Universal Analytical Forms for Modeling Image Probabilities

Anuj Srivastava, Xiuwen Liu, and Ulf Grenander

Abstract—Seeking probability models for images, we employ a spectral approach where the images are decomposed using bandpass filters and probability models are imposed on the filter outputs (also called spectral components). We employ a (two-parameter) family of probability densities, introduced in [11] and called **Bessel K forms**, for modeling the marginal densities of the spectral components, and demonstrate their fit to the observed histograms for video, infrared, and range images. Motivated by object-based models for image analysis, a relationship between the Bessel parameters and the imaged objects is established. Using L^2-metric on the set of Bessel K forms, we propose a pseudometric on the image space for quantifying image similarities/differences. Some applications, including clutter classification and pruning of hypotheses for target recognition, are presented.

Index Terms—Image probabilities, spectral analysis, Bessel K forms, clutter classification, target recognition, Gabor filters.

1 Introduction

Statistical techniques for image analysis and understanding require efficient and tractable probability models for analyzing the observed images. Given the tremendous variability associated with the imaged objects, detailed (e.g., 3D deformable templates) models are not feasible for "all possible objects." Therefore, one seeks a balance by designing low-level, coarse representations that are tractable and yet capture significant image variation. Here, we study a family of tractable, coarse probability models that can form building blocks of a larger image understanding system. Since the image space is very high-dimensional, a direct modeling of the joint probabilities is not possible, even if a large number of observations are provided, and some method for reducing dimensions is required. There are two general reductionist approaches adopted in the literature: 1) parameterize the probability densities using certain (low-dimensional) physical parameters (relating to the imaged objects) or 2) perform dimension reduction via purely numerical, nonphysical approaches.

In the first approach related to high-level vision, images are parameterized by the physical characteristics of the objects (such as shapes, textures, reflectance, illumination, and motion). These quantities are modeled mathematically, within some acceptable approximation and the resulting physical variables are used to analyze images. Probability models on images now consist of: 1) probability models on these physical variables and 2) the sensing models. An example of this idea is the *deformable template* theory [9] where images are studied through the transformations that match the physical templates to the observations. These models are detailed and capture sufficient variability to discriminate well even in challenging situations (cluttered scenes, low SNR, distant images, etc.). One drawback is that they are computationally expensive to implement, since they require synthesis of hypothesized images for comparison with the observed images. The second idea relates to low-level vision and involves one of many techniques that reduce dimensions using purely numerical considerations. That is, by not involving any physical consideration on the imaged objects, or any contextual knowledge, the images are treated as elements of a vector space and one seeks a low-dimensional subspace that best represents those numbers (under some chosen criterion). Principal components [16], independent components [5], [3], sparse coding [21], Fisher's discriminant [2], local linear embedding [24], and many other statistical learning algorithms are all instances of this idea. The main advantage here is the computational efficiency and the main drawback is knowledge deficiency. Lack of physical or contextual information leads to a limited performance, especially in challenging situations.

1.1 Models for Image Analysis

An important idea is to develop an adaptive strategy that balances these two levels of inferences. In this paper, we study a framework that provides some interaction between the pixel-based and the template-based inferences and is capable of shifting between the two depending on the available resources, both computational and informational. Consider a deformable template representation of the imaged objects, as laid out in [12], [26]. The basic idea is that images are made up of (images of) objects and their variability can be represented by physical variables. Using 3D models of objects (including polygonated surfaces, textures, and reflectance functions), all occurrences of these objects can be generated by applying similarity transformations. 3D scenes containing these transformed objects lead

- *A. Srivastava is with the Department of Statistics, Florida State University, Tallahassee FL 32306. E-mail: anuj@stat.fsu.edu.*
- *X. Liu is with the Department of Computer Science, Florida State University, Tallahassee, FL 32306. E-mail: liux@cs.fsu.edu.*
- *U. Grenander is with the Division of Applied Mathematics, Brown University, Providence, RI 02912. E-mail: ulf-grenander@home.com.*

Manuscript received 9 Nov. 2001; accepted 10 Feb. 2002.
Recommended for acceptance by D. Forsyth.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 115353.

to 2D images via occlusion and projection. What probability models on image can result from this model? As stated earlier, there are two possibilities:

1. The template approach where the set of possible objects is small and the 3D models are available for all objects. Also, the tools for synthesizing images of these objects under all transformations of pose and illumination are assumed available. Under these assumptions, a probability distribution on images can be written explicitly in terms of a joint probability on the transformation space and the set of objects. Object recognition is now solved on the product space [26], [19]. Inference procedures are laid out in the papers [10], [12], [19].
2. For general problems in image understanding, 3D models cannot be prestored for all objects, and furthermore, the transformations may not be identifiable in all conditions. Therefore, the representations tend to be less explicit and the image probability is motivated through empirical studies [20] and not via the physical parameters.

In this paper, we seek general models that retain some physical considerations, although not as explicitly as the template approach. Pursuing the second case, we replace 3D templates by their 2D profiles (call them **generators**) and denote them as gs. gs are the views (appearances, signatures, profiles) of randomly chosen objects, taken from random poses. Let $\mathcal{G}$ be the space of all possible generators associated with all objects, imaged from all angles. Random translations of 3D objects in a scene will be modeled by random placements and scalings of gs in an image.

Let $I : \mathbb{R}^2 \mapsto \mathbb{R}_+$ be the image map. Then, each object (with generator g_i) present in the scene contributes to the pixel value $I(z)$ according to $a_i g_i(\frac{1}{\rho_i}(z - z_i))$. Here, $z \in W \equiv [0, L] \times [0, L]$ is a variable for pixel location, $g_i : W \mapsto \mathbb{R}_+$ is a generator of a randomly chosen object, $\rho_i \in [0, L]$ is a random scale, and $a_i \in \mathbb{R}$ is the random weight associated with g_i. g_is are assumed to be drawn from $\mathcal{G}$ according to some probability dG. Although the physics of imaging dictates using occlusion models, we simplify the image formation by using the equation:

$$I(z) = \sum_{i}^{n} a_i g_i \left(\frac{1}{\rho_i}(z - z_i) \right), \; z, z_i \in W, \; a_i \in \mathbb{R}, \rho_i \in [0, L]. \tag{1}$$

Replacement of occlusion by superposition is a necessary simplification towards obtaining analytical probability models under linear filtering. Assume:

1. the weights a_is are i.i.d. standard normal,
2. the scales ρ_is are i.i.d. uniform on the interval $[0, L]$,
3. the locations z_is as samples from a 2D, homogeneous Poisson process with intensity $\lambda > 0$ on a compact set $W \subset \mathbb{R}^2$, and
4. a_is, g_is, z_is, and ρ_is are all assumed independent of each other.

Since g_is are assumed unknown, the related variables n, ρ_is, and z_is become indeterminable. However, we aim to derive probability models on I by implicitly incorporating their variability.

Motivated by a growing understanding of early human vision, a popular strategy has been to decompose images into their spectral components using a family of bandpass filters. Following that idea, our definition of a probability model on images is through its spectral representation. If certain low-dimensional statistics of these filtered components are found sufficient, then a significant reduction is achieved. Zhu et al. [30] have shown that the marginal distributions of spectral components, obtained using a collection of filters, sufficiently characterize homogeneous textures. The choice of histograms as sufficient statistics implies that only the frequencies of occurrences of (pixel) values in the filtered images are relevant and the location information is discarded [13], [31], [15]. Chubb et al. [4] also advocate the use of a histogram in texture representation. The main focus in these papers has been to model homogeneous textures, but we will apply spectral analysis to a general setting of image understanding. Portilla and Simoncelli [23] have suggested using the lower order statistics (mean, variance, skewness, kurtosis) to specify the marginal densities of the wavelet coefficients of the images. Wainwright et al. [28] have studied a family of Gaussian mixtures, resulting from different mixing densities, for modeling the observed histograms. Lee et al. [17]. have presented a synthesis model for capturing the statistics in the images of leaves. Donoho and Flesia [7] and Donoho and Huo [6] have proposed edge-based transforms to account for the patterns in the observed histograms.

Using a physical model for image formation, we have proposed a family of two-parameter probability densities [11], called Bessel K forms, to model the horizontal and the vertical derivatives of images. In this paper, we extend this model to a full spectrum of bandpass filters and arbitrary images. The two parameters associated with this family will form a sufficient statistic for a spectral component, denoting a significant reduction in the representation. The parameters can be directly estimated from the variance and the kurtosis of the filtered image pixels, thereby implying a simple estimation procedure. The main results presented here are: 1) demonstration of the success of Bessel K forms in modeling the spectral components for video, infrared, and range images of natural and artificial scenes, 2) derivation of an analytical expression for computing the L^2-metric, on the Bessel family, that leads to a pseudometric on image space, and 3) use the Bessel K forms (and the pseudometric) as a tool for clutter classification and for pruning possible hypothesis set for recognition of objects from their images. The last claim is based on: 1) motivating the model in (1) by relating it to the 3D deformable template representation, 2) relating Bessel parameters to certain physical characteristics of the imaged objects, and 3) an example of pruning hypothesis using the COIL database. In addition, we will present an asymptotic approximation of the Bessel K form that can potentially simplify future inference procedures.

This paper is organized as follows: Section 2 applies Bessel K forms to model spectral components of images and associates the estimated Bessel K parameters with the observed shapes. Section 3 derives an L^2-metric on the Bessel K forms and on the image space, while Section 4

applies this metric to clutter classification and target recognition.

2 Probability Models for Image Spectra

We start with some notation. Given an image I and a bank of filters $\{F^{(j)}, j = 1, 2, \ldots, K\}$, we compute for each filter $F^{(j)}$ a filtered image $I^{(j)} = I * F^{(j)}$, where $*$ denotes the 2D convolution operation. As an example, a Gabor filter is a bandpass filter with a Gaussian kernel centered around a specific wavenumber (see [14] for details). For a rotation $\theta \in [0, 2\pi)$, a Gabor filter is given by:

$$F(z) = \exp\left(-\frac{1}{2\sigma^2}\left(z_\theta(1)^2 + z_\theta(2)^2\right)\right)\exp\left(-j\frac{2\pi z_\theta(1)}{\sigma}\right),$$

where σ denotes the resolution associated with the filter and

$$z_\theta = \begin{bmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{bmatrix}\begin{bmatrix}z_1\\ z_2\end{bmatrix} \in \mathbb{R}^2.$$

Another filter suggested by Marr [18] to model early vision is the Laplacian of Gaussian filter whose operation on I is given by $(G * \Delta)I$, where G is a Gaussian kernel and Δ is the Laplacian operator. In addition to these filters, one can utilize a wide variety of filters: neighborhood operators, steerable filters, interpolation filters, etc. Each filter selects and isolates certain features present in the original image. In this paper, we do not address the issue of selecting filters to best accomplish a specific task. Instead, we will assume an arbitrary choice of filters as long as the resulting spectral components have marginals that are: 1) unimodal, 2) symmetric around the mode, and 3) are leptokurtic, i.e., their kurtosis are more than that of a Gaussian random variable with the same variance. Additionally, we want the filters such that the resulting representation is computationally efficient which holds for filters with smaller bandwidths.

Remark 1. In this paper, we force the mode to be at zero by setting $I^{(j)} = I^{(j)} - E[I^{(j)}]$.

