Brain responses strongly correlate with Weibull image statistics when processing natural images

H. Steven Scholte
Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

Sennay Ghebreab
Intelligent Systems Lab Amsterdam, University of Amsterdam, Amsterdam, The Netherlands

Lourens Waldorp
Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands

Arnold W. M. Smeulders
Intelligent Systems Lab Amsterdam, University of Amsterdam, Amsterdam, The Netherlands

Victor A. F. Lamme
Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands, & The Netherlands Institute of Neuroscience (NIN-KNAW), Amsterdam, The Netherlands

The visual appearance of natural scenes is governed by a surprisingly simple hidden structure. The distributions of contrast values in natural images generally follow a Weibull distribution, with beta and gamma as free parameters. Beta and gamma seem to structure the space of natural images in an ecologically meaningful way, in particular with respect to the fragmentation and texture similarity within an image. Since it is often assumed that the brain exploits structural regularities in natural image statistics to efficiently encode and analyze visual input, we here ask ourselves whether the brain approximates the beta and gamma values underlying the contrast distributions of natural images. We present a model that shows that beta and gamma can be easily estimated from the outputs of X-cells and Y-cells. In addition, we covaried the EEG responses of subjects viewing natural images with the beta and gamma values of those images. We show that beta and gamma explain up to 71% of the variance of the early ERP signal, substantially outperforming other tested contrast measurements. This suggests that the brain is strongly tuned to the image's beta and gamma values, potentially providing the visual system with an efficient way to rapidly classify incoming images on the basis of omnipresent low-level natural image statistics.

Keywords: natural image statistics, event-related potential, covariate, Weibull distribution, model, early vision, X-cell, Y-cell

Citation: Scholte, H. S., Ghebreab, S., Waldorp, L., Smeulders, A. W. M., & Lamme, V. A. F. (2009). Brain responses strongly correlate with Weibull image statistics when processing natural images. *Journal of Vision, 9*(4):29, 1–15, http://journalofvision.org/9/4/29/, doi:10.1167/9.4.29.

Introduction

Natural scenes are inherently complex, containing a wide variety of structural combinations. Still, the luminance, contrast, frequencies, and orientation of an arbitrary natural scene adhere to certain (statistical) regularities (Brumswik & Kamiya, 1953; Daugman, 1989). Upon closer inspection, our visual world therefore seems to be governed by simpler hidden structures. For instance, it has long been known that contrast histograms for natural images tend to be unimodal and skewed in the direction of lower absolute contrasts (Field, 1987; Ruderman & Bialek, 1994).

The distribution of contrasts in an image may be of particular importance for visual perception. Contrasts in an image carry key information, as they frequently coincide with the 2D silhouette projected off a 3D object. Moreover, contrast almost always reveals folds in the surface geometry and changes in the surface albedo. In

doi: 10.1167/9.4.29 Received October 1, 2008; published April 30, 2009 ISSN 1534-7362

addition, what is more, contrast does so independent of the accidental recording conditions.

This shows the importance of contrast per se, but the *distribution* of contrast may provide additional information. It has been shown that the more the geometry and the surface structure of a scene are coherent in space, the higher the correlation will be among the corresponding image contrast values (Rousselet, Pernet, Bennett, & Sekuler, 2008; Thomson, 2001). Contrast values in coherent scenes will therefore generally be highly correlated, and any device that records contrasts over a patch of the visual field (such as the receptive fields of neurons in the visual system) thus produces a sum over correlated values. As sums over correlated values follow a Weibull distribution (Meeker & Escobar, 1998), it can be predicted that the distribution of contrasts in natural images of coherent surface structures should follow a Weibull distribution when recorded with receptive fields of a given extent. Indeed, it has been observed (Geusebroek & Smeulders, 2002, 2005; Simoncelli, 1999) that the distributions of contrast values in natural images follow a Weibull function of the following form:

$$p(f) = c\ e^{\left(\frac{f-\mu}{\beta}\right)^{\gamma}}, \quad (1)$$

where c is a normalization constant that transforms the frequency distribution into a probability distribution. The parameter μ, denoting the origin of the contrast distribution, is generally close to zero for natural images. We normalize out this parameter by subtracting the smallest contrast value from the contrast data, leaving two free parameters per image, β (beta) and γ (gamma).

Observations (Geusebroek & Smeulders, 2002) on 45,000 stock photos of natural images, covering a wide variety of topics, revealed that 60% of all images have strict Weibull-shaped contrast distributions. The remaining percentage of images has a distribution close to Weibull or is highly regular. These images are typically composed of two (or more) parts in the image, each having a distinct Weibull contrast distribution.

Examples of natural images, and their Weibull-fitted distributions of contrasts, are shown in Figure 1, together with an artificial image whose contrast distribution does not adhere to the Weibull function. When correlation in the contrast values is low, gamma of the Weibull distribution approaches 2 and the distribution becomes a Gaussian distribution. This is known as the central limit theorem (Meerschaert & Scheffler, 2001). This condition occurs when the scene has limited coherence with many variations in geometrical depths and surface structures. In contrast, when the coherence in the scene is high, gamma approaches 1 and the distribution of contrasts conforms to a power law distribution. This is known as the law of fractals (Mandelbrot, 1983). The condition occurs when the statistical regularities of the scene are stochastic yet homogeneous in geometry and surface structure. For the majority of natural images, the contrast distribution will neither be Gaussian nor power law but will have beta and gamma values in between (Geusebroek & Smeulders, 2002).

If Weibull-shaped contrast distributions are so omnipresent, two questions arise: first, is the visual system adapted to it? Second, does the visual system exploit the parameters of the Weibull distribution (i.e., beta and gamma) to analyze image content? It is generally believed that, as a consequence of the hidden structures occurring in natural scenes, the visual system has adapted itself to these recurring patterns, which have been imprinted into the brain to achieve a more efficient encoding of ecologically relevant images (Daugman, 1989; Field, 1987). However, even if the brain is sensitive to environmental regularities, it does not necessarily follow that the brain exploits image statistics to characterize, categorize, or otherwise analyze incoming information. Some proposals have been made as to how the brain would employ the inherent statistical regularities of natural scenes, for example, to calculate the reflectance properties of objects from the distribution of luminance values (Motoyoshi, Nishida, Sharan, & Adelson, 2007). Likewise, the spatial frequency content (Field, 1987; Hsiao & Millane, 2005; Parraga, Troscianko, & Tolhurst, 2000; Torralba & Oliva, 2003), the luminance distribution (Fleming & Bulthoff, 2005; Victor, Chubb, & Conte, 2005), and the distribution of contrasts (Geisler, 2008; Ruderman & Bialek, 1994; Turiel, Mato, Parga, & Nadal, 1998) in complex scenes have been analyzed and shown to capture some aspect of the content of the scene. We propose that, because beta and gamma reflect the correlation hidden in the Weibull-shaped contrast distribution, these parameters could provide a very efficient encoding for the degree of coherence in the scene.

We first show how the beta and gamma values of an image can be closely approximated by combining the output of linear (X-cell-like) and non-linear (Y-cell-like) filtering systems, in order to show that the visual system is capable of computing these values. To address the idea that the human visual system is tuned to the parameters of the Weibull distribution, we test whether the evoked EEG responses of human subjects viewing natural images correlate with the beta and gamma values of the images. We conclude that indeed beta and gamma of the Weibull distribution could inform the brain about the spatial coherence in the perceived scene. We speculate that beta and gamma are ecologically relevant dimensions helping the brain to rapidly classify images.

