Griffin LD (2005) Feature classes for 1-D, 2nd order image structure arise from the maximum likelihood statistics of natural images. *Network: Computation in Neural Systems* 16(2/3):301-320

Feature classes for 1-D, 2nd order image structure arise from natural image maximum likelihood statistics

Lewis D Griffin

Computer Science
University College London, UK

l.griffin@cs.ucl.ac.uk

Much is understood of how quantitative aspects of image structure are measured by V1 simple cells, but less about how qualitative structure is determined from these measurements. We review Geometric Texton Theory (GTT) that aims to describe this step from quantitative to qualitative. GTT proposes that qualitative feature categories arise through consideration of the maximum likelihood (ML) explanations of image measurements. It posits that a pair of output vectors of an ensemble of co-localized neurons signal the same feature category if and only if the corresponding ML explanations are qualitatively similar. We present mathematical and empirical results relevant to GTT for the limited case of measurement by 1-D filters of up to 2nd order. The mathematical results identify the simplest explanations for measurements by such filters, while the empirical results identify the ML. We find that the ML explanations are not the most simple under any of the definitions of simple that we examined. However, the ML explanations do have properties predicted by GTT. In particular they change rapidly and qualitatively for certain narrow regions of measurement space, while remaining qualitative stable between those transition regions. Three feature categories arise naturally from the data: light bars, dark bars and edges. The results are consistent with GTT.

Keywords: Metamerism, hermite transform, textons, maximum likelihood, jet space

1. Introduction

This paper reports research on an approach to local image analysis in early vision – biological or machine - that we have previously named *Geometric Texton Theory* [1]. In this introduction we will review feature-based vision, models of image filtering and the transition from filter outputs to features. The introduction concludes with a preview of the body of the paper and a statement of notations used throughout.

1.1 Feature-based vision

Feature-based vision is founded on the hypothesis that one component of early visual processing is bottom-up computation of a representation of image structure in terms of *qualitative* descriptors (e.g. 'edge' or 'corner'). The motivation for the hypothesis is a belief that such feature-based description could be so successful at discarding the unimportant structure of an image and emphasizing the important that the process of vision would simply be, as David Marr wrote, that "you looked at the image, detected features on it, and used the features to classify and hence recognize what you were looking at." [2].

The Marr quote above is now more than twenty years old, and he traces the feature approach back a further thirty years [3]. Over the intervening decades there has certainly not been the constant and gradual improvement in models of and algorithms for the computation of features that might have been hoped for, rather the approach has been increasingly neglected. Of course, failure to be fruitful does not prove that the approach is misconceived, but it does lay a burden on those pursuing it to explain:

- how progress in machine vision has been possible without a theory of features, and
- why development of the feature approach has been difficult.

To the first question we would answer that a fundamental problem in vision arises from having to deal with the cross-product of two spaces of very high cardinality: (*I*) the space of all possible images, and (*C*) the space of possible image contents. Feature-based image description dramatically reduces the cardinality of *I*. In the absence of an effective feature-based approach, machine vision research has focused on producing working solutions in constrained domains where the cardinality of *C* instead can be restricted. To the second question we would say that a lack of ambition about how large a vocabulary of features can be stably computed has led to a misplaced pessimism about their potential efficacy. For example, to quote Marr again, "think of a 5 gradually changing into a 6 – a corner disappears, a gap narrows. Almost no single feature is necessary for any numeral." [2]

1.2 Image filters

Marr's pessimism seems to have been due to an incomplete picture of visual neuron properties leading him to consider only feature vocabularies of limited expressiveness. At the time he wrote *Vision*, the linear simple cells of V1 seemed to consist of 1^{st} and 2^{nd} order (in a derivative sense) filters only. With such filters all that it seemed possible to detect were 'edges' and 'bars' and it is clear that much work is indeed required to discriminate a '5' from a '6' with such a limited feature vocabulary. Marr's answer to this lack of expressivity was a process that stitched together simple local features (the raw primal sketch) into more complex multi-local features (the full primal sketch). However, it is now known that V1 simple cells frequently have more positive and negative sub-fields, and are more varied than Marr knew. So, (it is hypothesized [4-7]) an ensemble of co-localized visual neurons together have the power to characterize a substantial vocabulary of features without having to resort to the difficult process of multi-local feature combination.

A popular model of V1 simple cells is as Gabor functions which do indeed model the electrophysiological data very well. However gaussian derivatives (DtGs) of 1^{st} to 4^{th} or 5^{th} order provide an equally well-fitting set of models with much to recommend them [8-16]. We will refer to the measurements given by co-localised DtGs up to some order as a jet, and the

manifold of possible jets as the jet space. Although we appreciate the limited persuasiveness to some of such non-empirical considerations, to the author the most compelling points in favour of DtGs over Gabors are the interpretations of what they measure. In particular we note two. First, that measurement of the jet up to some order is equivalent to measurement of the initial terms of the Hermite Transform, which is a local analogue of the Fourier Transform [7, 17-19]. Second that the jet is also interpretable as the initial terms of the Taylor series of the image blurred to the same degree as the scale of the DtG filters [20]. Both interpretations are very appropriate for what we consider V1 to be i.e. a fully general-purpose system for local measurement.

1.3 From filters to features

Even though the 4^{th} or 5^{th} order jet, with its 15 or 21 dofs, seems to capture local image structure richly enough to be the basis of spatially complex features it is far from obvious how to define these features on the basis of the filter responses [21]. The majority of the relevant literature is to be found under the keyword 'textons' rather than features [22-25]. The most common position taken in this literature (implicitly or explicitly) is that textons/features correspond to clusters in the jet space. We agree with two parts of this position – that features correspond to regions of jet space and that natural image statistics somehow determine these regions, but we disagree that these regions are revealed as simple clusters. The simple fact is that if one forms a distribution over jet space for natural images there may well be a single strong cluster [26, 27] but no evidence for multiple clusters has ever been presented. Clusters have been found [28] in a non-linear re-parameterisation of the 2-D 2^{nd} order jet space, but in this case they are located nearer to category boundaries than centroids.

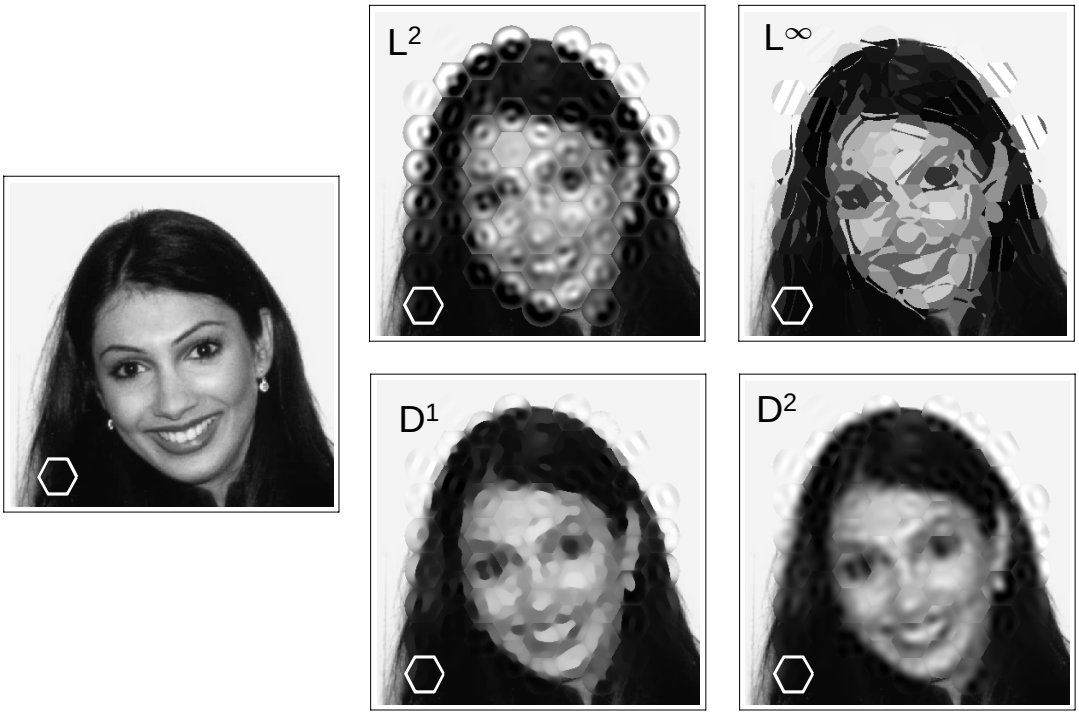
An alternative approach to the 'features from filters' problem has been pioneered by Koenderink [4-7] who has stressed the relevance of metamerism: that the jet does not uniquely determine the measured image even locally. In particular he has suggested that:

i. it is possible non-arbitrarily to associate with each point of jet space an iconic image from the metamery class of possible images, and
ii. the equivalence relation of 'qualitative identity of icons' gives rise to a partitioning of jet space into features.

This approach is illustrated in figure 1.

1. The figure illustrates Koenderink's proposal for feature classes. Even though any given local patch of the image will be measured by several V1 simple cells tuned to that position and scale, the cells' outputs fail to completely specify the image within the patch. Koenderink's proposal is that for each patch there is an iconic form that is consistent with the V1 measurements i.e. metameric with the original. Illustrative icons are shown for eight locations. These icons can then be grouped into classes based on identity of their qualitative structure. The equivalence classes of icons are the desired feature classes. In the example figure, the eight icons fall into three qualitative classes (based on the number and sign of their discontinuities) which are indicated by the colour of the contour surrounding the patch. The feature categories shown could be labelled: edge (grey), light bar (white) and dark bar (black).

We have examined several plausible rules for icon selection and have been unable to find any that perform convincingly especially at high order (see figure 2). Our failure to guess an icon selection rule motivated us to propose [1] *Geometric Texton Theory* (GTT) a refinement of Koenderink's proposal with the addition that icons should be defined as the ML (relative to natural images) elements of metamery classes. So our GTT extension ties features to natural image statistics in the way that the (incorrect) cluster idea of the texton approaches also does.