2.1 Analytical Models

Applying 2D convolution to both sides of (1), we obtain a spectral component

$$\begin{aligned} I^{(j)}(z) &\equiv \left(I * F^{(j)}\right)(z)\\ &= \sum_i a_i g_i^{(j)}\left(\frac{1}{\rho_i}(z - z_i)\right), \text{ where } g_i^{(j)} = F^{(j)} * g_i. \end{aligned} \tag{2}$$

The *conditional* density of $I^{(j)}(z)$, given the Poisson points $\{z_i\}$, the scales $\{\rho_i\}$, and the profiles $\{g_i\}$, is normal with mean zero and variance u, where

$$u \equiv \sum_i (g_i^{(j)}\left(\frac{1}{\rho_i}(z - z_i)\right)^2.$$

One departure here, from the model used in [11], is that the generators are now randomly selected and are included at random scales ρ_is in (2). Under this model and assuming u to be a scaled-Gamma random variable, the density function of the random variable $I^{(j)}(z)$ has been shown to be [11]: for $p > 0$, $c > 0$,

$$f(x; p, c) = \frac{1}{Z(p, c)}|x|^{p-0.5}K_{(p-0.5)}\left(\sqrt{\frac{2}{c}}|x|\right),\ x \in \mathbb{R}, \tag{3}$$

where K is the modified Bessel function (see (11) in Appendix A.1) and Z is the normalizing constant given by

$$Z(p, c) = \sqrt{\pi}\Gamma(p)(2c)^{0.5p+0.25}.$$

Remark 2. If $E[I^{(j)}] = \mu$ is not equal to zero then we can account for it by defining a three-parameter family:

$$f(x; \mu, p, c) = \frac{1}{Z(p, c)}|x - \mu|^{p-0.5}K_{(p-0.5)}\left(\sqrt{\frac{2}{c}}|x - \mu|\right),\ x \in \mathbb{R}.$$

As stated earlier, in this paper, we restrict to only two parameters by setting $\mu = 0$.

Let $\mathcal{D}$ be the space of all such densities:

$$\mathcal{D} = \{f(x; p, c) | p > 0, c > 0\}.$$

We refer to the elements of $\mathcal{D}$ as the **Bessel K forms** and the parameters (p, c) as the **Bessel parameters**. The elements of $\mathcal{D}$ have the following properties:

1. They are symmetric and unimodal with the mode at zero. For $p = 1$, $f(x; p, c)$ is the density of a double exponential. In general, it is the pth convolution power (for any $p > 0$) of a double exponential density. Therefore, it is unimodal with the mode at $x = 0$. For the same reason, it is symmetric around zero.
2. The kurtosis of a Bessel K form relates to the term
$$3 + \frac{\text{var}(u)}{(E[u])^2}.$$
Hence, the Bessel K forms are leptokurtic (the tails are heavier as compared to a normal curve with the same variance).
3. A Bessel K form is a specific kind of normal variance-mean mixture where the mixing variable is a scaled Gamma with parameters p and c. It becomes a special case of a larger family of self-reciprocal normal variance mixtures as described by Barndorff-Nielson et al. [1]. This connection opens the possibility of a larger family, namely, the *generalized hyperbolic distributions*, to be used in modeling image spectra, if needed. In addition to the shape and the scale, this family allows for a location parameter, skewness, and different rates (of exponential decay) on the two tails. The significance here for image analysis is that more filters, beyond the ones that lead to symmetric histograms, can also be included in the analysis.
4. The family of Bessel K forms is infinitely divisible, i.e., any random variable in this family can be written as a sum of two independent random variables from this family. However, if I_1 and I_2 are independent with densities $f(x; p_1, c_1)$ and $f(x; p_2, c_2)$, respectively, with $c_1 \neq c_2$, the density of $a_1 I_1 + a_2 I_2$ ($a_1, a_2 \in \mathbb{R}$) may not

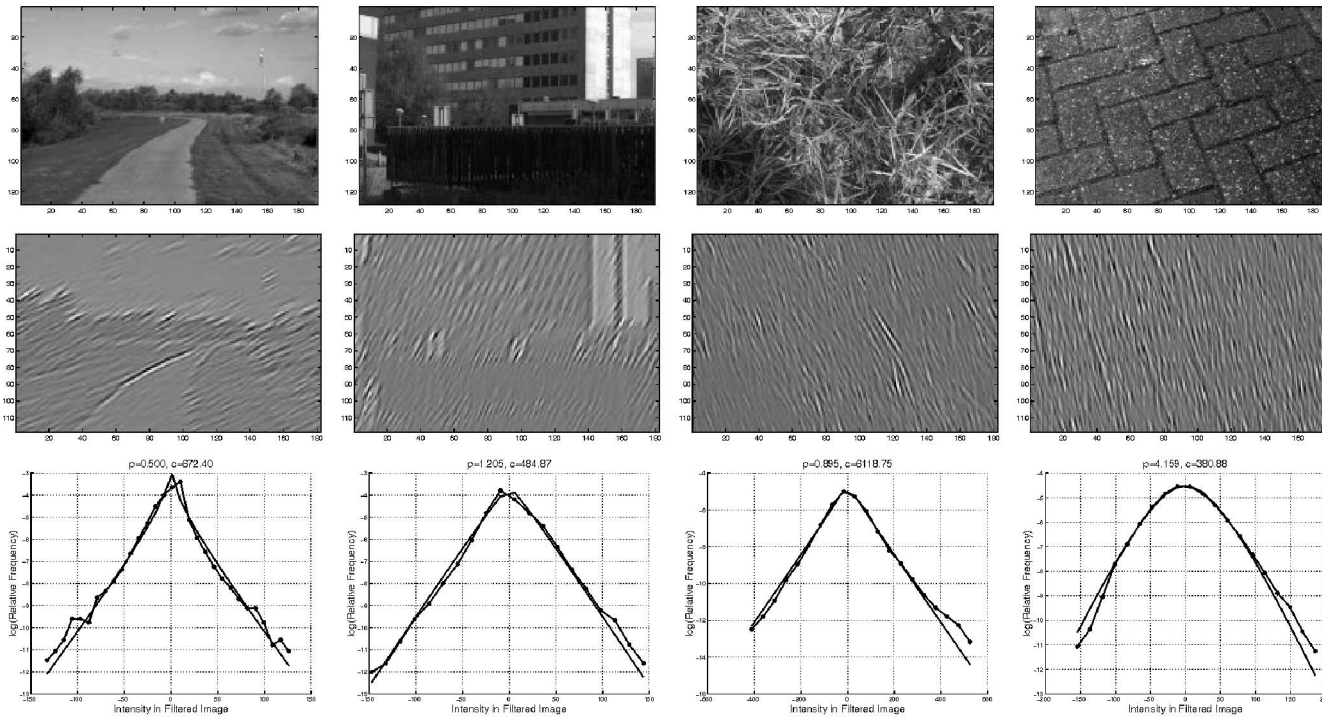


Fig. 1. Images (top panels), their Gabor components (middle panels), and the marginal densities (bottom panels). The observed densities are drawn in marked lines and the estimated Bessel K forms are drawn in solid lines.

be a Bessel K form but, in certain conditions, can be approximated by $f(x;p,c)$, where

$$p = \frac{(a_1^2 p_1 c_1 + a_2^2 p_2 c_2)^2}{a_1^4 p_1 c_1^2 + a_2^4 p_2 c_2^2}, \text{ and } c = \frac{a_1^4 p_1 c_1^2 + a_2^4 p_2 c_2^2}{a_1^2 p_1 c_1 + a_2^2 p_2 c_2}.$$

5. A Bessel K form is square-integrable only for $p > 0.25$. This property is due to the choice of Gamma density for u and it limits our later derivation of L^2-metric to the Bessel K forms with p-values larger than 0.25.

How does one estimate the Bessel K parameters for a given filtered image? p and c can be estimated using the equations:

$$\hat{p} = \frac{3}{\mathrm{SK}(I^{(j)}) - 3}, \ \hat{c} = \frac{\mathrm{SV}(I^{(j)})}{\hat{p}}, \tag{4}$$

where SK is the sample kurtosis and SV is the sample variance of the pixel values in $I^{(j)}$. A derivation of these estimators is omitted here as a similar result is presented in [11]. The computational task of estimating the marginal density is that of computing the second and the fourth moments of the filtered image. We illustrate some estimation results for a variety of images.

- Shown in the top panels of Fig. 1 are some images taken from the van Hateren [27] database. The middle panels display their specific filtered forms (or spectral components) for Gabor filters at arbitrarily chosen orientations and scales, and the bottom panels plot the marginal densities. On a log scale, the observed densities (histograms) are plotted in the marked (knotted) lines and the estimated Bessel K forms ($f(x;\hat{p},\hat{c})$) are plotted in the solid lines.
- For the image shown in the top panel of Fig. 2, we have estimated Bessel K forms for many Gabor filters. The middle panels show the marginals for different filter orientations ($\theta = 30, 90, 120$, and 150 degrees), while keeping the scale fixed at $\sigma = 4.0$ and the bottom panels are for different filter scales ($\sigma = 4, 6, 8$, and 10) keeping the orientation fixed at $\theta = 150$ degrees.
- Fig. 3 shows examples of estimation when the images are filtered by Laplacian of Gaussian filters. The top panels show some natural images from the van Hateren database [27] and the bottom panels show the corresponding estimated Bessel K forms.
- Fig. 4 shows estimation results for three infrared face images when filtered by Gabor filters. These results suggest the role of Bessel K forms in modeling images beyond the visual spectrum. For an application of Bessel K forms in infrared face recognition, please refer to the article [25].
- Shown in Fig. 5 are some examples of estimating marginal densities for the case of outdoor range images taken from the Brown range database of Lee and Huang. The three images shown in top panels are filtered using Gabor filters and the resulting densities are plotted in the bottom panels.

Instead of estimating Bessel parameters using moments, it may be preferable to use some robust estimation technique to account for the outliers. A simple idea is to consider a fraction (say one percent) of the tail as outlier and discard it in parameter estimation. Using this idea, we have found an improvement in estimation performance for the cases where p is quite small ($p < 0.1$). Other possibilities include maximum-likelihood estimation of $p,\ c$, or a robust estimation technique based on relating quartiles to p and c.

