# Toward a Full Probability Model of Edges in Natural Images

Kim S. Pedersen[1] and Ann B. Lee[2]

[1] DIKU, University of Copenhagen
Universitetsparken 1,
DK-2100 Copenhagen Ø, Denmark
[2] Division of Applied Mathematics,
Box F, Brown University,
Providence, RI 02912, USA

APPTS Technical Report #02-1, Division of Applied Mathematics,
Brown University, January 2002

**Abstract.** We investigate the statistics of local geometric structures in natural images. Previous studies [12,13] of high-contrast $3 \times 3$ natural image patches have shown that, in the state space of these patches, we have a concentration of data points along a low-dimensional non-linear manifold that corresponds to edge structures. In this paper we extend our analysis to a filter-based multiscale image representation, namely the local 3-jet of Gaussian scale-space representations. A new picture of natural image statistics seems to emerge, where primitives (such as edges, blobs, and bars) generate low-dimensional *non-linear* structures in the state space of image data.

**Keywords:** Natural image statistics, probability model of local geometry, scale-space, image features, biologically-inspired computational models

## 1 Introduction

The study of natural image statistics is an active research area and different approaches have been taken in this field [3,6,7,13,21,23,25]. It has previously been shown that natural image statistics — such as the marginal distributions of image intensity $L$ and the gradient magnitude $\|\nabla L\|$ — are highly non-Gaussian, and approximately invariant to changes in scale [3,16,21]. Roughly speaking, the research in natural image statistics can be divided into two related directions: Researchers such as Zetzsche et al. [26], Simoncelli [23], and Huang et al.[7,6] have looked at the 1D marginal and 2D joint statistics of filter responses for a fixed wavelet basis. They have, for example, explored complex dependencies between pairs of wavelet coefficients at nearby spatial positions, orientations and scales. Others have looked at the state space of image data and tried to find a set of directions (or projections of the data) that lead to an optimal image representation in some sense; see e.g. sparse coding [20] and ICA [2,8].

In this paper, we take a different approach to natural image statistics. We believe that in order to fully understand the statistics of natural images one needs to explore the *full probability distribution* of the salient structures of local image patches. The analysis is free from such restrictive assumptions as independent components, or even linear decompositions of an image into basis images.

Our work is inspired by David Marr's ideas [17] that the structure of images can be described by primitives such as edges, bars, blobs and terminations — the so called "primal sketch". The basic questions we ask are: *What are the probability distributions of Marr's primitives and how are these primitives represented geometrically in the state space of image data?* Can we develop models that will tell us how likely we are to observe a local geometric structure (such as an edge, ridge, blob, corner etc.) of a certain size in an image? Not much work has been done in this direction to our knowledge. Such probability models would, however, be useful as priors in image processing and computer vision applications as diverse as compression, feature detection, segmentation [15], image reconstruction [19] and enhancement [22,27].

To study the distribution of local geometric image structures, one first has to choose a representation that captures the image geometry in a neighborhood of some fiducial point. We need a set of local measurements of the luminance captured through a set of sensors — a sensorium. The concept of a sensorium makes sense both from a biological vision point of view (the receptive fields in the early visual system have been compared to feature detectors [10]) and from a mathematical point of view: Mumford and others [4,18] have argued that an image $I$ is not a function or a pointwise estimate $I(x, y)$ but a Schwartz distribution that can only be probed by averaging (or measuring) $\int\int I(\xi, \eta)\phi_j(x - \xi, y - \eta)d\xi d\eta$ through smooth "test" functions or sensors $\phi_j$.

In our previous study [12,13] of natural image statistics, the sensorium is defined by the sensors in the CCD camera used to collect the images. More specifically, we study the joint statistic of the *pixel* intensity values in high-contrast $3 \times 3$ natural image patches. We found (for optical images) that the state space of the patch data is extremely sparse, with most of the data concentrated around a continuous non-linear manifold in state space. This manifold corresponds to edges of different orientations $\alpha$ and positions $l$.

In this work, we investigate whether our previous results (such as the existence of an ideal edge manifold in state space and the concentration of natural image data around this manifold) generalize to different scales and more general image representations.

In the context of linear Gaussian scale-space theory [9], Koenderink and van Doorn [10] proposed the so-called *local jet* of an image as a biologically plausible representation of local image geometry. In this setting, the sensorium consists of partial derivatives of the Gaussian kernel function. Convolving an image with these kernels is equivalent to measuring the partial derivatives of a coarse-grained representation of the image — the so-called scale-space image. The jet space captures the local geometry in a neighborhood of a point in the image, where

the size of the neighborhood is determined by the standard deviation or "scale" of the Gaussian kernel.

In this work, we choose an image representation defined by Gaussian scale-space image derivatives up to third order — the 3-jet. This sensorium can distinguish between image structures such as edges, ridges, blobs and corners [10], but is blind to structures that require the descriptive power of image derivatives of order higher than 3.

We believe that such a representation of image data has certain advantages compared to many other types of multiscale representations. First of all, the Gaussian scale-space representation gives us a sensible way of defining image derivatives — the scale-space image derivatives. With these derivatives, we can use the language of differential geometry to define and interpret local features in images. The Gaussian kernel and its derivatives are furthermore similar to the receptive fields found in the mammalian visual system [9,10]. Both the mammalian receptive fields and the Gaussian scale-space derivatives are tuned to structures at different scales, orientations and frequencies.

As in [13], we focus our analysis on edge structures. We first define a model of an ideal edge in Gaussian scale-space, and show that the 3-jet representations of edges define a 2-dimensional differentiable manifold in jet space. We then study how empirical data, extracted from a large database of natural images, are distributed in 3-jet space with respect to this manifold. We find, in accordance with previous results in [13], that the natural image data are densely distributed around the edge manifold with a probability density function of the form $\theta^{-\gamma}$, where $\theta$ is the distance to the edge manifold and $\gamma$ is close to 0.7. We furthermore show that the results are approximately invariant to a change of scale.

