

Statistics of natural images as a function of dynamic range

Antoine Grimaldi

Department of Information and Communications
Technologies, Universitat Pompeu Fabra,
Barcelona, Spain



David Kane

Department of Information and Communications
Technologies, Universitat Pompeu Fabra,
Barcelona, Spain



Marcelo Bertalmío

Department of Information and Communications
Technologies, Universitat Pompeu Fabra,
Barcelona, Spain



The statistics of real world images have been extensively investigated, but in virtually all cases using only low dynamic range image databases. The few studies that have considered high dynamic range (HDR) images have performed statistical analyses categorizing images as HDR according to their creation technique, and not to the actual dynamic range of the underlying scene. In this study we demonstrate, using a recent HDR dataset of natural images, that the statistics of the image as received at the camera sensor change dramatically with dynamic range, with particularly strong correlations with dynamic range being observed for the median, standard deviation, skewness, and kurtosis, while the one over frequency relationship for the power spectrum breaks down for images with a very high dynamic range, in practice making HDR images not scale invariant. Effects are also noted in the derivative statistics, the single pixel histograms, and the Haar wavelet analysis. However, we also show that after some basic early transforms occurring within the eye (light scatter, nonlinear photoreceptor response, center-surround modulation) the statistics of the resulting images become virtually independent from the dynamic range, which would allow them to be processed more efficiently by the human visual system.

Introduction

Attneave (1954), was the first to apply principles of Shannon's information theory (Shannon, 1948) to the human visual system (HVS) by demonstrating the predictability of visual information in images for human observers. In turn, Barlow (1961) suggested that the role of early sensory neurons is to remove the

statistical regularities present in the visual signal. Two examples of this efficient representation theory are the finding by Field (1987) that the spatial frequency bandwidth and sampling of cortical cells are well matched to the spectral distribution of natural scenes such that the output of neighboring neurons will be decorrelated, while van Hateren and van der Schaaf (1998) showed that properties of the independent component filters obtained with a large set of natural images match well with the behavior of simple cells. With regard to the visual modality, the identification of statistical regularities and the concept of redundancy reduction have proven critical to the fields of vision science, image processing, and information coding (Simoncelli and Olshausen (2001)).

A large number of studies investigating the properties of natural images can be found in the literature, nearly all of which use regular, single-exposure photography. Dynamic range (DR) refers to the ratio of the maximum and minimum irradiance values in a scene or image; for instance, a uniformly illuminated scene may have a DR of less than two orders of magnitude, while a scene with direct sunlight and dark shadows may have a DR greater than seven orders of magnitude. In comparison, photography is limited by the DR of the camera sensor, which is typically less than three orders of magnitude, so a single picture of a high dynamic range (HDR) scene will suffer from the presence of underexposed and/or overexposed pixels. It was explicitly mentioned in the natural scenes databases of van der Schaaf and van Hateren (1996) and Geisler and Perry (2011) that care was taken to avoid over and under exposed pixels resulting mainly in the capture of low dynamic range (LDR) scenes. Consistent with this, we estimate that the average DR of an

Citation: Grimaldi, A., Kane, D., & Bertalmío, M. (2019). Statistics of natural images as a function of dynamic range. *Journal of Vision*, 19(2):13, 1–19, <https://doi.org/10.1167/19.2.13>.

<https://doi.org/10.1167/19.2.13>

Received March 29, 2018; published February 25, 2019

ISSN 1534-7362 Copyright 2019 The Authors



image to be just 72 in the van der Schaaf and van Hateren image database and 727 and in the Geisler and Perry image database.

The aim of HDR photography is to capture a wider range of light intensity values in any arbitrary scene, limiting as much as possible overexposed or underexposed areas in the images. A very popular strategy is to combine pictures of the same scene, taken with different exposure durations, into a single image Debevec and Malik, 1997. The resulting HDR image is a spatially varying weighted average of the single-exposure pictures. The weighting function is computed locally across the image and varies according the exact technique used, but will invariably ensure that overexposed and underexposed pixels receive a low weighting. Several studies have tackled the creation of HDR databases of natural images (e.g., Adams et al., 2016; Fairchild, 2007; Parmar, Imai, Park, & Farrell, 2008; Xiao, DiCarlo, Catrysse, & Wandell, 2002).

To date, the topic of DR and image statistics has received only very limited attention. To the best of our knowledge, just two studies have compared the statistics of low dynamic range (LDR) and HDR images. The first, by Dror, Leung, Adelson, and Willsky (2001), found that the properties of HDR images were similar in many respects to those of LDR scenes, except interestingly, that the well-established one over frequency statistic of the power spectrum of natural scenes breaks down for some images captured with HDR techniques, a point we shall return to later in this paper. The second, by Pouli, Cunningham, and Reinhard (2010), reported a much higher skewness and kurtosis for the HDR intensity distributions with respect to the LDR ones.

A major drawback of both these studies is that they compared the statistics of images captured with multi-exposure HDR techniques with the statistics of single-exposure images, but did not consider how the statistics of the images vary with the underlying DR of the captured scenes. Therefore, these works implicitly assume that an image in an HDR format, captured with a multi-exposure HDR technique, is “an HDR image,” while in fact, this image might be representing an LDR scene.¹

The motivation of this paper is to assess whether or not, by previously ignoring HDR scenarios, vision research has misrepresented or misunderstood the statistics of natural images. The contributions of the work are, first, to show that the statistics of natural images vary appreciably with the DR of the underlying scene, using for this purpose the recent Southampton-York Natural Scenes (SYNS) HDR image database created by Adams et al. (2016), and second, to highlight that early visual processes greatly reduce the impact of these variations. These results are consistent with efficient coding theory, in that by removing the



Figure 1. Tone mapped example of the equirectangular projection of an HDR image from the SYNS dataset.

dependence of intensity distributions on DR already at retinal level, the visual system gains in efficiency at not having to perform its processes for signals that widely vary in their statistical properties.

Methods

The SYNS dataset includes scenes captured from both rural and urban locations in and around the city of Southampton in the United Kingdom. A total of 92 panoramic images in HDR format were recorded with a SpheroCam HDR (SpheronVR, Waldfischbach-Burgalben, Germany) using a Nikkor 16 mm fish-eye lens covering a range of 26 *f*-stops. Images are provided as equirectangular projections of an HDR spherical image panorama and each pixel represents an angle in space of 4 minutes of arc. An example of a simply tone-mapped image is presented in Figure 1. The camera captures the full field of view by rotating about a point in space and uses multi-exposure capture and fusion so as to be able to record HDR content. Details about the database creation are provided in Adams et al. (2016). Exposure time, ISO and the aperture size that centered the median intensity value of each image are provided, and these values were used to get the irradiance maps.

Because equirectangular projections suffer from distortion, and in order to enlarge the number of images we use for the analysis, we perform sphere mapping using cubic projections; specifically, the *equi2cubic* function. All bottom faces show the presence of the tripod and were therefore discarded; the top faces were discarded as well due to the small amount of information provided by the upper views, mostly representing the sky, combined with some difficulties in reprojecting them properly, as observed in Figure 2. Given that humans rarely look straight upwards, and infrequently straight downwards, our analyses might better represent the statistics that humans typically



Figure 2. Tone mapped example of the cubic projection of an HDR image from the SYNS dataset.

encounter. Each $1,712 \times 1,712$ pixels square image corresponds to a $90^\circ \times 90^\circ$ field of view.

In order to get the luminance value at each pixel, sensor RGB values have to be converted to a standard color space by multiplication by a 3×3 matrix whose elements depend on the spectral sensitivities of the CCD sensors and are computed so as to minimize the error for a given set of color surfaces under a given illuminant (for details, see Bertalmio, 2014). In our case, we chose to convert sensor RGB values to CIE XYZ values, where Y is the luminance.

The spectral properties of the CCD sensors are available on the SYNS website, and the conversion matrix was computed using the reflectance spectra of 1,269 Munsell color chips (from uef.fi) under the standard CIE Illuminant D65 (obtained from cvrl.org). In short, since the luminosity function isn't within the space defined by the CCD spectral sensitivities, we act like we're illuminating the chips with a fixed amount of D65, use the luminosity function to get each chip's luminance, use the three CCD sensitivities to get what the CCD responses would be for each chip, and then run a regression to find the best linear combination of CCD responses to estimate the luminance.

Concerning the DR, it is computed as the contrast ratio $\frac{Y_{max}}{Y_{min}}$. In order to reduce the potential impact of noise, percentiles were used for the computation of the maximum and minimum values. To accommodate for

the absence of a noise model of the camera, we assume a greater impact of noise over the signal on the low luminance values (Granados et al., 2010) and use 0.1 percentile for Y_{min} and 99.99 percentile for Y_{max} . Please note that Y_{min} and Y_{max} are computed separately for each image.

The database was then divided into five different DR categories. The range was chosen such that each DR bracket contains either 73 or 74 images. The median DR and range of each DR bracket are shown in Table 1.

Throughout this paper we investigate how the image statistics vary in each DR bracket, and also how these statistics change as they undergo three transformations corresponding to some very well-established processes taking place in the eye (this is clearly not intended to be a complete model of retinal processing).

The first of these transforms concerns light scatter due to the optics of the human eye and is modeled using

DR category	Median DR	Minimum DR	Maximum DR
1	63	19	130
2	194	130	294
3	418	294	585
4	1,123	585	2,685
5	13,295	2,685	3,248,876

Table 1. DR values for the different categories.

the general glare equation provided in Vos et al. (1999), as the filter:

$$PSF(\theta) = \frac{10}{\theta^3} + \left(\frac{5}{\theta^2} + 0.1 \times \frac{p}{\theta} \right) \times \left(1 + \frac{a^4}{62.5} \right) + 2.5 \times 10^{-3} p \quad (1)$$

where θ corresponds to the viewing angle from the point from which the light is spread in degrees, and a and p are parameters depending respectively on the age and the pigmentation of the eye of the subject. For this study, an age of 25 years and a value of 0.5 for brown eyes were chosen arbitrarily. The resulting point spread function (PSF) filter was then applied by convolution to the images of our database with the same method as in McCann and Vonikakis (2017).

