WEBER'S LAW AND THE STATISTICS OF THE NATURAL ENVIRONMENT

Christoph Zetzsche[1], Florian Röhrbein[1,2] and Kerstin Schill[1]

[1]*Cognitive Neuroinformatics, Department FB3 Mathematics/Informatics, University of Bremen, 28334 Bremen, Germany*
[2]*NeuroComp Lab, Albert Einstein College of Medicine, Yeshiva University, Bronx, NY 10461, USA*

Abstract

Weber's law is one of the most fundamental properties of visual processing. This raises the question of why and how the underlying neural circuitry has developed. Here we propose that the emergence of Weber's law can be seen as a by-product of a more general evolutionary strategy for the development of sensory systems: the adaptation to the statistical regularities of natural scenes. Many basic properties of early vision have already been successfully explained within this framework. Here we extend this approach by measuring the joint statistics of neighbouring pixels of natural images under varying illumination conditions. We demonstrate that a linear decorrelating transform would leave significant statistical dependencies between the responses. We then show that the removal of these dependencies can be achieved by learning from the statistics a nonlinear gain control mechanism which can be implemented as ROG (ratio of Gaussian) filter. Weber's law is a direct consequence of this nonlinear operation. One single basic principle, the reduction of statistical dependencies between sensory messages, thus seems to be sufficient to derive all essential processing properties of early vision.

Investigations of the relation between natural scene statistics and the neural mechanisms of early vision have revealed that many basic properties, such as the frequency and orientation selectivity of neurons in the visual cortex, can be explained as an information-theoretically optimized adaptation to the statistical redundancies of the natural environment (for review see, e.g., Olshausen and Field, 2004, Simoncelli, 2003, Zetzsche and Krieger, 2001). This approach can also be extended to more complicated cortical processing properties, as in complex cells or in the extra-classical receptive field surround (e.g., Lyu and Simoncelli, 2009, Zetzsche and Nuding, 2005). According to the information-theoretic approach the neural operations represent a transformation of the state space coordinates which matches the representation to the structure of the multivariate probability distribution. One major criterion for a good match is the reduction of the statistical dependencies (ideally: statistical independence). Often this can be achieved by linear transforms, as in independent component analysis (ICA), but some statistical dependencies require nonlinear operations. Here we investigate whether this general goal of statistical independence can also explain how the nonlinear operations underlying Weber's law have developed in the visual system.

Natural Scene Statistics and Linear Filter Decompositions

For this, we first measured the joint statistical distribution of the responses of neighbouring retinal receptors to natural scenes with varying lightning conditions. 12 images were randomly taken from a database of natural images (van Hateren and van der Schaaf, 1998) under exclusion of non-natural objects or portions of sky. Before the statistics
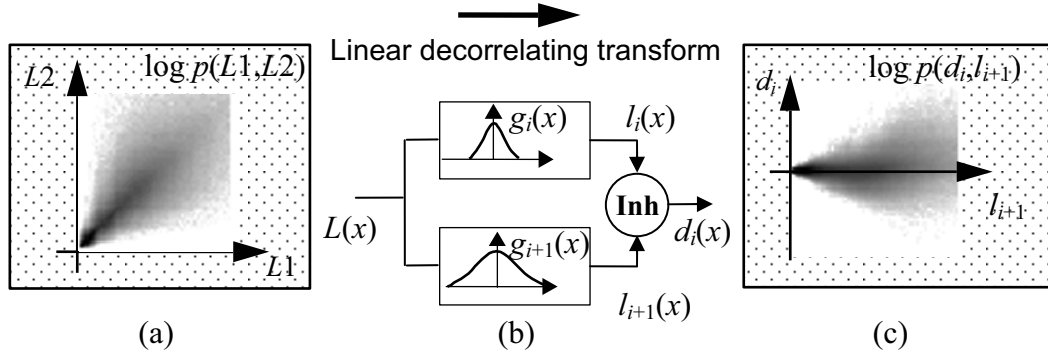
Fig. 1. (a) Joint pdf $p(L1, L2)$ of two neighboring pixels (with $L1 = L(x_0), L2 = L(x_0 + \Delta x)$) in scenes with spatially and/or temporally varying illumination (contrast enhanced). (b) Linear lateral inhibition (DOG). (c) Decorrelated joint statistics of linear DOG.

were computed, the images were converted to an absolute intensity scale by a mapping which takes into account the aperture and the exposure time used in recording each image. The resulting probability density function (pdf) of two neighboring pixels exhibits a typical shape: a high correlation between the pixel values and a systematic outward widening of the distribution towards the higher intensity values (Fig. 1a). Obviously, there exist strong statistical dependencies (redundancies) in the signal.

