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, D-28334 University of 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 in a highly efficient manner, even if contours are jittered, partially occluded, or reduced in luminance. In this contribution, we investigate algorithms and strategies for a reliable detection of contours, which are subjected to various plausible neuronal and environmental constraints, as e.g. synaptic noise or imperfect knowledge about the exact orientation of an edge at some position in a stimulus display. It is shown 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, the performance can increase when the information of the orientation of the contour elements becomes more uncertain.

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

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 open the question about the neuronal mechanisms which can explain the observed phenomena.

Theoretical investigations of possible mechanisms are based on two approaches: first, neural network models which employ intracortical horizontal connections

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

(4; 11; 10) to mediate excitatory interactions between edge elements forming the contour can quite successfully explain many psychophysical phenomena (7; 5). Second, a probabilistic approach (9; 12) provides an exact mathematical description of contour ensembles and thus an optimal algorithm for the detection of well defined contours. Here, the integration process is done by evaluating edge link probabilities against 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.

In our research project, we take a closer look at the relationship between these two computational paradigms. Probabilistic algorithms can yield higher detection performances than standard neural networks and are more robust against noise and uncertain information about the orientation of a stimulus at a specific location. Our brain has to cope with both influences, which are due to synaptic noise and due to the orientation tuning properties of typical visual cortical neurons: even if a localized high-contrast stimulus with a clearly defined orientation is presented, also neurons with preferred orientations different from the stimulus' orientation become activated (6).

These observations led us to the question, how robust a probabilistic model of contour integration can be against noise and uncertain information? In the following paper, we will tackle this question using large-scale computer simulations and analytical estimations. We will demonstrate that noise deteriorates detection performance, and we will show that in contrast to common belief, making the tuning of the input 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 $\rho(\mathbf{x}', \phi'|\mathbf{x}, \phi)$. ρ gives the probability that a contour passing through an edge in the visual field at position $\mathbf{x} = (x, y)$ with orientation ϕ will next pass through $\mathbf{x}' = (x', y')$ with orientation ϕ'. ρ is also called association field (AF) and can be used

- to randomly draw contour ensembles ('contour creation'),
- and to search for contours in a stimulus ('contour detection').

In this contribution, we use stimuli consisting of N edges at positions $\mathbf{x}_i$ with orientations θ_i. We furthermore assume that an observer does not know the exact θ_i, but instead is provided only with evidence $\mu(\mathbf{x}_i, \phi_k)$ over a set of $k = 1, \ldots, K$ possible orientations $\phi_k = k(2\pi/K)$. μ can be interpreted as the relative observation probability for an edge with orientation ϕ_k at $\mathbf{x}$ in the current stimulus. For each position i and orientation k, we now compute the relative probability $\mathbf{s}_{ik}^L$ ('saliency') to belong to an open contour of exactly L edges. The combinatorial expression for $\mathbf{s}^L$ is derived from the iterative scheme of Williams and Thornber (12) originally developed for closed contours,

$$\mathbf{s}^L = \sum_{l=1}^{L} \left(\sqrt{\mathbf{u}^T}\mathbf{Q}^{L-l}\right)\left(\mathbf{Q}^{l-1}\sqrt{\mathbf{u}}\right) \quad \text{with} \quad \mathbf{Q}^l_{i'k',ik} = \sqrt{\mathbf{u}_{i'k'}}\mathbf{P}_{i'k',ik}\sqrt{\mathbf{u}_{ik}}$$
$$\text{using} \quad \mathbf{u}_{ik} = \mu(\mathbf{x}_i, \phi_k) \quad \text{and} \quad \mathbf{P}_{i'k',ik} = \rho(\mathbf{x}_{i'}, \phi_{k'}|\mathbf{x}_i, \phi_k)\,. \quad (1)$$

with the notation Q^l for the l-th power of the matrix 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 the background. Subjects are asked to detect the contours, and their performance averaged over trials is recorded in dependence on various control parameters as e.g. element distance or orientation alignment jitter.

For our stimulations, 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 stimulus evidence for a specific visual field location i, $\mu(\mathbf{x}_i, \phi)$, will be termed the *afferent input.* In the next paragraph, we 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 $\mathbf{s}$ 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 $M(x, \mu, \kappa) = \exp(\kappa\cos(x-\mu))/(2\pi I_0(\kappa))$ with concentration parameter κ and circular mean μ. I_0 denotes a modified Bessel function of the first kind of order 0. Thus, we define the afferent input distribution centered around θ_i with a width of σ_{aff} as $\mathbf{u}_{ik} = M(2\phi_k, 2\theta_i, 1/{\sigma_{aff}}^2)$ To complicate contour detection, a jitter of amplitude $\eta_\theta = 0, 1, 2, \ldots$ can be applied to the contour which randomly rotates every edge by exactly $\pm\eta_\theta(2\pi/K)$. We assume the AF to be translational invariant and rotational invariant with respect to the source edge,

$$\rho\,(r, \alpha, \beta|0, 0, 0) = F(r)/2\,\Big[M(\gamma_\alpha, 0, 1/{\sigma_\alpha}^2)M(\gamma_\beta, 0, 1/{\sigma_\beta}^2)$$
$$+ M(\gamma_\alpha, \pi, 1/{\sigma_\alpha}^2)M(\gamma_\beta, \pi, 1/{\sigma_\beta}^2)\Big] \quad (2)$$