To model the nonlinearity of the photoreceptors' response to incident light, we use the Naka-Rushton equation (Naka & Rushton, 1966):

$$R = \frac{I}{I + I_s} R_{max} \quad (2)$$

where R is the response to the light stimulus, R_{max} the peak response, I the incident light, and I_s is the semi-saturation level (i.e., the intensity that causes half-maximum response); a global geometric average of the median and the mean was chosen for the computation of I_s as done in Ferradans, Bertalmío, Provenzi, and Caselles (2011).

Finally, to simulate the impact of center-surround modulation on the spatial frequency profile of neurons such as retinal ganglion cells, we use the model given by Enroth-Cugell and Robson (1966) that defines this filter:

$$CSF(f) = K\pi(k_c r_c^2 e^{\pi r_c f} - k_s r_s^2 e^{\pi r_s f}) \quad (3)$$

where f corresponds to the frequency in cycles/degree, r_c and r_s to the radii of the center and surround areas, and K , k_c and k_s are the parameters of the filter; r_c , r_s and the ratio $k_s r_s^2 / k_c r_c^2$ are taken from the first line of the table given in Enroth-Cugell and Robson (1966), k_c and K are respectively chosen as 1 and 10 to get transformed values in the same order of magnitude as the ones before, applying CSF.

To summarize, each input image from the database is considered as the real-world illumination reaching the eye. This signal is convolved with the PSF of Equation 1 simulating the optics of the human eye. Next, this new image is passed through the Naka-Rushton Equation 2, putatively modeling the response function of photoreceptors. Finally, we pass the image resulting from the previous step through the filter of Equation 3 to emulate the impact of lateral inhibition on contrast sensitivity.

For the statistical analysis, we choose to perform studies on single-pixel, derivative, and wavelet statistics with the exact same methodology used in Huang and Mumford (1999), which is a well-established work on natural image statistics based on the LDR database of van der Schaaf and van Hateren (1996). This will allow for a better comparison of the new results we obtain here for HDR images, and also to check the validity of our study since the statistics for the lower DR category of the database we are using should be consistent with those reported in Huang and Mumford (1999).

Results

Statistical moments

We begin by investigating how the statistical moments of median, standard deviation, skewness, and kurtosis vary as a function of DR. The results are presented in Figure 3; each subplot reports a separate moment, each dot represents an individual image, and each color denotes a different transform, with blue for the original images, red for the results of the Naka-Rushton transform, and yellow for the result of the center-surround modulation. For clarity we have omitted the plots for optical scatter as the transform (Equation 1) was deemed to have a very small impact upon the statistical moments. Each image is normalized (divided by its maximum computed with 99.99 percentile) prior to computing the moments. For the skewness, some negative values have been found and marked with a black cross on their absolute values plot in log axis.

For the first moment (Figure 3A) we observe a strong relationship between the median and DR. A highly negative correlation coefficient indicates this relationship. After the Naka-Rushton equation is applied, the effect is substantially reduced in strength but the correlation stays significant and it remains like this after the CSF transform.

In Figure 3B we can see that, for the original images, the standard deviation decreases with DR (recall that we first normalize the image between zero and one, otherwise standard deviation increases with DR). As was the case for the median, application of the Naka-Rushton equation strongly reduces this correlation and results in the standard deviation hovering around the 0.2 mark, irrespective of the DR; by comparison, we would expect a uniform distribution in the range [0,1] to have a standard deviation of 0.29. Again, the CSF filter keeps the correlation with DR low and simply reduces the magnitude of the moment.

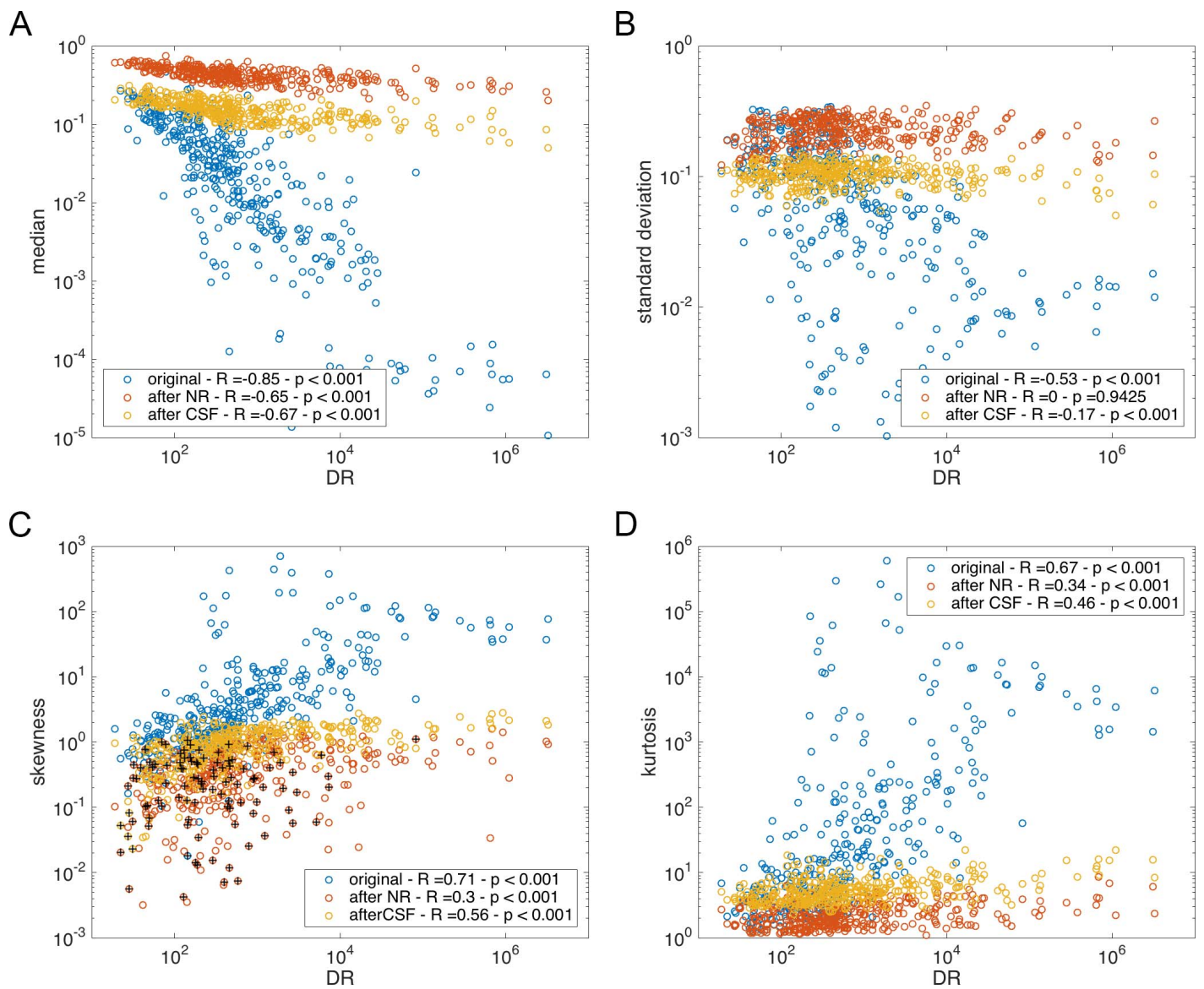


Figure 3. Statistical moments of normalized intensity distributions as a function of DR: median (A), standard deviation (B), skewness (C), and kurtosis (D).

Figure 3C shows that LDR natural scenes are typically positively skewed (concentrated toward low values), and the skew increases strongly with the DR. This is consistent with the normalized median decreasing with DR (assuming that intensity distributions are unimodal), which is what we saw in Figure 3A. After the Naka-Rushton equation is applied, the skewness no longer increases with DR, and after the CSF filter, skewness values are much less spread.

A similar phenomenon can be observed in Figure 3D for kurtosis. We note that kurtosis values are 3.0 for a normal distribution and 1.8 for a uniform distribution, and the figure shows that for LDR images, the kurtosis values are relatively close to those numbers, but then they increase very rapidly. Again, the impact of DR on kurtosis is greatly reduced by the application of the

Naka-Rushton equation, and the effect is maintained after the CSF transform.

The main observation is that the linear images, represented by the blue dots, have statistical moments that vary strongly with DR for each of the statistical moments evaluated, but that this effect size is greatly reduced by the application of the Naka-Rushton equation, as depicted by the red dots. Overall, we can conclude that the application of the Naka-Rushton equation greatly reduces the impact of DR on the statistical moments of the illuminance distributions. The band-pass filtering of the CSF maintains this relative independence from DR while producing other effects that can be better appreciated when we compute the single pixel, derivative, and wavelet statistics, in the following sections.