2. The figure allows comparison of several plausible rules for icon selection. The original image is at left. To produce each of the other panels, a hexagonal array of sites was set-up and the 4th order local structure was measured at each site at a scale indicated by the hexagons. Each hexagonal patch was then replaced by an icon with the same 4th order structure. The icon selection rule is indicated by the label at the top-left of each panel. These rules are further explained in section 2.

We have previously studied GTT for the 1-D, 1st order jet [1]. In this simplest of cases, the two dofs (0th and 1st order measurements) can be eliminated by affine scaling of profiles so that they all measure to the same values. Once this was done, we found that the ML form was a step edge. We claimed that this result, although very far from decisive, was supportive for GTT in that the ML profile had a simple qualitative structure. Simple structure is desirable because for the final completion of GTT it will be necessary to provide a definition of 'qualitative identity of icons'.

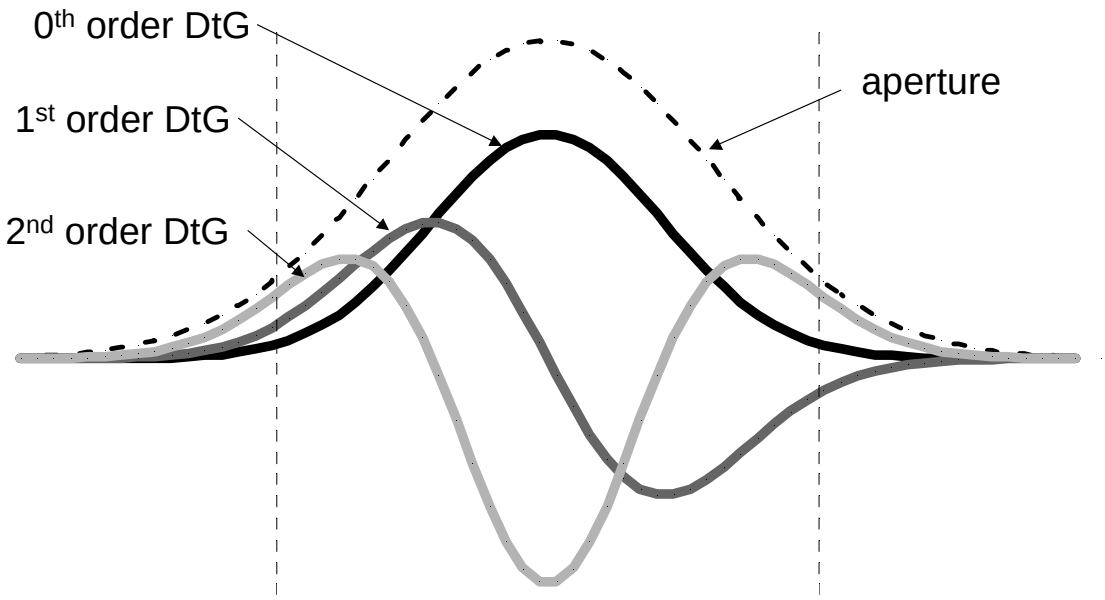
1.4 Paper notation and structure

In the remainder of the paper we continue our study of GTT but in this case for 1-D, 2nd order jets. There are three ways of interpreting such jets. First, as the result of measuring using the 0th, 1st and 2nd order DtG filters shown in figure 3. These filters are given by:

$$G_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}}, \quad G'_\sigma(x) = -\frac{x}{\sigma^2} G_\sigma(x) \qquad G''_\sigma(x) = \frac{1}{\sigma^2}\left(\frac{x^2}{\sigma^2} - 1\right) G_\sigma(x)$$

We denote the 2nd order jet of the image I measured at the origin, at a scale σ as $J^{(2)}_\sigma(I) = \langle m_0, m_1, m_2 \rangle = \langle G_\sigma, -G'_\sigma, G''_\sigma \rangle \cdot I$, where the dot product denotes $f \cdot g = \int fg$. The second interpretation of the jet arises because differentiation commutes with linear filtering,

so the jet is also equal to the initial terms of the Taylor series of the blurred image i.e. $J_\sigma^{(2)}(I) = \langle \delta, -\delta', \delta'' \rangle \cdot (G_\sigma \otimes I)$. Thirdly, the jet can be viewed as the initial terms of a Hermite Transform (HT) [19, 29]. To see this, we introduce $A_\sigma = G_{\sqrt{2}\sigma}$ the spatial weighting function of the HT and the Hermite functions $H_\sigma^{(0)} = G_\sigma / A_\sigma$, $H_\sigma^{(1)} = G_\sigma' / A_\sigma$, $H_\sigma^{(2)} = G_\sigma'' / A_\sigma$. Then we have $J_\sigma^{(2)}(I) = \langle H_\sigma^{(0)}, H_\sigma^{(1)}, H_\sigma^{(2)} \rangle \cdot (A_\sigma \times I)$. In this paper we mainly consider the DtG filtering picture of the jet although we will also make use of the aperture function from the HT view of things.



3. Shows the three DtG filters that we used to measure profiles for our 1-D, 2nd order study. The filters are of scale $\sigma = \sqrt{48} \approx 7$, and are plotted over the interval $x \in [-32, 32]$. The 'aperture' function is a 0th order Gaussian of scale $\sqrt{2}\sigma$, it will be relevant in section 3. The vertical dashed lines mark the outermost inflexions of the 2nd order filter. These dashes appear in many later figures of profiles and act as landmarks to allow comparison with the filters.

The structure of the paper is as follows. In section 2 we present analytical results about the simplest profiles to be found in 1-D, 2nd order metamery classes. By comparing the results for different definitions of simple we provide a context for section 3 in which we present the results on the ML profiles to be found in these metamery classes. In section 4 we show that the ML profiles give rise to a partitioning of the 1-D, 2nd order jet space into feature classes as hypothesized by GTT.

2. Norm-minimizing metameres

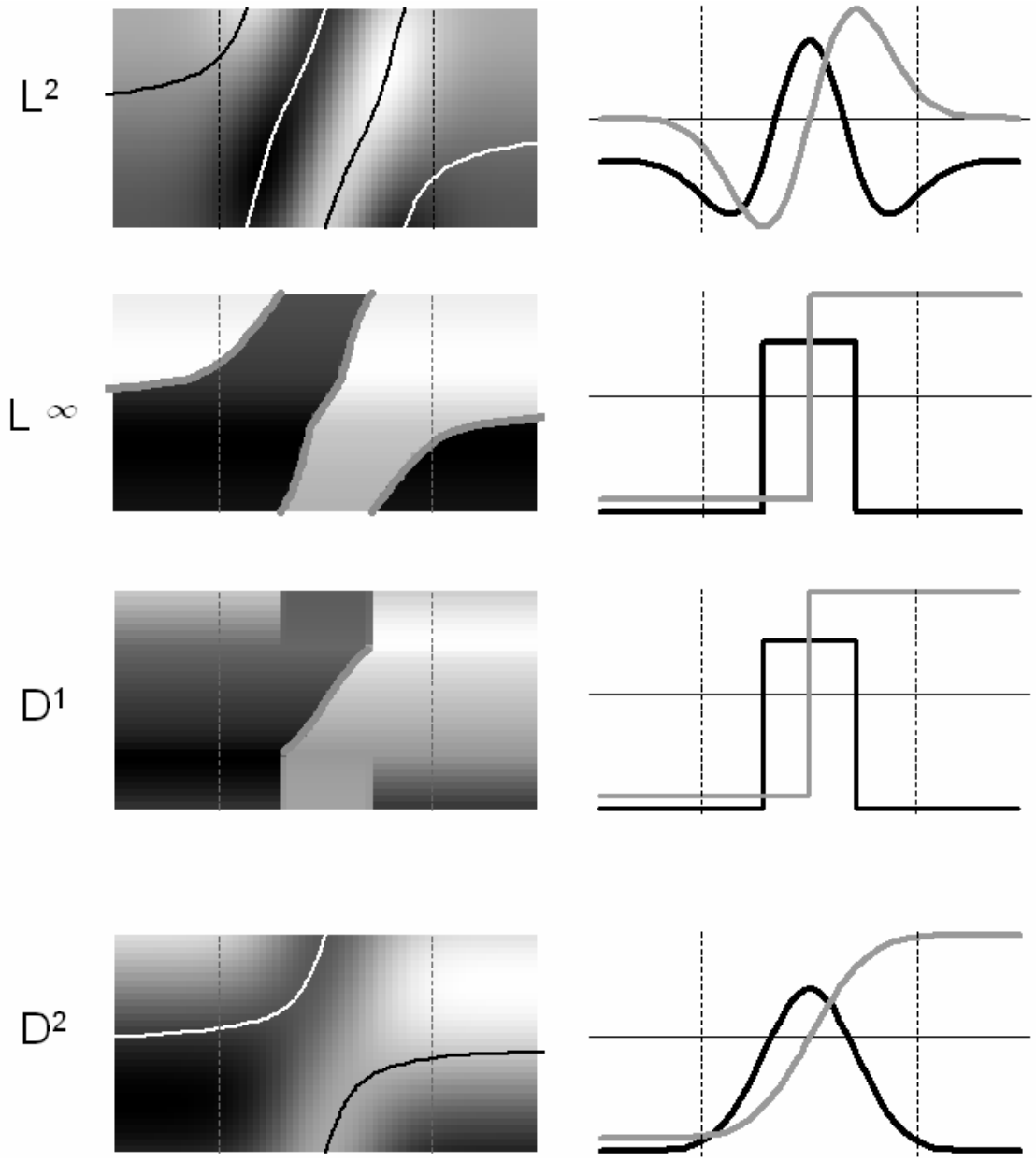
Specification of the jet at some point of an image does not fully determine the measured image, not even within the aperture corresponding to the jet measurement location and scale. The set of profiles that measure to a given jet constitute the metamery class for that jet. Many of the elements of a metamery class have wild oscillations at frequencies which are

sufficiently high that the DtG filters are 'blind' to them or attain huge values distant from the aperture [30], however there are also metameres which are more conservative in their excursions. In particular, one easily observes that natural image patches when considered in comparison to their metameres are of the conservative variety. Previously, this line of thought led us to pursue [31, 32] definitions of 'conservative/simple' in the hope that the within-metamery class minimizer of such a measure could play well the role of the icon that Koenderink has suggested should 'stand for' the metamery class or, equivalently, point in jet space (section 1.3). Although we now consider this 'icons are simplest' approach to be less well-motivated than the 'icons are the most likely' approach (see again section 1.3), that is the main subject of the paper, its study remains worthwhile. For it allows one to appreciate in what way ML icons are, or are not, simple.

