Robust integration and detection of noisy contours in a probabilistic neural model

Nadja Schinkel [a],* Klaus R. Pawelzik [a] Udo A. Ernst [a]

[a]*Institute for Theoretical Neurophysics, Otto-Hahn-Allee 1, University of Bremen, D-28334 Bremen, Germany.*

Abstract

Contour integration is an important step in the process of image segmentation and gestalt perception. Experimental evidence with monkeys and humans demonstrates that this specific computation is performed very fast and highly efficient, even if contours are jittered, partially occluded, or reduced in luminance. Here, we investigate the reliability of a probabilistic algorithm for contour detection under various neuronal and environmental constraints, as e.g. synaptic noise or imperfect knowledge about the exact orientation of an edge at some position in the visual field. We show that under most conditions there exists a range of tuning widths for orientation-specific neurons in the visual cortex which yields an optimum in contour detection performance. In particular, we demonstrate an increase of the performance when the information of the orientation of the contour elements becomes more uncertain.

Key words: gestalt perception, contour detection, noise, orientation preference

* E-Mail: nadja@neuro.uni-bremen.de

1 Introduction

One fundamental step during gestalt perception is the detection of the boundaries of an object. This process relies on the rule of 'good continuation' of line elements, linking those elements which are aligned colinearily (3; 8; 1). In psychophysical investigations it turns out that contour integration is both, remarkably fast, and robust against neuronal noise and jitter on the orientation of the contour elements (1; 2). These experimental results pose the question about the neuronal mechanisms underlying the contour integration process.

Theoretical investigations on possible mechanisms are based on two approaches: first, neural network models which employ intracortical horizontal connections (4; 11; 10) to mediate excitatory interactions between edge elements forming the contour can explain quite successfully many psychophysical phenomena (7; 5). Second, probabilistic models (9; 12) provide both, an exact mathematical definition of contour ensembles, and an optimal contour detection algorithm which evaluates edge link probabilities against the evidence for the presence of edges. One of the main differences between both approaches is that in typical network models, afferent input is *added* to the lateral feedback, while in the latter approach, evidence for oriented edges is bound *multiplicatively* to edge link probabilities.

We assume that the multiplicative nature of probabilistic algorithms can lead to higher detection performances than standard neural network approaches, thus being more robust against noise and uncertain information about the orientation of a stimulus at a specific location. Our brain has to cope with similar problems including synaptic noise and broad orientation tuning of typical neurons in visual cortex: even if a localized high-contrast stimulus

with a clearly defined orientation is presented, many neurons with a broad range of preferred orientations become activated (6).

In this paper, we quantify the robustness of a probabilistic model against noise and uncertain information using simulations and analytical estimates. We demonstrate that noise deteriorates detection performance as expected, but we also show that rendering the input tuning more unspecific can nevertheless lead to a profound increase in detection performance.

2   Theory

Probabilistic model. Based on concepts developed by (12), contour ensembles can be described by conditional probability densities $\rho(\vec{x}', \phi' | \vec{x}, \phi)$. ρ gives the probability that a contour passing through an edge in the visual field at position $\vec{x} = \{x_1, x_2\}$ with orientation ϕ will next pass through $\vec{x}' = \{x_1', x_2'\}$ with orientation ϕ'. ρ is also called association field (AF) and can be used both to randomly draw contour ensembles ('contour creation'), and to search for contours in a stimulus ('contour detection'). We now consider stimuli comprising N edges at positions $\vec{x}_i$ with orientations θ_i, and assume that observers do not know the exact orientation θ_i, but instead are only provided with evidence $\mu(\vec{x}_i, \phi_k)$ over a set of $k = 1, \ldots, K$ possible orientations $\phi_k = k(2\pi/K)$. μ is interpreted as the relative observation probability for an edge with orientation ϕ_k at position $\vec{x}$. For each position i and orientation k, we now compute the relative probability $\vec{s}^L = \{s_{ik}\}$ ('saliency') to belong to an open contour of exactly L edges. The combinatorial expression for $\vec{s}^L$ is derived from an iterative scheme originally developed for closed contours (12),

$$\vec{s}^L = \sum_{l=1}^{L} \left(\sqrt{\vec{u}^T} \mathbf{Q}^{L-l} \right) \left(\mathbf{Q}^{l-1} \sqrt{\vec{u}} \right) \quad \text{with} \quad Q_{i'k',ik}^l = \sqrt{u_{i'k'}^T} P_{i'k',ik} \sqrt{u_{ik}}$$

3

$$\text{using} \quad u_{ik} = \mu(x_i, \phi_k) \quad \text{and} \quad P_{i'k',ik} = \rho(\vec{x}_{i'}, \phi_{k'} | \vec{x}_i, \phi_k). \quad (1)$$

with the notation $\mathbf{Q}^l$ for the l-th power of the matrix $\mathbf{Q}$.

