

Statistical Invariants of Spatial Form: From Local AND to Numerosity

Christoph ZETZSCHE¹, Konrad GADZICKI and Tobias KLUTH
Cognitive Neuroinformatics, University of Bremen, Germany

Abstract Theories of the processing and representation of spatial form have to take into account recent results on the importance of holistic properties. Numerous experiments showed the importance of “set properties”, “ensemble representations” and “summary statistics”, ranging from the “gist of a scene” to something like “numerosity”. These results are sometimes difficult to interpret, since we do not exactly know how and on which level they can be computed by the neural machinery of the cortex. According to the standard model of a local-to-global neural hierarchy with a gradual increase of scale and complexity, the ensemble properties have to be regarded as high-level features. But empirical results indicate that many of them are primary perceptual properties and may thus be attributed to earlier processing stages. Here we investigate the prerequisites and the neurobiological plausibility for the computation of ensemble properties. We show that the cortex can easily compute common statistical functions, like a probability distribution function or an autocorrelation function, and that it can also compute abstract invariants, like the number of items in a set. These computations can be performed on fairly early levels and require only two well-accepted properties of cortical neurons, linear summation of afferent inputs and variants of nonlinear cortical gain control.

Keywords. shape invariants, peripheral vision, ensemble statistics, numerosity

Introduction

Recent evidence shows that our representation of the world is essentially determined by holistic properties [1,2,3,4,5,6]. These properties are described as “set properties”, “ensemble properties”, or they are characterized as “summary statistics”. They reach from the average orientation of elements in a display [1] over the “gist of a scene” [7,8], to the “numerosity” of objects in a scene [9]. For many of these properties we do not exactly know by which kind of neural mechanisms and on which level of the cortex they are computed. According to the standard view of the cortical representation of shape, these properties have to be considered as high-level features because the cortex is organized in form of a local-to-global processing hierarchy in which features with increasing order of abstraction are computed in a progression of levels [10]. At the bottom, simple and locally restricted geometrical features are computed, whereas global and complex properties are represented at the top levels of the hierarchy. Across levels, invariance is system-

¹Corresponding Author: Christoph Zetzsche, Cognitive Neuroinformatics, FB3, University of Bremen, P.O. Box 330 440, 28334 Bremen, Germany; E-mail: zetzsche@informatik.uni-bremen.de
Research supported by **DFG** (SFB/TR 8 Spatial Cognition, A5-[ActionSpace])

atically increased such that the final stages are independent of translations, rotations, size changes, and other transformations of the input. However convincing this view seems on first sight, it creates some conceptual difficulties.

The major difficulty concerns the question of what exactly is a low-level and a high-level property. Gestalt theorists already claimed that features considered high-level according to a structuralistic view are primary and basic in terms of perception. Further doubts have been raised by global precedence effects [11]. Similar problems arise with the recently discovered ensemble properties. The gist of a scene, a high-level feature according to the classical view, can be recognized in 150 msec [7,12,13,14] and can be modeled using low-level visual features [8]. In addition, categories can be shown to be faster processed than basic objects, contrary to the established view of the latter as entry-level representations [15]. A summary statistics approach, also based on low-level visual features, can explain the holistic processing properties in the periphery of the visual field [4,16,17]. What is additionally required in these models are statistical measures, like probability distributions and autocorrelation functions, from which it is not known how and on which level of the cortical hierarchy they can be realized.

One of the most abstract ensemble properties seems to be the number of elements in a spatial configuration. However, the ability to recognize this number is not restricted to humans with mature cognitive abilities but has also been found in infants and animals [9,18], recently even in invertebrates [19]. Neural reactions to numerosity are fast (100 msec in macaques [20]). And finally there is evidence for a “direct visual sense for number” since number seems to be a primary visual property like color, orientation or motion, to which the visual system can be adapted by prolonged viewing [21].

The above observations on ensemble properties raise a number of questions, from which the following are addressed in this paper: Sect. 1: Can the cortex compute a probability distribution? Sect. 2: And also an autocorrelation function? By which kind of neural hardware can this be achieved? Sect.3: Can the shape of individual objects also be characterized by such mechanisms? Sect. 4: What is necessary to compute such an abstract property like the number of elements in a spatial configuration? Can this be achieved in early sensory stages?