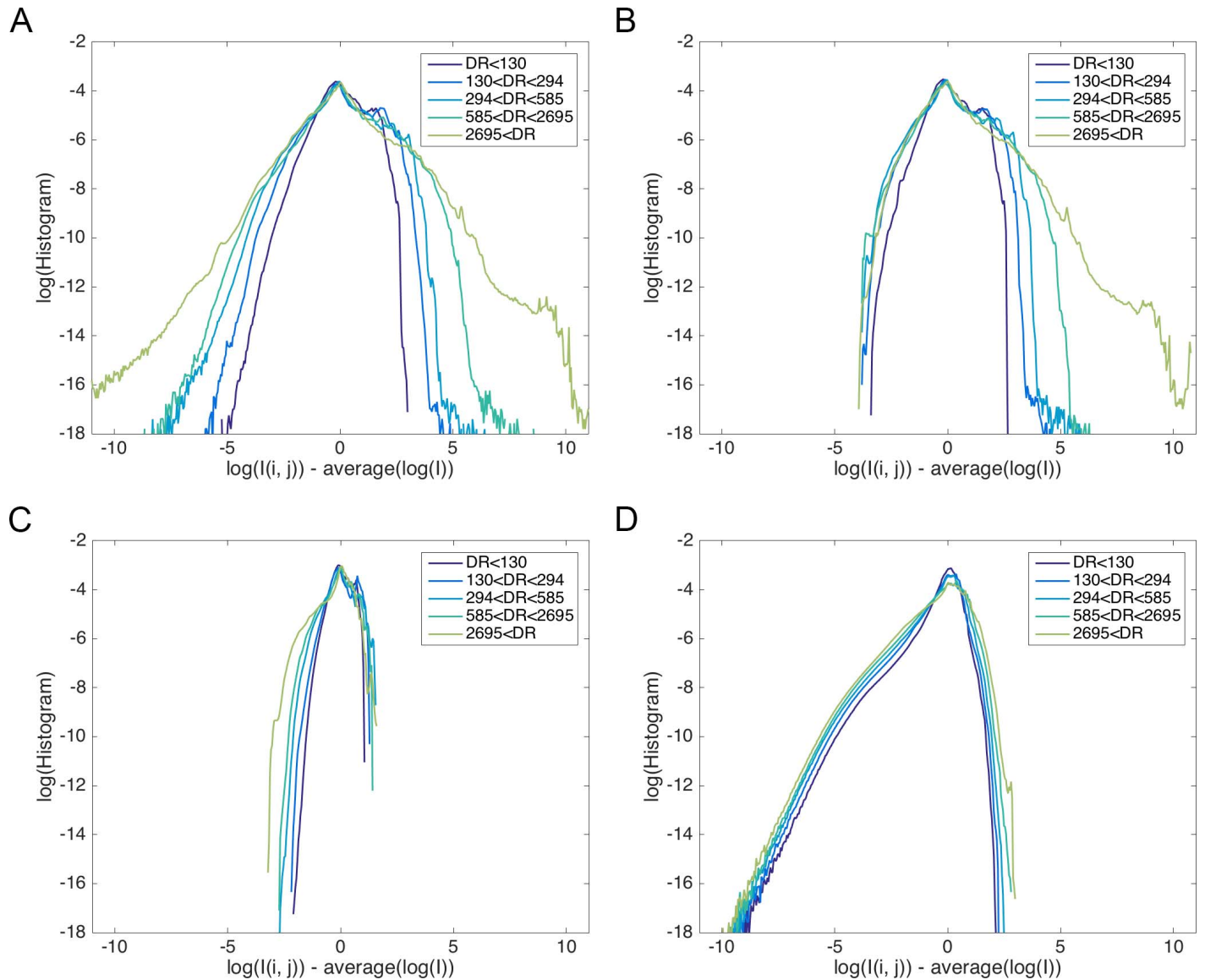


Figure 4. Median-subtracted histograms in log-log axes, by DR categories. (A) For the original values. (B) After light scatter, Equation 1. (C) After photoreceptor response, Equation 2. (D) After center-surround modulation, Equation 3.

Single pixel statistics

In the previous section we evaluated the image statistics as a function of their DR on an image-by-image basis. In this section we shall examine the average shape of the histogram for each of the five DR brackets described in the Methods section (Table 1). To do so, we use the methodology of Huang and Mumford (1999) in which each image is passed through a log transform before the median is subtracted, as follows, $H = \log(I(i,j)) - \text{median}(\log(I))$. Average histograms are then computed for each DR bracket and the results are plotted on log-log axes. In Figure 4 we plot the average histograms of the original linear image in subplot A and remaining transforms in subplots B, C, and D. Each color represents a different DR category.

To illustrate the effect observed on the figure and to compare these histograms with previous studies (Huang & Mumford, 1999, Dror et al., 2001) the statistical moments of the histograms are stored in Table 2 provided in the annex at the end of the document.

As an increase of standard deviation values in Table 2 demonstrates, the width of the histograms for the original linear images unsurprisingly increases with DR. Interestingly, we only observe straight edges in the histogram for the highest DR bracket we evaluate. This contrasts with the work of Huang and Mumford (1999), in which straight edges are found for the average histogram of low DR database of van der Schaaf and van Hateren (1996). As the tails of the histograms are better represented by the value of the

DR category	Mean	Standard deviation	Skewness	Kurtosis
Original (4A)				
1	0.1136	0.8614	0.1802	2.927
2	0.1903	1.153	0.1499	2.766
3	0.1305	1.2511	0.1917	3.268
4	0.1997	1.311	0.3977	3.7359
5	0.0739	1.381	0.4939	4.889
After PSF (4B)				
1	0.1419	0.8125	0.3548	2.863
2	0.1818	1.049	0.2364	2.800
3	0.1932	1.177	0.3445	2.977
4	0.2297	1.178	0.6526	3.599
5	0.1706	1.188	1.017	5.117
After NR (4C)				
1	0.0390	0.4169	−0.0187	2.686
2	0.0892	0.5317	0.0634	2.429
3	0.0172	0.5847	−0.1201	3.063
4	−0.0333	0.5999	−0.4797	3.249
5	−0.1645	0.6799	−0.9703	3.780
After CSF (4D)				
1	−0.1352	0.7105	−2.125	10.99
2	−0.1840	0.8328	−1.809	8.303
3	−0.2003	0.9193	−1.612	7.155
4	−0.2090	1.025	−1.329	5.719
5	−0.2376	1.115	−1.154	4.965

Table 2. Moments for the histograms of Figure 4. Note that the mean is not zero because the histograms are centered on the median.

kurtosis, the similarity of the kurtosis found by Huang and Mumford ($k = 4.56$; 1999) and the kurtosis found for the histogram of the highest DR bracket ($k = 4.89$, cf. Table 2) confirms this observation. We speculate upon why this is the case in the Discussion where we replicate the original findings.

Regarding the impact of the eye's PSF on illuminance distributions, represented in Figure 4B, one can observe that the impact is limited to raising the low illuminance values while having very little impact upon values above the median. This can also be seen as an increase of the skewness values of the histograms after applying the transform (cf. Table 2). The PSF models imperfections in the eye's medium that cause light scatter, and whose net effect is to increase the minimum intensity level of the light reaching the retina photoreceptors (Stiehl, McCann, & Savoy, 1983).

Figure 4C shows that after passing through the nonlinear transform of the photoreceptors, Equation 2, the average histograms for all DR categories become much more similar and also their DR is greatly reduced (the width of the histograms is much smaller). In Table 2, standard deviation values are decreased and the effect of the DR category on the moments appears to be reduced for the standard deviation, the skewness,

and the kurtosis. This can also be seen in Table 2, where standard deviation values decrease after the Naka-Rushton transform, and these results can be linked to the fact that the photoreceptor transform is a compressive nonlinearity that increases low intensity values much more than high intensity values, therefore reducing the DR (by increasing the minimum).

Finally, in Figure 4D we can see that application of the filter of Equation 3 produces images where the similarity of histograms across DR brackets is preserved, even enhanced, but now the DR has been notably increased. This is shown in Table 2 by the fact that the skewness is now significantly negative for all the categories and the standard deviation has increased substantially, meaning that the distributions now have more weight in the low intensity values and they have drifted further away from the high intensity values (i.e., the lower values have been reduced further and therefore the DR increases).

Derivative statistics

Derivative statistics correspond to the difference between two adjacent pixels. We find similar histograms for both horizontal and vertical derivative statistics, and as such we only plot results for the horizontal derivative distributions (computed with the formula $D = \log(I(i, j)) - \log(I(i, j + 1))$), represented in Figure 5. Again, the statistical moments of the histograms are stored in Table 3 provided in the annex at the end of the document.

For the original images, Figure 5A, we see that as the DR increases the adjacent pixels typically become more different and the width of the difference histograms widens. This result corresponds to the standard deviation globally increasing with DR, as Table 3 shows. In the same table, very high kurtosis values for all DR categories are consistent with the fact that in natural images neighboring pixels are highly correlated and therefore difference histograms present a high peak at zero (Huang & Mumford, 1999). The kurtosis value for the lower DR category, $k = 16.81$, is very similar to the one reported by Huang and Mumford, 1999, which was $k = 17.43$.

Light scattering has the effect of making these distributions more similar for all categories and also closer to zero values (see Figure 5B). This result was expected since the PSF filtering affects neighboring pixels and reduces differences between them. In Table 3, only the standard deviation is significantly modified, and it is greatly reduced by this transform.

The distributions for the five DR brackets become even more similar and their spread is reduced even further after applying the Naka-Rushton transform, see Figure 5C. This is consistent with what we observed

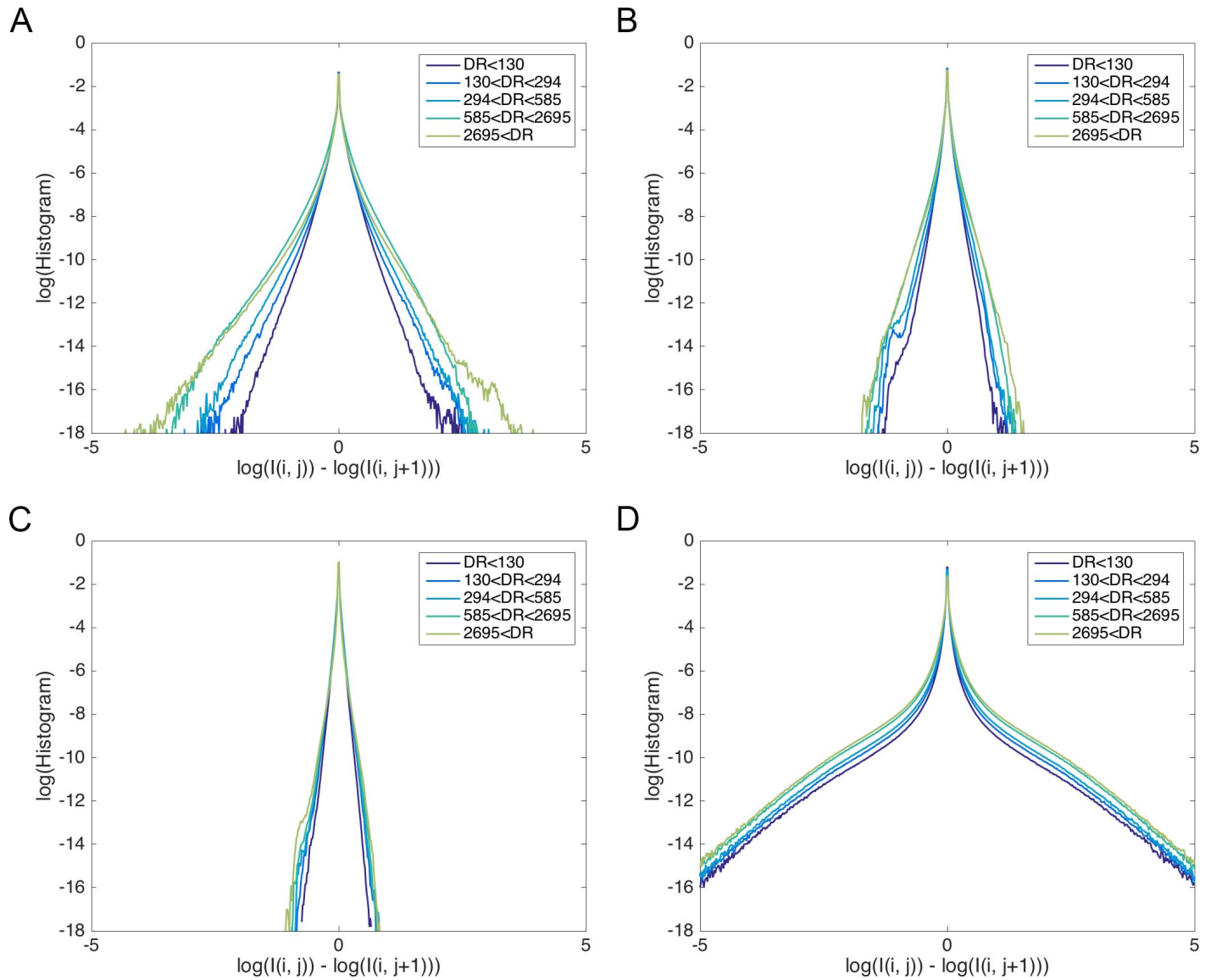


Figure 5. Histograms of the difference between adjacent pixels in linear-log axis. (A) For the original values. (B) After light scatter, Equation 1. (C) After photoreceptor response, Equation 2. (D) After center-surround modulation, Equation 3.

in Figure 4C, that showed that the DR was reduced, as well as the variability among the distributions in the different categories.