Methods

Images

For the experiments, we used a total of 1600 images with a resolution of 217 × 345 or 345 × 217 pixels

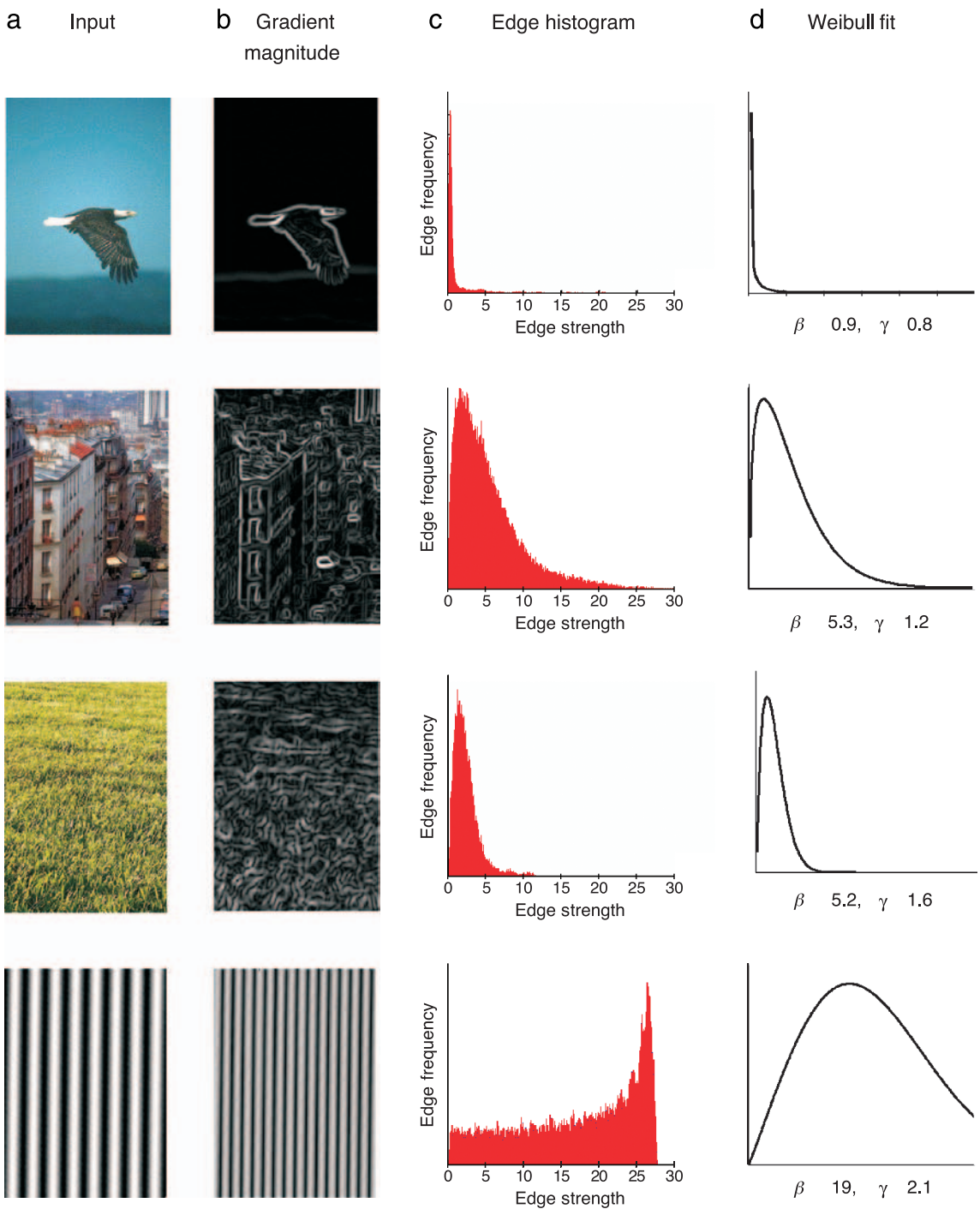


Figure 1. (a) Three distinct natural images with profoundly different visual appearance as reflected in their beta and gamma values and an example of an artificial image. The images have contrast distributions that deform from power law (eagle, first row), through Weibull (street, middle row) to Gaussian (grass, third row; but note that power law and Gaussian are special instances of the Weibull distribution). The artificial image (fourth row) has a contrast distribution that deviates from the Weibull distribution. (b) The beta and gamma values are estimated from the gradient magnitudes of each of these images. The gradient images are derived by filtering the images twice with differently oriented (90° difference) first-order directional Gaussian derivative filters and combining in the following way: $\sqrt{(dx^2 + dy^2)}$. (c) An edge histogram is generated on the basis of the gradient magnitude image. (d) A Weibull function is fitted to the contrast histogram (maximum likelihood) to approximate the beta and gamma parameters with a maximum likelihood fit.

(4.2° × 6.4°) and a bit depth of 24. Of these, 400 were pictures of animals in their natural habitat, and 400 were pictures of natural landscapes. These images were taken from Fabre-Thorpe, Delorme, Marlot, and Thorpe (2001). The other 800 images consisted of 400 indoor and 400 outdoor scenes that were completely uncontrolled with respect to factors such as illumination, occlusion, motion, and pose. These images are part of the National Institute of Standards and Technology news video corpus (Over, Leung, Ip, & Grubinger, 2004). One of the images was not used because its marker was lost during EEG acquisition.

Determining the beta and gamma parameters of the fitted Weibull function

The 1599 color images were converted to grayscale images $I(x, y)$ and subsequently convolved (denoted with ⊗) with first-order directional Gaussian derivative filters. These filters were designed to respond maximally to edges running vertically and horizontally relative to the image grid, one filter for each of the two perpendicular directions x and y. The filters were applied separately to the input images, producing separate measurements of the gradient component in each direction. The directional image derivatives were combined in the following way to obtain a gradient magnitude per image:

$$\nabla^1 I(x,y;\sigma) = \sqrt{\left(I(x,y)\otimes\frac{\partial G(x,y;\sigma)}{\partial x}\right)^2+\left(I(x,y)\otimes\frac{\partial G(x,y;\sigma)}{\partial y}\right)^2}, \quad (2)$$

where $G(x, y)$ is the two-dimensional Gaussian distribution function

$$G(x,y;\sigma)=\frac{1}{2\pi\sigma^2}e^{\frac{-(x^2+y^2)}{2\sigma^2}}. \quad (3)$$

The standard deviation of the Gaussian filters was set to 2 pixels (0.03°) for the estimation of the beta parameter and 5 pixels (0.075°) for the estimation of the gamma parameter.

Each of the 1599 gradient magnitude images was subsequently represented in a 256-bin histogram, summarizing the contrast distribution of the image. For each image, the parameters of the underlying Weibull distribution were estimated from the histogram. A three-parameter Weibull distribution was used:

$$p(f)=c\ e^{\left(\frac{f-\mu}{\beta}\right)^{\gamma}}, \quad (4)$$

where parameters μ, β (beta), and γ (gamma) represent the origin, scale, and shape of the distribution, respectively, and c is a normalization constant. We estimated μ and normalized it out to achieve illumination invariance, leaving only parameters beta and gamma.

The beta and gamma parameters were estimated using a maximum likelihood estimator (MLE) via the histogram to reduce computational load. Note that parameter estimation is only marginally sensitive to the effects of histogram quantization.

Estimating X- and Y-cell outputs

To estimate X-cell outputs, we convolved each grayscale image with a difference of Gaussian (DoG) filter. The DoG filter was defined by two Gaussian functions with standard deviations sigma = 1.2 and sigma = 7.2, yielding a spatial profile (center sigma = 0.03°, surround sigma = 0.18°) that corresponds to that of parvocellular X-cells for the primate visual system (Hubel & Wiesel, 1962). The convolution of the grayscale image $I(x, y)$, with this DoG filter resulted in the filtered images

$$\nabla^2 I(x,y;\sigma_1,\sigma_2)=I(x,y)\otimes(G(x,y;\sigma_1)-G(x,y;\sigma_2)). \quad (5)$$

The filtered images were rectified and fed into a contrast sensitivity profile (Out = C / (1.038 + C)) corresponding to that of parvocellular X-cells (Croner & Kaplan, 1995). Summation of the resulting intensity values yielded the "X-output."

To calculate the Y-output, we used a computation that is often used in models of the Y-system, where a signal from linear subunits is rectified and smoothed over space (Enrothcugell & Freeman, 1987). The size of the linear difference of Gaussian subunits was identical to that of the X-system specified above. Outputs were rectified and smoothed with a Gaussian window with a size (sigma = 0.1°) that corresponds to the receptive field size of magnocellular Y-cells of the primate visual system (Croner & Kaplan, 1995). The result was subsequently fed into a contrast sensitivity profile (Out = In / (0.169 + In)) corresponding to that of magnocellular cells. The "Y-output" was taken as the summed value of all filter outputs.