How can the sensory system get rid of them? A classical method to for this is a linear decorrelating transformation. The crucial operation for decorrelation is bandpass filtering, which in early vision is achieved by lateral inhibition, commonly formalized as Difference Of Gaussian (DOG) filter operation with $d_i(x) = l_i(x) - l_{i+1}(x)$ where l_i and l_{i+1} are the image signal filtered by Gaussians g_i and g_{i+1} with different spatial spread (Fig. 1b).

Decorrelation can be seen as a "rotation" of the coordinate system (Fig. 1c). However, this rotation by the linear transformation can only exploit the statistical dependencies of second order but cannot provide a separation of the higher-order dependencies which are reflected in the systematic dependence of the variance of the DOG-response d_i on the local average luminance l_{i+1} (Fig. 1c). Note that the complete pdf has more than the two dimensions shown and that the multivariate statistical dependencies extend both across spatial positions and spatial frequencies (spatial scales). Regarding position it is easy to see that the effect shown in Fig. 1c also holds for spatially neighboring samples of the DOG response $d_i(x)$. Since the average local luminance $l_i(x)$ varies only slowly over position x, the variances of a sample $d_i(x_0 - \Delta x/2)$ and of its neighbor $d_i(x_0 + \Delta x/2)$ depend both in the same way as shown in Fig. 1c on $l_{i+1}(x_0)$. So although the signal $d_i(x)$ is quite well decorrelated in the second-order sense, its samples $d_i(x_0 - \Delta x/2)$ and $d_i(x_0 + \Delta x/2)$ still have the shown substantial statistical dependencies.

For understanding the statistical dependencies across frequency we can decompose the complete signal $L(x)$ into a DOG pyramid (Burt and Adelson, 1983). For resolution level (or scale) i we can write

$$g_i = \sum_{j=i}^{n} d_j = \underbrace{(l_i - l_{i+1})}_{d_i} + \underbrace{(l_{i+1} - l_{i+2}) + (l_{i+2} - l_{i+3}) + ... + (l_{n-1} - l_n) + l_n}_{\sum_{i+1}^{n} d_j = l_{i+1}}. \qquad (1)$$
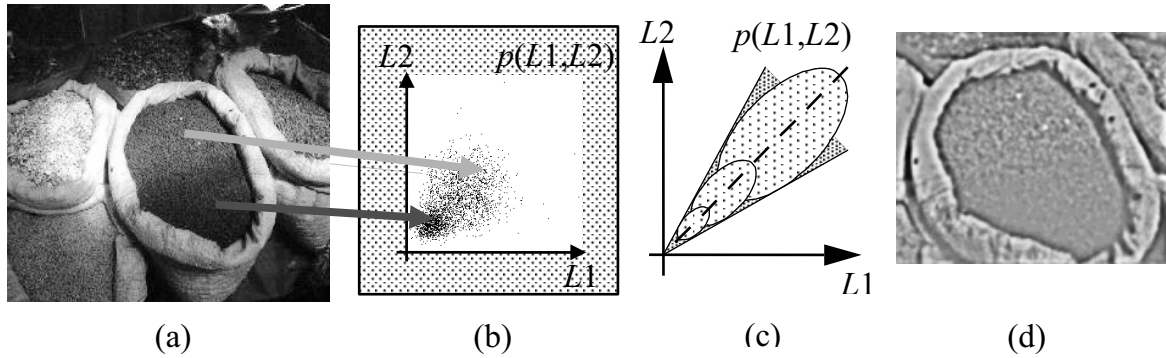
Fig. 2. (a-c) The influence of varying illumination on the pdf $p(L1, L2)$ and (d) the response of a linear inhibition scheme (DOG).

Here d_i denotes a DOG channel with resolution i and $\sum d_j$ the sum of the other DOG channels with lower frequencies $(j \geq i+1)$, which turns out to be equivalent to the local mean l_{i+1}. The statistical dependencies between a channel d_i and the other channels d_j in the DOG representation can hence also be described by the two-dimensional pdf $p(d_i, l_{i+1})$ in Fig. 1c. Together this demonstrates that the multivariate pdf of natural images is not separable in *linear* Cartesian coordinates (neither by PCA or ICA).

