Second order image statistics in computer graphics

Erik Reinhard[*]
University of Central Florida

Peter Shirley[†]
University of Utah

Michael Ashikhmin[‡]
SUNY Stony Brook

Tom Troscianko[§]
University of Bristol

Abstract

The class of all natural images is an extremely small fraction of all possible images. Some of the structure of natural images can be modeled statistically, revealing striking regularities. Moreover, the human visual system appears to be optimized to view natural images. Images that do not behave statistically as natural images are harder for the human visual system to interpret. This paper reviews second order image statistics as well as their implications for computer graphics. We show that these statistics are predominantly due to geometric modeling, while being largely unaffected by the choice of rendering parameters. As a result, second order image statistics are useful for modeling applications, which we show in direct examples (recursive random displacement terrain modeling and solid texture synthesis). Finally, we present an image reconstruction filter based on second order image statistics.

1 Introduction

The scenes that we observe in daily life contain structure and content that is easy for the human visual system to interpret. Natural images appear to deviate from random images in specific ways to form a sparse subset of all randomly formed images. The statistics of natural images have been studied to understand how their properties influence the human visual system. In this paper we study how these properties may be applied in computer graphics.

We focus on second order image statistics which are explained in Sections 2 and 3. We present several practical experiments which show which aspects of modeling and rendering influence these statistics in Section 4. We then show in Section 5 how statistical knowledge of images may be applied in computer graphics. Examples include image reconstruction, parameter optimization for creating fractal terrains and Perlin noise. We draw conclusions in Section 6.

2 Image statistics

Natural image statistics can be characterized by their order. In particular, first, second and higher order statistics are distinguished [van der Schaaf 1998]. We focus on the most remarkable

---

[*]reinhard@cs.ucf.edu

[†]shirley@cs.utah.edu

[‡]ash@cs.sunysb.edu

[§]tom.troscianko@bristol.ac.uk

and salient natural image statistic that has so far been discovered, which is that the slope of the power spectrum tends to be close to negative two. This is a second order statistic. The power spectrum of an M by M image is computed with [Nikias and Petropulu 1993]:

$$S(u,v) = \frac{|F(u,v)|^2}{M^2},$$ (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 power as function of frequency f lies approximately on a straight line [Burton and Moorhead 1987; Field 1987; Ruderman and Bialek 1994; Ruderman 1997; van der Schaaf 1998]. 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 negative two for natural images:

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

Here, $\alpha \approx -2$ is the spectral slope, η is its deviation from -2 and constant A describes the overall image contrast. This result is true for ensembles, but does not always hold when applied to individual images [Langer 2000; Balboa et al. 2001].

Although this spectral slope varies subtly between different studies [Burton and Moorhead 1987; Tolhurst et al. 1992; Field 1993; Ruderman and Bialek 1994; Dong and Atick 1995], 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.

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 [Torralba and Oliva 2003]. 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 [Ruderman 1997; van der Schaaf 1998]. 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 [Nikias and Petropulu 1993]. 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 [van der Schaaf 1998]:

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

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 [Field and Brady 1997]. 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, chromatic channels are shown to exhibit similar spectral behavior [Párraga et al. 1998], and therefore all subsequent qualitative arguments are expected to be true for color as well.

In this paper we study which aspects of rendering and modeling influence second order statistics. We will show that second order statistics are influenced most by the geometry of the scene and much less by the various aspects of the rendering algorithms employed. This decoupling of sensitivity to geometry and rendering is a very fortunate result, since it opens up the possibility of using second order statistics for assessing modeling efforts without the need to pay careful attention to the details of the lighting simulation.

3 Power spectrum computation

For the computation of the spectral slope, we follow the method in [van der Schaaf 1998]. 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:

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

where $w(x,y)$ a weight factor which is explained below. Next, the images were pre-filtered to avoid boundary effects. This is accomplished by applying a circular Kaiser-Bessel window function (with parameter $\alpha = 2$) to the image [Harris 1978]:

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

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:

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

This windowing function was chosen for its near-optimal trade-off between side-lobe level, main-lobe width and computability [Harris 1978]. The resulting images were then Fourier transformed:

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

Finally, the power spectrum was computed as per equation 1 before plotting the resulting data points. Although frequencies of up to 256 cycles per image are computed, only the 127 lowest frequencies were used to estimate the spectral slope. Higher frequencies may

Figure 1: *Left: Room scene, modeled to a spatial resolution of 1 mm. Right: Scanned data consisting of over 396.000 polygons.*

suffer from aliasing, noise and low modulation transfer [van der Schaaf 1998]. 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 [Hill 1975] and the scaling method [Crovella and Taqqu 1999] to maintain compatibility with [van der Schaaf 1998]. 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 The spectral slope of rendered images

In this section we study how various aspects of modeling and rendering affect the spectral slope of the resulting images. For specific image statistics to be useful, one would like them to be variant under only a small subset of these aspects. We will show in the following that this is indeed the case. Since the power spectrum is computed on images, in the following section we first address the issue of image manipulations such as gamma correction, anti-aliasing and lossy compression. We will show that second order image statistics are largely invariant under these transformations.