The definitions of 'simplest' that we consider here are norm-minimizers. For an image *I*, we define its luminance norms as $L^r(I) = \min_{\mu_r \in \mathbb{R}} \left(\int_{\vec{x} \in \mathbb{R}^n} \left| I(\vec{x}) - \mu_r \right|^r \right)^{\frac{1}{r}}$, and its gradient norms as $D^r(I) = \left(\int_{\vec{x} \in \mathbb{R}^n} \left| \nabla I(\vec{x}) \right|^r \right)^{\frac{1}{r}}$. In both cases, infinity norms are defined by taking the limit $r \to \infty$ and *n* denotes the dimensionality of the image domain.

For both types of norm, addition of a constant value to the image leaves the norm unchanged, whereas multiplication of the image by a constant similarly multiplies the norm i.e. $L^r(\alpha + \beta.I) = \beta.L^r(I)$ and $D^r(\alpha + \beta.I) = \beta.D^r(I)$. Therefore to study the norm-minimizers with respect to the 2nd order jet we need only study a 1-D factor subspace of jet space modulo the two irrelevant affine scaling dofs. We choose $\langle 0, \cos\theta, \sigma^{-1}\sin\theta \rangle$, indexed by the phase variable θ.

The norms we have previously identified as of interest are L^2, L^∞, D^1 and D^2. The minimizers of these norms show a degree of resemblance to structures found in natural images (or to put it the other way, image patches tend to be low in these norms when compared to their metameres). This is not true for all norms. For example, the L^1 norm minimizers consist of very 'unnatural' small collections of weighted delta functions. For our norms of interest, we have determined analytically the minimizers for each value of phase. Figure 4 shows the results and the following subsections provide detail.



4. Shows the norm-minimizing profiles for different norms and phases. The norm for the row is indicated at left. The middle column shows the minimizers as density plots. Each row of the density plot corresponds to a minimizing profile; phases varies vertically from $-\pi/2$ (2^{nd} order filter response is –ve, 1st order response is zero) at the bottom, through 0 (2^{nd} order zero, 1^{st} order +ve) in the middle, to $\pi/2$ (2^{nd} order +ve, 1^{st} order zero) at top. To assist visualization, lines have been added to the density plots to show the location of extrema and discontinuities. The right hand column shows the cosine-phase (black) and the sine-phase minimizers as regular plots. For the middle and right columns the vertical dotted lines are the same as in figure 3.

2.1 L^2: The variance-minimizers

As we have previously noted [1], the form of variance-minimizing metameres can be determined by the method of Lagrangian multipliers to be a weighted sum of a constant function and the DtG filters (other than the 0^{th} order) that measure the jet. As can be seen

from figure 4 (top, right) the pure cosine-phase variance-minimizing metamere is just the 2nd order DtG, the pure sine-phase the 1st order DtG, and intermediate phases are a mixture of the two. The distinctive feature of variance-minimizers is their return to a constant value outside the aperture. This is the reason why in figure 2 the variance minimizing patches are joined by seams that are more visible than for the other norms.

2.2 L^{∞}: The range-minimizers

As we have proved elsewhere [31] the range-minimizers are binary-valued functions with one or two discontinuities. For all phases apart from sine-phase, there are two discontinuities positioned at $x = \alpha, \beta$ such that $\alpha\beta = -\sigma^2$. For sine-phase there is only a single minimizer at the origin. For cosine-phase the discontinuities are symmetrically placed at $\pm\sigma$. These are shown in figure 4 (second row).

The general strategy for proving the form of these minimizers comes originally from an argument as to the form of optimal spectral reflectance functions in colour science [33]. The strategy, which is to show by *reductio ad absurdum* that the form of the solution is restricted, is as follows. Suppose the jet specifying the metamery class is $\vec{m}$ and hypothesis that the minimizer *s* has a certain type of form. If one can show that *s* can always be perturbed to make s^* which has (i) the same norm value as *s*, and (ii) has jet $(1+\varepsilon)\vec{m}$, where $\varepsilon > 0$, then it follows that $(1+\varepsilon)^{-1} s^*$ will have jet $\vec{m}$ but will have a lower norm that *s*. This contradicts the hypothesis that *s* was a minimizer, therefore the minimizer cannot have the hypothesized form. This strategy can be used to show that (i) the range-minimizer must be binary-valued, and then that (ii) the number of discontinuities is restricted. That the appropriate types of perturbation are always possible for the proof to work follows from the full rank of the Gaussian derivatives [6].

When one considers range minimizers for different orders of jet one discovers that what remains constant is that they are binary-valued, while what varies is the number and complexity of the transitions between the two values. For low orders the binary valuedness of the range minimizers does not seem to disqualify them from being effective icons (the figure 1 icons are range minimizers) but at higher orders it makes them noticeably unnatural (figure 2).

2.3 D^1: The total variation (TV) minimizers

The proof strategy outlined in the previous sub-section applies also to TV minimization. First the strategy is used to show that the TV-minimizers must be piece-wise constant. Next it is used to show that the TV-minimizers have at most two discontinuities. Then, for each possible value of the jet, one identifies the piecewise constant profiles with at most two discontinuities that measure to the correct values, and one computes which of these possibilities has the lowest TV. The resulting profiles are shown in figure 4 (third row). For phases in the interval $\theta \in \left[-\pi/4, \pi/4\right]$ (i.e. closer to sine- than cosine-phase) the TV-minimizer has a single discontinuity somewhere in the interval $x \in [-\sigma, \sigma]$. For phases

nearer cosine- than sine-phase ($|\theta| > \pi/4$), the TV-minimizing form has two discontinuities at $x = \pm\sigma$. In contrast to the range minimizers, the jump of the two discontinuities is not generally equal. As figure 4 shows, for pure sine- and cosine-phase the TV-minimizing profiles are identical to the range-minimizers. This is not true for all other phases.

We are unable to extent the proof strategy so as to determine analytically the form of the TV-minimizers for 2-D jets. However for at least some cases one can argue that the 2-D TV-minimizers are not piecewise continuous as they always are in 1-D. For example, for the 2-D 1st order jet the TV-minimizer is not a 2-D equivalent of the step edge that it is for the 1-D 1st order jet. Instead, the TV minimizer is a straight edge which is sharp and of high amplitude at the point where the jet observation is located, but of reducing amplitude and sharpness as one moves away from this point. Compare this with the range minimizer for the 1-D, 1st order jet which is an infinite straight step edge.

These differences between the 1-D and 2-D TV minimizers does not seem quite right and one suspects that this commonly used definition of 2-D TV is faulty in some way. One insight into the fault is that a nice property of TV in 1-D – to wit that the D^1 definition is exactly equivalent to a measure that sums the absolute height difference between all pairs of adjacent extrema – does not carry over to 2-D. This issue has been previously noted and addressed by others [34] but further work is needed to define a workable alternate 2-D TV definition. This would be very worthwhile given the widespread and effective use of TV minimization as a general prior for natural images [35-37]. Indeed, even in its faulty form, the icons that it selects are the best of any norm we have explored (see figure 2).

2.4 D^2: The roughness-minimizers

In the 1-D case here studied, the problem of roughness-minimization with respect to measurements by G' and G'' can be shown to be related by integration to variance-minimization with respect to measurement by G and G'. So one can show that the roughness minimizers must be a weighted sum of a constant function, a 0th order DtG and an error function of the same scale as the DtG filters defining the jet. These minimizers are shown in the bottom row of figure 4. Analogously to the variance-minimizers, the pure cosine-phase minimizers are gaussians, the pure sine-phase minimizer is an error function and the minimizers for intermediate phases are a mixture of the two.

3. Maximum likelihood (ML) metameres

In this section we present results of an investigation into the form of the ML (relative to natural images) metameres for the 1-D, 2nd order jet. That is to say we have determined the answer to questions of the form: if a randomly selected profile from a natural image has a 2nd order jet that measures to $\vec{m}$ what is the most likely form of the profile? The method we use is similar but not identical to our earlier study of the 1-D, 1st order jet [1].

3.1 Methods

We subdivide presentation of our method for computing ML profiles into: extraction of profiles, intensity scaling of profiles, computation of the ML, and control computations carried out.

3.1.1 Extraction of profiles

As a source of natural image data we used 1220 images from the 4000 image van Hateren natural image database (.iml) of linear, 1536×1024, images of woods, open landscapes and urban areas [38]. In a change from our previous method, for each image we created a 5-layer quad tree by blurring and sub-sampling. The bottom level of each tree (the original image) had a psf width of approximately 0.8 pixels (determined by fitting blurred step edges as previously described [1]). The additional blurring that we imposed on the other levels was such as to keep the ratio between the width of the psf and the pixel sampling constant. At all levels of the quad tree we took measures to reduce quantization effects that we have previously described [1].

We extracted 2700 1-D profiles from each image, for a total of 3.3×10^6. Each profile was extracted from a random level of the quad tree (with probability proportional to the number of pixels in the level) at a random location and orientation (all real-valued) using nearest-neighbour interpolation. Each profile was 64 samples in length.