Visual stimulation, ERP data acquisition, and data analysis

Stimuli were presented on a 19″ Ilyama monitor with a resolution of 1024 * 768 pixels and a frame rate of 100 Hz. Subjects were seated 90 cm from the monitor. During EEG acquisition, a stimulus was presented on

average every 1500 ms (range 1000–2000 ms) for 100 ms. Each stimulus was presented 2 times for a total of 3200 presentations per subject. Stimuli from the first data set (scenes with and without animals) were shown in alternating blocks with images from the second data set (indoor and outdoor scenes). Subjects had to indicate whether (a) the images contained an animal or not or whether (b) the image depicted an indoor or outdoor scene, depending on which block was being presented. Recordings were made with a Biosemi 52-channel Active Two EEG system (Biosemi Instrumentation BV, Amsterdam, The Netherlands). Data were sampled at 256 Hz. Data analysis was identical to Scholte, Witteveen, Spekreijse, and Lamme (2006) with the exception that the high-pass filter was placed at 0.1 Hz (12 db/octave) and the pre-stimulus baseline activity was taken between −100 and 0 ms with regard to stimulus onset. Trials were averaged over subjects per individual stimulus resulting in 1599 averages (one marker was lost in recording), consisting of 20 to 32 averages per individual image (median of 30 averages per image) since some ERP trials were rejected on the basis of amplitude (when amplitude exceeded 75 or −75 μV) or gradient (a larger voltage step than 50 μV/sample). The average responses (ERPs) were converted to a Current Source Density (CSD) response. The CSD conversion gives a signal that is more localized in space than a regular ERP and therefore has the advantage of reflecting more reliably the activity of neural tissue directly underlying the recording electrode (Nunez & Srinivasan, 2006).

To test whether the evoked EEG responses of human subjects viewing natural images correlate with the beta and gamma values of the Weibull distribution of the images, the 1599 averaged responses per image were covaried with the beta and gamma values of the images as estimated from their contrast histograms.

We also correlated the EEG responses with other frequently used image statistical parameters. The maximum, mean, and median contrasts were computed from the contrast histograms. Gaussian and logarithmic distributions were fit to the contrast histograms, and their parameter values were covaried with the ERPs. The ERPs were additionally covaried with the spatial frequency characteristics of each image, computed from the power spectrum of its largest concentric square portion: the analysis involved transformation of the cropped image into the frequency domain using Fast Fourier Transform, derivation of the Fourier power spectrum, and rotational averaging of this power spectrum over orientation. On log–log paper, the average power spectrum as a function of spatial frequency approximated a line. A line was fit to the one-dimensional summary of the power spectrum using least squares and estimated the intercept and slope. These two parameters were covaried with the ERP data.

Finally, to compare the contrast values represented by the Weibull distribution with other measures of contrast, we calculated two standard contrast measures for each grayscale image: Michelson and Root Mean Square (RMS). Michelson contrast was based on the highest and lowest intensity values in the image, (Lmax − Lmin) / (Lmax + Lmin). The RMS contrast was taken to be the standard deviation of the intensity values of all pixels in the image divided by the mean intensity.

Modeling the maximum amount of explainable data

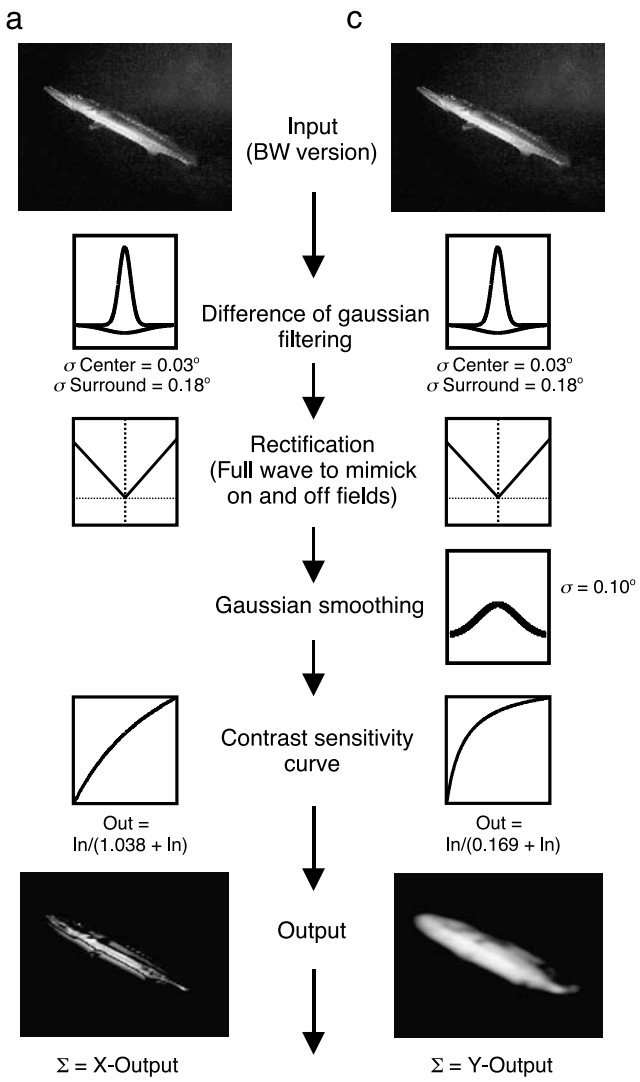
The EEG data are relatively noisy given the fact that each of the 1599 used images was presented only twice to each subject, in addition to the individual differences between subjects themselves. To address this issue, the maximum amount of explainable data (signal) was calculated per subject, and then averaged over all subjects. A subspace model was used to determine the maximum amount of explainable variance in the EEG data, independent of beta and gamma. See Appendix A for a complete description of this method.

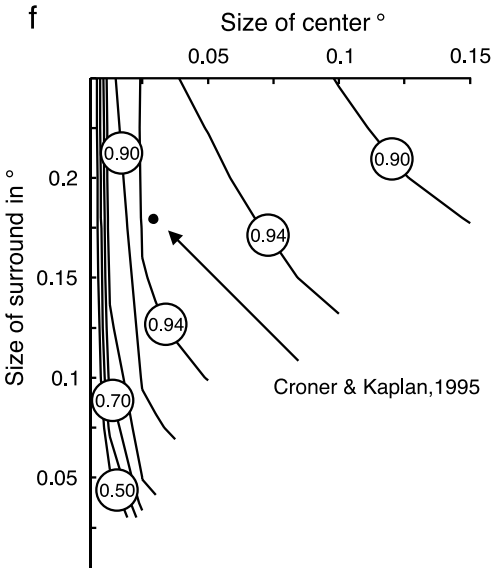
Results

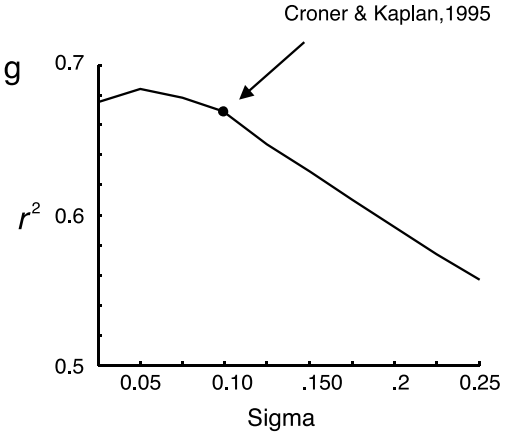
Neural computation of Weibull parameters

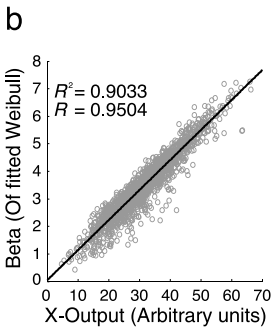
We are interested in testing our neural model for the brain, which approximates the beta and gamma values of the Weibull distribution of an image. Originally, the parameters, beta and gamma, of the fitted Weibull function to the contrast distribution were determined by maximum likelihood estimation (MLE, see Methods section) for each of the 1599 images. This mathematical method is biologically hardly plausible (see Figure 1). Here, we show how the beta and gamma parameters underlying the contrast distribution of a natural image can be closely approximated using biologically plausible filters (Figure 2). We tested the accuracy of approximating beta using the X-output of each image by first calculating beta using Weibull curve fitting, and then calculating X-output, for 1599 natural images (for image databases, see Methods section). The correlation between the X-output and the actual value of beta, estimated from the contrast distribution of each individual image by fitting a Weibull distribution, was remarkably high (r = 0.95; $p(1598)$ = 3.69e − 237; see Figure 2b).