This work is an attempt to develop a *full probability model of edges* in natural images that is universal, i.e. independent of scale and image representation. In the future, we plan to extend the analysis to representations of other image primitives (such as bars, blobs, and T-junctions).

The organization of the paper is as follows: In Sec. 2 we provide the necessary background on jet spaces and linear Gaussian scale-space theory. We introduce the Gaussian edge model in Sec. 3. In Sec. 4, we describe our image data set, the whitening and contrast normalization of this data, and the results of our analysis. Finally, we finish with concluding remarks (Sec. 5).

## 2 Multi-scale Local Jet

Linear Gaussian scale-space was proposed among others by Koenderink [9] as a sound theoretical framework for doing multi-scale image analysis. The Gaussian scale-space $L : \Omega \mapsto \mathbb{R}$ of an image $f : \Omega \mapsto \mathbb{R}$ (where $\Omega \subseteq \mathbb{R}^2$) can be defined as the solution to the heat diffusion equation

$$\frac{\partial L}{\partial t} = \frac{1}{2}\nabla^2 L$$

under the constraint $L(x, y; s = 0) = f(x, y)$. The *scale* $s \geq 0$ is related to $t$ by $t = s^2$. The Gaussian scale-space representation of an image is thus a convolution

$$L(x, y; s) = \iint_{\Omega} f(\xi, \eta)\phi(x - \xi, y - \eta; s)\, d\xi\, d\eta \tag{1}$$

with a Gaussian kernel function $\phi : \mathbb{R}^2 \mapsto \mathbb{R}$ where

$$\phi(x, y; s) = \frac{1}{2\pi s^2} e^{-\frac{(x^2+y^2)}{2s^2}}. \tag{2}$$

We interpret the parameter $s$ as the measurement scale of the scale-space image $L(x, y; s)$ as it corresponds to the width (or standard deviation) of the "smoothing filter" $\phi(\cdot)$.

An image is in general not a differentiable function, but in scale-space we obtain a family of smoothed versions of the image which are $C^\infty$-differentiable. We can compute partial derivatives $\partial_{x^n}\partial_{y^m} (\equiv \frac{\partial^{n+m}}{\partial x^n \partial y^m})$ of the scale-space representation by convolving the image $f(\cdot)$ with partial derivatives of the Gaussian kernel function $\phi(\cdot; s)$, as

$$L_{x^n y^m}(\cdot; s) \equiv \partial_{x^n}\partial_{y^m}(f * \phi) = f * (\partial_{x^n}\partial_{y^m}\phi)\ .$$

Note that the scale-space derivatives $L_{x^n y^m}$ constitute a scale-space as they also satisfy the previous heat diffusion equation.

Various approaches exist for discretization of scale-space representations (see e.g. [14]). Here we evaluate the convolution in Eq. (1) by multiplying the discrete Fourier transform of the discrete image with discretized Fourier transformed Gaussian derivatives $\partial_{x^n}\partial_{y^m}\phi$.

To be able to compare scale-space derivatives at different scales, it is convenient to use dimensionless coordinates (the so-called natural coordinates) and scale-normalized differential operators [14]. In the $(x, y)$-coordinate system the dimensionless coordinates are given by

$$(x', y') = \left(\frac{x}{s}, \frac{y}{s}\right) \tag{3}$$

and the scale-normalized partial derivatives are

$$L_{x'^n y'^m}(x, y; s) = \partial_{x'^n}\partial_{y'^m} L(x, y; s) = s^{n+m}\partial_{x^n}\partial_{y^m} L(x, y; s). \tag{4}$$

In the rest of this paper we will assume that *all* scale-space derivatives $L_{x^n y^m}$ are scale normalized.

We can describe the local geometry of an image by the so-called local jet [5,10]. Since the scale-space image $L(x, y; s)$ is a smoothed, differentiable version of an image, we can use a Taylor series to describe its behavior around a point $(x_0, y_0)$. For $(x_0, y_0) = (0, 0)$, for example, we have

$$\begin{aligned} L(x, y; s) = L + L_x x + L_y y + \frac{1}{2}(L_{xx}x^2 + 2L_{xy}xy + L_{yy}y^2) \\ + \frac{1}{6}(L_{xxx}x^3 + 3L_{xxy}x^2 y + 3L_{xyy}xy^2 + L_{yyy}y^3) + \ldots \end{aligned} \tag{5}$$

where the scale-space derivatives $L, L_x, L_y, \ldots$ are evaluated at $(x_0, y_0) = (0, 0)$. Consider now the truncated Taylor expansion of degree $k$. The so-called *local k-jet* (of $L(x, y; s)$ at $(x_0, y_0)$) is an equivalence class of smooth functions with respect to the map $j^k L : \mathbb{R}^2 \mapsto \mathcal{J}^k(\mathbb{R}^2 \mapsto \mathbb{R}) \subset \mathbb{R}^N$, $N = (2+k)!/(2k!)$, where

$$j^k L(x, y; s) = (L(x, y; s), L_x(x, y; s), L_y(x, y; s), \ldots, L_{x^n y^m}(x, y; s))^T \quad (6)$$

and $n + m = k$. The space $\mathcal{J}^k(\mathbb{R}^2 \mapsto \mathbb{R})$ of all $k$-jets of functions $\mathbb{R}^2 \mapsto \mathbb{R}$ is sometimes called a *k-jet space*. Images that belong to the same $k$-jet (i.e. the same point in $\mathcal{J}^k(\mathbb{R}^2 \mapsto \mathbb{R})$) "look" the same up to order $k$, in the sense that we can not distinguish between them by only looking at scale-space derivatives up to order $k$. Koenderink and van Doorn [11] named this class a metamer inspired by the terminology of Schrödinger's theory of colorimetry.

