Natural image statistics for computer graphics

Erik Reinhard, Peter Shirley and Tom Troscianko

UUCS-01-002

School of Computing
University of Utah
Salt Lake City, UT 84112 USA

March 26, 2001

Abstract

The class of all natural images is an infinitely small fraction of all possible images. The structure of natural images can be statistically modeled, revealing striking regularities. The human visual system appears to be optimized to view natural images, as opposed to any possible image, and therefore expects to interpret images which conform to these statistics. Research has shown that images that do not statistically behave as natural images are harder for the human visual system to interpret. This paper reviews the statistics of natural images and the implications for computer graphics in general are assessed. We argue that these statistics are important for graphics applications and finally, we provide a direct application of these findings to random subdivision terrain modeling.

Natural image statistics for computer graphics

Erik Reinhard* Peter Shirley† Tom Troscianko‡

Abstract

The class of all natural images is an infinitely small fraction of all possible images. The structure of natural images can be statistically modeled, revealing striking regularities. The human visual system appears to be optimized to view natural images, as opposed to any possible image, and therefore expects to interpret images which conform to these statistics. Research has shown that images that do not statistically behave as natural images are harder for the human visual system to interpret. This paper reviews the statistics of natural images and the implications for computer graphics in general are assessed. We argue that these statistics are important for graphics applications and finally, we provide a direct application of these findings to random subdivision terrain modeling.

CR Categories: G.3 [Mathematics of Computing]: Probability and Statistics; I.3.7 [Computing Methodologies]: Computer Graphics—3D Graphics; I.4.10 [Computing Methodologies]: Image Processing and Computer Vision—Image Representation

Keywords: Natural image statistics, Visual perception, Power spectra, Fractal terrains

1 Introduction

The set of all 1000 by 1000 pixel images occupies a million dimensional space. The set of natural images, which are those that appear in our world, form a sparse subset of all possible images that can be identified using statistical approaches. The statistics of natural images have been studied to understand how their properties influence the human visual system (HVS). To assess invariances in natural images, a large set of calibrated images are collected into an ensemble and statistics are computed on these ensembles. The ensemble should be chosen such that it is representative for all natural images. This cannot generally be achieved with a single image.

Natural image statistics can be characterized by their order. In particular, first, second and higher order statistics are distinguished [33]:

First order statistics treat each pixel independently, so that for example the distribution of intensities encountered in natural images can be estimated.

Second order statistics measure the dependencies between pairs of pixels, which are usually expressed in terms of the power spectrum (see section 2).

Higher order statistics are used to extract properties of natural scenes which can not be modeled using first and second order statistics. These properties include lines and edges.

First order statistics capture little structure that aids in understanding the HVS, nor does it distinguish natural images from arbitrary images very well. We therefore choose to largely ignore first order statistics.

In this paper we review current knowledge of image statistics, and argue that this knowledge is likely to be useful for some computer graphics applications. We first review the basics of image statistics (Section 2). We then discuss the mechanics of power spectra computation and how to analyze their structure (Section 3). This is followed by experiments on image ensembles (Section 4). Next, we provide evidence that second order statistics are sensitive to geometric variations rather than differences between lighting simulations (Sections 5 and 6). The results are then applied to fractal terrain modeling (Section 7), followed by conclusions (Section 8).

2 Background: second order statistics

The most remarkable and salient natural image statistic that has been discovered so far is that the slope of the power spectrum tends to be close to 2. The power spectrum of an M by M image is computed as:

$$S(u,v) = \frac{|F(u,v)|^2}{M^2}, \tag{1}$$

where F is the Fourier transform of the image. By representing the two-dimensional frequencies u and v in polar coordinates ($u = f\cos\phi$ and $v = f\sin\phi$) and averaging over all directions ϕ and all images in the ensemble, it is found that on log-log scale amplitude as function of frequency f lies approximately on a straight line [7, 11, 31–33]. This means that spectral power as function of spatial frequency behaves according to a power law function. Moreover, fitting a line through the data points yields a slope α of approximately 2 for natural images:

$$S(f) \propto \frac{A}{f^\alpha} = \frac{A}{f^{2-\eta}}. \tag{2}$$

Here, A is a constant determining the overall image contrast, α is the spectral slope and η is its deviation from 2. This result is true for ensembles, but may not hold for the lowest few frequencies when applied to individual images [3, 20].

Although this spectral slope varies subtly between different studies [7, 10, 12, 32, 40], it appears to be extremely robust against distortions and transformations and it is therefore concluded that this spectral behavior is a consequence of the images themselves, rather than of particular methods of camera calibration or exact computation of the spectral slope.

*University of Utah
†University of Utah
‡University of Sussex

However, the precise value of the spectral slope depends somewhat on the type of scenes that make up the ensemble. Most studies in this field use images of natural objects such as trees and shrubs because it is argued that the HVS evolved when only natural objects were present. Some studies show that the spectral slope for scenes containing man-made objects is slightly different [42]. Even if no manufactured objects are present, the statistics vary dependent on what is predominantly in the images. The second order statistics for sky are for example very different from those of trees.

One way in which this becomes apparent is when the power spectra are not circularly averaged, but when the log average power is plotted against angle. For natural image ensembles all angles show more or less straight power spectra, although most of the power is concentrated in horizontal and vertical angles [31, 33] (see also Figure 5). The horizon and the presence of tree-trunks are said to be factors in this, although this behavior is also likely to occur in man-made environments.