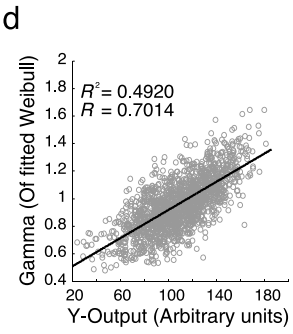
To arrive at a biologically plausible estimate for gamma, we tested the accuracy of approximating gamma using the Y-output of each image in the same way as the X-output. The value of the Y-output correlated strongly with the actual value of gamma estimated for each image

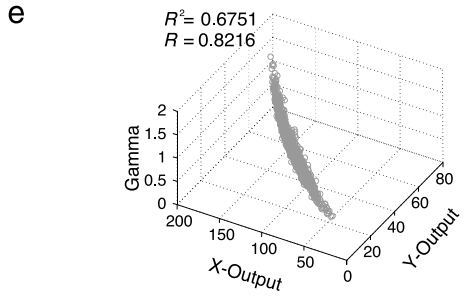

a
c
Input
(BW version)
Difference of gaussian
filtering
σ Center = 0.03°
σ Surround = 0.18°
σ Center = 0.03°
σ Surround = 0.18°
Rectification
(Full wave to mimick
on and off fields)
Gaussian smoothing
σ = 0.10°
Contrast sensitivity
curve
Out =
In/(1.038 + In)
Out =
In/(0.169 + In)
Output
Σ = X-Output
Σ = Y-Output


f
Size of center °
0.05
0.1
0.15
Size of surround in °
0.2
0.15
0.1
0.05
0.90
0.90
0.94
0.94
0.70
0.50
Croner & Kaplan,1995


Croner & Kaplan,1995
g
0.7
r^2
0.6
0.5
0.05
0.10
.150
.2
0.25
Sigma


b
Beta (Of fitted Weibull)
R^2 = 0.9033
R = 0.9504
X-Output (Arbitrary units)


d
Gamma (Of fitted Weibull)
R^2 = 0.4920
R = 0.7014
Y-Output (Arbitrary units)


e
R^2 = 0.6751
R = 0.8216
Gamma
X-Output
Y-Output

individually (r = 0.70; p(1598) = 3.69e − 237; see Figure 2d). An even better approximation of gamma could be obtained by combining the outputs of both the X- and Y-systems. Figure 2e shows how gamma is defined by a combination of X- and Y-outputs. Stepwise regression of gamma versus X- and Y-outputs yielded a combined correlation of r(1597) = 0.82.

We used biologically plausible receptive field sizes of X- and Y-type filters to estimate beta and gamma. However, it could be that these are not optimally suited for this goal, which would weaken our claim that beta and gamma are relevant to the brain. To examine to what degree the response properties of X- and Y-cells are optimal in extracting the beta and gamma parameters, we calculated the correlation between X-output and beta for a wide range of potential center and surround values of the first stage of filtering. Figure 2f shows a series of contour plots of the landscape indicating the correlation for each combination of center and surround sizes. Results indicate that the physiological center and surround sizes of the X-cells as reported in Croner and Kaplan (1995); and that we used) fall within the optimal range, but also that this optimal space is relatively wide (the area between the two r = 0.94 contour lines in Figure 2f). We examined the same for the size of the Gaussian smoothing, suggested here to be executed by the Y-cells (see Figure 2g). Results indicate that the physiological sizes of the Y-cells are close to the optimum needed to estimate gamma.

Our results show how the brain may calculate good approximations to the beta and gamma parameters underlying an image's contrast distribution on the basis of filters that are biologically realistic in shape, sensitivity, and size. Our results also show that a wider range of filters are capable of doing this, but that the receptive field sizes reported for the macaque visual system are close to optimally suited for this goal.

Figure 2. Model to calculate X- and Y-cell outputs for the stimuli used in this study. The images are converted from color to gray scale, filtered with a difference of Gaussians (DoG) then rectified. (a) To simulate the X-cells, the pixel values of the resulting images are transformed with a contrast gain function adequate for P-cells. Finally, pixel values are summated to yield the X-estimates. (b) The output of the X-cells correlates strongly (r = 0.95, p(1598) = 3.69e − 237) with beta. (c) To simulate the Y-cells, the rectified images are passed through a Gaussian smoothing function and pixel values are subsequently transformed with a contrast gain function adequate for M-cells. The resulting values are summated to yield the Y-estimate. (d) The output of the Y-cells correlates strongly (r = 0.70, p(1598) = 3.69e − 237) with gamma. (e) A combination of X- and Y-outputs correlates even better with gamma (r(1597) = 0.82). (f) Overview of the correlation of different combinations of different center and surround sizes for the estimation of the beta parameter via neurons. (g) Overview of different sizes of Gaussian smoothing in combination with the output of simulated Y-cells to estimate gamma.

Neural sensitivity to Weibull parameters

That the brain is able to compute the contrast distribution parameters does not necessarily imply that it actually performs this computation. To test whether the brain may compute these parameters, we presented the 1600 images used in our calculations to human subjects (n = 16) while EEG was recorded and calculated a grand-average per image. Next, the Weibull parameters, estimated from the individual stimuli, were correlated with the ERP responses.

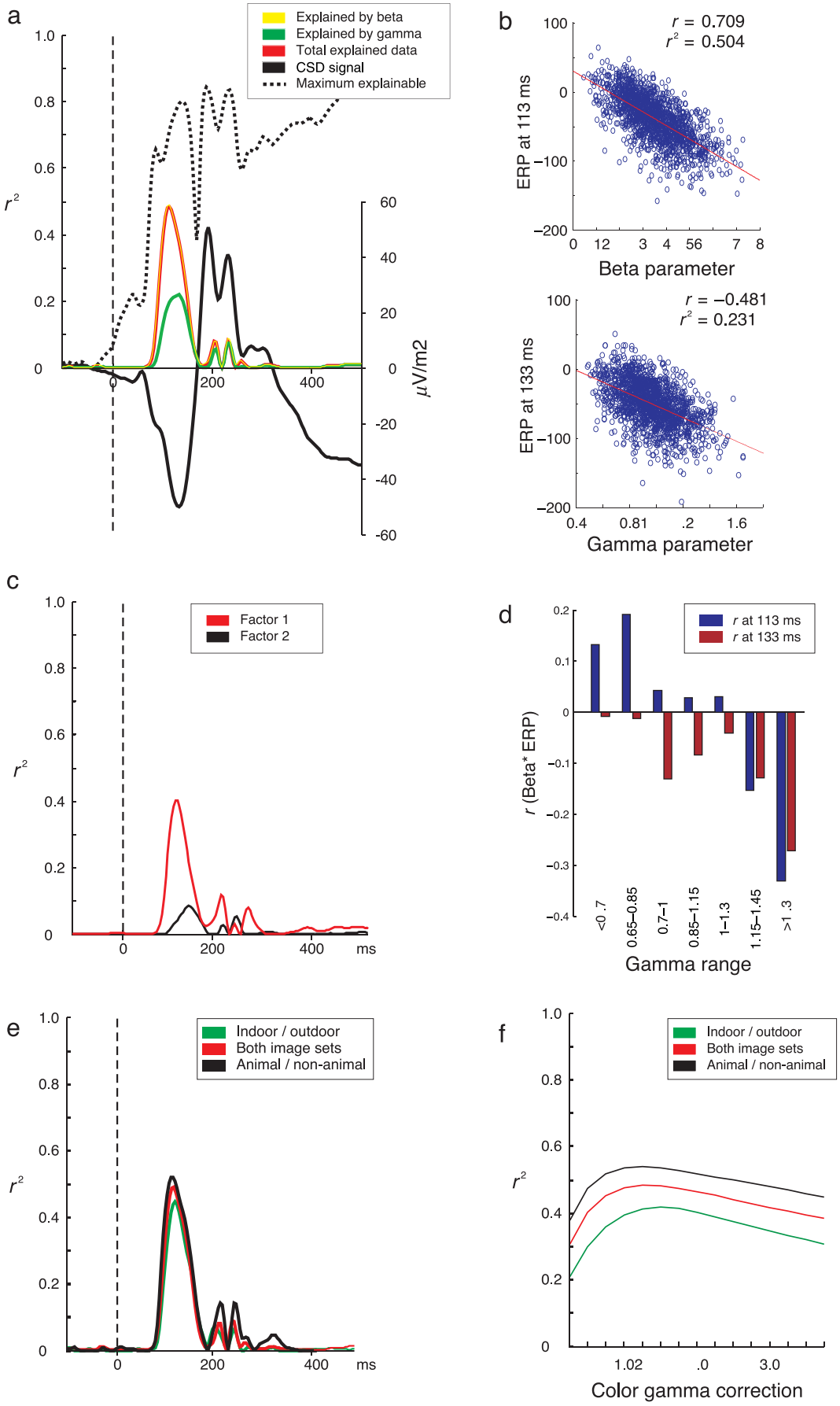
The results of this correlation show that beta and gamma explain up to 50% of the variance of (i.e., correlate up to r = −0.71 with) the measured ERP signal recorded between 80 and 200 ms at electrode Iz, which is overlying the early visual cortex. Beta explained the most variance at 113 ms after stimulus onset (r = −0.71; p(1598) = 3.0e − 245; r^2 = 0.504) while gamma explained most variance at 133 ms (r = −0.481; p(1598) = 3.6e − 234; r^2 = 0.231). Figure 3a shows the correlations of beta and gamma with the ERP signal over time. Figure 3b shows scatter plots of these correlations.