Can we understand why the statistical structure arises which causes the deficit of linear decorrelation? This process is illustrated in Figure 2. First, it is important that the retinal input $L(x)$ results from a nonlinear, multiplicative combination of an illumination component $I(x)$ and an reflectance component $R(x)$, i.e. $L(x)=I(x)R(x)$. Second, natural scenes are illuminated by different illumination functions $I_k(x)$, which vary across space and time. Fig. 2a shows a typical configuration with spatially varying illumination. The grain in the sack is illuminated partially by direct bright light, and partially by indirect dim light. This varying illumination will cause that similar or even identical local reflectance functions $R_i(x_i + \Delta x) \approx R_j(x_j + \Delta x)$ (in our example the homogenous grain which is identical in the bright and the dark area) are transformed into different luminance functions $L_i(x_i + \Delta x) \neq L_j(x_j + \Delta x)$. The corresponding statistical contributions $p_i\big(L_i(x_i + \Delta x)\big)$ and $p_j\big(L_j(x_j + \Delta x)\big)$ are scaled versions of each other, and constitute together the pdf $p(L(x))$ (in our example the pdf $p(L1, L2)$ of neighbouring pixels from both areas of the grain region (Fig. 2b)). The principle is schematically illustrated in Fig. 2c. The combination of various scaled subpopulations in the final pdf is the reason for the statistical dependency of the DOG response d_i on the mean that has been revealed in Fig. 1c. The crucial point is that a linear decomposition cannot separate the nonlinear interaction of the reflectance component and the illumination component (Fig. 2d). The linear response is proportional to the linear local differences in the image. The response to the grain texture is hence smaller in the dimly illuminated region than in the bright region, i.e. it is dependent on the illumination. It is sometimes assumed that a mere suppression of the low frequencies by a linear band-pass filter, or similarly, a sole shift of the operating point is already sufficient for illumination invariance but this is clearly not the case (Fig. 2d).

Nonlinear Removal of Statistical Dependencies

The removal of these dependencies requires a nonlinear transformation. A suitable locus for an appropriate nonlinearity is the inhibitory interaction, labeled as **Inh** in Fig. 1b. We thus
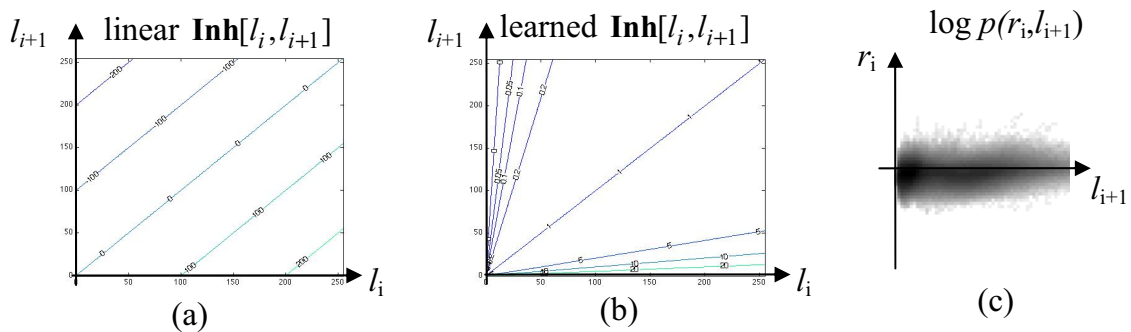
Fig. 3. Comparison of linear and learned nonlinear inhibition: (a) linear inhibition (DOG), (b) nonlinear inhibition learned by minimizing statistical dependendencies in the neural representation, and (c) joint pdf resulting from the divisive nonlinear inhibition.

have investigated whether the biological system can learn an appropriate nonlinear form of this inhibition on the sole basis of the statistical regularities of natural scenes. For this we again used the images from the van Hateren database and optimized the nonlinear interaction for the inhibition in order to minimize the statistical dependencies. The nonlinear operation we optimized can take various shapes, ranging between the extremes of a linear and a divise behavior, with a continuum of intermediate forms in between.

The learned inhibition is not longer a linear subtraction as in the DOG model (Fig. 3a) but becomes stronger in proportion to the low-resolution luminance l_{i+1} (Fig. 3b).

Evaluation shows that it comes very close to an exact division, i.e. $Inh[l_i, l_{i-1}] \approx r_i = \dfrac{l_i}{l_{i+1}}$.

