

Statistical characterization of real-world illumination

Ron O. Dror

Massachusetts Institute of Technology, Cambridge, MA, USA



Alan S. Willsky

Massachusetts Institute of Technology, Cambridge, MA, USA



Edward H. Adelson

Massachusetts Institute of Technology, Cambridge, MA, USA



Although studies of vision and graphics often assume simple illumination models, real-world illumination is highly complex, with reflected light incident on a surface from almost every direction. One can capture the illumination from every direction at one point photographically using a spherical illumination map. This work illustrates, through analysis of photographically acquired, high dynamic range illumination maps, that real-world illumination possesses a high degree of statistical regularity. The marginal and joint wavelet coefficient distributions and harmonic spectra of illumination maps resemble those documented in the natural image statistics literature. However, illumination maps differ from typical photographs in that illumination maps are statistically nonstationary and may contain localized light sources that dominate their power spectra. Our work provides a foundation for statistical models of real-world illumination, thereby facilitating the understanding of human material perception, the design of robust computer vision systems, and the rendering of realistic computer graphics imagery.

Keywords: lighting, illumination, material perception, natural image statistics, wavelets, environment map

Introduction

Computer vision, computer graphics, and studies of human perception have traditionally relied on idealized models of illumination, such as a single point light source, a small set of point light sources, or a uniform hemispherical source. Everyday real-world illumination, on the other hand, is highly complex and variable. Surfaces are illuminated not only by luminous sources, such as the sun, sky, or indoor lights, but also by light reflected from other surfaces in the environment.

In everyday life, we are usually successful at recognizing both objects and materials across a wide range of lighting conditions. After all, if we were dependent on a particular distribution of illumination to see things correctly, then we would be in trouble when the lighting changed.

Under atypical lighting conditions, however, human perception proves much less reliable. This is particularly true for the perception of material properties, such as sur-

face reflectance. [Figure 1](#) shows a shiny white plastic scoop under two different patterns of illumination. The shape of the object is easily recognizable in both photographs, but the scoop on the left looks glossy, whereas that on the right looks matte. In the image at left, small light sources lead to sharp specular highlights, whereas in the image at right, the broad diffuse lighting prevents such highlights. Producing the photograph at right required a fair amount of effort using specialized photographic equipment, because standard extended sources like fluorescent fixtures and bounced flash still have enough structure that they produce specular cues that give a sense of gloss.

[Figures 2](#) and [3](#) provide additional examples of cases where material perception becomes difficult under atypical illumination conditions. [Figure 2](#) shows two images of the same surface. The image at left was rendered under complex real-world illumination, whereas that at the right was rendered under point source illumination. Both images include specular highlights, but the image rendered under

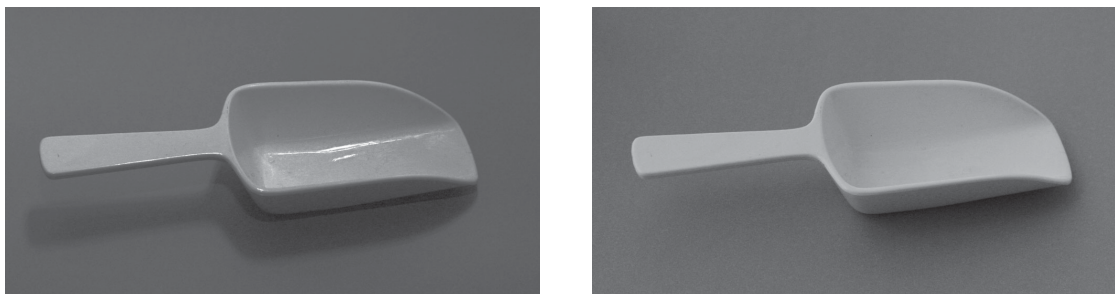


Figure 1. Two photographs of the same plastic scoop under different illumination conditions.

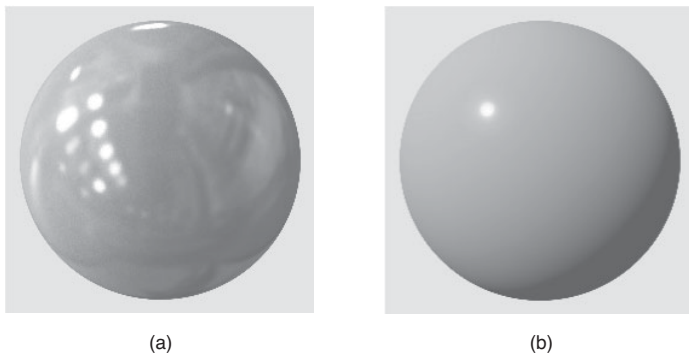


Figure 2. (a). A shiny sphere rendered under photographically acquired real-world illumination. (b). The same sphere rendered under illumination by a point light source.

realistic illumination provides a much stronger sense of the glossy reflectance than the image rendered under a point source.

Figure 3 compares a photograph of a metal sphere to a negative of the same photograph. The original photograph has the characteristic appearance of a metal sphere viewed in an everyday scene. The sphere simply produces a distorted and slightly blurred image of the world around it. The negative image could, in principle, also be a photograph of the same sphere, if it happened to be placed in a world with the appropriate distribution of light and dark. There is no physical reason why this scene could not exist, and a determined photographer could build it on purpose, but it would never occur in ordinary life. This negative image does not look like a metal sphere; in fact, it hardly

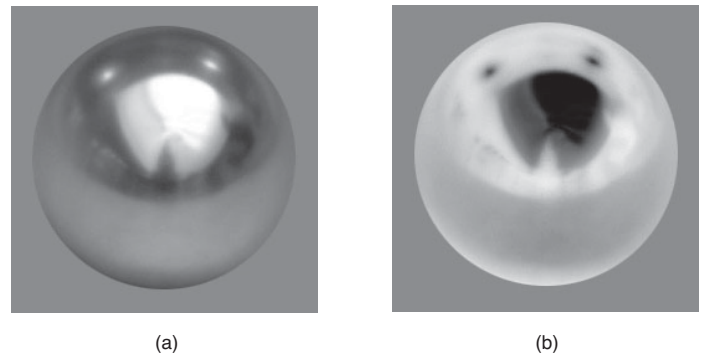


Figure 3. (a). A photograph of a metal sphere. (b). The negative of the same photograph.

looks like a realistic photograph of any ordinary sphere.

These demonstrations show that some illumination patterns lead to significant errors in human material perception. In ordinary life, however, we rarely encounter such phenomena. Figure 4 shows four spheres, each photographed in two locations. The images of different spheres in the same setting are more similar in a pixelwise sense than images of the same sphere in different settings. Yet, we easily recognize the various spheres under different everyday illumination conditions. In an experiment where subjects were asked to match reflectance properties under different illumination conditions, they consistently performed better when given two complex real-world illuminations than when given one real-world illumination and one simple synthetic illumination (Fleming, Dror, & Adelson,

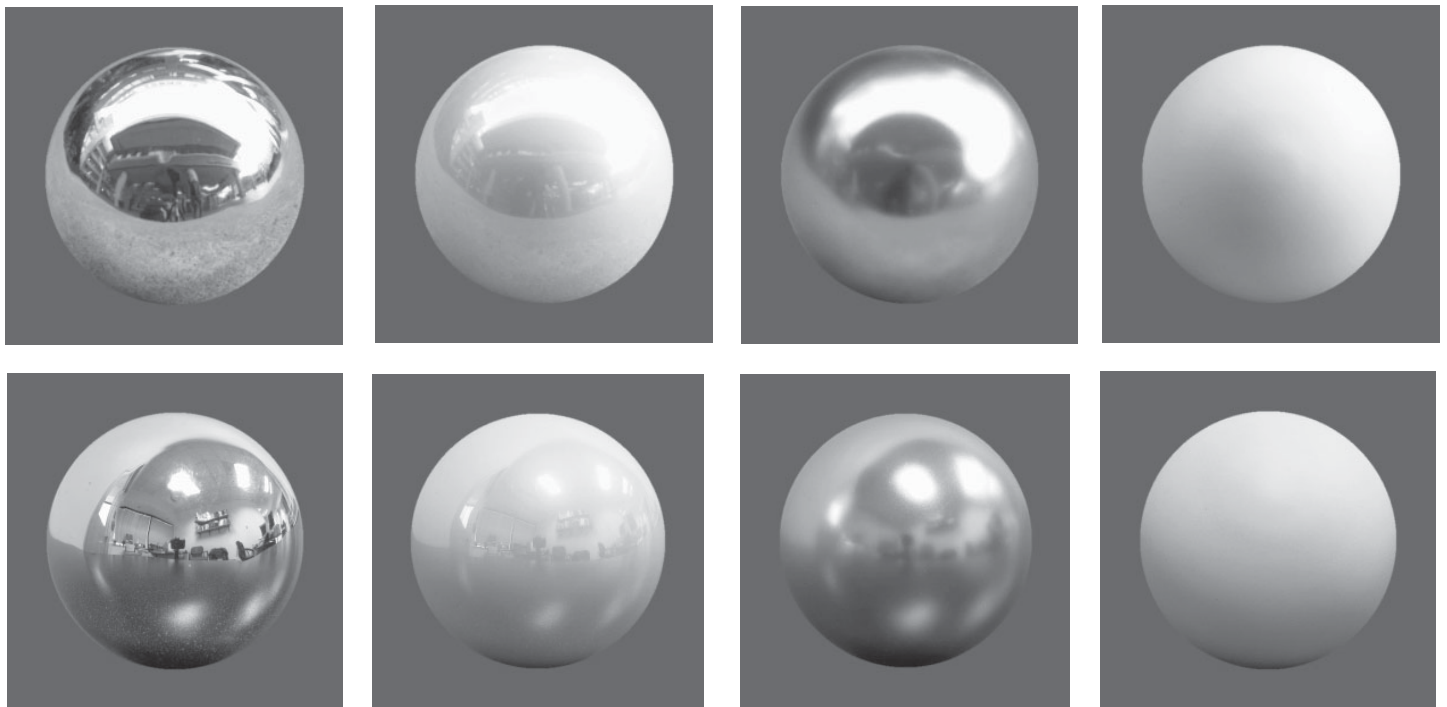


Figure 4. The two images in each column are photographs of the same sphere. The four images in each row were photographed in the same location, under the same illumination.

