Multiresolution Histograms and Their Use for Recognition

Efstathios Hadjidemetriou, Michael D. Grossberg, *Member, IEEE*, and
Shree K. Nayar, *Member, IEEE*

Abstract—The histogram of image intensities is used extensively for recognition and for retrieval of images and video from visual databases. A single image histogram, however, suffers from the inability to encode spatial image variation. An obvious way to extend this feature is to compute the histograms of multiple resolutions of an image to form a multiresolution histogram. The multiresolution histogram shares many desirable properties with the plain histogram including that they are both fast to compute, space efficient, invariant to rigid motions, and robust to noise. In addition, the multiresolution histogram directly encodes spatial information. We describe a simple yet novel matching algorithm based on the multiresolution histogram that uses the differences between histograms of consecutive image resolutions. We evaluate it against five widely used image features. We show that with our simple feature we achieve or exceed the performance obtained with more complicated features. Further, we show our algorithm to be the most efficient and robust.

Index Terms—Multiresolution histogram, scale-space, image sharpness, Fisher information, shape feature, texture feature, histogram matching, histogram bin width, feature parameter sensitivity, feature comparison.

✦

1 INTRODUCTION

HISTOGRAMS have been widely used to represent, analyze, and characterize images. One of the initial applications of histograms was the work of Swain and Ballard for the identification of 3D objects [73]. Following that work, various recognition systems [22], [72] based on histograms were developed. Currently, histograms are an important tool for the retrieval of images and video from visual databases [1], [52], [80], [81]. Some of the reasons for their importance are that they can be computed easily and efficiently, they achieve significant data reduction, and they are robust to noise and local image transformations. For many applications, however, the histogram is not adequate, since it does not capture spatial image information. In this work, intensity, together with spatial image information, is combined using the multiresolution histogram [27].

1.1 Contributions

The multiresolution decomposition of an image is computed with Gaussian filtering [41], [78]. The image at each resolution gives a different histogram. The multiresolution histogram, $\mathcal{H}$, is the set of intensity histograms of an image at multiple image resolutions. In this work, the multiresolution decomposition of an image is implemented with a pyramid for efficiency. The multiresolution histogram can be computed and stored efficiently also. Moreover, it can also be matched very fast using the L_1 norm. The

multiresolution histogram not only combines intensity with spatial information, but it also preserves the efficiency, simplicity, and robustness of the plain histogram.

The bottom row of Fig. 1 shows two images with identical histograms. The first and fourth columns show the multiresolution decomposition of these two images. The second and third columns show the multiresolution histograms of the two original images. In each multiresolution histogram, the histograms of corresponding lower resolutions of the two images are different. This is because the spatial information in the two original images is different. Note that the multiresolution histogram is an image representation because multiresolution decomposition is applied to the image. It is different from representations where multiresolution decomposition is applied exclusively to the histogram [13], [18].

Translations, rotations, and reflections preserve the multiresolution histogram. In general, however, the transformations of an image affect its multiresolution histogram. This effect is addressed using as analytical tools the image functionals called generalized Fisher information measures. These functionals relate histogram density values with spatial image variation. For some classes of images, it is shown that spatial image variation depends on parameters of shapes or properties of textures. Some of these shape and texel parameters are their size, elongation, boundary complexity, and placement pattern [27].

A matching feature based on the multiresolution histogram is examined. The feature uses the histogram of the original image together with the differences between histograms of consecutive image resolutions. The upper left part of Fig. 2 shows an image. Next to it are the histograms of three of its consecutive resolutions together with the two corresponding difference histograms. The feature is applied to three databases. The first is a database of synthetic images. The second database consists of Brodatz textures [8], and the third database consists of CUReT textures [16]. The performance of the multiresolution histogram also depends on the bin width and the

- E. Hadjidemetriou is with the Yale Image Processing and Analysis Group, Yale University School of Medicine, Department of Diagnostic Radiology, Brady Memorial Laboratory Room 332, 310 Cedar Street, PO Box 208042, New Haven, CT 06520-8042. E-mail: stathis@noodle.med.yale.edu.
- M.D. Grossberg and S.K. Nayar are with the Department of Computer Science, Columbia University, 1214 Amsterdam Avenue, Mailcode 0401, NY 10027-7003. E-mail: {mdog, nayar}@cs.columbia.edu.
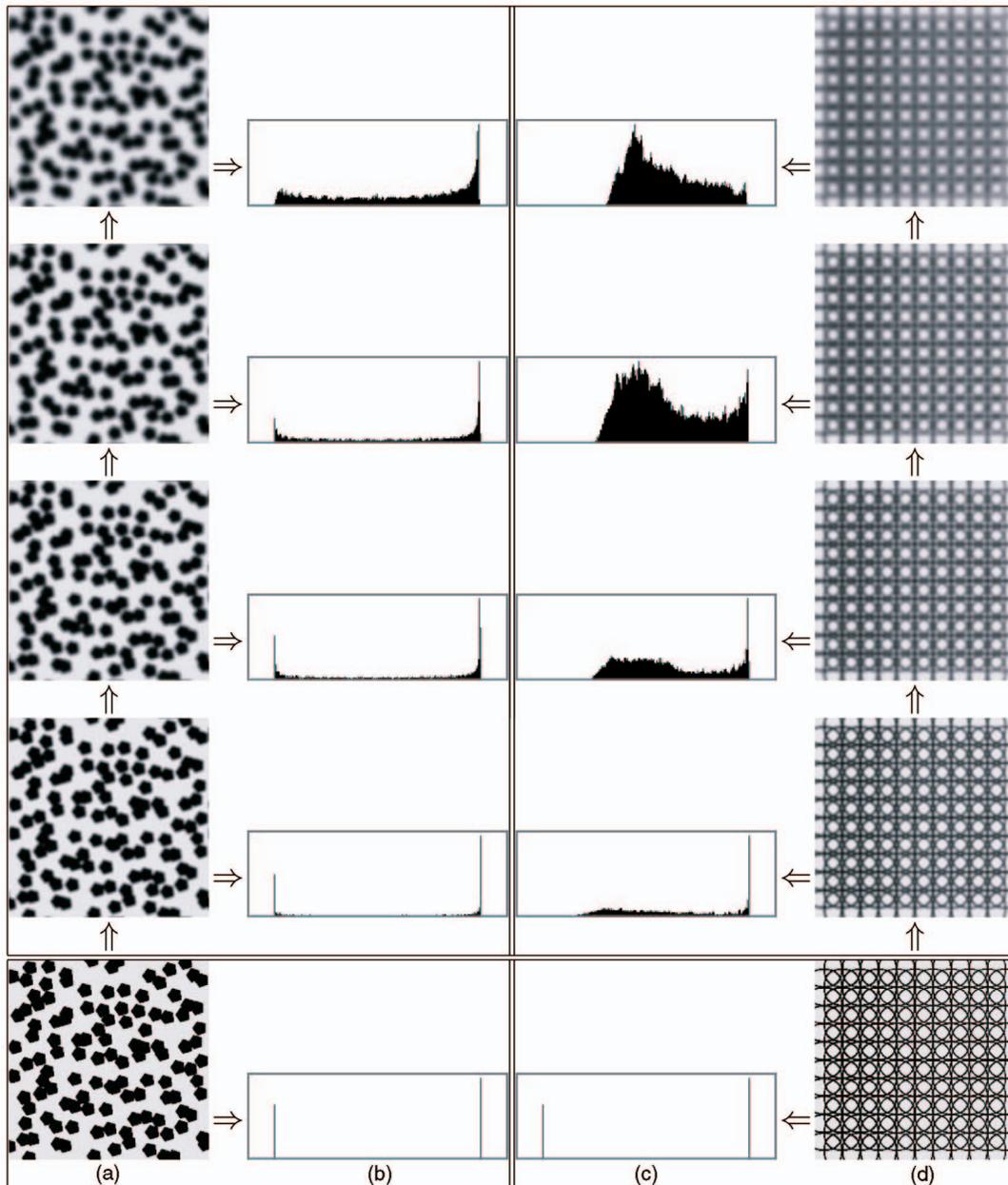
Fig. 1. Examples of two multiresolution histograms. Columns (a) and (d) show the multiresolution decomposition of two images. The bottom row shows the original images. Columns (b) and (c) show their multiresolution histograms, respectively. *The histograms of the two original images are identical*, but the two multiresolution histograms are different.

histogram smoothing. The multiresolution histograms are found to be very robust with respect to rotation, noise, intensity resolution, and database size. In general, they can discriminate between images with different spatial patterns alone without the help of any other filters or features [27].

The performance of the multiresolution histogram as an image feature is evaluated by comparing it with five commonly used image features. The five image descriptors are Fourier power spectrum features [77], Gabor wavelet features [38], Daubechies wavelet packets energies [43], autocooccurrence matrices [34], and Markov random field parameters [44], [50]. The two databases of natural textures are used in this comparative study, namely, the database of Brodatz textures and the database of CUReT textures. The

features are compared in terms of their computation cost and experimentally.

The sensitivity of the multiresolution histograms to their parameters is examined. For a fair comparison, the sensitivity of the image features to their parameters is also examined. The best performing parameters of each feature are used in the comparison. The matching performance of multiresolution histograms is comparatively robust to illumination, database size, number of classes, and pose. They were also found to be the most efficient.

1.2   Previous Work on Histogram Extensions

To discriminate between images, which have identical or similar histograms, several features have been suggested that extend plain histograms. Some algorithms have used local
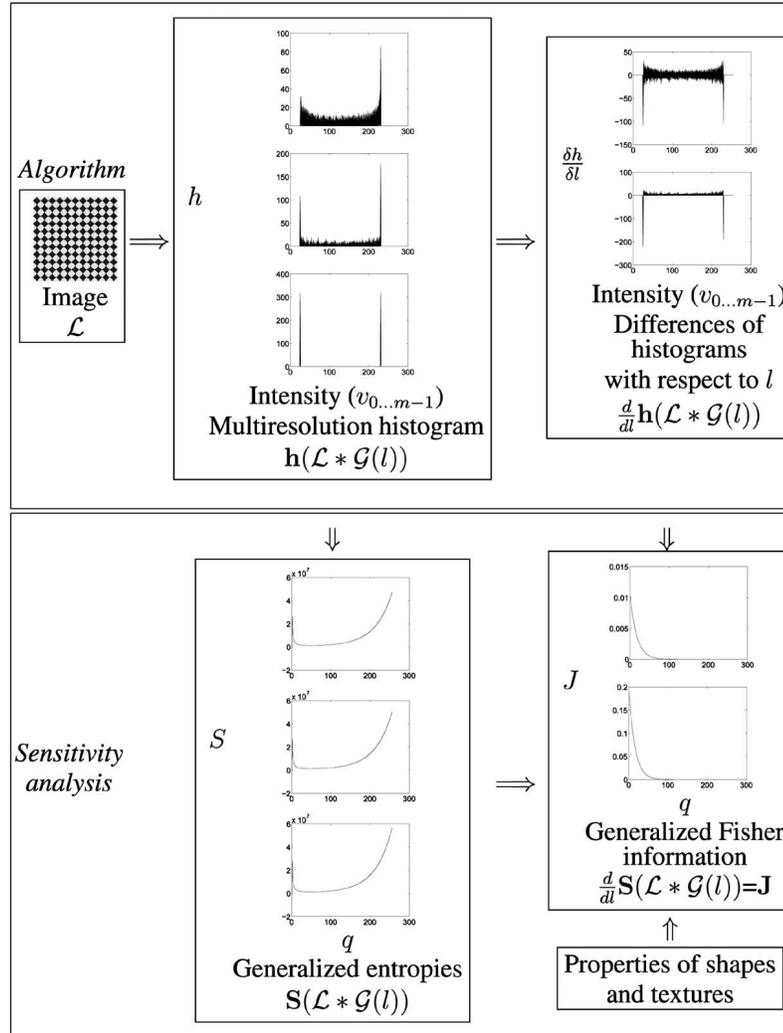
Fig. 2. The first column shows the original image. In the second column is the transformation of the multiresolution histogram into the generalized image entropies. The third column shows the transformation of the rate at which the histogram changes with image resolution into the generalized Fisher information measures. The third column demonstrates that the generalized Fisher information measures link the rate at which the histogram changes with image resolution to properties of shapes and textures. The lower part shows the quantities used only for the analysis.

intensity histograms rather than global ones. Local histograms have been combined with explicit image coordinates [12], [33], [47], [67], [71]. Another representation that combines image scale together with the histogram is the locally orderless histograms suggested by Griffin [24] as well as Koenderink and Van Doorn [42]. Sporring et al. [68] and Kadir and Brady [40] compute local histograms over regions of varying size. The local histograms are often related to the hard problem of region segmentation. Another limitation of local histograms is that they do not represent image structure.

