphrey. The striate visual cortex of the tree shrew. In E. G. Jones and A. Peters, editors, *Cerebral Cortex*, pages 157–205. New York: Plenum, 1985.
- [81] M. Maertens and S. Pollmann. fMRI reveals a common neural substrate of illusory contours and real contours in V1 after perceptual learning. *Journal of Cognitive Neuroscience*, 17:10:1553–1564, 2005.
- [82] R. Malach, Y. Amir, M. Harel, and A. Grinvald. Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. *Proceedings of the National Academy of Sciences USA*, 90:10469–10473, 1993.
- [83] S. Mandon and A. K. Kreiter. Rapid contour integration in macaque monkeys. *Vision Research*, 45:291–300, 2005.
- [84] B. Mathes. Neuronale Korrelate der Konturintegration, PhD-Thesis, Universität Bremen, Germany, 2006.
- [85] C. McAdams and J. Maunsell. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *Journal of Neuroscience*, 19:431–441, 1999.
- [86] G. Mitchison and F. Crick. Long axons within the striate cortex: Their distribution, orientation, and patterns of connection. *Proceedings of National Academy of Sciences USA*, 79:3661–3665, 1982.
- [87] B. N. Mora, G. J. Carman, and J. M. Allman. In vivo functional localization of the human visual cortex using positron emission tomography and magnetic resonance imaging. *Trends in Neuroscience*, 12:282–286, 1989.
- [88] D. Mumford. Elastica and computer vision. In C. L. Bajaj, editor, *Algebraic geometry and its applications*, pages 491–506. Springer-Verlag, New York, 1994.
- [89] M. E. Nelson. A mechanism for neuronal gain control by descending pathways. *Neural Computation*, 6:242–254, 1994.
- [90] H. Neumann and W. Sepp. Recurrent V1-V2 interactions in early visual boundary processing. *Biological Cybernetics*, 81:425–444, 1999.

Bibliography

- [91] H. C. Nothdurft, J. L. Gallant, and D. C. Van Essen. Response modulation by texture surround in primate area V1: correlates of 'popout' under anesthesia. *Visual Neuroscience*, 16:15–34, 1999.
- [92] A. K. Nugent, R. N. Keswani, R. L. Woods, and E. Peli. Contour integration in peripheral vision reduces gradually with eccentricity. *Vision Research*, 43:2427–2137, 2003.
- [93] J. Ostrovsky and P. Sinha. *personal communication*.
- [94] P. Parent and S. W. Zucker. Trace inference, curvature consistency, and curve detection. *IEEE transaction on pattern analysis and machine intelligence*, 2(8):823–839, 1989.
- [95] M. Persike and G. Meinhardt. Synergy of features enables detection of texture defined figures. *Spatial Vision*, 19:77–102, 2006.
- [96] M. W. Pettet. Shape and contour detection.
- [97] M. W. Pettet, S. P. McKee, and N. M. Grzywacz. Constraints on long range interactions mediating contour detection. *Vision Research*, 38:865–879, 1998.
- [98] U. Polat and D. Sagi. Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research*, 33:993–999, 1993.
- [99] U. Polat and D. Sagi. The architecture of perceptual spatial interaction. *Vision Research*, 34:73–78, 1994.
- [100] A. Pouget and T. J. Sejnowski. Spatial transformations in the parietal cortex using basis functions. *Journal of Cognitive Neuroscience*, 9:222–237, 1997.
- [101] A. Pouget and L. Snyder. Computational approaches to sensorimotor transformations. *Nature Neuroscience [Suppl]*, 3:1192–1198, 2000.
- [102] D. L. Ringach, R. M. Shapeley, and M. J. Hawken. Orientation selectivity in macaque V1: diversity and laminar dependence. *Journal of Neuroscience*, 22(13):5639–5651, 2002.
- [103] K. Rockland and J. S. Lund. Widespread periodic intrinsic connections in the tree shrew visual cortex. *Science*, 215(19):1532–1534, 1982.
- [104] K. S. Rockland and J. S. Lund. Intrinsic laminar lattice connections in primate visual-cortex. *Journal of Comparative Neurology*, 216:303–318, 1983.

- [105] K. S. Rockland, J. S. Lund, and A. L. Humphrey. Anatomical binding of intrinsic connections in striate cortex of tree shrews (*tupaia glis*). *Journal of Comparative Neurology*, 209:41–58, 1982.
- [106] W. D. Ross, S. Grossberg, and E. Mingolla. Visual cortical mechanisms of perceptual grouping: Interacting layers, networks, columns, and maps. *Neural Networks*, 13:571–188, 2000.
- [107] J. Rovamo and V. Virsu. An estimation and application of the human cortical magnification factor. *Experimental Brain Research*, 37:495–510, 1979.
- [108] P. A. Salin and J. Bullier. Corticocortical connections in the visual system: structure and function. *Physiological Review*, 75:107–154, 1995.
- [109] P. A. Salin, P. Girard, and J. Kennedy, H. Bullier. Visuotopic organization of corticocortical connections in the visual system of the cat. *Journal of Comparative Neurology*, 320:415–434, 1992.
- [110] E. Salinas and L. Abbott. Coordinate transformations in the visual system: how to generate gain fields and what to compute with them. *Progress in Brain Research*, 130:175–190, 2001.
- [111] E. Salinas and L. F. Abbott. Transfer of coded information from sensory to motor networks. *Journal of Neuroscience*, 15:6461–6474, 1995.
- [112] E. Salinas and L.F. Abbott. Invariant visual responses from attentional gain fields. *Journal of Neurophysiology*, 77:3267–3272, 1997.
- [113] M. Sambin. Angular margins without gradients. *Italian Journal of Psychology*, 1:355–361, 1974.
- [114] K. E. Schmidt, R. Goebel, S. Lowel, and W. Singer. The perceptual grouping criterion of collinearity is reflected by anisotropies of connections in the primary visual cortex. *European Journal of Neuroscience*, 9:1083–1089, 1997.
- [115] F. Sengpiel, A. Sen, and C. Blakemore. Characteristics of surround inhibition in cat area 17. *Experimental Brain Research*, 116:216–228, 1997.
- [116] M. I. Sereno, A. M. Dale, J. B. Reppas, and K. K. et al Kwong. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, 268:889–893, 1995.
- [117] S. Sherwood. A computational model for contour detection in the primary visual cortex, 2005. TITRA-NA-E05037, Master-Thesis, Royal Institute of Technology, Stockholm, Sweden.