The variance that is explained by both beta and gamma (blue line in Figure 3a) is only slightly larger than the variance explained by beta alone (red line). Formal testing reveals that gamma explains a significant, albeit little, amount of additional activity at 113 ms (for beta F(1, 1596) = 910,289, p = 1.3e − 158, for gamma F(1, 1596) = 10.9, p = 0.001, r^2 = 0.51) and 133 ms (for beta F(1, 1596) = 419.8, p = 7.5e − 083 and gamma F(1, 1596) = 7.6, p = 0.006, r^2 = 0.41).

This is partially due to the fact that, in natural images, beta and gamma correlate significantly (see Figure 5 and below), which reduces the amount of variance that gamma can explain independent of beta.

We performed a factor analysis (covariance matrix, Varimax rotation) on the beta and gamma parameters calculated from each image to see to what degrees they constitute 2 factors. The first factor explained 83% of the variance in beta and gamma and correlates mainly with beta (r = 0.928), the second factor explained 16% of the variance and correlated mainly with gamma (r = 0.928). We subsequently correlated these factors with the EEG signal. This revealed that the "beta"-like factor peaks at 121 ms, in channel Iz (beta like F(1, 1596) = 1504.891, p = 1.9e − 232, gamma like F(1, 1596) = 2.7e − 31, p = 1.1e − 33, r^2 = 0.507) and that the "gamma"-like factor peaks at 145 ms in channel Oz (beta like F(1, 1596) = 1002.726, p = 3.5e − 171, gamma like F(1, 1596) = 214.611, p = 1.1e − 45, r^2 = 0.41). This confirms that beta and gamma independently explain a substantial amount of variance (see Figure 3c).

Furthermore, it appears that there is additional explained variance in the ERP signal when we consider non-linear interactions between beta and gamma. For example, the strength and sign of the correlation of beta with the ERP depends on the value of gamma (Figure 3d).


a
Explained by beta
Explained by gamma
Total explained data
CSD signal
Maximum explainable
r^2
μV/m2
b
r = 0.709
r^2 = 0.504
ERP at 113 ms
Beta parameter
r = –0.481
r^2 = 0.231
ERP at 133 ms
Gamma parameter
c
Factor 1
Factor 2
ms
d
r at 113 ms
r at 133 ms
r (Beta*ERP)
<0.7
0.65–0.85
0.7–1
0.85–1.15
1–1.3
1.15–1.45
>1.3
Gamma range
e
Indoor / outdoor
Both image sets
Animal / non-animal
f
Color gamma correction

The brain could utilize these non-linear interactions by interpreting the beta values differently for different values of gamma.

Several other features of the results are remarkable. ERP responses toward individual stimuli consist of just one or two per subject. This will have induced a large amount of variability between the ERPs to each image that cannot be explained by stimulus variations but instead should be attributed to inter-individual differences, ongoing brain processes, and other sources of "noise." Therefore, there is a theoretical upper bound to the amount of ERP variability that can be explained by the variations in stimulus content, i.e., by beta and gamma. We calculated this upper bound (see Appendix A) and it is given in Figure 3a (dotted line). This is an important observation as the maximum explainable signal is neither small nor constant over time. Compared to this upper bound, beta and gamma explained 71% of the ERP variance recorded at 113 ms at electrode Iz. In other words, the two Weibull parameters together account for almost the entire stimulus-driven variance in the early ERP signal.

Where the ERP has several peaks and troughs well into the 500-ms epoch, which is typical of visual ERPs to meaningful stimuli, beta and gamma only explain the earliest components of this response up to about 200 ms. This suggests that beta and gamma are probably only relevant to the brain in the early phases of visual processing. Later on, other (probably higher level) features of the image take over.

Figure 3. (a) Data explained in channel Iz by beta (red), gamma (green), and both (yellow) parameters. The dotted lines show the maximum percentage of the data that can be explained. Notice that the ERP (black) continues to show a signal after beta and gamma stop explaining data, which is also shown by the dotted line. (b) Scatter plot of the beta parameter and the ERP at the moment of maximum correlation with beta (113 ms) and the maximum correlation with gamma (133 ms). (c) Explained ERP variance in channel Iz by first and second factors obtained from factor analysis. The first factor behaves similarly as (and is strongly correlated with) beta. The second factor behaves similarly as (and is strongly correlated with) gamma. (d) Explained ERP data for the beta parameter within a range of gamma values. Notice that the correlation of the ERP with beta is positive for low values of gamma (for instance the correlation of beta with ERP for images with gamma values lower than 0.7) and negative for high values of gamma (for instance higher than 1.3) at 113 ms. This pattern changes at 133 ms, where beta does not correlate with the ERP for low values of gamma and correlates negatively with the ERP with high values of gamma. (e) Explained ERP variance in channel Iz for the two subsets of images used. The explained variance for the animal/non-animal data is higher than the explained variance for indoor/outdoor scenes. (f) Total explained variance in the ERP for channel Iz with different gamma correction values applied to the images. Notice that even for very extreme gamma corrections the explained variance remains rather high.

Our data set consisted of 2 relatively different image sets. The first set consisted of images of natural scenes without motion artifacts. The second set consisted of images that were entirely captured from television. Both sets show very similar correlations of ERPs with beta and gamma (see Figure 3e). The explained ERP variance with relation to the animal/non-animal images is slightly larger (7.5%) at their maximum than the explained variance with relation to the television-captured images. This minor quantitative difference is most likely due to the "noise" edges present in the second data set, a result of the capture process. In addition, the two data sets may have been captured using different color gamma encoding (to increase dynamic range). In our experiments, we did not explicitly correct for these, potentially creating artificial differences between the contrast distributions of the two data sets. We applied color gamma correction (which influenced the color encoding, not related to the gamma value from the Weibull function) to each of the 1599 images with correction values ranging from 0.25 to 4 to assess to what extent the differences between real-world, modeled, and displayed luminance values affect our results. For the animal/non-animal data set, the maximum of the explained variance peaks when the correction value is 1.25, whereas for the indoor/outdoor data, a correction value of 1.5 yields the highest explained variance. The peak of the maximum explained variances near correction value 1 (no correction) and the shallow decay with correction values larger than 1 suggest that our data sets have a natural appearance and that our results are robust for effects of luminance compression and expansion.