A class of methods compute statistics of patterns of intensities. One example is the cooccurrence matrix [30], [34], [56]. Another example is the coherence vector [55] which represents region connectedness. Also, some researchers have used statistics of explicit geometric information, such as angles between neighboring line segments and ratios of neighboring line segments [19], [35].

1.3 Previous Work on Combining Histograms and Image Multiresolution

The dominant types of multiresolution decompositions have been constructed with derivative filters as well as orientation and frequency selective filters [10], [79]. Some of these filters have been differences of Gaussians [62], differences of offset Gaussians [48], [79], differences of offset differences of Gaussians [48], Gabor filters [7], [37], wavelets [14], [43], [59], and steerable filters [2], [23].

The Gabor filters combine spatial and frequency localization. An overview of their significance was given by Porat and Zeevi [60]. The wavelets achieve not only frequency and spatial localization [14], [43], [59], but also they can be implemented with critical subsampling. Thus, they are very efficient. In general, the histograms, or energies of one or more filtered images, have been used as features [37], [38], [51], [65]. For the case of Gabor filters, the feature can be the power spectrum at every pixel [21], or the maximum Gabor coefficient value of a pixel [7]. These features were used either exclusively [65], or together with the histogram of the original image [51], [52]. The features extracted from derivative filtering are sensitive to noise and even limited image deformations.

Gabor-based algorithms are sensitive to rotations, energy, and texel density. Thus, they are more suitable and

have been extensively used for texture segmentation. This has been demonstrated by Fogel and Sagi [21], as well as by Jain and Farrokhnia [37]. They have also been used for texture discrimination by Faugeras [20], and Coggins and Jain [15]. Some researchers have attempted to extract rotationally robust features from Gabor filtered images [29], [49]. The wavelet algorithms extract a feature vector for the entire image rather than for individual pixels. Thus, they are suitable and have been used for texture indexing [14], [43]. The Gaussian derivatives and steerable pyramids have also been used both for segmentation of images, image indexing, and object recognition [2], [23].

Decompositions obtained with derivative filters have been preferred because there has been a general belief that derivative multiresolution decompositions capture spatial information as opposed to Gaussian multiresolution decompositions which only introduce erroneous bias [2], [62]. This has prevented histograms of exclusively Gaussian multiresolution decompositions from being fully exploited. One application has been to compute individual histograms of low image resolutions to expedite retrieval [45]. Single histograms, however, of the original image, or lower resolutions, suffer from the inability to encode spatial image information. The histograms of multiple image resolutions have been used sequentially for texture synthesis by De Bonet [6] and Heeger and Bergen [31]. Another representation that employs Gaussian filtering has been the flexible histograms [5], [6]. Finally, Sablak and Boult [64] have implemented image multiresolution decomposition optically.

Several researchers have realized that the histogram bin width affects matching performance. A suggestion to ameliorate the binning problem has been to form clusters of bins [46], [61]. Cluster formation and the computation of their distance [28], [63], however, are computationally expensive. Thus, they deprive the histogram from its main advantages, efficiency and simplicity. Gaussian multiresolution scheme has also been combined with Markov random fields [44], [50]. Multiresolution Markov random field parameters have been used for texture discrimination [44].

2 BACKGROUND AND DEVELOPMENT OF ANALYTICAL TOOLS

Spatial image information is related robustly to weighted averages of the rates of change of histogram densities. The weighted averages are the Fisher information measures. The analysis starts with the lemma that the histogram can be transformed into a vector of generalized image entropies [69]. Generalized entropies are robust image features amenable to analysis. The rates at which the histogram bins change with image resolution can be transformed into the rates at which generalized entropies change with image resolution. The change of the generalized entropies with image resolution are given by the generalized Fisher information measures. In this analysis, the domains $\mathbf{D}$ of images $\mathcal{L}$ are taken to be continuous with coordinates $\mathbf{x} = (x, y)$. The domain is also assumed to be infinite [57].

2.1 Relation between Histogram and Tsallis Entropies of an Image

The Tsallis generalized entropies of an image $\mathcal{L}$ depend on a continuous parameter q and are given by

$$S_q = \int_{\mathbf{D}} \frac{\mathcal{L}(\mathbf{x}) - \mathcal{L}^q(\mathbf{x})}{q - 1} d^2 x, \qquad (1)$$

where image $\mathcal{L}$ has unit L_1 norm and $\mathcal{L}(\mathbf{x})$ is the intensity value at pixel $\mathbf{x}$. In the limit $q \to 1$ the Tsallis generalized entropies reduce to the Shannon entropy. In (1), the intensities at all points $\mathbf{x}$, denoted by $\mathcal{L}(\mathbf{x})$, can be substituted directly by their values: $v_0, v_1, \ldots, v_{m-1}$, where m in the total number of gray levels. The union of all the regions in the domain with identical intensity, v_j, gives the value of histogram density j, h_j. That is, (1) becomes [76]

$$S_q = \sum_{j=0}^{m-1} \left(\frac{v_j - v_j^q}{q - 1} \right) h_j. \qquad (2)$$

That is, the Tsallis generalized entropies can be expressed as a linear function of the histogram. Note that the transformation of the histogram into a generalized entropy may be considerably more complicated than linear. For example, the transformation of the histogram into a Renyi entropy [69] is logarithmic.

Consider a vector $\mathbf{S} = [S_{q_0} S_{q_1} S_{q_2} \ldots S_{q_{m-1}}]^{\mathrm{T}}$ consisting of any m different Tsallis entropies. Each element of this vector, using (2), can be expressed in terms of values of histogram densities. That is, vector $\mathbf{S}$ can be expressed as a function of the histogram $\mathbf{h} = [h_0 h_1 h_2 \ldots h_{m-1}]^{\mathrm{T}}$ to give the linear proportionality relation

$$\mathbf{S}(\mathcal{L}) \propto \mathbf{h}(\mathcal{L}). \qquad (3)$$

The algebra involved in obtaining relation (3), which is the matrix form of (2), is provided in the appendix [69]. The second column of Fig. 2 shows the linear transformation of the three histograms to the three corresponding vectors of generalized entropies. The histogram is a function where the domain is the intensity range or index of a bin. We replace the functional dependence on a particular bin with a variable q. The function at a value of q, for example, $q = 1$, has an aggregate statistical meaning, namely, the Shannon entropy.

2.2 Relation between Multiresolution Histogram and Generalized Fisher Information Measures

To decrease image resolution, we use a Gaussian filter $\mathcal{G}(l)$,

$$\mathcal{G}(l) = \frac{1}{2\pi l \sigma^2} exp\left(-\frac{x^2 + y^2}{2l\sigma^2} \right), \qquad (4)$$

where σ is the standard deviation of the filter [41], [78], and l is the resolution. A filtered image, $\mathcal{L} * \mathcal{G}(l)$, has histogram $\mathbf{h}(\mathcal{L} * \mathcal{G}(l))$ and entropy vector $\mathbf{S}(\mathcal{L} * \mathcal{G}(l))$. The rate at which the histogram changes with image resolution can be related to the rate at which image entropies change with image resolution. This relation is obtained by differentiating (2) with respect to l to obtain

$$\frac{dS_q(\mathcal{L} * \mathcal{G}(l))}{dl} = \sum_{j=0}^{m-1} \left(\frac{v_j - v_j^q}{q - 1} \right) \frac{dh_j(\mathcal{L} * \mathcal{G}(l))}{dl}. \qquad (5)$$

The rate at which the Tsallis generalized entropies change with image resolution, l, on the left hand side of (5), are related to closed form functionals of the image. These functionals are the generalized Fisher information measures [4], [58], [70],

$$J_q(\mathcal{L}) = \frac{\sigma^2}{2} \frac{dS_q(\mathcal{L} * \mathcal{G}(l))}{dl}. \qquad (6)$$

We were unable to find, in the literature, such closed form expressions for the rate of change of other families of generalized entropies with image resolution.

The substitution of (5) into the right-hand side of (6) gives

$$J_q(\mathcal{L}) = \frac{\sigma^2}{2} \sum_{j=0}^{m-1} \left(\frac{v_j - v_j{}^q}{q-1} \right) \frac{dh_j(\mathcal{L} * \mathcal{G}(l))}{dl}. \tag{7}$$

This equation reveals that J_q is linearly proportional to the rate at which the histogram densities change with image resolution. The proportionality factors, in (7), for $q > 1$ weigh heavier the rate of change of the histogram densities that have large intensity values and vice-versa [58], [76]. The proportionality factors of the Fisher information, J_1, weigh approximately equally all histogram densities. The rate of change of the histogram with image resolution can be transformed into the $m \times 1$ vector $\mathbf{J} = [J_{q_0} J_{q_1} J_{q_2} \ldots J_{q_{m-1}}]^{\mathrm{T}}$. Differentiating (3) with respect to resolution l and subsequently combining with (6) gives

$$\mathbf{J}(\mathcal{L}) \propto \frac{\sigma^2}{2} \frac{d\mathbf{h}(\mathcal{L} * \mathcal{G}(l))}{dl}. \tag{8}$$

The box in the first row and third column of Fig. 2 shows the differences between the histograms of consecutive image resolutions. The third column of Fig. 2 shows the transformation of the rate of histogram change with image resolution to the generalized Fisher information measures.

The generalized Fisher information measures, J_q, of an image with unit L_1 norm can also be computed directly from the image using [4], [58], [70]

$$J_q(\mathcal{L}) = \int_{\mathbf{D}} \left| \frac{\nabla \mathcal{L}(\mathbf{x})}{\mathcal{L}(\mathbf{x})} \right|^2 \mathcal{L}^q(\mathbf{x}) d^2 x. \tag{9}$$

The *sharpness*, or spatial variation, at a pixel is defined as

$$\left| \frac{\nabla \mathcal{L}(\mathbf{x})}{\mathcal{L}(\mathbf{x})} \right|^2.$$

The generalized Fisher information measures are nonlinear weighted averages of image *sharpness*. The average sharpness as can be seen from (9) is J_1, namely, the Fisher information [70]. The Fisher information is monotonically decreasing with Gaussian filtering. For a fixed value of variance, it achieves its minimum for a Gaussian image [4], [58], [70]. The third column of Fig. 2 shows that $\mathbf{J}$ relates the differences between histograms of consecutive image resolutions to image properties. This relation will be investigated in the following two sections. The component of vector $\mathbf{J}$ that will be used to a larger extent is the Fisher information, J_1.

3 MULTIRESOLUTION HISTOGRAMS OF SHAPES

To analyze the effect of shape parameters on the multiresolution histogram, we use J_q. The functional J_q is convex. Its single minimum is achieved for a radially symmetric Gaussian image [11], [58], [76]. As an image diverges from a Gaussian, its J_q values increase. Several classes of transformations and warps can deviate an image from a Gaussian. The value of J_q and its sensitivity to some of these classes of transformations and for some classes of images will be quantified. The value of J_q is preserved by translations, rotations, and reflections. These transformations commute with Gaussian filtering.

