The Complex Statistics of High-Contrast Patches in Natural Images

Ann B. Lee †, Kim S. Pedersen ‡, and David Mumford ∗

† Department of Physics, Box 1843, Brown University, Providence, RI 02912, USA

‡ DIKU, Department of Computer Science, University of Copenhagen, Universitetsparken 1, DK-2100 Copenhagen Ø, Denmark

∗ Division of Applied Mathematics, Box F, Brown University, Providence, RI 02912, USA

Abstract

Recently, there has been a great deal of interest in modeling the non-Gaussian structures of natural images. However, despite the many advances in the direction of sparse coding and multi-resolution analysis, a description of the full probability distribution of pixels in a neighborhood is still missing. In this study, we explore the space of data points representing the values of 3×3 high-contrast patches from optical and 3D range images. We find that the distribution of data is extremely "sparse" with the majority of the data points concentrated in clusters and non-linear low-dimensional manifolds. Furthermore, a more detailed analysis of probability densities reveals clear differences between optical and range images, which have similar second-order statistics and scaling properties. Our work indicates the importance of studying the complex statistics of natural images, and the need to understand the intrinsic dimensionality of natural data.

1 Introduction

There have been two main discoveries in natural image statistics the past decade that have changed the way we think about images.

The first is the empirical observation that natural image statistics scale [9, 24, 27]. This means that any local statistic calculated on $n \times n$ images and on block averaged $2n \times 2n$

images is approximately the same, or equivalently, that the probability of seeing an image $I(x, y)$ is approximately that of seeing $I(\sigma x, \sigma y)$. Scale invariance implies a surprising stability for image statistics, and creates a link between local image models describing small-scale structures and global image models describing the large-scale structures in images.

The second observation is that natural image statistics are extremely non-Gaussian [9, 19]. The non-Gaussian nature is apparent if one computes 1D histograms of virtually any mean-zero filter on virtually any database of images: the histogram will, essentially always, have kurtosis greater than 3 (the kurtosis of Gaussian distributions). In projection pursuit [11, 15] and sparse coding analysis [21], "interesting" directions are often equivalent to non-Gaussian directions (see, for example, Hyvärinen [16] for a survey of Independent Component Analysis and related methods). Furthermore, many denoising schemes, for example, wavelet shrinkage [8], are directly based on the highly kurtotic shape of the 1D marginal distributions of natural images. The full complexity of the non-Gaussian structures is perhaps most apparent if one studies the 2D joint statistics of images in the wavelet domain; see e.g. Zetzsche et al. and Simoncelli [31, 25] for an analysis of the marginal and pairwise joint histograms of wavelet coefficients at adjacent spatial locations, orientations, and spatial scales. Huang and Mumford [14] have used an approach similar to Buccigrossi and Simoncelli [1] but with Haar filters which are less smooth than the wavelets usually preferred in applications. They found that the equiprobable contours of the joint histograms of these filters are highly irregular with cusps and corners that indicate common geometric structures in the original images. These contour plots are even more striking for range images. In Huang et al. [13], the authors directly connect features in the joint Haar wavelet histograms of range images to object-like structures in 2×2 and 2×4 patches. Some of these structures are perhaps best classified as "clutter", i.e. small objects or markings (some partially resolved by the camera) that are not important to the understanding of the scene. A better statistical model for clutter, which may not be interesting in itself, is necessary to develop better image enhancement algorithms that can deal with non-Gaussian noise components.

Choosing the best representation for a given input signal is a central problem [4, 5, 25, 28]. For some time there has been a trend towards incorporating geometrical structures of images, e.g. [2, 3, 7, 10, 17, 18, 20, 30]. Despite the recent advances in computer vision and image understanding, we still do not know what the *full* probability distribution of naturally occurring images, or patches of these images, looks like. It seems that without this knowledge we are not able to fully take advantage of the sparseness of the space of natural images. No doubt any knowledge about how the data is distributed in higher-dimensional spaces is useful when developing better representation and compression schemes. It is especially interesting to find high-density clusters and low-dimensional manifolds in the image space, as these structures greatly reduce the dimensionality of the problem.

In this paper we will analyze the probability distribution of 3×3 high-contrast log-intensity image patches. The elliptical contours in the center part of contour plots of Haar wavelet joint filter reactions [14] possibly indicate that fluctuations around low-contrast image regions are more Gaussian in nature. Furthermore, it is commonly believed that image regions with high contrast contain a large proportion of the salient structures of a scene. Reinagel and Zador [22] have shown, for natural images with a variety of cognitive content, that humans tend to focus their eye movements around high-contrast regions — thus significantly biasing the effective input that reaches the early stages of the visual system towards these types of image regions. Geman and Koloydenko [12] have found that edges are "the most probable non-background microimage configurations". In their work they analyze 3×3 blocks in images by a modified order statistic. The patches are first quantized according to the differences between the pixel values in a patch, and then grouped into "equivalence classes" based on photometry, complexity and geometry. Our work is related to Geman et al's study of microimages, but our focus is on searching for non-linear low-dimensional manifolds and clusters of data points embedded in the space of image patches. After preprocessing, consisting of subtracting the mean of each 3×3 patch and then sphering the data, the dimensionality of this space is 7. Furthermore, we try to find a systematic way of distinguishing patches from different types of images, that have seemingly similar statistics. A comparison is made between the probability distributions of optical and 3D range images. The latter data has been collected with a laser range finder with very different sub-resolution pixel properties — a fact that is not directly apparent or easily expressed mathematically when analyzing 1D marginal distributions and other simple statistics.

Figure 1: Samples from the van Hateren optical image database. The gray values code for log-intensity values.

The organization of the paper is as follows: In Section 2, we describe the two data sets extracted from an optical image database by van Hateren [29] and a range image database by Huang and Lee [13]. In Section 3, we describe the preprocessing of the data sets. Our analysis is divided into three parts: In Section 4 we study the distribution of our data with respect to a Voronoi sampling of the space of data points. We proceed in Section 5 with a study of the probability density of high-contrast optical image patches around a 2D manifold of ideal step edges. Finally, in Section 6 we analyze the probability density of range data around binary patches. The results are discussed in Section 7.

2 Optical and Range Data Sets

We extract two data sets of high-contrast 3×3 patches, or "microimages", from optical and range images, respectively. These patches are then preprocessed according to Section 3.

- The *optical* data set contains about $3.1 \cdot 10^6$ high-contrast log-intensity patches.

Figure 2: Samples from the Huang-Lee range image database. The gray values code for log-range values.

These are extracted from van Hateren's still image collection [29] of 4167 linear [1] calibrated 1024×1536 images; see Figure 1 for samples. From each image, we randomly select 15000 3×3 patches, and keep the top 5 percent, i.e. 750 patches, with the highest contrast in log-intensities (see Section 3 for definition of contrast or "D-norm").

- The *range* data set contains about $6.9 \cdot 10^5$ high-contrast log-range patches. These are extracted from the Huang-Lee database [13] of around 200 444×1440 range images with mixed outdoor and indoor scenes; see Figure 2 for samples. We divide each image into disjoint 3×3 patches, and discard all patches with out-of-range data. From the remaining patches we select the top 5 percent patches, in each image, with highest contrast in log-range values.

The optical and range images have quite different subresolution properties. In optical images, the subpixel details are *averaged* by the point-spread function of the camera. The pixel values in a range image, on the other hand, usually correspond to the *minimum* of the subresolution details. The field of view of the scanner is $80°$ vertically and $259°$ horizontally. The beam divergence of the laser range finder is approximately 3 mrad. If the laser footprint hits two targets with a range difference larger than 3 meters, the returned value is the range to the nearest target. If the range difference is less than 3 meters, the returned value is roughly a weighted average of both ranges where the weight depends on the reflectivity of the two targets.

3 Preprocessing

We want to compare and group image patches based on their *geometrical structure*. In a natural scene, the reflectance and the shape of a surface are largely invariant, while the absolute distance to, and the illumination of, a point in a scene can vary widely. We thus work with the *logarithm*, rather than the absolute values, of intensity [2] or range. Each

[1] The pixel values in these images are approximately linearly proportional to the scene luminance.

[2] The logarithmic compression of intensities is consistent with the inverse relationship between ambient illumination and human sensitivity. According to Weber's law, the ratio $\Delta L/L$ of the just noticeable difference ΔL and the ambient luminance L, is constant for a wide range of luminances. It is argued that the human

high-contrast patch is furthermore centered and contrast-normalized before the analysis.