2003). This observation suggests that despite their complexity, real-world illumination patterns must possess stable properties that the visual system can exploit.

What, then, characterizes the patterns of illumination that occur in everyday life? These patterns of illumination are quite varied, occurring indoors and outdoors, in natural environments, such as a forest or a meadow, and in man-made environments, such as a kitchen or an outdoor plaza. Yet they all tend to share certain statistical properties, some of which are apparently used by the human visual system in estimating the reflective properties of materials. To understand material perception, we must understand what real-world illumination “looks like” (i.e., what statistical features are common for the environments we encounter in ordinary life).

Our purpose in this work is to establish the basic statistical properties of real-world illumination. We use the word “real-world” rather than “natural” to emphasize the fact that we include man-made environments. We examine the regularity and variability of real-world illumination patterns using distributions of illumination intensities (Section 3.1), spherical harmonic power spectra (Section 3.2), and distributions of bandpass filter pyramid coefficients (Section 3.3). We find widespread similarity between the statistics of real-world illumination and those previously described for photographs, with a few significant differences. Many of the statistical regularities of real-world illumination, like those of photographs and textures, can be described through marginal and joint distributions of bandpass filter coefficients. Some of these regularities correspond to intuitive notions, such as the presence of edges or bright light sources. Preliminary results of this study were presented in a conference paper (Dror, Leung, Willsky, & Adelson, 2001).

An understanding of real-world illumination statistics is important not only to elucidating the mechanisms of human perception, but also to designing computer vision systems that operate robustly in the real world. Current computer vision systems, for example, are limited in their ability to recognize materials. Recognition of surface reflectance depends heavily on illumination. Borrowing a term from estimation theory, one might view reflectance recognition as a “system identification” problem, where the input is illumination from all directions, the output is an observed image, and surface reflectance is an unknown property to be identified. To solve this problem robustly, given only the system output (the image), one must rely on predictable statistical properties of the input (the illumination). By taking advantage of the regularity of real-world illumination statistics, we have developed a system for classifying reflectance robustly under unknown everyday illumination conditions (Dror, Adelson, & Willsky, 2001; Dror, 2002).

The statistical characterization of real-world illumination finds further applications in computer graphics. The use of “natural lighting” patterns to render synthetic images is becoming increasingly common, because these renderings

appear more realistic than traditional renderings under synthetic illumination (Debevec, 1998; Hwang, 2004). Statistical properties of real-world illumination could be used to design compact representations of natural lighting patterns and efficient methods to render scenes under such lighting (Ng, Ramamoorthi, & Hanrahan, 2003). One might also use these properties to synthesize artificial illuminations that lead to realistic rendered images.

2. Methods

2.1 Measuring illumination as an image

One can measure the illumination incident from every direction at a particular point in the real world using a camera whose optical center is located at the point of interest. By combining photographs taken in different directions, one can compose a spherical map describing illumination at that point (Figure 5). Such spherical images are used as environment maps in computer graphics (Debevec, 1998). If all sources of direct and indirect illumination are relatively distant, the illumination map changes slowly as the hypothetical camera moves through space.

An illumination map is a type of image. However, accurate real-world illumination maps differ from typical photographs in several regards. First, illumination maps cover a much wider view angle, spanning the entire sphere instead of a narrow view angle near the horizontal. Second, accurate illumination maps must possess a much higher dynamic range than typical photographs to capture accurately the luminance of both the brightest and darkest areas. This is particularly true for illumination maps that contain localized primary light sources, such as incandescent lights or the sun.

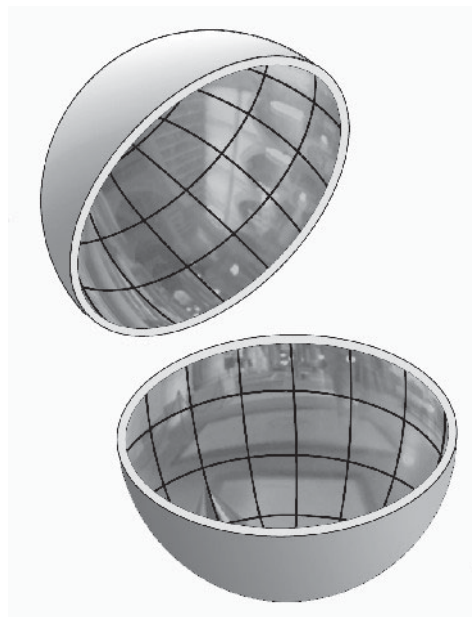


Figure 5. A photographically acquired illumination map, illustrated on the inside of a spherical shell.

A number of researchers have devoted a great deal of effort to capturing statistics of typical photographs or “natural image” statistics (Field, 1987; Tolhurst, Tadmor, & Chao, 1992; Ruderman, 1994; Huang & Mumford, 1999; Simoncelli, 1999; Buccigrossi & Simoncelli, 1999; Olshausen & Field, 2000). They have found that normal photographs of indoor and outdoor scenes display a great deal of regularity, particularly in power spectra and distributions of bandpass filter pyramid coefficients. These statistics have led to effective image denoising and compression schemes (Simoncelli & Adelson, 1996; Portilla, Strela, Wainwright, & Simoncelli, 2001; Buccigrossi & Simoncelli, 1999) as well as computational methods to recognize and synthesize texture (Heeger & Bergen, 1995; Portilla & Simoncelli, 2000), detect hidden messages in images (Farid, 2002), detect edges (Konishi, Yuille, Coughlan, & Zhu, 2003), and recognize transparency (Levin, Zomet, & Weiss, 2002). Natural image statistics have also helped explain the architecture of biological vision systems (Field, 1987; Laughlin, 1981; Simoncelli & Olshausen, 2001; Olshausen & Field, 2000). This work describes both similarities and differences between traditional natural image statistics and the statistics of illumination maps.

Whereas image statistics have previously been analyzed on a planar domain, illumination maps are naturally defined on a sphere. We found that storing illumination maps in equal-area cylindrical projection (Canters & Deleir, 1989) facilitated certain computations described in

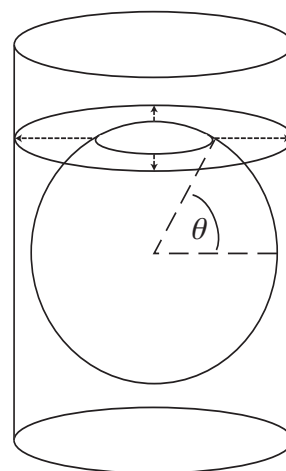


Figure 6. To produce the equal-area cylindrical projection of a spherical map, one projects each point on the surface of the sphere horizontally outward onto the cylinder, and then unwraps the cylinder to obtain a rectangular “panoramic” map.

Sections 3.1, 3.2, and 3.3. To construct this projection, one places the sphere at the center of a vertically oriented cylinder and projects each point on the spherical surface horizontally outward to the surface of the cylinder (Figure 6). One then unwraps the cylinder to obtain a rectangular map of finite extent. Regions of equal area on the sphere map to regions of equal area on the cylinder. Figure 7 displays illumination maps in equal-area projection with $k = 2/\pi$,

Teller Images



(a)



(c)

Debevec Images



(b)



(d)

Figure 7. Examples of the illumination maps we used, shown as panoramas in equal-area cylindrical projection. (a) and (c) are drawn from Teller's data set, whereas (b) and (d) are drawn from Debevec's. Dynamic range has been compressed for display purposes. The illumination map in (d) is identical to that in Figure 5.

where k is the ratio of the radius of the cylinder to the radius of the sphere. In particular, an infinitesimal patch on the sphere at latitude θ will find itself expanded by a factor of $k/\cos\theta$ in the horizontal direction and reduced by a factor of $\cos\theta$ in the vertical direction. Because the product of these two factors is a constant k , this projection preserves areas, even though it heavily distorts angles near the poles.

2.2 Data sets

We worked with two different sets of illumination maps, each consisting of high dynamic range images that represent the radiance incident at a point in the real world. The first set consisted of 95 illumination maps based on imagery acquired by Teller et al. (2001) in the environs of the MIT campus (<http://city.lcs.mit.edu/data>). The second set consisted of nine maps from Debevec's Light Probe Image Gallery (<http://www.debevec.org/Probes/>) (Debevec et al., 2000). Debevec's maps represent diverse lighting conditions from four indoor settings and five outdoor settings. Two examples from each data set are shown in Figure 7.

The images in both sets were acquired by combining photographs at multiple exposures to obtain pixel values that are linear in luminance, using the technique of Debevec and Malik (1997). We converted them all to gray-scale images with pixel values proportional to luminance. Debevec's illumination maps, which were computed from photographs of a chrome ball, cover the entire sphere. Teller's illumination maps were each mosaiced from multiple calibrated narrow-angle images. These mosaics cover the entire upper hemisphere as well as a band below the equator.

We compare our results to those of previously published studies of the statistics of traditional restricted-angle photographs. Huang and Mumford (1999) performed a number of statistical analyses on a particularly large set of images, consisting of 4000 photographs collected and calibrated by van Hateren and van der Schaaf (1998). These images were collected outdoors, but include photographs of buildings and roads as well as more "natural" scenes. Other image sets, such as that of Tolhurst et al. (1992), include indoor images.