Edges, ridges, blobs and corners are all common geometric structures found in images. In this paper, we will focus on the statistics of edges. We limit our analysis to the partial derivatives parameterizing the 3-jet, as the 3-jet captures the characteristics of the above mentioned geometric structures (edges, ridges, blobs, and corners) [10,11].

In the following, we will study the statistics of images mapped into 3-jet space by $\tilde{j}^3 L : \mathbb{R}^2 \mapsto \tilde{\mathcal{J}}^3$, where

$$\tilde{j}^3 L(x, y; s) = (L_x, L_y, L_{xx}, L_{xy}, L_{yy}, L_{xxx}, L_{xxy}, L_{xyy}, L_{yyy})^T \quad (7)$$

and the measurement scale $s > 0$. The scale-space derivatives $L_x, L_y, L_{xx}, \ldots$ are evaluated at $(x, y; s)$, and are *scale normalized* according to Eq. (4), i.e. $L_{x^n y^m}(x, y; s) = s^{n+m} \partial_{x^n} \partial_{y^m} L(x, y; s)$. In the above 3-jet representation, we have excluded the intensity $L(x, y; s)$ of the blurred image, as we are only interested in variations in the local image geometry. The tilde notation is to indicate that $\tilde{\mathcal{J}}^3 \subset \mathcal{J}^3(\mathbb{R}^2 \mapsto \mathbb{R})$.

## 3 The Edge Manifold

Geometric structures corresponding to edges can in jet space be modeled by the scale-space of an ideal step edge. In this section, we show that the contrast-normalized 3-jet representations of edges of different orientations, positions and scales trace out a differentiable 2D submanifold in the jet space $\tilde{\mathcal{J}}^3$.

For convenience, we define the edge model in the local orthonormal $(u, v)$-coordinate system where the $v$-axis has the direction of the local gradient at any point $P_0$ and the $u$-axis is perpendicular, i.e. we define unit vectors

$$e_v = (\cos\alpha, \sin\alpha)^T = \frac{1}{\sqrt{L_x^2 + L_y^2}} (L_x, L_y)^T \Big|_{P_0} \quad (8)$$

and

$$e_u = (\sin\alpha, -\cos\alpha)^T \ . \quad (9)$$

In this coordinate system, an ideal step edge (defined on $\mathbb{R}^2$) has the form

$$f(u,v;l) = \begin{cases} 1 & \text{if } v \geq l \\ 0 & \text{if } v < l \end{cases} \tag{10}$$

where $l \in \mathbb{R}$ is the displacement of the edge from the origin in the $v$-direction. The scale-space representation of the ideal step edge is (according to Eq. (1)) given by

$$G(u,v;l,s) = f(u,v;l) * \phi(u,v;s) = \int_{v'=-\infty}^{v} \psi(v';l,s)\, dv' \tag{11}$$

where $\psi(v;l,s) = \frac{1}{\sqrt{2\pi s^2}} e^{-\frac{(v-l)^2}{2s^2}}$ is a one-dimensional Gaussian kernel centered at $l$. For the scale-normalized partial derivatives of the edge model $G(u,v;l,s)$ along the $u$ and $v$-axes, we have

$$\begin{aligned} G_{u^n}(u,v;l,s) &= 0 \\ G_{v^n}(u,v;l,s) &= s^n \partial_{v^{n-1}} \psi(v;l,s) \end{aligned} \tag{12}$$

for $n \geq 1$.

We now map the edge model $G(u,v;\alpha,l,s)$ into the jet space $\tilde{\mathcal{J}}^3$ by computing the nine components of the map $\tilde{j}^3 G(0,0;\alpha,l,s)$ defined by Eq. (7). Since

$$\begin{aligned} \partial_x &= \cos\alpha\, \partial_v + \sin\alpha\, \partial_u \\ \partial_y &= \sin\alpha\, \partial_v - \cos\alpha\, \partial_u \ , \end{aligned} \tag{13}$$

we get that

$$\begin{aligned} G_{x^m,y^n}(0,0;\alpha,l,s) &= \cos^m\alpha\ \sin^n\alpha\, G_{v^{m+n}}(u,v;l,s)\Big|_{(u,v)=(0,0)} \\ &= s^{m+n} \cos^m\alpha\ \sin^n\alpha\, \partial_{v^{m+n-1}} \psi(v;l,s)\Big|_{v=0} \ . \end{aligned}$$

Denote the map that takes the edge model to the 3-jet space $\tilde{\mathcal{J}}^3$ by $\mathcal{E} : [0,2\pi) \times \mathbb{R} \times \mathbb{R}_+ \backslash \{0\} \mapsto \tilde{\mathcal{J}}^3$, where

$$\begin{aligned} \mathcal{E}(\alpha,l,s) = (&G_x(0,0;\alpha,l,s), G_y(0,0;\alpha,l,s), G_{xx}(0,0;\alpha,l,s), \\ &G_{xy}(0,0;\alpha,l,s), G_{yy}(0,0;\alpha,l,s), G_{xxx}(0,0;\alpha,l,s), \\ &G_{xxy}(0,0;\alpha,l,s), G_{xyy}(0,0;\alpha,l,s), G_{yyy}(0,0;\alpha,l,s))^T. \end{aligned} \tag{14}$$

Although the edge map $\mathcal{E}$ is a function of three variables (the angle $\alpha$, the displacement $l$ and the scale $s$), the loci of all points $\mathcal{E}(\alpha,l,s)$ trace out a *2-dimensional* $C^\infty$ differentiable manifold in $\mathbb{R}^9$ that only depends on $\alpha$ and the ratio $l/s$ (see Appendix A). Note that the edge manifold is periodic in $\alpha$ for fixed $l/s$.

In Fig. 1, we show some image samples that correspond to the edge manifold $\mathcal{E}(\alpha,l,s)$. These samples are for illustration purposes only and correspond to a Taylor expansion around the origin (Eq. (5)) truncated beyond the 3rd order terms.