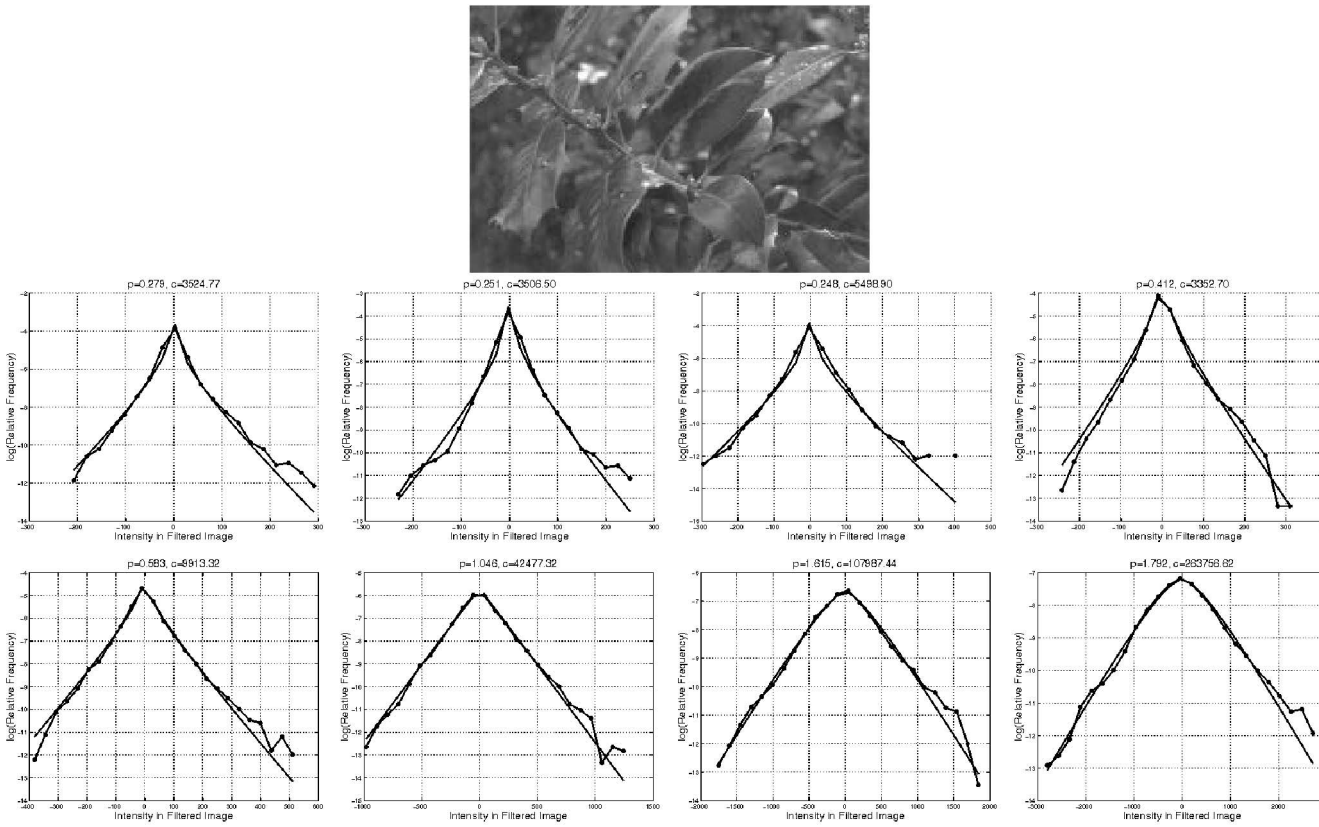


Fig. 2. Plots of observed and estimated marginals (on a log scale) of the spectral components of a given image (top panel). Middle panels depict the marginals for different filter orientations while the bottom panels are for different filter scales.

2.2 Performance Analysis of Bessel K Forms

To quantify the performance in modeling observed histograms by estimated Bessel K forms, a number of quantities can be used and we choose the Kullback-Leibler (KL) divergence. For any two density functions f_1 and f_2, the KL divergence is defined as the quantity:

$$KL(f_1, f_2) = \int_{\mathbf{R}} \log\left(\frac{f_1(x)}{f_2(x)}\right) f_1(x)dx.$$

We have computed it by discretizing at the center points of the histogram bins. To illustrate KL divergence values for our applications, we start with some examples. Shown in Fig. 6 are six plots, each containing a pair of densities: one observed (marked line) and one estimated (solid line), and the KL divergence between them is listed on the top. These six cases show a decreasing match, in going from top left to bottom right. The KL divergence, in that order, is found to be 0.0008, 0.0058, 0.0105, 0.00591, 0.247, and 0.3341, respectively.

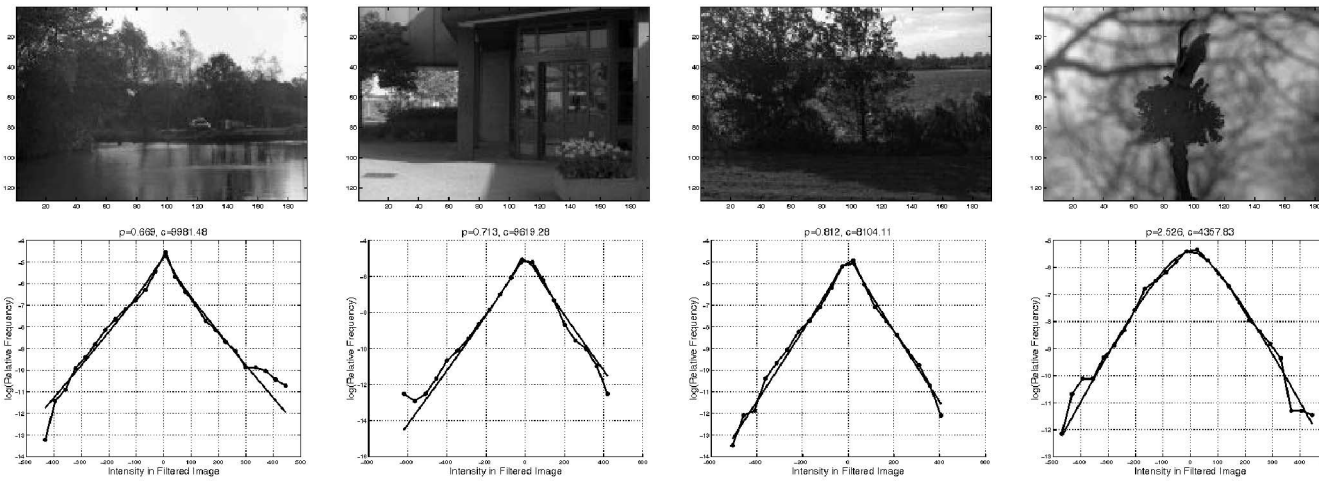


Fig. 3. Estimated Bessel K forms for natural images (top panels) when filtered by Laplacian of Gaussian filters.

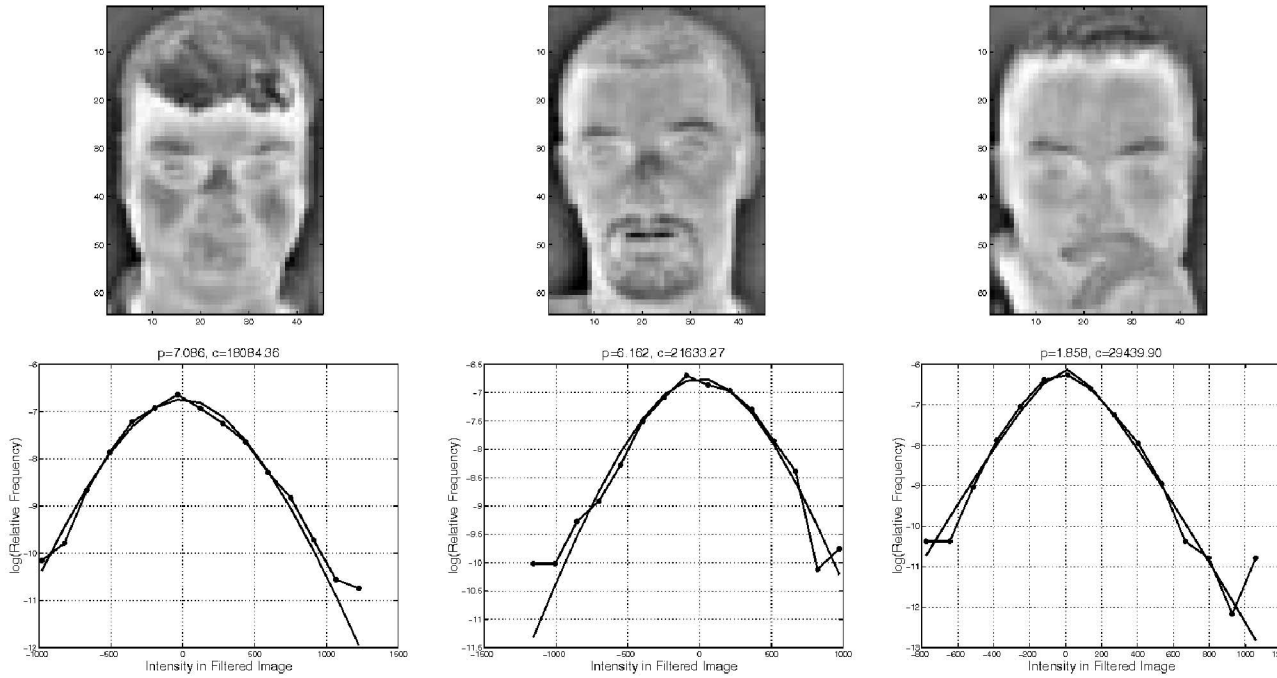


Fig. 4. Observed and estimated marginal densities (bottom panels) for the IR face images (top panels) and arbitrary Gabor filters.

Using KL divergence for evaluating matches between the observed and the estimated densities, we have computed the performance over two large databases. In each case, for a large combination of images and filters drawn randomly, we have averaged the KL divergence over thousands of resulting filtered marginals. The first database is made up of 300 natural video images downloaded from the van Hateren natural image database, and the second database is made up of 220 infrared face pictures. Shown in Fig. 7 are the convergence plots of the average KL divergence, plotted against the sample size. The left plot is for the natural video images with a limiting value of 0.0719, while the right plot is for the infrared images with a limiting value of 0.0479. A comparison of these values with the examples in Fig. 6 underscores a high degree of match between the observed histograms and the estimated Bessel K forms.

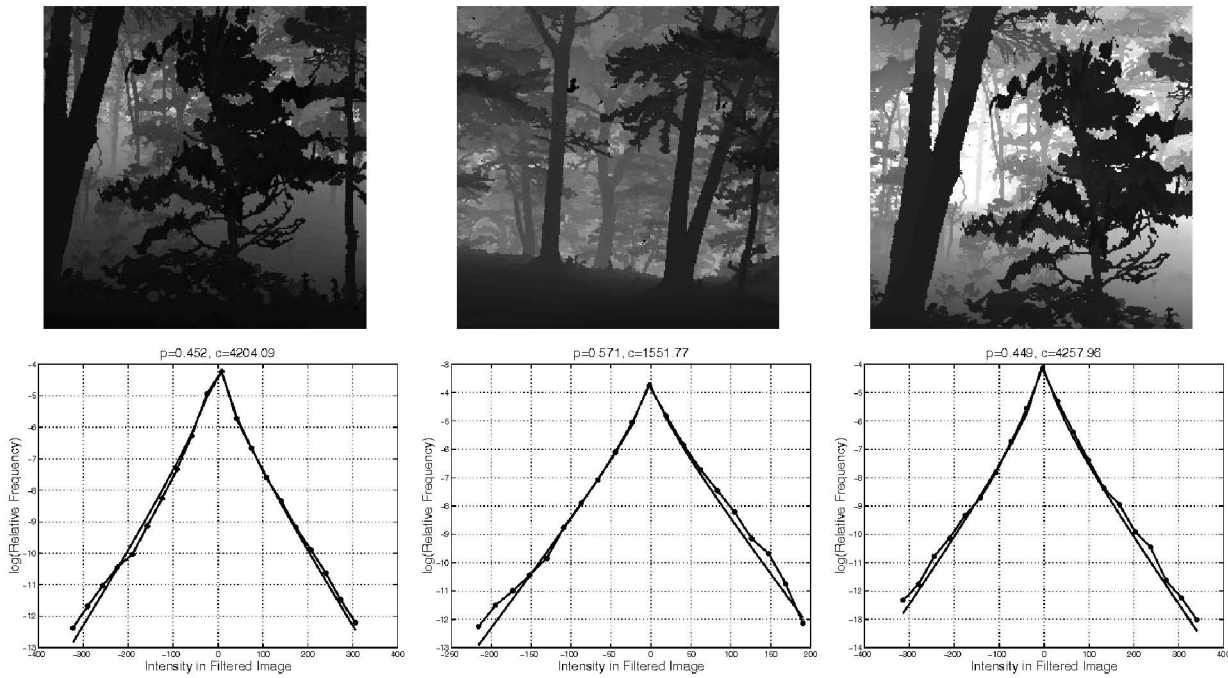


Fig. 5. Top panels: range images of a forest. Bottom panels: corresponding observed (marked lines) and estimated marginal densities (solid lines).

