How ideal do macaque monkeys integrate contours?

Udo A. Ernst [a,*] Sunita Mandon [b] Klaus R. Pawelzik [a]
Andreas K. Kreiter [b]

[a] *Institute for Theoretical Physics, University of Bremen, Germany*
[b] *Institute for Brain Research, University of Bremen, Germany*

Abstract

In this contribution, we present a joint study combining psychophysical investigations with macaque monkeys and information theoretical modelling of contour integration. We study the detection of S- and U-shaped contours of aligned gabor patches among a set of unaligned distractors. The approach allows to compare *quantitatively* the performance of monkeys with the optimal performance predicted by a stochastic model. We find that the performance of the animals is very high, and that it closely approaches the performance of the ideal observer.

Key words: gestalt perception, psychophysics, macaque monkey, stochastic model.

1 Introduction

Detecting the boundaries of an object is one of the fundamental steps during the process of object recognition. Although in complex real world environments the contours of an object are often partially occluded by other objects, they are easily detected by the visual system. Central to the contour-integration process is the rule of 'good continuation', first described 1912 by Max Wertheimer. Detailed psychophysical studies gave a quantitative description of the effect of several parameters in perception of continuity. In particular, studies of a contour-detection paradigm, where subjects have to find a continous path of Gabor elements embedded in noise have shown that local parameters of the contour elements such as alignment or spacing (6; 7; 8) as well as global parameters such as length and geometry (8; 9; 10; 11) influence the process of contour integration. Based on neuroanatomical and physiological properties of the mammalian primary visual cortex (V1) many authors

* E-Mail: uernst@physik.uni-bremen.de

suggested a prime importance of this area for object delineation and contour perception (6; 11; 12). It is generally accepted that in particular long-range connections, preferentially linking neurons with similar orientation (13; 14) and collinear aligned receptive fields (15) contribute to contour perception. Further evidence is provided by the finding that orientation-specific neurons in V1 are markedly fascilitated by a single collinear line element placed outside the classical receptive field (16; 17; 18).
Contour integration has also been investigated in numerous modeling studies, which mostly employ a 'classical' neural architecture closely resembling the typical connection patterns found in primary visual cortex. Many of these models have been quite successful in explaining the psychophysical phenomena, see e.g. (2). In some cases, given a careful calibration of the model parameters, it is even possible to quantitatively predict the performance of observers, see e.g. (4; 3).
However, there is still a gap between the heuristical descriptions of the Gestalt critera, and a solid mathematical framework or definition of Gestalt properties. Due to this lack of a theoretical foundation, it is e.g. often not possible to judge whether the performance of an observer in a contour detection task is high or low, compared to the optimal performance.
Recently, some theoretical work by Williams and Thornber (1) has bridged this gap by providing a stochastic framework for contour integration. They define a 'contour' as the outcome of a stochastic process with a known probability distribution. Due to this definition, it is straightforward to invert this process and to define an algorithm, which takes an arbitrary ensemble of oriented elements, and computes for each element the relative probability to belong to a contour. Thus, for every contour which is hidden among a set of distracting elements, one can determine the contour detection performance of an optimal observer. This upper limit can be compared to the detection ability of real observers; in our case, to macaque monkeys – allowing to judge how close to the optimum the brain is operating. Similarily, this comparison can also constrain the structure and the dynamics of the underlying neuronal circuitry.

2 Experimental Methods

We trained two monkeys to perform a two-alternative forced-choice paradigm. The monkeys sat in a primate chair in front of a 21" CRT screen (1152 by 864 pixels, refresh rate 100 Hz). Each trial started with the appearance of a small fixation spot in the display center. After a fixation period of 800–1000 ms an image was presented containing either a S- or an U-shaped path of Gabor elements embedded in a random array of Gabor elements of identical contrast and spatial frequency (Fig. 1). The Gabor-element stimulus was displayed for 600 ms and then followed by a gray screen of same mean luminance containing saccade targets on the left and right side. Monkeys indicated by means of a saccade the side of the display containing the contour. Only correct trials were rewarded with a drop of fruit juice. If the monkey's fixation deviated more