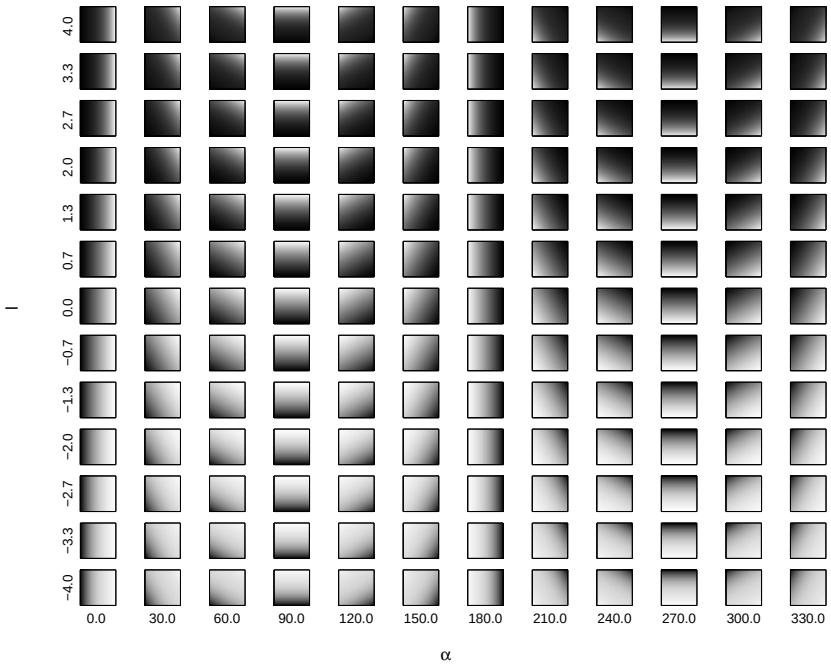


**Fig. 1.** Gaussian edge images. Each image is of size $65 \times 65$ pixels and corresponds to a 3rd order truncated Taylor expansion (Eq. (5)) of a Gaussian edge at scale $s = 1$ pixel.

## 4 Statistics of Edge Structures

### 4.1 The Empirical Data Set

In our experiments, we use the van Hateren still image collection consisting of 4167 $1024 \times 1536$ pixels gray-scale images[1] [24] (see Fig. 2 for samples from the database). Before doing any processing of the images $f(x, y)$ in the database, we compress the intensity range by taking the logarithm $\log(f(x, y) + 1)$ of the intensity.

For each scale-space image $L^{(i)}(x, y; s)$ at a fixed scale $s$, where $s = 1, 2, 4, 8, 16, 32$ pixels and $i = 1, \ldots, 4167$, we extract a random set of 1000 spatial coordinates $X^{(i,s)} \subseteq \Omega$ ($\Omega \subseteq \mathbb{R}^2$ denotes the image domain). At these spatial coordinates, we compute the 3-jet representations defined according to Eq. (7). This gives us data sets

$$J_s = \left\{ \tilde{j}^3 L^{(i)}(x, y; s) \subseteq \tilde{\mathcal{J}}^3 \,\middle|\, (x, y) \in X^{(i,s)}; i = 1, \ldots, 4167 \right\} \tag{15}$$

[1] We use the raw image set (.iml) where the intensity values have been linearized by the camera's lookup table.

**Fig. 2.** Sample images from the van Hateren still image collection. We show the log-transformed intensity values, $\log(f(x,y)+1)$.

where $s = 1, 2, 4, 8, 16, 32$ pixels. Elements in each set $J_s$ are points in $\tilde{\mathcal{J}}^3$ that have been sampled from different spatial positions and different images at a fixed scale $s$. The total number of data points [2] in each set $J_s$ is $|J_s| \approx 4.1 \cdot 10^6$.

### 4.2 Whitening and Contrast Normalization

The lighting conditions may vary across and between images. We are interested in variations in the local *geometry* of the image and would like to disregard variations caused by changing lighting.

Before contrast-normalizing (or more precisely dividing out the multiplicative factors in the scale-space derivatives) we need to first whiten the data. This will lead to a vector representation of the 3-jets where the elements are uncorrelated and of the same order of magnitude.

Assume that $\boldsymbol{x} \in J_s$ where $J_s$ is our data set (Eq. (15)). The covariance or correlation matrix $C = < \boldsymbol{x}\boldsymbol{x}^T >$ is scale invariant, so we can get a robust estimate of $C$ from the joined data set $\bigcup_s J_s$ where $s$ denotes the scale. The mean $< \boldsymbol{x} > = \mathbf{0}$ due to the convolution with mean-zero scale-space filters.

The first step in the data preprocessing is to define transformed input variables

$$\boldsymbol{y} = \Lambda^{-1/2} U^T \boldsymbol{x} \ , \tag{16}$$

where $U$ is a $9 \times 9$-matrix with the normalized eigenvectors of $C$ as columns, and $\Lambda$ is a diagonal $9 \times 9$-matrix with the corresponding eigenvalues of $C$ as diagonal elements. The transformed data $\boldsymbol{y}$ is "white" in the sense that the covariance matrix $< \boldsymbol{y}\boldsymbol{y}^T > = \mathbf{1}$.

The second step is to contrast-normalize the data according to

$$\hat{\boldsymbol{p}} = \frac{\boldsymbol{y}}{\|\boldsymbol{y}\|} \tag{17}$$

[2] To prevent numerical problems during contrast normalization (see Sec. 4.2 and Eq. (17)), we discard data points $\boldsymbol{y}$ with a norm $\|\boldsymbol{y}\|$ that is close to zero after whitening. This corresponds to 1% of all data points.

so that scale-space images of similar geometric structure have the same representation. The whitened and contrast-normalized data points $\hat{\boldsymbol{p}}$ all lie on a 8-dimensional unit sphere

$$\mathrm{S}^8 \equiv \{\hat{\boldsymbol{p}} \mid \|\hat{\boldsymbol{p}}\| = 1\} \in \mathbb{R}^9 \,. \tag{18}$$

The 8-sphere $\mathrm{S}^8$ is the *state space of whitened and contrast-normalized* 3-jet representations. The whitened and contrast-normalized data set $J_s$ at a fixed scale $s$ is denoted by $\hat{J}_s \in \mathrm{S}^8$.