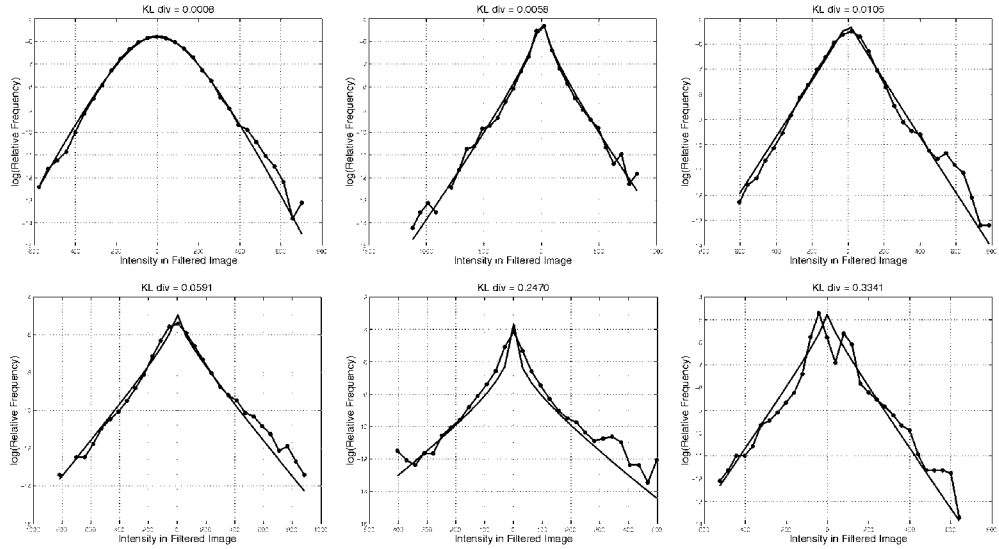


Fig. 6. Examples of Kullback-Leibler divergence: The divergence values for the six plots are listed on the top.

2.3 Relating the Shape Parameter to the Imaged Objects

Before we present some applications of these Bessel K forms, we first deal with an interesting question. How do the Bessel K parameters p and c, estimated for a spectral component of an image, relate to the objects that are present in that image? The physical characteristics of the imaged objects and the filter used in generating a spectral component, should dictate the resulting Bessel K form. Since c is essentially a scale parameter relating to the range of pixel values in I, its role is not as important as p. Here, we derive a relationship between the imaged generators and the estimated shape parameter. We will perform the analysis for the pixel $I(z)$ as composed of the generators $\{g_i(\frac{1}{\rho_i}(z - z_i))\}$ (1), although the analysis remains the same for any spectral component $I^{(j)}(z)$ composed of the generators $\{g_i^{(j)}(\frac{1}{\rho_i}(z - z_i))\}$ (2). Let the characteristic function of the random variable a_i be given by $\phi(\omega)$. (Later, we assume a specific ϕ by choosing a_is to be standard normal.) The conditional characteristic function of $I^{(j)}(z)$, given the Poisson points $\{z_i\}$s, the scales $\{\rho_i\}$s, and the generators $\{g_i\}$s, is:

$$\Psi_I(\omega|\{z_i\},\{\rho_i\},\{g_i\}) = \prod_{i=1}^{n} \phi\left(\omega g_i\left(\frac{1}{\rho_i}(z - z_i)\right)\right),$$

using the i.i.d. nature of a_is. Integrating out the uniform, independent placement of z_is, for a given n, we obtain the conditional characteristic function:

$$\Psi_I(\omega|n,\{g_i\},\{\rho_i\}) \propto \prod_{i=1}^{n}\left(\int_W \phi\left(\omega g_i\left(\frac{1}{\rho_i}(z - z_i)\right)\right)dz_i\right).$$

Similarly, integrating out the scales, we get:

$$\Psi_I(\omega|n,\{g_i\}) \propto \prod_{i=1}^{n}\left(\int_W \int_0^L \phi\left(\omega g_i\left(\frac{1}{\rho_i}(z - z_i)\right)\right)d\rho_i dz_i\right).$$

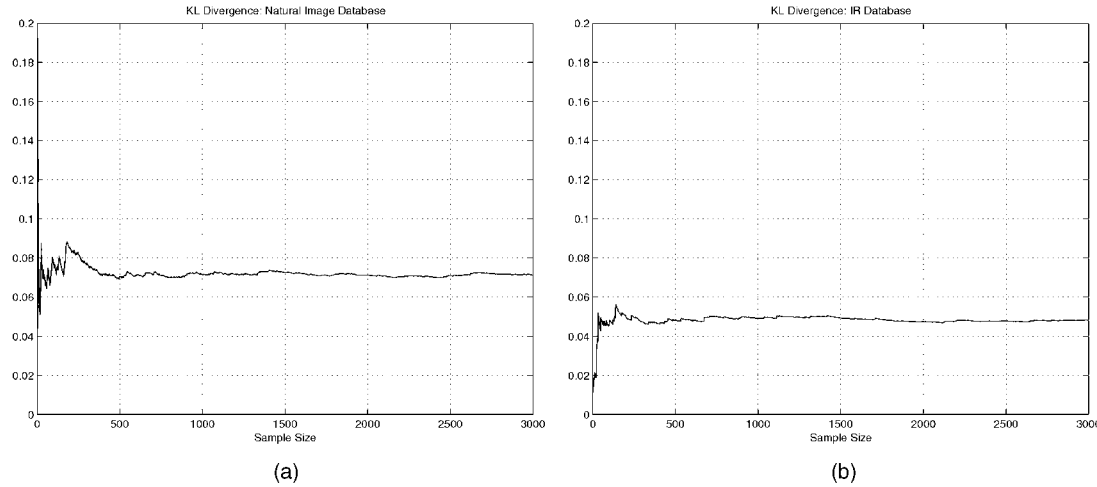


Fig. 7. Convergence of average KL-divergence between the observed and the estimated densities as the sample size increases. (a) Average for the van Hateren database of natural images. (b) Average for the FSU infrared face database.

Now, integrate out the random selection of the generators g_is. Recall that each g_i is drawn independently from the generator space $\mathcal{G}$ according to some measure dG. This gives,

$$\Psi_I(\omega|n) \propto \left(\int_{\mathcal{G}}\int_W\int_0^L \phi\left(\omega g_1\left(\frac{1}{\rho_1}(z-z_1)\right)\right)d\rho_1 dz_1 dG_1\right)^n.$$

The last step is to integrate with respect to n. For the image defined over the set W, n is a Poisson random variable with mean $\lambda|W|$. To simplify notation, we use λ to denote $\lambda|W|$ in the following:

$$\Psi_I(\omega) \propto \sum_{n=0}^{\infty} \left(\frac{\exp(-\lambda)\lambda^n}{(n!)}\left(\int_{\mathcal{G}}\int_W\int_0^L \phi\left(\omega g_1\left(\frac{1}{\rho_1}(z-z_1)\right)\right)d\rho_1 dz_1 dG_1\right)^n\right)$$
$$= \exp\left(\lambda\left(\int_{\mathcal{G}}\int_W\int_0^L \phi\left(\omega g_1\left(\frac{1}{\rho_1}(z-z_1)\right)\right)d\rho_1 dz_1 dG_1 - 1\right)\right).$$

Without loss of generality, we can substitute $z=0$. To simplify, assume that the generators $g \in \mathcal{G}$ are all even functions. Then,

$$\Psi_I(\omega) \propto \exp\left(\lambda\left(\int_{\mathcal{G}}\int_W\int_0^L \phi\left(\omega g_1\left(\frac{z_1}{\rho_1}\right)\right)d\rho_1 dz_1 dG_1 - 1\right)\right). \tag{5}$$

To find the cumulants of I, we use the relation:

$$\mu_k = \frac{d^k \log(\Psi_I(\omega))}{d\omega^k}|_{\omega=0}.$$

The two cumulants that we need are:

$$\mu_2 = \lambda\phi''(0)\left(\int_{\mathcal{G}}\int_W\int_0^L g_1\left(\frac{z_1}{\rho_1}\right)^2 d\rho_1 dz_1 dG_1\right),$$
$$\mu_4 = \lambda\phi^{(iv)}(0)\left(\int_{\mathcal{G}}\int_W\int_0^L g_1\left(\frac{z_1}{\rho_1}\right)^4 d\rho_1 dz_1 dG_1\right).$$

The kurtosis of $I(z)$, according to this model, is given by:

$$\text{kurtosis}(I) = \frac{\mu_4}{\mu_2^2} = \frac{\phi^{(iv)}(0)\left(\int_{\mathcal{G}}\int_W\int_0^L g_1\left(\frac{z_1}{\rho_1}\right)^4 d\rho_1 dz_1 dG_1\right)}{\lambda\left(\phi''(0)\left(\int_{\mathcal{G}}\int_W\int_0^L g_1\left(\frac{z_1}{\rho_1}\right)^2 d\rho_1 dz_1 dG_1\right)\right)^2}.$$

Comparing this with an earlier result that $p = \frac{3}{\text{kurtosis}(I)-3}$, derived in [11], we get

$$p = \frac{1}{\frac{\kappa}{3\lambda}-1}, \text{ where } \kappa = \frac{\phi^{(iv)}(0)\left(\int_{\mathcal{G}}\int_W\int_0^L g_1\left(\frac{z_1}{\rho_1}\right)^4 dz_1 dG_1\right)}{(\phi''(0))^2\left(\int_{\mathcal{G}}\int_W\int_0^L g_1\left(\frac{z_1}{\rho_1}\right)^2 d\rho_1 dz_1 dG_1\right)^2}. \tag{6}$$

Remark 3. One can replace g_1 by $g_1^{(j)}$ to obtain this relation for any spectral component of I. When $a_i \sim N(0,1)$, ρ_1 is fixed to be 1.0, and all the $g_i \equiv g$ (i.e., a fixed generator), then κ simplifies to

$$\kappa = \frac{\left(\int_W g(z_1)^4 dz_1\right)}{\left(\int_W g(z_1)^2 dz_1\right)^2}.$$

This equation provides an important relationship between a generator g and the parameter p. According to (6), $p<1$ occurs when $\lambda < \frac{\kappa}{6}$. If the generator g has sharp, distinct boundaries (i.e., κ is larger), then the p value is small unless the frequency of occurrence (λ) is large. Specifically, if a filter $F^{(j)}$ is used to extract a particular feature (e.g., oriented edges, junctions, bands, etc.) from the image I, then the value of p is dictated by the **distinctness** (κ) and the **frequency of occurrence** (λ) of that feature in the image. For example, shown in Fig. 8 is a variation of p value when the images are filtered for extracting vertical edges ($\theta = 90$). The top row shows images with increasing frequency of vertical edges in going from left to right. Correspondingly, the estimated p value shows an increase (0.102, 0.541, 1.473, and 8.83). Summarizing the relation between p and κ, we have:

$$\text{If } \begin{cases} 0 < \lambda < \kappa/6 & \text{then } p < 1 \\ \kappa/6 < \lambda < \kappa/3 & \text{then } p > 1 \end{cases}.$$

2.4 Asymptotic Approximation of Bessel K Forms

Although the proposed Bessel K forms model the observed histograms very well, their functional form is not easy to work with. Following the discussion in [1], it is possible to approximate the tails of Bessel K forms using a gamma density as follows: For large values of x, the modified Bessel function $K_\rho(x)$ can be approximated by the function $\sqrt{\frac{\pi}{2x}}\exp(-x)$ uniformly over compact sets of ρ. Since small values of ρ imply a large tail, this approximation holds well for a large part of the domain whenever p is small. The tails of a Bessel K form with the parameters p and c can be approximated well by the function:

$$\tilde{f}(x;p,c) = \frac{1}{\Gamma(p)}\left(\frac{2}{c}\right)^{p/2}|x|^{p-1}\exp\left(-\sqrt{\frac{2}{c}}|x|\right). \tag{7}$$

This result is useful as it provides a simpler expression although only in an asymptotic setting. Since for $p>1$, the maxima of $\tilde{f}(x;p,c)$ are attained at the points $x = \pm(p-1)\sqrt{\frac{2}{c}}$ and not at $x=0$, this approximation is valid only for $p << 1.0$. To illustrate this approximation on the filtered marginals, three example are shown in Fig. 9. The top panels show original images and the bottom panels show the log densities for arbitrary Gabor filters. The observed histograms are plotted in solid lines, the Bessel K forms $f(x;,p,c)$ are plotted in lines with + signs (- + -), and the asymptotic approximations $\tilde{f}(x;p,c)$ are plotted in lines with * (-*-).