First, the effect of shape elongation is examined analytically for four classes of images. The histogram of an elongated shape of the classes examined changes faster with resolution than that of a radially symmetric shape. Stretching can elongate radially symmetric shapes and is given by matrix

$$\begin{pmatrix} \sqrt{\rho} & 0 \\ 0 & 1/\sqrt{\rho} \end{pmatrix},$$

where ρ is the elongation. That is, this transformation is the mapping $x \to x\sqrt{\rho}$ and $y \to \frac{y}{\sqrt{\rho}}$. The determinant of the transformation is equal to unity. Hence, this family of transformations does not affect the histogram of an image [26]. The relation of J_q to elongation ρ is quantified for four families of shapes. The images in Fig. 3 show one instance from each shape family examined. For these shapes ρ is the ratio of the parameters along the axes and k is the product of the parameters along the axes. For example, for the Gaussian aligned with the axes of the image in Fig. 3b $\rho = \frac{\sigma_x}{\sigma_y}$ and $k = \sigma_x \sigma_y$, where σ_x and σ_y are the standard deviations along the two image axes. The expressions for J_q are computed using (9) [25]. For the four image classes in Fig. 3, the minimum value of J_q happens to correspond to a symmetric shape for which $\rho = 1$ and J_q increases in proportion to $(\rho + \frac{1}{\rho})$.

The value of J_q, for a class of superquadric images, depends on how complicated the boundary is. The value of J_q for this class is larger when the boundary is complicated. The class of superquadric images examined numerically is given by $\mathcal{L} = (R^\eta - x^\eta - y^\eta)^{0.15}$. The shape of the boundary depends on parameter η, which is varied. The base area of the shapes is kept fixed as n varies and the intensity within the boundary for this family of shapes is almost flat because of the value of the exponent, 0.15. Thus, these images have approximately the same histogram. Some members of the family of shapes are shown in Figs. 4a, 4b, 4c, 4d, and 4e: a pinched diamond for $\eta < 1$, a diamond for $\eta = 1$, a circle for $\eta = 2$, and a square with curved corners for $\eta > 2$.

The plots of J_1, J_2, and J_3 per pixel as functions of η are shown in Figs. 4f, 4g, and 4h, respectively. They are computed directly from the images by discretizing the integral of (9). The minimum of J_1, J_2, and J_3 correspond to a circle shown in Fig. 4d for which $\eta = 2$. The values of J_1, J_2, and J_3 increase rapidly as the shape varies from a circle to a pinched diamond as η decreases.

The value of J_q for the family of circular diffuse shapes also varies depending on the diffuseness of the intensities within the shape. This dependence is examined analytically. For this class of images, smooth intensity changes across their boundaries minimize the rate of change of the histogram with resolution. A member of the family of circular diffuse shapes for $\epsilon = 0.7$ is given in Fig. 3c. The same figure gives the expression of the generalized Fisher information measures of the family [25]. The diffuseness is varied with exponent ϵ, but the base area is kept constant. The change in the diffuseness with increasing ϵ, shown in Figs. 5a, 5b, 5c, 5d, and 5e, gives a step transition for $\epsilon = 0$, a hemisphere for $\epsilon = 0.5$, a paraboloid for $\epsilon = 1.0$, nearly a Gaussian for $\epsilon > 1.0$, and tends to an impulse as ϵ increases further.

The Fisher information as a function of the diffuseness ϵ is plotted in Fig. 5f. The shape that has minimum Fisher
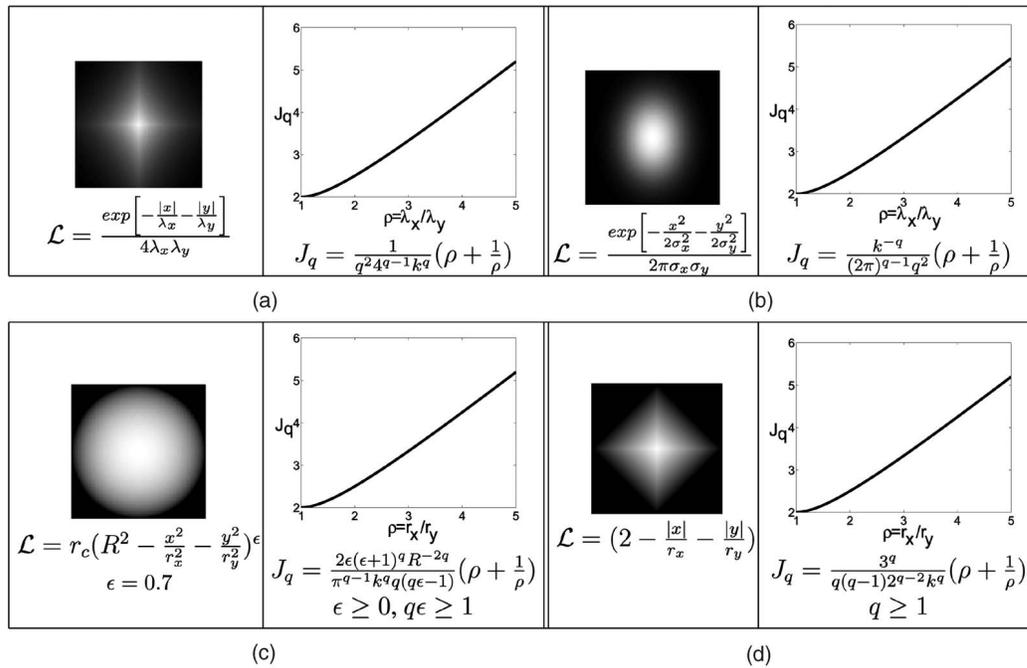
Fig. 3. The values of J_q for four image classes $\mathcal{L}$ as a function of elongation ρ, where ρ is the ratio of the shape parameters along the axes. The product k of the shape parameters and q are fixed. An image $\mathcal{L}$ from each family is shown together with the plot. For these image classes, the minimum value of J_q corresponds to radially symmetric or regular shapes for which $\rho = 1$ and increases in proportion to $\left(\rho + \frac{1}{\rho}\right)$. (a) Laplace. (b) Gaussian. (c) Circular diffuse shape. (d) Pyramid.
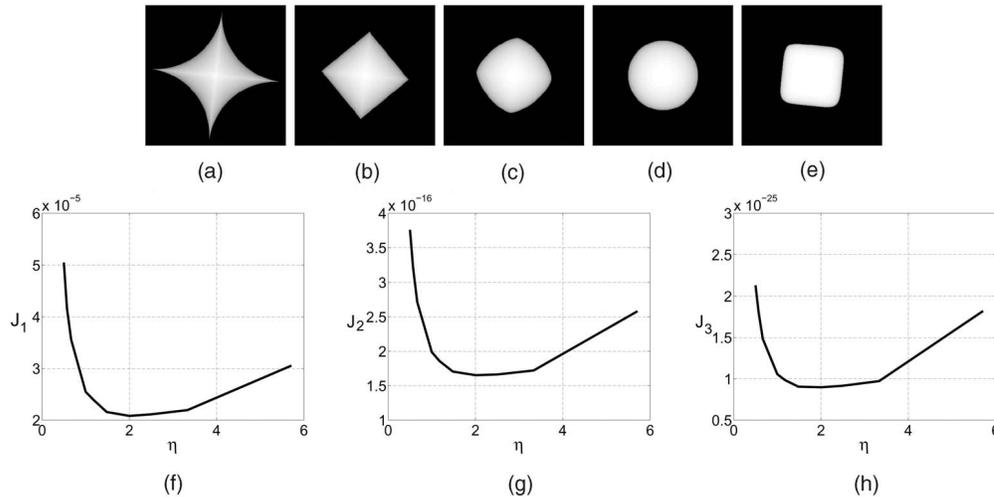
Fig. 4. The plots of J_1, J_2, and J_3 per pixel for the superquadric family of shapes $\mathcal{L} = (R^\eta - x^\eta - y^\eta)^{0.15}$ as a function of parameter η. Some members of this family of shapes are a pinched diamond, a diamond, a circle, and a square with curved corners. The plots are in (f), (g), and (h). The minima are attained for the circle shown in (d) for which $\eta = 2$. The largest values of J_1, J_2, and J_3 are attained for the pinched diamond in (a). (a) $\eta = 0.56$. (b) $\eta = 1.00$. (c) $\eta = 1.48$. (d) $\eta = 2.00$. (e) $\eta = 6.67$. (f) J_1 versus η. (g) J_2 versus η. (h) J_3 versus η.

information over the possible diffuseness patterns is that shown in Fig. 5d [25]. The diffuseness of this shape is similar to that of a Gaussian. The Fisher information increases rapidly as the shape changes from nearly Gaussian to a step boundary as ϵ decreases. For the example classes of images in this section, the histogram change with image resolution is minimized for shapes with rounded boundaries and smooth intensity transitions across the boundary.

4   MULTIRESOLUTION HISTOGRAMS OF TEXTURES

Several investigators have observed experimentally that the increase in the entropy of an image with filtering depends

on properties of image regions and image textures [36], [54], [69], [74]. This section quantifies the dependence of the multiresolution histogram on texture parameters using J_q for three different texture properties. Two example classes of images are used. The first example consist of texels which are Gaussian distributions and the second example consists of texels which are pinched diamonds such as that shown in Fig. 4a.

The dependence of J_q on the number of texels within a texture segment of fixed area is first investigated analytically. The value of J_q is expected to increase with the number of texels in a fixed area, since smaller texels are also *sharper*. Consider a simple texture model where the texel models are the shapes
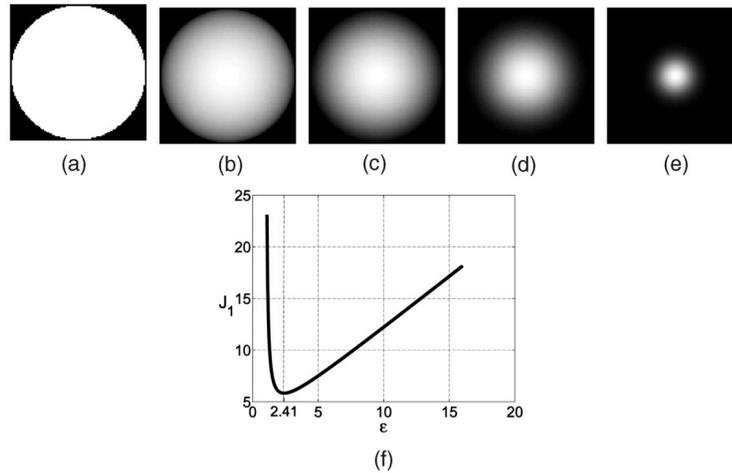
Fig. 5. The Fisher information, J_1 for a family of shapes of varying diffuseness. An image of this family for $\epsilon = 0.7$ is given in Fig. 3c. Some members of the family are a step transition, a hemisphere, a paraboloid, nearly a Gaussian, and an impulse. In (f) is the Fisher information J_1 as a function of ϵ computed analytically [25]. The minimum corresponds to the nearly Gaussian image in (d) and is largest for the step transition in (a). (a) $\epsilon = 10^{-4}$. (b) $\epsilon = 0.5$. (c) $\epsilon = 1.0$. (d) $\epsilon = 2.41$. (e) $\epsilon = 9$. (f) J_1 versus ϵ.
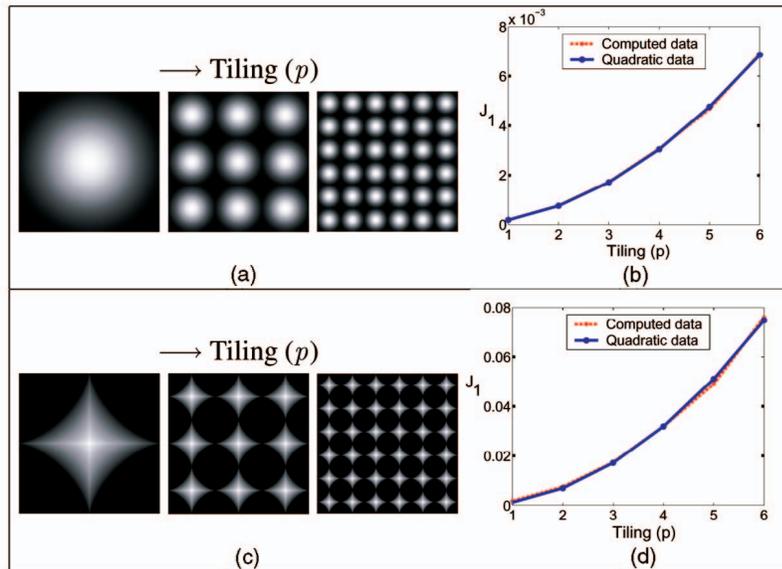
Fig. 6. The Fisher information as a function of the tiling parameter p of the textures. The images in (a) and (c) show two shapes as well as the textures resulting by minifying and tiling them. Next to the images, in (b) and (d), are the plots of the Fisher informations as a function of tiling parameter p. Each plot shows the data obtained directly from the images as well as the quadratic fit.

discussed in the previous section. The shapes are repeated to form a regular $p \times p$ pattern of identical texels. That is, the texture results by tiling a texel $r = p^2$ times. To preserve the size of the texture, the texels are also contracted by a transformation A whose determinant is given by the inverse of the repetition factor, that is $det A = 1/r$. The factor by which the area changes for textures with different parameter p is equal to unity since $(det A)r = 1$. Thus, textures for all p have the same histogram [26]. It can be shown that [25]:

$$J_q(\mathcal{L}_{texture}) = p^2 J_q(\mathcal{L}_{shape}), \tag{10}$$

where $\mathcal{L}_{shape}$ is the image of the original shape-texel, and $\mathcal{L}_{texture}$ is the image of the corresponding $p \times p$ regular texture. Equation (8) shows that the histogram change with resolution is multiplied by the same factor, p^2.