Finally, the contrast enhancement effect of the center-surround modulation provided by filtering with Equation 3 can be clearly seen in Figure 5D, where the distributions remain very similar for all DR categories but the spread is now greatly increased. This corresponds, in Table 3, to a notable increment of the standard deviation for all DR categories.

Power spectra

The literature about natural image statistics has consistently demonstrated that the power spectra of

natural scenes is linear when plotted on log-log axes. This corresponds to a $1/f^{2+\eta}$ power law relationship, with η called the “anomalous exponent,” usually small. An implication of this feature is that natural images are scale-invariant (Burton & Moorhead, 1987; Field, 1987; Ruderman & Bialek, 1994; van der Schaaf & van Hateren, 1996; Tolhurst, Tadmor, & Chao, 1992).

In the previous works of Dror et al. (2001) and Pouli et al. (2010) it was found that this rule was not valid for some HDR images. They associated this finding with the presence of direct light sources and/or specularities in the image. In this study, we investigate the effect of the DR on the averaged power spectrum of each image after collapsing across orientation.

Figure 6A shows the mean of power spectra for the five DR brackets. Substantial curvature is observed for

DR category	Mean	Standard deviation	Skewness	Kurtosis
Original (5A)				
1	0	0.1065	−0.0167	16.81
2	0	0.1132	0.0369	24.96
3	0	0.1313	−0.13297	23.83
4	0	0.1799	−0.2482	18.82
5	0	0.1469	−0.2529	34.19
After PSF (5B)				
1	0	0.0584	−0.0680	15.45
2	0	0.0591	−0.1531	24.75
3	0	0.0700	−0.2053	22.21
4	0	0.0904	−0.2073	18.15
5	0	0.0832	−0.1985	24.71
After NR (5C)				
1	0	0.0398	−0.0767	12.45
2	0	0.0459	−0.0329	13.09
3	0	0.0436	−0.1104	17.53
4	0	0.0442	−0.0771	18.62
5	0	0.0410	−0.2630	31.19
After CSF (5D)				
1	0	0.1967	−0.0015	102.0
2	0	0.2280	−0.0051	75.59
3	0	0.2509	−0.0034	62.40
4	0	0.3031	−0.0043	43.02
5	0	0.3262	−0.0012	37.34

Table 3. Moments for the histograms of Figure 5.

the highest DR category of the dataset, corresponding to a flattening of the distribution at low spatial frequencies. Some flattening may also be seen in the second-highest DR category. The effect of the eye's PSF (Figure 6B), results in a small reduction in the amplitude of the high spatial frequencies. After application of the Naka-Rushton equation we can see in Figure 6C that the $1/f^2$ relationship is recovered. Finally, Figure 6D depicts the result of the center-surround transform, which decreases the energy in the lower frequencies and also in the frequencies over 100 cycles/image. The CSF acts like a band pass filter resulting in the effect observed in Figure 6D.

We wanted to focus on the recovery of the scale-invariance property of natural scenes. As mentioned in Dror et al. (2001), the curvature present in the power spectrum of HDR images can be due to strong localized light sources producing very high illuminance values over small regions. To study the image-per-image evolution of this statistic, a second order polynomial was fit to the averaged power spectrum of each natural image, $P(x) = ax^2 + bx + c$, where $a = 0$ would correspond to the scale-invariant case where the power spectrum can be approximated by a linear function, and $a \neq 0$ implies some curvature in the fit. These fits were performed up to 200 cycles/image to avoid an overestimation of the fitting error that would be

produced by the drop on the high frequencies due to the PSF.

Figure 7 plots the term a as a function of DR, where negative values of a correspond to an inverted U-shape. The blue dots show the fits to the original images, and it can be seen that the coefficient a becomes more negative while increasing DR. The correlation coefficient plotted in Figure 7 is significant and the effect is clear on the figure. The red dots show the fits to the data after the image has been passed through the eye's PSF and the Naka-Rushton equation. As the figure shows, the fits do not exhibit the same degree of negative curvature at high DR values. The correlation coefficient is greatly reduced and one can observe that the a values remain around zero. In Figure 8A, we plot the fitting error for a first and second order polynomial fit respectively in blue and red for the camera sensor output, and in Figure 8B, we plot the error for the images after modeling the effect of the light scatter and the photoreceptor response. The main observation that can be made is that the curvature of the power spectra is reduced after the application of the Naka-Rushton equation. Indeed, in Figure 7 the absolute value of the a coefficients are below 0.1. And, although Figure 8A shows that even a second-order fit may have significant error, the application of the Naka-Rushton equation allows for the power spectra to be well approximated by a first-order polynomial, as the small errors of Figure 8B suggest.

In Figure 7, one can see that the curvature is evolving continuously as a function of the DR for the original values. HDR scenes tend to have a curved power spectrum, and while this curvature can be due to high specularities as reported before (Dror et al., 2001), that isn't necessarily the case. For instance, in Figure 7 some images with a high DR have a value of a close to zero or a small error of the first order polynomial fit in Figure 8A, showing that some HDR images comply with the $1/f$ rule. An example will be shown in the Discussion section.

Wavelets

Following Huang and Mumford (1999) we study joint statistics in the wavelet domain, deconstructing the image into subbands of different orientations and spatial scales. Again we want to compare with their results, so we do the exact same analysis, using Haar wavelets and doing the same type of plots. Joint histograms are shown in Appendix Figures A1, A2, and A3 as contour plots of log-histograms of eight wavelet coefficients pairs. The finest possible scale was used for the joint distributions, so *horizontal component*, *vertical component*, and *diagonal component* refer to the wavelet coefficients of the first sub-band for each orientation.

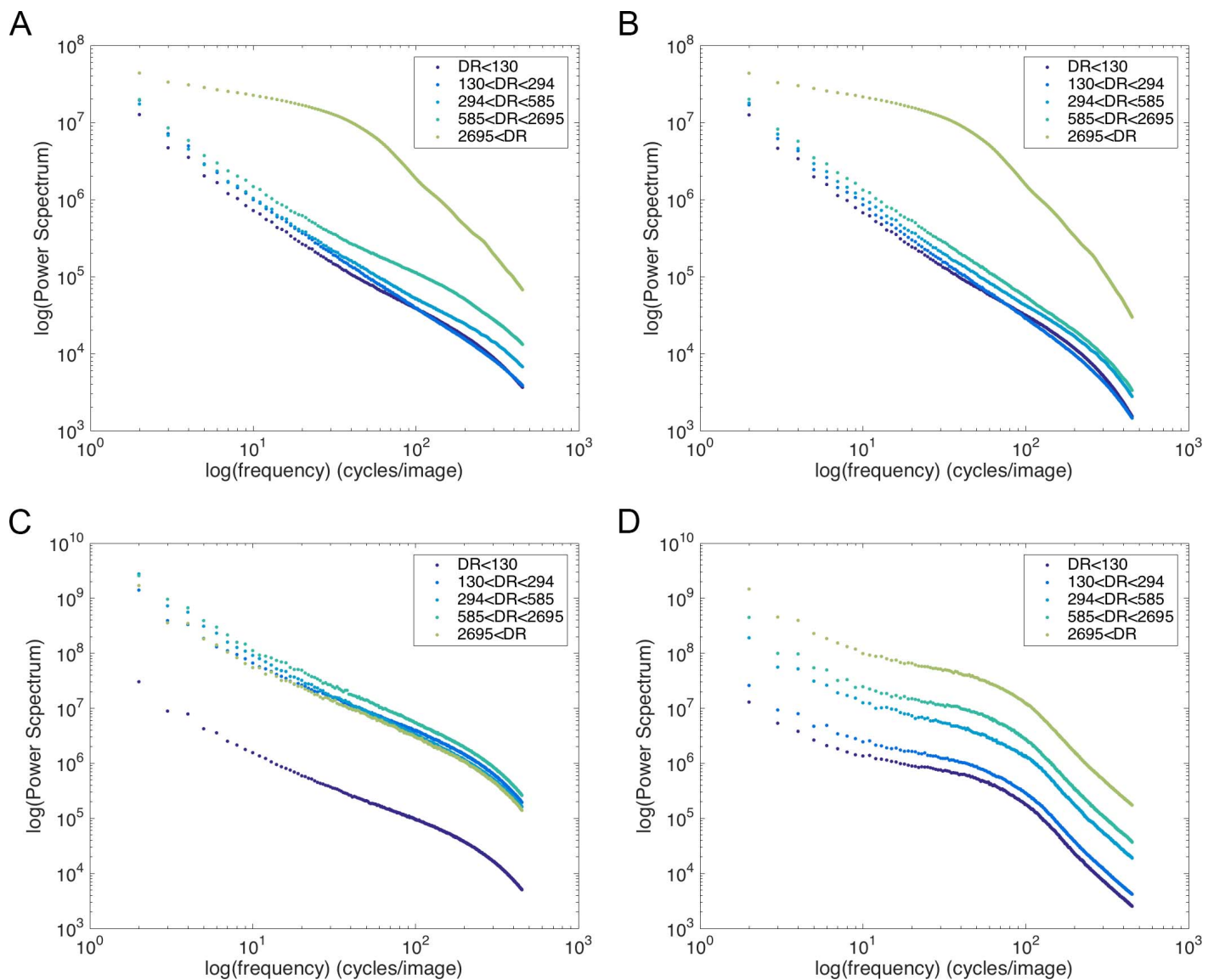


Figure 6. Power spectrum as a function of frequency on log-log axes. (A) For the original values. (B) After light scatter, Equation 1. (C) After photoreceptor response, Equation 2. (D) After center-surround modulation, Equation 3.