1. Neural Computation of a Probability Distribution

Formally, the probability density function $p_e(e)$ of a random variable \mathbf{e} is defined via the cumulative distribution function: $p_e(e) \triangleq \frac{dP_e(e)}{de}$ with $P_e(e) = Pr[\mathbf{e} \leq e]$. Their empirical counterparts, the histogram and the cumulative histogram, are defined by use of *indicator functions*. For this we divide the real line into m bins $(e^{(i)}, e^{(i+1)})$ with bin size $\Delta e = e^{(i+1)} - e^{(i)}$. For each bin i , an indicator function is defined as

$$Q_i(e) = 1_i(e) = \begin{cases} 1, & \text{if } e^{(i)} < e \leq e^{(i+1)} \\ 0, & \text{else} \end{cases} \quad (1)$$

An illustration of such a function is shown in Fig. 1a. From N samples e_k of the random variable \mathbf{e} we then obtain the histogram as $h(i) = \frac{1}{N} \sum_{k=1}^N Q_i(e_k)$. The cumulative histogram $H_e(e)$ can be computed by changing the bins to $(e^{(1)}, e^{(i+1)})$ (cf. Fig. 1b), and by performing the same summation as for the normal histogram. The reverse cumulative

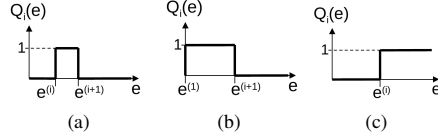


Figure 1. Indicator functions. Basic types are: (a) indicator function for computation of a classical histogram. (b) indicator function for a cumulative histogram. (c) indicator function for a reverse cumulative histogram.

histogram $\bar{H}(i)$ is simply the reversed version of the cumulative histogram. The corresponding bins are $\Delta e_i = (e^{(i)}, e^{(i+1)}]$ and the indicator functions are defined as (Fig. 1c)

$$Q_i(e) = 1_i(e) = \begin{cases} 1, & \text{if } e \geq e^{(i)} \\ 0, & \text{else} \end{cases} \quad (2)$$

The corresponding system is shown in Fig. 2.

The three types of histograms have identical information content since they are related to each other as

$$h(i) = H((i+1)) - H(i) = \bar{H}(i) - \bar{H}(i+1) \quad \text{and} \quad H(i) = 1 - \bar{H}(i) = \sum_{j=1}^i h(j). \quad (3)$$

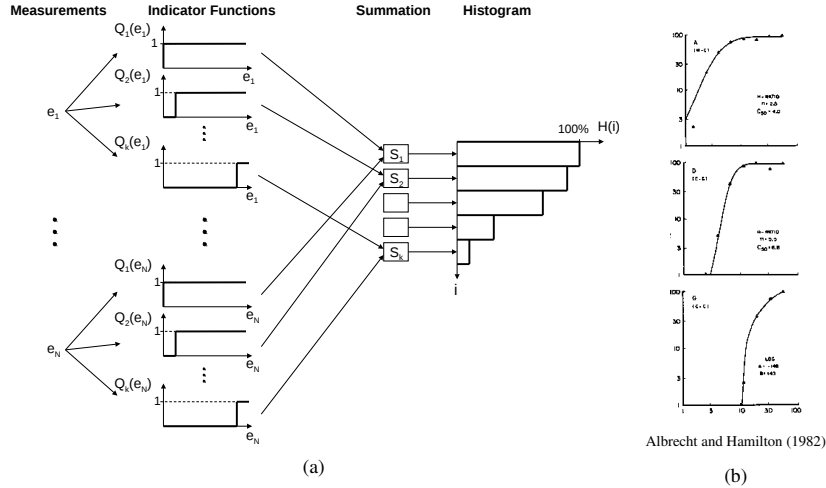


Figure 2. Computation of the reverse cumulative histogram. (a) shows the set of input variables e_1 to e_n over which the histogram should be computed. Each of these variables is input to a set of *indicator functions* $Q_i(e_k)$. For each bin of the histogram there is a summation unit S_i which sums over all indicator function outputs with index i , i.e. over all $Q_i(e_k)$.