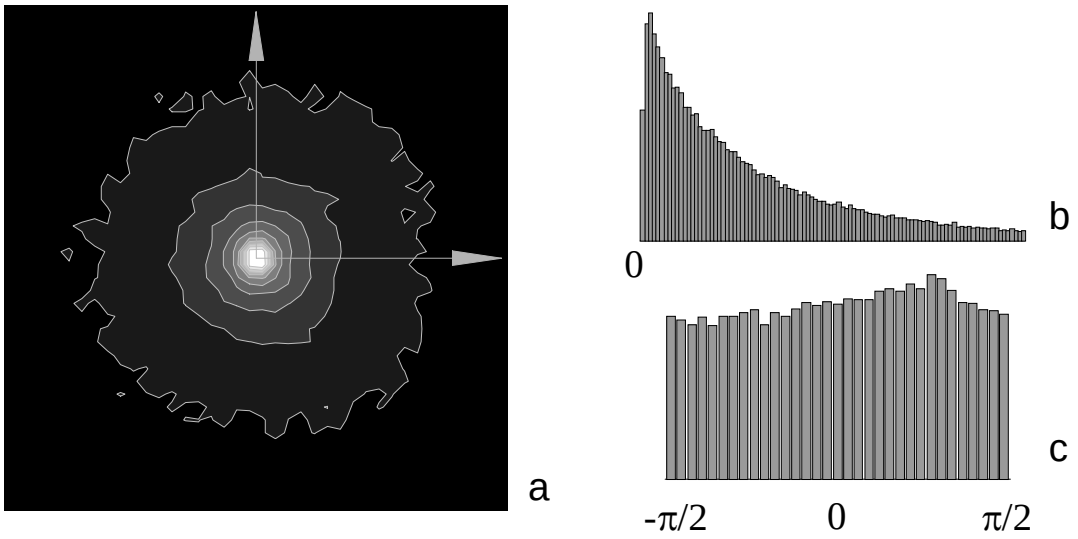
For comparison with natural image profiles we also constructed two sets of 3.3×10^6 synthetic profiles. One set – the Gaussian set - had each of its 64 values drawn independently from a normal distribution. The other set – Brownian profiles [39] – were generated by setting the value at one end of the profile with a normally distributed random variable ($N\left(0,100\right)$) then in turn setting each of the remaining 63 samples to be equal to the previous sample plus an independently generated normally distributed offset ($N\left(0,1\right)$).

Every database of profiles – natural and noise – was recreated three times using different random seeds, so that the uncertainty of subsequent processing could be determined.

3.1.2 Scaling of profiles

As in our previous study, we have chosen to factor out an affine component of image structure that we believe obscures the structure that we are really interested in. We do this by transforming ($P \to \alpha + \beta.P$) each profile that we examine so that its jet has the canonical form $\left\langle 0,\ \cos\theta,\ \sigma^{-1}\sin\theta \right\rangle$. The exact value of the factor σ^{-1} is not critical to our results but figure 5 shows that it is a natural choice, in that it causes the histogram of natural image profiles as a function of phase to be nearly flat. Note that the figure also shows that there is no hint of density clusters that could form the basis of feature categories.

After extraction, the jet of each profile (*P*) was measured ($\langle m_0, m_1, m_2 \rangle$) by forming the dot product with DtG filters of scale $\sigma = \sqrt{48} \approx 7$ and order 0, 1 and 2 (see figure 3). Each profile then underwent a two-step normalization process. First, any profile that had $m_1 < 0$ had its 64 samples reversed in order. Second, each profile was individually affinely scaled according to ($P \to \frac{P - m_0}{\sqrt{m_1^2 + \sigma^2 m_2^2}}$) so that it had a canonical jet. This normalization step was applied to both natural and noise profiles.



5. The density plot at left shows the distribution across jet space for natural image profiles. The horizontal axis is the 1st order filter response. The vertical axis is σ times the 2nd order filter response. The 0th order response is not plotted. The density has been square root transformed to improve visibility. On the right are plotted the polar marginal histograms; at top for the radial variable and at bottom for the phase. Note the flatness of the phase histogram.

3.1.3 Computation of ML profiles

In seeking the ML profile we implicitly assume the existence of a density over the space of possible profiles that reflects how often different profiles appear in natural images. We wish to find the profile where this density is maximum. However we do not have the density itself but only samples from it. Thus our problem is one of mode estimation. Our task has two differences from the majority of mode estimation problems dealt with in the literature. First is that the space from which are samples come is multi-dimensional. Second is that we use a non-standard metric on the space. We deal in turn with these issues and how they are tackled by our new algorithm.

3.1.3.1 Changes compared to previous method

Profile space is 64-D as we represent it. This multi-dimensionality has implications for how we estimate the mode. First off, consider that the density of a region of the space is naturally defined as the rate at which profiles within the region occur divided by the volume of the

region. Using this definition we can then define point density as the limit of region density for a nested sequence of convex regions containing the point. If we were able to take the limit all the way to zero-sized regions then the shape of the regions would be irrelevant. But in practice we are not able to take the limit to zero as we only have access to a finite number of samples rather than the density itself and small regions only contain a few samples and so the occurrence frequency within them cannot be reliably estimated. So our task is not to find the location with the highest point density, but rather to find the region with the largest reliably estimated region density. The centre of that region will be our mode estimate.

The second complication arises with regard to defining the volume of regions. This has to be done relative to some metric. We have previously argued that the obvious choice of the standard Euclidean metric $d(p,q)^2 = \int_x \left(p(x) - q(x) \right)^2$ is not correct for this problem as it gives an unwelcome dependency on the spatial cut-off of our representation of the profiles i.e. we might get a different ML profile if we used a representation that extended out twice as far from the jet measurement location. The solution is to use an apodized Euclidean metric, which weights differences between profiles depending on where the difference is. The apodisation function (aperture) that we use is the same that arises when the jet is interpreted as Hermite Transform coefficients, thus $d^*(p,q)^2 = \int_x A_\sigma \left(p(x) - q(x) \right)^2$.

The method of mode estimation that we have used will be presented and evaluated in detail elsewhere. But in summary it is a multiscale version of the popular mean shift algorithm [40]. The mean shift algorithm as it is generally used estimates the mode by iteratively shifting an aperture of fixed size and shape around until maximal region density is achieved. By using a fuzzy gaussian aperture, the shift rule is particularly simple and robust. The shift rule can be understood as implementing a gradient ascent process (with inbuilt step size selection) over a kernel-estimated density landscape. Our algorithm, which we have derived by generalizing the gradient ascent property of mean shift, modifies this procedure in three ways. First is that although we use gaussian apertures as in the original mean shift we do not require them to be isotropic. Second is that our shift rule alters not only the position of the aperture but also its size and shape. Third, we downgrade the count of the number of samples within an aperture in such a way that small, and thus unreliable, numbers are dramatically reduced. This modification ensures that we do not track down to region sizes where the density cannot be reliably estimated. The method for this is different from the method [41] we used in our previous study but with the same aim.

In comparison to our previous algorithm, the new one produces more accurate results and is faster. In section 3.2.1 we present results on noise profiles that confirm the correct operation of the algorithm.

3.1.3.2 Binning of phase

In our previous study, affine scaling brought all profiles into a single canonical metamery class. In the present study, scaling brings profiles into a 1-D family of canonical metamery classes indexed by phase. There is no way that we can determine the ML profile for any particular, exact value of phase, as chances are that we will have extracted only at most one profile with that value of phase. Instead we narrowly bin phase into 33 equally spaced bands and determine the ML profile for each of the 33 subsets of our collected profiles. Thus each

mode estimation is based on roughly 10^5 profiles. Even though our phase bins are narrow, each of the 10^5 profiles has a slightly different phase (thus slightly different jet) which could cause some undesired smearing of the density. To lessen the impact of this, each profile is projected to the nearest point in the null space of the DtG filters by subtracting a weighted sum the filters themselves. These 'black-projected' profiles are the input to our mode estimation algorithm. Having found the ML profile for a particular phase bin, a weighted sum of the DtG filters is added to it to bring it out of the null space and into the canonical metamery class indexed by the phase value which is the centre of the phase bin.