Next, we consider the impact of rendering on the value of the spectral slope, showing that rendering has a negligible effect on the power spectrum. Hence we argue that geometric considerations turn out to have a profound effect, and as such it is determined that the power spectrum could be used as a modeling tool. The validity of this assessment is demonstrated in Section 5, where it is shown that certain applications can directly benefit from measuring the power spectrum.

Display We empirically evaluate which aspects of image post-processing affect the spectral slope. Possible artifacts may arise from lossy compression for file formats such as jpeg, gamma correction and aliasing. Each of these is discussed in turn, using high quality renderings created with Radiance [Ward Larson and Shakespeare 1998], as shown in Figure 1.

The lighting simulation included diffuse inter-reflection and soft shadows. The images were 512x512 pixels with 64 samples per pixel, and were subsequently converted to PPM. The following tests all pertain to these images, which have a measured spectral slope of $\alpha = -2.36$ for the room model and $\alpha = -2.15$ for the head scan.

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 JPEG (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 differ-

	Room		Head	
Super-samples	α	σ	α	σ
1x1	-2.23	0.15	-2.06	0.10
2x2	-2.32	0.19	-2.12	0.11
4x4	-2.35	0.20	-2.13	0.11
8x8	-2.36	0.20	-2.15	0.11

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.*

		Room		Head	
Size	Energy	α	σ	α	σ
1/8	64	-2.35	0.20	-2.13	0.11
1/4	16	-2.35	0.20	-2.13	0.11
1/2	4	-2.35	0.20	-2.13	0.11
1	1	-2.36	0.20	-2.15	0.11
2	1/4	-2.36	0.20	-2.16	0.11
4	1/16	-2.37	0.21	-2.17	0.11

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

ent 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.

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. 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.

Aliasing is yet 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 power spectrum computations (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. To minimize aliasing artifacts affecting the spectral slope, 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.

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 answer the question whether second order statistics apply to lighting simulation or modeling. The scenes from the previous section (Figure 1) were used for rendering images using different lighting simulations: we compare with and without diffuse inter-reflection as well as with hard versus soft 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 both scenes, shadows do not appear to be an overly important factor for computing image statistics.

Second, we have assessed the influence of diffuse inter-reflection on second order statistics. As diffusely reflecting surfaces are unlikely 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 term 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.

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 in the following section.

5 Sample applications

In this section we will first show that a measure of the power spectrum of rendered images can be used for parameter tuning in procedural modeling. We show this for both fractal terrains and Perlin turbulence applications. In addition, a user study is presented that shows that people prefer to look at images with natural statistics (i.e., with a spectral slope close to -2) when asked which image out of a set looks most realistic.

As the final sample application, reconstruction filters are re-examined. Normally, for many graphics applications, images are assumed to show band-pass behavior. In this paper, we have shown that natural images are not band-pass limited, but have an amplitude spectrum of around -1, which is equivalent to a power spectrum of approximately -2. This insight leads to a different type of image reconstruction filter, which we derive below. This filter has the desirable property that it preserves the spectral slope during filtering, which leads to a better reconstruction of images.

Procedural terrains We expect natural image statistics to be useful in any realistic modeling application that involves parameter tuning. One such application is procedural terrains where parameters affect each stage of subdivision [Mandelbrot 1983; Peitgen and Saupe 1988].

We implemented the midpoint subdivision algorithm of Fournier et al. [Fournier et al. 1982]. 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 4 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 [Ward Larson and Shakespeare 1998]. The resulting images have an approximately linear relationship between spatial frequency and spectral slope. Figure 2 demonstrates this behavior, along with the observation that for all values of parameter
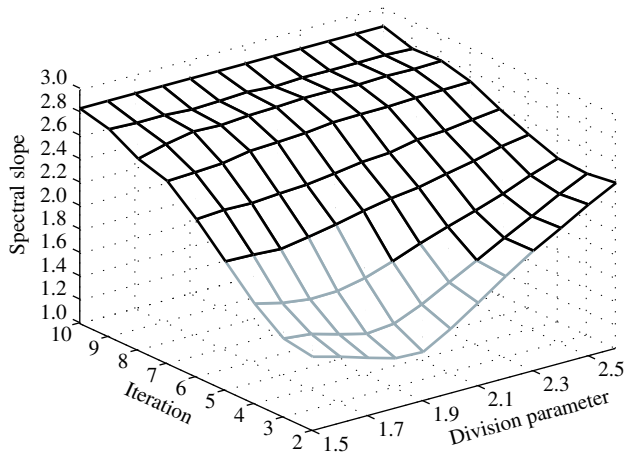
Figure 2: *The spectral slope for each iteration of the terrain generation process and for each of the 12 resulting terrains of Figure 4. The boundary between black and grey shows which parameters cause a spectral slope of around -1.87.*

Figure 3: *Responses from 52 participants using prints and 84 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 4.*

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.