The power spectrum is related to the auto-correlation function through the Wiener-Khintchine theorem, which states that the auto-correlation function and the power spectrum form a Fourier transform pair [21]. Hence, power spectral behavior can be equivalently understood in terms of correlations between pairs of pixel-intensities.

A related image statistic is contrast, normally defined as the standard deviation of all pixel intensities divided by the mean intensity (σ/μ). This measure can either be computed directly from the image data, or through Parceval's theorem it can be derived from the power spectrum [33]:

$$\frac{\sigma^2}{\mu^2} = \sum_{(u,v)} S(u,v). \tag{3}$$

This particular contrast computation can be modified to compute contrast in different frequency bands. Frequency-conscious variance can then be thresholded, yielding a measure which can detect blur [13]. This is useful as lack of contrast can also be caused by the absence of sufficient detail in a sharp image.

The above second order statistics are usually collected for luminance images only, as luminance is believed to carry the greatest amount of information. However, chrominance channels are shown to exhibit similar spectral behavior [24], and therefore all subsequent qualitative arguments are expected to be true for color as well.

2.1 Interpretation

The fact that the spectral behavior of natural images yields a straight line with a slope of around 2 is important. Recent unrelated studies have found that in image interpretation tasks, the HVS performs best when the images conform to this second order statistic. In one such study, images of a car and a bull were morphed into each other, with varying distances between the images in the sequence [25]. Different sequences were generated with modified spectral slopes. The minimum distance where participants could still distinguish consecutive images in each morph sequence was measured. This distance was found to be smallest when the spectral slope of the images in the morph sequence was close to 2. Deviation of the spectral slope in either direction reduced the ability to distinguish between morphed images.

In a different study the effect of spectral slope on the detection of mirror symmetry in images was assessed [29]. Here, white noise patterns with varying degrees of vertical symmetry were created and subsequently filtered to alter the spectral slope. This experiment, in which participants had to detect if symmetry was present, revealed that performance was optimal for images with a spectral slope of 2.

These studies confirm the hypothesis that the HVS is tuned to natural images. This is consistent with psychophysical measurements, which indicate that the HVS expects to see images with a $1/f^2$ power spectrum and subsequently whitens them non-adaptively [1]. Whitening, or flattening the power spectrum, results in a spectral slope of $\alpha = 0$, which can be obtained by letting $S'(f) = f^2 S(f)$. Hence, it makes sense for computer graphics applications to anticipate this type of visual processing and ensure that images have $1/f^2$ power spectra. Investigation of second order image statistics is therefore the main goal of this paper.

2.2 Higher order statistics

One of the disadvantages of using amplitude information in the frequency domain is that phase information is completely discarded, thus ignoring the position of edges and objects. For this reason higher order statistics have been applied to natural image ensembles. The simplest global *nth*-order statistics that may capture phase structure are the third and fourth moments, commonly referred to as skew and kurtosis. They are computed as [21]:

$$s = \frac{E\{x^3\}}{E\{x^2\}^{3/2}},$$
$$k = \frac{E\{x^4\}}{E\{x^2\}^2} - 3.$$

These dimensionless measures are by definition zero for pure Gaussian distributions. Skew can be interpreted as an indicator for the difference between the mean and the median of a dataset. Kurtosis is based on the size of a distribution's tails relative to Gaussian. A positive value, associated with long tails in the distribution of intensities, is usually associated with natural image ensembles. This is for example evident when plotting log-contrast histograms, which plot the probability of a particular contrast value appearing. These plots are typically non-Gaussian with positive kurtosis [31].

Thomson [39] pointed out that for kurtosis to be meaningful for natural image analysis, the data should be decorrelated, or whitened, prior to any computation of higher order statistics. This can be accomplished by flattening the power spectrum, which ensures that second order statistics do not also capture higher order statistics. Thus, skew and kurtosis become measures of variations in the phase spectra. Regularities are found when these measures are applied to the pixel histogram of an image ensemble [38, 39]. This appears to indicate that the HVS may exploit higher order phase structure after whitening the image. Understanding the phase structure is therefore an important avenue of research in the field of natural image statistics.

2.3 Response properties of cortical cells

One of the reasons to study the statistics of natural images is to understand how the HVS codes these images. Because natural images are not completely random, there is a significant amount of redundancy. The HVS may represent such scenes using a sparse set of active elements. These elements will then be largely independent.

In this section it is assumed that an image can be represented as a linear superposition of basis functions. Efficiently encoding images now involves finding a set of basis functions that spans image space and ensures that the coefficients across an image ensemble are statistically as independent as possible [23]. The resulting basis functions (or filters) can then be compared to the response properties of cortical cells in order to explain the early stages of human visual processing. This has resulted in a number of different representations for basis functions, including Principal Components Analysis

(PCA) [2, 15], Independent Components Analysis (ICA) [4, 5, 22], Gabor basis functions [11, 12] and wavelets [12, 18, 23].

The PCA algorithm finds a set of basis functions which maximally decorrelates pairs of coefficients, which is achieved by computing the eigenvalues of the covariance matrix (between pairs of pixel intensities). The corresponding eigenvectors represent a set of orthogonal coefficients, whereby the eigenvector with the largest associated eigenvalue accounts for the greatest part of the covariance. By using only the first few coefficients, it is possible to encode an image with a large reduction in free parameters [11]. If the image statistics are the same in all regions of an image, i.e. if stationarity is assumed, then PCA produces coefficients which are similar to the Fourier transform [6]. Indeed, PCA is strictly a second order method assuming Gaussian signals.