Simulation paradigm. In psychophysical experiments (1; 2), typical stimuli consist of N oriented Gabor elements or single bars. L elements are aligned to form a contour within $N - L$ randomly oriented edges as background. Subjects are asked to detect the contour. Their performance averaged over trials is recorded in dependence on various control parameters as e.g. element distance or orientation alignment jitter.

For our simulations, we choose the same approach but place the stimuli on a hexagonal grid with periodic boundary conditions in order to exclude boundary effects and hints to the location of contours from the inter-element distance statistics. Contours can be positioned either horizontal, left oblique or right oblique within the random background.

In order to establish a connection between the probabilistic algorithms and the structures in the brain, we identify certain parameters and functions with their possible physiological counterpart. The evidence $\mu(\vec{x}_i, \phi)$ for a stimulus at a specific visual field location i, will be termed the *afferent input*. In the next paragraph, we will choose μ to mimick the typical input tuning provided by the LGN to a hypercolumn in the primary visual cortex. Furthermore, we interpret the saliencies $\vec{s}$ as the *activation* of the cortical columns, and the AF ρ as the pattern of *horizontal intracortical connections* linking neurons with similar orientation preferences.

Choice of functions, distributions, and parameters. For defining a suitable association field and afferent input, we use von-Mises functions (i.e. Gaussians defined on a circular manifold) $M(x, \mu, \kappa) = \exp(\kappa \cos(x - \mu))/(2\pi I_0(\kappa))$

4

with concentration parameter κ and circular mean μ. I_0 denotes a modified Bessel function of the first kind of order 0. The afferent input distribution $u_{ik} = M(2\phi_k, 2\theta_i, 1/\sigma_{aff}^2)$ is centered around θ_i with a width of σ_{aff}. To make contour detection more difficult, a jitter of amplitude $\eta_\theta = 0, 1, 2, \ldots$ which randomly rotates every edge by exactly $\pm\eta_\theta(2\pi/K)$ can be applied to the contour. We assume the AF to be translational and rotational invariant with respect to the source edge,

$$\rho\left(r, \alpha, \beta | 0, 0, 0\right) = 0.5\, F(r) \sum_{\delta=\{0,\pi\}} M(\gamma_\alpha, \delta, 1/\sigma_\alpha^2) M(\gamma_\beta, \delta, 1/\sigma_\beta^2) \qquad (2)$$

where $r = |\vec{x}' - \vec{x}|$ is the distance between two contour elements, α the view angle under which the destination edge is seen from the source edge, and β the orientation difference between the two edges. For simplicity, the distance function is $F(r) = 1$ for the six nearest neighbors and 0 otherwise. γ_α and γ_β are defined as $\gamma_\alpha = \beta/2 - \alpha$ and $\gamma_\beta = \beta/2$. For our simulations we choose physiologically plausible values of $\sigma_\alpha = 2 \cdot \sigma_\beta = \pi/12$. As information processing in the brain is noisy, additive noise η_u can be applied on the afferent input $\vec{u}$. This noise is choosen uniformly distributed between 0 and $\eta_u \max\{\vec{u}\}$. Our simulations used $N = 324$ and $L = 9$ together with $K = 24$ or $K = 72$.

3 Simulations and Results

To determine whether the contour is detected or not, we take into account the $L_{max} = 1$, 3, or 5 elements with the highest saliencies. If more than half of these saliencies belong to contour elements, the contour is assumed to be detected. As we find comparable performances for all L_{max} only performances determined with $L_{max} = 5$ are shown, averaged over sets of $N_{stim} = 100$ stimuli.