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 -1.87, because this was the value found by van der Schaaf on a large ensemble of natural images [van der Schaaf 1998], and we replicated his spectral slope computation.

A spectral slope of -1.87 corresponds to a value of k close to 2.2. To evaluate this selection of k we asked 52 people to pick the most realistic image from Figure 4. The images, rendered at a resolution of 1024^2 pixels, were each printed on a single sheet at a resolution of 300 dpi, yielding 8.67^2 cm^2 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 3. 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 van der Schaaf's natural image ensemble ($\alpha = -1.87 \pm 0.43$ s.d.) [van der Schaaf 1998], 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.
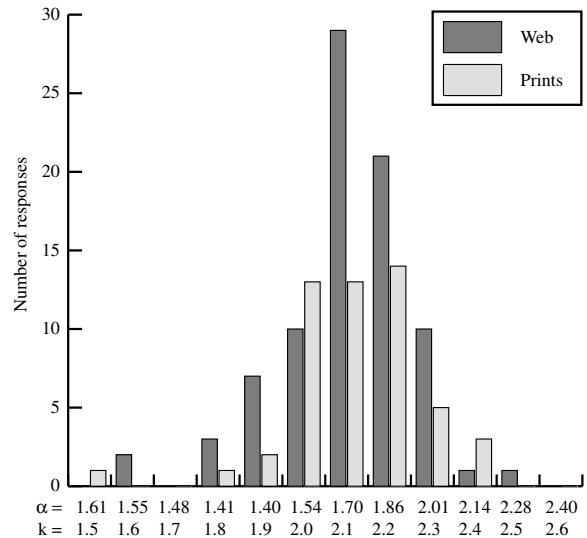
Perlin turbulence A second parameter optimization problem is encountered when generating Perlin turbulence [Perlin 1985]. The output of that method is sometimes used directly as texture and in that case it should look natural. This algorithm takes two parameters and varying these will lead to different procedural textures. Perlin turbulence is the sum of absolute Perlin noise functions that are scaled in both space and intensity:

$$\text{turbulence}(\vec{x}) = \sum_i \text{noise}(a^i \vec{x})/b^i. \tag{7}$$

Applying our spectral analysis tool directly to the textures, the influence of varying these two parameters a and b can be assessed. This is shown in Figure 5. Conventional wisdom says that the optimal choice of parameters for Perlin turbulence is around 2.0 for both of them (e.g., [Upstill 1990]). Choosing a texture with scale invariant properties, i.e. with a spectral slope of around negative two, would result in a rather different choice of parameters, as indicated by the boundary between black and grey data points in this graph. From Figure 5 we deduce that fixing parameter b to be around $1.2 - 1.3$ and keeping parameter a within a reasonable range, such as $1.0 - 4.0$, yields Perlin noise textures with natural image statistics. Examples of such parameter selections are given in Figure 6, as well as a more conventional choice of 2 for both parameters.