Unfortunately, decorrelation does not guarantee independence. Also, intensities in natural images are do not have a Gaussian distribution, and therefore PCA yields basis functions which do not capture higher order structure very well [23]. In particular, the basis functions tend to be orientationally and frequency sensitive, but have global extent. This is in contrast to cells in the human primary cortex which are spatially localized (as well as being localized in frequency and orientation).

In contrast to PCA, ICA constitutes a transformation resulting in basis functions which are non-orthogonal, localized in space, frequency and orientation and aims to extract higher order information from images [4, 5]. ICA finds basis functions which are as independent as possible [19]. To avoid second order statistics to influence the result of ICA, the data is usually first decorrelated (also called whitened), for example using a PCA algorithm. Filters can then be found that produce extrema of the kurtosis [33]. A kurtotic amplitude distribution is produced by cortical simple cells, leading to sparse coding. Hence, ICA is believed to be a better model than PCA for the output of simple cortical cells.

The receptive fields of simple cells in the mammalian striate cortex are localized in space, oriented and bandpass. They are therefore similar to the basis functions of wavelet transforms [12, 23]. For natural images, strong correlations between wavelet coefficients at neighboring spatial locations, orientations and scales have been shown using conditional histograms of the coefficients' log magnitudes [35]. These results were successfully used to synthesize textures [27, 36] and to denoise images [34].

Figure 1: *Example images drawn from ensemble "A".*

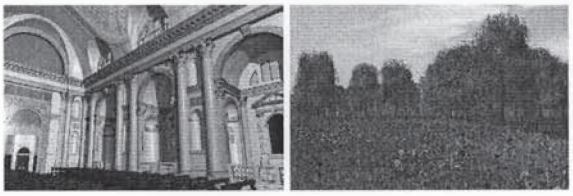

Figure 2: *Example images drawn from ensemble "B".*

Figure 3: *Example images drawn from ensemble "C".*

2.4 Motivation

For statistical approaches to be useful in computer graphics applications, it is important to be able to attach a physical meaning to the statistics that are found. For the higher order methods in the previous sections, we would be interested in the statistics of the coefficients, rather than the appearance of the basis functions. In the absence of sufficient data on these statistics, we choose to ignore these results for the moment and instead focus on the widely accepted second order statistics as discussed in section 2. In this paper we study which aspects of rendering and modeling influence second order statistics.

3 Calculation of image statistics

To verify our approach, three different image ensembles were used. Two of these ensembles consist of synthetic images, while the third ensemble consists of natural images from the van Hateren database [22]. This last ensemble consists of just over 4000 calibrated images of outdoor scenes, 133 of which were randomly chosen to form our natural image ensemble. Below, we will refer to this as ensemble "A".

For the artificial image ensembles, research images from various internet sites were collected. Images that had obvious copyright notices were rejected, as well as images that were smaller than 512x512 pixels. Further, both the rendering and the modeling had to be of high subjective quality. Images which had obvious markings of post-processing, as well as camera specific artifacts (depth of field), were also rejected. The images thus collected were subdivided into two groups, where the classification was based on the subjective quality of the modeling. The high quality modeling ensemble (ensemble "B") consists of 30 images and the ensemble with slightly simpler modeling has 18 images (ensemble "C"). Example images of all three ensembles can be found in Figures 1 to 3.

The images in the natural image ensemble are available as luminance images. The synthetic images are usually given in some RGB format. Where appropriate, these were converted to a 24 bit uncompressed file format (the influence of lossy compression on the power spectra is assessed in Section 5). All subsequent computations were performed on the Y (luminance) channel, after conversion to YUV.

For these image ensembles (as well as for individual images), the power spectra were computed and the spectral slope was estimated. The power spectrum computation proceeds as follows (after [33]): For images that are larger than 512x512 pixels, a window of this size was cut out of the middle of the image upon which further processing was applied. Then, the weighted mean intensity μ was subtracted to avoid leakage from the DC-component of the image, with μ defined as:

$$\mu = \frac{\sum_{(x,y)} L(x,y)w(x,y)}{\sum_{(x,y)} w(x,y)}. \quad (4)$$

Next, the images were prefiltered to avoid boundary effects. This is accomplished by applying a circular Kaiser-Bessel window function (with parameter $\alpha = 2$) to the image [16]:

$$w(x,y) = \frac{I_0\left(\pi\alpha\cdot\sqrt{1.0-\left(\frac{x^2+y^2}{(N/2)^2}\right)}\right)}{I_0(\pi\alpha)} \quad : \quad 0 \le \sqrt{x^2+y^2} \le \frac{N}{2}.$$

Here, I_0 is the modified zero-order Bessel function of the first kind and N is the window size (512 pixels). In addition, this weight function was normalized by letting:

$$\sum_{(x,y)} w(x,y)^2 = 1. \quad (5)$$

This windowing function was chosen for its near-optimal trade-off between side-lobe level, main-lobe width and computability [16]. The resulting images were then Fourier transformed:

$$F(u,v) = \sum_{(x,y)} \frac{L(x,y)-\mu}{\mu} w(x,y) e^{2\pi i(ux+vy)}. \quad (6)$$

Finally, the power spectrum was computed as per equation 1 and the resulting data points plotted. Although frequencies up to 256 cycles per image are computed, only the 127 lowest frequencies were used to estimate the spectral slope. Higher frequencies may suffer from aliasing, noise and low modulation transfer [33]. The estimation of the spectral slope was performed by fitting a straight line through the logarithm of these data points as function of the logarithm of $1/f$. This method was chosen over other slope estimation techniques such as the Hill estimator [17] and the scaling method [9] to maintain compatibility with [33]. In addition, the number of data points (127 frequencies) is insufficient for the scaling method, which requires at least 1,000 data points to yield reliable estimates.