Let $\mathbf{x} = [x_1, \ldots, x_9]^T = [I_{11}, I_{21}, I_{31}, I_{12}, \ldots, I_{33}]^T \in \mathbb{R}^9$ be a non-constant vector with the log-values of the original patch. Centering and contrast normalization will then lead to a new vector

$$\mathbf{y} = \frac{\mathbf{x} - \frac{1}{9}\sum_{i=1}^{9} x_i}{||\mathbf{x} - \frac{1}{9}\sum_{i=1}^{9} x_i||_D} \tag{1}$$

in a 7-dimensional subspace $\tilde{\mathrm{S}}^7 \subset \mathbb{R}^9$. The norm $||\mathbf{x}||_D$, or the "D-norm", is defined by

$$||\mathbf{x}||_D{}^2 = \sum_{i \sim j} (x_i - x_j)^2 \ . \tag{2}$$

The sum is taken over the 12 bonds of 4-connected neighbors in the 3×3 patch. In matrix form, we have

$$||\mathbf{x}||_D{}^2 = \mathbf{x}^T D \mathbf{x}, \tag{3}$$

where

$$D = \begin{bmatrix} 2 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 3 & -1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & 0 & 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 3 & -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 4 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 & 3 & 0 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 & 0 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 & -1 & 2 \end{bmatrix} . \tag{4}$$

The "D-norm" is a finite version of $\int\int ||\nabla I||^2 d\xi d\eta$, which is the unique scale-invariant norm on images $I(\xi, \eta)$. The D-norm is also a "natural" measure of the contrast in the patch, in the sense that it can be related to the correlation matrix $\Sigma = E[\mathbf{x}'\mathbf{x}'^T]$, where the vector $\mathbf{x}' = (\mathbf{x} - \frac{1}{9}\sum_{i=1}^{9} x_i)$ represents the centered data. Both the matrices Σ and D are of rank 8, with the null space spanned by the vector $\mathbf{x}_0 = [1, 1, 1, 1, 1, 1, 1, 1, 1]^T$. Assume that $\tilde{\Sigma}$ and $\tilde{D}$ are 8×8 matrix representations in the subspace orthogonal to $\mathbf{x}_0$.

visual system uses this adaptation scheme to preserve the relative reflectances (brightness ordering) of an array of surfaces in a scene.

For scale invariant images, we have that

$$\tilde{\Sigma} \propto \tilde{D}^{-1} \,. \tag{5}$$

Empirical observations show that the relation holds approximately for both the optical and range datasets.

The geometrical interpretation of the preprocessing stage (defined by Equation 1) is the following: First, we map the data points $\mathbf{x}$ onto the hyperplane

$$\mathrm{L}^8 = \{\mathbf{x} \in \mathbb{R}^9 \mid x_1 + \ldots + x_9 = 0\} \tag{6}$$

that is orthogonal to $\mathbf{x}_0 = [1,1,1,1,1,1,1,1,1]^T$. Then, we project the points to a 7-dimensional Euclidean ellipsoid

$$\tilde{\mathrm{S}}^7 = \{\mathbf{y} \in \mathrm{L}^8 \subset \mathbb{R}^9 \;:\; \mathbf{y}^T D \mathbf{y} = 1\} \,. \tag{7}$$

It is convenient to make a change of basis to a coordinate system where the data points lie on a *Euclidean* sphere. The matrix D can be diagonalized with the 2-dimensional Discrete Cosine Transform (DCT) basis of a 3×3 image patch. In vector form, we write the 8 non-constant DCT basis vectors as

$$\begin{array}{lcl}
\mathbf{e}_1 & = \frac{1}{\sqrt{6}}[& 1, \;\; 0, \;\; -1, \;\; 1, \;\; 0, \;\; -1, \;\; 1, \;\; 0, \;\; -1]^T \\
\mathbf{e}_2 & = \frac{1}{\sqrt{6}}[& 1, \;\; 1, \;\; 1, \;\; 0, \;\; 0, \;\; 0, \;\; -1, \;\; -1, \;\; -1]^T \\
\mathbf{e}_3 & = \frac{1}{\sqrt{54}}[& 1, \;\; -2, \;\; 1, \;\; 1, \;\; -2, \;\; 1, \;\; 1, \;\; -2, \;\; 1]^T \\
\mathbf{e}_4 & = \frac{1}{\sqrt{54}}[& 1, \;\; 1, \;\; 1, \;\; -2, \;\; -2, \;\; -2, \;\; 1, \;\; 1, \;\; 1]^T \\
\mathbf{e}_5 & = \frac{1}{\sqrt{8}}[& 1, \;\; 0, \;\; -1, \;\; 0, \;\; 0, \;\; 0, \;\; -1, \;\; 0, \;\; 1]^T \\
\mathbf{e}_6 & = \frac{1}{\sqrt{48}}[& 1, \;\; 0, \;\; -1, \;\; -2, \;\; 0, \;\; 2, \;\; 1, \;\; 0, \;\; -1]^T \\
\mathbf{e}_7 & = \frac{1}{\sqrt{48}}[& 1, \;\; -2, \;\; 1, \;\; 0, \;\; 0, \;\; 0, \;\; -1, \;\; 2, \;\; -1]^T \\
\mathbf{e}_8 & = \frac{1}{\sqrt{216}}[& 1, \;\; -2, \;\; 1, \;\; -2, \;\; 4, \;\; -2, \;\; 1, \;\; -2, \;\; 1]^T
\end{array} \tag{8}$$

where the normalization is chosen such that $||\mathbf{e}_1||_D = \ldots ||\mathbf{e}_8||_D = 1$. Let the DCT basis vectors above be the columns of a 9×8-matrix $A = [\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_8]$, and introduce a diagonal 8×8-matrix Λ with the diagonal elements equal to $1/||\mathbf{e}_1||^2, 1/||\mathbf{e}_2||^2, \ldots, 1/||\mathbf{e}_8||^2$. A change of basis taking $\mathbf{e}_i$ to be the unit vectors according to

$$\mathbf{v} = \Lambda A^T \mathbf{y} \ , \tag{9}$$

or equivalently, $\mathbf{y} = A\mathbf{v}$, will then transform the ellipsoid $\tilde{\mathrm{S}}^7$ in Equation 7 to a 7-dimensional *Euclidean* sphere

$$\mathrm{S}^7 = \{\mathbf{v} \in \mathbb{R}^8 \ : \ ||\mathbf{v}|| = 1\} \ . \tag{10}$$

To measure the distance between two 3×3 image patches P_1 and P_2, we simply calculate the *angular distance* between the corresponding two points $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^8$ on the 7-dimensional unit sphere. In other words, our distance measure is given by

$$\begin{aligned} \mathrm{dist}(P_1, P_2) &= \arccos(\mathbf{v}_1 \cdot \mathbf{v}_2) \\ &= \arccos\left(\frac{B\mathbf{y}_1 \cdot B\mathbf{y}_2}{||B\mathbf{y}_1|| \ ||B\mathbf{y}_2||}\right) \ , \end{aligned} \tag{11}$$

where the matrix $B = \Lambda A^T$, and the vectors $\mathbf{y}_1, \mathbf{y}_2 \in \mathbb{R}^9$ represent the centered 3×3 image patches P_1 and P_2.

4 A First Exploration of the Probability Distribution on the 7-Sphere

As a first exploration of our two data sets, we divide the 7-sphere $S^7 \subset \mathbb{R}^8$ into Voronoi cells, and analyze how the data points are distributed with respect to these cells.

A Voronoi cell V_i around a sampling point $\mathbf{x}_i \in \mathrm{S}^7$ is defined as the set of all points of $\mathbf{x} \in \mathrm{S}^7$ that are at least as close to $\mathbf{x}_i$ as to any other sampling point $\mathbf{x}_j$ that is

$$V_i = \{\mathbf{x} \in \mathrm{S}^7 \mid dist(\mathbf{x}, \mathbf{x}_i) \leq dist(\mathbf{x}, \mathbf{x}_j) \text{ for any } \mathbf{x}_j \in \mathrm{S}^7\}, \tag{12}$$

where $dist(\mathbf{x}, \mathbf{x}_i)$ is the angular distance between the two points as defined in Equation 11.

We use a sampling scheme that divides the 7-sphere into 2592 Voronoi cells with roughly the same size (see [6] for sampling schemes in high-dimensional spaces). The sampling points $\mathbf{x}_i$ are given by the following four 8-vectors and their permutations and sign changes:

1. The 16 unit vectors along the axes, $(1,0,0,0,0,0,0,0)^T$.
2. The 112 permutations and sign changes of $\frac{1}{\sqrt{2}}(1,1,0,0,0,0,0,0)^T$.
3. The 1344 permutations and sign changes of $\frac{1}{\sqrt{2/3}}(1,1/2,1/2,0,0,0,0,0)^T$.
4. The 1120 permutations and sign changes of $(1/2,1/2,1/2,0,0,0,0,0)^T$.

From a Monte Carlo estimate (with 1 million random points on the 7-dimensional unit sphere), we get that the volumes of the 4 types of Voronoi cells above are approximately 0.0046, 0.0055, 0.0085, and 0.018, respectively.

We bin our high-contrast optical and range patches into the 2592 Voronoi cells using the definition in Equation 12. We define the density of the Voronoi cell around sample point $\mathbf{x}_i$ as

$$\rho(V_i) = \frac{N(V_i)/\sum_j N(V_j)}{\mathrm{vol}(V_i)/\mathrm{vol}(\mathrm{S}^7)}, \tag{13}$$

where $N(V_i)$ is the number of patches of the Voronoi cell V_i associated with the point $\mathbf{x}_i$, and $\mathrm{vol}(V_i)$ is the volume of that cell, and $\mathrm{vol}(\mathrm{S}^7) = \sum_i \mathrm{vol}(V_i) = \pi^4/3$ is the total volume of the 7-sphere.

