PERGAMON

Vision Research 41 (2001) 585–598

Vision Research

www.elsevier.com/locate/visres

Sensitivity to contrast histogram differences in synthetic wavelet-textures

Frederick A.A. Kingdom [a,*], Anthony Hayes [b], David J. Field [c]

[a] *McGill Vision Research Unit, Department of Ophthalmology, McGill University, Room H4-14, 687 Pine Avenue West, Montreal, Quebec, Canada H3A 1A1*

[b] *Department of Psychology, University of Melbourne, Parkville, Victoria 3052, Australia*

[c] *Department of Psychology, Cornell University, Ithaca, NY 14853, USA*

Received 13 December 1999; received in revised form 27 October 2000

Abstract

Recent research on texture synthesis suggests that characterisation of those properties of textures to which human observers are sensitive may be provided by the histograms of the coefficients of a wavelet decomposition. In this study we examined the properties of wavelet histograms that affect texture discrimination by measuring observer sensitivity to differences in the wavelet histograms of synthetic textures. The textures, generated via Gabor micropattern synthesis, were broadband, with amplitude spectra that are characteristic of natural images, i.e. $1/f$. We measured texture-difference thresholds for three moments of the wavelet histograms — variance, skew and kurtosis — by manipulating the contrast, phase, and density, of the Gabor elements used to construct the textures. Observers discriminated more efficiently between textures that had differences in kurtosis, than between textures that had differences in either variance or skew. Performance was compared to two model observers; one used the pixel-luminance histogram, the other used the histogram of the output of wavelet-filters. The results support the idea that the visual system is relatively sensitive to the kurtosis, or 4th moment, of the wavelet histogram of textures. We argue that higher than 4th-order moments will, in practice, become increasingly difficult for the visual system to represent because the lack of a perfect match between the elements and the receptive fields effectively blurs the response histogram, thereby attenuating higher moments.

Keywords: Texture discrimination; Texture synthesis; Fractal textures; Kurtosis

1. Introduction

Although there is no universally accepted definition of texture, there is general agreement that textures consist of aggregations of small elements that are uniform in at least some of their image statistics. These statistics may include orientation, colour, spatial frequency, density, and contrast. Understanding which statistical properties are important for the perceptual discrimination of textures is pertinent to both vision science and computer graphics.

Recently, a number of studies have reported considerable success in synthesising complex natural textures (e.g. Heeger & Bergen, 1995; De Bonet, 1997; Simoncelli & Portilla, 1998; Portilla & Simoncelli, 1999). In these studies a particular subset of the statistical properties of a complex natural texture (e.g. a field of grass) are measured. These statistics are then used in the synthesis of a new image that contains what it is hoped are the important texture properties, though not necessarily all the properties, of the original pattern. To the extent that the method is successful, the synthesised textures will appear to be examples of textures drawn from the same population (e.g. similar fields of grass). In many cases, the synthetic patterns are perceptually equivalent to the originals; at least prior to attentive scrutiny. However, the approach, which may succeed for some textures, fails for others; there is often a dependency on the structure of the particular natural texture that is not reflected in the measured subset of statistics. Although our goal here is not to produce, or

* Corresponding author. Tel.: +1-514-8421231, ext. 4804; fax +1-514-8431691.

E-mail address: fkingd@po-box.mcgill.ca (F.A.A. Kingdom).

0042-6989/01/$ - see front matter © 2001 Published by Elsevier Science Ltd.
PII: S0042-6989(00)00284-4

test, a model of texture synthesis, the variability in success of current methods for the efficient cloning of new examples from a given texture invites the question of which statistics are important for perceptual equivalence.