4 Image ensembles

For the three image ensembles, second order statistics were extracted as detailed above. The 1.87 ($\pm$0.43 standard deviation) spectral slope for the van Hateren database was confirmed: we found 1.88 ± 0.42 s.d. for our subset of 133 images. The deviations from this value for the artificial image ensembles are depicted in Figure 4. These results show that our synthetic image ensembles produce straight power spectra, although the total amount of contrast and the value of the slope vary. It is also interesting to note that the subjective classification into two separate synthetic image ensembles using quality of modeling as criterion, has produced power spectra that are markedly different.

Especially the area under the data points, which can be interpreted as a measure for contrast, is an order of magnitude higher for the high quality synthetic ensemble. It is possible that the subjective selection criterion was subconsciously guided by the human visual system's tendency to prefer viewing areas with high contrast [30].

The angular distribution of power tends to show peaks near horizontal and vertical angles. Both the natural image ensemble "A" and the high quality synthetic image ensemble "B" show this (Figure 5). The lower quality synthetic ensemble "C" shows a distinct lack of power in horizontal directions, but has peaks present in vertical directions.

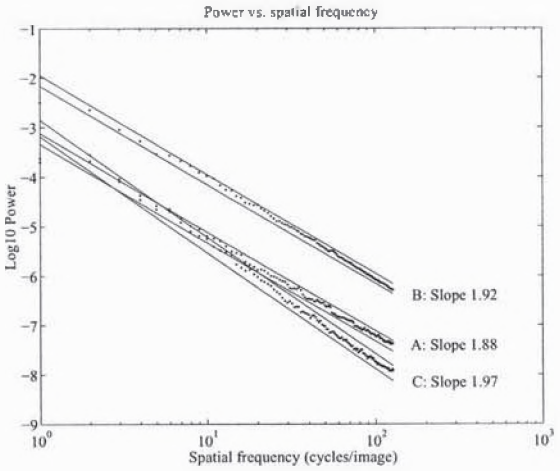

Figure 4: *Spectral slopes for the three image ensembles: A natural image ensemble, B high quality synthetic ensemble and C lower quality synthetic ensemble. The double lines indicate $\pm$2 standard deviations for each ensemble.*

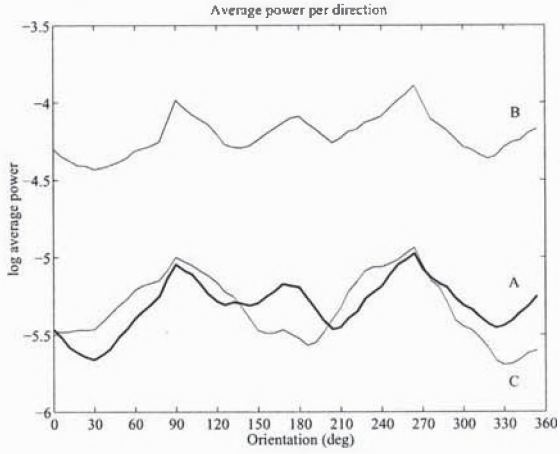

Figure 5: *Log average power as function of angle.*

Differences between the artificial image ensembles and the natural ensemble may be partially explained by the fact that rendered imagery predominantly depicts indoor scenes, whereas ensemble "A" is exclusively outdoors.

While the two synthetic ensembles produce straight lines with spectral slopes that fall well within the range of slopes observed in natural images, individual images tend to deviate fairly strongly from the average. In the following sections, we assess which parts of the rendering pipeline affect second order statistics, which goes towards establishing these statistics as a tool for graphics applications.

5 Image manipulation

In this and the following sections we empirically evaluate which aspects of rendering and modeling affect the spectral slope. Starting with issues relating to image space, possible artifacts may arise from lossy compression for file formats, such as gif and jpeg, gamma correction and aliasing. Each of these is discussed in turn, using a high quality rendering created using Radiance [41]. The

Figure 6: *Test scene, modeled to a spatial resolution of 1 mm.*

modeling for this image (depicted in Figure 6) was done by hand to a resolution of 1 mm. This provides a detailed scene as encountered in many graphics applications.

The lighting simulation included diffuse inter-reflection and soft shadows. An image was created at 512x512 pixels with 64 super-samples per pixel, which was subsequently converted to PPM. The tests in the following subsections all pertain to this rendered image, which has a measured spectral slope of $\alpha = 2.36$.

5.1 File formats

Lossy compression, as employed by various different file formats, may cause the frequency content of images to change. To see whether substantial modifications occur, the above PPM image was converted to GIF and JPEG (the latter using different levels of quality and smoothing) using XV. With the exception of smoothing, which destroys high frequency content, the effect of file conversion on the spectral slope is generally benign with deviations less than 1%. For different levels of smoothing, the spectral slope varies linearly with the amount of smoothing applied during file conversion.