Figs. 6a and 6c show two shapes as well as the textures formed by contracting and tiling them. The texels in Fig. 6a

are Gaussians and the texels in Fig. 6c are superquadrics. Figs. 6b and 6d show their Fisher informations, respectively, as a function of p. Each of the plots shows the Fisher information computed directly from the images as well as the quadratic fit expected from (10). For both examples, the quadratic fit almost perfectly agrees with the data.

Subsequently, *the effect of overlap between neighboring texels on the Fisher information is discussed*. Textures with overlapping texels are less *sharp* and have a smaller Fisher information. Analytically, Gaussian filtering monotonically increases the size of texels and decreases their Fisher information [3], [39]. Fig. 7a shows a texture consisting of a mixture of Gaussians of linearly increasing standard deviation. Fig. 7b shows that the Fisher information monotonically decreases with texel width, which in this case is the standard deviation of the Gaussians. In Fig. 7, the standard deviation is shown as a percentage of the texel
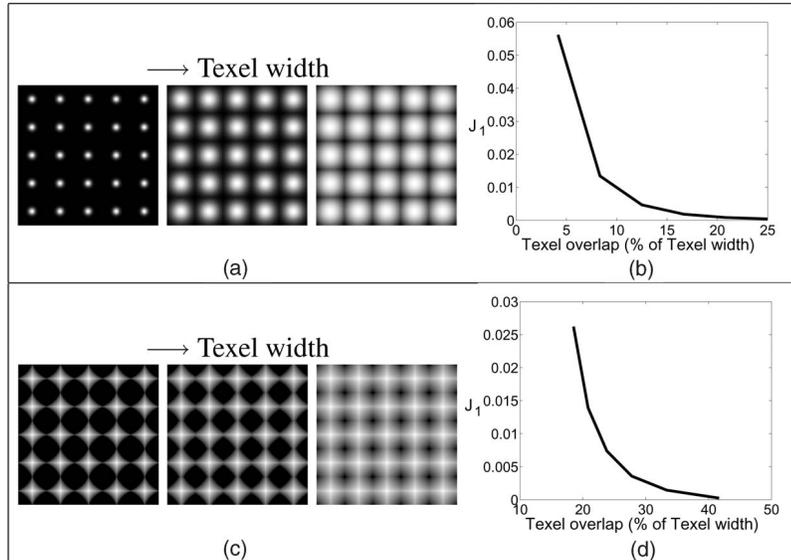
Fig. 7. The Fisher information as a function of the overlap between neighboring texels. The images in (a) and (c) show two textures with texels of increasingly larger width. Next to the images, in (b) and (d), respectively, are the plots of the Fisher informations as a function of the overlap between neighboring texels. The overlap is shown as a percentage of the width of the texels.
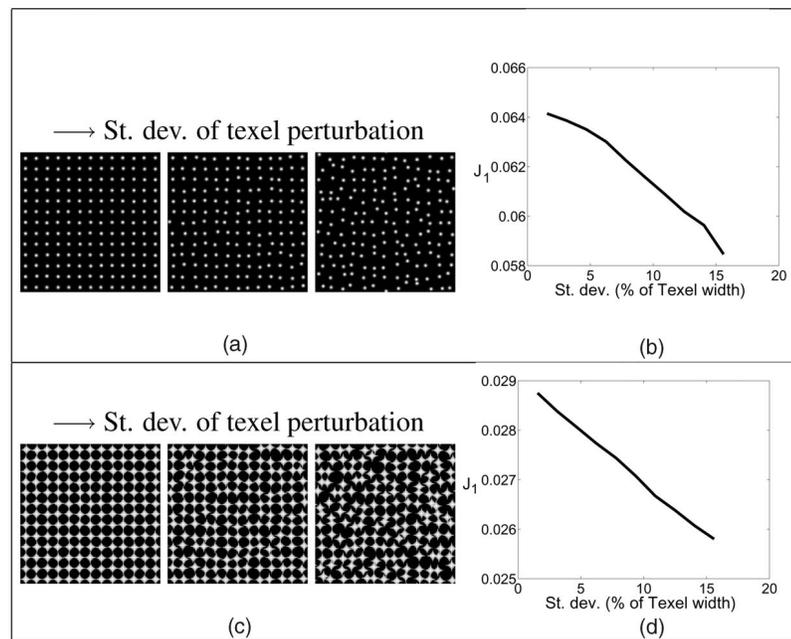
Fig. 8. The Fisher information as a function of the randomness in the placement of the texels. The images in (a) and (c) show two textures with increasingly larger randomness in the placement of the texels. Next to the images, in (b) and (d), respectively, are the plots of the Fisher informations as a function of the standard deviation of the Gaussian noise in texel placement. The standard deviation is shown as a percentage of the texel width.

width, which is 50 pixels. It is verified that the Fisher information is monotonically decreasing. Fig. 7c shows a texture with superquadric texels of linearly increasing width. Their Fisher information is shown in Fig. 7d. The overlap is shown as a percentage of the texel width. The width of the texels is also 50 pixels. Again, the Fisher information monotonically decreases with texel width.

In the remainder of this section textures whose texel placement is random are examined. Randomness, on average, monotonically decreases Fisher information [3], [11], [39]. This is verified experimentally for two classes of textures. The positions of the texels in a regular texture are perturbed with Gaussian noise of linearly increasing standard deviation. The Fisher

information is measured as a function of the standard deviation of the perturbation noise. This experiment is performed 20 times and the average values of the Fisher information measures are computed. The results of the experiments are shown in Fig. 8. The texels in Fig. 8a are Gaussian and the texels in Fig. 8c are superquadric. Figs. 8b and 8d show the average Fisher information as a function of the standard deviation of the Gaussian noise in pixel placement. The standard deviation is shown as a percentage of the texel width in the perturbed texture. The width of the texels is 19 pixels. In both cases, the Fisher information decreases monotonically as expected.

The rates at which histogram densities change with resolution increase linearly with the number of texels for the two classes of images examined. Overlap among neighboring texels and randomness in the texel placement decreases the histogram change with image resolution.

5 MATCHING ALGORITHM USING MULTIRESOLUTION HISTOGRAMS AND ITS COMPLEXITY

Multiresolution histograms are sensitive to image structure. Thus, they are used for image matching. The multiresolution is implemented with an 8 bit/pixel Burt-Adelson pyramid [9]. The multiresolution histogram is the vector $\mathcal{H} = [\mathbf{h}_0, \mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_{l-1}]$, where $\mathbf{h}_i$ is a row vector corresponding to the histogram of resolution i, and $l - 1$ is the lowest image resolution. The histograms are normalized to unit L_1 magnitude to become independent of the size of the original image and the subsampling factor of the pyramid.

The next step is to compute the cumulative histograms corresponding to each image resolution. Subsequently, the differences between the cumulative histograms of consecutive image resolutions are computed. A pyramid with l levels has $l - 1$ difference histograms. The difference histograms are proportional to the discrete versions of the Fisher information measures. The difference histograms are concatenated to form a feature vector. The distance between the feature vectors is the L_1 norm.

The bin width of the histograms affects their matching performance. The appropriate bin width for the histogram of an image depends mainly on the number of pixels in the image and on the standard deviation of the histogram [66], [75]. In turn, the standard deviation of the histogram depends on the dynamic range of the histogram as well as the pyramid level. More precisely, the optimal bin width for L_1 norm for a Gaussian or nearly Gaussian histogram is given by [17], [75]

$$w(\mathbf{h}) = (8\pi)^{1/6} \hat{\sigma}(\mathbf{h}) n^{-1/3}, \qquad (11)$$

where $w(\mathbf{h})$ is the bin width of histogram $\mathbf{h}$, and $\hat{\sigma}(\mathbf{h})$ is an estimate of the standard deviation of the histogram $\mathbf{h}$. The true underlying histogram density is very rarely Gaussian in this work. Thus, (11) gives only the order of the histogram bin width [66] and the ratio for the bin widths between histograms of images of consecutive pyramid levels. The latter ratio is the subsampling factor of the intensity resolution of the histograms in a multiresolution histogram. It is obtained by combining the relation $n_{i+1} = n_i/4$ resulting from the image pyramid together with (11) to get

$$\frac{w(\mathbf{h}_{i+1})}{w(\mathbf{h}_i)} = \frac{\hat{\sigma}(\mathbf{h}_{i+1})}{\hat{\sigma}(\mathbf{h}_i)} 2^{2/3} \overset{1}{\leq} 2^{2/3}, \qquad (12)$$

where $w(\mathbf{h}_i)$ is the bin width of the histogram of pyramid level i, $\hat{\sigma}(\mathbf{h}_i)$ is the standard deviation of the histogram of pyramid level i, and pyramid level $i + 1$ is of lower resolution than pyramid level i. If the standard deviation of the histogram monotonically decreases, then $\frac{\hat{\sigma}(\mathbf{h}_{i+1})}{\hat{\sigma}(\mathbf{h}_i)} \leq 1$. The latter leads to Step 1 of (12). Increasing the bin width of a histogram decreases the length of the histogram. To normalize with respect to the length of the histograms, the amplitudes of the histograms are multiplied by their intensity subsampling factors.

Prior to decreasing the intensity resolution of a histogram it is necessary to low-pass filter it to prevent aliasing. The additional filtering is done prior to decreasing the intensity resolution and also prior to computing the cumulative histograms or difference histograms. It may even be necessary to low-pass filter the histogram of the original image. This is because the minimum bin width, or maximum frequency, that an image can support in its histogram may be larger than the finest possible intensity change of the image. The low-pass filter used in this work is Gaussian. Filtering the histogram of an image with a Gaussian is equivalent to adding uncorrelated Gaussian noise directly to the image [32].

The cost of computing the Burt-Adelson pyramid is of order $O(n\lambda)$, where n is the number of pixels, and λ is the width of the separable Gaussian filter. The cost of computing the histograms is of order $O(n)$. The costs of filtering, computing the cumulative histograms, and the difference histograms are of order $O(tl)$, where t is the number of intensity quantization levels. That is, the total cost of computing the multiresolution histogram feature vectors is of order $O(n\lambda)$. The distance computation and storage cost are of order $O(t(l-1))$. The number of pyramid levels is given by $l = log_2 \sqrt{n}$, where n is the number of pixels in the image. Both the cost of computing the multiresolution histogram and the cost of computing the distance between multiresolution histograms is low.

6 MATCHING EXPERIMENTS USING MULTIRESOLUTION HISTOGRAMS

The matching performance of the multiresolution histogram was tested extensively with three databases; namely, a database of synthetic images, a database of Brodatz textures [8], and a database of CUReT textures [16]. The largest of these databases is the database of CUReT textures which consists of a total of 8,046 images of 61 different physical textures.

The experiments were performed with three databases; namely, a database of 108 synthetic images, a database of 91 Brodatz textures [8], and a database of 8,046 CUReT textures [16]. The last two databases contain natural textures and consist of several image classes. Each class has images of different instances of the same texture. All images have intensity resolution of 8 bit/pixel.