(b) The response functions of three neurons in the visual cortex [22]. They show a remarkable similarity to the indicator functions for the reverse cumulative histogram. First, they come with different sensitivities. Second, they exhibit an independence on the input strength: once the threshold and the following transition range is exceeded the output remains constant and does no longer increase when the input level is increased.

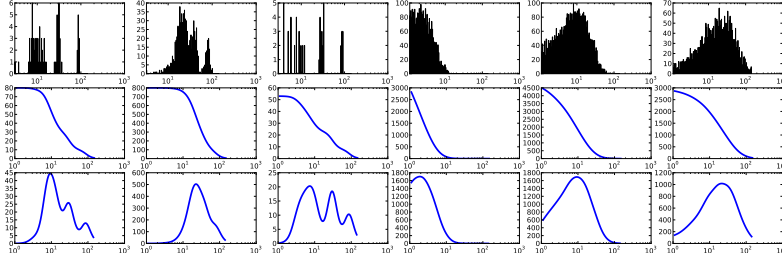


Figure 3. Neurobiological computation of a reverse cumulative histogram. The upper row shows several examples of input probability distributions. The second row shows the corresponding reverse cumulative histograms computed by a dense set of simulated neurons. The third row shows the estimated probability distributions as derived from the neural representation by use of Eq. (3).

How does all this relate to visual cortex? Has the architecture shown in Fig. 2a any neurobiological plausibility? The final summation stage is no problem since the most basic capability of neurons is computation of a linear sum of their inputs. But how about the indicator functions? They have two special properties: First, the indicator functions come with different *sensitivities*. An individual function does only generate a non-zero output if the input e exceeds a certain level, a kind of threshold, which determines the sensitivity of the element $e^{(i)}$ in Eq. (2) and Fig. 1c. To cover the complete range of values, different functions with different sensitivities are needed (Fig. 2a). Second, the indicator functions exhibit a certain *independence of the input level*. Once the input is clearly larger than the threshold, the output remains constant (Fig. 1c).

Do we know of neurons which have such properties, a range of different sensitivities, and a certain independence of the input strength? Indeed, cortical gain control (or normalization), as first described in early visual cortex (e.g. [22]) but now believed to exist throughout the brain [23], yields exactly these properties. Gain-controlled neurons (Fig. 2b) exhibit a remarkable similarity to the indicator functions used to compute the reverse cumulative histogram, since they (i) come with different sensitivities, and (ii) provide an independence of the input strength in certain response ranges.

The computation of a reverse cumulative histogram thus is well in reach of the cortex. We only have to modify the architecture of Fig. 2a by the smoother response functions of cortical neurons. The information about a probability distribution available to the visual cortex is illustrated in Fig. 3. The reconstructed distributions, as estimated from the neural reverse cumulative histograms, are a kind of Parzen-windowed (lowpass-filtered) versions of the original distributions.

2. Neural Implementation of Auto- and Cross-Correlation Functions

A key feature of the recent statistical summary approach to peripheral vision [4,6,24,16] is the usage of auto- and cross-correlation functions. These functions are defined as

$$h(i) = \frac{1}{N} \sum_{k=-N/2+1}^{N/2} e(k) \circ g(i+k), \quad (4)$$

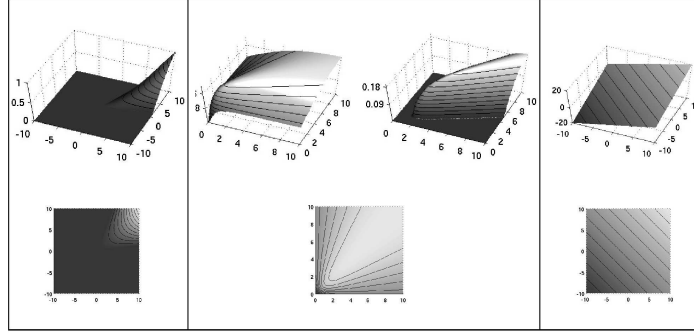


Figure 4. Different types of AND-like functions. Each function is of the type $g_k = g(s_i, s_j)$, i.e. assigns an output value to each combination of the two input values. The upper row shows the functions as surface plots, the lower row as iso-response curves. Left: Mathematical multiplication of two inputs. Center: AND-like combinations that can be obtained by use of cortical gain control (normalization). The upper left figure shows the classical gain control without additional threshold. The upper right figure shows the same mechanism with an additional threshold. This results in a full-fledged AND with a definite zero response in case that only one of the two inputs is active. Right: The linear sum of the two input values for comparison purposes.