This is reasonable because the crucial factor that causes the statistical dependencies in the linear setting is the direct *proportionality* of the filter response *to the local mean* l_{i+1}. The nonlinear transform has to get rid of this proportionality, i.e. it has to *reduce* the gain of the system *in proportion to the local mean*. The learned inhibition thus represents an adaptive gain control mechanism and a formalization of such a mechanism is a divisive interaction by a "ratio of Gaussians" (ROG) operator (Sperling, 1970, Zetzsche and Hauske, 1989). (A logarithmic transducer function would formally have a similar effect but would be much less suited for the processing of a wide dynamic range). The resulting joint pdf does no longer exhibit the statistical dependency but is approximately separable, which amounts to statistical independence of the components (Fig. 3c). In addition, such a nonlinear operation also can successfully recover the reflectance component (Fig. 4).

The nonlinear transformation that can remove the statistical dependencies yields Weber's law as a direct consequence. It produces an input-output relation in which a
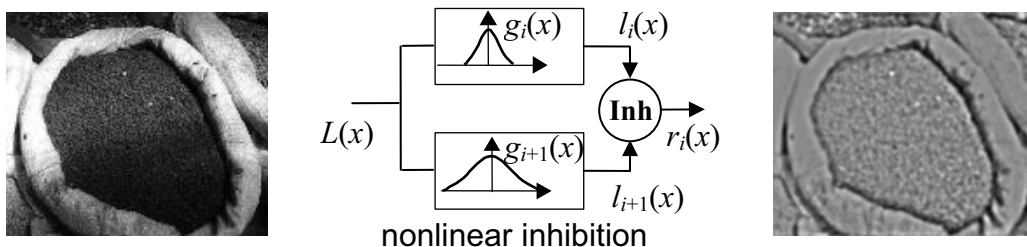

Fig. 4. Processing of a scene with varying illumination by nonlinear gain control. The operator can separate the influence of the illumination by responding only to the reflectance pattern (which is a homogeneous texture of grain in this example, cf. Fig. 2d).
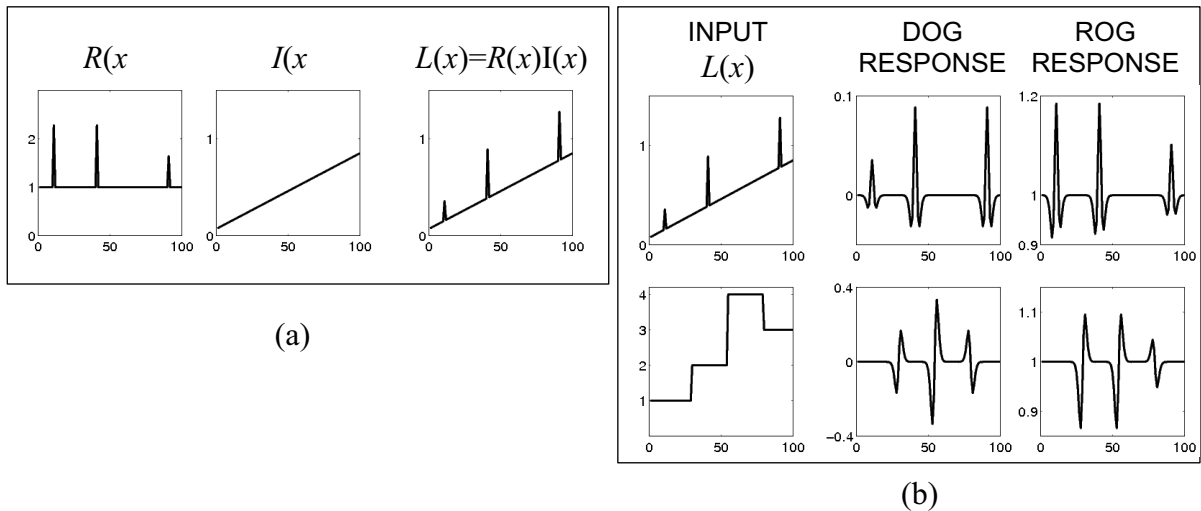
Fig. 5. Nonlinear gain control by the ROG mechansim and Weber's law. (a) The input $L(x)$ results from a multiplicative combination of the reflectance $R(x)$ and the illumination $I(x)$. (b) A linear DOG operator responds in proportion to the linear signal differences. The nonlinear ROG operator yields constant response increments for input ratios $\Delta L/L_0$=const., i.e., yields Weber's law. Below is a second example with steps as input.

fixed output increment is caused by all those inputs for which the input increment ΔL is proportional to the local mean L_0. This is just Weber's law $\Delta L/L_0$=const. (Fig. 5).