Dependence of performance on afferent input width σ_{aff}. We investigated static and dynamic noise with $\eta_u = 0.001$ and $\eta_u = 0.05$ with no contour jitter ($\eta_\theta = 0$). In both cases the performance was better for dynamic noise than for static noise, because dynamic noise is averaged out over time (see Fig. 1 (left)). With broader afferent input, the performance decreases as
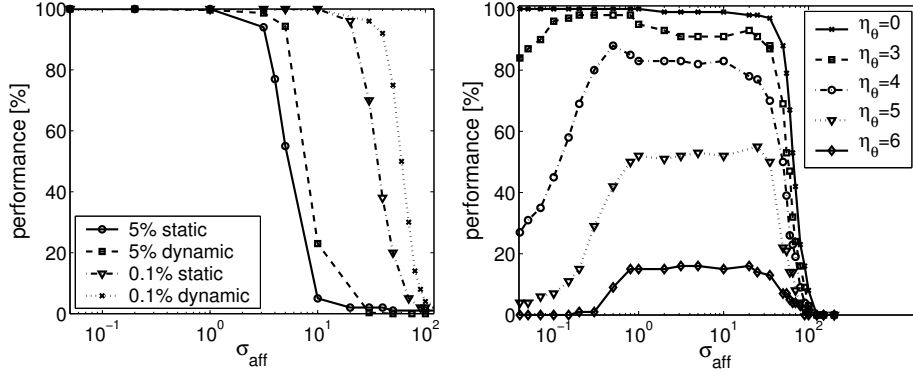

Fig. 1. Contour detection performance in percent correct in dependence on the afferent input width σ_{aff} for different noise levels and types (left), and for different orientation jitter strengths η_θ (right). Performance gets better for lower noise levels, and static noise is impeding the performance stronger than dynamic noise (left). On the right, performance curves become non-monotonous for higher values of η_θ showing that broader afferent input can lead to an *increase* in detection performance.

the information about the orientation of the edge becomes more and more uncertain. The performance starts to decrease at a critical width of $\sigma_{aff} \approx 10$ ($\sigma_{aff} \approx 4$) for noise levels of $\eta_u = 0.001$ ($\eta_u = 0.05$) which reach or exceed the amplitude of the afferent input tuning curve. Performance finally drops to chance level where a distinction between a noisy Von Mises-function and an equidistribution is no longer possible.

Dependence of performance on orientation jitter η_θ. Further simulations were performed varying the orientation jitter η_θ with a discretization of $K = 72$. Again, we see a decrease of performance for higher σ_{aff} due to the

influence of noise which is nearly independent of the orientation jitter. For jitters up to $\eta_\theta = 2$, performance starts at 100%, monotonously decreasing to chance level (See Fig. 1 (right)). For jitters from $\eta_\theta = 3$ the performances start below 100%, then increase significantly with σ_{aff} reaching a plateau at between $\sigma_{aff} = 0.1$ and $\sigma_{aff} = 1.0$, and finally decrease to chance level.
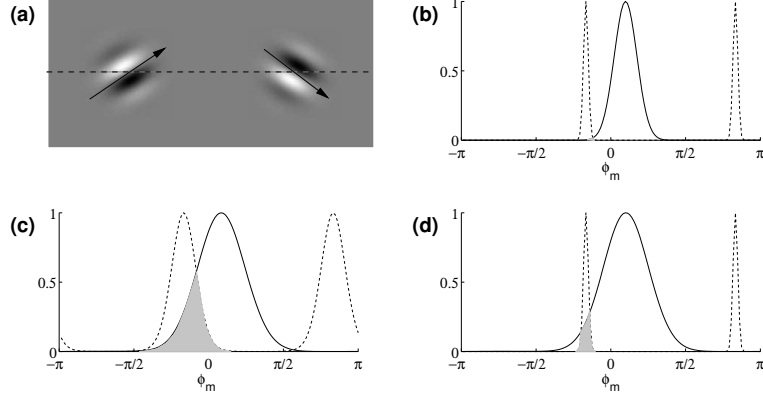

Fig. 2. (a) Stimulus situation considered in the text: two contour elements i and i' are shown with an orientation jitter of $-\eta_\theta$ and $+\eta_\theta$ (arrows), respectively. (b)-(d) show the afferent input distribution $u_{i'k'}$ at destination element i' (dashed line), together with the lateral input $v_{i'k'i}$ (solid line) in different situations, namely; (b) with both narrow σ_{aff} and narrow AF, (c) with broad σ_{aff} and narrow AF, and (d) with narrow σ_{aff} and broad AF. The shaded areas depict the overlap of the two curves symbolizing the relative link probability of i' and i.