3 Pseudometrics for Comparing Images

We have chosen to represent images via the Bessel parameters of their spectral components. One distinct advantage, of having such analytical forms for the marginals of the

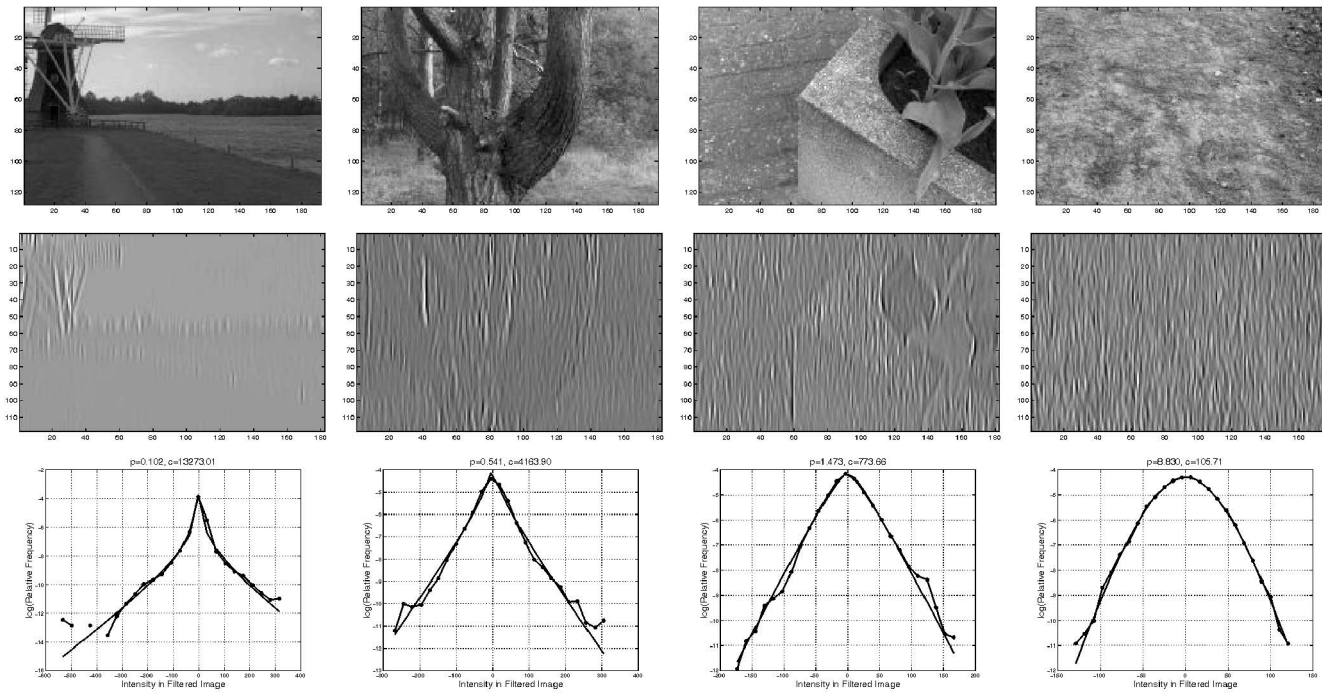


Fig. 8. Variation of p-values for extracting vertical edges ($\theta = 90$). Top panels are the original images, middle panels are the filtered images, and the bottom panels are the densities (log-scale). The estimated p-values are: 0.102, 0.541, 1.473, and 8.8, respectively.

spectral components, is the resulting theoretical framework for image analysis. For instance, we would like to be able to compare images by directly comparing their respective Bessel parameters. An analytical form is very useful in the sense that we do not need to estimate the densities for this comparison.

To quantify the distance between two Bessel K forms, we have chosen the L^2-metric on $\mathcal{D}$. It is possible that other metrics, such as the Kullback-Leibler divergence, Renyi's α-divergence, or even the L^1 metric, may prove more useful in certain situations. Since we are restricting ourselves to only $\mathcal{D}$, and not the full set of pdfs, we suggest that many of these choices will provide similar results, especially if the task is classification or hypothesis pruning. The main drawback of choosing L^2 is that Bessel K forms are not in L^2 for $p < 0.25$. In the case of natural images, the p-values are mostly larger than 0.25, while, for images of objects with sharp, well-defined edges, p can sometimes be below 0.25.

Remark 4. In cases where an image-filter combination leads to $p < 0.25$, we can choose one of following: 1) drop that filter, 2) replace p by $0.25 + \epsilon$, and then compute the L^2-metric, or 3) compute the L^2-metric numerically using the quadrature integration at a certain resolution.

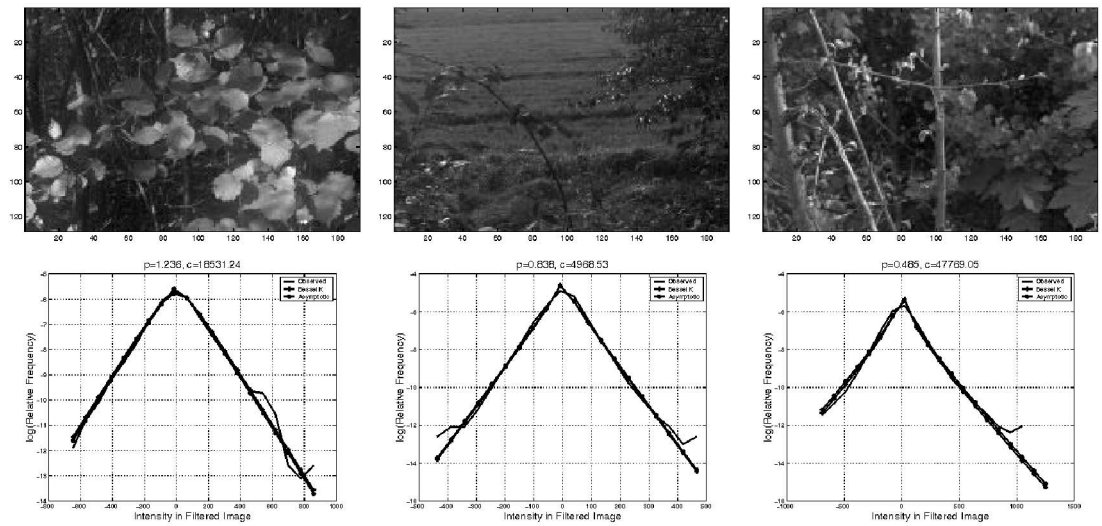


Fig. 9. Asymptotic approximations of Bessel forms: For the images shown in top panels, the bottom panels plot the observed histogram (solid), Bessel K form (-+-), and the approximation (-*-).

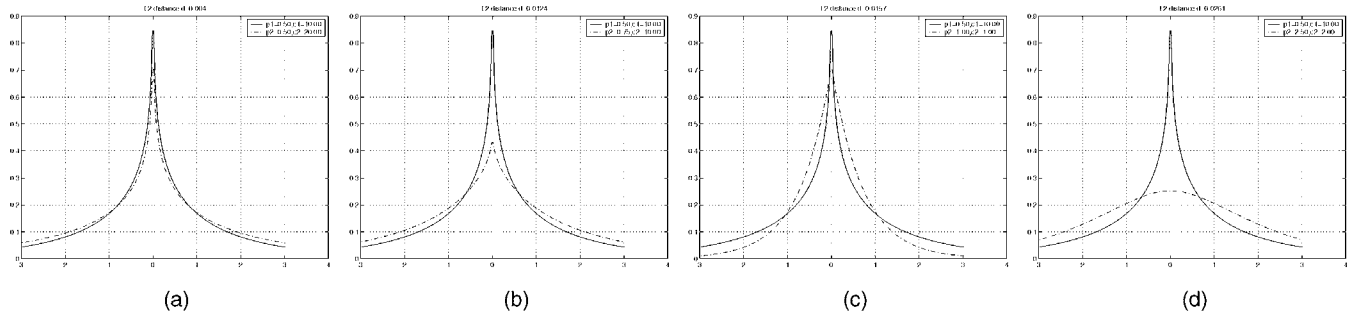


Fig. 10. Examples of the L^2-metric on $\mathcal{D}$: (a) $p_2 = 0.5,\ c_2 = 20,\ d = 0.063$, (b) $p_2 = 0.75,\ c_2 = 10,\ d = 0.114$, (c) $p_2 = 1.0,\ c_2 = 1.0,\ d = 0.125$, (d) $p_2 = 2.5,\ c_2 = 2.0,\ d = 0.161$. $p_1 = 0.5$ and $c_1 = 10.0$ are held constant.

For $f(x; p_1, c_1)$ and $f(x; p_2, c_2)$ in $\mathcal{D}$, the L^2-metric is

$$d(p_1, c_1, p_2, c_2) = \sqrt{\int_x (f(x; p_1, c_1) - f(x; p_2, c_2))^2 dx}.$$

This metric can be computed in a closed form, under certain restrictive conditions, as follows.

Theorem 1. *The L^2-distance between the two Bessel K densities, parameterized by (p_1, c_1) and (p_2, c_2), respectively, is given by: for $p_1, p_2 > 0.25$, $c_1, c_2 > 0$,*

$$d(p_1, c_1, p_2, c_2) = \left(\frac{1}{2\sqrt{2\pi}}\Gamma(0.5)\left(\frac{\mathcal{G}(2p_1)}{\sqrt{c_1}} + \frac{\mathcal{G}(2p_2)}{\sqrt{c_2}} - \frac{2\mathcal{G}(p_1+p_2)}{\sqrt{c_1}}(\frac{c_1}{c_2})^{p_2}\mathcal{F}\right)\right)^{\frac{1}{2}}, \tag{8}$$

where $\mathcal{G}(p) = \frac{\Gamma(p-0.5)}{\Gamma(p)}$ and $\mathcal{F} = F((p_1+p_2-0.5), p_2; p_1+p_2; 1-\frac{c_1}{c_2})$ (F is the hypergeometric function). If $c_1 = c_2 = c$, then the metric simplifies to:

$$\left(\frac{1}{2\sqrt{2\pi}}\frac{\Gamma(0.5)}{\sqrt{c}}(\mathcal{G}(2p_1) + \mathcal{G}(2p_2) - 2\mathcal{G}(p_1+p_2)\right).$$

Proof. Please refer to the Appendix A.2. □

Shown in Fig. 10 are some examples of this metric. Each plot shows two Bessel K forms with the parameters (p_1, c_1) and (p_2, c_2), respectively. In these examples, $p_1 = 0.5$ and $c_1 = 10.0$ are held fixed, while p_2, c_2 are varied to demonstrate different cases. The distance $d(p_1, c_1, p_2, c_2)$ increases from left to right.