Similarly, we define the map $\hat{\mathcal{E}} : [0, 2\pi) \times \mathbb{R} \mapsto \mathrm{S}^8$ that takes the edge map $\mathcal{E}(\alpha, l, s)$ (Eq. (14)) to the state space of whitened and contrast normalized 3-jet representations by

$$\hat{\mathcal{E}}(\alpha, l/s) = \frac{\Lambda^{-1/2} U^T \mathcal{E}(\alpha, l/s, 1)}{\|\Lambda^{-1/2} U^T \mathcal{E}(\alpha, l/s, 1)\|}. \tag{19}$$

We measure the *distance* between two data points $\hat{\boldsymbol{p}}_0, \hat{\boldsymbol{p}}_1 \in \mathrm{S}^8$ by their angular separation, i.e.

$$\mathrm{dist}(\hat{\boldsymbol{p}}_0, \hat{\boldsymbol{p}}_1) \equiv \arccos(\hat{\boldsymbol{p}}_0^T \hat{\boldsymbol{p}}_1). \tag{20}$$

### 4.3 Density Results for the Empirical Data Set

In Sec. 3, we described the theoretical manifold of edges in 3-jet space. In this section, we verify that the *empirical* data from natural images (i.e. the whitened and contrast-normalized data in sets $\hat{J}_s$ from Sec. 4.1 and Sec. 4.2) are really densely distributed around the manifold of edges. By putting parallel bins around the manifold and computing histograms of the data, we can get an estimate of the functional form of the probability density around the surface.

First we divide the whitened and constrast-normalized edge manifold $\hat{\mathcal{E}}(\alpha, l/s)$ of Eq. (19) into a mesh of spherical triangles in the same fashion as in [13]. We sample the edge manifold with parameters $l \in [-4s, 4s]$ and $\alpha \in [0, 2\pi)$. We refine the mesh of triangles until no vertices in a triangle are more than 11 degrees apart. This gives us a triangulated mesh with a total of 22944 triangles. We use the triangulated mesh of the manifold to estimate the distances between the data points of $\hat{J}_s$ and the manifold of edges $\hat{\mathcal{E}}(\alpha, l/s)$ on the 8-sphere $\mathrm{S}^8$. The distance $\mathrm{dist}(\boldsymbol{x}, \hat{\mathcal{E}}(\alpha, l/s))$ between a data point $\boldsymbol{x} \in \hat{J}_s$ and the edge manifold $\hat{\mathcal{E}}(\alpha, l/s)$ is approximated by the distance to the center point of the closest triangle in the mesh.

Fig. 3 (top) shows a normalized histogram of the number of whitened and contrast-normalized data points $\hat{\boldsymbol{p}}_n^s \in \hat{J}_s$ $(n = 1, \ldots, |\hat{J}_s|)$ versus the distance $\theta$ to the edge manifold $\hat{\mathcal{E}}(\alpha, l/s) \subset \mathrm{S}^8$. Let

$$N(\theta; s) = \# \left\{ n \middle| \theta - \frac{\Delta\theta}{2} \leq \mathrm{dist}(\hat{\boldsymbol{p}}_n^s, \hat{\mathcal{E}}(\alpha, l/s)) < \theta + \frac{\Delta\theta}{2} \right\} \tag{21}$$

where $\Delta\theta$ is the histogram bin width and $\mathrm{dist}(\hat{\boldsymbol{p}}_n^s, \hat{\mathcal{E}}(\alpha, l/s))$ is the angular distance (Eq. (20)) from the data point $\hat{\boldsymbol{p}}_n^s$ to the closest point on the triangulated mesh of the edge manifold $\hat{\mathcal{E}}(\alpha, l/s)$.

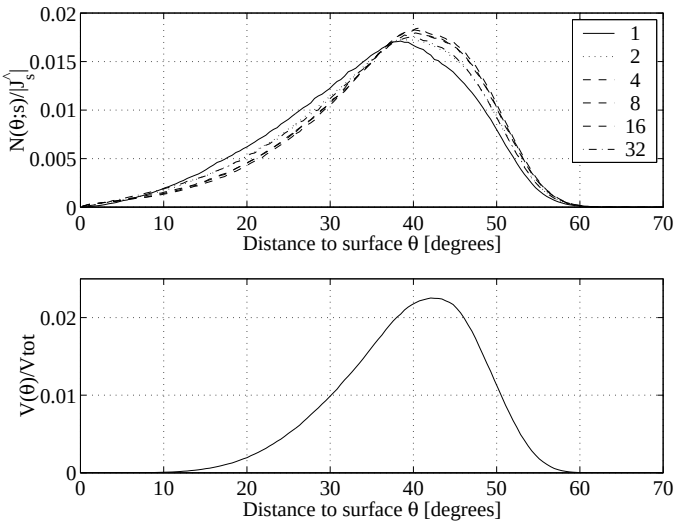


**Fig. 3.** (Top) The normalized histograms $N(\theta; s)/|\hat{J}_s|$ of the data sets $\hat{J}_s$ ($|\hat{J}_s| \approx 4.1 \cdot 10^6$ points), where $s = 1, 2, 4, 8, 16, 32$ (see legend) and $\theta$ is the distance to the edge manifold $\hat{\mathcal{E}}(\alpha, l/s) \subset \mathrm{S}^8$. (Bottom) Normalized histogram $V(\theta)/V_{\mathrm{tot}}$, which corresponds to the Monte Carlo estimated volume on $\mathrm{S}^8$ of the histogram bins of $N(\theta; s)$ ($V_{\mathrm{tot}} = 10^7$ points).