Hence, moderate levels of compression do not have an appreciable effect on the measured spectral slope. We therefore conclude that the different file formats used in image ensembles B and C above, did not significantly affect our results.

5.2 Gamma correction

Gamma correction is a non-linear transformation to adapt the appearance of an image to a particular display device. As such, the spectral information present in the image may be affected due to the non-linear nature of the transformation. However, our measurements using the gamma correction option in XV show that the spectral slope is only weakly dependent on gamma correction value, as can be seen in Figure 7. The largest deviations occur for extreme gamma correction values that strongly darken the image

It is therefore concluded that gamma correction does not constitute a significant factor in the determination of spectral slopes and therefore does not substantially influence our results of Section 4.

5.3 Aliasing

Aliasing is another factor which may affect the spectral slope by projecting frequencies above the Nyquist limit to lower frequencies. Despite careful consideration of this issue by using only the 127 lowest frequencies in the spectral computations, as explained in Section 3, the rendering process itself may still cause aliasing at lower frequencies. Images with different numbers of super-samples were computed, suppressing aliasing by different amounts. Table 1 shows how the spectral slope depends on the number of super-samples. In order to minimize aliasing artifacts in the Fourier domain, it appears that at least 16 super-samples per pixel are needed. All the renderings in this and following section use 64 super-samples, eliminating aliasing as a possible factor.

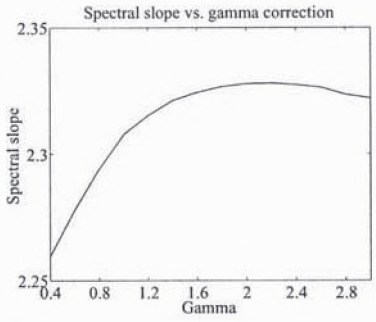

Figure 7: *Spectral slope vs. gamma correction value.*

Super-samples/pixel	α	σ
1x1	2.23	0.15
2x2	2.32	0.19
4x4	2.35	0.20
8x8	2.36	0.20

Table 1: *Spectral slope α and standard deviation σ as function of super-sampling. Slope α changes less than 1% if 16 or more super-samples are computed per pixel.*

6 Modeling or rendering?

From the previous section it is clear that most of the distortions regularly applied to synthetic images do not unduly affect our spectral analysis. In this section we extend our analysis to the spatial domain and answer the question whether second order statistics apply to lighting simulation or modeling. The scene as described in the previous section was used for rendering images using different lighting simulations: with and without diffuse inter-reflection as well as with and without shadow rays.

6.1 Shadows

Depending on the nature of the scene rendered, shadows can make an important contribution to the overall appearance of a scene. Hence, the accuracy with which shadows are rendered may affect the statistics of the resulting image. By varying the size of the light sources and adjusting their emission to maintain constant light levels, the effect of varying soft shadows on second order statistics was measured. Table 2 shows that for our room scene, shadows do not appear to be an overly important factor for computing image statistics. However, it should be noted that special cases could be constructed which prove the contrary.

6.2 Diffuse inter-reflection

We have also assessed the influence of diffuse inter-reflection on second order statistics. As diffusely reflecting surfaces are unlikely

Size	Energy	α	σ
1/8	64	2.35	0.20
1/4	16	2.35	0.20
1/2	4	2.35	0.20
1	1	2.36	0.20
2	1/4	2.36	0.20
4	1/16	2.37	0.21

Table 2: *Spectral slope α and its standard deviation σ as function of light source size and energy (both multiplication factors).*

to produce high spatial frequencies due to their illumination, we expect an even smaller effect than for the light source tests above. The largest difference occurs when switching from lighting simulations without diffuse inter-reflection to those with diffuse inter-reflection. This changed the spectral slope by 5.5%.

As adding diffuse inter-reflection to the lighting computations produced a marked effect, we were wondering whether this is due to an overall increase in energy in the environment. Ray tracing allows a constant ambient to be added to each shading result, which increases image intensities. It was found that this produced the same effect. Hence, we conclude that absolute energy levels are more important than the low-frequency distribution associated with diffuse inter-reflection.

6.3 Discussion

It appears that the particular details of the lighting simulation, whether it be soft or hard shadows, diffuse inter-reflection etc. do not significantly influence second order image statistics. Differences between renderers, such as ray tracing and OpenGL rendering with Phong shading, also did not prove essential (data not shown). We therefore conclude that these statistics are invariant to rendering details. Although second order statistics can not be used to differentiate between lighting simulations, this makes them ideal tools for assessing quality of modeling. We provide evidence for this hypothesis through the example of fractal terrain modeling.

7 Sample application: fractal terrains

We would expect natural image statistics to be useful in any realistic rendering application that normally involves parameter tuning such as solid texture creation [26], procedural plant modeling [28] or displacement mapping [8, 37]. We have chosen procedural terrain modeling as our test for whether natural image statistics can be useful for automatic parameter selection.