3.1.4 Control computations

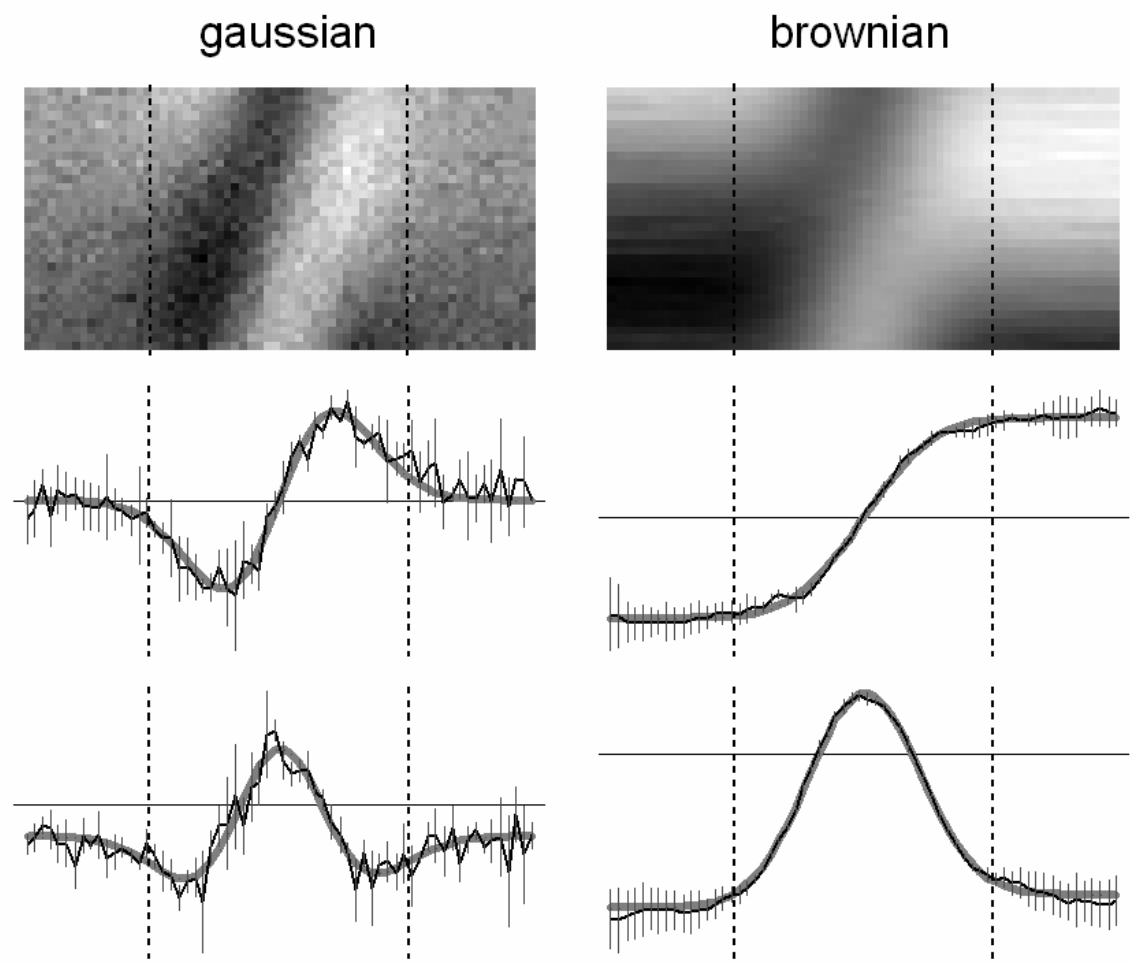
In our previous study we performed control computations that we used to assess the stability of our results with respect to various aspects of the data and algorithm. The aspects tested were filter scale, sampling relative to filter scale, interpolation scheme, log-transformation of image values, and dataset of images used. We found that none of these aspects had an effect on the results. In this study we restricted our control computations to checking on the effect of using a different dataset of images and checking scale invariance. As before, we used the BT dataset (ftp://ftp.vislist.com/IMAGERY/BT_scenes/) as an alternate to the van Hateren. There are only 98 images in the BT set, so more profiles were extracted from each image to compensate. We reassessed scale invariance because, unlike our previous study, we extract profiles at multiple levels of scale. The check we perform was to compare the use of all five levels of the quad tree to (i) just using the bottom level (equivalent to our previous single scale strategy), and (ii) using all but the bottom level.

3.2 Results

In the following sections we present results on ML noise profiles, ML natural image profiles, and control computations.

3.2.1 Noise profiles

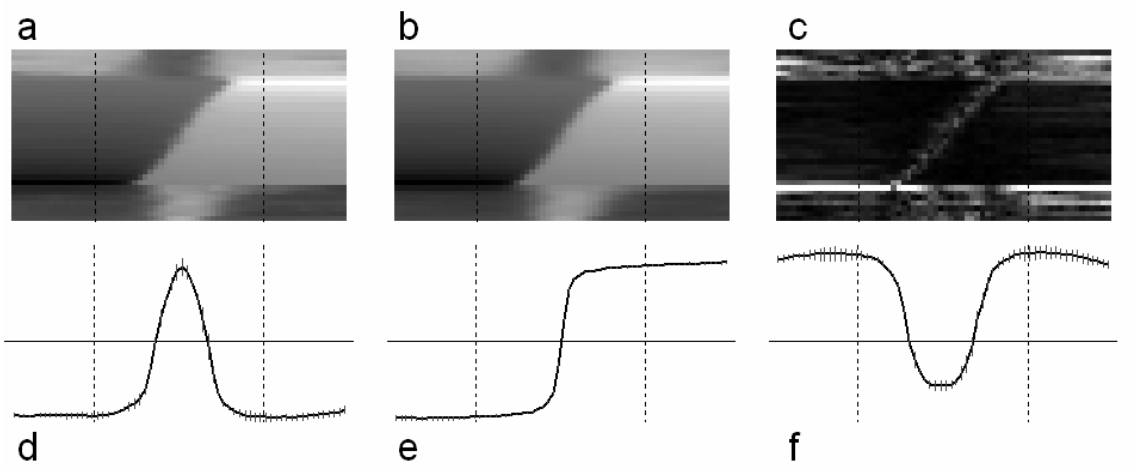
Figure 6 shows the ML profiles for gaussian and brownian noise. The figure shows that the profiles are very well-fitted by the variance-minimizing profiles (for gaussian noise) and the roughness-minimizing profiles (for brownian noise). This agrees with the analytical prediction [1] and so validates the performance of our mode estimation algorithm for this quantity and dimensionality of data. Note how the scatter across the three repeat computations is larger for the gaussian than for the brownian noise. In fact the scatter for the gaussian noise is the largest out of all the profiles we have examined (shown and not shown). This makes sense as the gaussian has very low kurtosis, thus finding modes of it should be particularly difficult.



6. Shows the ML profiles for the noise data. The left column is for gaussian noise, the right for brownian noise. The upper panels are laid out the same as in figure 4. The bottom two rows show the ML profiles for particular phases: the middle row is sine-phase corresponding to the central row of the top panel, the bottom row is cosine-phase corresponding to the bottom row of the top panel. The ML profiles are in black with error bars showing one sd of the scatter across the three repeat mode estimations. The thick grey curves are the L^2 and D^2 minimizers of the appropriate phase.

3.2.2 Natural image profiles

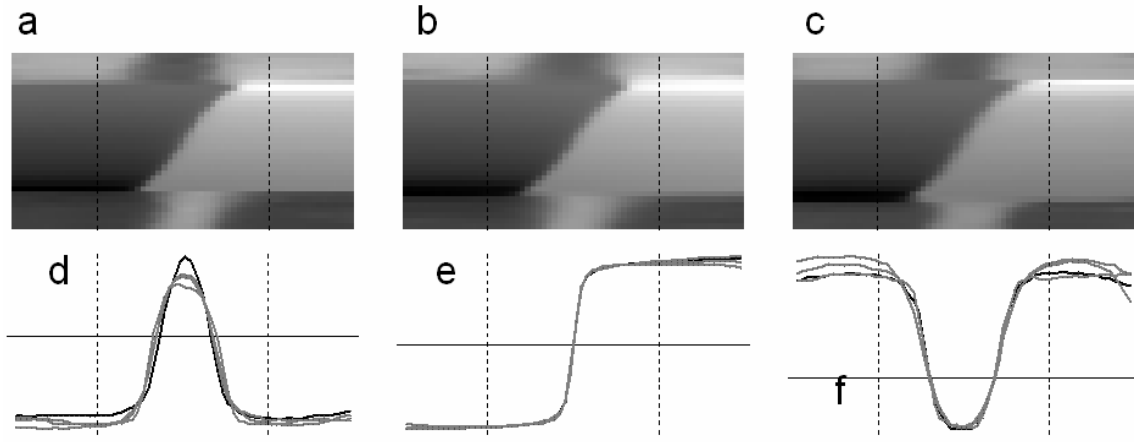
Figure 7 shows the ML profiles for the natural images (van Hateren, all 5 levels of the quad tree used). The figure shows that the scatter across the three repeat computations is small for phases near sine-phase and larger but still modest for phases near cosine-phase. Comparison with the norm-minimizing profiles in figure 4 shows that the ML natural profiles are clearly different from the variance- and roughness-minimizers but there are similarities with the TV- and range-minimizers. Near sine-phase the best agreement is with the TV-minimizers while near cosine-phase it is with the range-minimizers. This is reminiscent of the suggestion that different norms are appropriate as image priors for edge-like and blob-like regions [42].



7. Shows the ML profiles for natural images. Panel a is the result of a single mode estimation at each phase. Panel b is the mean of three mode estimations using different profiles. Panel c shows the sd of the scatter across the three repeats. Panels d-f show the sine- and cosine-phase profiles corresponding to the top, middle and bottom rows of the density plots. The error bars show 1 sd of scatter across the three repeat computations.

3.2.3 Control computations

Figure 8 shows the results of our three control computations – only fine scales, only coarse scales and alternate image dataset. It is clear that the differences with each other and the main results of figure 7 are minor (we do not claim statistical indistinguishability). The only clear difference is for the alternate image dataset. This will be discussed further in section 4.



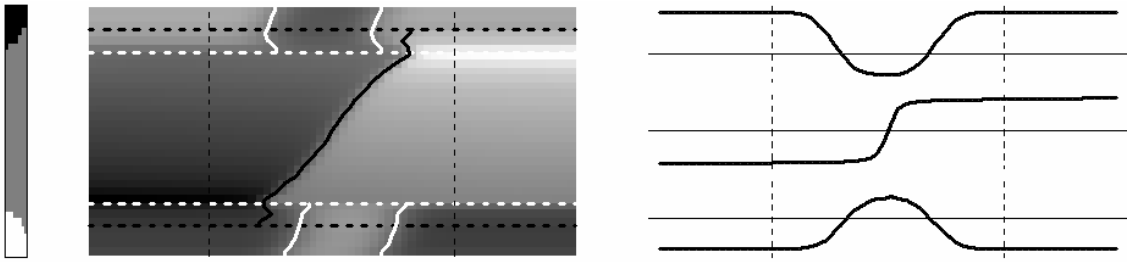
8. Shows the results of our three control computations. Panels a-c are density plots of the average of three ML computations and should be compared to 7b. Panel a shows the ML profiles when the computation has used only profiles extracted from the finest level of the quad tree for each image. Panel b is for profiles taken from the coarser levels (2-5) of the quad tree. Panel c is for profiles taken from the alternate image dataset. Panels d-f show plots of the ML profiles at cosine- and sine-phases for the main computation (black) and the three control computations (grey).

4. Induced feature classes

We have found a simple model that describes well the ML natural images profiles and their variation with phase (i.e. the data of figure 7b). The model consists of three template profiles (figure 9 right) – roughly an edge, a light bar and a dark bar – that by spatial shifting, intensity scaling and adding can be made to fit the ML profile at each phase. For phases near cosine-phase one or other of the unmixed bar templates fits the data. For phases around sine-phase the unmixed edge template fits the data. Only for six of our 33 phase bins is it necessary to use a bar-edge mixture. These six are arranged in two groups of three and separate the pure edges from the pure bars. The relative weighting of the bar and edges in these transition zones follow a monotonic pattern from pure edge to pure bar (figure 9 left). The centre of figure 9 shows the model profiles overlaid with lines that indicate how the bars and edges are shifted for different phases.