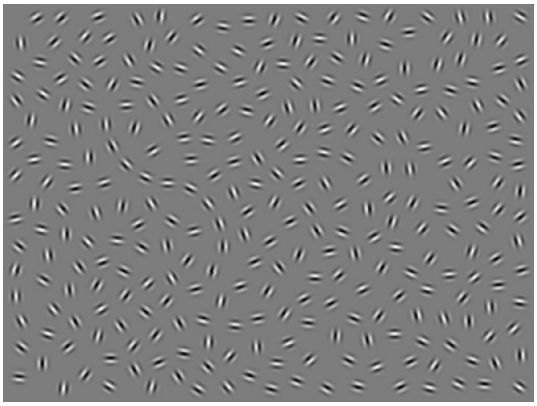

Fig. 1. Example of a gabor element stimulus with S-shaped contour and orientation jitter of $\alpha = 0$.

than 0.6^o of visual angle, or if they responded either too early or too late, the trial was automatically aborted without reward.
The path of the contour was computed using the method described in (6). In short, the path was constructed by joining together nine line elements to form a S- or an U-shape. The length of the line was four times the period of the modulating sinusoid of the Gabor elements. A random jitter between $\pm 5\%$ was added to each line length. In addition, the shapes were distorted by adding a random jitter between $\pm 30^o\%$ to the angle that would lead to a regular S- or U-shape ($22.5^o \pm 6.7^o$). Then, a Gabor element was placed at the halfway of each line. The contour was turned and moved to an arbitrary position on the left or right side of the image. After the positions of the contour elements had been determined, background elements were added to the image by using a method adopted from (8). This method ensures same pair-distribution functions for contour and background elements. To investigate the influence of colinearity on contour salience, the orientation of the path elements was varied relative to the orientation of the path. The orientation of each contour element was given a random change of $\pm\alpha$, where α was either $8^o, 10^o, 12^o, 14^o, 16^o, 18^o, 20^o$ in monkey 1 and $8^o, 10^o, 12^o, 14^o, 16^o, 36^o$ in monkey 2.

3 Theory

In the present article, we used the framework of Williams and Thornber to compute the saliency of the Gabor elements in the stimuli used in the experiments. We briefly summarise the method which is explained in detail in (1).
Stochastic model. A contour consisting of L elements at positions $\mathbf{x}_l$ with orientations ϕ_l can be described as the result of a stochastic process with transition probabilities $p(\mathbf{x}, \phi|\mathbf{y}, \theta)$. p gives the probability to find the next element of a contour at position $\mathbf{x}$ with orientation ϕ, given that the previous element of the contour had an orientation θ at position $\mathbf{y}$. p is often referred to as the 'association field', which is depicted in Fig. 2. Now assume a stimulus

consisting of $M = L + K$ elements $(\mathbf{x}_m, \phi_m)$ with L targets forming a contour, and K distractors. Then $P_{m,n} := p(\mathbf{x}_m, \phi_m | \mathbf{x}_n, \phi_n)$ with $m, n \in \{1, \ldots, M\}$ defines a probability matrix which can be used to compute the 'saliencies' c_m of the elements, which is defined as the relative probability that element m is a part of a contour within the display. According to (1), c_m is given by the expression

$$c_m = \frac{(\mathbf{s}\,\overline{\mathbf{s}}^T)_{m,m}}{\mathbf{s}^T\overline{\mathbf{s}}} \quad \text{with} \quad \lambda\mathbf{s} = P\mathbf{s} \quad \text{and} \quad \lambda\overline{\mathbf{s}} = P^T\overline{\mathbf{s}}\,. \tag{1}$$

Association field. We approximate the statistics of the relative positions and orientations of the elements of the contour by using a closed mathematical expression