We implemented the midpoint subdivision algorithm of Fournier et al. [14]. That algorithm iteratively adds smaller and smaller random displacements to smaller and smaller spatial scales. As the scale decreases by a factor of 2, the magnitude of displacement decreases by a factor k. Smaller values of k result in rougher terrain, and larger values result in smoother terrain. Figure 10 shows twelve terrain models using $k = 1.5$ to $k = 2.6$ after applying 10 iterations. The terrains, each consisting of 524,288 triangles, were rendered in Radiance with diffuse inter-reflection and an 11 o'clock sky model [41]. The resulting images have an approximately linear relationship between spatial frequency and spectral slope. Figure 8 demonstrates this behavior, along with the observation that for all values of parameter k, the spectral slope decreases predictably with each iteration of the algorithm. The relationship between division parameter k and spectral slope shows after 10 iterations a minimum of 1.40 for parameter $k = 2.1$. The value of the spectral slope increases slightly for rougher terrains, which we believe to be caused by self-shadowing. However, this effect is small and does not interfere with our results.

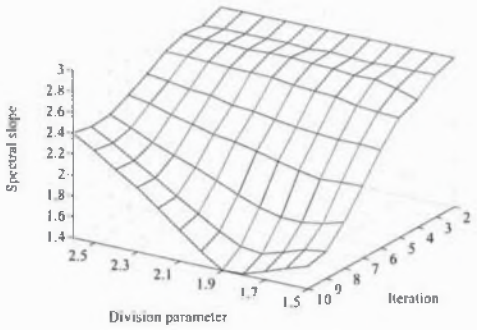

Figure 8: *The spectral slope for each iteration of the terrain generation process and for each of the 12 resulting terrains of Figure 10.*

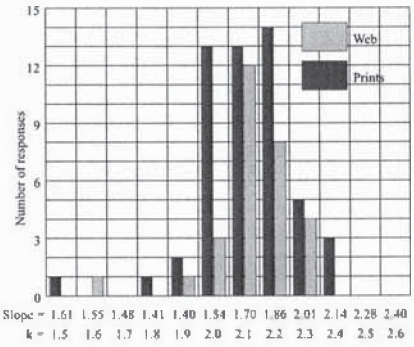

Figure 9: *Responses from 52 participants using prints and 31 participants who performed the web based experiment. The histogram shows both the spectral slope and the division parameter k associated with the images in Figure 10.*

If we were to set k automatically to produce an image with statistics tuned to the human visual system, we would use a k which produces images with spectral slopes near 1.88, as this was the value found for the natural image ensemble "A" (Section 4). This corresponds to a value of k close to 2.2. To evaluate this selection of k we asked 30 people to pick the most realistic image from Figure 10. The images, rendered at a resolution of 1024^2 pixels, were each printed on a single sheet at a resolution of 300dpi, yielding 8.67^2 cm. prints. These prints were then presented to the participants in a randomly ordered pile. We asked the subjects to select the image that looked most realistic. To make sure that geographic location did not bias the participants, the experiment was repeated using participants from both North America and Europe. This latter experiment was conducted using a web page to show the images. Again, the question asked was to selected the most realistic image.

The results of both experiments are shown in Figure 9. This histogram shows the number of responses obtained for each terrain. The peaks of the histograms lie close to $k = 2.1$ and $k = 2.2$. This corresponds to a spectral slope of between 1.70 and 1.86. The selection that our participants made correlates to the selection based on spectral analysis of the terrain images (which would be $k = 2.2$). Additionally, the mean value lies well within one standard deviation of the natural image ensemble "A" ($\alpha = 1.88 \pm 0.42$ s.d.), suggesting that people would select images with "natural" statistics if they were given the choice.

We therefore conclude that this analysis can be successfully used for parameter selection in fractal terrains. The agreement between

user and algorithmic image selection suggests that this statistical approach lends itself to wider usage in graphics applications, especially those applications where computer generated imagery needs to be evaluated for realism or parameters need to be tuned.

8 Conclusions

Because computer graphics applications should produce cues appropriate for the HVS, it is important to understand what kind of images the HVS expects. We have surveyed natural image statistics and the most salient one, $1/f^2$ power spectra, was investigated further. We established that this image statistic is sensitive to geometry, while the entire rendering pipeline seems to not significantly affect this statistic. Hence, modeling applications can benefit from applying second order statistics, as shown by an example application in the form of fractal terrain modeling. We envisage these statistical tools to be able to provide criteria for a wide range of graphics applications, as argued in the previous section.

While this paper has shown that second order statistics are useful, it is by no means the only statistic and we deem adherence to these particular statistics necessary but perhaps not sufficient for conveying information to the HVS. We therefore intend to continue this line of research by better understanding the mechanisms that cause these and other statistics.

References

[1] J. J. Atick and N. A. Redlich, *What does the retina know about natural scenes?*, Neural Computation, 4 (1992), pp. 196–210.

[2] R. J. Baddeley and P. J. B. Hancock, *A statistical analysis of natural images matches psychophysically derived orientation tuning curves*, Proc. Roy. Soc. Lond. B, 246 (1991), pp. 219–223.

[3] R. M. Balboa, C. W. Tyler, and N. M. Grzywacz, *Occlusions contribute to scaling in natural images*, Vision Research, (2001). In press.

[4] A. J. Bell and T. J. Sejnowski, *Edges are the 'independent components' of natural scenes*, Advances in Neural Information Processing Systems, 9 (1996).

[5] ———, *The independent components of natural scenes are edge filters*, Vision Research, 37 (1997), pp. 3327–3338.