3. Results

3.1 Illumination intensity distribution

3.1.1 Marginal distribution of intensity

Although light is typically incident on a real-world surface from every direction, the strongest illumination usually comes from primary light sources in a few directions. To quantify this intuition, we examined the marginal distribution of illumination intensity for our sets of illumination maps. This distribution is effectively just a histogram of pixel values. To compute it accurately, we must take into account the solid angle corresponding to each pixel of the

illumination map. For an equal-area projection, this solid angle is constant, so we can compute the marginal distribution of illumination intensities with an unweighted pixel histogram.

Figure 8 shows total illumination intensity distributions for the 95 Teller images and for the 9 Debevec images. Panels (a) and (b) show the distribution of linear luminance values, whereas panels (c) and (d) show the distribution of log luminance values. The linear luminance distribution plots reveal the general trend we expect — a majority of pixels at low intensity, with a heavy positive tail corresponding to pixels of much higher intensities. A typical digital photograph stored in 8-bit format necessarily lacks this heavy positive tail due to limited dynamic range.

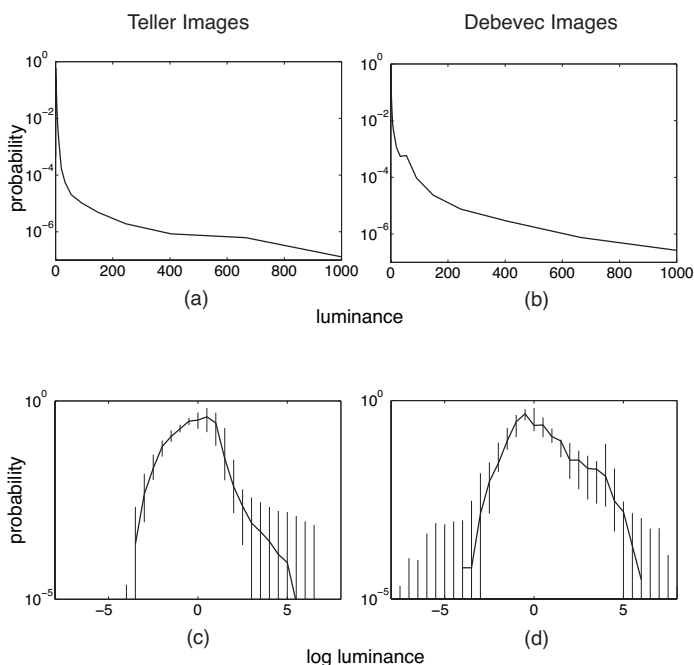


Figure 8. Illumination intensity distributions. (a) and (b) show mean histograms of linear luminance values for the 95 Teller images and the 9 Debevec images, respectively. (c) and (d) show median histograms of natural log luminance values for the two image sets. The vertical bars extend from the 20th percentile to the 80th percentile of the distribution values over the image set. For all analysis in Section 3.1, the pixel values in each image were scaled linearly before analysis such that their mean log value was 0 (i.e., such that their geometric mean was 1).

The log luminance histograms of Figure 8(c) and 8(d) show that a majority of pixels fall near the mean log luminance, with a smaller proportion of particularly dark or bright pixels. Huang and Mumford (1999) attributed the asymmetry in the distribution of log luminance values for the 12-bit images they analyzed to the presence of sky in many of their images. Our distributions exhibit more striking asymmetries, partly because both the Teller and Debevec data sets contain not only sky but also more localized

light sources. The distribution for the Teller set is particularly asymmetric due to the presence of the sun in many images and to underexposure in the imaging system at very low intensities.

The distribution of log luminance values for the Teller image set has SD $\sigma = 1.04$, kurtosis¹ $k = 4.04$, and differential entropy² $H = 2.06$ bits. The Debevec image set has $\sigma = 1.32$, $k = 12.49$, and $H = 2.21$ bits. Huang and Mumford found $\sigma = 0.79$, $k = 4.56$, and $H = 1.66$ bits. The kurtosis values are influenced heavily by individual outliers. The SD s and entropies of the distributions are higher for our data sets than for those of traditional photographs, due to the higher dynamic range and the presence of concentrated illumination sources.

Despite the aforementioned overall trends, intensity distributions vary a great deal from one illumination to the next. The degree of variation in the distribution between images is summarized by the vertical lines in Figure 8(c) and 8(d), which extend from the 20th percentile to the 80th percentile of the distribution values over all the images. Table 1 provides summary statistics on the SD , kurtosis, and differential entropy of log luminance values for individual images in each data set. Kurtosis varies more from one image to another than SD and differential entropy.

	Teller Images			Debevec Images			
	σ	k	H	σ	k	H	
Mean	1.02	5.15	1.64	Mean	1.27	8.83	1.90
SD	0.21	4.20	0.33	SD	0.39	6.82	0.39
Min	0.57	1.69	0.80	Min	0.73	2.26	1.30
Max	1.81	19.88	2.43	Max	1.82	21.46	2.44

Table 1. Statistics on the distribution of log luminance values in individual images in each data set. The columns correspond to SD (σ), kurtosis (k), and differential entropy (H) of pixel values for an individual image. The rows correspond to the mean, SD , minimum, and maximum of that image statistic across all images in the data set.

3.1.2 Nonstationarity

Most researchers in image processing treat images as samples of a stationary statistical process. That is, they assume that all parts of the image possess identical statistical properties; therefore, they treat each part of the image in the same way. Illumination maps clearly violate this stationarity assumption, if only because primary light sources, such as the sun, sky, and indoor lights, are more likely to appear in the upper hemisphere. Illumination maps with randomized orientation would, of course, be stationary, but in practice their orientation is not arbitrary; human and machine vision systems typically know which way is up.

Figure 9(a) and 9(b) show mean luminance as a function of elevation for the two data sets. As expected, illumination generally increases with elevation. Interestingly, the mean intensity reaches a local minimum at the horizontal

view direction. Both data sets contain illumination maps in which the ground reflects a significant amount of light from above, whereas visible surfaces in the horizontal direction are shadowed [e.g., Figure 7(b)]. Torralba (A. Torralba, personal communication, August, 2001; 2001) observed that images of large-scale scenes viewed from a horizontal direction also have nonstationary means. He aligned large sets of images with respect to a feature of interest, such as a person, and averaged the images within each set pixelwise to obtain “average images,” such as that shown in Figure 10. In most outdoor urban and natural settings, the average images exhibit a dip in intensity near the horizon (A. Torralba, 2001), similar to the dip we observed for illumination maps in Figure 9(a) and 9(b).

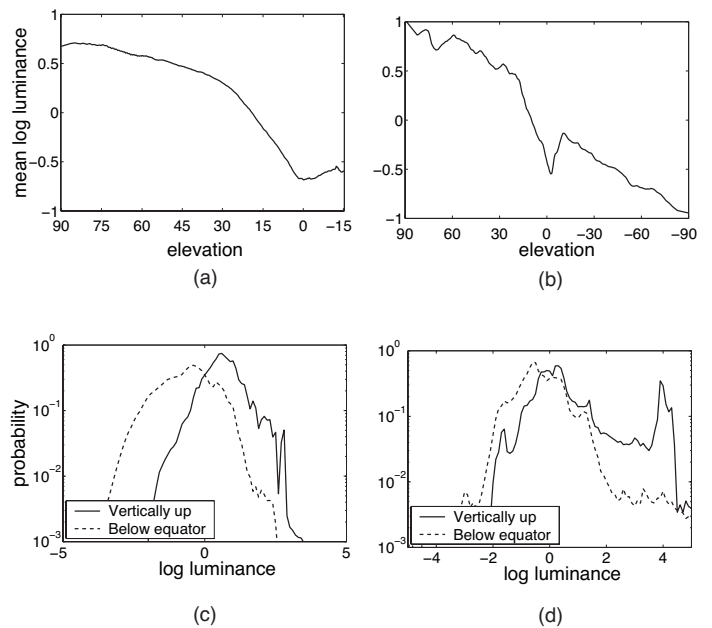


Figure 9. Dependence of illumination on elevation. (a) and (b) show mean log luminance as a function of elevation. (c) and (d) each show two histograms of illumination intensities, one for directions within 30° of the upward vertical and the other for directions from 0° to 15° below the equator. Images were normalized as in Figure 8.

Figure 9(c) and 9(d) each show two illumination intensity histograms at different ranges of elevations. The marginal distributions for higher view directions have a larger mean as well as heavier positive tails, reflecting the larger probability of bright localized sources at higher elevations.

3.1.3 Joint distribution of illumination from adjacent directions

To describe the spatial structure of real-world illumination maps, we must use statistics that depend on joint distributions of multiple pixels. The simplest way to do this is to examine the empirical joint histograms of pairs of pixels



Figure 10. This image, described by Torralba, represents the pixelwise mean of over 300 images of outdoor scenes containing a person whose head spans approximately two pixels. The images are aligned with respect to the person's head before averaging, so that a humanlike shape is visible in the center. The remainder of the average image is of nonuniform intensity, with increased intensity near the top of the image and a noticeable dip in intensity near the horizon. Reprinted from A. Torralba and P. Sinha (2001) with author's permission.

with some specific spatial relationship. Figure 11 shows contour plots of the joint histograms of horizontally adjacent pixels from all of the Teller illumination maps and from all of the Debevec maps. We define the horizontal direction in the global coordinate frame such that “horizontally adjacent” pixels lie along the same line of latitude. We divide each line of latitude into 512 “adjacent” pixels.

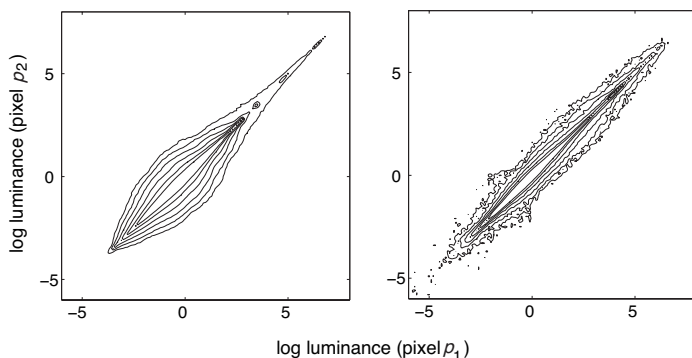


Figure 11. Joint histograms of log luminance at horizontally adjacent pixels p_1 and p_2 in the Teller images (left) and Debevec images (right). Images were normalized as in Figure 8.