According to the nomenclature used in Huang and Mumford, 1999, these components are cousins to one another, coefficients at adjacent spatial locations in the same sub-band are called brothers, and coefficients that correspond to the same spatial location at different scales are called parent and child.

Figure A1 plots joint wavelet histograms computed on the original images for the lower DR category (top panel) and the higher DR category (bottom panel). In the lower DR category, the contour plots for all the wavelet coefficient pairs shown have very similar shapes to the corresponding plots reported in the work of Huang and Mumford (1999) for LDR images, which is another check of the validity of our study: the energy is equally distributed between the horizontal and vertical components but there is more energy for the horizontal

wavelet coefficients than for the diagonal ones, joint histograms for horizontal brothers are elongated along the identity axis, and at this scale parents appear to have more energy than children. These qualitative properties are also observed for the higher DR category, Figure A1B, but these plots are quite different from the corresponding ones in Figure A1A, showing how wavelet statistics are also affected by DR: here the histograms are much less spread, which may be due to the fact that many HDR images have a strong light source that dominates the distribution and makes, in comparison, most of the Haar wavelet coefficients very small.

Figure A2 plots joint wavelet histograms computed on the images after applying the Naka-Rushton equation (the stage with the PSF is omitted because of

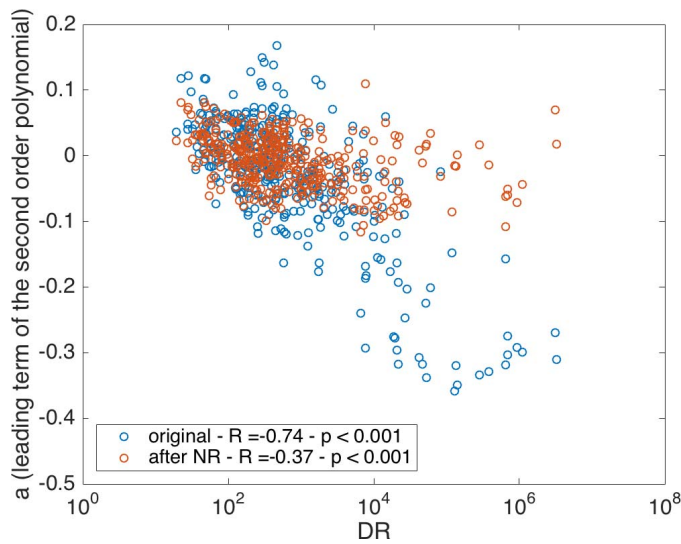


Figure 7. Leading term coefficient of the second order polyfit as a function of DR for the original values (blue) and at the photoreceptors' output (red).

its small effect on the statistics). The overall shapes of the plots resemble those of the original images, but after photoreceptor transduction the wavelet statistics look more similar among the low DR and the high DR categories; some differences remain like the sharpness of the vertices.

The last transform, the center-surround modulation, whose wavelet statistics are represented in Figure A3, makes the joint histograms even more similar between the two extreme DR categories. The action of the photoreceptor response, first, and the center-surround transform next, are enough to make the wavelet statistics virtually independent of the DR of the input. Let us

point out, however, that now the shapes are changed significantly; for instance, both the joint histogram for the horizontal component and its upper brother and the joint histogram for the vertical component with its left brother seems to be extended along the $f(x) = -x$ axis: neighbors become then less similar.

Discussion

The results reported above show that the statistical analysis of image databases that include mid to high DR images will lead to dramatically different results depending on the DR of the image in question.

We have observed, however, that these differences disappear or are substantially weaker after we pass the images through some of the transformations that occur in retinal processing. For instance, the first to fourth statistical moments vary substantially less with DR, the average histograms become more similar across the DR brackets evaluated, and the curvature seen in the power spectra of the highest DR category disappears after a nonlinear transform representing a process as early as photoreceptor transduction.

However, past studies on natural image statistics addressing later processes in the visual system are not invalidated, because images with a high DR will have the same properties as LDR images after being absorbed by the photoreceptors. Furthermore, if one assumes that the HVS is adapted for natural scenes as studied before with single-exposure images, this property can extend to HDR images regarding the similarity of the intensity distributions after the

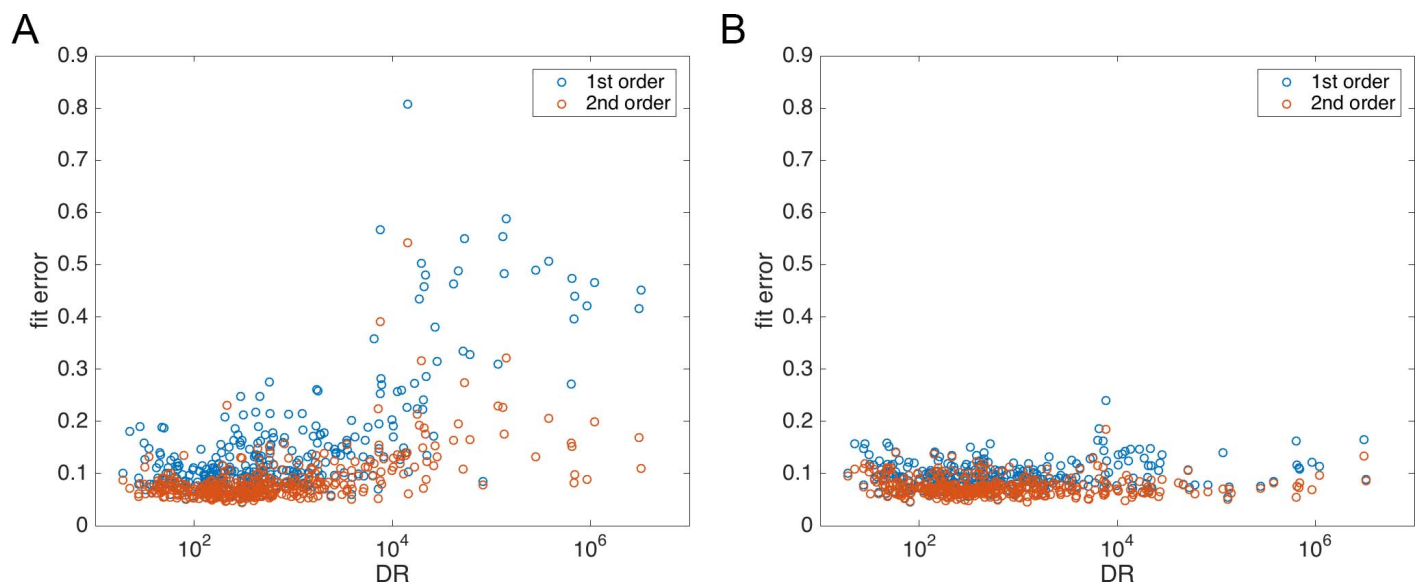


Figure 8. Fitting error for the second and first order polyfit as a function of DR for the original images (A) and the photoreceptors' output (B).

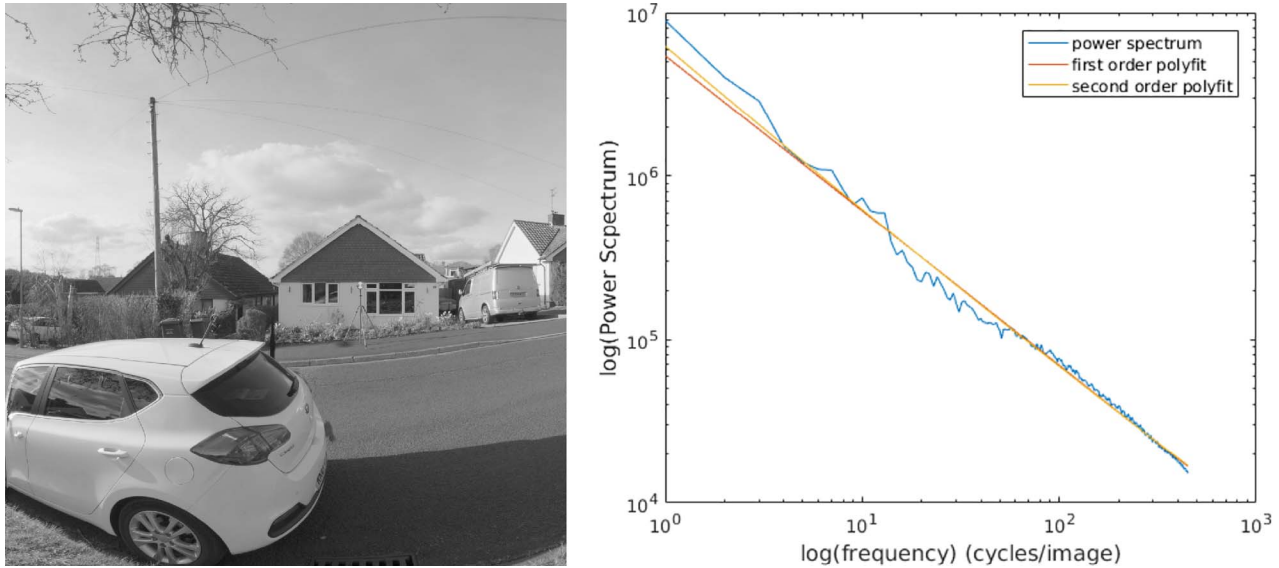


Figure 9. Example of an image (tone mapped version on the left) with a high DR ($DR = 92683$) and a power spectrum (on the right) following the $1/f$ rule.

application of retinal transforms as shown in this paper. This is consistent with the efficient coding hypothesis, that predicts that an efficient visual encoder must adapt to the full shape of the intensity distribution (Smirnakis, Berry, Warland, Bialek, & Meister, 1997). The visual system should perform quite elaborated processes in order to directly adapt to image inputs whose statistics widely vary, as in the case with natural images of different DR. Therefore, the visual system gains in efficiency at having to perform simpler processes that are tailored to a more limited range of statistical variation, when the inputs have their dependence on DR removed early on and are already at retinal level.