Choosing texture parameters such that the result has a power spectrum with a slope close to -2 implies that the texture is scale invariant and will therefore appear natural at all viewing distances. This is not the case for other textures. To confirm that this is indeed the case, two Perlin noise textures were generated at a resolution of 4096^2 pixels using the same parameters as those in Figure 6. First a window of 512^2 pixels was cut out of the middle of each and the spectral slope was computed for these windows. Then the texture was halved in size and this process was repeated a number of times, yielding a sequence of spectral slopes as given in Table 3. The texture with parameters $a = 2.8$ and $b = 1.1$ yielded similar spectral slopes for each image size, indicating scale invariance, while the other texture was not invariant to scale. This confirms that our method to assess the naturalness of textures is valid and that a Perlin noise texture with natural statistics may be more useful than other
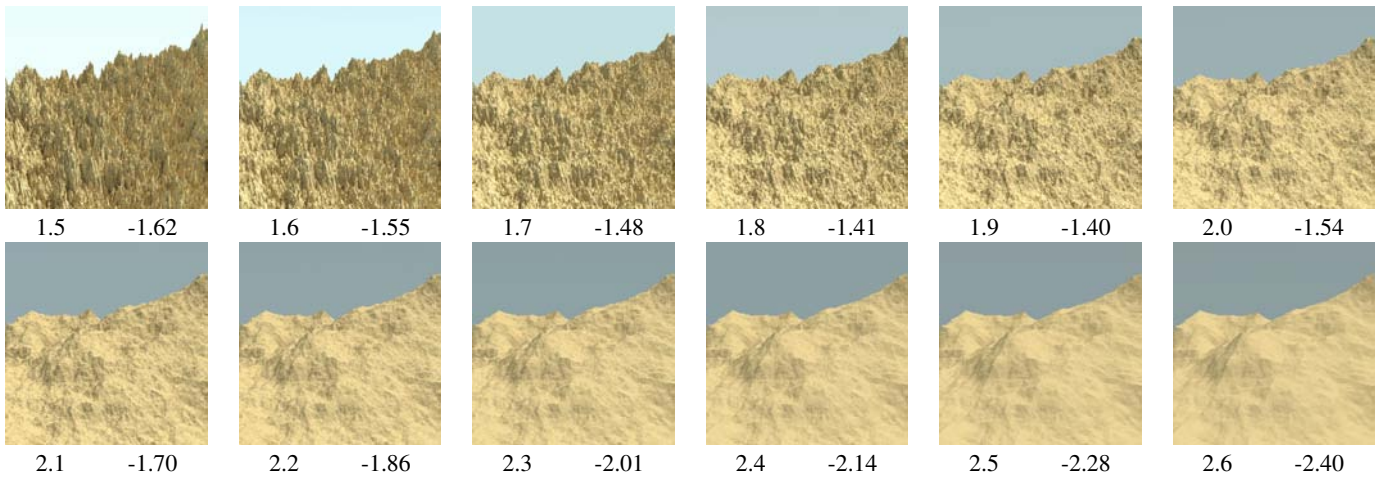
Figure 4: *Fractal terrains. The numbers for each image are the division parameter k (left) and the spectral slope of the image (right).*
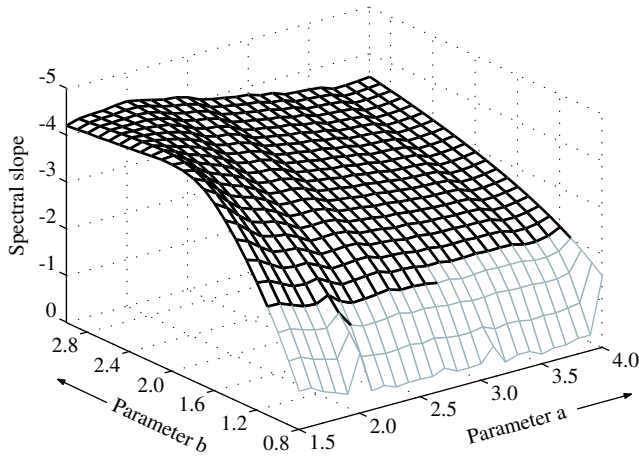

Figure 5: *Spectral slope as function of the parameters to the Perlin turbulence function. The boundary between black and grey indicates the range of parameters where a spectral slope of about -1.87 is obtained.*

Figure 6: *Perlin noise examples. Underneath each image on the left are its parameters and on the right its associated spectral slope. The bottom right image has standard parameters of 2.0 and 2.0, leading to a non-natural image statistic, while the other three yield more or less natural spectral slopes.*

similarly generated textures.

Finally, it should be noted that Perlin turbulence takes two input parameters, both of which are not directly linked to the image statistics that they yield. Alternative noise texture synthesis algorithms have exploited power spectral considerations [Lewis 1989; Heeger and Bergen 1995; De Bonet and Viola 1998], although never directly in the Fourier domain. Noise images with a given spectral slope may be constructed by randomizing both the amplitude spectrum A and phase spectrum P. Then the pixels are scaled by a factor that determines the spectral slope of the resulting image. The assignments for each pixel in the Fourier domain are as follows: $A(x,y) = r_1 f^{-\alpha/2}$, $P(x,y) = r_2 A(x,y)$, with r_1 and r_2 random variables, α the desired spectral slope and $f = \sqrt{x^2 + y^2}$ the frequency. We experimented with both uniformly distributed variables and Gaussian distributed random variables (obtained from uniformly distributed random numbers using the Box-Muller transform [Box and Muller 1958]), leading to practically identical results. The image is finally generated by taking the inverse Fourier transform, producing images as shown in Figure 7. These images
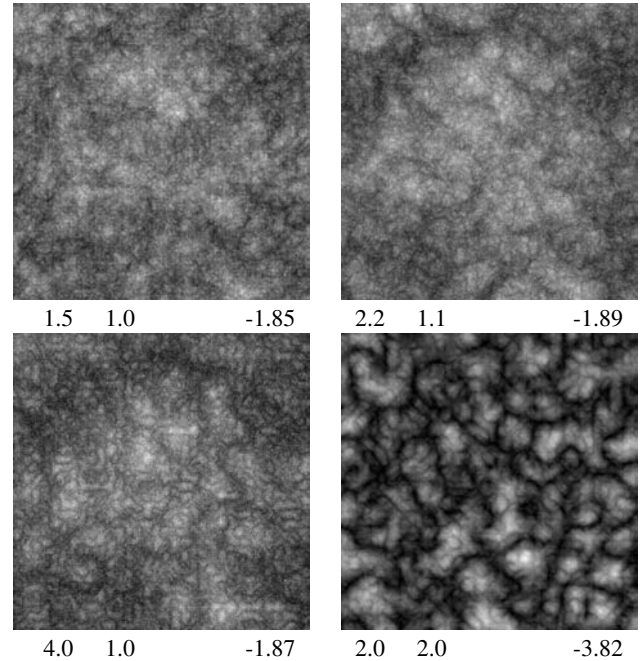
are qualitatively different from Perlin turbulence images with the same spectral slope, which is due to a different randomization of the phase spectrum (compare Figures 6 and 7).