In Figure 3 and 4 we show the density $\rho(V_i)$ of the Voronoi cells for high-contrast optical and range patches, together with the cumulative percentage of volume $\sum_{j\leq i}\mathrm{vol}(V_j)$ occupied by the cumulative percentage of patches $p_{\mathrm{cum}}(V_i) = \sum_{j\leq i} p(V_j)$, where $p(V_j) = N(V_j)/\sum_k N(V_k)$. We find that the probability distribution of data on S^7 is extremely sparse, with the majority of data points concentrated in a few regions on the sphere. For both the optical and range data, half of the patches can be divided into an optimal set of Voronoi cells that occupy less than 5% of the total volume of the 7-sphere.

The most frequent Voronoi cells for optical patches are horizontal, vertical, and diagonal step edges with a gradual transition in gray level intensity around the edge (Figure 5). For high-contrast range patches, the most frequent cells are horizontal and vertical binary edges (Figure 6).

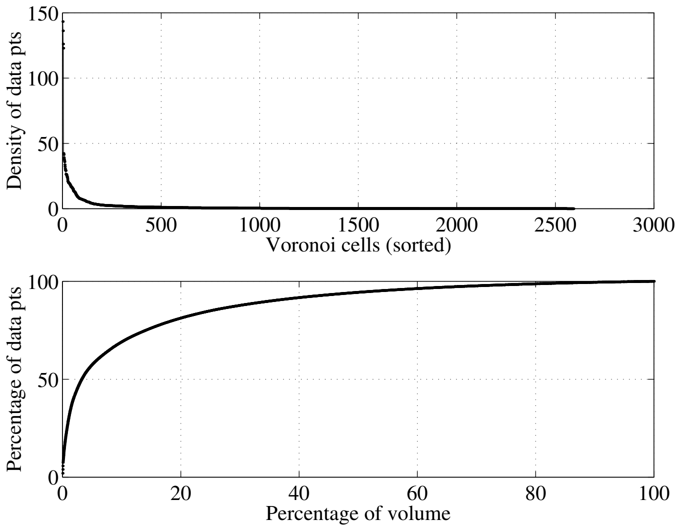

Figure 3: (Top) Density of high-contrast optical patches in Voronoi cells that are sorted according to decreasing density. (Bottom) Cumulative percentage of optical patches in the Voronoi cells above versus the cumulative percentage of volume in S^7 that are occupied by these cells.

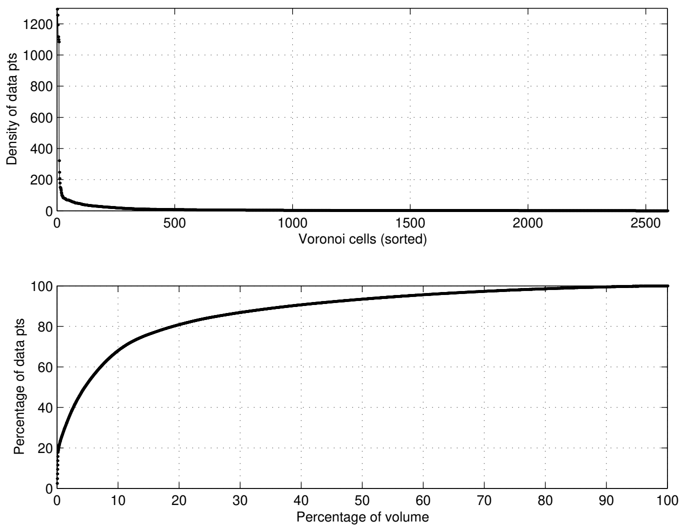

Figure 4: (Top) Density of high-contrast range patches in Voronoi cells that are sorted according to decreasing density. (Bottom) Cumulative percentage of range patches in the Voronoi cells above versus the cumulative percentage of volume in S^7 that are occupied by these cells.

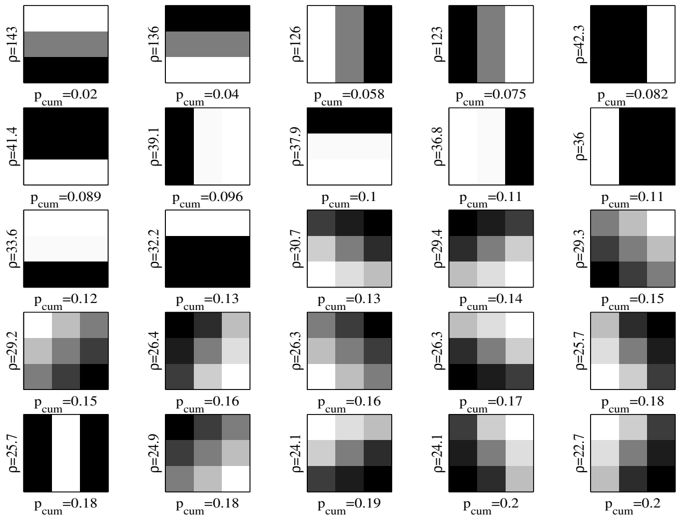

Figure 5: The 25 most frequent Voronoi patches for optical images sorted after their densities ρ. For each Voroni cell, we display the 3×3 patch corresponding to the sample point $\mathbf{x}_i$. The cumulative percentage of patches is shown as p_{cum}.

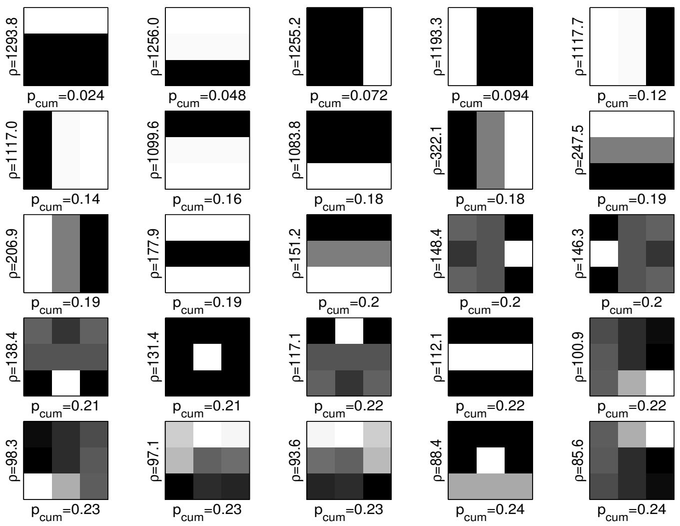

Figure 6: The 25 most frequent Voronoi patches for range images sorted after their densities ρ. For each Voronoi cell, we display the 3×3 patch corresponding to the sample point $\mathbf{x}_i$. The cumulative percentage of patches is shown as p_{cum}.

5 Optical Data: Probability Density Around a 2D Manifold of Step Edges

The analysis of Voronoi cells on the 7-sphere indicates that step edges are common high-contrast patterns. In Section 5.1, we present an ideal model for edges in optical images that takes into account the *averaging* effects of the optics of the camera. In Sections 5.2 and 5.3, we test the model with real data. Finally, in Section 5.4 we apply this model to range image patches.

5.1 The Ideal Surface

We represent the 3×3 image patch by a square $SQ = \{(x, y) : -3/2 \leq x, y \leq 3/2\}$. The pixels in the patch are given by

$$S_{ij} = \{(x, y) :\ (j - 3/2) \leq x \leq (j - 1/2),\ (1/2 - i) \leq y \leq (3/2 - i)\}\ , \tag{14}$$

where $i, j = 0, 1, 2$.

The pixel value $I(i, j)$ in an optical image is approximately an average of the underlying scene $\phi(x, y)$ recorded at each pixel S_{ij}, i.e. the pixel intensity

$$I(i, j) = \int_{S_{ij}} \phi(x, y) dx dy\ . \tag{15}$$

For an ideal step edge,

$$\phi_{\alpha,l}(x, y) = \begin{cases} a & \text{if} \quad -x \sin\alpha + y \cos\alpha \geq l \\ b & \text{if} \quad -x \sin\alpha + y \cos\alpha < l \end{cases} \tag{16}$$

where $a > b$. The parameter $\alpha \in [0, 2\pi)$ is the angle that the direction perpendicular to the edge makes with the y-axis, and the parameter $l \in (-3/2, 3/2)$ is the displacement of the edge from the origin (Figure 7). Thus, pixels in $I_{\alpha,l}(i, j)$ strictly above the edge have intensities a, pixels strictly below the edge have intensities b, and pixels along the edge is an weighted average of a and b. After centering and contrast normalization (see Section 3)

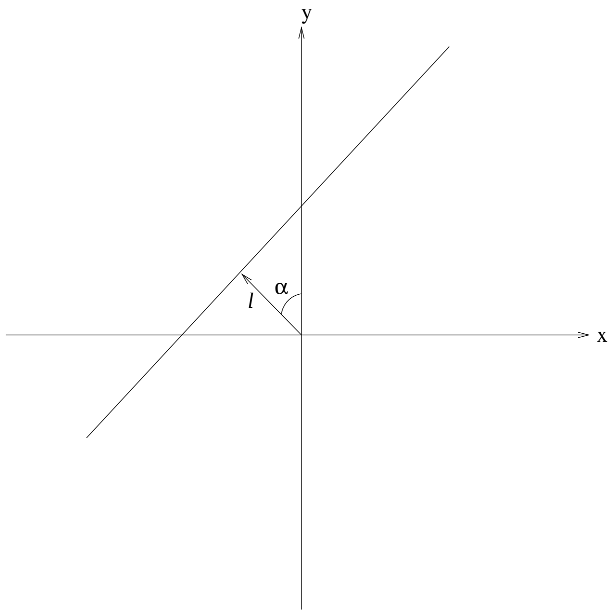

Figure 7: The parameter $\alpha \in [0, 2\pi)$ is the angle that the direction perpendicular to the edge makes with the y-axis, and the parameter $l \in (-3/2, 3/2)$ is the displacement of the edge from the origin.

of the edge patches, we arrive at a set of points $\mathbf{v}_{\alpha,l} \in \mathrm{S}^7 \subset \mathbb{R}^8$. It can be shown that the loci of these points, with $\alpha \in [0, 2\pi)$ and $l \in (-3/2, 3/2)$, define a C^1 2-dimensional manifold, M^2, embedded in the 7-dimensional sphere.