Beta versus other contrast values

There are many other measures of contrast formulated over the years. If indeed the Weibull parameters are relevant to the brain, the ERP should correlate more strongly with beta than with any other parameters of the contrast distribution. Beta reflects a specific summary of the distribution of local contrast values. A Weibull distribution with high beta will have its median shifted to the right, i.e., have a greater number of high contrast values. Likewise, from its proposed neural approximation (see Neural computation of Weibull parameters section), it follows that beta reflects the presence of many high contrast edges (see also Figure 1).

In Figure 4a and Table 1, we explore the correlation of beta with the ERP at all electrodes and at all time samples. The explorations of parameters of other distributions that could be fitted to the contrast values are also shown in Figure 4a: a half-sided Gaussian (free parameter: sigma) and a simple exponential (free parameter: mean). These functions are special cases of the Weibull distribution (with gamma fixed at 1 and 2, respectively) and could thus provide a simpler one-parameter model for the brain to summarize the contrast distribution. Figure 4b shows

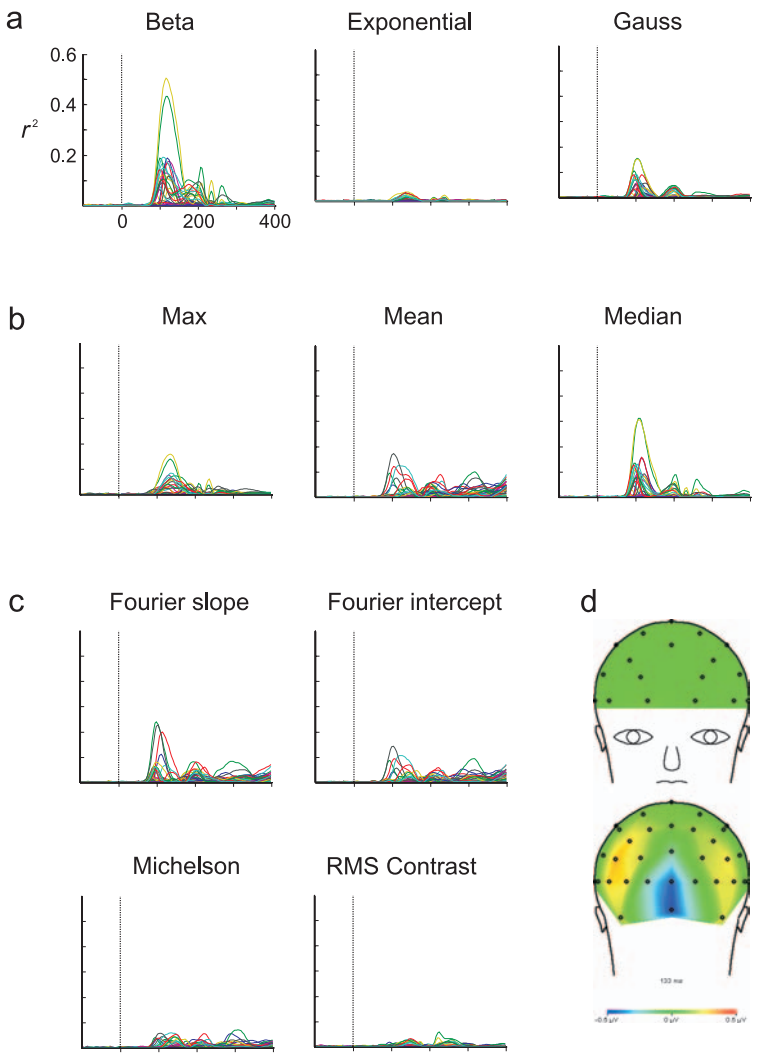


Figure 4. Explained variance for all channels displayed over time in the same plot for the (a) Weibull beta parameter, the exponential and Gaussian distribution, (b) the maximum, mean, and median of the edge histogram, and (c) the Fourier slope, the Fourier intercept, the Michelson contrast, and Root Mean Square contrast. Notice that beta outperforms all other measures significantly. (d) Spatial distribution of beta explained variance at t = 113 ms.

correlations between ERP signals and other simple parameters of the contrast distribution: the mean, median, and maximum. Finally, in Figure 4c, we show correlations between ERP and traditional contrast measurements like RMS contrast, Michelson contrast, Fourier intercept, and Fourier slope. The moment in time and channel of the maximum amount of variance each of these correlations explain is given in Table 1. It is obvious from these results that neither of the alternative parameters explains as much variance of the ERP as beta does, by far. In sum, all evidence leads us to conclude that the brain is particularly sensitive to the parameters of the Weibull distribution underlying the image it perceives.

Finally, we propose that calculating beta and gamma opens the possibility to reliably balance natural image sets. Apparently, the beta and gamma values of images provide an excellent model of the way low-level visual

	r^2	ms	Channel
Beta	0.50	113	Iz
Gamma	0.23	129	Iz
Exp	0.04	133	Iz
Gaussian	0.15	102	Oz
Max	0.16	133	Iz
Median	0.31	109	Oz
Mean	0.17	102	POz
Michelson	0.07	305	Oz
RMS	0.06	223	Oz
Slope	0.24	98	Oz
Intercept	0.14	102	POz

Table 1. The maximum explained variance of different statistics from the contrast histogram, Fourier intercept and slope, and some contrast measures, in combination with the channel and time at which this maximum occurs.

areas respond to natural images. When sets of images are balanced with respect to beta and gamma, one can be very certain that low-level features and early visual cortical responses are reasonably balanced as well, which is important when studying higher level aspects of visual processing.

Discussion

Origin of the Weibull parameters sensitivity

We have shown that the brain is capable of approximating the beta and gamma underlying the contrast distribution in natural images. This can be achieved in a biologically plausible way, by filtering the image with a combination of linear and non-linear receptive fields. We used published sizes of difference of Gaussians receptive field profiles and published contrast sensitivities of primate LGN cells (Croner & Kaplan, 1995) to approximate beta and gamma of a set of 1599 natural images. Given these restrictions, we obtained remarkably high correlations between our estimates and the actual values of beta and gamma of each image. Next, we showed that beta and gamma of each image explain up to 71% of the stimulus-driven variance of ERPs evoked in human subjects viewing natural images. Furthermore, the Weibull parameters explain more variance of the ERP than any other measurement of contrast or any other parameter fitted to the contrast histogram. These findings together make it a distinct possibility that the brain actually uses the Weibull parameters in the evaluation of images. This begs the question why this would be the case.

There are several reasons that the contrast histogram of natural images can be approximated by a Weibull function. The first comes from the tendency of nature to increase entropy so that small details are occurring more often in an image than large structures (Koenderink, 1984). Causality furthermore implies that large structures are made up of many smaller ones, a property of natural objects that occurs at a wide range of resolving powers (Mandelbrot, 1983). An image of a single, sufficiently complex, object against a highly uniform background, as the picture of the eagle in Figure 1, will therefore have a power law distribution of contrasts (many zero contrasts, few high contrasts). Viewing such an object with fixed sets of finite receptive fields will yield a Weibull contrast distribution, in this case with a low value of gamma. When many objects or object parts are present (some large, many small), the distribution will arise from summing over correlated contrast values, which yields a Weibull distribution with a high value of gamma (as in the picture of grass in Figure 1). Here, very low contrast values are rare and the peak number of contrasts has shifted toward intermediate levels of contrast. Under constant lighting conditions, and at a fixed power of resolution, that is with the same receptive field sizes, going from one to many objects will increase the value of beta (Geusebroek & Smeulders, 2005).

Are the Weibull parameters ecologically relevant?

From their dependence on the number of objects in the image, a potentially powerful ecological relevance of the Weibull parameters arises. It implies that a difference in object distribution or difference in perceptual grouping results in a different shape of the Weibull distribution and, hence, in different values of beta and gamma. This is all the more apparent when we look at the beta and gamma parameters of our set of natural images. In Figure 5a, we plot a thumbnail of each image at its respective beta and gamma values, of a random subset of our 1599 images (for the sake of clarity; see Supplementary Figure 1 for a high-resolution complete version of this image space).