where autocorrelation results if $e(k) = g(k)$ and where \circ indicates multiplication. With respect to their neural computation, the outer summation is no problem, but the crucial function is the *nonlinear multiplicative interaction between two variables*. A neural implementation could make use of the Babylonian trick $ab = \frac{1}{4}[(a+b)^2 - (a-b)^2]$ [25,26,27], but this requires two or more neurons for the computation and thus far there is neither evidence for such a systematic pairing of neurons nor for actual multiplicative interactions in the visual cortex. However, exact multiplication is not the key factor: a reasonable statistical measure merely requires provision of a matching function such that $e(k)$ and $g(i+k)$ generate a large contribution to the autocorrelation function if they are similar, and a small contribution if they are dissimilar. For this, it is sufficient to provide a neural operation which is AND-like [27,28]. Surprisingly, such an AND-like operation can be achieved by the very same neural hardware as used before, the cortical gain control mechanism, as shown in [28]. Cortical gain control [22,29] applied to two different features $s_i(x, y)$ and $s_j(x, y)$ can be written as

$$g_k(x, y) = g(s_i(x, y), s_j(x, y)) := \max \left(0, \frac{s_i + s_j}{(\sqrt{s_i^2 + s_j^2} + \varepsilon)\sqrt{2}} - \Theta \right) \quad (5)$$

where $k = k(i, j)$, ε is a constant which controls the steepness of the response and Θ is a threshold. The resulting nonlinear combination is comparable with an AND-like operation of two features and causes a substantial nonlinear increase of the neural selectivity, as illustrated in Fig. 4.

Of course there will be differences between a formal autocorrelation function and the neurobiological version, but the essential feature, the signaling of good matches in dependence of the relative shifts will be preserved (Fig. 5).

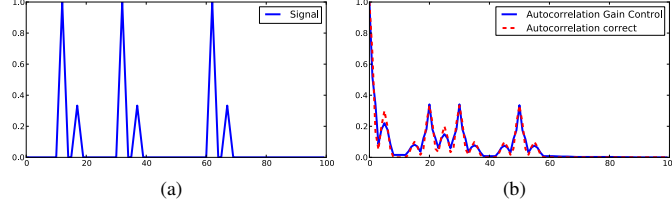


Figure 5. Mathematical and neurobiological autocorrelation functions. (a) shows a test input and (b) the corresponding mathematical (red dotted) and neurobiological (blue) autocorrelation function.

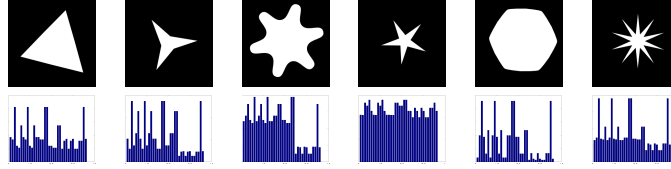


Figure 6. Different shapes and the corresponding integral features. We used parameter combinations of six different orientations $\theta_i = (i-1)\pi/6$, $i = 1, \dots, 6$, and four different scales $r_i = 2^{-i}$, $i = 1, \dots, 4$. The radial half-bandwidth was set to $f_{r,h} = \frac{1}{3}r$ and the angular half-bandwidth was constant with $f_{\theta,h} = \pi/12$. Each parameter combination creates pairs of variables for each x,y-position which are AND-combined by the gain control mechanism described in Eq. (5) as $g_k(x,y) = g(s_i(x,y), s_j(x,y))$.