Figure 8 shows a few examples of step edges between $0 \leq \alpha \leq 180$ degrees and $1.5 < l < -1.5$ on the manifold. The step edges are here chosen on a triangular grid with $\Delta_\alpha = 15$ degrees and $\Delta_l = 1/4$ pixel units. Numbers between patches represent the angular distances in degrees between nearest neighbors in this grid. Although we use an even sampling of grid points in the (α, l)-coordinate system, the distances between the nearest neighbors vary widely. Note that, because of the contrast normalization, the patches are degenerate for $-1.5 < l < -0.5$ and $0.5 < l < 1.5$, when $\alpha = 0$, 90, 180, and 270 degrees (horizontal and vertical edges). In the figure, the distances in the "l-direction" between these patches are zero.

Figure 9 shows the geometry of the surface of blurred step edges more clearly: Here we estimate the surface metric $f(\alpha, l) = \frac{dA}{d\alpha\, dl}$ numerically with a grid that is rectangular and more finely spaced than in the example above. We divide the surface into rectangular

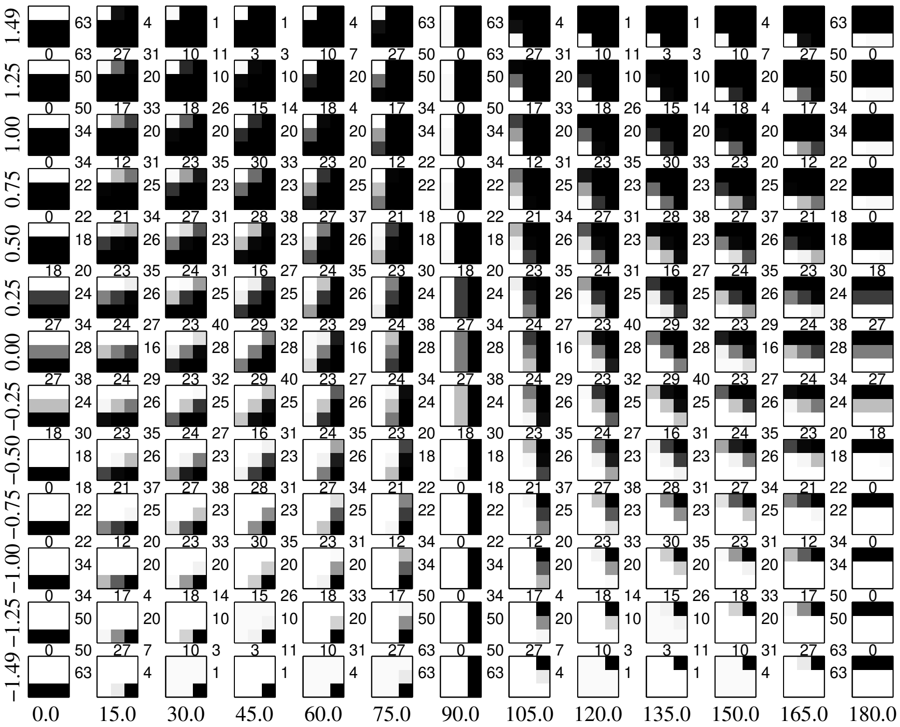

Figure 8: Examples of step edges between $0 \leq \alpha \leq 180$ degrees (horizontal axis) and $1.5 < l < -1.5$ (vertical axis). The step edges are here chosen on a triangular grid with $\Delta_\alpha = 15$ degrees and $\Delta_l = 1/4$ pixel units. Numbers between patches represent the angular distances in degrees between neighbors in the horizontal, vertical and (lower left - upper right) diagonal directions in the (α, l)-grid.

bins with sides $\Delta_\alpha = 3.75$ degrees and $\Delta_l = 1/16$ pixels, and calculate the area of the flat polygons in $\mathbb{R}^8$ that have 4 neighboring grid points as vertices [3]. Because of the contrast normalization, $f(\alpha, l) = 0$ for horizontal and vertical edges ($\alpha = 0$, 90, 180, and 270 degrees) when $-1.5 < l < -0.5$ or $0.5 < l < 1.5$ pixel units.

5.2 Density as a Function of the Distance to the Surface

We can estimate the probability density around the ideal surface from our data. In the following experiment, we use 3 million ($N_{\text{tot}} = 3 \cdot 10^6$) high-contrast patches from optical images; see "optical dataset" in Section 2. We model the surface with a triangulated mesh where the vertices are blurred step edges. We start with a triangulation on a

[3]This leads to an underestimate of the surface area on the 7-sphere $S^7 \subset \mathbb{R}^8$.

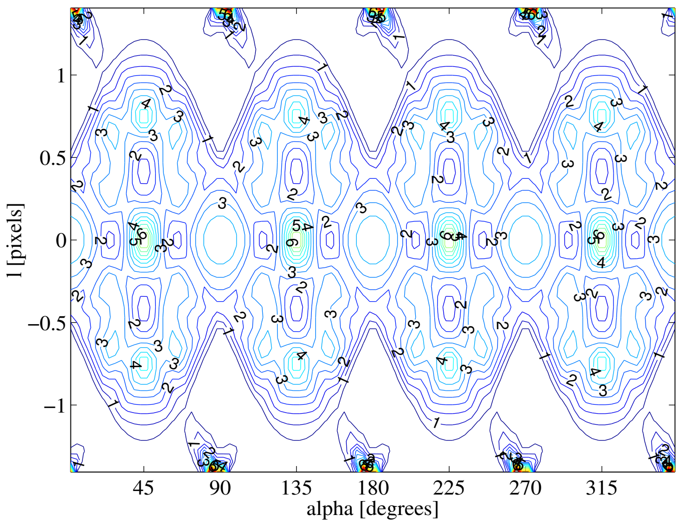

Figure 9: Contour plot of the metric $f(\alpha, l) = \frac{dA}{d\alpha\, dl}$ for a surface of step edges. For the calculation of the area A, we first divide the surface into rectangular bins with sides $\Delta_\alpha = 3.75$ degrees and $\Delta_l = 1/16$ pixels, and then calculate the area of the flat polygons that have 4 neighboring grid points as vertices.

uniform (α_i, l_i)-grid with $\Delta_\alpha = 7.5$ degrees and $\Delta_l = 1/8$ pixel units. The mesh is then successively refined with Delaunay triangulation until any two points in a triangle is less than 10 degrees. In this case, the final mesh has 4392 triangles.

For each data point $\mathbf{x}_n$ $(n = 1, 2, \ldots, N_{\text{tot}})$, we calculate the distance to the closest triangle (orthogonal projection) on the mesh. Figure 10(a), top, shows a normalized histogram of the number of data points as a function of the distance θ to the surface in degrees. Let

$$N(\theta) = \# \left\{ n : \theta - \frac{\Delta_\theta}{2} \leq \text{dist}(\mathbf{x}_\text{n}, \text{M}^2) < \theta + \frac{\Delta_\theta}{2} \right\} , \tag{17}$$

where Δ_θ is the bin width of the histogram. Linear regression (Figure 10(b), top) gives that $N(\theta) \propto \theta^{1.2}$ for small θ.

For the density calculation, we also need an estimate of the volume of the set

$$B_\theta = \left\{ \mathbf{x} \in \text{S}^7 : \theta - \frac{\Delta_\theta}{2} \leq \text{dist}(\mathbf{x}, \text{M}^2) < \theta + \frac{\Delta_\theta}{2} \right\} . \tag{18}$$

Figure 10(a), bottom, and Figure 10(b), bottom, show the results from a Monte Carlo simulation with $V_{\text{tot}} = 10^7$ sample points, $\mathbf{s}_n$ $(n = 1, 2, \ldots, V_{\text{tot}})$, that are uniformly randomly distributed on a 7-dimensional unit-sphere. The number of sample points in B_θ, i.e.

$$V(\theta) = \#\left\{ n : \theta - \frac{\Delta_\theta}{2} \le \mathrm{dist}(\mathbf{s}_\mathrm{n}, \mathrm{M}^2) < \theta + \frac{\Delta_\theta}{2} \right\} . \tag{19}$$

is directly proportional to the *volume* of B_θ. As expected[4], the number of random points V increases as $V(\theta) \propto \theta^4$ for small θ. The curve for $V(\theta)$ has a maximum around $\theta = 43$ degrees, after which it drops. The drop may indicate folds in the surface where part of the 7-sphere are at equal distance to different points of the surface. Furthermore, the plot shows that all points in S^7 are within approximately 60 degrees of M^2.