Requiring that each pixel pair be separated by a fixed distance on the sphere results in virtually identical histograms.

Figure 11 shows that log luminance values at horizontally adjacent pixels p_1 and p_2 are highly correlated. Much of the mass of the joint histogram concentrates near the diagonal where $p_1 = p_2$. In agreement with Huang and Mumford (1999), we found that $p_1 + p_2$ and $p_1 - p_2$ are more nearly independent than p_1 and p_2 . In particular, the mutual information³ of p_1 and p_2 is 2.41 bits for the Teller images and 3.25 bits for the Debevec images, whereas that of $p_1 + p_2$ and $p_1 - p_2$ is only 0.10 bits for the Teller images and 0.07 bits for the Debevec images. Hence, the percentage difference between the luminance incident from two horizontally adjacent spatial directions is roughly independent of the mean luminance from those two directions.

The variability of marginal pixel histograms from image to image leads to variability in the joint pixel histogram from image to image. The ensemble pixel histograms of Figure 11 also vary between the two data sets. In both panels of Figure 11, the increased extent of the joint distributions in the upper right quadrant compared to the lower left reflects the asymmetry of the marginal distribution illustrated in Figure 8.

The utility of joint pixel histograms for examining spatial illumination structure is limited by the difficulty of visualizing joint histograms of three or more pixels. In addition, the histograms vary from one illumination map to another. We wish to identify the statistical regularities in illumination. We therefore turn to two image-processing techniques that have formed the basis for statistical characterization of spatial properties of natural images frequency domain analysis and bandpass filter pyramid analysis.

3.2 Spherical harmonic power spectra

Much early work on natural image statistics focused on the regularity of their power spectra. A number of authors (Field, 1987; Tolhurst et al., 1992; Ruderman, 1994) have observed that two-dimensional power spectra of natural images typically fall off as $1/f^{2+\eta}$, where f represents the modulus of the frequency and η is a small constant that varies from scene to scene. A power spectrum of this form is characteristic of self-similar image structure. If one zooms in on one part of the image, the power spectrum will typically change only by an overall scaling factor.

The natural spherical equivalent of the planar Fourier transform is a spherical harmonic decomposition. The spherical harmonics form an orthonormal basis for square integrable functions on the sphere. Associated with each basis function is an order L , a non-negative integer analogous to frequency. The $2L + 1$ spherical harmonics of order L span a space that is closed under rotation (Inui, Tanabe, & Onodera, 1996).

Just as planar white noise has a flat two-dimensional power spectrum, white noise on the sphere has equal power in every spherical harmonic. Similarly, if the self-similarity properties observed in the natural image statistics literature

carry over to spherical illumination maps, the average power of the spherical harmonics at order L will fall off as $1/L^{2+\eta}$.

We computed spherical harmonic coefficients for the illumination maps in both data sets using the formulas given by Inui et al. (1996). We obtained average power at each order L as the mean of squares of the coefficients at that order. Teller's data lack information about the lowest portion of the illumination hemisphere. We applied a smooth spatial window to these illumination maps before transforming them to the spherical harmonic domain.

Figure 12 shows the relationship between average power and harmonic order for the four illumination maps of Figure 7, when pixel value is proportional to log luminance. All four images have power spectra that lie close to a straight line of slope -2 on log-log axes, corresponding to a power spectrum of the form k/L^2 . We fit a straight line on log-log axes to the power spectrum of each image in the Teller data set. The best-fit lines had slopes ranging from -1.88 to -2.62 , with a mean of -2.29 . All 95 regressions gave R-square values of at least 0.95, with 86 of them above 0.97 and a mean R-square value of 0.98, indicating excellent fits. When we fixed the slope to -2 in all regressions, we also found good fits, with a minimum R-square value of 0.93 and a mean of 0.96. Fixing the slope to -2.29 gave a minimum R-square value of 0.91 and a mean of 0.98.

We obtain qualitatively different results for the same illuminations when we compute power spectra for illumination maps whose pixel values are linear in luminance. Illumination maps that lack concentrated primary light sources, such as those of Figure 7(a) and 7(b), have spherical harmonic spectra that are well approximated by $k/L^{2+\eta}$ with η small. On the other hand, illumination maps that contain intense, localized light sources have smooth power spectra that remain flat at low frequencies before falling off at higher frequencies. The illuminations of Figure 7(c) and 7(d) both display this behavior; the power spectrum of a linear luminance version of Figure 7(c) is shown in Figure 13. In these images, one or a few luminous sources, such as the sun or incandescent lights, dominate the power spectrum. Because these light sources approximate point sources, their spectra are flat at low frequencies. If one clips the brightest pixel values in these images, the power spectra return to the familiar $k/L^{2+\eta}$ form (Figure 13).

Figure 14 shows the mean spherical harmonic power spectrum of all the illuminations in the Teller data set, with vertical bars indicating the variability from one image to another. Panels (a) and (b) represent the spectra of linear luminance images, whereas (c) represents the spectra of log luminance images, and (d) represents the spectra of images where the brightest pixel values have been clipped. In panel (a), the images were normalized to have identical mean luminance values before computation of the power spectra. The power spectra exhibit a great deal of variability, but this results predominantly from differences in the total

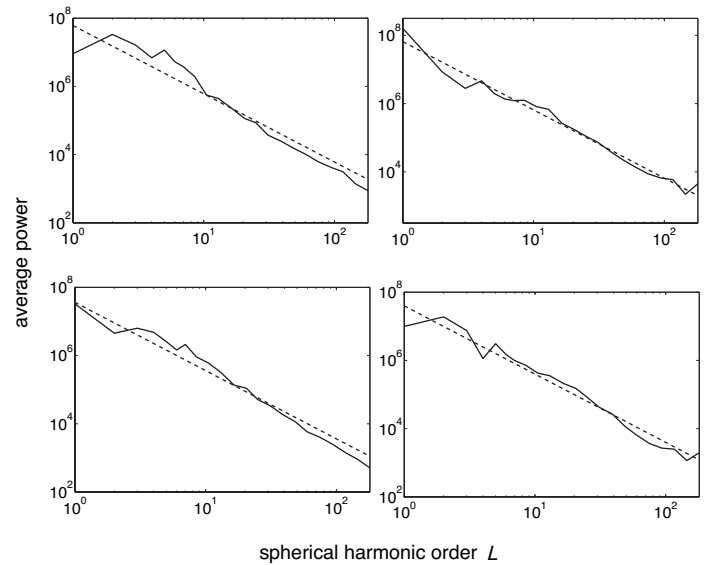


Figure 12. Spherical harmonic power spectra (solid lines) of illumination maps (a), (b), (c), and (d) in Figure 7 with pixel value proportional to log luminance. Each data point represents the average power of an interval containing one or more discrete frequencies, with the intervals approximately equally spaced on log axes. The dotted lines of slope -2 correspond to power spectra of the form k/L^2 .

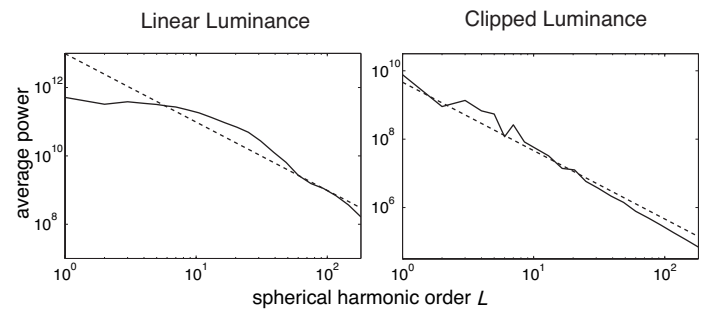


Figure 13. Left. The spherical harmonic power spectrum of the illumination map in Figure 7(c), with pixel values linear in luminance. Right. The corresponding spectrum after the pixel values corresponding to the sun have been clipped to a luminance value only slightly greater than that of the sky. Clipping these extremely bright pixels reduces power at all frequencies and produces a more linear power spectrum. The dotted lines of slope -2 correspond to power spectra of the form k/L^2 .

variance (power) of the different images. If the images are normalized for total variance instead, the variability of the power spectra decreases. The error bars are still quite large at low frequencies, however, because images dominated by one or a few point sources have flat power spectra at low frequencies. Clipping the brightest luminances or log transforming the image leads to more regularly shaped power spectra, as indicated by the smaller error bars of (c) and (d).