3. Figural Properties from Integrals

We extracted different features $s_{r,\theta}$ from the image luminance function $l = l(x,y)$ by applying a Gabor-like filter operation $s_{r,\theta}(x,y) = (l * \mathcal{F}^{-1}(H_{r,\theta}))(x,y)$ where \mathcal{F}^{-1} denotes the inverse Fourier transformation and the filter kernel $H_{r,\theta}$ is defined in the spectral space. We distinguish two cases (even and odd) which can be seen in the following definition in polar coordinates:

$$H_{r,\theta}^{even}(f_r, f_\theta) := \begin{cases} \cos^2\left(\frac{\pi}{2} \frac{f_r - r}{2f_{r,h}}\right) \cos^2\left(\frac{\pi}{2} \frac{f_\theta - \theta}{2f_{\theta,h}}\right) & , (f_r, f_\theta) \in \Omega_{r,\theta} \\ 0 & , \text{else,} \end{cases}$$

with $\Omega_{r,\theta} := \{(f_r, f_\theta) | f_r \in [r - 2f_{r,h}, r + 2f_{r,h}] \wedge f_\theta \in [\theta - 2f_{\theta,h}, \theta + 2f_{\theta,h}] \cap [\theta + \pi - 2f_{\theta,h}, \theta + \pi + 2f_{\theta,h}]\}$, where $f_{r,h}$ denotes the half-bandwidth in radial direction and $f_{\theta,h}$ denotes the half-bandwidth in angular direction. $H_{r,\theta}^{odd}$ is defined as the Hilbert transformed even symmetric filter kernel.

Various AND combinations of these oriented features (see caption Fig. 6) are obtained by the gain-control mechanism described in Eq. (5). The integration over the whole domain results in *global* features $F_k := \int_{\mathbb{R}^2} g_k(x,y) d(x,y)$ which capture basic shape properties (Fig. 6).

4. Numerosity and Topology

One of the most fundamental and abstract ensemble properties is the number of elements of a set. Recent evidence (see Introduction) raised the question at which cortical level

the underlying computations are performed. In this processing, a high degree of invariance has to be achieved, since numerosity can be recognized largely independent of other properties like size, shape and positioning of elements. Models which address this question in a neurobiologically plausible fashion, starting from individual pixels or neural receptors instead of an abstract type of input, are rare. To our knowledge, the first approach in this direction has been made in [30]. A widely known model [31] has a shape-invariant mapping to number which is based on linear DOG filters of different sizes, which substantially limits the invariance properties. A more recent model is based on unsupervised learning but has only employed moderate shape variations [32]. In [30] we suggested that the necessary invariance properties may be obtained by use of a theorem which connects local measurements of the differential geometry of the image surface with global topological properties [30,33]. In the following we will build upon this concept.

The key factor of our approach is a relation between surface properties and a topological invariant as described by the famous Gauss-Bonnet theorem. In order to apply this to the image luminance function $l = l(x, y)$ we interpret this function as a surface $S := \{(x, y, z) \in \mathbb{R}^3 | (x, y) \in \Omega, z = l(x, y)\}$ in three-dimensional real space. We then apply the formula for the Gaussian curvature

$$K(x, y) = \frac{l_{xx}(x, y)l_{yy}(x, y) - l_{xy}(x, y)^2}{(1 + l_x(x, y)^2 + l_y(x, y)^2)^2}, \quad (6)$$

where subscript denotes the differentiation in the respective direction (e.g. $l_{xy} = \frac{\partial^2 l}{\partial x \partial y}$). The numerator of (6) can also be written as $D = \lambda_1 \lambda_2$ where $\lambda_{1,2}$ are the eigenvalues of the Hessian matrix of the luminance function $l(x, y)$ which represent the partial second derivatives in the principal directions. The values and signs of the eigenvalues give us the information about the shape of the luminance surface S in each point, whether it is elliptic, hyperbolic, parabolic, or planar. Since Gaussian curvature results from the multiplication of the second derivatives $\lambda_{1,2}$ it is zero for the latter two cases. It has been shown that this measure can be generalized in various ways, in particular towards the use of neurophysiologically realistic Gabor-like filters instead of the derivatives [27,30]. The crucial point, however, is the need for *AND combinations of oriented features* [27,30] which can be obtained as before by the neural mechanism of cortical gain control [28].