Two observations can be made. The first is that gamma and beta are not fully independent yet correlate in the sense that low gamma values coincide with low beta values and high gamma with high beta ($r = 0.71$, for all 1599 images). This strong correlation is not a statistical necessity, but a property of natural images, borne out from the originating mechanism of the contrast distribution discussed above. Scrambled versions of the same images yield a substantial lower correlation ($r = 0.33$). So, the correlation between beta and gamma is the result of spatial coherence.

Second, it appears that images in beta–gamma space go from isolated objects in the lower left to scenes with multiple objects, or almost texture-like scenes, with many similar objects, in the upper right corner. This axis seems to represent image complexity. We estimated perceptual image complexity by converting the images to JPEG format. High compressibility (resulting in small JPEG file sizes) typically occurs when scenes are simple, while complex scenes (or very noisy ones) are typically difficult to compress (resulting in high JPEG file sizes). Figure 5b shows the same as Figure 5a, i.e., beta and gamma on the orthogonal axes, with the difference that we now plot the JPEG file size in color code for all images. High beta (and gamma) correlate with high JPEG file size, i.e., high complexity of the scene (stepwise regression of JPEG file size vs. Beta and Gamma yields $r = 0.828$; Beta only vs. JPEG size yields $r = 0.819$).

Orthogonal to this "complexity" axis, the images seem to separate along the dimension of texture similarity, in the sense that images with similar textures are found in the lower right, while images with substantial difference in texture (resulting in more perceived depth) are found in the upper left side.

The full potential of sorting images according to beta and gamma of their contrast distribution remains to be

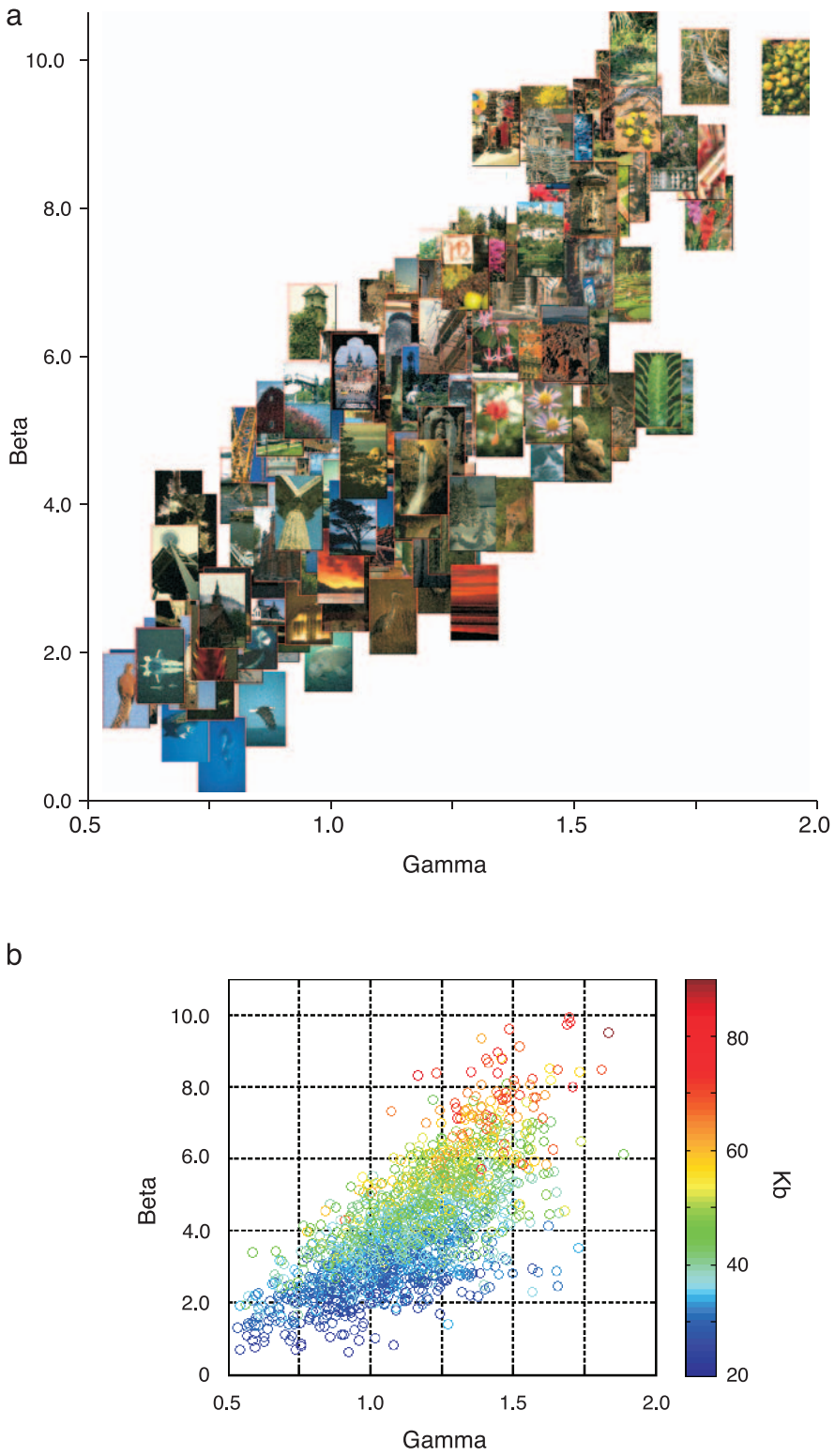


Figure 5. (a) A mapping of a subset of our 1599 images onto the beta–gamma space. An image's position in this space is fixed by beta and gamma values. The axis going from bottom left to top right shows an increase in image cluttering and a decrease in figure-ground segmentation. The axis going from top left to bottom right shows a change from images with dissimilar textures to images with similar textures (see Supplementary Figure 1 for a version of this space with all animal/non-animal images). (b) The values of beta and gamma with their corresponding JPEG compression factor. Both beta and gamma capture different parts of the complexity of an image as encoded by the JPEG algorithm.

resolved. Previously, it was shown that in the case of natural homogeneous yet stochastic texture (e.g., grains of rice), beta and gamma inform the viewer about aspects such as texture coarseness and roughness, and illumination intensity and angle (Geusebroek & Smeulders, 2002, 2005). We have shown that beta and gamma are accurate descriptions of the clutter, spatial coherence, and complexity of an image. Furthermore, the present research suggests (Figure 5) that other ecologically relevant categorization on the basis of beta and gamma will be possible, for example, in scene segmentation, determining to what extent objects are close or far away, and even in the detection of (elements of) the perceptual gist. In this way, the brain could efficiently shift between modes, and the scale of analysis required for rapid visual processing (Oliva & Schyns, 1997; Schyns & Oliva, 1994). Given that the brain is highly sensitive to the Weibull parameters, further investigation into the computational potential of exploiting this omnipresent natural image statistic seems warranted.

Appendix A

Modeling the maximum explainable data

In order to see what could maximally be explained for each sample in the data for all subjects simultaneously, an independent method from the Weibull model to estimate the *R*-squared value is required. The algorithm we used is a variation on the subspace methods used in, for example, Stoica and Nehorai (1989). Such algorithms assume a model for the data in which the signal and noise are uncorrelated and the noise is uncorrelated in time. They then use an eigenvalue decomposition of the second-order moment matrix to estimate the signal and noise space. We can use this decomposition to project the data into the noise space, thus creating residuals based on a minimal model. These residuals are then transformed to *R*-squared values to be compared with the *R*-squared values obtained from the data.

Three problems of this method, which are particular to our situation, need to be addressed:

i. The noise is correlated in time,
ii. the amount of signal in the data should not be underestimated, and
iii. multiple subjects should be combined.

Splitting up the data into two parts of single trials solves the first problem: one part is used to estimate the noise correlations. Then, these correlations from the first part are removed from the second part of the data. The second problem is more difficult but equally important. If we are to indicate the maximum of signal to be explained in terms of *R*-squared, then we must never underestimate the amount of signal in the data, so that our maximum would end up lower than it should. We have used a hypothesis testing method applied to the eigenvalues of the second-order moment matrix to deal with this issue. To tackle the third problem, we chose to detect the noise space for each subject separately because there are individual differences. Then, the *R*-squared values for all 16 subjects were averaged.