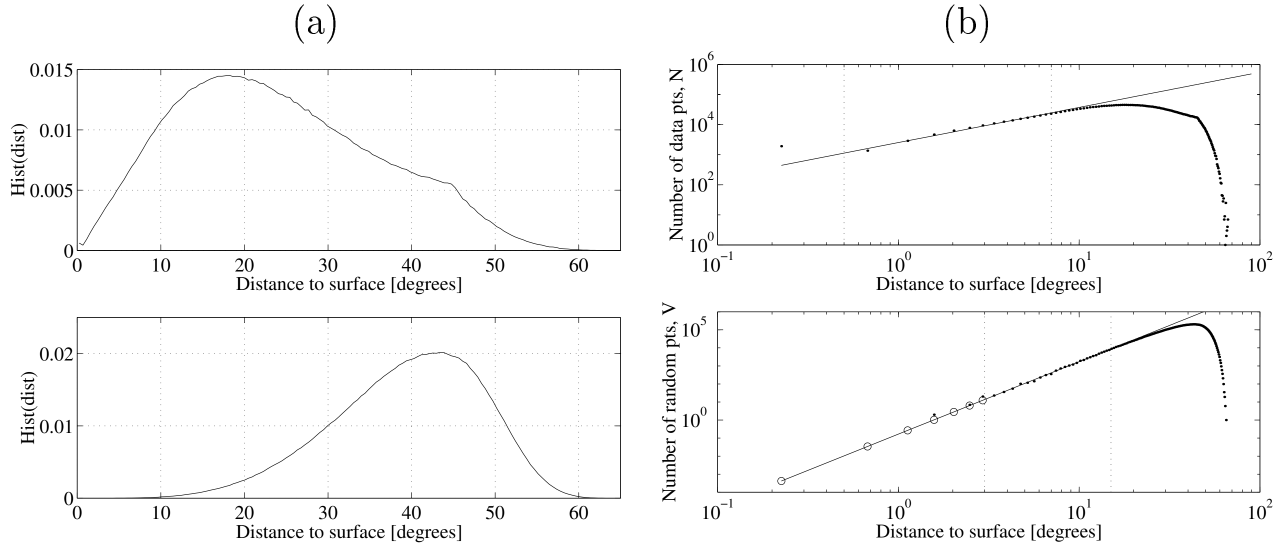

Figure 10: (a) *Top*: Normalized histogram of the number of data points, N, as a function of the distance, θ, to the surface. *Bottom*: Normalized histogram of the number of random samples, V, as a function of θ. (b) *Top*: Log-log plot of N versus θ. Linear regression in the interval between 0.5 and 7 degrees gives $N \sim \theta^{1.2}$ (solid line). *Bottom*: Log-log plot of V versus θ. Linear regression in the interval between 3 and 15 degrees gives $V \sim \theta^{4.0}$ (solid line). The circles represent extrapolated values of V for $\theta \le 3$ degrees. The total number of natural data points is $N_{\mathrm{tot}} \approx 3 \cdot 10^6$. The total number of random samples is $V_{\mathrm{tot}} = 10^7$. The random samples are uniformly distributed on a 7-dimensional sphere, and give a Monte Carlo estimate of the volume of the space of the data points.

[4]For a single point $\mathbf{x}_0 \in \mathrm{S}^7$, the volume

$$\mathrm{vol}\{\mathbf{x} \in \mathrm{S}^7 : \theta \le \mathrm{dist}(\mathbf{x}, \mathbf{x}_0) \le \theta + \mathrm{d}\theta\} = \mathrm{C}_6\, (\sin\theta)^6 \mathrm{d}\theta\ ,$$

where C_6 is the volume (surface area) of a 6-dimensional unit sphere. The 2-dimensional manifold $\mathrm{M}^2 \subset \mathrm{S}^7$ has 5 normal directions. Hence, for small θ, where the sub-manifold is locally linear, $\mathrm{vol}(X_\theta) \propto (\sin\theta)^4\, d\theta$.

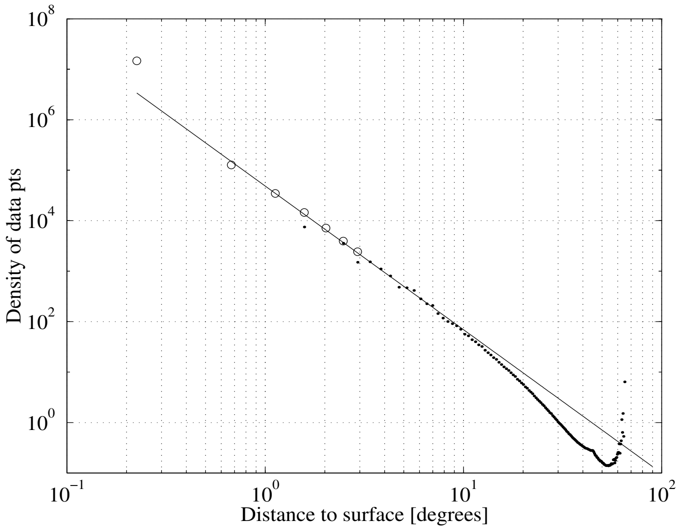

Figure 11: Log-log plot of the density function $p(\theta)$ for optical data. For $\theta \lesssim 10$ degrees, $p(\theta) \sim \theta^{-2.8}$ (solid line). The circles represent the values of the ratio $\frac{N(\theta)/N_{\mathrm{tot}}}{V(\theta)/V_{\mathrm{tot}}}$ for extrapolated values of $V(\theta)$ when θ is small (Figure 10b, bottom).

Figure 11 shows the density function $p(\theta) = \frac{N(\theta)/N_{\mathrm{tot}}}{V(\theta)/V_{\mathrm{tot}}}$. For $\theta \lesssim 10$ degrees,

$$p(\theta) \sim \theta^{-2.8} \ . \tag{20}$$

This result strongly supports the idea that there exists a 2-dimensional manifold in the 7-sphere where the data points are concentrated. In fact, we here see evidence of an *infinite* probability density at the surface of step edges.

In Figure 12, top, we have plotted the percentage of data points that are *within* a certain distance θ of the surface. The curve shows that about 50% of the data points are within a tubular neighborhood

$$K_\theta = \{\mathbf{x} \in \mathrm{S}^7 : \mathrm{dist}(\mathbf{x}, \mathrm{M}^2) \leq \theta\} \tag{21}$$

of the surface with width $\theta = 22$ degrees. This neighborhood corresponds to only 4% of the total volume of S^7 (Figure 12, bottom).

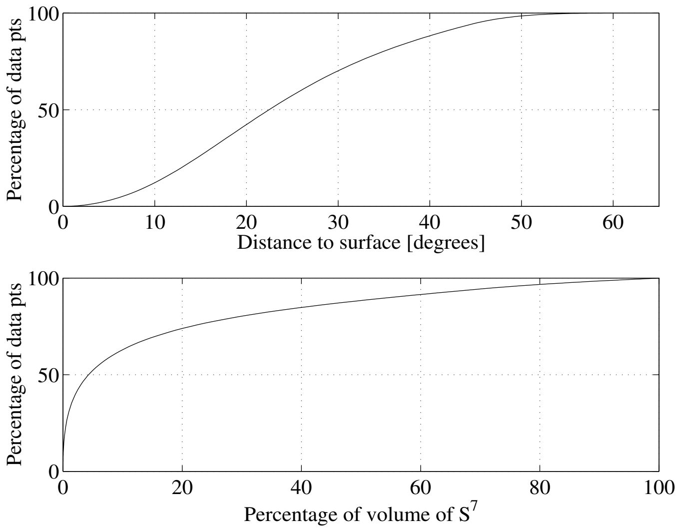

Figure 12: (a) Percentage of data points that are within a tubular neighborhood K_θ of the surface with width θ ("distance to surface"). (b) Same curve plotted versus the ratio $\frac{100\cdot\mathrm{vol}(K_\theta)}{\mathrm{vol}(S^7)}$ ("percentage of volume of S^7").

5.3 Density as a Function of the Surface Coordinates

We next study how the data points are spread out along the manifold. For the position calculation, we choose blurred step edges $\mathbf{v}_{\alpha_i,l_i}$ on a rectangular (α_i, l_i)-grid with the uniform grid spacing $\Delta_\alpha = 3.75$ degrees and $\Delta_l = 1/16$ pixel units.

For each data point $\mathbf{x}_n$ in our data set, we find the closest grid point $\mathbf{v}_{\alpha_n,l_n}$ in the mesh. We compute the 2D histogram

$$N(\alpha_i, l_i) = \#\left\{n\,:\,d(\mathbf{x}_n,\mathbf{v}_{\alpha_n,l_n})\leq\theta_{\max},\ \alpha_n\in\left[\alpha_i-\frac{\Delta_\alpha}{2},\alpha_i+\frac{\Delta_\alpha}{2}\right),\ l_n\in\left[l_i-\frac{\Delta_l}{2},l_i+\frac{\Delta_l}{2}\right)\right\}\ , \tag{22}$$

where $\theta_{\max} = 20$ degrees.