The following corollary from the Gauss-Bonnet theorem is the basis for the invariance properties in the context of numerosity.

Corollary 4.1 *Let $S \subset \mathbb{R}^3$ be a closed two-dimensional Riemannian manifold. Then*

$$\int_S K \, dA = 4\pi(1 - g) \quad (7)$$

where K is the Gaussian curvature and g is the genus of the surface S .

We consider the special case where the luminance function consists of multiple objects (polyhedra with orthogonal corners) with constant luminance level. We compare the surface of this luminance function to the surface of a cuboid with holes that are shaped like the polyhedra. The trick is that the latter surface has a genus which is determined by the number of holes in the cuboid and which can be determined by the integration of the local curvature according to Eq. (7). If we can find the corresponding contributions of

the integral on the image surface, we can use this integral to count the number of objects. We assume the corners to be locally sufficiently smooth such that the surfaces are Riemannian manifolds. The Gaussian curvature K then is zero almost everywhere except on the corners. We hence have to consider only the contributions of the corners. It turns out that these contributions can be computed from the elliptic regions only if we use different signs for upwards and downwards oriented elliptic regions. We thus introduce the following operator which distinguishes the different types of ellipticity in the luminance function. Let $\lambda_1 \geq \lambda_2$, then the operator $N(x, y) := |\min(0, \lambda_1(x, y))| - |\max(0, \lambda_2(x, y))|$ is always zero if the surface is hyperbolic and has a positive sign for positive ellipticity and a negative one for negative ellipticity. We thus can calculate the numerosity feature which has the ability of counting objects in an image by counting the holes in an imaginary cuboid as follows:

$$F = \int_{\Omega} \frac{N(x, y)}{(1 + l_x(x, y)^2 + l_y(x, y)^2)^{\frac{3}{2}}} d(x, y). \quad (8)$$

The crucial feature of this measure are contributions of fixed size and with appropriate signs from the corners. The denominator can thus be replaced by a neural gain control mechanism and an appropriate renormalization. For the implementation here we use a shortcut which gives us straight access to the eigenvalues. The numerator $D(x, y)$ of (6) can be rewritten as

$$D(x, y) = l_{xx}l_{yy} - \frac{1}{4}(l_{uu} - l_{vv})^2 = \frac{1}{4}[(l_{xx} + l_{yy})^2 - \underbrace{((l_{xx} - l_{yy})^2 + (l_{uu} - l_{vv})^2)}_{=: \varepsilon^2}] = \frac{1}{4}(\Delta l^2 - \varepsilon^2) \quad (9)$$

with $u := x \cos(\pi/4) + y \sin(\pi/4)$ and $v := -x \sin(\pi/4) + y \cos(\pi/4)$. The eigenvalues then are $\lambda_{1,2} = \frac{1}{2}(\Delta l \pm |\varepsilon|)$ and we can directly use them to compute $N(x, y)$. Application of this computation to a number of test images is shown in Fig. 7.

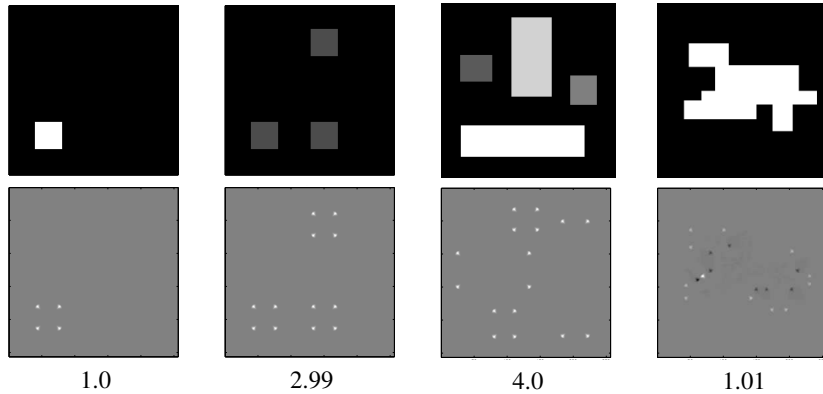


Figure 7. Based on a close relation to topological invariants the spatial integration of local curvature features can yield highly invariant numerosity estimates. The numerical values in the last row are the normalized integrals of the filter outputs (middle row).