Reconstruction Aliasing is one of the fundamental problems in image synthesis. This topic is covered in great detail in the literature (see for example [Glassner 1995]), but knowledge about the statistical nature of the imagery rarely incorporated in the design of anti-aliasing filters. A notable exception is that for the case of Gaussian stimulus ensembles signal statistics can be effectively used to recover an under-sampled signal beyond what is allowed

Image size	2.8/1.1 α	2.8/1.1 σ	2.0/2.0 α	2.0/2.0 σ
8192	1.89	0.13	3.18	0.34
4096	1.87	0.14	3.67	0.21
2048	1.88	0.13	3.86	0.15
1024	1.89	0.10	3.84	0.18
512	1.87	0.14	3.82	0.18

Table 3: *Spectral slope α and standard deviation σ as function of image size (effectively viewing distance) - results are given for two Perlin turbulence textures: with parameters $a = 2.8$ and $b = 1.1$ and with parameters $a = 2.0$ and $b = 2.0$.*

Spectral slope -1.88   Spectral slope -3.82
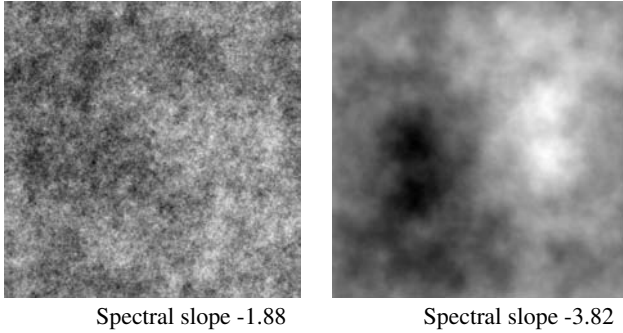
Figure 7: *Noise images produced in the Fourier domain.*

by the Nyquist theorem [Ruderman and Bialek 1992]. We leverage this idea for the case of natural images. Other domain knowledge, such as the location of edges can also be applied to image reconstruction [Li and Orchar 2001], but these methods are complicated relative to the hands-off approach that we present next.

There are two different causes for aliasing in computer graphics: pre-aliasing and post-aliasing. Pre-aliasing is due to an inadequate sampling frequency. Let a signal have Fourier transform $F(f)$. If it is sampled with rate f_s (by applying a regularly spaced sampling train), the spectrum of the sampled signal will consist of replicas of $F(f)$ positioned at frequencies kf_s, $k = 0, \pm 1, \pm 2, \dots$. If the signal is band-limited ($F(f) = 0$ for $f > f_{max}$), it can be completely reconstructed from its samples if $f_s > 2f_{max}$ by applying a box filter in frequency space (equivalent to convolution with a *sinc* function in the spatial domain), see also Figure 8 (top).

As most signals in computer graphics are not band-limited, they can not be perfectly reconstructed by this procedure. On the other hand, it is known from methods such as ray tracing that adding more samples per pixel will not visually improve the result beyond a certain point (see also Table 1). This suggests that although the signals may not be band-limited in the strict mathematical sense, in practice they can be treated as having some effective cut-off frequency. This allows us to restrict our attention to only a few central replicas of the original signal, which is a common simplification used in anti-aliasing. In fact, aliasing is usually assumed to be caused mostly by the overlap of the central copy of the signal with its two closest neighbors and, moreover, attention is restricted to the area up to and around the Nyquist frequency, $f_N = \pm 0.5 f_s$.

Post-aliasing is due to a poor choice of reconstruction filter. Different models of the underlying signal would suggest different reconstruction filters and deviation of the real signal from the model can render the suggested filter ineffective. This is for example the case for a band-limited model where the associated "ideal" sinc function is hardly ever used. We present a model based on image statistics, leading to an alternative filtering function.