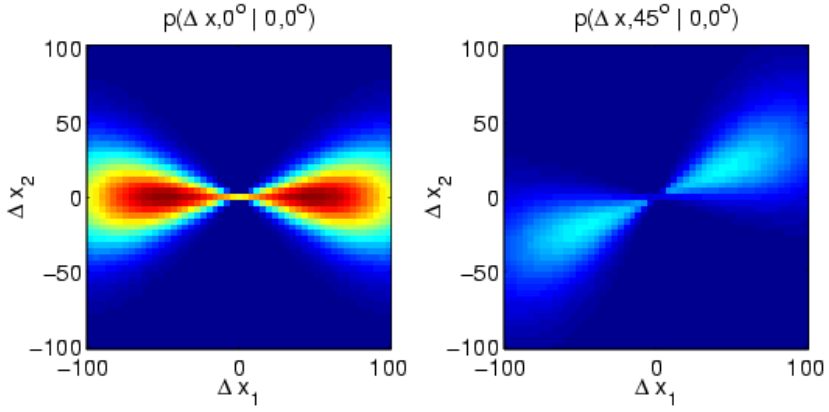

Fig. 2. Examples of the association field used in this article. On the left, we show the association field of an element with orientation 0^o located in the center to all elements with the same orientation at every other location. On the right, we show the association field of the same element in the center to all elements with 45^o-orientation in the surround.

$$p(\mathbf{x}, \phi | \mathbf{y}, \theta) = C \exp\left(\frac{(|\mathbf{x}-\mathbf{y}| - \sigma_d)^2}{2\sigma_d^2}\right) \exp\left(\frac{||\phi - \theta||_{2\pi}^2}{2\sigma_a^2}\right) \exp\left(\frac{||\delta||_{2\pi}^2}{2\sigma_c^2}\right) \tag{2}$$

where C is a normalisation constant, $\sigma_d = 50$ is the inter-element distance, σ_a and σ_c are length scales for the alignment and curvature of the association field, and δ is the angle difference between the orientation at $\mathbf{x}$, resulting from an ideal continuation of a contour starting at $\mathbf{y}$, and ϕ. The norm $||a||_b$ is defined as $a = a'$ for $a' < b/2$, and $a = b - a'$ for $a' \geq b/2$ with $a' = mod(a, b)$. We adjusted the length scales to closely reflect the statistical properties of the target contours, thus $\sigma_a = \pi/4$ and $\sigma_c = \pi/4$.

Computation of performance. Having obtained the saliences $\mathbf{c}$ via Eq.(1), the performance of an ideal observer has been computed from statistical detection theory. In particular, the observer has to decide on which semi-field of the stimulus display, one of it consisting of only K_1 distractors, and the other one consisting of both K_2 distractors and L targets, the contour had been

displayed ($K = K_1 + K_2$). The K_1 and K_2 distractor saliencies define probability distributions $p_{K_1}(c)$ and $p_{K_2}(c)$, respectively, while the L targets define a probability distribution $p_L(c)$. These distributions can be integrated to form the cumulative probabilities $q_{K_1}(c_0)$, $q_{K_2}(c_0)$, $q_L(c_0)$ describing the probability of a saliency smaller than c_0. Assume that we choose a convenient c_0 as our decision criterion, and compare it to the maximum saliency c_{max} of the elements in a semifield. If $c_{max} > c_0$, we then decide that the semi-field contains a contour. If contours can be found with probability 0.5 in each semifield, the probability of a correct decision $p_{correct}(c_0)$ is then given by the expression

$$p_{correct}(c_0) = \frac{1}{2}(1 - q_{K_1}(c_0)^{K_1} + q_{K_2}(c_0)^{K_2} q_L(c_0)^L) \tag{3}$$

The maximum of $p_{correct}$ over all possible c_0 then defines the performance of an optimal observer.