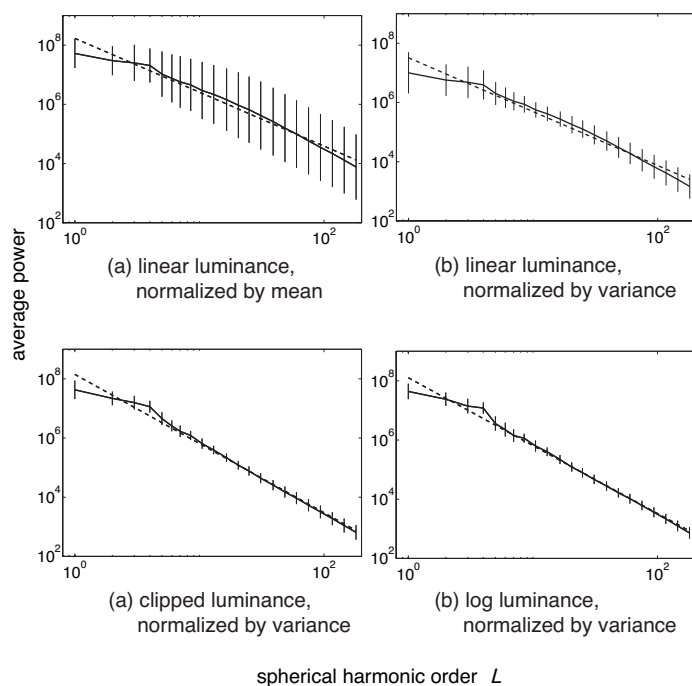


Figure 14. Mean power spectra of the 95 Teller images. Heavy solid lines indicate the mean of the individual power spectra at each spherical harmonic order, whereas each vertical bar extends both above and below this line by one SD . The power spectra of (a) and (b) were computed on images whose pixel values were linear in luminance. In (a), images were scaled to have the same mean, whereas in (b), images were scaled to have the same pixelwise variance (i.e., the same total non-DC power). In (c), power spectra were computed for “clipped” images, which were linear in luminance up to a ceiling value slightly greater than the typical luminance of the sky. The power spectra of (d) were computed for log luminance images. The images of (c) and (d) were scaled to have the same variance. The dotted lines are best-fit lines corresponding to power spectra of the form $k/L^{2+\eta}$, where η is -0.18 in (a) and (b), 0.34 in (c), and 0.29 in (d). Each point on the heavy solid curve represents the average power of an interval containing one or more discrete frequencies. Note that the vertical lines are not traditional error bars, because they represent SD rather than SEM . These SD s were computed on log power values.

Previous work on natural images has reported $1/f^{2+\eta}$ power spectra whether pixel values are linear or logarithmic in luminance (Ruderman, 1994). These results on linear luminance images differ from ours because most previous researchers have avoided photographs of point-like luminous sources and have used cameras of limited dynamic range, such that a few maximum intensity pixels could not dominate the image power spectra. A natural illumination map, on the other hand, may be dominated by light sources occupying a small spatial area. Once the relative strength of such sources is reduced through clipping or a logarithmic transformation, illumination maps have power spectra similar to those of typical photographs.

3.3 Bandpass filter pyramid statistics

The fact that a single bright source can dominate the power spectrum of an illumination map represents a shortcoming of frequency domain analysis. Multiscale bandpass filter pyramids, such as wavelets, allow a more localized analysis; a single point-like source will affect only a few wavelet coefficients. Indeed, such analysis forms the basis for most recent work in the natural image statistics literature (Ruderman, 1994; Simoncelli & Olshausen, 2001; Wainwright, Simoncelli, & Willsky, 2001). The distributions of pyramid coefficients at various scales and orientations capture not only power spectral properties, but also the non-Gaussian nature of real-world images. These distributions tend to be highly kurtotic, with many small coefficients and a few larger ones, indicating that bandpass filter pyramids provide a sparse representation of natural images. The scale-invariant properties of natural images translate into predictable relationships between pyramid coefficient distributions at different scales. The regular nature of these distributions facilitates image denoising (Portilla et al., 2001; Simoncelli & Adelson, 1996), image compression (Buccigrossi & Simoncelli, 1999), and texture characterization (Heeger & Bergen, 1995; Portilla & Simoncelli, 2000), and has also proven useful in understanding neural representations in biological visual systems (Simoncelli & Olshausen, 2001; Schwartz & Simoncelli, 2001).

Previous analysis of natural images and textures has assumed that the data are defined on a planar domain. Because illumination maps are defined as functions of orientation, they are most naturally analyzed in a spherical domain. To this end, we utilized the spherical wavelet framework introduced by Schröder and Sweldens (1995). These transforms operate on data defined on a subdivided icosahedron whose vertices are quasi-regular on the surface of the sphere. Such transforms are known as second-generation wavelet transforms because the basis functions are not exact translates and dilates of a single function (Schröder & Sweldens, 1995). We used a transform described by Amaratunga and Castrillon-Candas (2001), based on second-generation wavelets with vanishing zero-order moments and approximately vanishing first-order moments. These wavelets are constructed from simple hat functions using a linear lifting scheme.

Figure 15 shows marginal distributions of spherical wavelet coefficients at three successive scales for the 95 Teller images. The distributions are highly kurtotic, with the great majority of coefficients near zero and a few much larger coefficients. Figure 16 summarizes the variation from image to image for the distribution at one scale, for both linear luminance and log luminance images. The distributions are remarkably similar from one image to another, although the distributions associated with the linear luminance images exhibit variations in the overall scale of the wavelet coefficient distribution. The sun and other bright localized sources that dominate the entire power spectra of some of the illumination maps (Section 3.2) have a less no-

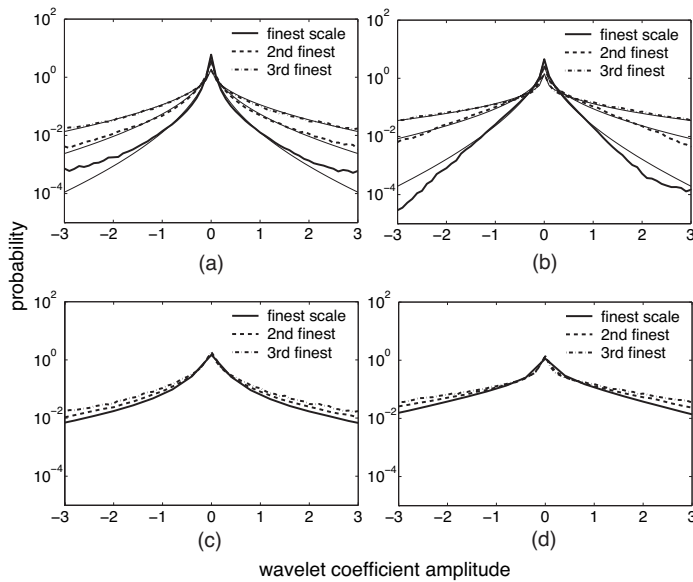


Figure 15. Distributions of spherical wavelet coefficients at successive scales (thick lines), along with generalized Laplacian fits [thin lines in (a) and (b)], for the 95 Teller images. In (a) and (b), as elsewhere in this work, the spherical wavelet basis functions are normalized to have identical power at every scale. In (c) and (d), their amplitudes are divided by 4 at the finest scale and by 2 at the next finest scale. (a) and (c) were computed on images whose pixel values were linear in luminance, whereas (b) and (d) were computed on log luminance images. The α parameters of the generalized Laplacian fits ranged from 0.50 to 0.52 for the linear luminance images, and from 0.41 to 0.59 for the log luminance images. We windowed the illumination maps as described in Section 3.2 before computing the wavelet transform, and discarded wavelet coefficients corresponding to the absent portions of the illumination map. We divided each linear luminance image by its mean before computing wavelet coefficients.

ticeable effect on the distributions of wavelet coefficients because they influence only a handful of wavelet coefficients. The variance of wavelet coefficients at a particular scale provides a measure of spectral power in some frequency band. A single localized light source can greatly influence this variance by contributing a few large outlying wavelet coefficients. However, it will have a relatively small effect on the shape of the histogram.

Several authors have observed that generalized Laplacian distributions of the form $P(x) \propto \exp(-|x/s|^\alpha)$ accurately model the wavelet coefficient distributions of typical photographs and of ensembles of photographs (Buccigrossi & Simoncelli, 1999; Huang & Mumford, 1999). Panels (a) and (b) of Figure 15 show maximum likelihood fits of this form to the ensemble histogram of wavelet coefficients from the Teller images. The fits are reasonably accurate, although they tend to underestimate the actual distribution for high wavelet coefficient magnitudes. We observed similar behavior for fits to empirical wavelet coefficient distri-

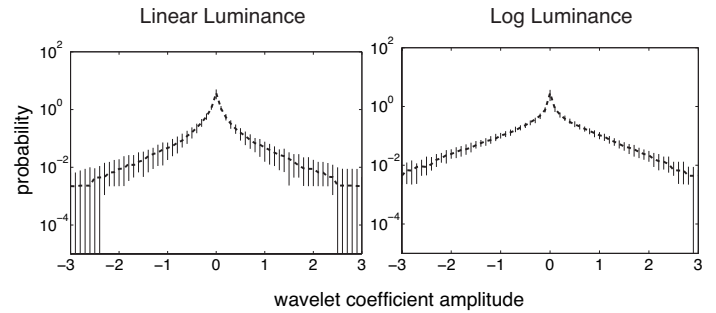


Figure 16. Variation in marginal distributions of wavelet coefficients from one image to another, for the second-finest scale band of Figure 15. The heavy dashed lines indicate the median of the histogram values across the 95 images. The vertical bars extend from the 20th percentile to the 80th percentile of the distribution values across images. We divided each linear luminance image by its mean before computing wavelet coefficients but did not normalize either linear or log luminance images for variance.

butions for individual illumination maps. This discrepancy from results reported in the natural image statistics literature may be due to the higher dynamic range of the illumination maps we analyzed.

The wavelet coefficient distributions of Figure 15 also exhibit evidence of scale invariance in illumination maps. Distributions of coefficients at different scales are similar apart from an overall normalization constant. Scale invariance requires that all statistics computed on an ensemble of images $I(\mathbf{x})$ be identical to those computed on normalized, rescaled versions of the images $\beta^\nu I(\beta \mathbf{x})$, where the exponent ν is independent of the scale β (Ruderman, 1994). An exponent $\nu=0$ leads to two-dimensional power spectra of the form $1/f^2$, where f is the modulus of frequency. More generally, a nonzero exponent ν leads to power spectra of the form $1/f^{2-\nu}$. For a scale-invariant image ensemble, the variance of wavelet coefficient distributions will follow a geometric sequence at successively coarser scales. If the wavelet basis is normalized such that wavelets at different scales have constant power, as measured by the L^2 norm, then the variance will increase by a factor of $2^{2+\nu}$ at successively coarser scales. If we increase the amplitude of the basis functions by a factor of 2 at each coarser scale, then the variance of the coefficients will increase by a factor of only 2^ν at successively coarser scales. Panels (c) and (d) of Figure 15 illustrate the results of such rescaling. Because ν is small, the distributions change little from one scale to the next. Note that linear-luminance illumination maps are not strictly scale invariant, as evidenced by the fact that their power spectra often deviate significantly from the $1/f^{2-\nu}$ form. The distributions of wavelet coefficients at successive scales suggest, however, that illumination maps do possess scale-invariant properties apart from the contributions of bright localized light sources.