The method

Let y denote the vector containing the t time samples, A be a gain matrix of the signal space, and Σ be the covariance matrix of the noise with scaling factor, σ^2. Then, the population model for the second-order moment matrix containing all time samples is the t by t matrix

$$E\{yy'\} = AA' + \sigma^2\Sigma. \tag{A1}$$

For our purpose, it is irrelevant what A exactly is, only the assumption that A is uncorrelated to the noise is important. From this second-order moment matrix, the noise subspace can be determined by considering the number of multiplicities of eigenvalues. If the noise is uncorrelated, then the eigenvalues equal to the scaling factor σ^2 of the noise correspond to the noise subspace. For this to work, we need to make Σ the identity matrix (with ones on the diagonal and zeros elsewhere). If Σ were known, then we would simply use the inverse of Σ and decorrelate

$$E\{yy'\}\Sigma^{-1} = AA'\Sigma^{-1} + \sigma^2 I. \tag{A2}$$

A very good estimate of Σ can be obtained from the data because we have many trials available. By splitting up the data in two sets of trials of size t_1 and t_2, we can use the first set to estimate Σ. An unbiased estimate of Σ is

$$S_e = \frac{1}{t_1 - 1}\sum_{j=1}^{t_1}\left(y_j - \bar{y}\right)\left(y_j - \bar{y}\right)', \tag{A3}$$

where $\bar{y}$ is the average over trials. On the sample level, we now have $S_y S_e^{-1}$, where

$$S_y = \frac{1}{t_2 - 1}\sum_{j=1}^{t_2} y_j y_j', \tag{A4}$$

with the data y from the second batch of trials. Next, we compute the eigenvalues $l_1, l_2, \ldots, l_n$ of $S_y S_e^{-1}$ to determine the noise space by considering which of these eigenvalues are equal to the noise scaling factor, σ^2. If we use a sufficiently large number t_1 of trials for S_e, then we can treat this matrix as fixed and assume that the distribution of the eigenvalues of $S_y S_e^{-1}$ is the same as the distribution of the eigenvalues of S_y. We used 65% of

the total number of trials for estimation of S_e and 35% for estimation of S_y. With 1599 trials in this experiment, this amounts to $t_1 = 1039$ and $t_2 = 560$. We can sequentially test for k eigenvalues of the signal space, or equivalently, we can test for $q = n - k$ eigenvalues equal to the noise scaling factor. We begin with $q = 0$ (all noise) and continue to test the null hypothesis H_q: $l_{k+1} = l_{k+2} = \ldots = l_q = \sigma^2$ until the test says that there is no difference between the eigenvalues anymore. At this point, we have found that the noise space can be constructed by q eigenvectors. We use the Bartlett corrected test of the likelihood ratio test for equality of likelihood (Muirhead, 1982):

$$-\left(t_2 - k - \frac{(q+2)(q-1)}{6q}\right)\ln\left(\frac{\prod_{i=k+1}^{n} l_i}{\frac{1}{q}\sum_{i=k+1}^{n} l_l}\right). \quad \text{(A5)}$$

The null distribution is $\chi^2_{(q+2)(q-1)/2}$ with $(q+2)(q-1)/2$ degrees of freedom. In order to cover the signal space as best as possible (so as not to underestimate the maximum value in R-squared), we used a significance level of 0.05 for each test.

Acknowledgments

We would like to thank Michèle Fabre-Thorpe for providing us with the animal and non-animal stimuli and Nikolaus Kriegeskorte for helpful comments and suggestions.

Commercial relationships: none.
Corresponding author: H. Steven Scholte.
Email: h.s.scholte@uva.nl.
Address: Roeterstraat 15, A625 1018 WB, Amsterdam, The Netherlands.

References

Brumswik, E., & Kamiya, J. (1953). Ecological cue-validity of proximity and of other Gestalt factors. *American Journal of Psychology, 66,* 20–32. [PubMed]

Croner, L. J., & Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Research, 35,* 7–24. [PubMed]

Daugman, J. G. (1989). Entropy reduction and decorrelation in visual coding by oriented neural receptive-fields. *IEEE Transactions on Biomedical Engineering, 36,* 107–114. [PubMed]

Enrothcugell, C., & Freeman, A. W. (1987). The receptive-field spatial structure of cat retinal Y cells. *The Journal of Physiology, 384,* 49–79.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *Journal of Cognitive Neuroscience, 13,* 171–180. [PubMed]

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America A, Optics and Image Science, 4,* 2379–2394. [PubMed]

Fleming, R. W., & Bulthoff, H. H. (2005). Low-level image cues in the perception of translucent materials. *ACM Transactions on Applied Perception, 2,* 346–382.

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59,* 167–192. [PubMed]

Geusebroek, J. M., & Smeulders, A. W. M. (2002). *A physical explanation for natural image statistics*. Paper presented at the 2nd International Workshop on Texture Analysis and Synthesis, Heriot-Watt University.

Geusebroek, J. M., & Smeulders, A. W. M. (2005). A six-stimulus theory for stochastic texture. *International Journal of Computer Vision, 62,* 7–16.

Hsiao, W. H., & Millane, R. P. (2005). Effects of occlusion, edges, and scaling on the power spectra of natural images. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 22,* 1789–1797. [PubMed]

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of Physiology, 160,* 106–154. [PubMed] [Article]

Koenderink, J. J. (1984). The structure of images. *Biological Cybernetics, 50,* 363–370. [PubMed]

Mandelbrot, B. (1983). *The fractal geometry of nature*. New York: W. H. Freeman.

Meeker, W. Q., & Escobar, L. A. (1998). *Statistical methods for reliability data*. New York: John Wiley.

Meerschaert, M. M., & Scheffler, H. P. (2001). *Limit distributions for sums of independent random variables: Heavy tails in theory and practice*. New York: John Wiley.

Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature, 447,* 206–209. [PubMed]

Muirhead, R. J. (1982). *Aspects of multivariate statistical theory*. Hoboken, NJ: John Wiley & Sons.

Nunez, P. L., & Srinivasan, R. (2006). *The neurophysics of EEG* (2nd ed.). Oxford, UK: Oxford University Press.

Oliva, A., & Schyns, P. G. (1997). Coarse blobs or fine edges? Evidence that information diagnosticity changes the perception of complex visual stimuli. *Cognition, 34,* 72–107. [PubMed]

Over, P., Leung, C., Ip, H., & Grubinger, M. (2004). Multimedia retrieval benchmarks. *IEEE Multimedia, 11,* 80–84.

Parraga, C. A., Troscianko, T., & Tolhurst, D. J. (2000). The human visual system is optimized for processing the spatial information in natural visual images. *Current Biology, 10,* 35–38. [PubMed] [Article]

Rousselet, G. A., Pernet, C. R., Bennett, P. J., & Sekuler, A. B. (2008). Parametric study of EEG sensitivity to phase noise during face processing. *BMC Neuroscience, 9,* 98. [PubMed] [Article]

Ruderman, D. L., & Bialek, W. (1994). Statistics of natural images: Scaling in the woods. *Physical Review Letters, 73,* 814–817. [PubMed]

Scholte, H. S., Witteveen, S. C., Spekreijse, H., & Lamme, V. A. F. (2006). The influence of inattention on the neural correlates of scene segmentation. *Brain Research, 1076,* 106–115. [PubMed]

Schyns, P. G., & Oliva, A. (1994). From blobs to boundary edges: Evidence for time-scale dependent and spatial-scale dependent scene recognition. *Psychological Science, 5,* 195–200.

Simoncelli, E. P. (1999). Modeling the joint statistics of images in the wavelet domain. *Proceedings of SPIE, 3813,* 8.

Stoica, P., & Nehorai, A. (1989). MUSIC, maximum likelihood, and Cramer–Rao bound. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 37,* 720–741.

Thomson, M. G. (2001). Beats, kurtosis, and visual coding. *Network, 12,* 271–287. [PubMed]

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network, 14,* 391–412. [PubMed]

Turiel, A., Mato, G., Parga, N., & Nadal, J. P. (1998). Self-similarity properties of natural images resemble those of turbulent flows. *Physical Review Letters, 80,* 1098–1101.

Victor, J. D., Chubb, C., & Conte, M. M. (2005). Interaction of luminance and higher-order statistics in texture discrimination. *Vision Research, 45,* 311–328. [PubMed]