We define the density $\rho(\alpha_i, l_i)$ of data points along the surface as

$$\rho(\alpha_i, l_i) = \frac{N(\alpha_i, l_i)}{f(\alpha, l)\,\Delta_\alpha\Delta_l}\ , \tag{23}$$

where $f(\alpha_i, l_i)$ is given in Section 5.1.

Figure 13 shows the results for regions where $f(\alpha_i, l_i) > 1$ (see Figure 9 for the surface metric). Although the data points are spread out along the whole surface there is a clear

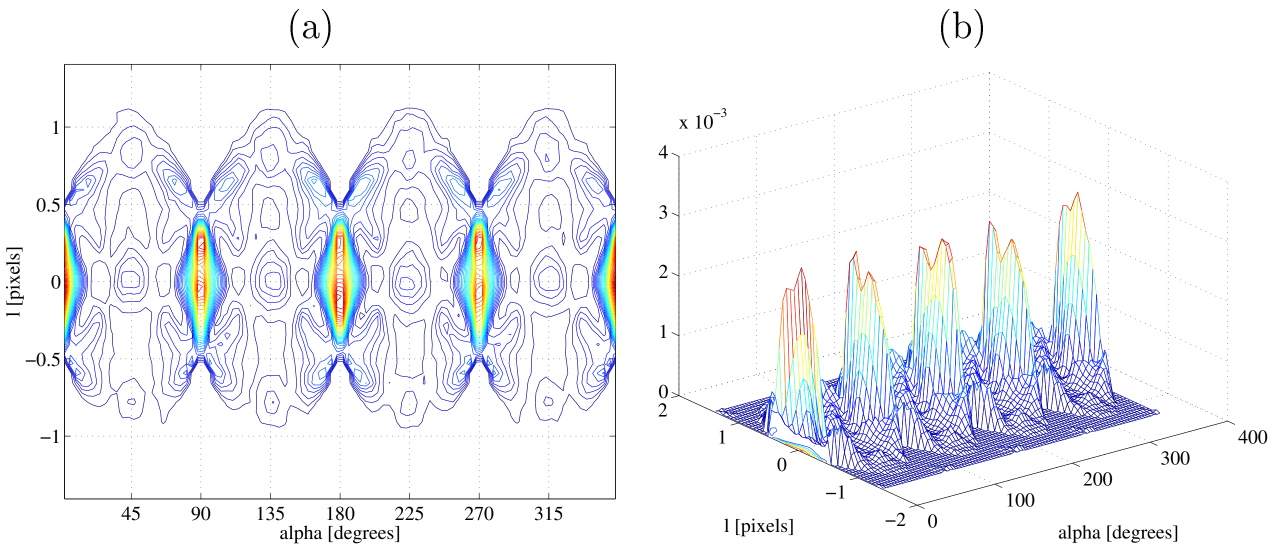

Figure 13: (a) Two-dimensional contour plot and (b) three-dimensional mesh of the density $\rho(\alpha, l)$ of data points along the surface of step edges. In the density calculation, we only include data points that are, at a position (α_i, l_i) where $f(\alpha_i, l_i) > 1$, and in a tubular neighborhood of the surface with width $\theta_{\max} = 20$. The bin widths $\Delta_\alpha = 3.75$ degrees and $\Delta_l = 1/16$ pixel units.

concentration of data points at $\alpha = 0$, 90, 180, and 270 degrees (vertical and horizontal edges).

5.4 Range Data: Probability Density as a Function of Distance to the Surface

A probability density estimation of range data around the surface of step edges leads to very different results.

In Figure 14 we show the normalized histogram $N(\theta)$ of range patches as a function of the distance θ to the surface; compare with Figure 10(a), top, for optical data. The histogram for range data has sharp peaks at $\theta \approx 0, 11, 22, 24, \ldots$ degrees. These peaks in the density function indicate that there are *clusters* in S^7 where the data points are concentrated. A more detailed analysis shows that the positions of the local maxima approximately correspond to the distances between binary patches and the surface of step edges.

For 3×3 patches, there are 510 binary patches which belongs to 50 symmetry groups,

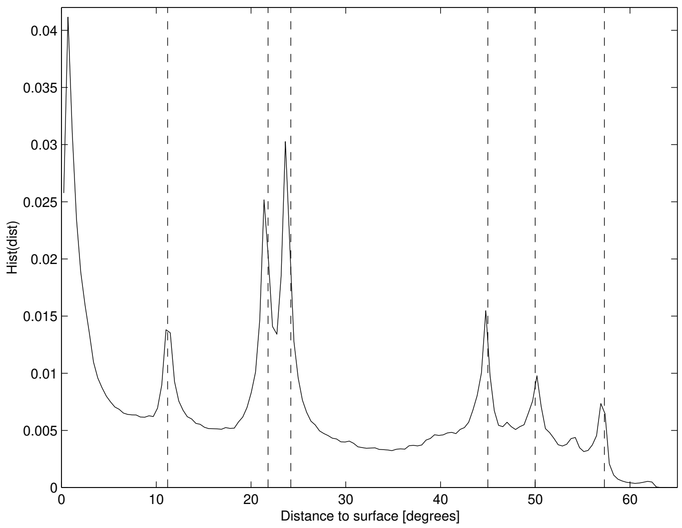

Figure 14: Normalized histogram $N(\theta)$ for range patches as a function of distance θ to the surface of step edges. (Dashed lines) Binary symmetry groups no. 3, 4, 5, 13/17, 27, and 42 (see also Figure 15).

[12], with respect to rotation, reflection, and intensity inversion. Grouping the 510 binary patches into these symmetry groups lets us focus on the structural differences of binary patches. We project the 510 binary patches onto S^7 by centering and sphering as discussed in Section 3 and then compute their distances to the surface of step edges.

In Figure 15, we have sorted the symmetry groups according to their distances to the surface of step edges. Note that groups 1 and 2 are exactly on the surface. Group 3 is at the same distance from the surface as the second local maximum at $\theta \approx 11$ (Figure 14), group 4 and 5 correspond to the third and fourth local maxima at $\theta \approx 22$ and $\theta \approx 24$. The peak at $\theta \approx 45$ may be due to, for example, symmetry groups 13 (horizontal and vertical bars) and 17 (blobs). The peaks at $\theta \approx 50$ and $\theta \approx 57$ could be signs of patterns similar to symmetry group 27 (diagonal edges) and 42 (single dots), respectively.

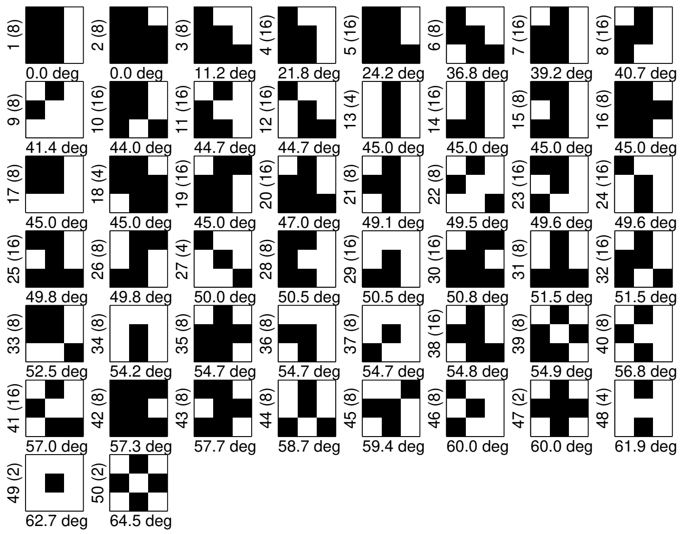

Figure 15: The 50 binary symmetry groups arranged according to increasing θ-values, where θ is the angular distance to the closest point on the surface of step edges. The number below each displayed patch represents this distance. The number in parenthesis left of each patch represents the number of binary patches in the symmetry group.

6 Range Data: Probability Density Around Binary Patches

In the previous section we saw that high-contrast range image patches are concentrated in clusters both on and around the surface of blurred step edges. Furthermore, these clusters appear to be centered around binary patches. This motivates us to model the density of high-contrast range patches around binary patches.