Authors in the natural image statistics literature have noted that even though bandpass filter pyramid coefficients are approximately decorrelated, coefficients that are near one another in position, scale, or orientation exhibit codependencies that are remarkably reproducible for different images (Simoncelli, 1999; Buccigrossi & Simoncelli, 1999; Huang & Mumford, 1999). These codependencies are due largely to edgelike structures in images, so oriented filter pyramids are important in analyzing them. The spherical wavelet basis used to generate Figures 15 and 16, on the other hand, consists of wavelet functions with approximate radial symmetry. Because oriented pyramid transforms for spherical data domains are not readily available, we applied planar pyramid analysis to equal-area cylindrical projections of the Teller and Debevec illumination maps. This projection introduces spatially varying distortion that may affect the image statistics, but it allows direct comparison of our results to the existing literature on natural image statistics

and texture analysis. Horizontal lines in the projected images correspond to lines of latitude on the sphere, whereas vertical lines correspond to lines of longitude. We used an 8-tap quadrature mirror filter (QMF) pyramid described by Johnston (1980) and implemented by Simoncelli and Adelson (1990), and we used $k = 2/\pi$ in the equal-area projection. We confirmed that the coefficient distributions for both vertically and horizontally oriented filters at successive scales are similar to those observed for spherical wavelets in Figure 15.

Figure 17 shows the conditional distributions of the horizontal QMF coefficients of the Teller illumination maps given the values of several nearby coefficients. These distributions are shown as images, with each column representing the distribution of the horizontal coefficient given a particular value of a related coefficient. Brighter pixels represent higher probabilities, with the probabilities in each column summing to one.

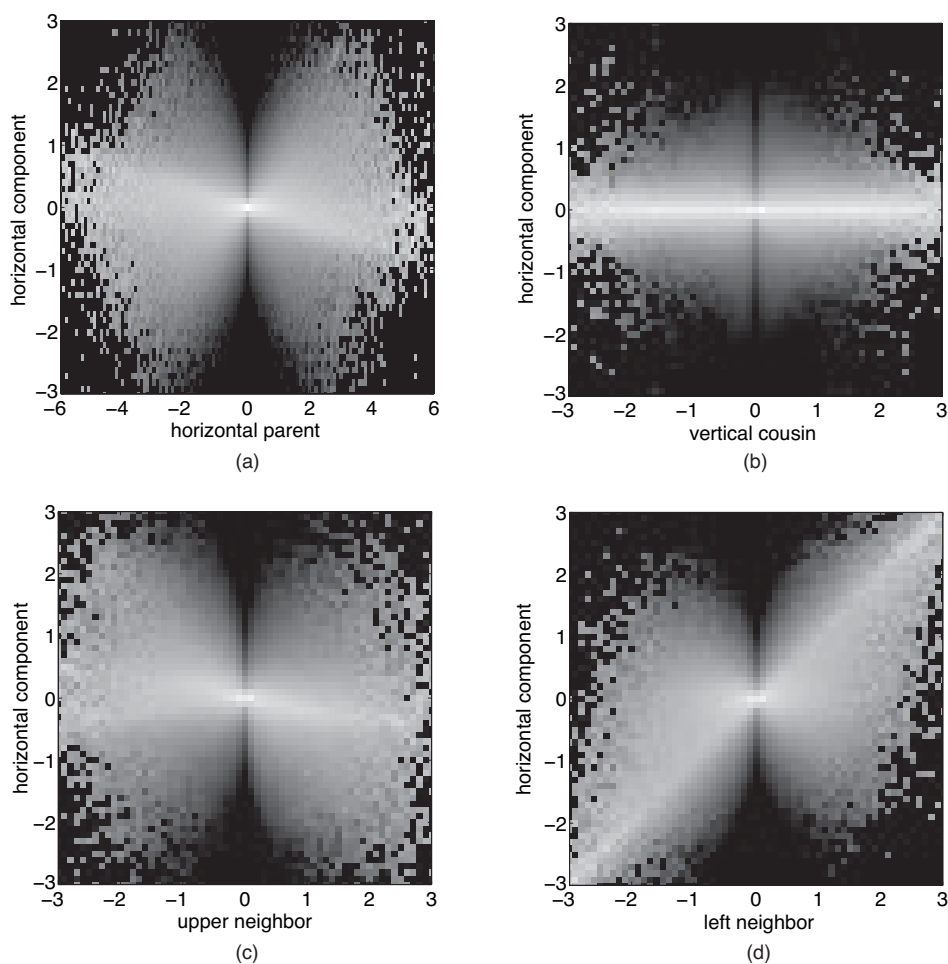


Figure 17. Conditional histograms for a horizontal filter coefficient given the values of its neighbors. The brightness of each pixel indicates a probability; the probabilities in each column sum to unity. The vertical axis is a fine-scale horizontal coefficient of an 8-tap QMF decomposition. The horizontal axis represents (a) the horizontal coefficient at the same position but at the next coarser scale, (b) the vertical coefficient at the same scale and position, (c) a vertically adjacent horizontal coefficient at the same scale, and (d) a horizontally adjacent horizontal coefficient at the same scale. The conditional histograms represent average distributions over the 95 Teller log luminance images.

All four of the joint distributions exhibit a “bow tie” shape characteristic of natural images (Simoncelli, 1999; Buccigrossi & Simoncelli, 1999). The variance of a filter coefficient increases with the magnitude of neighboring coefficients at the same scale and orientation, and also with the magnitude of coefficients of other scales and orientations at the same spatial location. Intuitively, edges and bright sources tend to produce large coefficients at multiple scales and orientations and at nearby positions. Figure 17(d) shows that two horizontally adjacent, horizontally oriented coefficients at the same scale also exhibit significant correlation. This correlation reflects the tendency of edges in an image or illumination map to continue in the same direction; horizontally oriented filters respond strongly to horizontal edges.

4. Discussion

4.1 Applications of illumination statistics

The properties of real-world illumination are important in vision and graphics because illumination, together with the reflectance properties and geometry of a surface, determines the appearance of a surface in an image. In graphics, one must specify an illumination to render an image. In vision, one must make assumptions about illumination to recognize reflectance or geometry. The statistical regularities discussed in this work may, therefore, find application in several areas.

Understanding human vision

We have found that humans are able to match surface reflectance properties from isolated images of surfaces under different unknown real-world illuminations (Fleming, Dror, et al., 2003). In the absence of assumptions about illumination, this problem is underconstrained; different combinations of illumination and reflectance could produce exactly the same image, even if one assumes that surface geometry is known. Indeed, humans perform much worse in reflectance matching tasks given images rendered under simple synthetic illumination maps. Our experimental evidence suggests that humans may depend on the statistical properties discussed here to judge surface reflectance properties (Fleming, Dror, et al., 2003). Hartung and Kersten (2003) and Fleming, Torralba, Dror, and Adelson (2003) have found evidence that humans take advantage of similar properties of illumination to recognize shape under unknown illumination. Statistical characterization of illumination is an essential component of a Bayesian approach to object and material perception (Kersten, Mamassian, & Yuille, 2004).

Computer vision

We have been able to take advantage of the statistical regularity of real-world illumination to design a computer

vision system that classifies surface reflectance from images of a surface under unknown illumination (Dror Adelson, & Willsky, 2001; Dror, 2002). The regularity of illumination patterns translates into predictable relationships between certain features of an image of a surface and the reflectance of that surface. In particular, we found that statistics summarizing the distributions of pixel intensities and bandpass filter coefficients of the observed image facilitated classification of surface reflectance. More generally, an understanding of illumination statistics may allow recognition of materials and material properties by a computer vision system.

Shape-from-shading algorithms depend on the relationship between surface orientation and reflected surface radiance, which in turn depends on illumination. Statistical priors on illumination may allow computer vision systems to recognize surface geometry under unknown illumination, even for specular surfaces. Such statistical priors may also facilitate accurate motion estimation in the presence of specularities.

Computer graphics

Researchers in computer graphics have recently devoted considerable effort to rendering scenes using real-world illumination to achieve greater realism (Debevec, 1998; Debevec et al., 2000; Ramamoorthi & Hanrahan, 2001). Performing such renderings quickly and with reasonable storage requires compact representations for real-world illumination and efficient methods for rendering under such illumination. One may be able to exploit statistical properties of real-world illumination to achieve these goals. For example, Ng et al. (2003) found that a wavelet-based lighting approximation proves more effective than one based on spherical harmonics.

The illumination statistics discussed here might also be used to recover illumination maps from sparse or incomplete measurements, or to create synthetic illumination maps that lead to realistic rendered images.

4.2 Comparison of illumination maps and typical photographs

We have found that the statistical properties of real-world illumination maps are similar in many ways to those of typical photographs. This might be expected, given that illumination maps can also be thought of as photographs of the real world. The structures that contribute to the statistics of typical photographs, such as edges, surfaces, and textured regions, are also present in illumination maps. On the other hand, we have observed a number of differences between the statistics of illumination maps and those reported in the natural images statistics literature. These stem from several differences between illumination maps and typical photographs:

- Illumination maps have a much greater angular extent than typical photographs.
- Photographs are typically taken in a nearly horizontal direction, matching the experience of human vision. Illumination maps are omnidirectional, with most power typically incident from above. Illumination maps often include primary light sources, such as the sun; photographs tend to avoid these.
- Illumination maps have an intrinsic sense of orientation, which photographs may or may not share.
- Illumination maps generally have a much higher dynamic range than typical photographs.
- Illumination maps are linear in luminance, whereas most photographic devices compress the luminance range in a nonlinear and often uncharacterized fashion.