Theorem 1 provides a metric between two Bessel K forms or between two spectral marginals. It can be extended to a pseudometric on the image space as follows: For any two images, I_1 and I_2, and for a given bank of filters, $F^{(1)}, \ldots, F^{(K)}$, let the Bessel parameter values be given by: $(p_1^{(j)}, c_1^{(j)})$ and $(p_2^{(j)}, c_2^{(j)})$, respectively, for $j = 1, 2, \ldots, K$. Then, the L^2-distance, between the spectral representations of the two images, is defined as:

$$d_I(I_1, I_2) = \sqrt{\left(\sum_{j=1}^{K} d\left(p_1^{(j)}, c_1^{(j)}, p_2^{(j)}, c_2^{(j)}\right)^2\right)}. \tag{9}$$

Note that d_I is not a proper metric on the image space because two different images can have $d_I = 0$ between them. Also, d_I is dependent upon the choice of filters. It has been established in the literature that different spectral components of the same image are often correlated and, therefore, this Euclidean form may not be appropriate. In such cases, another choice such as the *max* of all components, $d_I(I_1, I_2) = \max_j d(p_1^{(j)}, c_1^{(j)}, p_2^{(j)}, c_2^{(j)})$, may be pursued. Since the Gabor filters are sensitive to the image orientations, the resulting metric also depends upon the orientations. To enforce rotational invariance, one can define a metric which selects the supremum over all possible image orientations. The other choice is to restrict to filters that are rotationally invariant.

4 Application of Bessel K Representations

Now, we present some examples of applying Bessel K formulations and the metric d_I to image understanding problems. We have selected examples from: 1) clutter classification, 2) target recognition, and 3) texture synthesis.

4.1 Clutter Classification

An important application of this Bessel K representation is in the classification of clutter for ATR (automated target recognition) scenarios. In particular, given an observed image of a target imaged in a cluttered environment, one would like to characterize the clutter to the extent that it improves the ATR performance. Some knowledge of the clutter type, whether it is grass, buildings, trees, or roads, can help improve the target recognition performance. In this section, we will utilize Bessel K forms to represent the image spectra, and will employ d_I as defined in (9) to classify the clutter types from their images. To illustrate the idea, we will use d_I to cluster some images of natural clutter shown in Fig. 11. For a simple illustration, let the images in the top row be the training images that are already classified, and the bottom row be images that are to be classified. Using 27 small-scale Gabor filters ($K = 27$, for nine orientations at three scales each), we have computed the pairwise distances d_Is.

Using the nearest-neighbor approach and the metric d_I, one can perform clutter classification. To illustrate a classification of clutter types, we have plotted a clustering chart in the left panel of Fig. 12 using the dendrogram function in matlab. This function generates a clustering tree for points in a high-dimensional space when their pairwise distances are given. The clustering of I_1 with I_2, I_3 with I_4, etc., demonstrates the success of this representation and the metric chosen. For comparison, we ran the same clustering

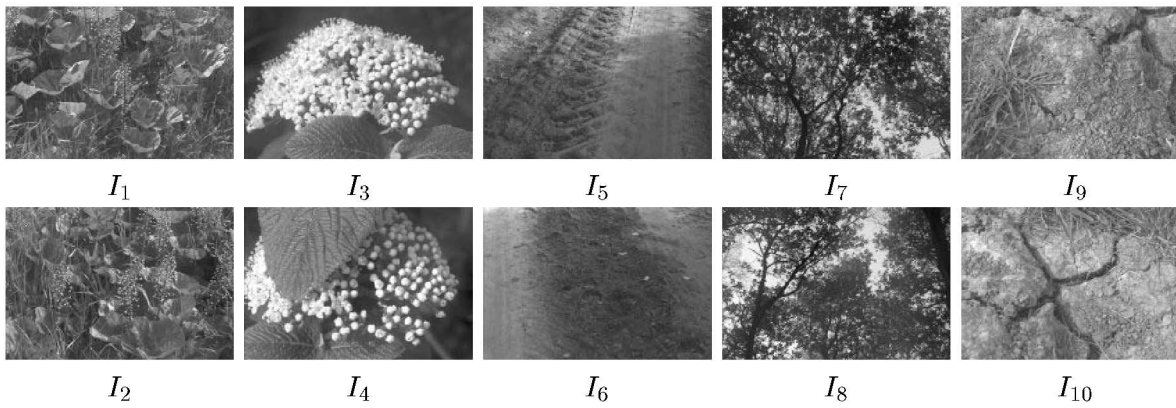


Fig. 11. Ten natural images from the van Hateren database: top row consists of the training images and bottom row consists of the test images.

program using an Euclidean metric on a principal subspace of the given images. We extracted nonoverlapping patches of size 20×30 from the original images, performed principal component analysis (PCA) in $\mathbb{R}^{600}$, and retained only the first 40 components. Images are then projected onto this linear subspace to compute the coefficients and the resulting pairwise Euclidean distances. Shown in the right panel is a dendrogram clustering based on this metric.

Shown in Fig. 13 is another example of image clustering using d_I.

4.2 Pruning Hypotheses for Target Recognition

Recognition of objects from their observed images corresponds to the selection of hypothesis in presence of the nuisance parameters [12]. As stated under Case 1 in Section 1.1, this hypothesis selection is often performed using detailed models involving physical shapes, texture, pose, and motion [26], [10], [12]. Such methods are based on low- and high-dimensional deformations of targets' templates in order to match their synthesized images with the observed images. The deformations capture the variability in pose, motion, illumination, etc., and form the set of nuisance parameters, call it S, for hypothesis selection; they typically are computationally expensive to implement. Given an image, the task of searching over all possible templates is demanding and can benefit from a pruning that places significant probability only on a small subset of possible hypotheses.

Let $\mathcal{A}$ be the set of all possible objects. Define a probability mass function on $\mathcal{A}$ according to: for $\alpha \in \mathcal{A}$,

$$P(\alpha|I) = \frac{\exp\left(-\min_{s\in S}\left(\sum_{j=1}^{K} d\left(p_{obs}^{(j)}, c_{obs}^{(j)}, p_{\alpha,s}^{(j)}, c_{\alpha,s}^{(j)}\right)^2\right)/T\right)}{\sum_{\alpha'\in\mathcal{A}} \exp\left(-\min_{s\in S}\left(\sum_{j=1}^{K} d(p_{obs}^{(j)}, c_{obs}^{(j)}, p_{\alpha',s}^{(j)}, c_{\alpha',s}^{(j)}\right)^2\right)/T\right)}, \tag{10}$$

where T controls our confidence (analogous to the temperature in Gibbs' energies) in this probability. Here, $(p_{obs}^{(j)}, c_{obs}^{(j)})$ are the estimated parameters for the image I and filter $F^{(j)}$, and $(p_{\alpha,s}^{(j)}, c_{\alpha,s}^{(j)})$ are the estimated parameters for the filter $F^{(j)}$ and the target α rendered at the nuisance variable $s \in S$. Note that $(p_{\alpha,s}^{(j)}, c_{\alpha,s}^{(j)})$ can be precomputed offline for all $\alpha \in \mathcal{A}$, $s \in S$, and $j \in \{1, 2, \ldots, K\}$.

To illustrate this idea, consider the following experiment. Shown in Fig. 14 are some sample images of objects from the Columbia object image library (COIL) [22]. This database consists of 72 images each of a total of 100 objects, taken at five degree separation in azimuth, and has been widely used in

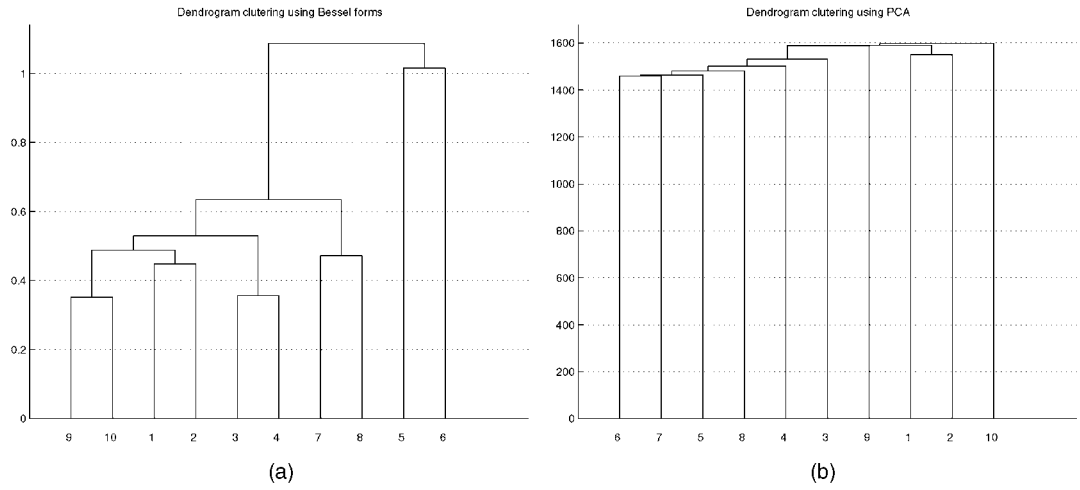


Fig. 12. Dendrogram clustering of images in Fig. 11 using d_I (a) and using a Euclidean metric on PCA (b).

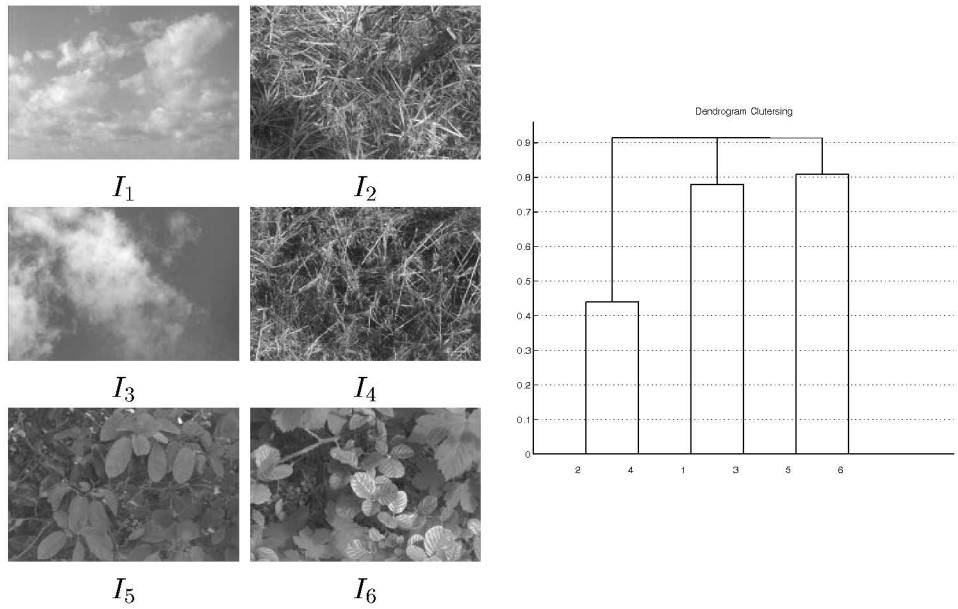


Fig. 13. For the images shown in left, a dendrogram clustering plot using Bessel K forms is shown in the right panel.