6.1 Density as a Function of Distance to Nearest Binary Patch

We want to calculate the density of high-contrast range patches $\mathbf{x}$ as a function of the angular distance θ to the *closest* binary patch $\mathbf{x}_0 \in \mathrm{S}^7$. We define the density of range patches as a function of distance as

$$p(\theta_i) = \frac{N(\theta_i)}{V(\theta_i)}, \tag{24}$$

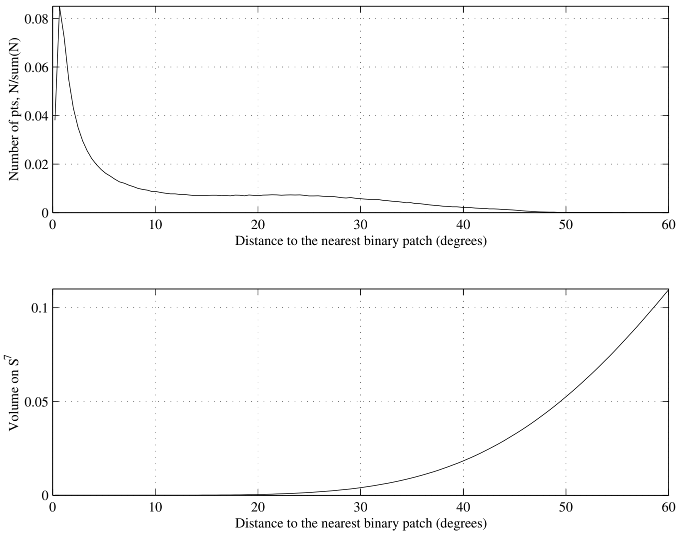

Figure 16: (Top) Histogram $N(\theta_i)$ of high-contrast range patches with respect to distance θ_i to nearest binary patch. (Bottom) Volume $V(\theta_i)$ of bin θ_i.

where

$$N_i = \# \left\{ n : \theta_i - \frac{\Delta\theta}{2} \leq \mathrm{dist}(\mathbf{x}_\mathrm{n}, \mathbf{x}_0) < \theta_\mathrm{i} + \frac{\Delta\theta}{2} \right\} \tag{25}$$

is a histogram of range patches $\mathbf{x}_n$ ($n = 1, 2, \ldots, N_{\mathrm{tot}}$) that are at a distance $\theta_i - \Delta\theta/2 < \phi \leq \theta_i + \Delta\theta/2$ to the nearest binary patch, and N_{tot} is the total number of range patches.

The volume

$$V_i = \mathrm{vol}\{\mathbf{x}_n \in \mathrm{S}^7 : \theta_i \leq \mathrm{dist}(\mathbf{x}_\mathrm{n}, \mathbf{x}_0) < \theta_\mathrm{i} + \mathrm{d}\theta\} \tag{26}$$

of bin θ_i can be computed according to

$$V(\theta_i) = \frac{16\pi^3}{15} \int_{\theta_i - \Delta\theta/2}^{\theta_i + \Delta\theta/2} \sin^6(\phi)\, d\phi. \tag{27}$$

In Figure 16 we plot the bin volume $V(\theta_i)$ versus the distance θ_i.

For each high-contrast range patch we have computed the angular distance (Equation 11) to each of the 510 binary patches. In Figure 16 we show the histogram $N(\theta_i)$ of range patches with respect to the distance to the nearest binary patch.

In Figure 17 we show the log-log plot of the density $p(\theta_i)$ of high-contrast range patches with respect to the distance θ_i to the nearest binary patch. The curve is straight for more

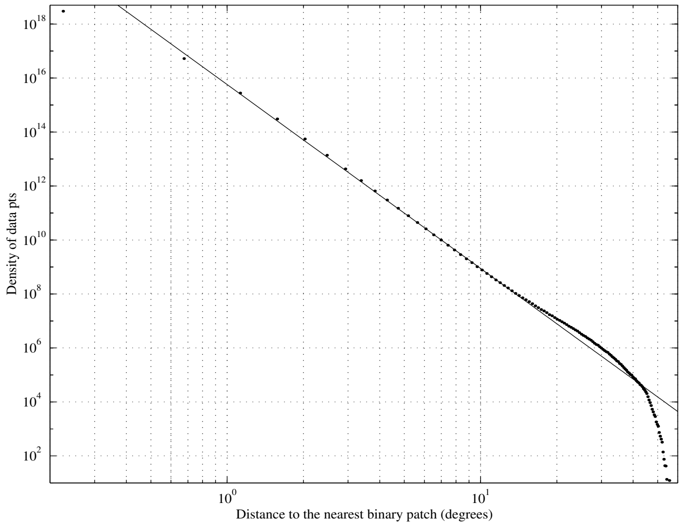

Figure 17: Log-log plot of the density $p(\theta_i)$ of high-contrast range patches with respect to θ_i, the distance to the nearest binary patch. Linear regression in the interval 0.6° and 10° gives $p(\theta) \sim \theta^{-6.8}$ (see solid line).

than a decade of distances, from 0.6 to 10 degrees. Linear regression gives that

$$p(\theta) \sim \theta^{-6.8}. \tag{28}$$

This plot indicates that the probability density is infinite at the 510 binary patches.

In Figure 18 we show the cumulative sum of the number of patches $N_{\text{cum}}(\theta_i) = \sum_{j \leq i} N(\theta_j)$ with respect to the distance to the nearest binary patch, as well as the cumulative volume $V_{\text{cum}}(\theta_i) = \sum_{j \leq i} V(\theta_j)$ versus $N_{\text{cum}}(\theta_i)$. We see that 90% of the high-contrast patches are within a neighborhood of 30° of one of the 510 binary patches. The neighborhoods of these 90% patches occupy only 0.3% of the total volume of S^7.

6.2 Distribution of Range Patches Across the 50 Binary Symmetry Groups

We want to study how the geometrical structure varies among high-contrast range patches. This is done by analyzing the distribution of the range patches with respect to the 50 binary symmetry groups. As mentioned previously, each symmetry group describes binary patches with a certain geometry.

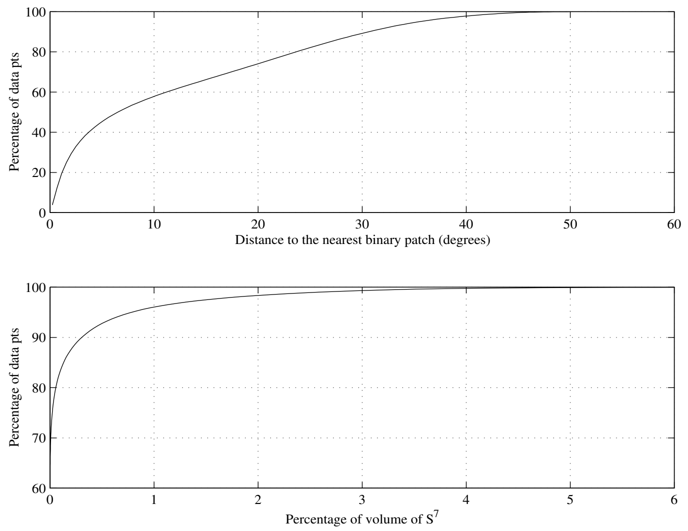

Figure 18: (Top) The cumulative sum of the number of patches $N_{\mathrm{cum}}(\theta_i)$ with respect to the distance to the nearest binary patch. (Bottom) The cumulative volume $V_{\mathrm{cum}}(\theta_i)$ versus $N_{\mathrm{cum}}(\theta_i)$.

For each high-contrast range patch we determine the closest binary patch and its corresponding symmetry group. We then count the number of range patches closest to each of the 50 binary symmetry groups; see the normalized histogram in Figure 19. The figure shows that high-contrast range patches cluster around only a few of the 50 symmetry groups, in fact, 80% of the patches are closest to 8 of the symmetry groups.

Figure 20 shows the 50 symmetry groups ordered after the percentage of range patches closest to each group. For each symmetry group we plot the most frequent patch of the binary patches within that group. The most common structures among high-contrast patches are horizontal edges followed by slanted edges and corner-like structures. The least probable structures are checkerboard and cross patches.

7 Summary and Conclusions

We have analyzed the probability distribution of 3×3 high-contrast patches from optical and 3D range images. In the preprocessing stage, we centered and sphered the log-intensity or log-range values, respectively, of these patches, thereby representing each patch with a point on a 7-dimensional unit sphere.

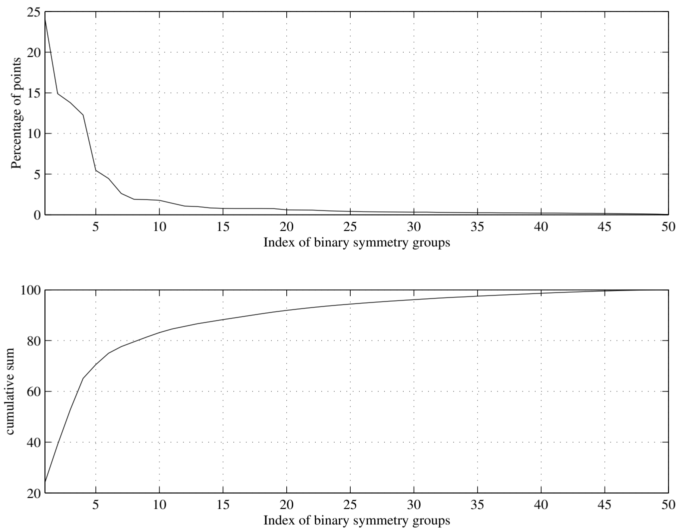

Figure 19: (Top) Histogram of high-contrast range patches closest to each of the 50 binary symmetry groups. (Bottom) Cumulative sum of the number of closest patches versus the index of the symmetry groups.