Some of these differences, such as the limited dynamic range and nonlinear response of typical photographs, might be viewed as limitations of the recording device. If one wishes to use image statistics for image processing or computer vision tasks, however, the relevant statistics are those of the actual images to be processed, regardless of the fidelity with which they represent the real world. For illumination maps, on the other hand, accurate representation of the dynamic range is critical. Using illumination maps with compression or truncation of the dynamic range for rendering purposes will lead to a lack of realism in the resulting rendered images. In fact, the use of illumination maps or “light probes” for rendering purposes has motivated recent developments in high dynamic range photography (Debevec & Malik, 1997; Debevec, 1998).

4.3 Illumination maps as textures

The domains in which we have characterized natural illumination statistics – distributions of intensities, power spectra, and distributions of wavelet coefficients – are also used to characterize texture (Heeger & Bergen, 1995; Portilla, Strela, Wainwright, & Simoncelli, 2001). Indeed, we might think of illumination patterns as types of textures. We can test the extent to which a set of statistics captures the perceptually essential characteristics of real-world illumination by applying texture synthesis algorithms to generate novel illuminations whose statistics match those of real-world illuminations. Panel (a) of Figure 18 shows a sphere rendered under the photographically acquired illumination map of Figure 7(d). Panels (b), (c), and (d) show identical spheres rendered under synthetic illumination maps. The illumination map of (b) consists of Gaussian noise with a $1/f^2$ power spectrum; although the power spectrum resembles that of natural illumination, the resulting sphere does not look realistic at all.⁴ The illumination map of (c) was synthesized to have a pixel intensity distribution and mar-

ginal wavelet coefficient distributions identical to those of (a), using the texture synthesis technique of Heeger and Bergen (1995). This sphere looks much more realistic, and human observers are able to recognize that its reflectance properties are similar to those of the sphere in (a) (Fleming, Dror, et al., 2003). Finally, the illumination map of (d) was created using the texture synthesis technique of Portilla and Simoncelli (2001), which ensures that not only its pixel intensity distribution and marginal wavelet coefficient distributions but also certain properties of its joint wavelet coefficient distributions match those of (a). This synthetic illumination map captures the presence of edges in the real illumination map, leading to a sphere whose apparent reflectance properties are even more similar to that of (a). This suggests that the statistical properties of natural illumination described in this chapter play an important role in reflectance estimation by the human visual system (as discussed in Fleming, Dror, et al., 2003). It also suggests that one may be able to produce realistic renderings using properly synthesized illumination.

4.4 Future directions

One could extend our treatment of real-world illumination by considering how an illumination map tends to change as the camera recording it moves through space. That is, one might consider the statistics of the plenoptic function, which describes all the rays of light passing through every point in a three-dimensional volume (Adelson & Bergen, 1991). The five-dimensional plenoptic function can be characterized as the set of two-dimensional spherical illumination maps at every point in a three-dimensional volume. Because image-based rendering involves resampling the plenoptic function (McMillan & Bishop, 1995), statistical priors on this function could facilitate image-based rendering with sparse data.

We carried out our analysis using only illumination intensity information – that is, we essentially analyzed gray-scale illumination maps. One could extend this treatment by considering color. Rough preliminary analysis suggests that the statistical properties discussed in this work are similar for different color channels.

5. Conclusions

The illumination distributions we encounter in everyday life are highly variable and complex. At the same time, they exhibit a great deal of statistical regularity. In fact, one might view illumination patterns as complicated textures with clearly recognizable characteristics.

We examined the statistics of a set of illumination maps recorded photographically in everyday indoor and outdoor scenes. The pixel intensity distributions of these illumination maps peak at low intensities, with fewer pixels of much higher intensity. The frequency spectra, computed using spherical harmonics, fall off at a predictable rate at

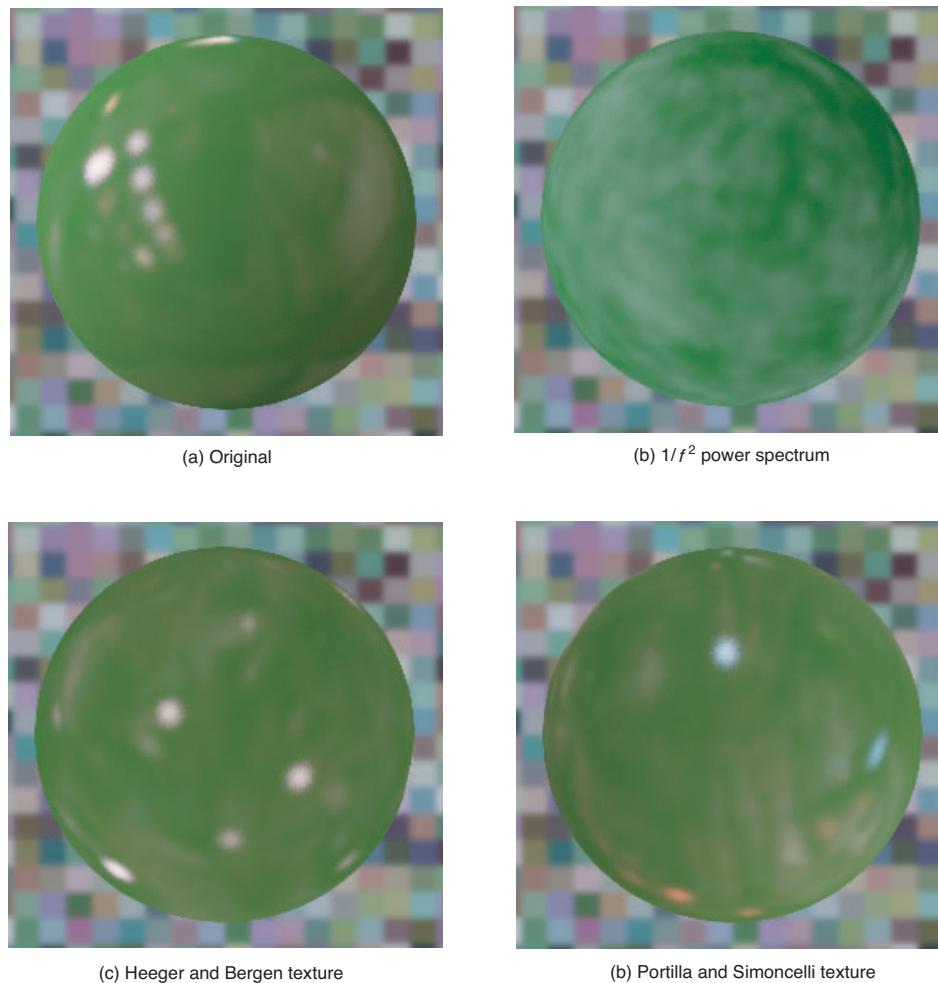


Figure 18. Spheres of identical reflectance properties rendered under a photographically acquired illumination map (a) and three synthetic illumination maps (b-d). The illumination in (b) is Gaussian noise with a $1/f^2$ power spectrum. The illumination in (c) was synthesized with the procedure of Heeger and Bergen (1995) to match the pixel histogram and marginal wavelet histograms of the illumination in (a). The illumination in (d) was synthesized using the technique of Portilla and Simoncelli (2000), which also enforces conditions on the joint wavelet histograms. The illumination map of (a) is due to Debevec et al. (2000).

high frequencies. Bandpass filter pyramid coefficients at each scale and orientation have highly kurtotic distributions of a predictable shape. Coefficients of filters adjacent in scale, orientation, and position exhibit strong statistical dependencies. Although the coefficients themselves are roughly uncorrelated, their magnitudes are heavily correlated. These predictable statistics correspond to intuitive notions, such as the presence of sharp edges at different scales in real-world illumination patterns.

Many of the regularities observed through earlier studies of low dynamic range, restricted field-of-view photographs carry over to real-world illumination maps. Unlike the photographs analyzed in most of the natural image statistics literature, however, the illumination maps we analyzed have a very wide field of view and contain primary light sources represented with high dynamic range. This leads to several significant differences between the statistics of illumination maps and those typically reported in the natural image statistics literature. The presence of strong

point-like light sources in some scenes leads to high variability in the power spectra of illumination maps, particularly at low frequencies. In particular, the power spectra may deviate significantly from the $1/f^{2+\eta}$ model, violating scale invariance. Illumination maps display nonstationary statistical properties, such as different distributions of illumination intensity at different elevations. Typical photographs may also lack stationarity, but their nonstationary properties have received little attention in the literature. Wavelet coefficient distributions are fairly regular from one illumination map to another, but fits to generalized Laplacian distributions are less tight than those previously observed for more typical photographs (Buccigrossi & Simoncelli, 1999; Huang & Mumford, 1999).

The characteristics of real-world illumination play an essential role in the perception of material properties. A description of these statistics also facilitates the rendering of realistic computer graphics imagery and the design of robust computer vision systems able to recognize materials.

Footnotes

¹The kurtosis of a random variable X with probability density $f(x)$ is defined as

$$k = \frac{\int (x - \bar{x})^4 f(x) dx}{(\int (x - \bar{x})^2 f(x) dx)^2}.$$

The kurtosis of a Gaussian is 3, and distributions with kurtosis higher than 3 are often referred to as heavy-tailed.