All the images of the database of synthetic images are bivalue and have the same histogram. Thus, they cannot be matched based on their histograms. *The databases of natural textures were histogram equalized to cancel variations due to illumination.* Thus, again the images of these databases cannot be discriminated based on their histograms. They can be matched, however, based on their multiresolution histograms. The differences between the multiresolution histograms of the various images in the database is caused simply because of differences in the shape and texture of the images.

The histograms of the original images of the databases of natural textures were equalized. The histograms of lower resolutions of these images form a single distribution that is nearly Gaussian. Thus, (11) applies and (12) can give the ratio of histogram bin widths between consecutive image resolutions. Two values based on (11) were used in the experiments. One is the maximum subsampling factor $\frac{w(\mathbf{h}_{i+1})}{w(\mathbf{h}_i)} = 2^{2/3} = 1.59$ and the other is $\frac{w(\mathbf{h}_{i+1})}{w(\mathbf{h}_i)} = 2^{1/2} = 1.41$. Equation (11) does not apply to the database of synthetic textures since the

(1) Add Gaussian noise to the image

(2) Construct image pyramid

(3) Compute histograms

(4) Normalize histograms with L_1 norm

(5) Compute cumulative histograms

(6) Compute difference histograms

(7) Subsample difference histograms and renormalize

(8) Concatenate the difference histograms to form the feature vector

(9) Compare using L_1 norm
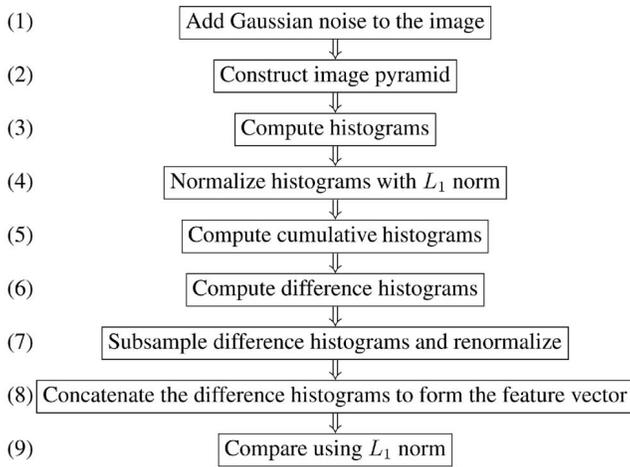
Fig. 9. The steps of the matching algorithm. Bypassing Step 1 avoids smoothing of the database images.

histograms of the bivalue images of the database consist of two spikes and not a single nearly Gaussian distribution.

The experiments examined both matching without noise and with noise. To corrupt an image, Gaussian noise of standard deviation 8 bit/pixel was superimposed. In the experiments under noise the sensitivity with respect to Gaussian noise in the test image and the database images were both examined. This is equivalent to testing the sensitivity with respect to smoothing of the histogram with a Gaussian filter. The effective standard deviation of the image noise was different from that of the original superimposed noise due to clipping within the finite intensity range of the image. The matching algorithm is summarized in Fig. 9. Note that the algorithm has two normalization steps, Step 4 and Step 7, to account for image and histogram subsampling, respectively.

6.1 Database of Synthetic Textures

Many of the images in the database consist of texel shapes and texel placements that were explored in previous sections. More precisely, the texels include dots, circles, triangles, squares, and superquadrics. The placement of the texels is regular in some images and random in others. The database consists of 108 images of size 320×320. Some images from the database are shown in Fig. 10. The histogram consists of 40 percent of the pixels of gray level 25 and 60 percent of the pixels of gray level 230.
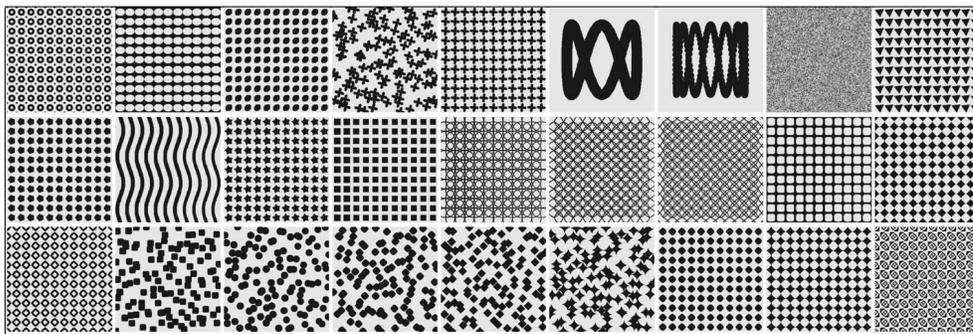
Eight test images were corrupted with Gaussian noise of effective standard deviation 15 gray levels. They were then matched against the database with multiresolution histograms of 8 bit/pixel. Each test image together with the first three matches are shown in Fig. 11. The percentage of correct matches to the corresponding original image as a function of Gaussian noise in the test image is shown in Fig. 12. In Fig. 12a, the matching was performed without smoothing of the database histograms. In Fig. 12b, the matching was performed with smoothing of the database histograms. The intensity resolution of the histograms varied from 8 bit/pixel to 3 bit/pixel in both figures. The matching rate for each intensity resolution is indicated by a different plot.

The plots in Fig. 12 show that multiresolution histograms are robust with respect to both noise and intensity resolution. The best performance in Fig. 12 was obtained for the highest intensity resolution, 8 bit/pixel, due to the large number of pixels in the image, 320×320. The performance is improved with smoothing of the database histograms. This database demonstrates in Fig. 11 the ability of multiresolution histograms to match similar images.

6.2 Database of Brodatz Textures

The second database consists of 13 of the Brodatz textures digitized under seven different rotation angles $0°, 30°, 60°, 90°, 120°, 150°,$ and $200°$. Thus, the total number of images in the database is 91. The size of the images is 179×179 pixels and are histogram equalized. Some of the database images are shown in Fig. 13.

Four equalized test images were corrupted with Gaussian noise of effective standard deviation 15 gray levels. They were matched against the database using 8 bit/pixel multiresolution histograms. Each test image together with the first three matches are shown in Fig. 14. The percentages of correct matches as a function of Gaussian noise are shown in Fig. 15. The matching rates in Fig. 15 leave out the database image corresponding to the test image. The plots in Fig. 15a were obtained without smoothing of the database histograms and constant bin width across image resolution. The plots in Fig. 15b were obtained with smoothing of the database histograms and constant bin width. The intensity resolutions varied from 8 bit/pixel to 3 bit/pixel in both figures. The highest intensity resolutions of 7 bit/pixel and 8 bit/pixel perform the best due to the relatively large number of pixels, 179×179. The plots in Fig. 15c were obtained with histogram smoothing and histogram bin width dependent on image resolution. The initial intensity resolution was 8 bit/pixel. The

Fig. 10. Several samples from the database of synthetic images.
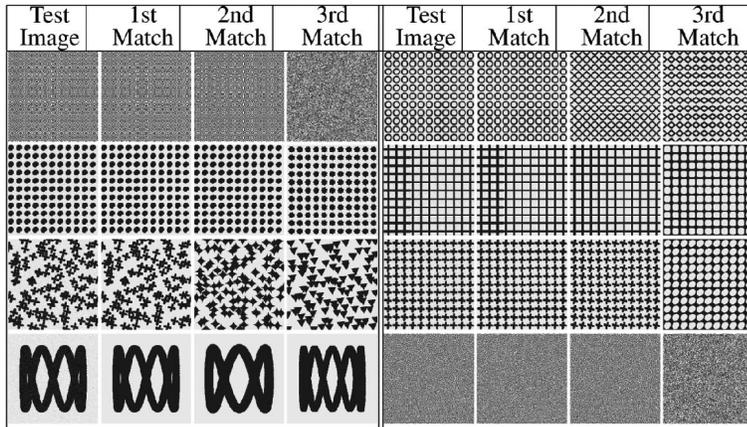
Fig. 11. Matching of synthetic images using 8 bit/pixel multiresolution histograms. The test images are shown in the columns marked Test Image and are corrupted with Gaussian noise of effective standard deviation 15 gray levels. The first three matches are shown in consecutive columns.
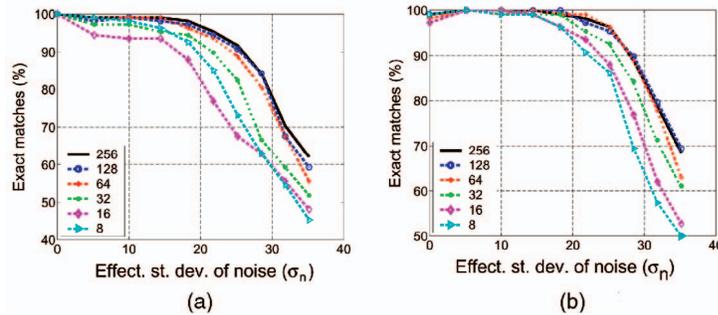
Fig. 12. Matching percentage of multiresolution histograms under Gaussian noise for the database of synthetic images. The sensitivity to intensity resolution was also examined. The plots in (a) were obtained without smoothing of the database histograms, and the plots in (b) were obtained with smoothing of the database histograms.

subsampling factors were $2^{2/3}$ and $2^{1/2}$. Fig. 15c also contains the plot for constant bin width of 8 bit/pixel for comparison.

The plot obtained with the highest subsampling factor in Fig. 15c had the highest performance. Adaptive bin width not only improves performance, but also reduces storage and matching costs. This database demonstrates the robustness of multiresolution histograms to noise and intensity resolution as well as their effectiveness in class matching. It also demonstrates their invariance to rotations.

6.3 Database of CUReT Textures

The third database contains natural textures and is a subset of the CUReT database [16]. It consists of the 61 physical textures with 131 or 132 instances of each physical texture under different illumination and viewing conditions. The total number of images is 8,046 and are all histogram equalized. The subset was selected so that the projection of the textures in the original CUReT database images is at least the size of the images in the database used in this

work, 100×100 pixels. Some of the database images are shown in Fig. 16.

Ten images from the database were corrupted with Gaussian noise of effective standard deviation 15 gray levels. They were matched against the database using 8 bit/pixel multiresolution histograms. In matching the database images corresponding to the test images were excluded. Each test image together with the first three matches are shown in Fig. 17. The matches in the last row and right column, excluding the identical image, are incorrect. The mismatching images, however, are perceptually similar to the test image.