We want to create a larger image while preserving $1/f$ behavior through an appropriate choice of reconstruction filter may yield vi-
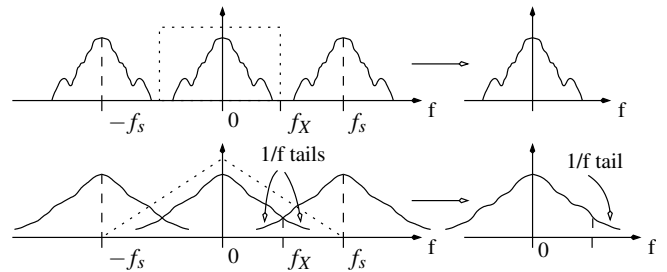

Figure 8: *Optimal filtering for a band-limited signal (top) and a natural image statistics preserving filtering (bottom). Filter functions are shown by dotted lines. Note that in the second case some distortions are introduced to the reconstructed spectrum.*

sually improved results. To derive such a filter $g(f)$, in frequency space the $1/f$ statistic should be preserved (Figure 8, bottom):

$$\frac{1}{|f|} = g(f) \left(\frac{1}{|f|} + \sum_{k=1}^{n} \frac{1}{|f \pm kf_s|} \right) \qquad (8)$$

which means that the result of applying filter $g(f)$ to the sampled signal should produce a function with $1/f$ frequency behavior. This expression is exact for $n \to \infty$, but as argued above in practice a good approximation is obtained by keeping only the first few terms. Given the shape of the power spectrum of natural images, $1/|f \pm kf_s|$ is a good general form for the copies in the region we are interested in (up to and around the Nyquist frequency). Another reason to approximate $g(f)$ is that its Fourier transform can not be expressed in closed form. Note that we use absolute values instead of the square root of the power spectrum to avoid the use of highly image-specific phase information. Function $g(f)$ and its approximations are illustrated in Figure 9. Details of the functional form of $g(f)$, such as the small deviation of the slope f from -1, as well as the exact number of terms we omit, make only minor differences in practice. Hence, we approximate $g(f)$ with a triangle in Fourier space, leading to a *sinc*^{2} function in the spatial domain:

$$g(x) = \frac{\sin^2(\pi x)}{(\pi x)^2} \qquad (9)$$

where a single pixel is assumed to have width one. Note that this result is already normalized to avoid the frequency rippling effect [Mitchell and Netravali 1988]. In 2D, we follow common practice and use a direct product $g(x)g(y)$ to perform filtering. Preserving natural image statistics during reconstruction thus leads to a *sinc*$^2(x)$ filter.

To compare this result with other commonly used filters, images were reduced in size and then different filters were applied while magnifying them back to their original size [Mitchell and Netravali 1988]. Four different reconstruction filters are compared in Figures 10 and 11: box, sinc, sinc^2 and the Mitchell filter (with optimal parameters) [Mitchell and Netravali 1988]. The *sinc*^{2} filter produces visually superior results to both box and *sinc* filters and is comparable in quality with Mitchell's filter. Note that Mitchell's filter involved a significant amount of parameter tuning while our recommendation is a direct consequence of applying domain knowledge of the spectral behavior of natural images. To ensure that the perceived quality of the *sinc*^{2} filter is not simply due to a possible increase in power or the non-negative nature of the filter, we also experimented with other variations of *sinc* and its powers. All such attempts produced unsatisfactory results. Adding noise during reconstruction would be an alternative approach to preserving the $1/f$
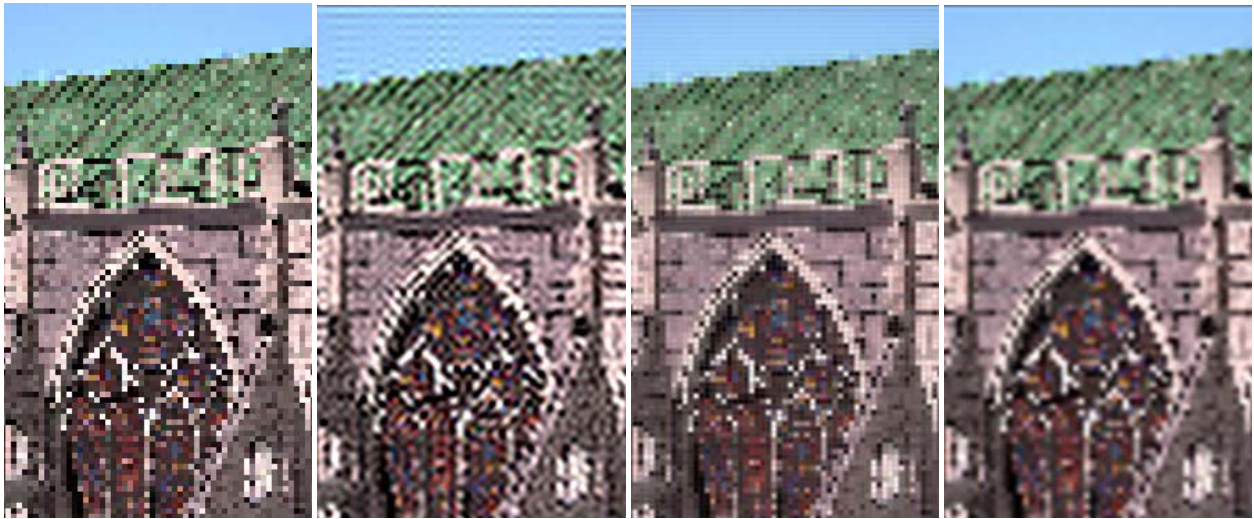
Figure 10: *Left to right: box, sinc, sinc^2 and Mitchell filters with a magnification factor of 4.*