The process that appears to be contributing more to make the signal distributions independent of DR is the nonlinear photoreceptor transform, which we have modeled with the Naka-Rushton equation. This result was to be expected, as one of the main purposes of adaptation at photoreceptor level is to increase efficiency by handling “the very extensive range of light levels presented to the eye by nature” (Shapley & Enroth-Cugell, 1984, p. 272), that is, to reduce the DR of the signal. In fact, in the image processing and computer graphics fields, very successful methods for reducing the DR of images are based on the Naka-Rushton equation (e.g., see Ferradans et al., 2011, and references therein).

Center-surround modulation, on the other hand, *increases* the DR, as observed in Figure 4D, which is consistent with contrast enhancement properties of the visual system (Martinez, Molano-Mazon, Wang, Sommer, & Hirsch, 2014) and psychophysical studies of the DR of the visual system (Kunkel & Reinhard, 2010). The CSF transform also widens the derivative

histograms (Figure 5D), flattens the power spectra for frequencies below 100 cycles/image (Figure 6D), and reduces the similarity among neighbors (Figure A3). These properties support the decorrelation hypothesis (Atick & Redlich, 1992) and the response equalization hypothesis (Field, 1987).

We are very much aware that the first stages of visual processing performed by the HVS are way more complex than the transforms described in this study. Many different types of ganglion cells exist with functions and size varying substantially (Sanes & Masland, 2015) and one can argue that strict spatial information is incomplete when studying the HVS and still images do not represent the fundamental spatio-temporal information reaching the eye (Rucci & Victor, 2015). Still, with basic models of processes occurring in the retina one can have a hint of the functions of the different stages.

Assumptions regarding the nature of HDR scene illumination distributions can also be made with these results. When observing the evolution of the moments in Figure 3, one can see that illumination maps go to highly positively skewed and leptokurtic distributions when the DR increases, which could correspond to images with highlights. The curvature of the power spectra of most of the HDR images can also support the hypothesis that HDR scenes are characterized by specularities and bright light sources that flatten the low frequency part of the power spectrum, as pointed out by Dror et al. (2001). Although, it is important to notice that some HDR images do not present a curved spectrum. In Figure 7, they correspond to blue dots in the top-right corner, with a relatively high a and a high DR. An example of this image with its power spectrum is presented in Figure 9. In order to be

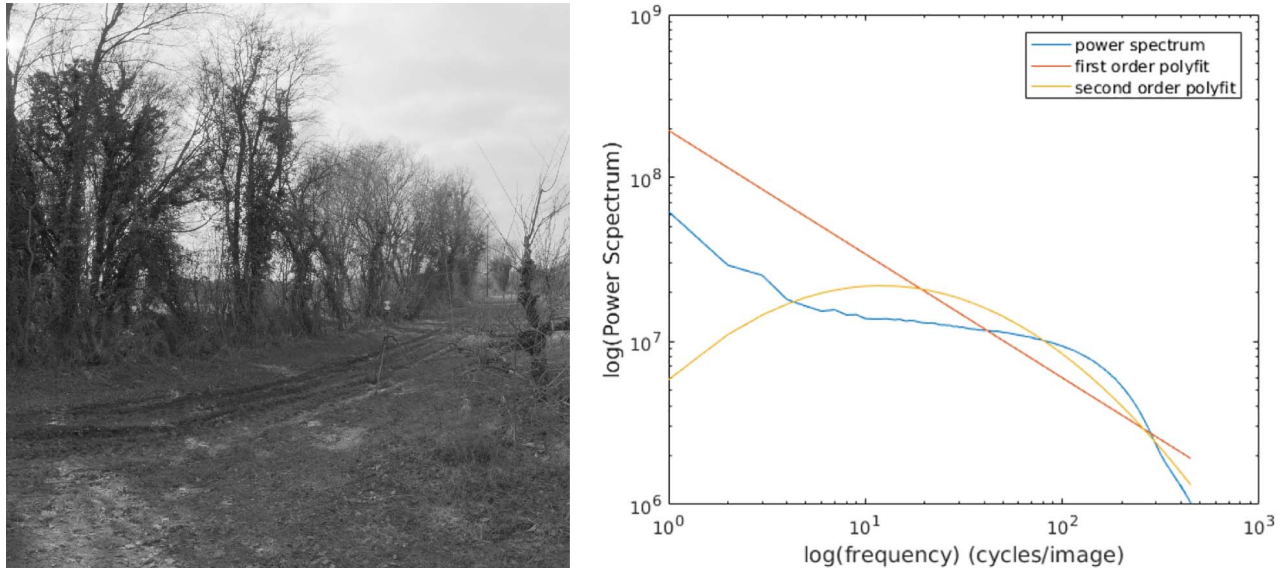


Figure 10. Example of an image (tone mapped version on the left) with a high DR ($DR = 453400$) and a power spectrum (on the right) that can't be well approximated by a second order polynomial.

displayed properly, the image is tone mapped so one can't see the specularities in the car and the truck in the background. This scene may abide by the $1/f$ law because of the multiple specularities and their distribution in space, as opposed to a local light source that would indeed flatten the power spectrum.

Some HDR images can have also a more complex shape than this flattening in the low frequencies. Most images have a power spectrum well represented by a second order polynomial but some exceptions with a relatively high fitting error are noticed in Figure 8A. We show one example of this complex power spectrum in Figure 10. Here, the sun is present in the top-left corner of the image forming a local light source. The geometry of the scene can affect the power spectrum and make it take a shape that can only be approximated with a third order polynomial. No further study on this topic is presented here but one can agree on the fact that specular highlights affect the power spectrum of natural scenes, as stated in a previous study (Dror et al., 2001).

The histogram analysis conducted in Figure 4 followed the methodology used in Huang and Mumford (1999) to compute and visualize the average illuminance histogram of the van der Schaaf and van Hateren image database, showing that it was well approximated in piece-wise linear formulation when plotted on log-log axes. There was however, some ambiguity as to how the original histogram was computed, as Huang and Mumford (1999) stated only that images underwent the following transformation: $\log(I) - \text{average}(\log(I))$. Thus, it was not clear whether average referred to the mean, the median, or some other computation. Additionally, the van der Schaaf and van Hateren image database contains two images

sets denoted *.imi* and *.imc*, the former being images that are linearly related to the sensor values and the latter having a correction for the optics of the camera applied. In Figure 11, we test the four possible combinations using the van der Schaaf and van Hateren image database. The results demonstrate that we only obtain the characteristic linear slopes when we use the *.imc* optically corrected image dataset and subtract each image using the image median. As such, we use the median to compute the histograms in Figure 4. The gray dashed box highlights the region illustrated in the original study by Huang and Mumford (1999), which is substantially smaller than the region we plot here, even if one can argue that these parts of the histograms contain a very small number of pixels. When plotted over this greater range we do not obtain a straight line over the full range of positive values. In the SYNS dataset, images are not optically corrected and it may be a reason why we do not observe linear parts on the histograms of small DR categories. It is to be noted that the number of images per category is much smaller in this study than in the previous one; we have about 70 images for each category, while in Huang and Mumford (1999), the dataset contains more than 4,000 natural scenes. The type of scenes can then affect the histograms; for example, in Figure 4A, one can observe a peak in the positive part of the histograms. This peak may correspond to the pixels forming the sky as the SYNS dataset contains a lot of scenes with the horizon creating bimodal intensity distributions.

However, a piece-wise linear histogram is observed for the HDR category and one explanation would be that HDR images are less affected by optical scatter regarding the distribution they have. It is hard to make

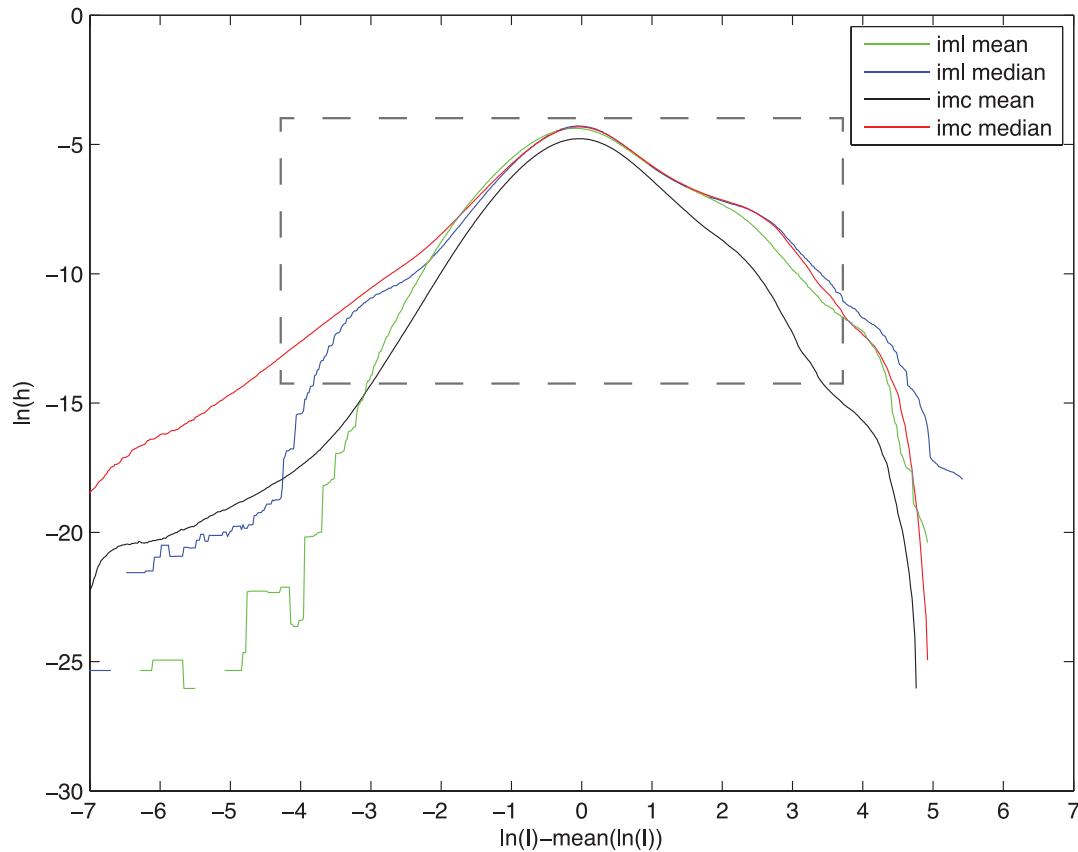


Figure 11. Histograms computed from the LDR van der Schaaf and van Hateren image database.

a strong conclusion about the real world illuminations, given the limitations imposed by the capture devices. One solution is to use physically based computer graphics methods, which bypass the need for an optical system. Additionally, one can gain complete access to the reflectance and illuminance distributions, potentially allowing us to develop a deep understanding of why natural scenes have the illuminance distributions they do.