5. Conclusion

Recent evidence shows that ensemble properties play an important role in perception and cognition. In this paper, we have investigated by which neural operations and on which processing level statistical ensemble properties can be computed by the cortex. Computation of a probability distribution requires *indicator functions* with different sensitivities, and our reinterpretation of cortical gain control suggests that this could be a basic function of this neural mechanism. The second potential of cortical gain control is the computation of *AND-like feature combinations*. Together with the linear summation capabilities of neurons this enables the computation of powerful invariants and summary features. We have repeatedly argued that AND-like feature combinations are essential for our understanding of the visual system [27,30,34,35,36,28]. The increased selectivity of nonlinear AND operators, as compared to their linear counterparts, is a prerequisite for the usefulness of integrals over the respective responses [30,28]. We have shown that such integrals of AND features are relevant for the understanding of texture perception [37], of numerosity estimation [30], and of invariance in general [28]. Recently, integrals over AND-like feature combinations in form of auto- and cross-correlation functions have been suggested for the understanding of peripheral vision [4,16,17].

A somewhat surprising point is that linear summation and cortical gain control, two widely accepted properties of cortical neurons, are the only requirements for the computation of ensemble properties. These functions are already available at early stages of the cortex, but also in other cortical areas [23]. The computation of ensemble properties may thus be an ubiquitous phenomenon in the cortex.

Acknowledgement

This work was supported by DFG, SFB/TR8 Spatial Cognition, project A5-[ActionSpace].

References

- [1] S. C. Dakin and R. J. Watt. The computation of orientation statistics from visual texture. *Vision Res*, 37(22):3181–3192, 1997.
- [2] D. Ariely. Seeing Sets: Representation by Statistical Properties. *Psychol Sci*, 12(2):157–162, 2001.
- [3] Lin Chen. The topological approach to perceptual organization. *Visual Cognition*, 12(4):553–637, 2005.
- [4] B. Balas, L. Nakano, and R. Rosenholtz. A summary-statistic representation in peripheral vision explains visual crowding. *J Vis*, 9(12):13.1–18, 2009.
- [5] G. A. Alvarez. Representing multiple objects as an ensemble enhances visual cognition. *Trends Cog Sci*, 15(3):122–31, 2011.
- [6] R. Rosenholtz, J. Huang, and K. Ehinger. Rethinking the role of top-down attention in vision: effects attributable to a lossy representation in peripheral vision. *Front Psychol*, 3:13, 2012.
- [7] S. Thorpe, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 381(6582):520–522, 1996.
- [8] A. Oliva and A. Torralba. Modeling the Shape of the Scene: A Holistic Representation of the Spatial Envelope. *International Journal of Computer Vision*, 42(3):145–175, 2001.
- [9] E. M. Brannon. The representation of numerical magnitude. *Curr Opin Neurobiol*, 16(2):222–9, 2006.
- [10] J. Hegde and D.J. Felleman. Reappraising the Functional Implications of the Primate Visual Anatomical Hierarchy. *The Neuroscientist*, 13(5):416–421, 2007.
- [11] D. Navon. Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology*, 9(3):353–383, 1977.
- [12] M. R. Greene and A. Oliva. The Briefest of Glances. *Psychol Sci*, 20(4):464–472, 2009.