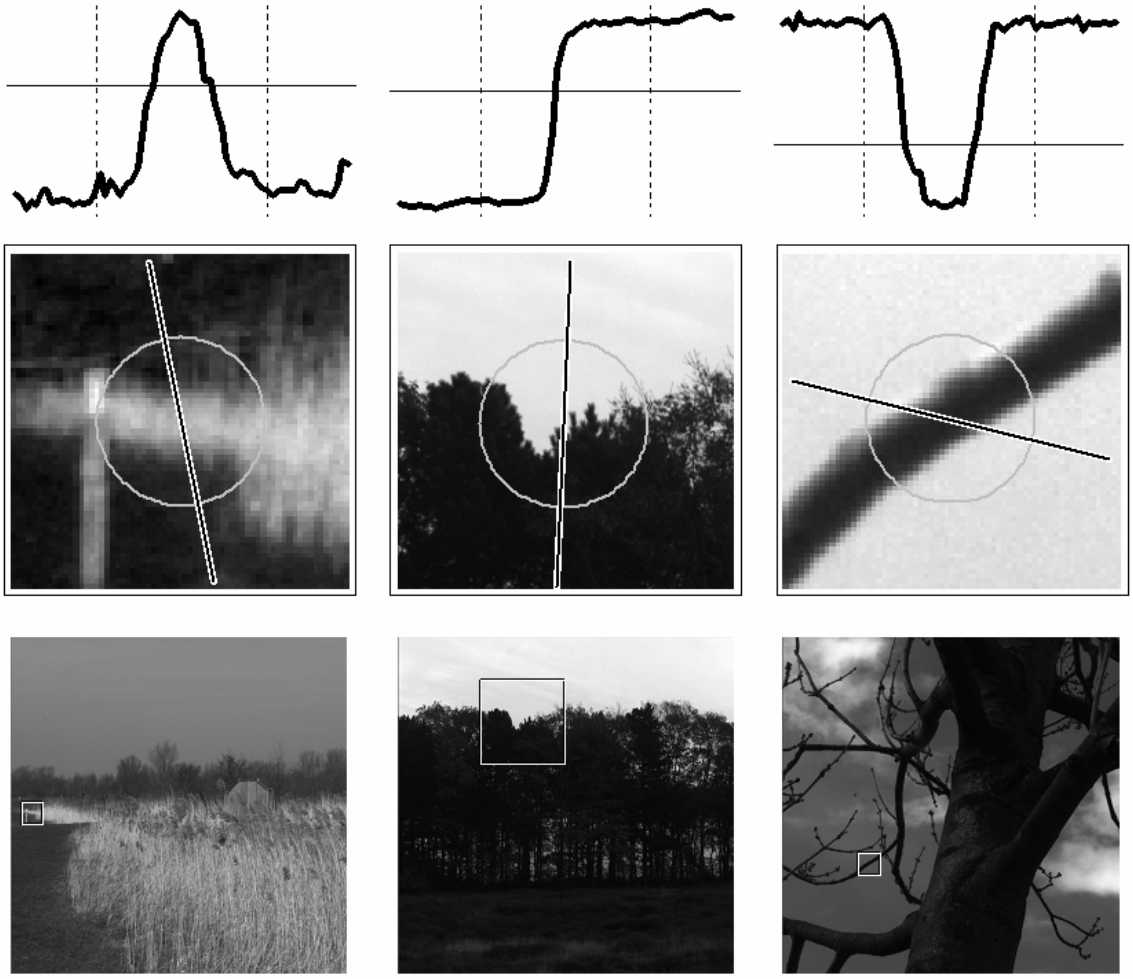
The light bar template was created by blurring a bar function with sharp discontinuities. As free parameters it had bar half-width and blur sigma. Similarly for the dark bar. The parameterisation of the edge template requires six parameters. The main structure of the edge template is determined by a single parameter that described the blur applied to a zero-centred step edge. A second parameter controlled the scale of a zero-centred 3rd order DtG, and a third parameter controls the weight of this addition to the blurred step edge. The effect of the 3rd order DtG is to slightly alter the shape of the edge roll-off. The remaining three parameters for the edge template control a linear, cubic and quintic terms whose effect is to slightly adjust the template away from the edge.

The model also has parameters that control how the templates are spatially shifted and intensity scaled for each of the 33 phase bins. In total our model had 119 parameters whereas the data has $33 \times 64 = 2112$ dofs. We optimised the fit of the model using a gradient descent technique. We failed to achieve a statistically satisfactory fit: our chi-squared score was 5.0 sds out from the null hypothesis value (which actually is not too bad for such high-dimensional model fitting [43]). However our desire here is more for a descriptive than a precise model. To score the descriptive success of the model we compute normalized differences between model and data profiles, where the normalized difference between two profiles is defined to be the apodized distance between the two divided by the average of the apodized lengths of the two profiles. Using this definition, we calculate that the RMS difference between the model and the mean of the three ML computations is 10%. To put this value in context, the RMS pair-wise difference between the three repeat ML computations was 15%.



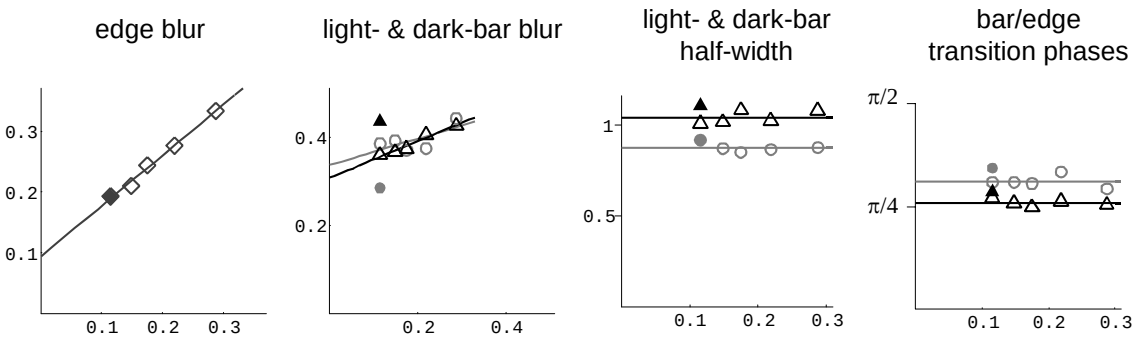
9. Illustrates the model that well describes the ML natural image profile data. The background density plot of the central panel shows the model profiles and can be compared to fig. 7b. The overlaid white lines show the twin edges of the light- and dark-bar templates, the solid black line shows the position of the centre of the edge template. The dotted lines show the phase extent of the different parts of the model. Where these extents overlap, the model profiles are a mixture of the two types of template. The vertical bar at the left of the figure shows the amount of each template used at each phase. The grey fill is for edges, the black fill for dark bars, and the white fill for light bars. The right of the figure shows the three templates. Observe that although the edge template has six control parameters, in form it is only subtly different from a blurred step edge. Also note the slight difference between the light- and dark-bar templates, this is mostly due to a difference in the bar half-widths rather than the degree of blur.

In figure 10 we present actual natural image profiles that have our three template forms and show how they arose in the images from which they were extracted. The example having the light bar form arose from a foreshortened view of a region of the ground plane containing lighter vegetation than the surround. The dark bar arose from the simpler situation of a branch silhouetted against the sky. The edge example is another silhouette, but this time of a tree canopy against the sky.



10. Shows examples of profiles that are close to the template forms in the model. The left column is for the light bar template, the middle for the edge, and the right for the dark blur. The top row shows the profiles, the middle row shows the image patch that the profile can from, and the bottom row shows where the patch came from in its parent image. Note that the image intensities in the middle and bottom rows are scaled differently.

In our previous 1-D, 1st order study we showed that the slight blur of the step edge that we found to be the ML form was attributable to the non-zero width of the point spread function (psf) of the van Hateren images. This was shown by blurring the database of images to increase the width of their psf and showing that this increased the blur of the ML profile. We have conducted a similar study for the 1-D, 2nd order jet. The variables we have investigated are: the light and dark bar half-widths, the light and dark bar blur sigmas, the edge blur sigma and the phases at which the transitions between edge and bar occur. We used four levels of psf width in addition to the original width. The results are plotted in figure 11 together with the results for our alternate image dataset (with its original psf).



11. Shows graphs of model parameters (vertical) plotted against psf width (horizontal). For all horizontal axes and the vertical axes of the three leftmost plots the graphed quantities are expressed as a fraction of filter scale. The leftmost plot concerns the edge template, the others concern the light bars (grey circles) and dark bars (black triangles). For all plots, outline symbols are for the van Hateren dataset, and the solid symbols for the alternate dataset. In the leftmost plot, the one solid symbol hides an outline symbol below it. The linear and constant functions shown are best fits to the van Hateren data only. For all graphs, the leftmost datapoints correspond to images without any additional blur applied.

The results for the edge blur sigma as a function of psf width are very similar to our previous study. They increase together with a slope near 1 and the extrapolated edge blur sigma for a zero psf width is 10% of the scale of the filters that measure the jet. The results for bar blur sigma are less clear. Both light and dark bars blur sigmas seem to follow the same linear relationship with psf width, but the slope is much less than one and the extrapolated blur sigma at zero psf width is a non-negligible 30% of the filter scale. In contrast, bar half-width and bar/edge phase transition are constant with psf width but they are different for light and dark bars. The light-bar/edge transition occurs at a phase of –0.97, the dark-bar/edge at 0.85. Both these transition phases are to the cosine side of the midway phase $\pi/4 \approx 0.79$, the difference being more extreme for dark- than light-bars.