1.1. Order statistics in texture discrimination

Some of the earliest work on texture segregation was that of Julesz and colleagues (e.g. Julesz, Frisch, Gilbert, & Shepp, 1973), who attempted to describe textures by means of a set of 'nth-order' combinatorial relationships among pixel values. Use of the 'order' of statistics is common to a number of scientific fields, and is common to a number of approaches to the description of image and texture properties. But the definition of an 'nth-order' statistic varies. For Julesz, a 1st-order statistic is a representation of the pixel histogram, a 2nd-order statistic is a measurement of the relationship among pairs, or dipoles, of pixels, a 3rd-order statistic is a measurement of the relationship among triplets (tripoles) of pixels, and so on. In his early work, Julesz suggested that all the statistics relevant to perceptual equivalence of textures were found in the 1st- and 2nd-order relationships. However, later, he and others demonstrated that textures differing only in higher-order Julesz statistics could be effortlessly discriminated (reviewed in Bergen, 1991). (A recently highlighted problem with these studies concerns the assumption that the texture pairs had identical dipole statistics. Chubb and Yellott (2000) claim that any discreet image can be completely specified by its dipole statistics (see also Gagalowicz, 1981; Yellott, 1993 and Section 4).

There is sometimes confusion between the definition of the order of statistics as described in the above paragraph and the moments of a distribution. For example, a statistical measure of '*n*th-order' can refer to the value of the largest exponent of the polynomial describing a function, often referred to as the moment of the function. Thus, the variance of a pixel histogram would be a 2nd-order statistic of the image because it is a function generated from the squares of the pixel amplitudes. By this measure, the variance, skew and kurtosis (the 2nd, 3rd, and 4th moments) are examples of 2nd-, 3rd- and 4th-order statistics since the definitions use the following powers of the exponents:

$$\text{Variance} = \sum_{i}^{N}(x_i - \mu)^2/N$$

$$\text{Skew} = \sum_{i}^{N}(x_i - \mu)^3/N$$

$$\text{Kurtosis} = \sum_{i}^{N}(x_i - \mu)^4/N$$

where x_i is the value of the ith member of the distribution, μ the mean of the distribution and N the number of values.

Although the above two definitions of '*n*th order statistic' are not equivalent, a more general description encompasses both of them. By allowing 'n' to represent the components, a 4th-order statistic would include (x_1, x_2, x_3, x_4) as with the relational statistics, or $(x_1, x_1\, x_1\, x_1)$ as with the 4th moment, and might also include $(x_1\, x_1, x_2, x_3)$ or $(x_1, x_1\, x_1, x_4)$. However, in this paper we want to make a distinction between what we describe as combinatorial statistics, describing activity dependencies between different vectors (e.g. different neurones), and histogram statistics as defined by the moments. If we take the example of neuron activity, the moments capture the activity distribution of all the cells, or the activity distribution within a class of cells, and the combinatorial statistics describe the relative activity of different cell types as a function of distance. In this paper we explore how much information about the histogram is available to the visual system. We ask how accurately the visual system represents the shape of this histogram.

When discussing the 'order' of statistics used to describe an image, it is important to consider the nature of the basis functions that are used to represent the image. For example, we can describe an image by the magnitudes of pixel values. We can also completely describe an image by the amplitudes of all of its complex Fourier coefficients. We can thus discuss the '1st-order' combinatorial statistics of the pixels as described by the pixel histogram, or we can discuss the 1st-order combinatorial statistics as described by the histograms of the Fourier amplitudes. However, these two 1st-order histograms do not provide the same information. What may be a high order combinatorial statistic for one basis set may be a relatively low order combinatorial statistic for another set. The proposal that the visual system uses only 1st- and 2nd-order combinatorial relationships (as first proposed by Julesz) may indeed be true, once the texture is defined using an appropriate basis set. When wavelet basis functions are used to describe a texture, the 1st-order (histograms) and 2nd-order dipoles (combinatorial relations) may well be sufficient to characterise the perceptually important properties of textures.