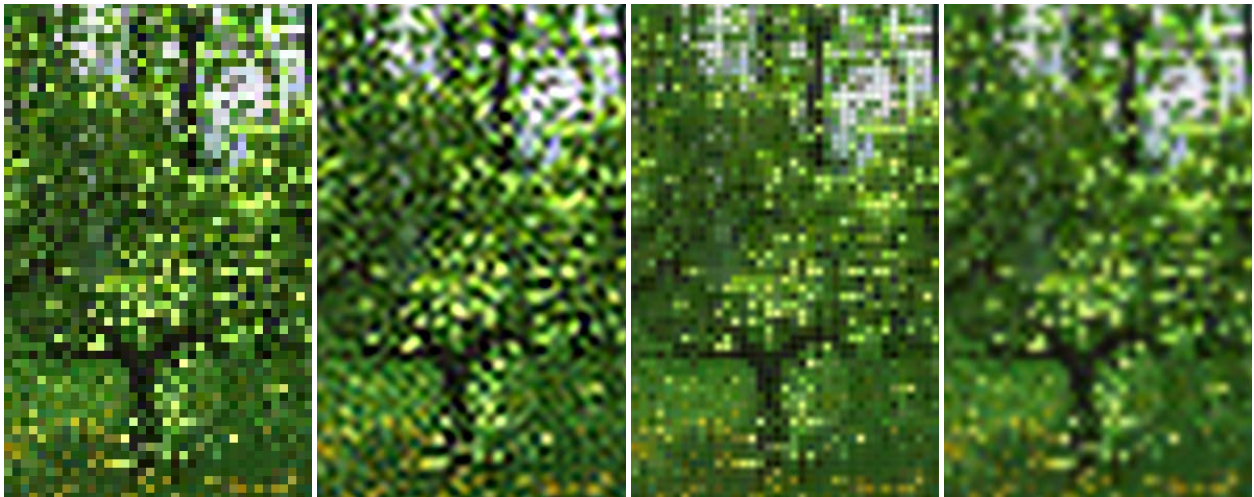Figure 11: *Left to right: box, sinc, sinc^2 and Mitchell filters with a magnification factor of 8.*
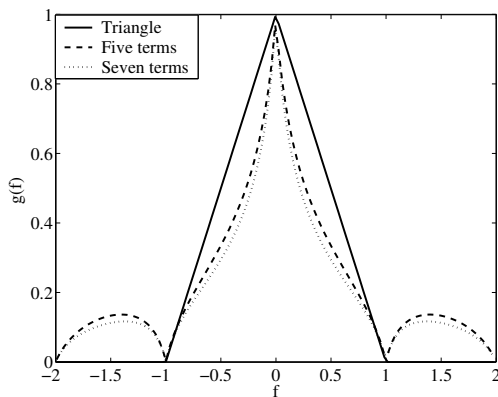

Figure 9: *Plot of $g(f)$ given by Equation 8 with five and seven terms kept and its approximation by a triangle (solid line). One unit on the frequency axis corresponds to the sampling frequency. Note that $f = \pm 0.5$ therefore corresponds to the Nyquist frequency.*

characteristic of the input signal. However, in our experience this method also produces inferior results.

The design of reconstruction filters is one of the more fundamental problems in computer graphics. Although many different factors which we do not consider here contribute to this issue (such as practical implementation problems for infinite *sinc*-style filters), our results demonstrate that image statistics can be an important factor in thinking about the fundamentals of image synthesis processes. Finally, we believe that the apparent reduction in aliasing artifacts as compared with the sinc filter is similar in nature to that described in [Ruderman and Bialek 1992].

6 Conclusions

Because computer graphics should contain cues appropriate for the HVS, it is important to understand what kind of images the HVS expects. We have surveyed the most studied natural image statistic: the $1/f^2$ power spectrum. This statistic is sensitive to geometry, while the entire rendering pipeline seems to not significantly affect this statistic. Hence, modeling applications may benefit from

applying second order statistics, as shown by two of the example applications. In addition, the theoretical importance of the power spectrum was demonstrated with reconstruction filters. We envisage these statistical tools to be able to provide criteria for a wide range of graphics applications.

While we have shown that second order statistics can be useful, it is by no means sufficient for conveying information to the HVS. While natural spectral slopes are deemed necessary to inject realism into rendered imagery, higher order statistics which are contained in the phase spectrum, require further analysis. We recognize that we have implied average characteristics of ensembles to create or modify individual images; the relationship between ensemble results and individual images deserves further study.