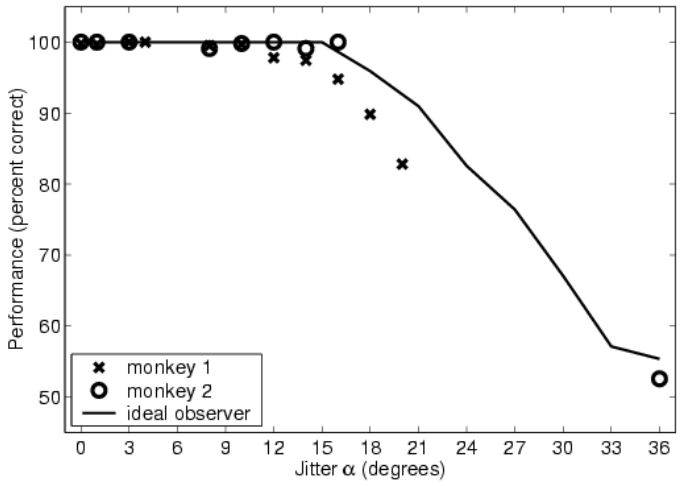

Fig. 3. Performance displayed as the percentage of correct detections, for monkey 1 (crosses) and monkey 2 (open circles), as well as for an ideal observer (solid line).

4 Results

Fig. 3 shows the detection performance of both monkeys as a function of the jitter α. The performance remains above 95% up to a jitter of $\alpha = 16^o$. For larger α values, the performance drops towards chance level, which is reached approximately around $\alpha = 36^o$.

For the comparison with the theoretical framework, we choose $\sigma_a = \pi/4$ and $\sigma_c = \pi/4$ as parameters for the association field. These values reflect the statistics of the geometrical dependencies between the target elements in the stimulus displays presented to the monkeys. We then computed the saliencies and best detection performances using the algorithm outlined in the Theory-section. The resulting curve is shown in the same graph together with the experimental results (Fig. 3). We observed that the experimental performance

closely approaches the predicted optimal performance. These results suggest that the neuronal mechanisms responsible for contour integration are highly optimised for this task.

5 Discussion

Our experiments show that monkeys are able to detect the contours with a similar or even better performance as human observers (6). In addition, the statistical approach outlined allows to investigate the phenomenon of contour integration on a solid mathematical background. In particular, one can compare the performance of monkeys to the performance of an optimal observer. Our simulations indicate that the detection ability of the animals closely approaches this optimum.
Currently, we optimise the detailed match between the algorithm generating the psychophysical stimuli and the statistical model underlying our theoretical studies. This is done by including additional information about the variation of the association field parameters, preferrably by using the statistical properties of contours in natural images. In addition, we are collecting further psychophysical data to span the full range of the psychometric curve.

References

[1] Williams L.R., Thornber K.K., Neural Computation 12, 1683–1711, 2001.
[2] Li Z., Netw. Comput. Neural Syst. 10, 187–212, 1999.
[3] Ernst U.A., Etzold A., Herzog M.H., Eurich C.W., Neurocomputing, in press, 2003.
[4] Ernst U.A., Eurich C.W., in The Handbook of Brain Theory, ed. Arbib M.A., MIT Press, 294–300, 2002.
[5] Z. Li, Neur. Comp. 13, 1749–1780, 2001.
[6] Field D.J., Hayes A., Hess R.F., Vision Res. 33, 173–193, 1993.
[7] Li W., Gilbert C.D., J. Neurophysiol. 88, 2846–2856, 2002.
[8] Braun J., Spatial Vision 12, 211–225, 1999.
[9] Kovács I., Julesz B., PNAS 90, 7495–7497, 1993.
[10] Pettet M.W., Vision Res. 39, 551–557, 1999.
[11] Pettet M.W., McKee S.P., Grzywacz N.M., Vision Res. 38, 865–879, 1998.
[12] Yen S.C., Finkel L.H., Vision Res. 38, 719–741, 1998.
[13] Gilbert C.D., Wiesel T.N., J. Neurosci. 9, 2432–2442, 1989.
[14] Stettler D.D., Das A., Bennett J., Gilbert C.D., Neuron 36, 739–750, 2002.
[15] Schmidt K.E., Goebel R., Löwel S., Singer W., Eur J. Neurosci. 9, 1083–1089, 1997.
[16] Kapadia M.K., Ito M., Gilbert C.D., Westheimer G., Neuron 15, 843–856, 1995.
[17] Ito M., Gilbert C.D., Neuron 22, 593–604, 1999.
[18] Polat U., Mizobe K., Pettet M.W., Kasamatsu T., Norcia A.M., Nature 391, 580–584, 1998.