where $r = |\mathbf{x}' - \mathbf{x}|$ is the distance between two contour elements, α denotes 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 F is choosen to be $F(r) = 1$ for the six nearest neighbors on the grid, 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 $\sigma_\alpha = \sigma_\beta/2 = \pi/12$, which are typical values for an AF. As information processing in the brain is noisy, additive noise η_u can be applied on the afferent input $\mathbf{u}$. This noise is uniformly distributed between 0 and $\eta_u \max\{\mathbf{u}\}$. In our simulations, we choose $N = 324$ and $L = 9$, together with $K = 24$ and also 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. We find that these three performances are comparable. In the figures, we only show the performance determined with $L_{max} = 5$, 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 $\eta_u = 0.05$, setting $\eta_\theta = 0$. In both cases the performance was better for dynamic noise than for static noise, as the dynamic noise averages out over time (see Fig. 1 (left)). With

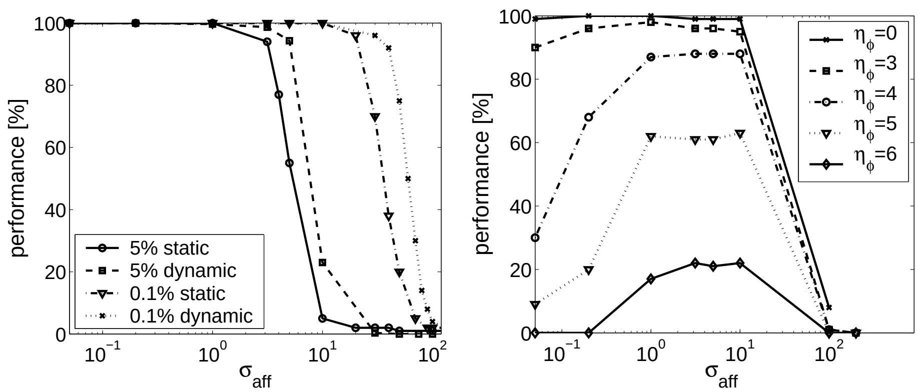

Fig. 1. Contour detection performance in percent correct in dependence on the afferent input width σ_{aff} for different noise levels and types (left), and for different orientation jitter strengths η_θ. Performance gets better for lower noise levels, and static noise is impeding the performance stronger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

broader afferent input, the performance decreases as 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 a noise of $\eta_u = 0.001$ ($\eta_u = 0.05$) where the difference between maximum and minimum of the afferent input tuning curve reaches and exceeds the size of the noise. Performance finally drops to chance level where a distinction between the noisy Von Mises-function and an equidistribution is no more 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 performance starts below 100%, then increases significantly with σ_{aff} (in our example between $\sigma_{aff} = 0.2$ and $\sigma_{aff} = 1$), and later decreases to chance level.

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 following condition

$$o_{i'i} = \sum_{k'=1}^{K} \mathbf{u}_{i'k'} \mathbf{v}_{i'k'i} \gg 0 \,. \tag{3}$$

$\mathbf{v}$ defined as $\mathbf{v}_{i'k'i} = \sum_{k=1}^{K} \mathbf{P}_{i'k',ik} \mathbf{u}_{ik}$ 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 curves $\mathbf{u}_{i'}$ and $\mathbf{v}_{i'i}$ 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 widths σ_α and σ_β of the AF 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 is 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

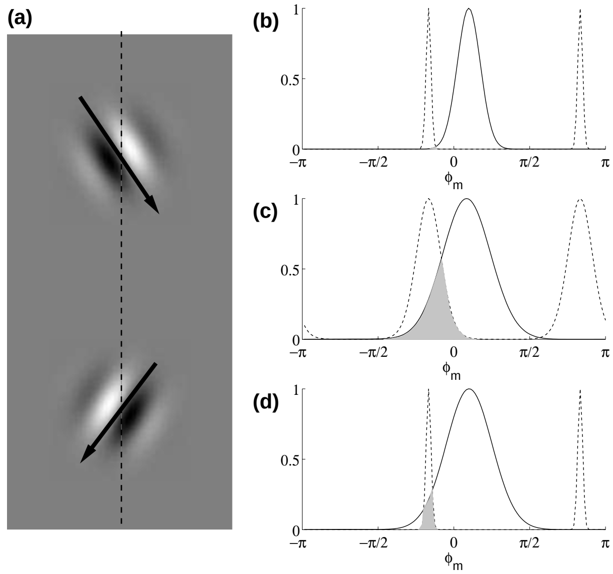

Fig. 2. (a) Stimulus situation considered in the text: two contour elements i and i' are shown with an orientation jitter of $-\eta_\theta$ and $+\eta_\theta$ (arrows), respectively. (b)-(d) show the afferent input distribution $\mathbf{u}_{i'k'}$ at destination element i' (dashed line), together with the lateral input $\mathbf{v}_{i'k'i}$ (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.

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 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 advantages and disadvantages 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.

Acknowledgements: We would like to thank Sunita Mandon and Andreas Kreiter for numerous fruitful discussions and valuable data. Udo Ernst was funded by the Volkswagen-Foundation, project 'Neuronal mechanisms of gestalt perception'.

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.