Fig. 14. Matching of Brodatz textures using 8 bit/pixel multiresolution histograms. The test images are shown in the column marked Test Image and are corrupted with Gaussian noise of effective standard deviation 15 gray levels. The first three best matches are shown in consecutive columns.
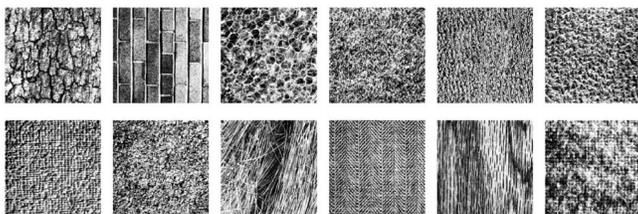
Fig. 13. Samples from the database of Brodatz textures.
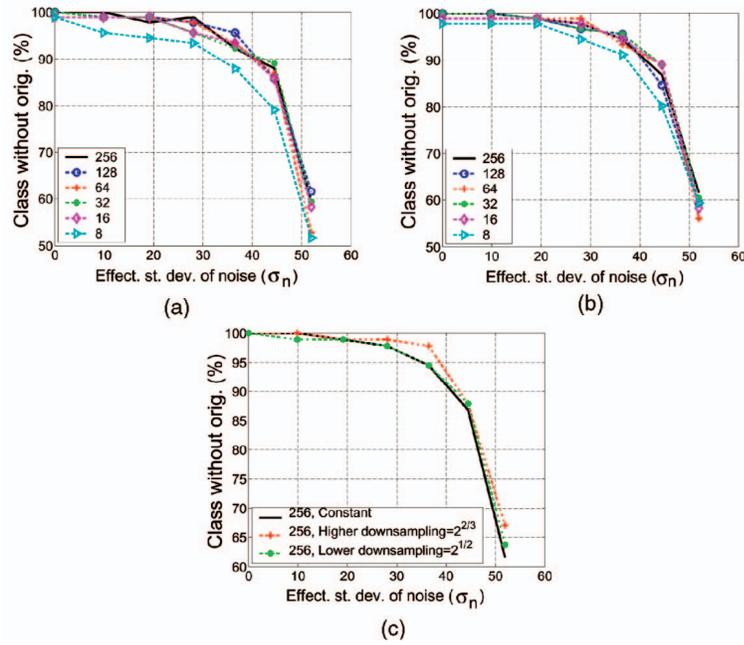
Fig. 15. The percentage of matching for multiresolution histograms with respect to image noise and histogram intensity resolution for the Brodatz textures database. The database image corresponding to the test image was left out in matching. The plots in (a) were obtained without smoothing of the database histograms. The plots in (b) were obtained with smoothing of the database histograms. The plots in (c) were obtained with histogram smoothing and adaptive bin width. The intensity subsampling factors were $2^{2/3}$, and $2^{1/2}$. The initial intensity resolution for all the plots in (c) was 8 bit/pixel.
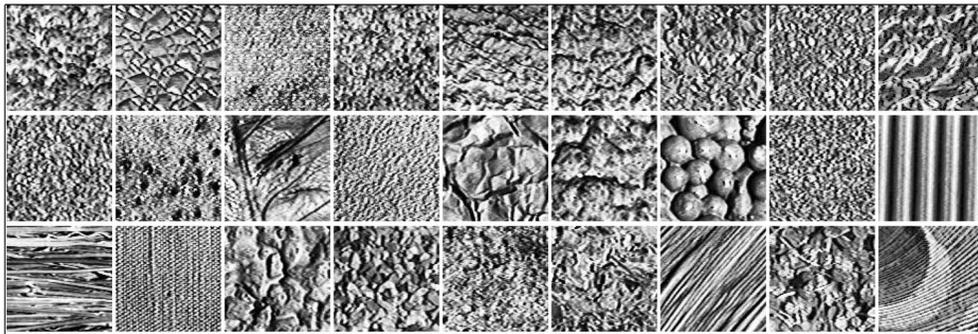
Fig. 16. A few samples from the database of CUReT textures [16].
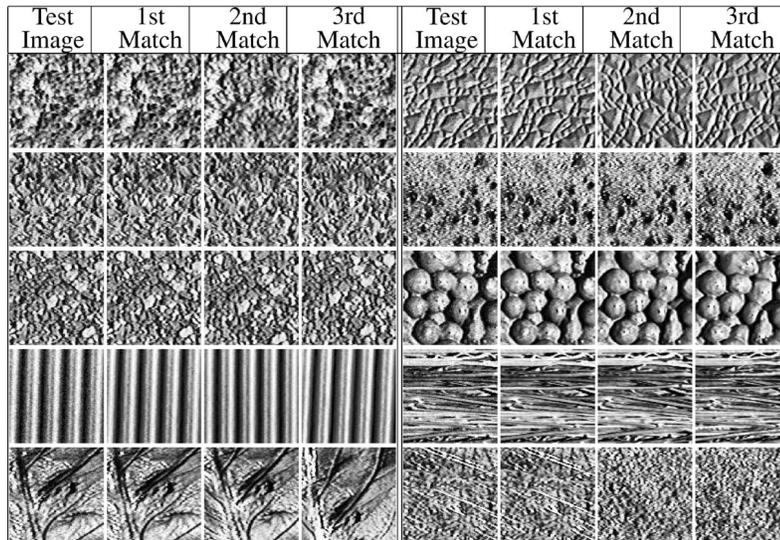
Fig. 17. Matching of CUReT textures using 8 bit/pixel multiresolution histograms. The test images are shown in the columns marked Test Image and are corrupted with Gaussian noise of effective standard deviation 15 gray levels. The first three best matches are shown in consecutive columns.
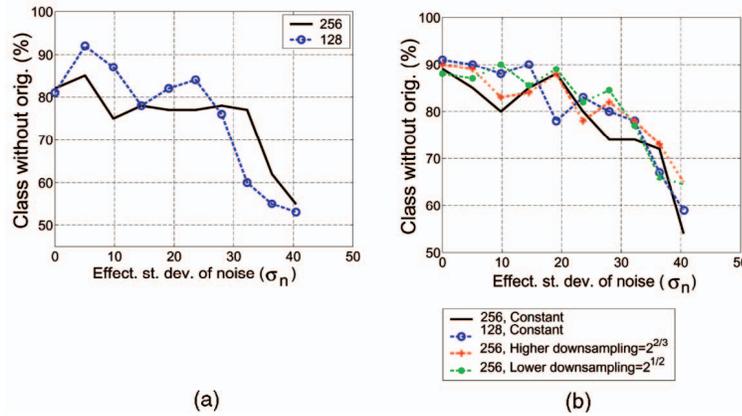
Fig. 18. The percentage of matching for multiresolution histograms with respect to image noise and intensity resolution for the CUReT textures database. The database image corresponding to the test image was left out in matching. The plots in (a) were obtained without smoothing of the database histograms and the plots in (b) with smoothing of the database histograms. In (b), both constant and adaptive bin widths were used. For the adaptive bin widths, the initial intensity resolution was 8 bit/pixel and the subsampling factors were $2^{2/3}$ and $2^{1/2}$.

The percentages of correct matches as a function of Gaussian noise are shown in Fig. 18. To compute the matching percentage for a specific level of noise 100 images were randomly selected from the database and used as test images. The matching rates in Fig. Fig. 18a were obtained with constant histogram bin width. The matching rates in Fig. 18b were obtained with smoothing of the database histograms and either constant or adaptive bin widths. In all plots for constant histogram bin widths the intensity resolutions varied from 8 bit/pixel to 7 bit/pixel. For the plot of the adaptive bin widths the initial intensity resolution was 8 bit/pixel and the subsampling factors were $2^{2/3}$ and $2^{1/2}$.

The multiresolution histograms of intensity resolution 7 bit/pixel performed better than the multiresolution histograms of maximum intensity resolution, 8 bit/pixel. This is more obvious in Fig. 18a and is due to the relatively small number of pixels in the equalized images, 100×100, and the extended dynamic range covered by the histogram. Equation (11) shows that both of these characteristics lead to a larger optimal bin width. The best performance is obtained for histogram smoothing and adaptive bin width.

In Fig. 18a, the best performance for all plots was obtained for nonzero noise because of aliasing. In Fig. 18b, the plots for adaptive bin width performed as well and even better than those with constant bin width. In addition, the algorithms with adaptive bin width have lower storage and matching cost requirements. The high matching performance of the multiresolution histograms for this database demonstrates the robustness with respect to noise and intensity resolution. These experiments also demonstrate the robustness of the multiresolution histograms toward database size, number of database classes, and illumination. In general, the multiresolution histograms are efficient and have been shown to be a robust image feature.

7 COMPARISON WITH OTHER IMAGE FEATURES

The robustness of the multiresolution histogram is verified by comparing it to five commonly used image features. These features are:

1. *Fourier power spectrum annuli*: The image was transformed into frequency domain with regular Fourier tranform that can be applied to images of any length and width. The Fourier power spectrum is segmented into annuli all of which have the same thickness. The feature vector consists of the sum of the values over the different annuli [77].

2. *Gabor wavelet features*: The parameters used were those suggested by Jain et al. [38]. That is, the Gabor filters had four different orientations and the difference between consecutive frequencies was one octave. The range of intensities from lowest to highest for an image with $\sqrt{n}$ pixels per row is of order $O(\log \sqrt{n})$. Filtering was implemented in the Fourier domain. The band-pass images were transformed back into the spatial domain to compute their L_1 norms. The transformations were implemented with fast Fourier transform [38].

3. *Daubechies wavelet packets features*: The feature vector consists of the L_2 norms of the images of the wavelets packets transform [43]. The wavelet transform was combined with spatial subsamling to give critical image sampling.

4. *Auto-cooccurrence matrices*: It was computed over a square window around all pixels. The side of the window was 11 pixels. The feature is the entire matrix [34].

5. *Markov random field parameters*: Each pixel is assumed to be a linear combination of the intensities in a window surrounding it [44], [50]. The size of the window is 3×3. Each image pixel gives one linear relation. The linear parameters are computed with least squares pseudoinverse that gives the shortest length. The pseudoinverse was computed with Householder transformations because they are robust [53].

The distance between the feature vectors was computed with the L_1 norm for all but one. The distance between wavelet packets feature vectors was the L_2 norm [43]. The matching performance of the features in this work are computed for the two databases described in Section 6. The first is the database of Brodatz images, samples of which are shown in Fig. 13. The second is the database of CUReT textures, samples of which are shown in Fig. 16.

All features except wavelets packets features are invariant with respect to translation. The only feature invariant to
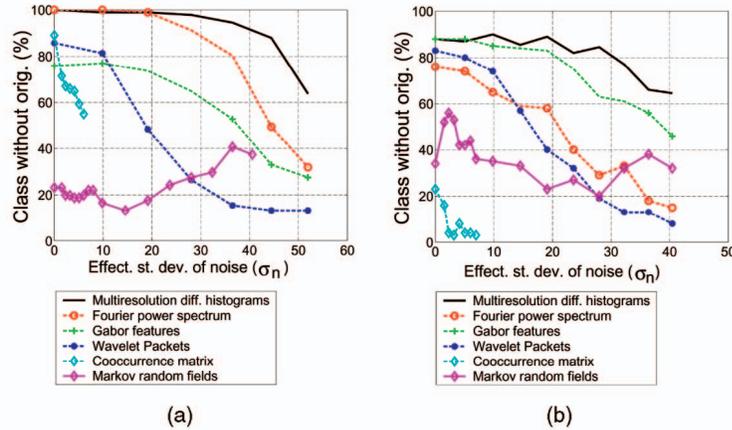
Fig. 19. The percentage of matching as a function of noise. In (a) are the results for the database of Brodatz textures and in (b) are the results for the database of CUReT textures. The matching excludes the database image corresponding to the test image. The multiresolution histograms are the most robust to rotations and noise.

rotation are the multiresolution histograms. The Fourier power spectrum annuli are robust to rotations. The auto-cooccurrence matrices are also relatively robust to rotations.

7.1 Comparison of Matching Performance

The experimental setup is the same as that described in the previous section where the multiresolution histograms were examined for their sensitivity to their parameters. To compute the matching rate the original database images corresponding to the test images were not considered. The sensitivity of the features to their parameters was also examined. This makes the comparison with the multiresolution histograms fair, since the sensitivity of the latter with respect to intensity resolution has already been examined in Section 6. The best performing parameters from each feature are used.

For the multiresolution histogram, histogram smoothing as well as adaptive bin width were used. The initial image resolution is the maximum, 256. The subsampling factor of the bin width was $2^{1/2}$. The same parameters of the multiresolution histograms were used for both databases. The best performing set of frequencies of the Gabor filters for both databases is the one where all the image harmonics are considered.

The first comparison is based on the *database of Brodatz images*. The number of annuli of the Fourier power spectrum was 20. The number of resolutions of the Daubechies wavelet packets transform was 3. The intensity resolution of the auto-cooccurrence matrices was 3 bit/pixel. The plot of the matching rate as a function of noise starting from zero is shown in Fig. 19a.

The matching rate at zero noise in Fig. 19a is an indication of the sensitivity to rotations. The matching rate of multiresolution histograms is 100 percent since they are invariant to rotations. The annular features of the Fourier power spectrum also have a matching rate of 100 percent. Thus, they are very robust to rotations. The performance of the auto-cooccurrence matrix with respect to rotations at zero level of noise is high relative to the performance of the other features. The wavelet packets features and the Gabor features are rotationally sensitive and have a very low matching performance.

The results in Fig. 19a for the Brodatz database demonstrate that the multiresolution histograms and the Gabor features are the most robust to noise. The auto-cooccurrence matrices and the Daubechies wavelet features are more sensitive to noise.