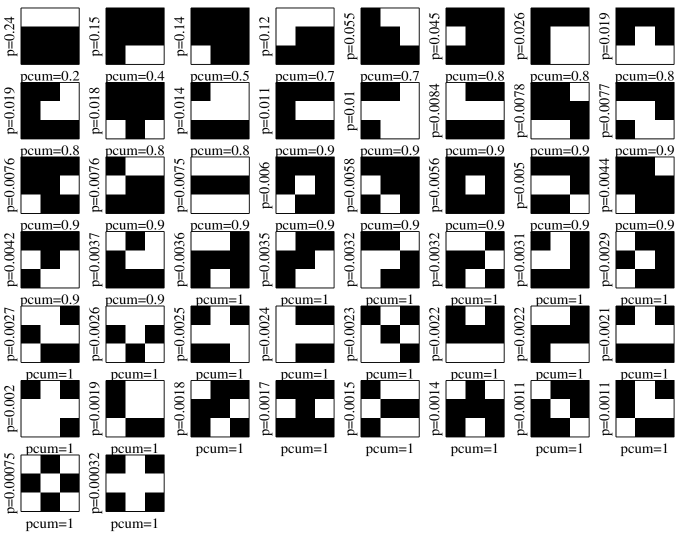

Figure 20: The 50 binary symmetry groups ordered after the percentage "p" of high-contrast range patches closest to one of the binary patches in the groups. The cumulative percentage "pcum" is also indicated. For each group, we display the binary patch which is most common among the high-contrast range patches.

As a first exploration, we examined how these patches distribute with respect to an approximately uniform Voronoi sampling of the 7-sphere. The analysis showed that both optical and range patches occupy a very small amount of the total surface area (volume) of the 7-sphere — in both cases half of the data can be divided into a set of approximately equally sized Voronoi cells with a total volume of less than 5% of the volume of the 7-sphere. For optical patches, the centers of the most densely populated Voronoi cells resemble blurred step edges. For range patches they resemble binary patterns.

A more detailed analysis showed clear differences in the probability distributions of optical and range patches:

The majority of high-contrast optical patches are concentrated around a 2-dimensional C^1 *submanifold* embedded in the 7-dimensional sphere. This surface is highly *non-linear* and corresponds to ideal step edges blurred by the optics of the camera. A density calculation shows that the probability density of optical patches is infinite on this ideal surface. About 50% of the optical data points are within a tubular neighborhood of the surface with a width that corresponds to only 4% of the total volume of the 7-sphere.

The majority of range patches, on the other hand, are concentrated in *clusters*, rather than on a smooth manifold. The centers of these clusters seem to correspond to binary patches, i.e. patches with only two range values. A density calculation around the 510 possible binary patches indicated an infinite probability density at these "hot spots" of the image space. About 90% of the high-contrast patches are in a neighborhood of these spots that correspond to only 0.3% of the total volume of the 7-dimensional unit sphere.

The differences between optical and range patches are partly due to differences in subresolution properties. In optical images, the subpixel details are averaged by the point-spread function of the camera. In range images, on the other hand, the pixel values usually correspond to the minimum of the subresolution details. We also believe that "objects", i.e. correlated regions with sharp edges between different regions, are often more well-defined in terms of range, than in terms of changes in intensity, color, texture, and lighting.

In this paper, we find a way of systematically distinguishing between different types of images that may have seemingly similar second-order statistics and scaling properties. Previous work on natural image statistics have mainly focused on 1D or 2D histograms of filter reactions. Few attempts have been made to understand the full probability distribution of images or image patches. The main idea behind this study is that "object-like" structures in the world create strong pixel correlations in natural images, generating data in clusters or (non-linear) low-dimensional manifolds. There is no doubt that a knowledge of the intrinsic dimension of the image space and the structures where the data is concentrated, will be useful in developing better image representation and compression schemes. Most current approaches, for example, involve a *linear* decomposition of images with respect to a suitable set of basis functions. Although the analysis in the current paper only deals with 3×3-patches, we believe that the more general results, such as the existence of low-dimensional non-linear manifolds and clusters, can be generalized to larger patches. An analysis of, for example, 5×5 patches, would certainly lead to a larger variety of structures (corresponding to bars, blobs, curved edges etc.) and can be performed with similar methods without any major computational efforts. Alternatively, one could also analyze the phase space formed by collections of local responses of multiscale filters of wavelet or scale-space theory. The latter methods would most likely lead to similar qualitative results as in this paper.

Our results indicate the importance of studying the complex statistics of natural images, and the need to understand the intrinsic dimensionality and embeddings of natural data. A complete description of the probability distribution of natural image patches, however, requires also modeling low-contrast patches and learning the embeddings and their coordinates in an unsupervised manner (see e.g. the dimensionality reduction algorithms in [23, 26]).

References

[1] R. W. Buccigrossi and E. P. Simoncelli. Image compression via joint statistical characterization in the wavelet domain. *IEEE Trans Image Processing*, 8(12):1688–1701, Dec. 1999.

[2] E. J. Candès and D. L. Donoho. Curvelets – a surprisingly effective nonadaptive representation for objects with edges. In L. L. Schumaker et al., editor, *Curves and Surfaces*. Vanderbilt University Press, Nashville, TN, 1999.

[3] E. J. Candès and D. L. Donoho. Ridgelets: A key to higher-dimensional intermittency? *Phil. Trans. R. Soc. Lond. A*, 1999. Submitted.

[4] R. R. Coifman and D. L. Donoho. Translation invariant de-noising. In A. Antoniadis, editor, *Wavelets and Statistics*. Springer-Verlag, Heidelberg, 1995.

[5] P. Comon. Independent component analysis, a new concept? *Biometrika*, 36:287–314, 1994.

[6] J. H. Conway and N. J. A. Sloane. *Sphere Packings, Lattices and Groups*. Number 290 in Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1988.

[7] D. L. Donoho. Orthonormal ridgelets and linear singularities. *SIAM Journal on Mathematical Analysis*, 31(5):1062–1099, 2000.

[8] D. L. Donoho and I. M. Johnstone. Ideal spatial adaption via wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

[9] D. J. Field. Relations between the statistics of natural images and the response proporties of cortical cells. *Journal of Optical Society of America*, 4(12):2379–2394, December 1987.

[10] L. Florack, B. ter Haar Romeny, M. Viergever, and J. Koenderink. The gaussian scale-space paradigm and the multiscale local jet. *International Journal of Computer Vision*, 18:61–75, 1996.

[11] J. H. Friedman. Exploratory projection pursuit. *J. of the American Statistical Association*, 82(397):249–266, 1987.

[12] D. Geman and A. Koloydenko. Invariant statistics and coding of natural microimages. In *Proc. of the IEEE Workshop on Statistical and Computational Theories of Vision*, 1999. Published on the Web.

[13] J. Huang, A. Lee, and D. Mumford. Statistics of range images. In *Proc. of CVPR'00*, 2000.

[14] J. Huang and D. Mumford. Statistics of natural images and models. In *Proc. of CVPR'99*, 1999.

[15] P. J. Huber. Projection pursuit. *The Annals of Statistics*, 13(2):435–475, 1985.

[16] A. Hyvärinen. Survey on independent component analysis. *Neural Computing Surveys*, 2:94–128, 1999.

[17] P. Johansen, S. Skelboe, K. Grue, and J. D. Andersen. Representing signals by their toppoints in scale space. In *Proceedings of the International Conference on Image Analysis and Pattern Recognition*, pages 215–217. IEEE Computer Society Press, 1986.

[18] J. Koenderink. The structure of images. *Biological Cybernetics*, 50:363–370, 1984.

[19] S. Mallat. A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 11:674–693, 1989.

[20] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 14(7):710–732, 1992.

[21] B. A. Olshausen and D. J. Field. Natural image statistics and efficient coding. *Network: Computation in Neural Systems*, 7(2):333–339, May 1996.

[22] P. Reinagel and A. M. Zador. Natural scene statistics at the centre of gaze. *Network: Computation in Neural Systems*, 10(4):341–350, November 1999.

[23] S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, 2000.

[24] D. L. Ruderman and W. Bialek. Statistics of natural images: Scaling in the woods. *Physical Review Letters*, 73(6):814–817, August 1994.

[25] E. P. Simoncelli. Modeling the joint statistics of images in the wavelet domain. In *Proc. SPIE, 44th Annual Meeting*, volume 3813, pages 188–195, 1999. Denver, CO, July 99.

[26] J. B. Tenenbaum, V. Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323, 2000.

[27] A. Turiel, G. Mato, and N. Parga. The self-similarity properties of natural images resemble those of turbulent flows. *Physical Review Letters*, 80(5):1098–1101, February 1998.

[28] A. Turiel and N. Parga. Multifractal wavelet filter of natural images. *Physical Review Letters*, 85(15):3325–3328, October 2000.

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

[30] A. P. Witkin. Scale space filtering. In *Proc. of the Eighth International Joint Conference on Artificial Intelligence*, volume 2, pages 1019–1023, Karlsruhe, Germany, 1983.

[31] C. Zetzsche, B. Wegmann, and E. Barth. Nonlinear aspects of primary vision: Entropy reduction beyond decorrelation. In *Int'l Symposium, Society for Information Display*, volume 24, pages 933–936, 1993.