There are several reasons why Gabor (or similar — the Gabor is a Gaussian windowed sinusoid) wavelets may be effective in characterising the perceptual properties of textures. They provide a reasonable match to what are understood to be the properties of 'simple' cells in primary visual cortex, and if the early stages of texture analysis begin here, then these units of description may allow a relatively efficient description. It has been argued that the wavelet description is a statistically efficient description of natural scenes (e.g. Field, 1994), allowing a scene to be represented with a relatively independent sparse number of active elements. For these reasons, it would not be surprising to dis-

cover that the perceptual properties of natural textures could be described with a relatively 'low-order' wavelet analysis.

In this study we focus on the 1st-order combinatorial statistics of textures using Gabor wavelet basis functions. Following from our own work (Field, Hayes, & Hess, 1993; Kingdom & Keeble, 1996, 1998; Field, Hayes, & Hess, 2000) and the work of others (Nothdurft, 1985, 1991; De Bonet, 1997), we believe there are reasons that the 2nd-order combinatorial statistics (relations among pairs of wavelets) are relevant to texture discrimination. However, here we address the question of what aspects of the 1st-order combinatorial statistics — the wavelet-contrast histogram — are relevant to texture discrimination.

1.2. Positional variation blurs the wavelet-contrast histogram

The term 'wavelet-contrast histogram' has two meanings in the context of the present study. The first is as a description of the distribution (i.e. relative frequencies) with which various contrasts of a particular wavelet are chosen when synthesising an image. Because our wavelet textures comprise elements of many sizes and orientations, it is also possible to create textures in which different element types have different contrast histograms. The first meaning of 'wavelet-contrast histogram', therefore, describes the histogram or histograms used to synthesise an image (i.e. the input). The second meaning of the term is as a description of the output of a wavelet transform of an image. In this case, the histogram describes the distribution of the various response-strengths (contrasts) of a given wavelet when convolved with the image.

The distinction between the 'input' and 'output' forms of the wavelet-contrast histogram is an important one that does not arise with pixel textures, whose input and output histograms are identical (the pixel-luminance histogram describes the frequencies of all pixel intensities in the image). Indeed, generating an image by summing a set of randomly positioned functions drawn from a particular wavelet-contrast histogram is not likely to generate images with the same output histogram, even when a wavelet transform of the same type is used as both the input and output. Consider the trivial case where an image consists of a single micropattern (e.g. a single Gabor function) on a blank — DC — field, that is then analysed using a wavelet transform whose basis-function shape is matched to the micropattern. If the neighbouring wavelet basis functions are orthogonal (two functions are said to be orthogonal when their summed cross product is zero) and a matching wavelet function is placed precisely over the micropattern, then the output of the wavelet transform will be represented by a single active unit. In this case the input (one micropattern) with equal the output (one active unit).

However, usually there is no reason to expect any input element to fall precisely on a filter. Furthermore, from the point of view of the visual system, there is no reason to expect the visual system's receptive fields to be orthogonal. If the basis set is non-orthogonal (like the Gabor basis set used here), or the input micropattern does not align precisely with the wavelet transform, the set of wavelets that overlap with the input element in space, orientation, and spatial frequency, will be activated. The wavelet-contrast histogram that goes in will then not be the histogram that comes out. Rather, the energy of the single input element will be distributed over neighbouring wavelet filters. Thus, the output wavelet-contrast histogram will be, essentially, a blurred version of the histogram used to generate the texture.