To get an estimate of the probability density of points around the edge manifold, we also need to calculate the volume of the bins $[\theta - \frac{\Delta\theta}{2}, \theta + \frac{\Delta\theta}{2})$ in the state space $\mathrm{S}^8$. We here estimate the bin volume by sampling $V_{\mathrm{tot}} = 10^7$ uniformly randomly distributed points $\boldsymbol{v}_n$ ($n = 1, \ldots, V_{\mathrm{tot}}$) on the 8-sphere. The histogram

$$V(\theta) = \# \left\{ n \middle| \theta - \frac{\Delta\theta}{2} \leq \mathrm{dist}(\boldsymbol{v}_n, \hat{\mathcal{E}}(\alpha, l/s)) < \theta + \frac{\Delta\theta}{2} \right\} \tag{22}$$

of the number of samples versus the distance $\theta$ to the surface of edges is a Monte Carlo estimate of the volume of the histogram bins. Fig. 3 (bottom) shows the normalized histogram $V(\theta)/V_{\mathrm{tot}}$.

We define the empirical *density* of data points around the edge manifold as

$$\rho(\theta; s) = \frac{N(\theta; s)/|\hat{J}_s|}{V(\theta)/V_{\mathrm{tot}}}. \tag{23}$$

Fig. 4 shows the calculated density for the data sets $\hat{J}_s$, where $s = 1, 2, 4, 8, 16, 32$ pixels. These results seem to indicate that the probability distribution of data points in the jet space has an *infinite* density at the surface of scale-space edges (where $\theta = 0$). This is consistent with the results on high-contrast $3 \times 3$ pixel image patches in [13]. Furthermore, the density function $\rho(\theta; s)$ is approximately

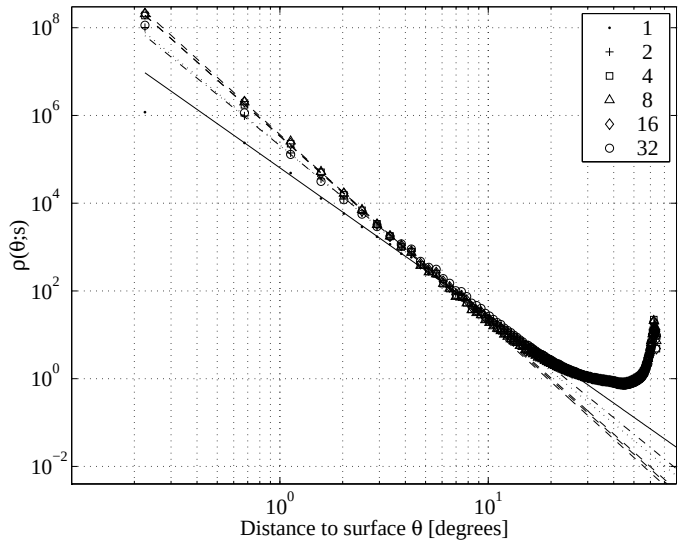


**Fig. 4.** Density $\rho(\theta; s)$ versus the distance $\theta$ to the edge manifold $\hat{\mathcal{E}}(\alpha, l/s) \subset \mathrm{S}^8$ for data points in $\hat{J}_s$. Each graph represents the density at a fixed scale $s$, where $s = 1, 2, 4, 8, 16, 32$ pixels (see legend). By linear regression for $\theta < 9$ degrees we get that $\rho_{\mathrm{fit}}(\theta; s) \sim \theta^{-\gamma_s}$, where $\gamma_s = 1.7, 1.0, 0.7, 0.7, 0.7, 1.1$ $(s = 1, 2, 4, 8, 16, 32)$.

*scale invariant* and seems to converge towards the functional form $\rho(\theta; s) \sim \theta^{-0.7}$ as we increase the scale $s$. The latter results are consistent with many of the previous empirical findings on scale invariance of natural image statistics; see e.g. [3,21].

In Fig. 5 (top), we calculate the cumulative sum $\sum_{\beta \leq \theta} N(\beta; s)/|\hat{J}_s|$ (in percent) of the number of data points as a function of the distance $\theta$ to the surface of edges. For all scales $s$, we get that 20% of all data points are within 29 degrees of the surface of edges, which corresponds to less than 12% of the total surface area of the 8-sphere $\mathrm{S}^8$ (see Fig. 5 (bottom)). In other words, points in these subsets of $\hat{J}_s$ are densely clustered around the low-dimensional manifold of edges.

To better illustrate the connection between the density function $\rho(\theta; s)$ and the image space, we end this section by computing $\rho(\theta; s)$ for pixels in the classical "Lena" image. Fig. 6 shows both scale-space images of Lena and the corresponding densities $\rho(\theta; s)$ for different scales $s$. In the density calculation, we first map the pixels in the scale-space of Lena into the jet space $\tilde{\jmath}^3$ in Eq. (7) and then whiten and contrast normalize according to Sec. 4.2. We subsequently compute the distance $\theta$ of these points to the edge manifold $\hat{\mathcal{E}}(\alpha, l/s) \subset \mathrm{S}^8$ and, finally, we look up the fitted density values $\rho_{\mathrm{fit}}(\theta; s) \sim \theta^{-\gamma_s}$ (see Fig. 4) corresponding to the computed $\theta$-values. The gray values in the second and fourth rows of Fig. 6