The second comparison was based on the *database of CUReT textures*. The number of annuli of the Fourier power spectrum was 20, and the number of levels of the wavelet packets transform was 4. The intensity resolution of the auto-cooccurrence matrices was 5 bit/pixel. Fig. 19b shows the percentage of matches. The size of the CUReT database is 8,046 images. Thus, the matching rate for this database provides information about the sensitivity of the features to database size. The multiresolution histograms, the wavelet packets features, and the Gabor features are robust with respect to the size of the database and the number of classes in the database. The Fourier power spectrum features are sensitive to database size. The auto-cooccurrence matrices and the Markov random field parameters are very sensitive with respect to the size of the database. The sensitivity to database size is summarized in Table 1.

The CUReT database also provides information about robustness to noise. The multiresolution histograms are the most robust to noise. The Gabor features and Daubechies wavelet features are also robust to noise. Auto-cooccurrence matrices and Markov random field parameters are very sensitive to noise. The results about the robustness to noise agree with those of the Brodatz database. The multiresolution histograms are robust to noise, since Gaussian filtering averages out noise, and image noise simply smoothes the histograms. The sensitivity of the features to noise and illumination is summarized in Table 1. The most sensitive to illumination is the auto-cooccurrence matrix. In general, the multiresolution histogram is the most robust feature.

7.2 Comparison of Computation Costs

The same parameters are used for all the features. The size of the images is 200×200, that is $n = 40,000$ pixels. The width of the window is $\lambda = 5$. The levels of resolution are $l = 5$. For the multiresolution histograms and the auto-cooccurrence matrices $t = 256$, for the Fourier power spectrum features $t = 40$, and for the Gabor features $t = 16$.

TABLE 1
The Experimental Sensitivity of the Texture Features with Respect to Superimposed Noise, Database Size, and Illumination

Feature	Gaussian noise	Database size	Illumination
Fourier power spectrum annuli	robust	sensitive	robust
Gabor features	robust	robust	robust
Daubechies wavelet energies	sensitive	robust	robust
Multiresolution histograms	robust	robust	robust
Auto-cooccurrence matrix	very sensitive	very sensitive	very sensitive
Markov random field parameters	very sensitive	very sensitive	sensitive

The computation cost of the multiresolution histograms was computed in Section 5. The computation costs of all the features are given in Table 2. The least expensive feature to compute are the multiresolution histograms. The most expensive feature to compute are the Markov random field parameters, since they involve the computation of a least squares pseudoinverse. The Gabor features are the most expensive to compute for large images. This is because both the number of features and the cost of the fast Fourier transform increase with image size.

The Gabor decomposition involves a high degree of oversampling since many bandpass images are computed. This increases the computational requirements. The wavelet packets decomposition samples the image critically. The latter could also be implemented without image subsampling. Avoiding subsampling would result in oversampling of the image. It would make, however, the wavelet decomposition invariant to translations, and might also improve its performance. The multiresolution histograms were implemented with the Burt-Adelson pyramid. They could also be implemented without subsampling.

8 CONCLUSION AND FUTURE WORK

The multiresolution histogram captures spatial image information. Moreover, it retains the simplicity, efficiency, and robustness of the plain histograms. The high matching performance of the multiresolution histograms was demonstrated experimentally. They were also shown to be very robust to image rotation, image noise, and intensity resolution. Smoothing the histogram and using a histogram bin

TABLE 2
The Features are Listed Top to Bottom in Order
of Decreasing Computation Cost

	Feature	Computation cost
1	Markov random field parameters	$O(n(\lambda^2-1)^2 - \frac{(\lambda^2-1)^3}{3})$
2	Gabor features	$O((\log\sqrt{n}+1)n\log n)$
3	Fourier power spectrum features	$O(n\sqrt{n})$
4	Auto-cooccurrence matrix	$O(n\lambda^2)$
4	Wavelet coefficient energies	$O(n\lambda l)$
6	Multiresolution histograms	$O(n\lambda)$

The most expensive to compute are the Markov random field parameters and the Gabor features.

width dependent on image resolution also improved performance. The multiresolution histogram was compared to five other image features. It was shown that the multiresolution histograms are the most efficient and robust.

The dependence of the multiresolution histogram as well as the Fisher information measures on image shape and image texture can be investigated further. The multiresolution histogram can also be computed over higher dimensional domains such as 3D data. The range of the multiresolution histograms can be of multiple dimensions such as color. The performance of multiresolution histograms formed with eccentric Gaussians could be examined. Such multiresolution histograms would be rotationally sensitive.

APPENDIX A

RELATION BETWEEN HISTOGRAM AND TSALLIS GENERALIZED ENTROPIES

Property. *A histogram of m gray levels is related linearly to a vector $\mathbf{S}$ of Tsallis entropies of m different orders.*

Proof. Take the orders of the m entropies to be $\mathbf{q} = [q_0 q_1 q_2 \ldots q_{m-1}]$ and the corresponding entropies to be given by the vector $\mathbf{S}$ of size m. Each element of vector $\mathbf{S}$ can be expressed in terms of the histogram, as shown in (2), to give

$$\mathbf{S} = \mathbf{W}(\mathbf{U} - \mathbf{V})\mathbf{h}. \qquad (13)$$

Matrix $\mathbf{W}$ is an $m \times m$ matrix with diagonal elements given by $w_{ii} = \frac{1}{q_i - 1}$. Matrix $\mathbf{U}$ has identical rows, each one given by $[v_0\ v_1\ v_2 \ldots v_{m-1}]$, which are the consecutive image gray levels. Matrix $\mathbf{V}$ is given by:

$$\mathbf{V} = \begin{pmatrix} v_0^{q_0} & v_1^{q_0} & \cdots & v_{m-1}^{q_0} \\ v_0^{q_1} & v_1^{q_1} & \cdots & v_{m-1}^{q_1} \\ v_0^{q_2} & v_1^{q_2} & \cdots & v_{m-1}^{q_2} \\ \cdots & \cdots & \ddots & \cdots \\ v_0^{q_{m-1}} & v_1^{q_{m-1}} & \cdots & v_{m-1}^{q_{m-1}} \end{pmatrix}.$$

The matrix $\mathcal{P} = \mathbf{W}(\mathbf{U} - \mathbf{V})$ in (13) is the linear proportionality matrix of (3) and (8). □

REFERENCES

[1] J. Bach, C. Fuler, A. Gupta, A. Hampapur, B. Horowitz, R. Humphrey, R. Jain, and C. Shu, "The Virage Image Search Engine: An Open Framework for Image Management," *Proc. SPIE Conf. Storage and Retrieval for Image and Video Databases IV,* vol. 2670, pp. 76-87, 1996.

[2] D. Ballard and L. Wixson, "Object Recognition Using Steerable Filters at Multiple Scales," *Proc. IEEE Workshop Qualitative Vision,* pp. 2-10, 1993.

[3] A. Barron, "Entropy and the Central Limit Theorem," *The Annals of Probability,* vol. 14, no. 1, pp. 336-342, 1986.

[4] N. Blackman, "The Convolution Inequality for Entropy Powers," *IEEE Trans. Information Theory,* vol. 11, pp. 267-271, 1965.

[5] J. Bonet, P. Viola, and J. Fisher, "Flexible Histograms: A Multiresolution Target Discrimination Model," *Proc. of SPIE,* vol. 3370, 1998.

[6] J.D. Bonet, "Multiresolution Sampling Procedure for Analysis and Synthesis of Texture Images," *Proc. Ann. Conf. Computer Graphics (SIGGRAPH),* pp. 361-368, 1997.

[7] A. Bovik, M. Clark, and W. Geisler, "Multichannel Texture Analysis Using Localized Spatial Filters," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 12, no. 1, pp. 55-73, Jan. 1990.

[8] P. Brodatz, *Textures: A Photographic Album for Artists and Designers.* New York: Dover Publications, 1966.

[9] P. Burt and E. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Trans. Comm.,* vol. 31, no. 4, pp. 532-540, 1983.

[10] F. Campbell and J. Robson, "Application of Fourier Analysis to the Visibility of Gratings," *J. Physiology,* vol. 197, pp. 551-566, 1968.

[11] E. Carlen and A. Soffet, "Entropy Production by Block Variable Summation and Central Limit Theorems," *Comm. Math. Physics,* vol. 140, pp. 339-371, 1991.

[12] M. Carlotto, "Texture Classification Based on Hypothesis Testing Approach," *Proc. Int'l Japanese Conf. Pattern Recognition,* pp. 93-96, 1984.

[13] M. Carlotto, "Histogram Analysis Using a Scale-Space Approach," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 9, no. 1, pp. 121-129, 1987.

[14] T. Chang and C. Kuo, "Texture Analysis and Classification Using Tree Structured Wavelet Transform," *IEEE Trans. Image Processing,* vol. 4, pp. 863-870, 1993.

[15] J. Coggins and A. Jain, "A Spatial Filtering Approach to Texture Analysis," *Pattern Recognition Letters,"* vol. 3, pp. 195-203, 1985.

[16] K. Dana, B. Ginneken, S. Nayar, and J. Koenderink, "Reflectance and Texture of Real-World Surfaces," *ACM Trans. Graphics,* vol. 18, no. 1, pp. 1-34, 1999.

[17] L. Devroye and L. Gyorfi, *Nonparametric Density Estimation: The L_1 View.* John Wiley and Sons, 1985.

[18] J. Engel, "The Multiresolution Histogram," *Metrika,* vol. 46, pp. 41-57, 1997.

[19] A. Evans, N. Thacker, and J. Mayhew, "The Use of Geometric Histograms for Model-Based Object Recognition," *Proc. British Machine Vision Conf.,* pp. 429-438, 1993.

[20] O. Faugeras, "Texture Analysis and Classification Using a Human Visual Model," *Proc. Fourth Int'l Conf. Pattern Recognition,* pp. 549-552, 1978.

[21] I. Fogel and D. Sagi, "Gabor Filters as Texture Discriminator," *Biological Cybernetics,* vol. 61, pp. 103-113, 1989.

[22] B. Funt and G. Finlayson, "Color Constant Color Indexing," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 17, no. 5, pp. 522-529. May 1995.

[23] H. Greenspan, S. Belongie, R. Goodman, and P. Perona, "Rotation Invariant Texture Recognition Using a Steerable Pyramid," *Proc. 12th Int'l Conf. Pattern Recognition,* pp. 162-167, 1994.

[24] L. Griffin, "Scale–Imprecision Space," *Image and Vision Computing,* vol. 15, pp. 369-398, 1997.

[25] E. Hadjidemetriou, "Use of Histograms for Recognition." Columbia Univ., ftp://ftp.cs.columbia.edu/pub/CAVE/stathis/thesisHadjidemetriou.pdf, 2002.

[26] E. Hadjidemetriou, M. Grossberg, and S. Nayar, "Histogram Preserving Image Transformations," *Int'l J. Computer Vision,* vol. 45, no. 1, pp. 5-23, 2001.

[27] E. Hadjidemetriou, M. Grossberg, and S. Nayar, "Spatial Information in Multiresolution Histograms," *Proc. Computer Vision and Pattern Recognition Conf.,* vol. 1, pp. 702-709, 2001.

[28] J. Hafner, H. Sawhney, W. Equitz, M. Flickner, and W. Niblack, "Efficient Color Histogram Indexing for Quadratic Form Distance Functions," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 17, no. 7, pp. 729-736, July 1995.

[29] G. Haley and B. Manjunath, "Rotation-Invariant Texture Classification Using a Complete Space Frequency Model," *IEEE Trans. Image Processing,* vol. 8, no. 2, pp. 255-269, 1999.

[30] R. Haralick, "Statistical and Structural Approaches to Texture," *Proc. IEEE,* vol. 67, pp. 786-804, 1979.

[31] D. Heeger and J. Bergen, "Pyramid-Based Texture Analysis/Synthesis," *Proc. Ann. Conf. Computer Graphics (SIGGRAPH),* pp. 229-234, 1995.

[32] C. Helstrom, *Probability and Stochastic Processes for Engineers.* Toronto: Macmillan, 1991.