- [13] J. Hegdé. Time course of visual perception: coarse-to-fine processing and beyond. *Prog Neurobiol*, 84(4):405–39, 2008.
- [14] M. Fabre-Thorpe. The characteristics and limits of rapid visual categorization. *Front Psychol*, 2:243, 2011.
- [15] M. J-M Macé, O. R. Joubert, J-L. Nespoulous, and M. Fabre-Thorpe. The time-course of visual categorizations: you spot the animal faster than the bird. *PLoS one*, 4(6):e5927, 2009.
- [16] J. Freeman and E. P. Simoncelli. Metamers of the ventral stream. *Nature neuroscience*, 14(9):1195–1201, 2011.
- [17] H. Strasburger, I. Rentschler, and M. Jüttner. Peripheral vision and pattern recognition: a review. *J Vis*, 11(5):13, 2011.
- [18] A. Nieder, D. J. Freedman, and E. K. Miller. Representation of the quantity of visual items in the primate prefrontal cortex. *Science*, 297(5587):1708–11, 2002.
- [19] H. J. Gross, M. Pahl, A. Si, H. Zhu, J. Tautz, and S. Zhang. Number-based visual generalisation in the honeybee. *PLoS one*, 4(1):e4263, 2009.
- [20] J. D. Roitman, E. M. Brannon, and M. L. Platt. Monotonic coding of numerosity in macaque lateral intraparietal area. *PLoS biology*, 5(8):e208, 2007.
- [21] J. Ross and D. C. Burr. Vision senses number directly. *Journal of vision*, 10(2):10.1–8, 2010.
- [22] D. G. Albrecht and D. B. Hamilton. Striate cortex of monkey and cat: contrast response function. *J Neurophysiol*, 48(1):217–237, Jul 1982.
- [23] M. Carandini and D. J. Heeger. Normalization as a canonical neural computation. *Nature Reviews Neurosci*, 13:51–62, Jul 2012.
- [24] R. Rosenholtz, J. Huang, A. Raj, B. J. Balas, and L. Ilie. A summary statistic representation in peripheral vision explains visual search. *J Vis*, 12(4):1–17, 2012.
- [25] H.L. Resnikoff and R.O. Wells. *Mathematics in Civilization*. Popular Science Series. Dover, 1984.
- [26] E. H. Adelson and J. R. Bergen. Spatiotemporal energy models for the perception of motion. *J. Opt. Soc. Am. A*, 2(2):284–99, 1985.
- [27] C. Zetsche and E. Barth. Fundamental limits of linear filters in the visual processing of two-dimensional signals. *Vision Res*, 30(7):1111–1117, 1990.
- [28] C. Zetsche and U. Nuding. Nonlinear and higher-order approaches to the encoding of natural scenes. *Network*, 16(2–3):191–221, 2005.
- [29] D.J. Heeger. Normalization of cell responses in cat striate cortex. *Visual Neurosci*, 9(2):181–198, 1992.
- [30] C. Zetsche and E. Barth. Image surface predicates and the neural encoding of two-dimensional signal variations. In B. E. Rogowitz and Jan P. A., editors, *Proc SPIE*, volume 1249, pages 160–177, 1990.
- [31] S. Dehaene and J. P. Changeux. Development of elementary numerical abilities: a neuronal model. *J. Cogn. Neurosci.*, 5(4):390–407, 1993.
- [32] I. Stoianov and M. Zorzi. Emergence of a 'visual number sense' in hierarchical generative models. *Nat Neurosci*, 15(2):194–6, 2012.
- [33] M. Ferraro E. Barth and C. Zetsche. Global topological properties of images derived from local curvature features. In L. P. Cordella C. Arcelli and G. Sanniti di Baja, editors, *Visual Form 2001. Lecture Notes in Computer Science*, pages 285–294, 2001.
- [34] C. Zetsche, E. Barth, and B. Wegmann. The importance of intrinsically two-dimensional image features in biological vision and picture coding. In A. B. Watson, editor, *Digital images and human vision*, pages 109–138. MIT Press, Cambridge, MA, 1993.
- [35] G. Krieger and C. Zetsche. Nonlinear image operators for the evaluation of local intrinsic dimensionality. *IEEE Transactions Image Processing*, 5:1026–1042, 1996.
- [36] C. Zetsche and G. Krieger. Nonlinear mechanisms and higher-order statistics in biological vision and electronic image processing: review and perspectives. *J Electronic Imaging*, 10(1):56–99, 2001.
- [37] E. Barth, C. Zetsche, and I. Rentschler. Intrinsic 2D features as textons. *J. Opt. Soc. Am. A*, 15(7):1723–1732, 1998.