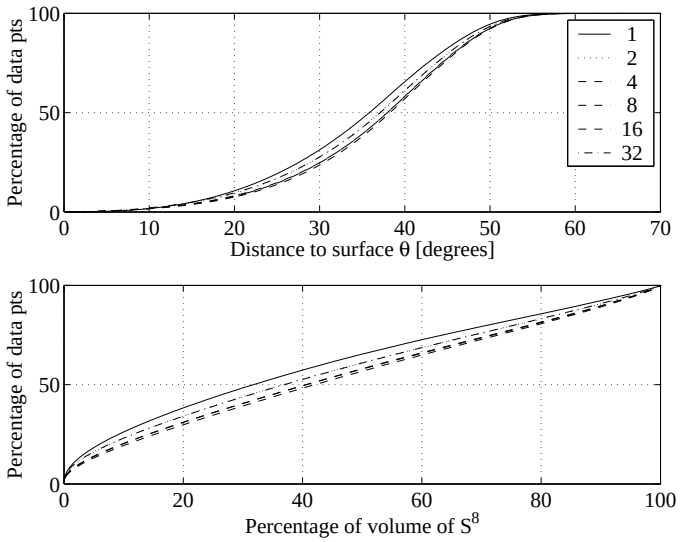


**Fig. 5.** (Top) Cumulative sum $\sum_{\beta \leq \theta} N(\beta; s)/|\hat{J}_s|$ (in percent) of the number of data points in $\hat{J}_s$ as a function of the distance to the surface of edges. (Bottom) Cumulative volume versus cumulative number of data points for data sets $\hat{J}_s$.

code for the magnitude of $\rho_{\text{fit}}(\theta; s) \sim \theta^{-\gamma_s}$ for different scales $s = 1, 2, 4, 8, 16, 32$. The first and third rows show the corresponding scale-space images.

## 5 Conclusions

We have extended the results in [13] from a pixel-based image representation to the jet space representation of linear Gaussian scale-space. The goal of this work is to investigate whether our previous findings on small image patches generalize well to larger scales and more general filter-based image representations.

In this work, we analyze Gaussian scale-space derivatives computed at randomly chosen points in (a large database of) natural images. At each chosen location, we compute the 3-jet representation (a 9-dimensional vector of up to 3rd order scale-space image derivatives) at different fixed scales. After whitening and a contrast-normalization, the data is on the surface of a unit 8-sphere in $\mathbb{R}^9$ centered at the origin.

Analysis shows that the probability distribution of empirical data has an infinite density at a 2-dimensional $C^\infty$-differentiable manifold in the 8-sphere (the state space of whitened and contrast-normalized 3-jet representations). This non-linear surface corresponds to the loci in jet space of Gaussian edges of different orientations $\alpha$, positions $l$ and scales $s$. Our results are approximately invariant to a change of scale. In fact, for increasing scales $s$, the density around the surface seems to converge towards the functional form $\rho(\theta) \sim \theta^{-0.7}$, where $\theta$ is the

**Fig. 6.** The first and third rows show scale-space images of "Lena" ($226 \times 226$ pixels) for scales $s = 1, 2, 4, 8, 16, 32$ pixels. The second and fourth rows show the corresponding log-densities $\log(\rho(\theta; s))$ in jet space (a bright pixel corresponds to a high density at that position in the image). The densities $\rho(\theta; s)$ are estimated at each pixel by first computing the distance $\theta$ to the manifold of Gaussian edges and then looking up the density values $\rho_{\text{fit}}(\theta; s)$ (see Fig. 4) that correspond to these $\theta$-values.

distance to the edge manifold. For all scales, we find that 20% of the randomly chosen image points have a 3-jet representation that are within 29 degrees of the edge manifold. This region around the manifold corresponds to less than 12% of the total surface volume of the 8-sphere.

The results above are consistent with our earlier findings in [13] for $3 \times 3$ natural image patches. In this work, we have studied the manifold of Gaussian edges parametrized by the orientation $\alpha$ and the scale-normalized position $l' = l/s$ of an edge. More generally, we believe that one can define a dictionary of probability models on representations of *general* primitives (edges, bars, blobs, T-junctions) parametrized by $\Phi = \{\phi_1, \phi_2, \ldots\}$ for *any* set of filters $f_1, \ldots, f_N$ in a sensorium. In the N-dimensional state space of the filter-based image representations, the

image primitives will define manifolds of the general form

$$M(\Phi) = [f_1(\cdot) * I(\cdot;\Phi), \ldots, f_N(\cdot) * I(\cdot;\Phi)]^T .$$

The picture that seems to emerge is that natural images are extremely sparse with most of the data in state space concentrated along these low-dimensional structures that correspond to edges, blobs, bars etc. One has to realize that these manifolds are in general highly *non-linear* — this makes our approach fundamentally different from, for example, ICA and sparse coding where one studies linear projections in state space. It should also be noted that the dimension of the state space of the image data is determined by the number of filters in the analysis (which is usually very large), while the dimension of the manifolds of image primitives is fixed and determined by the complexity of the primitives only. Because of the *low dimensionality* of the primitive manifolds (2 for edges, 3 for bars, etc), a "probabilistic primal sketch" of natural images may have important implications on the information-theoretic bounds one can put on compression of these images.

**Acknowledgments**

We would like to thank Associate Professor Mads Nielsen, Professor Peter Johansen and Professor David Mumford for their valuable comments.

## A Appendix: The Edge Manifold

**Theorem 1.** *The loci of all points $\mathcal{E}(\alpha, l, s)$ (Eq. (14)) trace out a 2-dimensional $C^\infty$-differentiable manifold in the jet space $\tilde{\jmath}^3 \subset \mathbb{R}^9$. This manifold of edge representations is parametrized by the angle $\alpha$ and the ratio $l/s$ between the displacement $l$ and the scale $s$.*

*Proof.* The edge map $\mathcal{E}$ (Eq. (14)) is *infinitely differentiable*, as

$$G_{x^m,y^n}(x, y; \alpha, l, s) = s^{m+n} \cos^m \alpha \ \sin^n \alpha \, \partial_{v^{m+n-1}} \psi(v; l, s)$$

and the functions $\cos\alpha$, $\sin\alpha$ and $\psi(v; l, s)$ are $C^\infty$-differentiable with respect to $\alpha$, $l$ and $s$ (for $s > 0$).