This distinction between input and output wavelet-contrast histograms is particularly important when one considers the higher moments of the wavelet-contrast histogram. The higher moments (>2nd) capture the various ways in which a particular histogram can vary from a Gaussian distribution of a particular mean and variance. In our wavelet textures, we have concentrated on the 2nd, 3rd, and 4th moments of the Gabor wavelet-contrast histogram. These moments are manipulable in a relatively straightforward manner – by varying the contrast, phase and density of the component wavelets. What of, say, the 5th, 6th and higher moments? How might these statistics be manipulated without affecting the lower moments? Chubb, Econopouly, and Landy (1994) (see also, Chubb & Landy, 1991) developed a technique which influenced the present study. Using pixel-noise textures, they independently manipulated up to the 7th moment of the pixel luminance histogram by pre-defining each histogram shape as an orthogonal (Legendre) polynomial probability density function, and then randomly selected luminances from them. Thresholds were measured for the discrimination of textures with histogram shapes that differed in a given moment. One of the key differences between their stimuli and ours is that their texture elements — pixels — are discrete and orthogonal. It is possible to render differences in particular higher-order moments of wavelet textures (e.g. only the 5th or 6th moments) providing those wavelet elements are discretely positioned and the positions and shapes of the filters are precisely matched to the positions and shapes of the elements. However, as we point out above, the effect of randomly positioning the Gabor elements in our textures is to blur the output wavelet-contrast histogram. Higher-order histogram moments in the output are more vulnerable than lower-order moments, and so selecting micropattern contrasts from a higher-order polynomial function is ineffective.

It is important also to realize that the aim of Chubb et al. (1994) was to derive the point-wise, luminance, transducer function, employed by the visual system to process noise patterns, rather than compare the discrimination efficiency for texture pairs with different pixel-luminance histogram shapes. Their data did reveal, though, that the higher the moment, the more difficult the discrimination. However, without a comparison of human data to some measure of the amount of physical histogram information available to the cortex, the relative efficiency with which different moment statistics can be processed cannot be determined. Moreover, we argue that the histograms of activity available to the cortex will show little evidence of these higher-order moments.

1.3. Aim of study

In the present study we measured perceptual sensitivity to the 2nd, 3rd, and 4th moments of synthetic Gabor-micropattern textures, examples of which are shown in Fig. 1. We also measured discrimination sensitivity to these textures relative to model observers that used pixel histograms and wavelet histograms based on Gabor filters with the same properties as the elements.

Our textures were constructed from the addition of Gabor elements of various orientations, spatial scales, phases, and contrasts. However, the elements at different scales were derived from the same mother element, with different scales and orientations created by altering only the overall size and orientation. These textures embody many of the properties of stochastic, natural textures; they are broadband, and they have an amplitude spectrum that falls proportionately with spatial frequency ($1/f$) (see Field, 1994). We used the paradigm of Chubb et al. (1994), and measured sensitivity in texture discrimination as a function of differences in the shapes of the histograms. Manipulating the properties of the elements themselves — their contrasts, phases, and densities — altered both the pixel-luminance and wavelet-contrast histogram shapes.