[6] T. Bossomaier and A. W. Snyder, *Why spatial frequency processing in the visual cortex?*, Vision Research, 26 (1986), pp. 1307–1309.

[7] G. J. Burton and I. R. Moorhead, *Color and spatial structure in natural scenes*, Applied Optics, 26 (1987), pp. 157–170.

[8] R. L. Cook, L. Carpenter, and E. Catmull, *The reyes image rendering architecture*, Computer Graphics (SIGGRAPH '87 Proceedings), 21 (1987), pp. 95–102.

[9] M. E. Crovella and M. S. Taqqu, *Estimating the heavy tail index from scaling properties*, Methodology and Computing in Applied Probability, 1 (1999), pp. 55–79.

[10] D. W. Dong and J. J. Atick, *Statistics of natural time-varying images*, Network: Computation in Neural Systems, 6 (1995), pp. 345–358.

[11] D. J. Field, *Relations between the statistics of natural images and the response properties of cortical cells*, J. Opt. Soc. Am. A, 4 (1987), pp. 2379–2394.

[12] D. J. Field, *Scale-invariance and self-similar 'wavelet' transforms: An analysis of natural scenes and mammalian visual systems*, in Wavelets, fractals and Fourier transforms, M. Farge, J. C. R. Hunt, and J. C. Vassilicos, eds., Clarendon Press, Oxford, 1993, pp. 151–193.

[13] D. J. Field and N. Brady, *Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes*, Vision Research, 37 (1997), pp. 3367–3383.

[14] A. Fournier, D. Fussell, and L. Carpenter, *Computer rendering of stochastic models*, Communications of the ACM, 25 (1982), pp. 371–384.

[15] P. J. B. Hancock, R. J. Baddeley, and L. S. Smith, *The principle components of natural images*, Network, 3 (1992), pp. 61–70.

[16] F. J. Harris, *On the use of windows for harmonic analysis with the discrete fourier transform*, Proc. IEEE, 66 (1978), pp. 51–84.

[17] B. M. Hill, *A simple general approach to inference about the tail of a distribution*, The Annals of Statistics, 3 (1975), pp. 1163–1174.

[18] J. Hurri, A. Hyvärinen, and E. Oja, *Wavelets and natural image statistics*, in Proc. of 10th Scandinavian Conference on Image Analysis, June 1997, pp. 13–18.

[19] A. Hyvärinen, *Survey on independent components analysis*, Neural Computing Surveys, 2 (1999), pp. 94–128.

[20] M. S. Langer, *Large-scale failures of $f^{-\alpha}$ scaling in natural image spectra*, J. Opt. Soc. Am. A, 17 (2000), pp. 28–33.

[21] C. L. Nikias and A. P. Petropulu, *Higher-order spectra analysis*, Signal Processing Series, Prentice Hall, 1993.

[22] H. van Hateren and A. van der Schaaf, *Independent component filters of natural images compared with simple cells in primary visual cortex*, Proc. R. Soc. Lond. B, 265 (1998), pp. 359–366.

[23] B. A. Olshausen and D. J. Field, *Emergence of simple-cell receptive field properties by learning a sparse code for natural images*, Nature, 381 (1996), pp. 607–609.

[24] C. A. Párraga, G. Brelstaff, and T. Troscianko, *Color and luminance information in natural scenes*, J. Opt. Soc. Am. A, 15 (1998), pp. 563–569.

[25] C. A. Párraga, T. Troscianko, and D. J. Tolhurst, *The human visual system is optimised for processing the spatial information in natural visual images*, Current Biology, 10 (2000), pp. 35–38.

[26] K. Perlin, *An image synthesizer*, Computer Graphics (Proceedings of SIGGRAPH 85), 19 (1985), pp. 287–296.

[27] J. Portilla and E. P. Simoncelli, *A parametric texture model based on joint statistics of complex wavelet coefficients*, Int'l Journal of Computer Vision, 40 (2000), pp. 49–71.

[28] P. Prusinkiewicz and A. Lindenmayer, *The Algorithmic Beauty of Plants*, Springer-Verlag, 1990.