At a first glance, this result is counterintuitive: a *more unspecific* information about the edge orientations leads to an *increase* in detection performance. To explain this phenomenon, we consider the worst-case situation in which two successive elements i and i' of a contour are jittered by $+\eta_\theta$ and $-\eta_\theta$, respectively (see Fig. 2(a)). In such a critical situation, the mean link probability between hypercolumns i and i' should exceed the average link probability of

two background elements, leading to the condition

$$o_{i'i} = \sum_{k'=1}^{K} u_{i'k'} v_{i'k'i} \gg 0 \,. \tag{3}$$

where $v_{i'k'i} = \sum_{k=1}^{K} P_{i'k',ik} u_{ik}$ denotes the 'lateral input' a column (i', k') gets from a neighbouring hypercolumn (i, k), $k = 1, \ldots, K$. The link saliency is large if the two vectors $\vec{u}_{i'}$ and $\vec{v}_{i'i}$ shown in Fig. 2 have a considerable overlap $o_{i'i}$. As this is the case with increasing σ_{aff}, detection performance increases. The intuitive explanation for this behaviour is that a broader afferent input can recruit horizontal interactions which were inactive before, in order to strengthen the effective link between the jittered contour elements. The same effect shows up when instead of σ_{aff}, the AF widths σ_{α} and σ_{β} are increased (cf. to Eq. 3). However, for a very wide AF, contour detection does not benefit from enlarging the σ's, because the probability for pseudo-contours in the background gets too large.

4   Discussion and Outlook

We showed that contour detection performance is nearly independent of the width of the afferent input, if the jitter on the contour alignment is small. In particular, the performance stays at 100% until the difference of maximum and minimum afferent input becomes smaller than η_u.

For larger jitter we get the surprising result that performance increases with the afferent input width σ_{aff} as the link saliency is improved due to the overlap of afferent and lateral input. This behaviour leads to a robust maximum in contour detection performance for σ_{aff} large enough to yield a considerable overlap, but not that large that the performance is significantly diminished by

8

orienation jitter η_θ and neuronal noise η_u. The range for this optimal performance depends on the length scales of the AF. Thus we find that the broad input tuning on orientation selective neurons in the visual cortex does not need to be a disadvantage, but rather can be beneficial for contour detection.

In collaboration with experimentalists, we are currently subjecting this framework to real, more irregular stimuli used in psychophysical experiments with monkeys and humans. First results can be found in (2). While traditional models of the cortex like the Wilson-Cowan model use an additive coupling of afferent and lateral input, our approach is based on multiplicative couplings. In the future, we will study the differences of these two approaches more closely. Furthermore, we will investigate how the combinatorial expression Eq. 1 can be mapped to a biologically realistic, iterative model achieving a comparable performance.

References

[1]   J. Braun, Spatial Vision 12 (1999) 211–225.

[2]   U.A. Ernst, S. Mandon, K.R. Pawelzik, and A.K. Kreiter, Neurocomputing (2004), in the press.

[3]   D.J. Field, A. Hayes A., R.F. Hess, Vision Res. 33 (1993) 173–193.

[4]   C.D. Gilbert, T.N. Wiesel, J. Neurosci. 9 (1989) 2432–2442.

[5]   M.H. Herzog, U.A. Ernst, A. Etzold, C.W. Eurich, Neur. Comp. 15 (2003) 2091–2113; U.A. Ernst, C.W. Eurich, in The Handbook of Brain Theory, ed. Arbib M.A., MIT Press (2002) 294–300.

[6]   D. Hubel and T. Wiesel, J. Physiology London 160 (1962) 106–154.

[7] Z. Li, Netw. Comput. Neural Syst. 10 (1999) 187–212; Neur. Comp. 13 (2001) 1749–1780.

[8] W. Li, C.D. Gilbert, J. Neurophysiol. 88 (2002) 2846–2856.

[9] D. Mumford, Elastica and Computer Vision, in: *Algebraic geometry and its applications*, New York, Springer-Verlag (1994).

[10] K.E. Schmidt, R. Goebel, S. Löwel, W. Singer, Eur J. Neurosci. 9 (1997) 1083–1089.

[11] D.D. Stettler, A. Das, J. Bennett, C.D. Gilbert, Neuron 36 (2002) 739–750.

[12] L.R. Williams, K.K. Thornber, Neur. Comp. 12 (2001) 1683–1711.