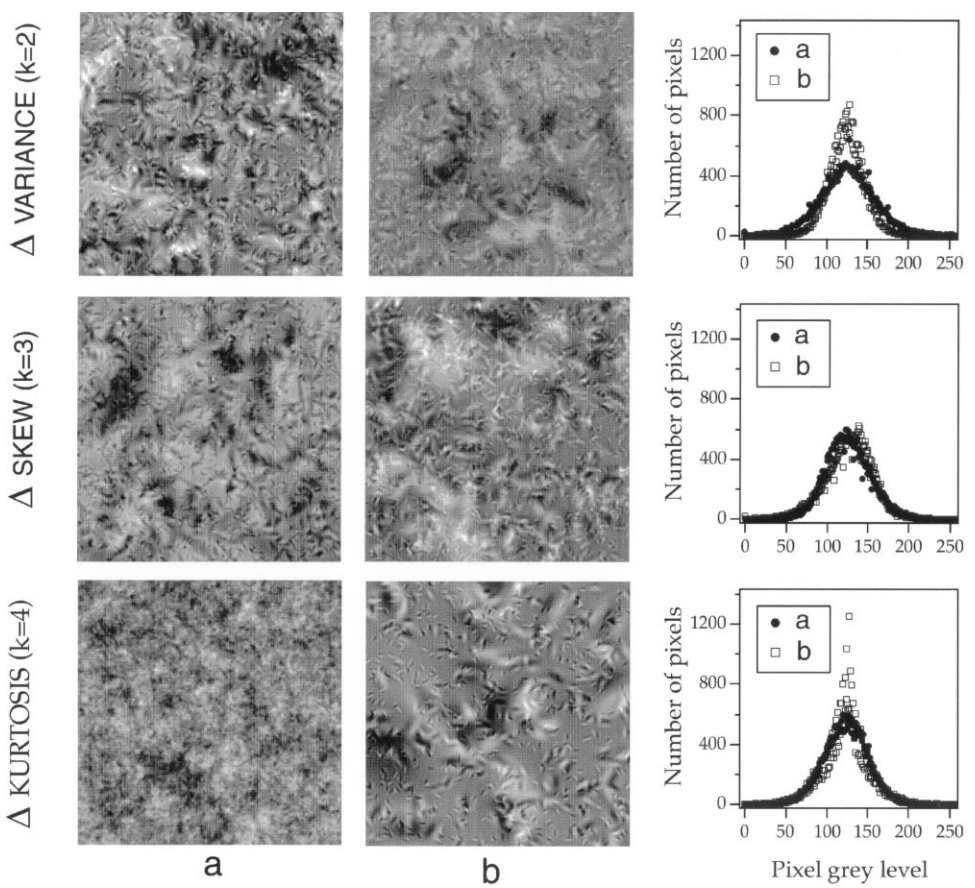

Fig. 1. Examples of wavelet textures used in the experiment, and their pixel luminance histograms. The top pair have the same mean luminance, but differ in r.m.s. contrast, and hence histogram variance. The middle pair are made from different phase Gabors; black-on-centre (left) and white-on-centre (right); their histograms have the same mean and variance, but differ in skew. The bottom pair have the same mean and r.m.s. contrast, but differ in wavelet density; their histograms have the same mean, variance, and skew, but differ in kurtosis. All textures have fractal, $1/f$, Fourier power spectra. Apart from the top two textures, the textures also have identical Fourier power spectra.

2. Methods

2.1. Stimuli

2.1.1. Generation

A VSG2/3F video-graphics card (Cambridge Research Systems), hosted by a Gateway 2000 P5 computer, generated the stimuli, which were displayed on a BARCO Calibrator monitor.

2.1.2. Calibration and contrast resolution

The VSG2/3F can display images with 256 intensity levels per RGB gun, selected from 12-bit (4096 levels) linearised colour look-up-tables (CLUTs). We separately calibrated the three RGB guns, and the CLUTs were generated such that when the three RGB 8-bit grey levels were set to be equal, the display was black-white. Calibration was carried out using the Cambridge Research Systems optical system, which generates the 12-bit γ-corrected CLUTs. The 12-bit CLUTs provided a contrast resolution of about 0.05% throughout the monitor range (see Kingdom & Whittle, 1996), which is sufficient for measuring contrast thresholds. Whatever the contrast of the stimulus, it was always displayed with the full 8-bits; the range of values was suitably selected from the 12-bit CLUTs.

2.1.3. Gabor elements

Multiplying a sine function by a two-dimensional Gaussian envelope generated the Gabor elements:

$$L(x, y) = M + A\cos[2\pi f(x\cos\theta - y\sin\theta) + \rho] \times \exp[-(x^2 + y^2)/2\sigma^2], \quad (1)$$

where M is the mean luminance, A is the amplitude, f is the luminance spatial frequency; θ is the orientation, ρ is the phase; and σ, is the S.D. of the Gaussian envelope (which was circularly symmetric). All Gabor elements had a spatial frequency bandwidth at half-height of 1.6 octaves and an orientation bandwidth at half-height of 60°. Contrast was defined as peak amplitude, A, divided by the mean, M. The mean luminance of the elements, M, and the background luminance, was set to 35 cd m^{-2}, unless otherwise specified. Full details of micropattern parameters are given with the description of each experiment.