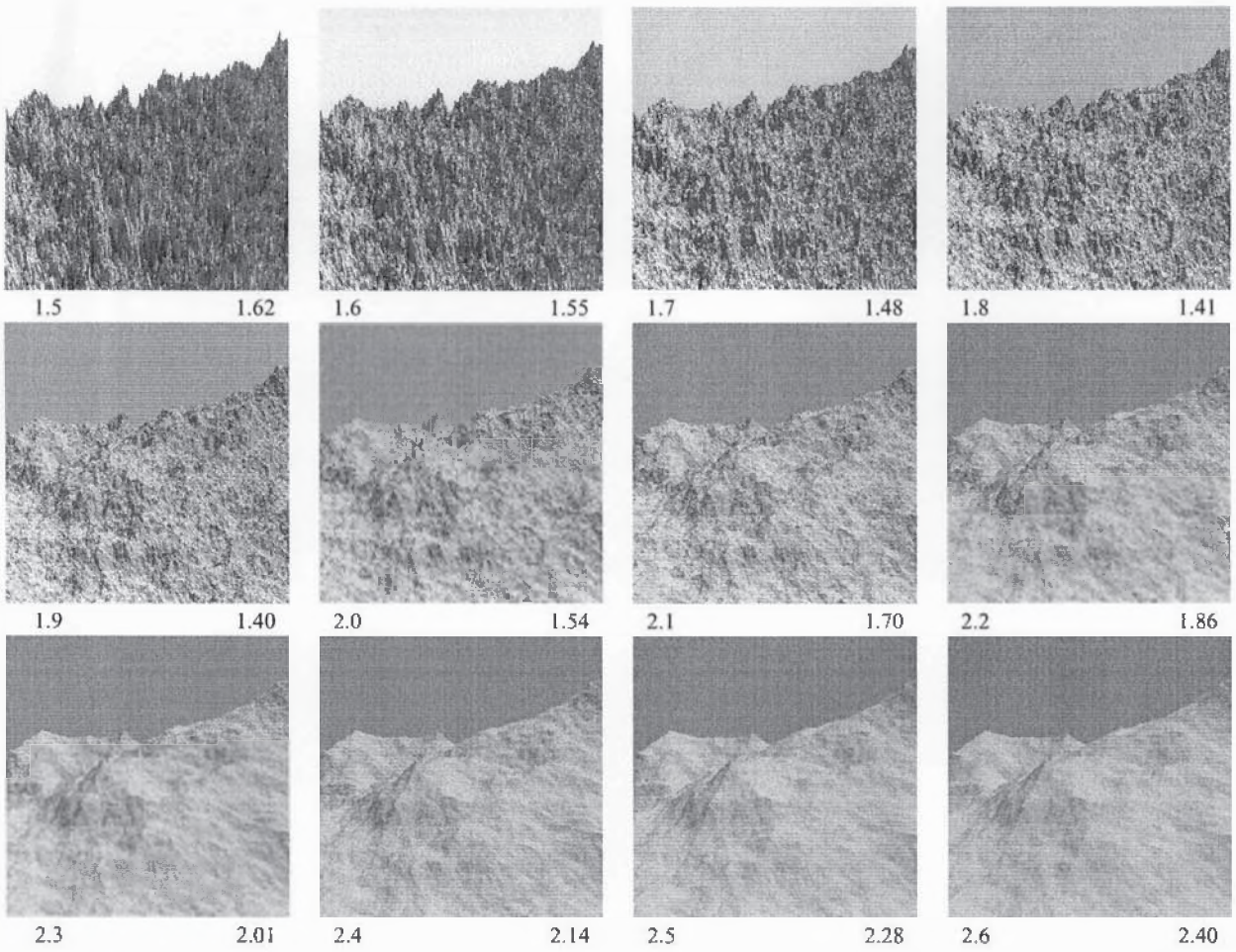

Figure 10: *Fractal terrains. The numbers for each image are the division parameter k (left) and the spectral slope of the image (right).*

[29] S. J. M. RAINVILLE AND F. A. A. KINGDOM, *Spatial-scale contribution to the detection of mirror symmetry in fractal noise*, J. Opt. Soc. Am. A, 16 (1999), pp. 2112–2123.

[30] P. REINAGEL AND A. M. ZADOW, *Natural scene statistics at the centre of gaze*, Network: Comput. Neural Syst., 10 (1999), pp. 1–10.

[31] D. L. RUDERMAN, *The statistics of natural images*, Network: Computation in Neural Systems, 5 (1997), pp. 517–548.

[32] D. L. RUDERMAN AND W. BIALEK, *Statistics of natural images: Scaling in the woods*, Physical Review Letters, 73 (1994), pp. 814–817.

[33] A. VAN DER SCHAAF, *Natural image statistics and visual processing*, PhD thesis, Rijksuniversiteit Groningen, The Netherlands, March 1998.

[34] E. P. SIMONCELLI, *Bayesian denoising of visual images in the wavelet domain*, in Bayesian Inference in Wavelet Based Models, P. Müller and B. Vidakovic, eds., vol. 141 of Lecture Notes in Statistics, New York, July 1999, Springer-Verlag, pp. 291–308.

[35] ———, *Modelling the joint statistics of images in the wavelet domain*, in Proc. SPIE 44th Anual Meeting, vol. 3813, July 1999, pp. 188–195.

[36] E. P. SIMONCELLI AND J. PORTILLA, *Texture characterization via joint statistics of wavelet coefficient magnitudes*, in Proc. 5th Int'l Conf. on Image Processing, October 1998.

[37] B. SMITS, P. SHIRLEY, AND M. M. STARK, *Direct ray tracing of displacement mapped triangles*, in Proceedings Eurographics Workshop on Rendering, Brno, Czech Republic, June 2000, pp. 307–318.

[38] M. G. A. THOMSON, *Higher-order structure in natural scenes*, J. Opt. Soc. Am. A, 16 (1999), pp. 1549–1553.

[39] ———, *Visual coding and the phase structure of natural scenes*, Network: Computation in Neural Systems, 10 (1999), pp. 123–132.

[40] D. J. TOLHURST, Y. TADMOR, AND T. CHIAO, *Amplitude spectra of natural images*, Ophthalmic and Physiological Optics, 12 (1992), pp. 229–232.

[41] G. WARD LARSON AND R. A. SHAKESPEARE, *Rendering with Radiance*, Morgan Kaufmann Publ., 1998.

[42] C. ZIEGAUS AND E. W. LANG, *Statistical invariances in artificial, natural and urban images*, Z Naturforsch, 53a (1998), pp. 1009–1021.