Figure 11 also shows the model parameters for the alternate data-set. The only definite difference from the main data results is the degree of blur of the bar-templates: the light-bars are less blurred and the dark-bars more blurred than for the main data.

5. Discussions & Conclusions

In the following section we discuss aspects of our method that have changed since the previous study, and issues to do with scaling profiles before ML computation. After that is a discussion of notable aspects of our results in particular those that are unexplained by the present study. Finally we draw conclusions.

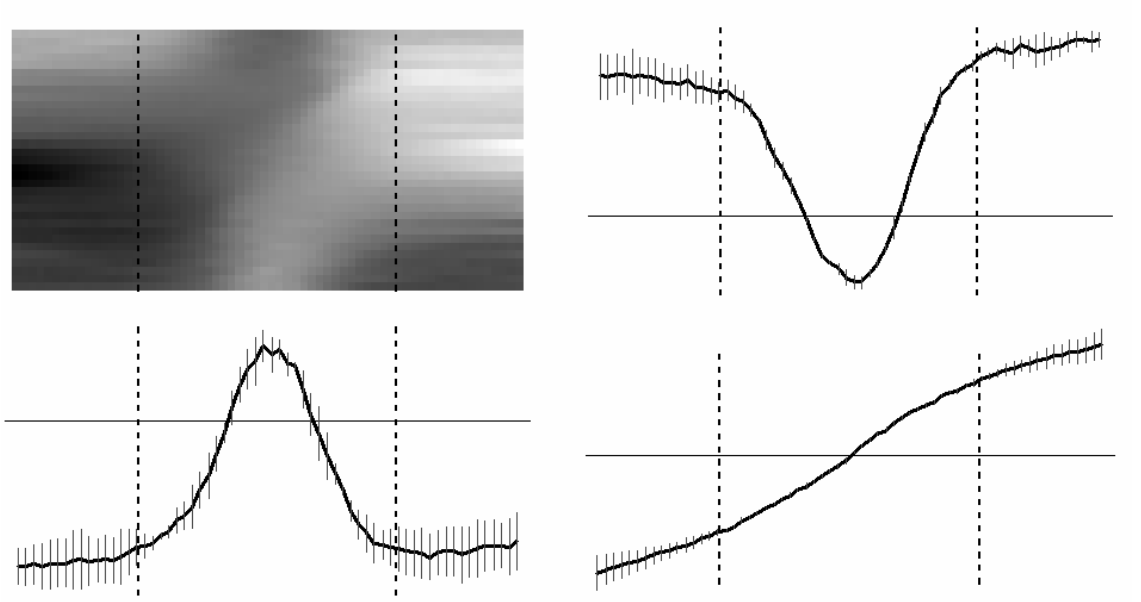
5.1 Discussion of Methods

There were two major changes in method from the previous study. First was that we sampled profiles across a range of scales by using a quad tree representation of each image (section 3.1.1). The motivation for this was that scale invariance had been found in the previous study

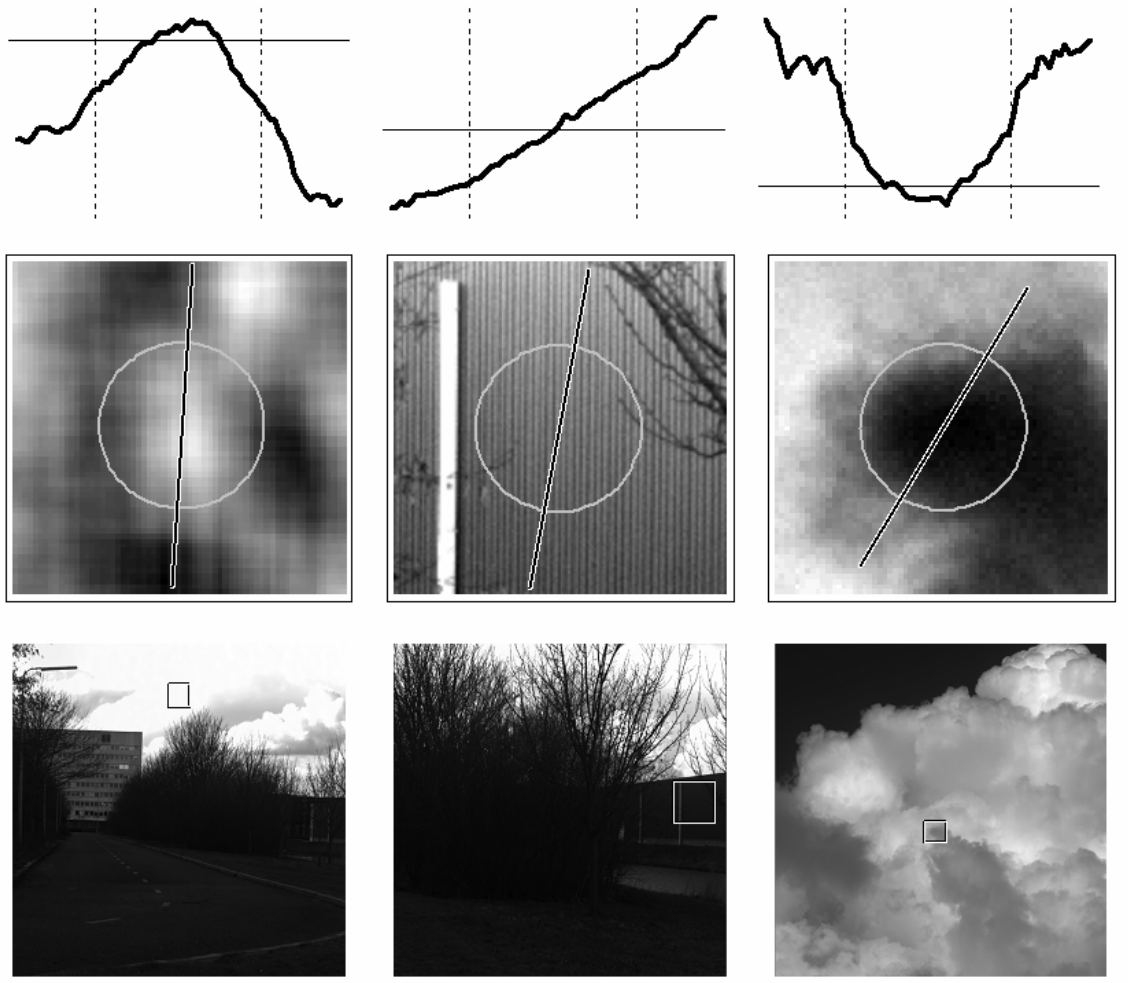
and so scale was another dimension (along with image position and orientation) that we could collapse to increase the generality of our results. Because scale invariance remains an interesting issue and because the change to the use of the quad tree is a significant algorithmic change, we performed a new test for scale invariance (section 3.1.4) and again found it to hold to a good level of approximation (section 3.2.3 and figure 8).

The second major change from our previous method was the mode estimation algorithm used. The new algorithm has been validated on noise profiles (sections 3.1.1 and 3.2.1) and we have also repeated the computations of our previous study. In summary, for the repeat of the previous study we found the same results as before except that the blur of the step edge that was the ML profile was greater using the new algorithm. In the pervious study the blur sigma of the ML form was 25% of the filter sigma, using the new mode estimation algorithm it was 38%. To these two blur sigmas should be joined a third for comparison, the blur sigma of the edge template in the current study was 19%. Thus the difference for the 1^{st} order results cannot be attributed to a consistent high blurring problem with the new mode estimation algorithm. The new algorithm is more principled than the previous, and on balance we prefer to trust its results.

An issue that we examined in some detail in the discussion of our previous study was the distinction between particular metamery classes (specified by an unmodified jet) and canonical metamery classes (where profiles have been intensity-scaled to a canonical form). There we noted that computing the ML profile of a canonical metamery class does not tell us the ML profiles of the particular metamery classes that transform to that canonical one. To see if this was more than a possibility, we computed ML profiles for subsets of the full population of profiles: light & flattish, light & steep, dark & flattish, and dark & steep. We found that the two steep subsets had ML forms like the full population, but the two flattish subsets were different. Although this result makes it very likely that particular metamery classes have different ML forms to the canonical we previously argued that this did not make our canonical results irrelevant. We wrote: “Our results on the canonical and the particular are not contradictory and may be phrased thus: if all that one knows of a randomly selected natural image profile is that it has non-zero first-order structure (true for 99.998% of profiles) then it is ML to be a step edge, if one knows something about the magnitude of the first-order structure that may alter what the ML form is.” Although we reassert this argument for the present study, it is clear that the particular/canonical difference remains unresolved and interesting. In an attempt to shed further light on to it, in the present study we also computed ML forms for restricted subsets of the population of profiles: light, dark, flattish and steep. Consistent with our previous findings, the only subset that gave different results from the full population was the flattish profiles. The ML form for these is shown in figure 12, and example flattish profiles with the ML form are shown, together with their image contexts, in figure 13. Further work is required to understand all the issues surrounding the particular/canonical distinction.



12. Shows examples of flattish profiles that are close to the ML forms for such profiles. The left and right columns are for cosine phase, the centre for sine phase. Like figure 10 the rows show the image context of the profiles.



13. Shows example flattish profiles having close to the ML form of the flattish profiles. The left and right columns are for cosine-phase, the centre for sine-phase. The rows of the figure are laid out like figure 11.

5.2 Discussion of Results

The first notable aspect of the ML results is the blur of the model template profiles. Why are the templates blurred at all, what determines the degree of blur and why is it different for edges and bars? In our previous study we concluded that the blur of the ML edges was due to the non-zero width of the image psf, was monotonically related to that quantity and plausibly went to zero with it in the extrapolated limit. Figure 11 (left) supports the same explanation for the blur of the edge templates in the present study. However the data plotted in figure 11 (2^{nd} column) suggest that only a small fraction of the bar template blur can be attributed to the non-zero width psf. We have observed that in general when fewer profiles are input into our mode estimation algorithm, one effect on the output is a general increase in the blur. Thus a plausible hypothesis for the higher blur of the bar templates is that more samples are needed for cosine than sine phases in order to achieve the same accuracy of mode estimation. This could be explained by the natural image distribution of profiles for cosine phases being less sharply peaked than for sine phases. This would also explain the higher scatter over repeat

computations that we observe for these phases (figure 7c). However why there should be such a difference between phases is unexplained and should be the subject of future work.