2.1.4. Wavelet textures

Examples of wavelet textures are shown in Fig. 1. They were each 192 × 192 pixels, and subtended 4.3 × 4.3° at the viewing distance of 100 cm. Their generation was based upon the following function:

$$I(x,y) = \delta\sum_{\theta}^{12}\sum_{\rho}^{4}\sum_{f_s}^{2(s-1)} G(A, f,\sigma,\theta,\rho,x - xr,y - yr), \quad (2)$$

where G represents a Gabor micropattern (a wavelet) with parameters A, f, σ, θ, ρ, as defined above, whose x–y position in the texture was randomised, as given by xr and yr. In the standard condition the total number of wavelet orientations (θ) and phases (ρ) was 12 and 4, respectively. The 12 orientations were set at 30° intervals staring at 0° (vertical). The four phases were 0, 90, 180, or 270°. The total number of different wavelet spatial frequencies, f, was eight, and these were, in cycles per image, $f = 1.0$, 1.74, 3.02, 5.25, 9.14, 15.88, 27.61 and 48.0.

In order to create wavelet textures with a $1/f$ spectrum, the relative numbers of wavelets at each spatial frequency must be proportional to the square of the wavelet spatial frequency, assuming wavelets of equal bandwidth and contrast. Put another way, different spatial frequency wavelets must provide equal coverage. The number of wavelets at each spatial frequency was controlled by the term $R^{2(s-1)}(s = 1, 2,\ldots, 8)$, where s indexes the spatial frequency f_s, and R is the ratio between adjacent spatial frequency bands, which in our textures was 1.74. This formulation leads to equal coverage for each spatial frequency, since $R^{2(s-1)}/f_s^2$ is a constant.

Finally, the parameter δ controlled the overall number of elements in the texture, N, which from Eq. (2) is given by:

$$N = 48\delta\sum_{s=1}^{8} R^{2(s-1)}. \quad (3)$$

The texture we employed with the largest N, N_{max}, had a total number of wavelet elements equal to the number of pixels in the 192 × 192 array; i.e. 36 864. To produce this texture, δ was set to 0.221. Because $\delta < 1$, it follows from Eq. (2) that not all types of wavelet elements (eight spatial frequencies, each at 12 different orientations, at each of four phases) can be represented, since from Eq. (3), $N = 167\ 136$ for $\delta = 1$ and $R = 1.74$. We avoided potential problems arising from this under-representation by sub-sampling the orientations of the wavelet elements in a lawful manner. For example, in the case of $\delta = 0.221$ ($N_{max} = 36\ 864$), the number of wavelet elements at the lowest spatial frequency was 11. Given 11 wavelet elements, it was impossible to have exemplars of all 12 orientations, each at four phases. Here, the orientations were selected from the set of 12 orientations such as to maximise the range of orientations covered, and the phases were allocated randomly.

Rather than use δ, we employ the parameter D to define the density of the textures, where $D = N/N_{max}$, expressed as a percentage. Thus the texture with N_{max} wavelet elements, shown in the bottom left panel of Fig. 1, has $D = 100\%$. All the other textures shown in Fig. 1 have $D < 100\%$.

The contrast of our wavelet-textures is best defined as the r.m.s. of the luminances of the image divided by the mean image luminance (Moulden, Kingdom & Gatley, 1990), and for all conditions except the test conditions