²The differential entropy H of X is defined as

$$H(X) = - \int f(x) \log_2 f(x) dx.$$

Differential entropy is a measure of information content for a continuous random variable. The distribution with variance σ^2 that maximizes differential entropy is the Gaussian, which has differential entropy

$$\frac{1}{2} \log_2 2\pi e \sigma \approx 2.05 + \log_2 \sigma$$

bits. On the other hand, a distribution that concentrates all probability density near a few discrete points could have an arbitrarily negative differential entropy.

³The mutual information of random variables X and Y is defined as $I(X, Y) = H(X) + H(Y) - H(X, Y)$, where $H(X)$ and $H(Y)$ are the differential entropies of X and Y , respectively, and $H(X, Y)$ is the differential entropy of their joint density.

⁴The illumination map of Figure 18(b) was synthesized in the spherical harmonic domain. The maps of (c) and (d) were synthesized in a rectangular domain corresponding to an equal-area cylindrical projection of the sphere. In (c) and (d), we performed principle component analysis in color space to produce three decorrelated color channels, each of which is a linear combination of the red, green, and blue channels. We then synthesized textures independently in each channel of this remapped color space, as suggested by Heeger and Bergen (1995). Unfortunately, the nonlinear dependencies between the decorrelated color channels are much more severe for high dynamic range illumination maps than for the 8-bit images common in the texture analysis literature. To reduce artifacts associated with these dependencies, we passed the original illumination maps through a compressive nonlinearity on luminance before wavelet analysis, and then applied the inverse nonlinearity to the synthesized illumination maps. The compressive nonlinearity leads to a less heavy-tailed distribution of pixel intensities.

Acknowledgments

Seth Teller, Neel Master, and Michael Bosse shared the data set from the MIT City Scanning Project and helped us use it to construct illumination maps. Thomas Leung contributed to our initial investigation of illumination statis-

tics. Roland Fleming and Antonio Torralba provided insightful suggestions, as did the anonymous reviewers. Julio Castrillon-Candas assisted us in using his fast lifted surface wavelet transform software. This work was supported by National Defense Science and Engineering Graduate Fellowship and Whitaker Fellowships to ROD, by National Institutes of Health Grant EY11005-04 and Office of Naval Research/Multi-University Research Initiative Contract N00014-01-0625 to EHA, by a grant from Nippon Telegraph and Telephone Corporation to the MIT Artificial Intelligence Lab, by a contract with Unilever Research, and by Office of Naval Research Grant N00014-00-1-0089 to ASW.

Commercial relationships: none.

Corresponding author: Ron O. Dror.

Email: rondror@ai.mit.edu.

Address: D. E. Shaw Research and Development, 120 W. 45th Street, New York, NY 10036.

References

- Adelson, E. H., & Bergen, J. R. (1991). The plenoptic function and the elements of early vision. In M. Landy and J. A. Movshon (Eds.), *Computational models of visual processing*. Cambridge, MA: MIT Press.
- Amaratunga, K., & Castrillon-Candas, J. E. (2001). Surface wavelets: A multiresolution signal processing tool for 3D computational modeling. *International Journal for Numerical Methods in Engineering*, 52, 239-271.
- Buccigrossi, R. W., & Simoncelli, E. P. (1999). Image compression via joint statistical characterization in the wavelet domain. *IEEE Transactions on Image Processing*, 8, 1688-1701.
- Canters, F. and Declair, H. (1989). *The world in perspective: A directory of world map projections*. New York: John Wiley & Sons.
- Debevec, P. E. (1998). Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography. *Proceedings of SIGGRAPH, 1998*, 189-198.
- Debevec, P. E., Hawkins, T., Tchou, C., Duiker, H.-P., Sarokin, W., & Sagar, M. (2000). Acquiring the reflectance field of a human face. *Proceedings of SIGGRAPH, 2000*, 145-156.
- Debevec, P. E., & J. Malik (1997). Recovering high dynamic range radiance maps from photographs. *Proceedings of SIGGRAPH, 1997*, 369-378.
- Dror, R. O. (2002). *Surface reflectance recognition and real-world illumination statistics* (AI Lab Technical Report, AITR-2002-009). Cambridge, MA: MIT Artificial Intelligence Laboratory. [Article]

- Dror, R. O., Adelson, E. H., & Willsky, A. S. (2001). Surface reflectance estimation and natural illumination statistics. *Proceedings of IEEE Workshop on Statistical and Computational Theories of Vision*, Vancouver, Canada.
- Dror, R. O., Leung, T., Willsky, A. S., & Adelson, E. H., (2001). Statistics of real-world illumination. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Kauai, Hawaii.
- Farid, H. (2002). Detecting hidden messages using higher-order statistical models. *International Conference on Image Processing*, Rochester, NY.
- 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, 2379-2394. [[PubMed](#)]
- Fleming, R. W., Dror, R. O., & Adelson, E. H. (2003). Real-world illumination and the perception of surface reflectance properties. *Journal of Vision*, 3(5), 347-368, <http://journalofvision.org/3/5/3/>, doi:10.1167/3.5.3. [[PubMed](#)][[Article](#)]
- Fleming, R. W., Torralba, A., Dror, R. O., & Adelson, E. H. (2003). How image statistics drive shape-from-texture and shape-from-specularity [[Abstract](#)]. *Journal of Vision*, 3(9), 73a, <http://journalofvision.org/3/9/73/>, doi:10.1167/3.9.73.
- Hartung, B., & Kersten, D. (2003). How does the perception of shape interact with the perception of shiny material? [[Abstract](#)] *Journal of Vision*, 3(9), 59a, <http://journalofvision.org/3/9/59/>, doi:10.1167/3.9.59.
- Heeger, D. J., & Bergen, J. R. (1995). Pyramid-based texture analysis/synthesis. *Proceedings of SIGGRAPH 1995*, 229-238.
- Huang, J., & Mumford, D. (1999). Statistics of natural images and models. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1, 541-547.
- Hwang, G. T. (2004). Hollywood's master of light. *Technology Review*, 107, 70-73.
- Inui, T., Tanabe, Y., & Onodera, Y. (1996). *Group theory and its applications in physics*. Springer: New York, Berlin, and Heidelberg.
- Johnston, J. D. (1980). A filter family designed for use in quadrature mirror filter banks. *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, 1980, 291-294.
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, 55, 271-304. [[PubMed](#)]
- Konishi, S. M., Yuille, A. L., Coughlan, J. M., & Zhu, S. C. (2003). Statistical edge detection: Learning and evaluating edge cues. *Pattern Analysis and Machine Intelligence*, 25, 57-74.
- Laughlin, S. B., (1981). A simple coding procedure enhances a neuron's information capacity. *Z. Naturforsch.*, 36c, 910-912.
- Levin, A., Zomet, A., & Weiss, Y. (2002). Learning to perceive transparency from the statistics of natural scenes. *Sixteenth Annual Conference on Neural Information Processing Systems*, Vancouver, Canada.
- McMillan, L., & Bishop, G. (1995). Plenoptic modeling: An image-based rendering system. *Proceedings of SIGGRAPH, 1995*, 39-46.
- Ng, R., Ramamoorthi, R., & Hanrahan, P. (2003). All-frequency shadows using non-linear wavelet lighting approximation. *Proceedings of SIGGRAPH, 2003*, 376-381.
- Olshausen, B. A., & Field, D. J. (2000). Vision and the coding of natural images. *American Scientist*, 88, 238-245.
- Portilla, J., & Simoncelli, E. P. (2000). A parametric texture model based on joint statistics of complex wavelet coefficients. *International Journal of Computer Vision*, 40, 49-71.
- Portilla, J., Strela, V., Wainwright, M. & Simoncelli, E. P. (2001). Adaptive Wiener denoising using a Gaussian scale mixture model in the wavelet domain. *Proceedings of the International Conference on Image Processing*, Thessaloniki, Greece.
- Ramamoorthi, R., & Hanrahan, P. (2001). An efficient representation for environment irradiance maps. *Proceedings of SIGGRAPH, 2001*, 159-170.
- Ruderman, D. L. (1994). The statistics of natural images. *Network-Computation in Neural Systems*, 5, 517-548.
- Schröder, P., & Sweldens, W. (1995). Spherical wavelets: Efficiently representing functions on the sphere. *Proceedings of SIGGRAPH, 1995*, 161-172.
- Schwartz, O., & Simoncelli, E. P. (2001). Natural signal statistics and sensory gain control. *Nature: Neuroscience*, 4, 819-825. [[PubMed](#)]
- Simoncelli, E. P. (1999, July). Modeling the joint statistics of images in the wavelet domain. *Proceedings SPIE, 44th Annual Meeting*, Denver, CO.
- Simoncelli, E. P., & Adelson, E. H. (1990). Subband transforms. In J. W. Woods (Ed.), *Subband Image Coding* (pp. 143-192). Norwell, MA: Kluwer Academic Publishers.
- Simoncelli, E. P., & Adelson, E. H. (1996). Noise removal via Bayesian wavelet coring. *Proceedings of the International Conference on Image Processing*, Lausanne, Switzerland.

- Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24, 1193-1216. [[PubMed](#)]
- Teller, S., Antone, M., Bosse, M., Coorg, S., Jethwa, M., & Master, N. (2001). Calibrated, registered images of an extended urban area. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Kauai, Hawaii.
- Tolhurst, D. J., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmology and Physiological Optics*, 12, 229-232. [[PubMed](#)]
- Torralla, A. & Sinha, P. (2001). *Contextual priming for object detection* (Memo 2001-020). Cambridge, MA: MIT Artificial Intelligence Laboratory. [[Link](#)]
- van Hateren, J. H., & van der Schaaf, A. (1998). Independent component filters of natural images compared with simple cells in primary visual cortex. *Proceedings of the Royal Society of London B*, 265, 359-366. [[PubMed](#)]
- Wainwright, M. J., Simoncelli, E. P., & Willsky, A. S. (2001). Random cascades on wavelet trees and their use in analyzing and modeling natural images. *Applied and Computational Harmonic Analysis*, 11, 89-123.