The second notable aspect of the ML results are the differences between light- and dark-bars. Figure 11 (right hand panels) shows that the light-bar templates are slightly wider than the dark-bar templates, and the phase extent of light bars is less than that of dark-bars. We have confirmed that this is not due to a programming error by computing ML profiles for intensity-negated images so it presumably reflects a real fact about natural images. A difference in the preponderance of light and dark bars in natural images would not be unexpected given the different types of scene geometry that cause these image structures e.g. there is no light bar equivalent of the silhouette (see figure 10).

The third notable aspect of the ML results is that the edge-bar phase transition zones each cover three phase bins rather just one. This present study cannot determine whether this is a real feature of the pattern of ML profiles or (as we think likely) a consequence of insufficient profile samples being used in our computations. Further work is needed to resolve this.

Finally we note the largest difference that we obtained in any of our control computations was for the blur of the bar templates for the alternate dataset (see figure 11, 2nd panel). Possibly the result is attributable to the small size of the alternate dataset, it has only 1.3% the number of pixels as the main set. For future studies we will use a more comparable dataset so that the significance of differences such as these can better be assessed.

5.3 Conclusions

Our previous study found that a step edge was the ML profile for the 1-D, 1st order jet. This was compatible with the ML profile for natural images being identical to the range- or the TV-minimizing profiles. Neither of these possibilities are borne out by the 2nd order results shown in figure 7. Initially the new results seemed negative with respect to our GTT proposal for features. However on closer examination we found a simple model that describes the results (figure 9). This model is fully compatible with our GTT proposal, which, to reiterate, is that the equivalence relation of qualitative equality between ML metameres induces feature classes on the jet space. With this proposal in mind, one can see that figure 9 shows that we have induced a categorization of the canonical 2nd order jet space into three classes separated by two fuzzy intermediate bands. Indeed the categories arise even more simply than we had expected as the appropriate relation of qualitative identity between ML profiles is very straightforward: equality modulo shifting and affine scaling. To further test the GTT proposal we need results for 2-D patches and ways to test whether the feature categories that derive from GTT are useful and/or correspond to features computed by the visual system [44].

6. References

1. Griffin, L.D., M. Lillholm, and M. Nielsen, *Natural image profiles are most likely to be step edges.* Vision Research, 2004. **44**(4): p. 407-421.
2. Marr, D., *Vision*. 1982, New York: W H Freeman & co.
3. Barlow, H.B., *Summation and inhibition in the frog's retina*. Journal of Physiology (London), 1953. **119**: p. 69-88.

4. Koenderink, J.J. and A.J. van Doorn, *Receptive Field Assembly Specificity.* Journal of Visual Communication and Image Representation, 1992. **3**(1): p. 1-12.
5. Koenderink, J.J., *What is a feature?* Journal of Intelligent Systems, 1993. **3**(1): p. 49-82.
6. Koenderink, J.J. and A.J. van Doorn, *Metamerism in complete sets of image operators,* in *Adv. Image Understan. '96*. 1996. p. 113-129.
7. Koenderink, J.J. and A.J. van Doorn, *Local Image Operators and Iconic Structure*, in *Algebraic Frames for the Perception-Action Cycle*, G. Sommer and J.J. Koenderink, Editors. 1997, Springer. p. 66-93.
8. Georgeson, M.A. and T.C.A. Freeman, *Perceived location of bars and edges in one-dimensional images: Computational models and human vision.* Vision Research, 1997. **37**(1): p. 127-142.
9. Young, R.A., *The Gaussian derivative model for spatial vision: I. Retinal mechanisms.* Spatial Vision, 1987. **2**: p. 273-293.
10. Young, R.A., R.M. Lesperance, and W.W. Meyer, *The Gaussian Derivative model for spatial-temporal vision: I. Cortical model.* Spatial Vision, 2001. **14**(3-4): p. 261-319.
11. Young, R.A. and R.M. Lesperance, *The Gaussian Derivative model for spatial-temporal vision: II. Cortical data.* Spatial Vision, 2001. **14**(3-4): p. 321-389.
12. Koenderink, J.J. and A.J. van Doorn, *Representation of Local Geometry in the Visual-System.* Biological Cybernetics, 1987. **55**(6): p. 367-375.
13. Koenderink, J.J., *Operational Significance of Receptive-Field Assemblies.* Biological Cybernetics, 1988. **58**(3): p. 163-171.
14. Koenderink, J.J. and A.J. van Doorn, *Receptive-Field Families.* Biological Cybernetics, 1990. **63**(4): p. 291-297.
15. Koenderink, J.J. and A.J. van Doorn, *Generic Neighborhood Operators.* Ieee Transactions on Pattern Analysis and Machine Intelligence, 1992. **14**(6): p. 597-605.
16. Florack, L.M.J., et al., *Families of Tuned Scale-Space Kernels,* in *Computer Vision - Eccv 92*. 1992. p. 19-23.
17. Debnath, L., *Integral Transforms and their Applications*. 1995: CRC Press.
18. Martens, J.B., *Local orientation analysis in images by means of the Hermite transform.* Ieee Transactions on Image Processing, 1997. **6**(8): p. 1103-1116.
19. Makram-Ebeid, S. and B. Mory, *Scale-space image analysis based on hermite polynomials theory*, in *Proc. Conf. on Scale Space Methods in Computer Vision*, L.D. Griffin and M. Lillholm, Editors. 2003, Springer. p. 57-71.
20. Koenderink, J.J., *The Structure of Images.* Biological Cybernetics, 1984. **50**(5): p. 363-370.
21. Georgeson, M.A., G.S.A. Barbieri-Hesse, and T.C.A. Freeman, *The primal sketch revisited: locating and representing edges in human vision via Gaussian-derivative filtering.* Perception, 2002. **31**: p. 1-1.
22. Leung, T. and J. Malik, *Representing and recognizing the visual appearance of materials using three-dimensional textons.* International Journal of Computer Vision, 2001. **43**(1): p. 29-44.
23. Liu, X.W. and D.L. Wang, *A spectral histogram model for texton modeling and texture discrimination.* Vision Research, 2002. **42**(23): p. 2617-2634.
24. Zhu, S.C., et al., *What are textons?*, in *Computer Vision - Eccv 2002, Pt Iv*. 2002. p. 793-807.
25. Julesz, B., *A brief outline of the texton theory of human vision.* Trends in Neuroscience, 1984. **7**(2): p. 41-45.

26. Lee, A.B., K.S. Pedersen, and D. Mumford, *The nonlinear statistics of high-contrast patches in natural images.* International Journal of Computer Vision, 2003. **54**(1-2): p. 83-103.
27. Pedersen, K.S., *Statistics of Natural Image Geometry*, in *Department of Computer Science*. 2003, University of Copenhagen: Copenhagen.
28. Koenderink, J.J. and A.J. van Doorn, *Local structure of gaussian texture.* IEICE Transactionson Information & Systems, 2003. **E86-D**(7): p. 1165-1171.
29. Debnath, L., *On Hermite Transforms.* Mathematicki Vesnik, 1964. **1**(16): p. 285-292.
30. van den Boomgaard, R., *Least squares and robust estimation of local image structure*, in *Proc. Scale Space Methods in Computer Vision*, L.D. Griffin and M. Lillholm, Editors. 2003. p. 237-254.
31. Tagliati, E. and L.D. Griffin, *Features in Scale Space: Progress on the 2D 2nd Order Jet*, in *LNCS*, M. Kerckhove, Editor. 2001, Springer. p. 51-62.
32. Griffin, L.D., *Local image structure, metamerism, norms, and natural image statistics.* Perception, 2002. **31**(3): p. 377-377.
33. Schrödinger, E., *Theorie der pigmente von grosster leuchtkraft.* Annalen den Physik, 1920. **62**: p. 603-622.
34. Kimmel, R. and A.M. Bruckstein, *Regularized Laplacian Zero Crossings as Optimal Edge Integrators.* International Journal of Computer Vision, 2003. **53**(3): p. 225-243.
35. Chan, T.F. and C.K. Wong, *Total variation blind deconvolution.* Ieee Transactions on Image Processing, 1998. **7**(3): p. 370-375.
36. Rudin, L.I., S. Osher, and E. Fatemi, *Nonlinear total variation based noise removal algorithms.* Physica D, 1992. **60**: p. 259-268.
37. Vogel, C.R. and M.E. Oman, *Fast, robust total variation-based reconstruction of noisy, blurred images.* Ieee Transactions on Image Processing, 1998. **7**(6): p. 813-824.
38. van Hateren, J.H. and D.L. Ruderman, *Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex.* Proceedings of the Royal Society of London Series B-Biological Sciences, 1998. **265**(1412): p. 2315-2320.
39. Mandelbrot, B. and J. van Ness, *Fractional brownian motions, fractional noises and applications.* SIAM Review, 1968. **10**(4): p. 422-437.
40. Comaniciu, D. and P. Meer, *Mean Shift: A robust approach toward feature space analysis.* IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002. **24**(5): p. 603-619.
41. Griffin, L.D. and M. Lillholm. *Mode Estimation by Pessimistic Scale Space Tracking*. in *Scale Space '03*. 2003. Isle of Skye, UK: Springer.
42. Lillholm, M., M. Nielsen, and L.D. Griffin, *Feature-based Image Analysis.* International Journal of Computer Vision, 2003. **52**(2): p. 73-95.
43. Press, W.H., et al., *Numerical Recipes in C++*. 2nd ed. 2002, Cambridge: Cambridge University Press.
44. Georgeson, M.A. and G.S. Barbieri-Hesse, *Bars and edges: What makes a feature for human vision?* Perception, 2003. **32**: p. 158-159.