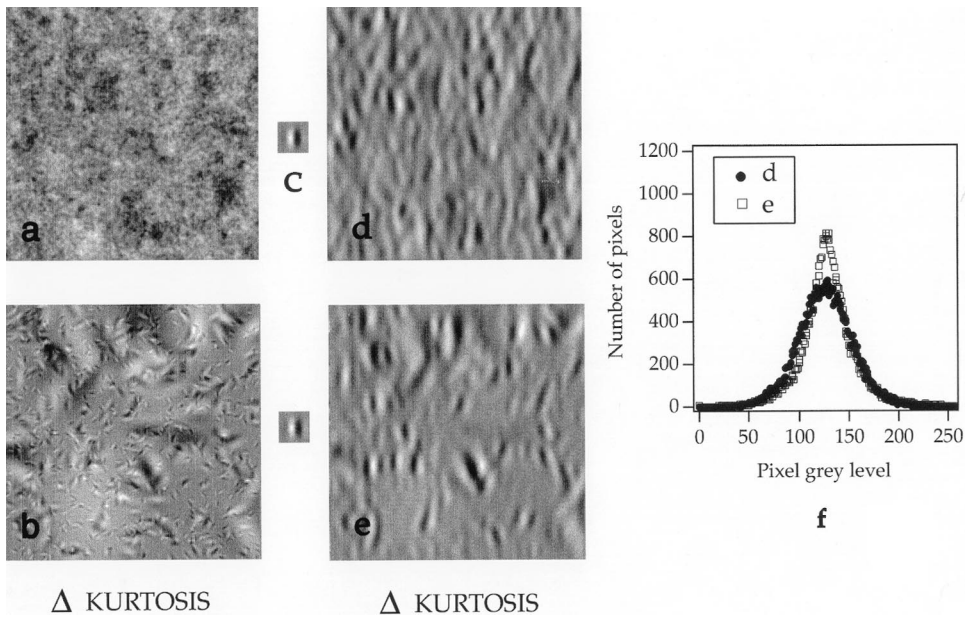

Fig. 2. Two wavelet textures, a and b, have the same variance (or r.m.s. contrast) but different numbers of wavelets (i.e. different wavelet densities), and hence different kurtosis. Texture a has a density D of 100%, i.e the same number of wavelets as pixels, while texture b has a density D of 2%, i.e. only 1/50th the number of wavelets as pixels. Both textures have been convolved with a vertically-oriented Gabor filter (c), to produce outputs (d) and (e). The histograms of pixel intensities in d and e are shown in f, and these constitute the wavelet-contrast histograms of the two textures. The difference in kurtosis between the pixel-luminance histograms of (a) and (b) (not shown) is reflected in the difference in kurtosis in the wavelet-contrast histograms of (d) and (e).

in the variance discrimination experiment, contrast was 22%, with a S.D. due to stochastic variation of 1.2%. We set the contrast of each texture by suitable mapping of the 8-bit grey-level image generated in computer memory onto the appropriate values in the linearised 12-bit display CLUTs.

2.1.5. Histogram statistics

In the experiments described below, observers were required to discriminate between pairs of wavelet textures that differed in the moments of both their pixel-luminance and the calculated output wavelet-contrast histograms. For the pixel-luminance histograms, moments were calculated from the distribution of pixel values. For the wavelet-contrast histograms, moments were calculated from the distribution of pixel values in the convolution output of a wavelet kernel with the texture.

The kth moment, M, of a histogram is defined as:

$$M_k = 1/N\sum_{i}^{N}(x_i - \mu)^k / M_2^{k/2} \quad (4)$$

where μ is the mean of the distribution, and N is the sample size. Variance is computed when k is set to 2, skew when k is set to 3, and kurtosis when k is set to 4. The term $M_2^{k/2}$ normalises the statistic.

The variance, skew, and kurtosis, of the pixel-luminance and wavelet-contrast histograms were varied as follows. Variance was manipulated by changing the contrast of the wavelet texture (see Section 2 for details). Skew was manipulated by varying the phase of the wavelet elements. In the experiments involving the manipulation of skew, only one phase of wavelet micropattern was present in each texture, unlike in all other experiments where four phases were present in equal proportions (0, 90, 180 and 270°). The two extremes of skew were achieved by, on the one hand, using bright-on-centre even-symmetric wavelet elements ($\rho = 0$), on the other hand, dark-on-centre even-symmetric wavelet elements ($\rho = 180$); intermediate skews were achieved by intermediate phases. Because the presence of skew slightly altered the mean luminance of each texture, we applied DC restoration for each texture. There are several ways to alter skew, and we have selected one that is relatively simple. Finally, kurtosis was manipulated by varying D, the density of wavelet elements; importantly, while holding the variance constant.