[33] W. Hsu, T. Chua, and H. Pung, "An Integrated Color-Spatial Approach to Content-Based Image Retrieval," *Proc. ACM Multimedia,* pp. 305-313, 1995.

[34] J. Huang, S. Kumar, M. Mitra, W. Zhu, and R. Zabih, "Spatial Color Indexing and Applications," *Int'l J. Computer Vision,* vol. 35, no. 3, pp. 245-268, 1999.

[35] B. Huet and E. Hancock, "Line Pattern Retrieval Using Relational Histograms," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, no. 12, pp. 1363-1370, Dec. 1999.

[36] M. Jagersand, "Saliency Maps and Attention Selection in Scale and Spatial Coordinates: An Information Theoretic Approach," *Proc. Int'l Conf. Computer Vision,* pp. 195-202, 1995.

[37] A. Jain and F. Farrokhnia, "Unsupervised Texture Segmentation Using Gabor Filters," *Pattern Recognition,* vol. 24, no. 12, pp. 1167-1186, 1991.

[38] A. Jain, N. Ratha, and S. Lakshmanan, "Object Detection Using Gabor Filters," *Pattern Recognition,* vol. 30, no. 2, pp. 295-309, 1997.

[39] O. Johnson, "Entropy Inequalities and the Central Limit Theorem," *Stochastic Processes and Their Applications,* vol. 88, pp. 291-304, 2000.

[40] T. Kadir and M. Brady, "Saliency, Scale and Image Description," *Int'l J. Computer Vision,* vol. 45, no. 2, pp. 83-105, 2001.

[41] J. Koenderink, "The Structure of Images," *Biological Cybernetics,* vol. 50, pp. 363-370, 1984.

[42] J. Koenderink and A.J. van Doorn, "The Structure of Locally Orderless Images," *Int'l J. Computer Vision,* vol. 31, nos. 2-3, pp. 159-168, 1999.

[43] A. Laine and J. Fan, "Texture Classification by Wavelet Packet Signatures," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 15, no. 11, pp. 1186-1191, Nov. 1993.

[44] S. Lakshmanan and H. Derin, "Gaussian Markov Random Fields at Multiple Resolutions," *Markov Random Fields: Theory and Application,* R. Chellapa and A. Jain, eds., pp. 131-157, 1993.

[45] J. Lee and B. Dickinson, "Multiresolution Video Indexing for Subband Coded Video Databases," *Proc. SPIE Conf. Storage and Retrieval for Image and Video Databases II,* vol. 2185, pp. 162-173, 1994.

[46] W. Leow and R. Li, "Adaptive Binning and Dissimilarity Measure for Image Retrieval and Classification," *Proc. Conf. Computer Vision and Pattern Recognition,* vol. 2, pp. 234-239, 2001.

[47] G. Lowitz, "Can a Local Histogram Really Map Texture Information?" *Pattern Recognition,* vol. 16, pp. 141-147, 1983.

[48] J. Malik and P. Perona, "Preattentive Texture Discrimination with Early Vision Mechanisms," *J. Optical Soc. Am. A,* vol. 7, no. 5, pp. 923-932, 1990.

[49] R. Manthalkar, P. Biswas, and B. Chatterji, "Rotation and Scale Invariant Texture Classification Using Gabor Wavelets," *Proc. Second Int'l Workshop Texture Analysis and Synthesis (in conjuction with ECCV),* pp. 87-90, 2002.

[50] J. Mao and A. Jain, "Texture Classification and Segmentation Using Multiresolution Simultaneous Autoregressive Models," *Pattern Recognition,* vol. 25, no. 2, pp. 173-188, 1992.

[51] B. Mel, "Combining Color Shape and Texture Histogramming in a Neurally Inspired Approach to Visual Object Recognition," *Neural Computation,* vol. 9, pp. 777-804, 1997.

[52] W. Niblack, "The QBIC Project: Querying Images by Content Using Color, Texture, and Shape," *Proc. SPIE Conf. Storage and Retrieval for Image and Video Databases,* vol. 1908, pp. 173-187, 1993.

[53] B. Noble and J. Daniel, *Applied Linear Algebra.* Prentice-Hall, 1988.

[54] A. Oomes and P. Snoeren, "Structural Information in Scale-Space," *Proc. Copenhagen Workshop Gaussian Scale-Space Theory,* vol. 96, no. 19, pp. 48-57, May 1996.

[55] G. Pass and R. Zabih, "Histogram Refinement for Content Based Image Retrieval," *Proc. IEEE Workshop Applications of Computer Vision,* pp. 96-102, 1996.

[56] G. Pass and R. Zabih, "Comparing Images Using Joint Histograms," *Multimedia Systems,* vol. 7, no. 3, pp. 234-240, 1999.

[57] F. Pennini and A. Plastino, "Fisher's Information Measure in a Tsallis' Nonextensive Setting and Its Application to Diffusive Processes," *Physica A,* vol. 247, pp. 559-569, 1997.

[58] A. Plastino, A. Plastino, and H. Miller, "Tsallis Nonextensive Thermostatistics and Fisher's Information Measure," *Physica A,* vol. 235, pp. 577-588, 1997.

[59] M. Popovic, "Texture Analysis Using 2D Wavelet Transform: Theory and Applications," *Proc. IEEE TELSIKS,* pp. 149-158, 1999.

[60] M. Porat and Y. Zeevi, "The Generalized Gabor Scheme of Image Representation in Biological and Machine Vision," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 10, no. 4, pp. 452-468, 1988.

[61] J. Puzicha, T. Hofmann, and J. Buchmann, "Histogram Clustering for Unsupervised Segmentation and Image Retrieval," *Pattern Recognition Letters,* vol. 20, no. 9, pp. 899-909, 1999.

[62] S. Ravela and R. Manmatha, "Retrieving Images by Appearance," *Proc. Int'l Conf. Computer Vision,* pp. 608-613, 1998.

[63] Y. Rubner, L. Guibas, and C. Tomasi, "The Earth Mover's Distance, Multi-Dimensional Scaling, and Color-Based Image Retrieval," *Proc. ARPA Image Understanding Workshop,* 1997.

[64] S. Sablak and T. Boult, "Multilevel Color Histogram Representation of Color Images by Peaks for Omni-Camera," *Proc. IASTED Int'l Conf. Signal and Image Processing,* 1999.

[65] B. Schiele and J. Crowley, "Recognition without Correspondence Using Multidimensional Receptive Field Histograms," *Int'l J. Computer Vision,* vol. 36, no. 1, pp. 31-50, 2000.

[66] D. Scott, "On Optimal and Data-Based Histograms," *Biometrica,* vol. 66, no. 3, pp. 605-610, 1979.

[67] J. Smith and S. Chang, "Tools and Techniques for Color Image Retrieval," *Proc. SPIE,* vol. 2670, pp. 1630-1639, 1996.

[68] J. Sporring, C. Colios, and P. Trahanias, "Generalized Scale-Space," *Proc. Int'l Conf. Image Processing,* vol. 1, pp. 920-923, 2000.

[69] J. Sporring and J. Weickert, "Information Measures in Scale-Spaces," *IEEE Trans. Information Theory,* vol. 45, no. 3, pp. 1051-1058, 1999.

[70] A. Stam, "Some Inequalities Satisfied by the Quantities of Information of Fisher and Shannon," *Information and Control,* vol. 2, pp. 101-112, 1959.

[71] M. Stricker and A. Dimai, "Color Indexing with Weak Spatial Constraints," *Proc. SPIE Storage and Retrieval for Image and Video Databases,* pp. 29-39, 1996.

[72] M. Stricker and M. Orengo, "Similarity of Color Images," *Proc. SPIE Conf. Storage and Retrieval for Image and Video Databases III,* vol. 2420, pp. 381-392, 1995.

[73] M. Swain and D. Ballard, "Color Indexing," *Int'l J. Computer Vision,* vol. 7, no. 1, pp. 11-32, 1991.

[74] M. Tanaka, T. Watanabe, and T. Mishima, "Tsallis Entropy in Scale-Spaces," *Proc. SPIE Conf. Vision Geometry VIII,* vol. 3811, pp. 273-281, 1999.

[75] C. Taylor, "Akaike's Information Criterion and the Histogram," *Biometrica,* vol. 74, no. 3, pp. 636-639, 1987.

[76] C. Tsallis, "Nonextensive Statistics: Theoretical, Experimental and Computational Evidences and Connections," *Brazilian J. Physics,* vol. 29, no. 1, 1999.

[77] J. Weszka, C. Dryer, and A. Rosenfeld, "A Comparative Study of Texture Measures for Terrain Classification," *IEEE Trans. Systems, Man, and Cybernetics,* vol. 6, no. 4, pp. 269-285, 1976.

[78] A. Witkin, "Scale-Space Filtering," *Proc. of Int'l Joint Conf. of Artificial Intelligence,* pp. 1019-1022, 1983.

[79] R. Young, "The Gaussian Derivative Theory of Spatial Vision: Analysis of Cortical Cell Receptive Field Line–Weighting Profiles," Technical Report GMR–4920, 1985.

[80] H. Zhang, A. Kankanhali, and S. Smoliar, "Automatic Partitioning of Full-Motion Video," *Multimedia Systems,* vol. 1, pp. 10-28, 1993.

[81] H. Zhang, C. Low, W. Smoliar, and J. Wu, "Video Parsing, Retrieval and Browsing: An Integrated and Content-Based Solution," *ACM Multimedia,* pp. 15-24, 1995.

Efstathios Hadjidemetriou received the BEng (Honors) degree in electrical engineering from McGill University, Montreal, in 1994. He received the MSc degree in electrical engineering from Columbia University in 1996 together with the Armstrong memorial award. In 2002, he received the PhD degree in computer science also from Columbia University. He is currently a postdoctoral associate at the Yale University School of Medicine, Department of Diagnostic Radiology and Electrical Engineering. His PhD thesis was concentrated on the use of histograms for recognition. He is currently working on the processing of confocal microscopy video sequences and on analyzing magnetic resonance spectroscopy data.

Michael D. Grossberg received the PhD degree in mathematics from the Massachusetts Institute of Technology in 1991. He is a research scientist with the Columbia Automated Vision Environment (CAVE), at Columbia University. His research in computer vision has included topics in the geometric and photometric modeling of cameras, and analyzing features for indexing. Dr. Grossberg was a lecturer in the Computer Science Department at Columbia University. He was also a Ritt Assistant Professor of Mathematics at Columbia University. He has held postdoctoral fellowships at the Max Plank Institute for Mathematics in Bonn and Hebrew University in Jerusalem. He has authored and coauthored papers that have appeared in ICCV, ECCV, CVPR. He has filed several US and international patents for inventions related to computer vision. He is a member of the IEEE.

Shree K. Nayar received the PhD degree in electrical and computer engineering from the Robotics Institute at Carnegie Mellon University in 1990. He is the T.C. Chang Professor of Computer Science at Columbia University. He currently heads the Columbia Automated Vision Environment (CAVE), which is dedicated to the development of advanced computer vision systems. His research is focused on three areas: the creation of cameras that produce new forms of visual information, the modeling of the interaction of light with materials, and the design of algorithms that recognize objects from images. His work is motivated by applications in the fields of computer graphics, human-machine interfaces, and robotics. Dr. Nayar has authored and coauthored papers that have received the Best Paper Honorable Mention Award at the 2000 IEEE CVPR conference, the David Marr Prize at the 1995 ICCV held in Boston, Siemens Outstanding Paper Award at the 1994 IEEE CVPR conference held in Seattle, 1994 Annual Pattern Recognition Award from the Pattern Recognition Society, Best Industry Related Paper Award at the 1994 ICPR held in Jerusalem, and the David Marr Prize at the 1990 ICCV held in Osaka. He holds several US and international patents for inventions related to computer vision and robotics. Dr. Nayar was the recipient of the David and Lucile Packard Fellowship for Science and Engineering in 1992, the National Young Investigator Award from the National Science Foundation in 1993, and the Excellence in Engineering Teaching Award from the Keck Foundation in 1995. He is a member of the IEEE.

▷ **For more information on this or any computing topic, please visit our Digital Library at** www.computer.org/publications/dlib.