Furthermore, if we introduce the dimensionless variables

$$u' = \frac{u}{s} , \quad v' = \frac{v}{s} , \quad l' = \frac{l}{s} , \quad s' = 1$$

and assume scale-normalized derivatives as in Eq. (12), we get that

$$G_{v^n}(u, v; l, s) = G_{v'^n}(u', v'; l', s') \tag{24}$$

for $n \geq 1$. Eq. (24) follows from the scaling properties of the Gaussian function $\psi(v; l, s)$: We have that

$$\psi(v'; l', s') = s\psi(v; l, s)$$

and [1]

$$\partial_{v^k}\psi(v;l,s) = \left(\frac{-1}{s\sqrt{2}}\right)^k H_k\left(\frac{v-l}{s\sqrt{2}}\right)\psi(v;l,s)$$

for $k \geq 0$. Hence,

$$\begin{aligned}\partial_{v^k}\psi(v;l,s) &= \frac{1}{s^k}\left(\frac{-1}{s'\sqrt{2}}\right)^k H_k\left(\frac{v'-l'}{s'\sqrt{2}}\right)\frac{1}{s}\psi(v';l',s') \\ &= \frac{1}{s^{k+1}}\,\partial_{v'^k}\psi(v';l',s')\ ,\end{aligned}$$

which is equivalent to (Eq. (24))

$$\underbrace{s^{k+1}\,\partial_{v^k}\psi(v;l,s)}_{G_{v^n}(u,v;l,s)} = \underbrace{s'^{k+1}\,\partial_{v'^k}\psi(v';l',s')}_{G_{v'^n}(u',v';l',s')}$$

where $n = k+1 \geq 1$.

Although the edge map $\mathcal{E}$ in Eq. (14) is a function of three variables (the angle $\alpha$, the displacement $l$ and the scale $s$), the loci of all points $\mathcal{E}(\alpha,l,s)$ trace out a *2-dimensional* manifold (in the jet space) which only depends on $\alpha$ and the dimensionless ratio $l/s$.

## References


1. G. Arfken. *Mathematical Methods for Physicists.* Computational Imaging and Vision. Academic Press, London, 1985.
2. A. J. Bell and T. J. Sejnowski. The "independent components" of natural scenes are edge filters. *Vision Research*, 37:3327–3338, 1997.
3. D. J. Field. Relations between the statistics of natural images and the response proporties of cortical cells. *J. Optic. Soc. of Am.*, 4(12):2379–2394, 1987.
4. L. Florack. *Image Structure.* Computational Imaging and Vision. Kluwer Academic Publishers, Dordrecht, 1997.
5. L. Florack, B. ter Haar Romeny, M. Viergever, and J. Koenderink. The gaussian scale-space paradigm and the multiscale local jet. *IJCV*, 18:61–75, 1996.
6. J. Huang, A. Lee, and D. Mumford. Statistics of range images. In *Proc. of CVPR'00*, 2000.
7. J. Huang and D. Mumford. Statistics of natural images and models. In *Proc. of CVPR'99*, 2000.
8. A. Hyvärinen. Survey on independent component analysis. *Neural Computing Surveys*, 2:94–128, 1999.
9. J. J. Koenderink. The structure of images. *Biol. Cybern.*, 50:363–370, 1984.
10. J. J. Koenderink and A. J. van Doorn. Representation of local geometry in the visual system. *Biol. Cybern.*, 55:367–375, 1987.
11. J. J. Koenderink and A. J. van Doorn. Metamerism in complete sets of image operators. In K. Bowyer and N. Ahuja, editors, *Advances in Image Understanding*, pages 113–129. IEEE Comp. Soc. Press, 1996.
12. A. B. Lee, K. S. Pedersen, and D. Mumford. The complex statistics of high-contrast patches in natural images. Technical Report APPTS Report #01-3, Brown University, USA, 2001.

13. A. B. Lee, K. S. Pedersen, and D. Mumford. The complex statistics of high-contrast patches in natural images. In *WWW Proc. of 2nd Int. IEEE Workshop on Statistical and Computational Theories of Vision*, Vancouver, Canada, 2001. http://www.cis.ohio-state.edu/~szhu/SCTV2001.html.
14. T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer, 1994.
15. J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *IJCV*, 43(1):7–27, 2001.
16. S. Mallat. A theory for multiresolution signal decomposition: The wavelet representation. *IEEE Transaction on PAMI*, 11:674–693, 1989.
17. D. Marr. *Vision*. W. H. Freeman, New York, 1982.
18. D. Mumford and B. Gidas. Stochastic models for generic images. *Quarterly of Applied Mathematics*, 59(11):85–111, March 2001.
19. M. Nielsen and M. Lillholm. What do features tell about images? In Michael Kerckhove, editor, *Scale-Space and Morphology in Computer Vision*, LNCS 2106, pages 39–50, 2001.
20. B. A. Olshausen and D. J. Field. Natural image statistics and efficient coding. *Network: Computation in Neural Systems*, 7(2):333–339, May 1996.
21. D. Ruderman and W. Bialek. Statistics of natural images: Scaling in the woods. *Physical Review Letters*, 73(6):814–817, August 1994.
22. E. P. Simoncelli. Bayesian denoising of visual images in the wavelet domain. In P. Müller and B Vidakovic, editors, *Bayesian Inference in Wavelet Based Models*, pages 291–308. Springer-Verlag, New York, 1999.
23. E. P. Simoncelli. Modelling the joint statistics of images in the wavelet domain. In *Proc. of SPIE, 44th Annual Meating*, volume 3813, pages 188–195, Denver, 1999.
24. J. H. van Hateren and A. van der Schaaf. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc. R. Soc. Lond. Series B*, 265:359 – 366, 1998.
25. M. J. Wainwright and E. P. Simoncelli. Scale mixtures of gaussians and the statistics of natural images. *Adv. Neural Information Processing Systems*, 12, 2000.
26. 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.
27. S. C. Zhu and D. Mumford. Grade: Gibbs reaction and diffusion equations. — a framework for pattern synthesis, image denoising, and removing clutter. In *Proc. of ICCV'98*, 1998.