In addition, we would like to extend the concept of image statistics to 3-d geometry. This would allow us to perform computations directly on the geometry without having to render images first. This would remove the trial-and-error aspect now witnessed in for example the parameter optimization algorithm for fractal terrains. Currently, to establish the optimal parameters to generate a good fractal terrain, we have to render a sequence of images of different terrains and select the terrain that produced the best image. With an equivalent 3-d "geometry statistic", this could be greatly simplified.

Acknowledgments

References

BALBOA, R. M., TYLER, C. W., AND GRZYWACZ, N. M. 2001. Occlusions contribute to scaling in natural images. *Vision Research 41*, 7, 955–964.

BOX, G. E. P., AND MULLER, M. E. 1958. A note on the generation of random normal deviates. *Annals Math. Stat 29*, 610–611.

BURTON, G. J., AND MOORHEAD, I. R. 1987. Color and spatial structure in natural scenes. *Applied Optics 26*, 1 (January), 157–170.

CROVELLA, M. E., AND TAQQU, M. S. 1999. Estimating the heavy tail index from scaling properties. *Methodology and Computing in Applied Probability 1*, 1, 55–79.

DE BONET, J., AND VIOLA, P. 1998. A non-parametric multi-scale statistical model for natural images. In *Proceedings of the 1997 conference on Advances in Neural Information Processing Systems 10*, 773–779.

DONG, D. W., AND ATICK, J. J. 1995. Statistics of natural time-varying images. *Network: Computation in Neural Systems 6*, 3, 345–358.

FIELD, D. J., AND BRADY, N. 1997. Visual sensitivity, blur and the sources of variability in the amplitude spectra of natural scenes. *Vision Research 37*, 23, 3367–3383.

FIELD, D. J. 1987. Relations between the statistics of natural images and the response properties of cortical cells. *J. Opt. Soc. Am. A 4*, 12 (December), 2379–2394.

FIELD, D. J. 1993. 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, 151–193.

FOURNIER, A., FUSSELL, D., AND CARPENTER, L. 1982. Computer rendering of stochastic models. *Communications of the ACM 25*, 6 (June), 371–384.

GLASSNER, A. S. 1995. *Principles of digital image synthesis*. Morgan Kaufmann, San Fransisco, CA.

HARRIS, F. J. 1978. On the use of windows for harmonic analysis with the discrete fourier transform. *Proc. IEEE 66*, 1, 51–84.

HEEGER, D. J., AND BERGEN, J. R. 1995. Pyramid-based texture analysis/synthesis. In *Proceedings of the 22nd annual conference on Computer Graphics and Interactive Techniques*, 229–238.

HILL, B. M. 1975. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics 3*, 5, 1163–1174.

LANGER, M. S. 2000. Large-scale failures of $f^{-\alpha}$ scaling in natural image spectra. *J. Opt. Soc. Am. A 17*, 1 (January), 28–33.

LEWIS, J. P. 1989. Algorithms for solid noise synthesis. *Computer Graphics 23*, 3 (July), 263–270.

LI, X., AND ORCHAR, M. T. 2001. New edge-directed interpolation. *IEEE Transactions on Image Processing 10*, 10, 1521–1527.

MANDELBROT, B. B. 1983. *The Fractal Geometry of Nature*. W.H. Freeman and Co.

MITCHELL, D. P., AND NETRAVALI, A. N. 1988. Reconstruction filters in computer graphics. *Computer Graphics 22*, 4 (August), 221–228.

NIKIAS, C. L., AND PETROPULU, A. P. 1993. *Higher-order spectra analysis*. Signal Processing Series. Prentice Hall.

PÁRRAGA, C. A., BRELSTAFF, G., AND TROSCIANKO, T. 1998. Color and luminance information in natural scenes. *J. Opt. Soc. Am. A 15*, 3, 563–569.

PEITGEN, H.-O., AND SAUPE, D., Eds. 1988. *The Science of Fractal Images*. Springer Verlag.

PERLIN, K. 1985. An image synthesizer. *Computer Graphics 19*, 3 (July), 287–296.

RUDERMAN, D. L., AND BIALEK, W. 1992. Seeing beyond the Nyquist limit. *Neural Computation 4*, 5, 682–690.

RUDERMAN, D. L., AND BIALEK, W. 1994. Statistics of natural images: Scaling in the woods. *Physical Review Letters 73*, 6, 814–817.

RUDERMAN, D. L. 1997. The statistics of natural images. *Network: Computation in Neural Systems 5*, 4, 517–548.

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

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

TORRALBA, A., AND OLIVA, A. 2003. Statistics of natural image categories. *Network: Comput. Neural Syst. 14*, 391–412.

UPSTILL, S. 1990. *The Renderman Companion*. Addison-Wesley, Reading, MA.

WARD LARSON, G., AND SHAKESPEARE, R. A. 1998. *Rendering with Radiance*. Morgan Kaufmann Publishers.