Conclusion

This study examines the impact of dynamic range on the statistics of real world images. The vast majority of studies into image statistics use single-exposure photography, which is constrained to less than three orders of DR without substantial over or under exposure artifacts. HDR techniques, on the other hand, allow the capture of scenes with a much wider DR, including images with direct light sources such as the sun or artificial lighting, which can often have an extremely high DR ($>10^7$), but also scenes with deep shadows and high contrast that span the mid-dynamic ranges.

The results in this paper demonstrate that the statistics of the linear image (i.e., the image received at the camera sensor) change dramatically with DR. The strongest effects are noted in the four statistical moments evaluated here and upon the one over frequency relationship for the power spectrum, which breaks down for images with a very high DR. Effects are also noted in the derivative statistics, the single pixel histograms, and the Haar wavelet analysis.

The second conclusion from this paper is that the early transformations of the human visual system, and in particular the Naka-Rushton nonlinearity that models photoreceptor responses, greatly reduce the impact of DR on the statistics, turning the four statistical moments virtually independent from DR and recovering the $1/f^2$ behavior for the power spectrum.

These results are in agreement with the efficient coding hypothesis. If image distributions have statistics that vary widely, the visual system should perform quite elaborate processes in order to directly adapt to these inputs. If, on the other hand, the inputs have their dependence on DR removed early on, already at retinal level, the visual system gains in efficiency at having to perform simpler processes, tailored to a more limited range of statistical variation.

Keywords: natural scenes, high dynamic range, statistics, retinal coding, spatial vision

Acknowledgments

This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement Number 761544 (project HDR4EU) and under Grant Agreement Number 780470 (project SAUCE); and by the Spanish government and FEDER Fund, Grantref. TIN2015-71537-P(MINECO/ FEDER, UE).

Commercial relationships: none.

Corresponding author: Antoine Grimaldi.

Email: antoine.grimaldi@upf.edu.

Address: Department of Information and Communications Technologies, Universitat Pompeu Fabra, Barcelona, Spain.

Footnote

¹ It is worth noting that the terminology in the literature is confusing, and the term dynamic range can refer to the capture technique, to the bit-depth of the resulting image, or to the DR of the resulting image.

References

- Adams, W. J., Elder, J. H., Graf, E. W., Leyland, J., Lutigheid, A. J., & Murry, A. (2016). The Southampton-York Natural Scenes (SYNS) dataset: Statistics of surface attitude. *Scientific Reports*, 6, 35805.
- Atick, J. J., & Redlich, A. N. (1992). What does the retina know about natural scenes? *Neural Computation*, 4(2), 196–210.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, 61(3), 183–193.
- Barlow, H. B. (1961). Possible principles underlying the transformations of sensory messages. In W. Rosenblith (Ed.), *Sensory communication* (pp. 217–234). Cambridge, MA: MIT Press.
- Bertalmío, M. (2014). *Image processing for cinema*. London, UK: CRC Press.
- Burton, G., & Moorhead, I. R. (1987). Color and spatial structure in natural scenes. *Applied Optics*, 26(1), 157–170.
- Debevec, P. E., & Malik, J. (1997). Recovering high dynamic range radiance maps from photographs. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques* (pp. 369–378). New York, NY: ACM Press/Addison-Wesley Publishing Co.
- Dror, R. O., Leung, T. K., Adelson, E. H., & Willsky, A. S. (2001). Statistics of real-world illumination. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001. CVPR 2001: Vol. 2* (pp. II–II). New York, NY: IEEE.
- Enroth-Cugell, C., & Robson, J. G. (1966). The contrast sensitivity of retinal ganglion cells of the cat. *The Journal of Physiology*, 187(3), 517–552.
- Fairchild, M. D. (2007). The HDR photographic survey. In *15th Color and Imaging Conference, 2007* (pp. 233–238). Albuquerque, NM: Society for Imaging Science and Technology.
- Ferradans, S., Bertalmío, M., Provenzi, E., & Caselles, V. (2011). An analysis of visual adaptation and contrast perception for tone mapping. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(10), 2002–2012.
- 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*, 4(12), 2379–2394.
- Geisler, W. S., & Perry, J. S. (2011). Statistics for optimal point prediction in natural images. *Journal of Vision*, 11(12):14, 1–17, <https://doi.org/10.1167/11.12.14>. [PubMed] [Article]
- Granados, M., Ajdin, B., Wand, M., Theobalt, C., Seidel, H.-P., & Lensch, H. P. (2010). Optimal HDR reconstruction with linear digital cameras. In *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2010. CVPR 2010* (pp. 215–222). New York, NY: IEEE.
- Huang, J., & Mumford, D. (1999). Statistics of natural images and models. In *Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1999. CVPR 1999, Vol. 1* (pp. 541–547). New York, NY: IEEE.
- Kunkel, T., & Reinhard, E. (2010). A reassessment of the simultaneous dynamic range of the human visual system. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization* (pp. 17–24). New York, NY: ACM.
- Martinez, L. M., Molano-Mazon, M., Wang, X., Sommer, F. T., & Hirsch, J. A. (2014). Statistical wiring of thalamic receptive fields optimizes spatial

- sampling of the retinal image. *Neuron*, 81(4), 943–956.
- McCann, J. J. M., & Vonikakis, V. (2017). Calculating retinal contrast from scene content: A program. *Frontiers in Psychology*, 8: 2079.
- Naka, K., & Rushton, W. (1966). S-potentials from colour units in the retina of fish (cyprinidae). *The Journal of Physiology*, 185(3), 536–555.
- Parmar, M., Imai, F., Park, S. H., & Farrell, J. (2008). A database of high dynamic range visible and near-infrared multispectral images. In J. M. DiCarlo & B. G. Rodricks (Eds.), *Proceedings of SPIE, the International Society for Optical Engineering, Digital Photography IV*, Vol. 6817 (p. 68170N). San Jose, CA: SPIE.
- Pouli, T., Cunningham, D., & Reinhard, E. (2010). Statistical regularities in low and high dynamic range images. In *Proceedings of the 7th Symposium on Applied Perception in Graphics and Visualization* (pp. 9–16). New York, NY: ACM.
- Rucci, M., & Victor, J. D. (2015). The unsteady eye: An information-processing stage, not a bug. *Trends in Neurosciences*, 38(4), 195–206.
- Ruderman, D. L., & Bialek, W. (1994). Statistics of natural images: Scaling in the woods. *Physical Review Letters*, 73(6), 814–817.
- Sanes, J. R., & Masland, R. H. (2015). The types of retinal ganglion cells: Current status and implications for neuronal classification. *Annual Review of Neuroscience*, 38, 221–246.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423.
- Shapley, R., & Enroth-Cugell, C. (1984). Chapter 9. Visual adaptation and retinal gain controls. *Progress in Retinal Research*, 3, 263–346.
- Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24(1), 1193–1216.
- Smirnakis, S. M., Berry, M. J., Warland, D. K., Bialek, W., & Meister, M. (1997, March 6). Adaptation of retinal processing to image contrast and spatial scale. *Nature*, 386(6620), 69–73.
- Stiehl, W. A., McCann, J. J., & Savoy, R. L. (1983). Influence of intraocular scattered light on lightness-scaling experiments. *Journal of the Optical Society of America A*, 73(9), 1143–1148.
- Tolhurst, D., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic and Physiological Optics*, 12(2), 229–232.
- van der Schaaf, A., & van Hateren, J. H. (1996). Modelling the power spectra of natural images: Statistics and information. *Vision Research*, 36(17), 2759–2770.
- van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings: Biological Sciences*, 265(1394), 359–366.
- Vos, J., van den Berg, T., et al. (1999). Report on disability glare. *CIE Collection on Glare*, 135(1), 1–9.
- Xiao, F., DiCarlo, J. M., Catrysse, P. B., & Wandell, B. A. (2002). High dynamic range imaging of natural scenes. In *Color and Imaging Conference, Vol. 10*, (pp. 337–342). Scottsdale, AZ: Society for Imaging Science and Technology.