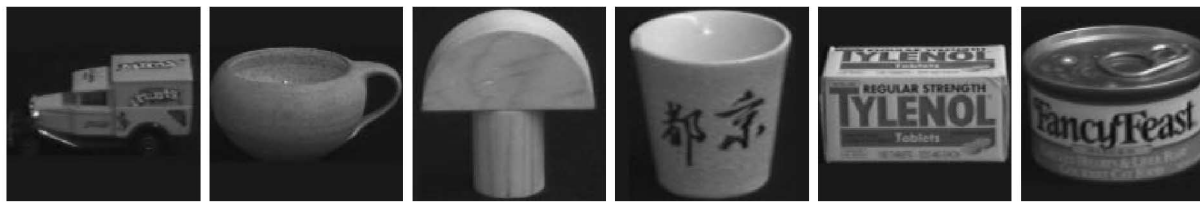


Fig. 14. Sample images of objects from COIL image database.

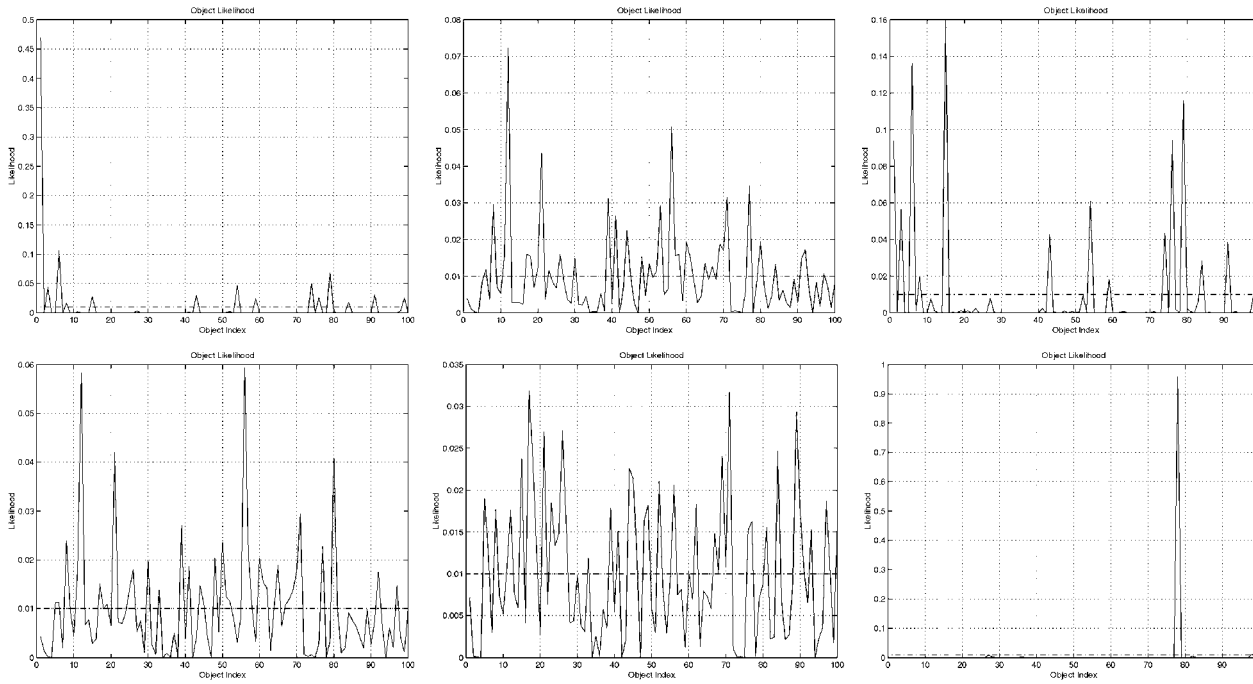


Fig. 15. Plots of $P(\alpha|I)$ versus α for six test images in the COIL database. The test images are of objects α_1, α_{12}, α_{15}, α_{12}, α_{15}, and α_{78}, respectively, for arbitrary orientations. Dotted lines suggest a threshold level for pruning.

testing object recognition algorithms. In this experiment, we divided 7,200 images into nonoverlapping training and test sets. Some of the images are used as training and the remaining for testing, similar to the work presented in [22]. We have used a bank of $K = 39$ filters, consisting of gradient filters, Laplacian of Gaussian filters, and Gabor filters. For each image of the object α at the pose s in the training set, we estimate $(p_{\alpha,s}^{(j)}, c_{\alpha,s}^{(j)})$, for each filter $F^{(j)}$. Then, given a test

TABLE 1
Correct Recognition Rate for the Full COIL-100 Data Set Using PCA, ICA, and Bessel Forms

Training/test per object	PCA	ICA	SNoW [29]	SVM [22]	Bessel Forms
36 / 36	98.58%	98.47%	95.81%	96.03%	99.89%
18 / 54	96.67%	96.52%	92.31%	91.30%	99.00%
8 / 64	87.23%	87.91%	85.13%	84.80%	92.44%
4 / 68	75.82%	76.03%	81.46%	78.50%	78.65%

image I, the estimated parameters $(p_{obs}^{(j)}, c_{obs}^{(j)})$ are used to compute the probability $P(\alpha|I)$ according to (10). Shown in Fig. 15 are six plots of $P(\alpha|I)$ versus α (at $T = 0.5$) for six different test images I in the COIL database. All the objects with probabilities larger than some threshold, say 0.01, can be shortlisted for detailed hypothesis testing. As an example, the plot in top left shows $P(\alpha|I)$ for an image I of the first object. In short-listing by thresholding, we are left with only 14 possible hypotheses, a significant reduction from 100. The bottom middle plot displays the worst case of the whole experiment and still shortlists 35 objects.

To support the use of Bessel K models in hypothesis pruning, we have actually used $P(\alpha|I)$ for object recognition and have compared results with some other recently proposed procedures: principal component analysis (PCA), independent component analysis (ICA), support vector machines (SVM), and SNoW. Pontil and Verri [22] have applied SVM (Support Vector Machines) method to 3D object recognition and have tested it on a subset of the COIL-100 data set with half for training and the other half for testing. As pointed out by Yang et al. [29], this dense sampling of training views simplifies the recognition problem. Hence, we have presented recognition results for different training to test ratios in splitting the COIL database. The number of components selected is such that complexity remains similar to that of Bessel representations. Table 1 summarizes that Bessel representations, in addition to being analytic and parametric, generally outperform these other methods.

4.3 Texture Synthesis

To further illustrate the strength of Bessel representations, we present some examples of texture synthesis. For homogeneous textures, it is possible to sufficiently characterize them using their spectral responses. As described in [31], [30], choosing the filtered marginals for several filters as sufficient statistics leads to a Gibbs' distribution on the image space. This points to a natural Gibbs' type MCMC sampling method to generate high probability images from the probability model. Using the same sampling scheme, except the observed marginals are now replaced by the **estimated Bessel K forms**, we have generated high probability samples on the image space. The Gibbs' distribution on the image is space is given by:

$$f_I(I) = \frac{1}{Z}\exp\left(-\sum_{j=1}^{K}\|H\left(I^{(j)}\right) - f\left(x; p^{(j)}, c^{(j)}\right)\|^2/T\right),$$

where H denotes the observed histogram and Z is the normalizing constant involved. Shown in the top panels of Fig. 16 are real texture images used to estimate the Bessel K parameters (for 39 filters) and shown in the bottom panels are the corresponding samples from a distribution based on the estimated parameters.

5 Conclusion

We have applied Bessel K forms to model the probability densities of the filtered marginals. The estimated parametric

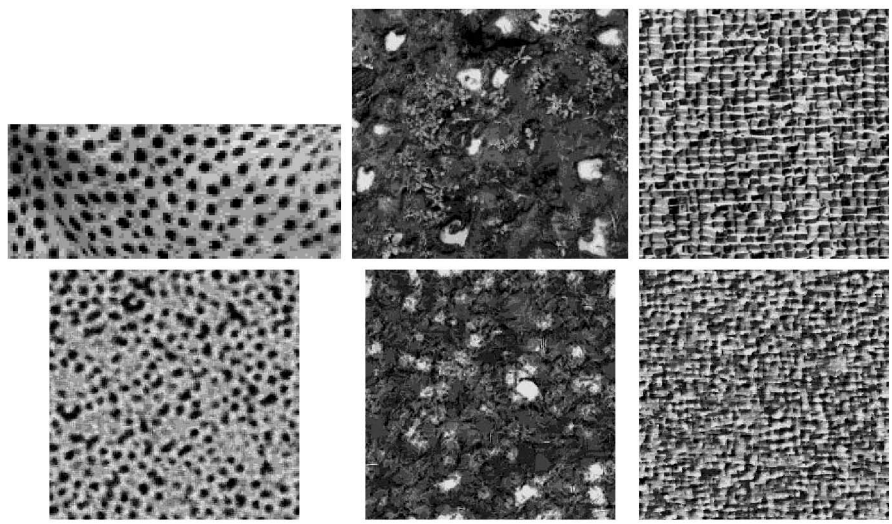


Fig. 16. Top row: observed images of the textures. Bottom row: synthesized images using the Bessel K densities.

forms are shown to match well with the observed histograms for a variety of images: video, infrared, and range, for gradient, Gabor, and Laplacian of Gaussian filters. Given the assumptions behind this construction, we expect this model to perform well in other imaging modalities such as MRI, PET, and radar imaging. Bessel parameter p is related to the distinctness and the frequency of occurrence of the filtered characteristics of imaged objects. We have used L^2 metric on the set of Bessel forms (restricted to $p > 0.25$) to derive a pseudometric on the image space. This metric can be used for, among other things, clutter classification and target recognition. Although the performance of Bessel representations in challenging object recognition situations remains to be tested, their ability to prune possible hypotheses and feed to a more detailed recognition model seems promising.

APPENDIX

A.1 Bessel Function

The modified Bessel function used in this paper is defined as:

$$K_\nu(xy) = \frac{\Gamma(\nu+0.5)(2y)^\nu}{\Gamma(0.5)x^\nu}\int_0^\infty \frac{\cos(xz)}{(z^2+y^2)^{\nu+0.5}}dz, \qquad (11)$$

for $\Re(\nu) > -0.5$, $x > 0$, and $|arg(y)| < \frac{\pi}{2}$.

A.2 Proof of Theorem 1

To establish the theorem, we will need the integral formula ([8, p. 676, (4)]:

$$\begin{aligned}&\int_0^\infty x^{-\lambda}K_\mu(ax)K_\nu(bx)dx = \\ &\frac{2^{-2-\lambda}a^{-\nu+\lambda-1}b^\nu}{\Gamma(1-\lambda)}\Gamma\left(\frac{1-\lambda+\mu+\nu}{2}\right)\Gamma\left(\frac{1-\lambda-\mu+\nu}{2}\right) \\ &\Gamma\left(\frac{1-\lambda+\mu-\nu}{2}\right)\Gamma\left(\frac{1-\lambda-\mu-\nu}{2}\right) \\ &F\left(\frac{1-\lambda+\mu+\nu}{2},\frac{1-\lambda-\mu+\nu}{2};1-\lambda;1-\frac{b^2}{a^2}\right),\end{aligned} \qquad (12)$$

where $\Re(a+b) > 0$, $\Re(\lambda) < 1 - |\Re(\mu)| - |\Re(\nu)|$. F is the hypergeometric function; it is an infinite series in its last argument.