Discussion

In this paper we have investigated how a biological system can make use of the specific statistical dependencies of natural scenes to learn a nonlinear inhibitory processing scheme which (i) minimizes the statistical dependencies of the nonlinear filter response on the local mean lumincance, (ii) is able to approximately separate the reflectance component from the illumination component, and (iii) produces Weber's law. In an earlier investigation (Röhrbein and Zetzsche, 2002) we have not explicitly addressed the learning of the nonlinear inhibition from the statistics, but we already observed that the ROG response is statistically independent of the local mean luminance. This is in agreement with a similar observation by Mante et al. (2005), who found that local rms contrast and local mean luminance are statistically independent in natural images. This accordance seems reasonable given that the nonlinear ROG response and the rms contrast are closely related. However, the conceptual status of the entities being studied is different in the two approaches. Simply said, what in our approach is an empirical result of our statistical investigations is in their approach a definition of a variable. Our aim is to derive the structure of the neural transformation from the specific form of the statistical dependencies in images. (Here we addressed luminance gain control, for cortical contrast gain control see Schwartz and Simoncelli (2001) and Zetzsche et al. (1999)). Mante et al. focus on the *relation* between the mechanisms, and test the hypothesis that, due to the statistical independence of the measured variables, the two gain control mechanisms should operate independently of each other. This is plausible, given the fact that the statistical dependencies that are exploited by contrast gain control differ from those analysed here. Finally, it should be noted that there remain spatially distributed statistical dependencies between luminance and contrast (Lindgren et al., 2008).

     In conclusion, Weber's law describes a fundamental nonlinearity of the visual system. Here we suggested a simple explanation for the development of the underlying neural machinery: Substantial statistical redundancies are a typical characteristic of sensory

messages, and the exploitation of such redundancies seems to be a universal strategy for an efficient representation of sensory information. Our statistical analysis has shown that there exist significant statistical dependencies between early sensory signals which cannot be exploited by classical linear decorrelation schemes. However, the system can learn from these dependencies how to develop a nonlinear gain control that yields approximate illumination invariance and Weber's law. Like other basic visual functions, Weber's law can thus be seen as a consequence of one single principle: the visual system seeks to exploit the statistical redundancies of natural scenes.

Acknowledgements

References

Burt, P.J. & Adelson, E.H. (1983). The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications, 31(4),* 532-540.

Lindgren, J.T., Hurri, J. & Hyvärinen, A. (2008). Spatial dependencies between local luminance and contrast in natural images. *Journal of Vision, 8(12),* 1-13.

Lyu, S. & Simoncelli, E.P. (2009). Nonlinear extraction of 'Independent Components' of natural images using radial Gaussianization. *Neural Computation, 21(6),* 1485-1519.

Mante, V., Frazor, R.A., Bonin, V., Geisler, W.S. & Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nature Neuroscience, 8,* 1690 – 1697.

Olshausen, B.A. & Field, D.J. (2004) Sparse coding of sensory inputs. *Current Opinion in Neurobiology, 14,* 481-487.

Röhrbein, F. & Zetzsche, C. (2002). The statistics of natural scenes and Weber's law. In P. Würtz & M. Lappe (Eds.), *4. Workshop Dynamic Perception,* Berlin: AKA/ Amsterdam: IOS Press; pp. 233-238.

Schwartz, O. & Simoncelli, E. (2001) Natural signal statistics and sensory gain control. *Nature Neuroscience, 4(8),* 819-825.

Shapley, R. & Enroth-Cugell, C. (1984) Visual Adaptation and Retinal Gain Controls, In: N. Osborne & G. Chader (Eds*.), Progress in Retinal Research, vol. 3,* London: Pergamon; pp. 263-346

Simoncelli,E.P. (2003). Vision and the statistics of the visual environment. *Current Opinion in Neurobiology, 13(2),* 144-149.

Sperling, G. (1970). Model of visual adaptation and contrast detection. *Perception and Psychophysics, 8,* 143-157.

van Hateren, J. & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. Lond. B, 265,* 359-366.

Zetzsche, C. & Hauske, G. (1989). Multiple channel model for the prediction of subjective image quality. In: *Proc. SPIE, 1077,* Bellingham, WA: SPIE; pp. 209-216.

Zetzsche, C., Krieger, G. & Wegmann, B. (1999). The atoms of vision. *J. Opt. Soc. Am. A, 16,* 1554-1565

Zetzsche, C & Krieger, G. (2001) Nonlinear mechanisms and higher-order statistics in biological vision and electronic image processing: review and perspectives. *J. Electronic Imaging, 10(1),* 56-99.

Zetzsche, C. & Nuding, U. (2005). Nonlinear and higher-order approaches to the encoding of natural scenes. *Network, 16*(*2-3*), 191-221.