Appendix

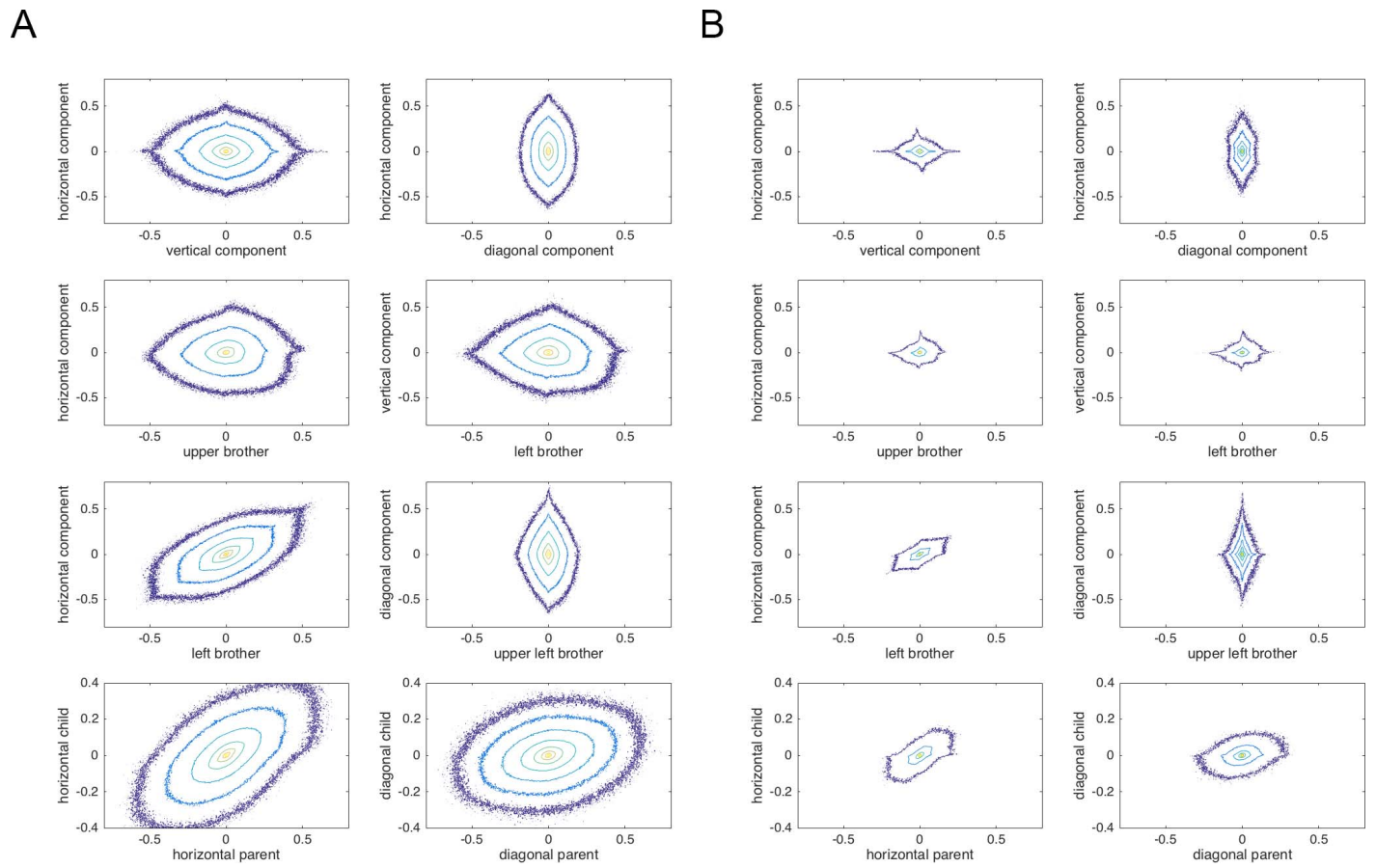


Figure A1. Wavelet analysis of original values for the (A) lowest DR category (<262) and for the (B) highest DR category (>11200).

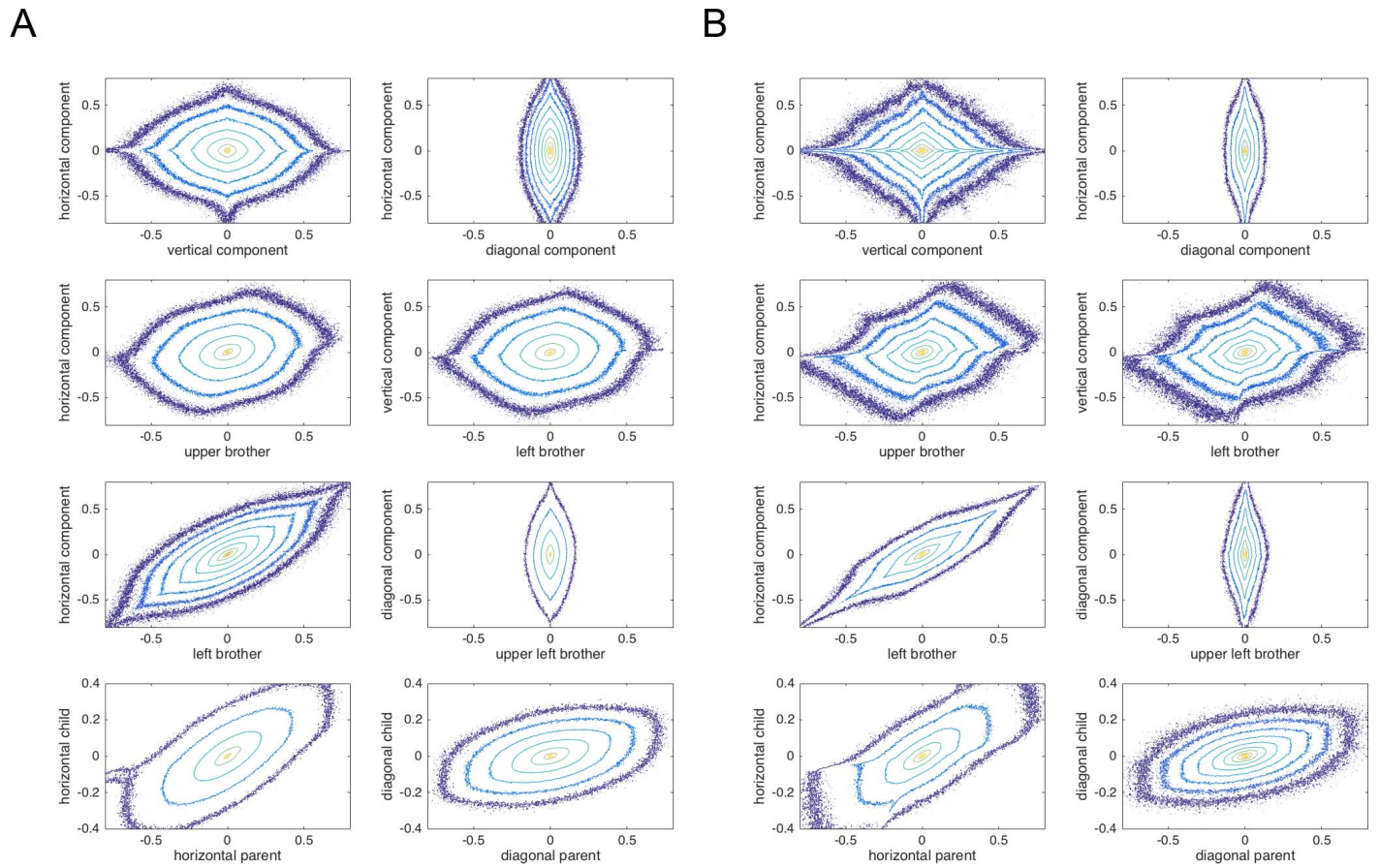


Figure A2. Wavelet analysis after NR for the lowest DR category (A) and for the highest DR category (B).

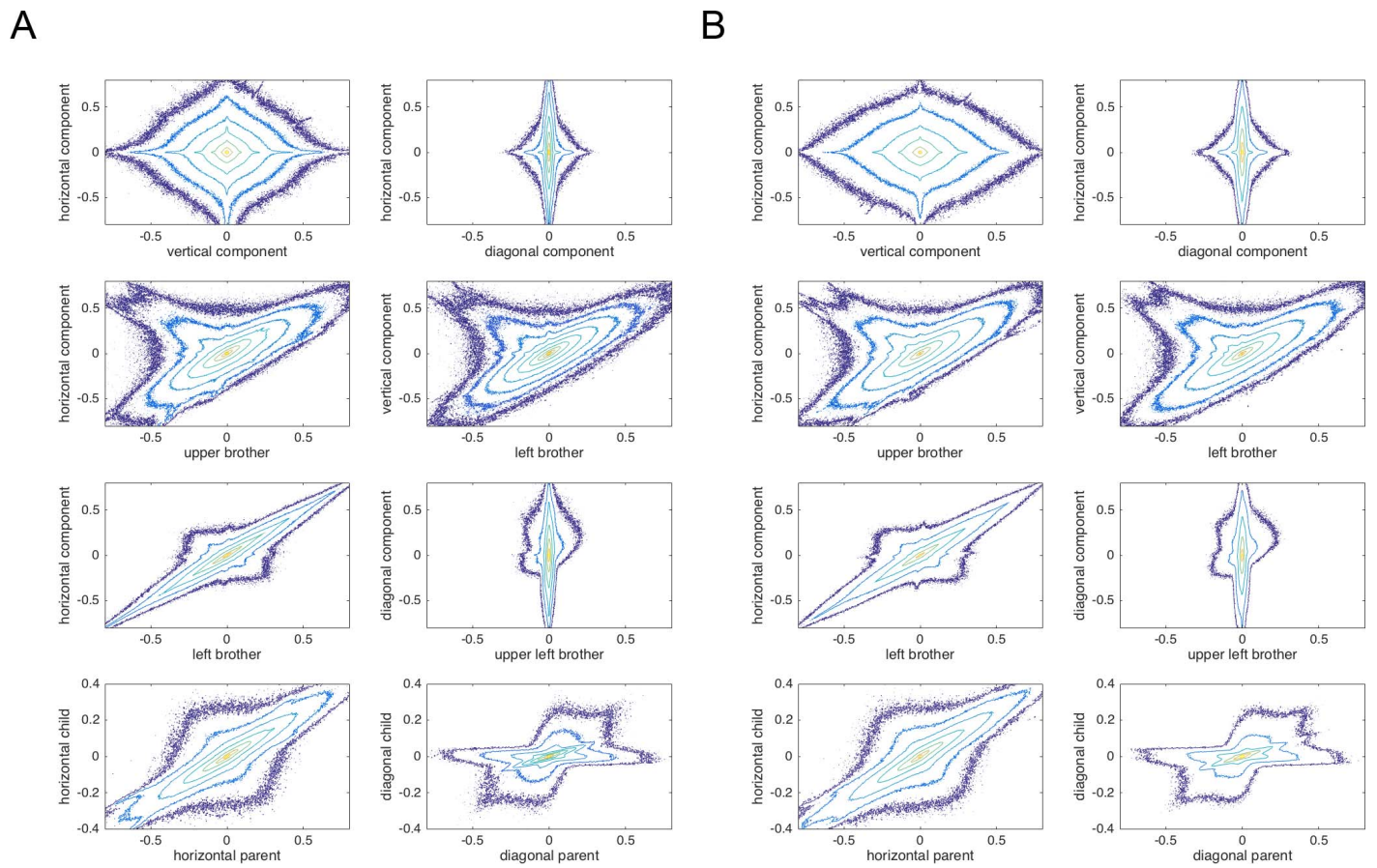


Figure A3. Wavelet analysis after CSF for the lowest DR category (A) and for the highest DR category (B).