To derive the L^2-metric, we start with its square:

$$\begin{aligned}d(p_1,c_1,p_2,c_2)^2 = &\int_x f(x;p_1,c_1)^2dx + \int_x f(x;p_2,c_2)^2dx \\ &- 2\int_x f(x;p_1,c_1)f(x;p_2,c_2)dx.\end{aligned}$$

Consider these terms one by one, starting with the first term:

$$\begin{aligned}&\frac{1}{Z(p_1,c_1)^2}\int_x x^{2p_1-1}K_{(p_1-0.5)}\left(\sqrt{\frac{2}{c_1}}x\right)K_{(p_1-0.5)}\left(\sqrt{\frac{2}{c_1}}x\right)dx \\ &= \frac{1}{Z(p_1,c_1)^2}2^{-3}(2c_1)^{p_1}\frac{\Gamma(2p_1-0.5)}{\Gamma(2p_1)}\Gamma(p_1)^2\Gamma(0.5) \\ &= \frac{\sqrt{2}}{4\pi}\Gamma(0.5)\left(\frac{1}{\sqrt{c_1}}\frac{\Gamma(2p_1-0.5)}{\Gamma(2p_1)}\right),\end{aligned}$$

using the integral formula (12). Similarly, the second term becomes

$$\frac{\sqrt{2}}{4\pi}\Gamma(0.5)\left(\frac{1}{\sqrt{c_2}}\frac{\Gamma(2p_2-0.5)}{\Gamma(2p_2)}\right).$$

Substituting for the integral in the cross term gives:

$$\begin{aligned}&\frac{1}{Z(p_1,c_1)Z(p_2,c_2)} \\ &\int_x x^{p_1+p_2-1}K_{(p_1-0.5)}\left(\sqrt{\frac{2}{c_1}}x\right)K_{(p_2-0.5)}\left(\sqrt{\frac{2}{c_2}}x\right)dx \\ &= \frac{1}{Z(p_1,c_1)Z(p_2,c_2)}2^{-3}2^{p_1}c_1^{p_2}\left(\sqrt{\frac{2}{c_1}}\right)^{-p_1+0.5}\left(\sqrt{\frac{2}{c_2}}\right)^{p_2-0.5} \\ &\quad\frac{\Gamma(p_1+p_2-0.5)}{\Gamma(p_1+p_2)}\Gamma(p_1)\Gamma(p_2)\Gamma(0.5) \\ &\quad F\left((p_1+p_2-0.5),p_2;p_1+p_2;1-\frac{c_1}{c_2}\right) \\ &= \frac{\sqrt{2}}{4\pi}\Gamma(0.5)\left(\frac{1}{\sqrt{c_1}}\left(\frac{c_1}{c_2}\right)^{p_2}\frac{\Gamma(p_1+p_2-0.5)}{\Gamma(p_1+p_2)}\mathcal{F}\right).\end{aligned}$$

Combining these three terms, the result in (8) follows: It should be noted that the metric is symmetric in the parameters (p_1,c_1) and (p_2,c_2), even though it does not appear that way from the expression in (8). The condition associated with the formula (12) implies that p_1 and p_2 must be greater than 0.25.

ACKNOWLEDGMENTS

This research was supported in part by the grants Army Research Office DAAD19-99-1-0267, NMA 201-01-2010, and US National Science Foundation DMS-0101429. The images used in the experiments are taken from the Groningen image database, the COIL database, Brown range database, and the Florida State University infrared face database. The authors are grateful to the producers of these databases for making them public. The authors also thank Professor J. Sethuraman for some useful discussions on this paper.

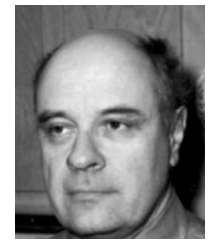
REFERENCES

[1] O. Barndorff-Nielsen, J. Kent, and M. Sorensen, "Normal Variance-Mean Mixtures and z Distributions," *Int'l Statistical Rev.*, vol. 50, pp. 145-159, 1982.
[2] P.N. Belhumeur, J.P. Hepanha, and D.J. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition Using Class Specific Linear Projection," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711-720, July 1997.
[3] A.J. Bell and T.J. Sejnowski, "The "Independent Components" of Natural Scenes are Edge Filters," *Vision Research*, vol. 37, no. 23, pp. 3327-3338, 1997.
[4] C. Chubb, J. Econopouly, and M.S. Landy, "Histogram Contrast Analysis and the Visual Segregation of iid Textures," *J. Optical Soc. Am. A*, vol. 11, pp. 2350-2374, 1994.
[5] P. Comon, "Independent Component Analysis, a New Concept?" *Signal Processing*, vol. 36, no. 3, 1994.
[6] D.L. Donoho and A.G. Flesia, "Can Recent Innovations in Harmonic Analysis 'Explain' Key Findings in Natural Image Statistics," http://www-stat.stanford.edu/donoho/Reports, 2001.
[7] D.L. Donoho and X. Huo, "Beamlets and Multiscale Image Analysis," http://www-stat.stanford.edu/donoho/Reports, 2001.
[8] I.S. Gradshteyn and I.M. Ryzhik, *Table of Integral Series and Products*, A. Jeffrey, ed., Academic Press, 2000.
[9] U. Grenander, *General Pattern Theory*. Oxford Univ. Press, 1993.

[10] U. Grenander, M.I. Miller, and A. Srivastava, "Hilbert-Schmidt Lower Bounds for Estimators on Matrix Lie Groups for ATR," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 20, no. 8, pp. 790-802, Aug. 1998.
[11] U. Grenander and A. Srivastava, "Probability Models for Clutter in Natural Images," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 4, pp. 424-429, Apr. 2001.
[12] U. Grenander, A. Srivastava, and M.I. Miller, "Asymptotic Performance Analysis of Bayesian Object Recognition" *IEEE Trans. Information Theory,* vol. 46, no. 4, pp. 1658-1666, 2000.
[13] D.J. Heeger and J.R. Bergen, "Pyramid-Based Texture Analysis," *Proc. SIGGRAPH,* pp. 229-238, 1995.
[14] B. Jahne, H. Haubecker, and P. Geibler, *Handbook of Computer Vision and Applications, Vol. 2.* Academic Press, 1999.
[15] B. Julesz, "A Theory of Preattentive Texture Discrimination Based on First-Order Statistics of Textons," *Biological Cybernetics,* vol. 41, pp. 131-138, 1962.
[16] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, no. 1, pp. 103-108, Jan. 1990.
[17] A.B. Lee and D. Mumford, "Occlusion Models for Natural Images: A Statistical Study of Scale-Invariant Dead Leaves Model," *Int'l J. Computer Vision,* vol. 41, nos. 1 and 2, 2001.
[18] D. Marr, *VISION: A Computational Investigation into the Human Representation and Processing of Visual Information.* New York: W.H. Freeman and Company, 1982.
[19] M.I. Miller, A. Srivastava, and U. Grenander, "Conditional-Expectation Estimation via Jump-Diffusion Processes in Multiple Target Tracking/Recognition," *IEEE Trans. Signal Processing,* vol. 43, no. 11, pp. 2678-2690, Nov. 1995.
[20] D. Mumford, "Empirical Investigations into the Statistics of Clutter and the Mathematical Models It Leads To," Lecture for the Review of ARO Metric Pattern Theory Collaborative, 2000.
[21] B.A. Olshausen and D.J. Field, "Sparse Coding with an Over-Complete Basis Set: A Strategy Employed by V1?" *Vision Research,* vol. 37, no. 23, pp. 3311-3325, 1997.
[22] M. Pontil and A. Verri, "Support Vector Machines for 3D Object Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 20, no. 6, pp. 637-646, June 1998.
[23] J. Portilla and E.P. Simoncelli, "A Parametric Texture Model Based on Join Statistics of Complex Wavelet Coeeficients," *Int'l J. Computer Vision,* vol. 40, no. 1, pp. 49-70, 2000.
[24] S.T. Roweis and L.K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *Science,* vol. 290 pp. 2323-2326, 2000.
[25] A. Srivastava, X. Liu, B. Thomasson, and C. Hesher, "Spectral Probability Models for Infrared Images and Their Applications to IR Face Recognition," *Proc. Workshop Computer Vision Beyond Visual Spectrum,* 2001.
[26] A. Srivastava, M.I. Miller, and U. Grenander, "Bayesian Automated Target Recognition," *Handbook of Image and Video Processing,* pp. 869-881, Academic Press, 2000.
[27] H. van Hateren, "Natural Stimuli Collection: A Public Database of Natural Images," http://hlab.phys.rug.nl. Apr. 2002.
[28] M.J. Wainwright, E.P. Simoncelli, and A.S. Willsky, "Random Cascades on Wavelet Trees and Their Use in Analyzing and Modeling Natural Images," *Applied and Computational Harmonic Analysis,* vol. 11, pp. 89-123, 2001.
[29] M.H. Yang, D. Roth, and N. Ahuja, "Learning to Recognize 3D Objects with SNoW," *Proc. Sixth European Conf. Computer Vision,* vol. 1, pp. 439-454, 2000.
[30] S.C. Zhu, X. Liu, and Y.N. Wu, "Statistics Matching and Model Pursuit by Efficient MCMC," *IEEE Trans. Pattern Recognition and Machine Intelligence,* vol. 22, pp. 554-569, 2000.
[31] S.C. Zhu, Y.N. Wu, and D. Mumford, "Minimax Entropy Principles and Its Application to Texture Modeling," *Neural Computation,* vol. 9, no. 8, pp. 1627-1660, Nov. 1997.

Anuj Srivastava received the PhD degree in electrical engineering from Washington University in 1996 and was a visiting researcher in the Division of Applied Mathematics at Brown University during the year 1996-1997. He is an assistant professor in the Department of Statistics at Florida State University (FSU), Tallahassee. He directs the Laboratory for Computational Vision at FSU that engages in research on image understanding using multiple imaging modalities. His research interests lie in the areas of statistical signal processing and image analysis, statistics on manifolds, and computational statistics. His research has been supported by the US Army Research Office, the US National Science Foundation, the National Imagery and Mapping Agency, and FSU research funds.

Xiuwen Liu received the BEng degree in computer science in 1989 from Tsinghua University, Beijing, China, the MS degrees in geodetic science and surveying in 1995 and in computer and information science in 1996, and the PhD degree in computer and information science in 1999 from The Ohio State University, Columbus. Currently, he is an assistant professor in the Department of Computer Science at the Florida State University, leading the Florida State Vision group. His current research interests include low dimensional representations of images, image classification and segmentation, statistical computer vision, neural networks, and computational models of vision.



Ulf Grenander received the PhD degree in 1950 from the University of Stockholm on a thesis dealing with statistical inference in stochastic processes. His later work includes Toeplitz forms, time-series analysis, and probabilities on algebraic structures. His main scientific contribution is pattern theory, a mathematical theory of patterns that he initiated in 1966. It has been applied to digital anatomy, analysis of micrographs, tracking of objects and laguage theory. He is the L. Herbert Ballou University Professor at Brown University.