Bibliography

- [118] H. Z. Shouval, D. H. Goldberg, J. P. Jones, M. Beckerman, and L.N. Cooper. Structures of long-range connections can provide a scaffold for orientation maps. *Journal of Neuroscience*, 20:1119–1128, 2000.
- [119] A. M. Sillito, K. L. Grieve, H. E. Jones, J. Cudeiro, and Davis J. Visual cortical mechanisms detecting focal orientation discontinuities. *Nature*, 378:492–496, 1995.
- [120] L. C. Sincich and G. G. Blasdel. Oriented axon projections in primary visual cortex of the monkey. *Journal of Neuroscience*, 21(12):4416–4126, 2001.
- [121] S. D. Slotnick, S. A. Klein, T. Carney, and E. E. Sutter. Electrophysiological estimate of human cortical magnification. *Clinical Neurophysiology*, 112:1349–1356, 2001.
- [122] A. T. Smith, K. D. Singh, A. L. Williams, and M. W. Greenlee. Estimating receptive field size from fMRI data in human striate and extrastriate cortex. *Cerebral Cortex*, 11(12):1182–1190, 2001.
- [123] E. Spelke, K. Breinlinger, K. Jacobson, and A. Phillips. Gestalt relations and object perception: A developmental study. *Perception*, 22:1482–1501, 1993.
- [124] L. Strother and M. Kubovy. On the surprising saliency of curvature in grouping by proximity. *Journal of Experimental Psychology: Human Perception and Performance*, 32(2):226–234, 2006.
- [125] S. Thorpe, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 381:520–522, 1996.
- [126] R. B. Tootell, M. S. Silverman, E. Switkes, and R. L. De Valois. Deoxyglucose analysis of retinotopic organization on primate striate cortex. *Science*, 218:902–904, 1982.
- [127] A. M. Treisman and G. A. Gelade. A feature-integration theory of attention. *Cognitive Psychology*, 12:97–136, 1980.
- [128] S. Treue and J. Martinez-Trujillo. Feature based attention influences motion processing gain in macaque visual cortex. *Nature*, 399:575–579, 1999.
- [129] D. Ts'o, C. D. Gilbert, and T. N. Wiesel. Relationships between horizontal connections and functional architecture in cat striate cortex as revealed by crosscorrelation analysis. *Journal of Neuroscience*, 6:1160–1170, 1986.

- [130] T. Tversky, W. S. Geisler, and J. S. Perry. Contour grouping: closure effects are explained by good continuation and proximity. *Vision Research*, 44:2769–2777, 2004.
- [131] M. A. Ursino and G. E. La Cara. A model of contextual interactions and contour detection in primary visual cortex. *Neural Networks*, 17:719–735, 2004.
- [132] R. Van Rullen, A. Delorme, and S. J. Thorpe. Feed-forward contour integration in primary visual cortex based on asynchronous spike propagation. *Neurocomputing*, 38–40:1003–1009, 2001.
- [133] M. Wertheimer. *Laws of organization in perceptual forms*. London: Hartcourt Brace Jovanovich, 1938.
- [134] L. R. Williams and D. W. Jacobs. Local parallel computation of stochastic completion fields. *Neural Computation*, 9(4):859–881, 1997.
- [135] L. R. Williams and D. W. Jacobs. Stochastic completion fields: A neural model of illusory contour shape and salience. *Neural Computation*, 9(4):837–858, 1997.
- [136] L. R. Williams and K.K. Thornber. Orientation, scale, and discontinuity as emergent properties of illusory contour shape. *Neural Computation*, 13:1683–1711, 2001.
- [137] S.-C. Yen and L. Finkel. Extraction of perceptually salient contours by striate cortical networks. *Vision Research*, 38:719–741, 1998.
- [138] S.M. Zeki. Uniformity and diversity of structure and function in the monkey prestriate visual cortex. *Journal of Physiology*, 277:273–290, 1978.
- [139] D. Zipser and R. A. Andersen. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature*, 331:679–684, 1988.
- [140] S. W. Zucker, A. Dobbins, and L. Iverson. Two stages of curve detection suggest two styles of visual computation. *Neural Computation*, 1:68–81, 1989.