Fig. 1 shows example wavelet textures that differ in variance ($k = 2$), skew ($k = 3$), and kurtosis ($k = 4$), in the ways described above. The pixel histograms of each pair of textures are shown on the right-hand side of each texture pair. Fig. 2 shows two wavelet textures differing in kurtosis that have been convolved with a Gabor wavelet. The pixel histograms of both convolution images are also shown, and these histograms constitute the wavelet-contrast histograms.

2.2. Procedure

A bias-free 2IFC (two interval, forced choice) procedure was employed with a 'two-up, one-down' adaptive staircase that established the threshold difference in the variable of interest at the 70.7% correct level. This corresponds to a d'_e (e stands for 'empirical') measure of 0.78 for a 2AFC task (Macmillan & Creelman, 1991), and this was the value used in all measures of efficiency (see below). In the variance-discrimination condition, observers were required to respond to the interval with the highest contrast. In the skew condition, observers were required to respond to the interval in which the texture had more white and/or less black 'flecks'. In the kurtosis discrimination condition, observers were required to respond to the texture that was sparser. Feedback — a short tone — was given for an incorrect decision.

2.3. Model observer analysis

The purpose of the main experiment was to assess the ability of observers to discriminate texture pairs that differed in the variance, skew and kurtosis of their pixel-luminance and output wavelet-contrast histograms. We compared human performance with that of two model observers, whose performances were based on the appropriate histogram moment. In the first model, the histogram was of pixel luminance values, and we term this the pixel-luminance-histogram model. In the second model, the histogram was of the pixel values in the convolution response of a wavelet kernel to the texture, in other words the output wavelet-contrast histogram. We term this the wavelet-contrast-histogram model. The kernel was matched in profile to that of the vertically oriented, highest spatial frequency, wavelet micropattern used in the textures. Example convolution responses and their associated histograms are shown in Fig. 2.

Model-observer performance was calculated as follows. For each condition, a large number of test and comparison texture pairs at the empirically determined threshold level were generated, and the appropriate statistic (variance, skew, or kurtosis) calculated for each member of the pair, according to Eq. (4). The mean and S.D.s of the statistic for both test and comparison were then calculated over the whole sample. If the mean values of the statistic k for the test and comparison stimuli are μ_{tk} and μ_{ck}, and their S.D.s are σ_{kt} and σ_{kc}, then, assuming that the statistics are normally distributed over the sample (as expected by the central limit theorem), the performance of a model-observer, who chooses the interval for each texture pair which gives the higher (or lower) value of the statistic, can be described by the signal-detection-measure d'_m (m = model observer), where:

$$d'_m = \frac{\mu_{tk} - \mu_{ck}}{\sqrt{[(\sigma^2_{k.a} + \sigma^2_{k.b})/2]}}. \quad (5)$$

To illustrate this procedure, Fig. 3 shows the distributions of kurtosis for two textures with densities 3.0 and 2.5%. The $D = 3\%$ texture was one of the three test condition densities (see below), and the $D = 2.5\%$ texture was the one just discriminable from the test when kurtosis was the test variable. The kurtosis values in the top figure have been calculated from the pixel-luminance histograms, while those in the bottom figure have been calculated from the wavelet-contrast histograms., as described above. Best fitting Gaussian functions have been fitted to each distribution, and show that the distributions are very close to normal. Model observer performance calculated using Eq. (5) essentially provides an estimate of the separation between each pair of distributions-the difference between their means divided by a measure of their average standard deviation. Note how the distributions of kurtosis values derived from

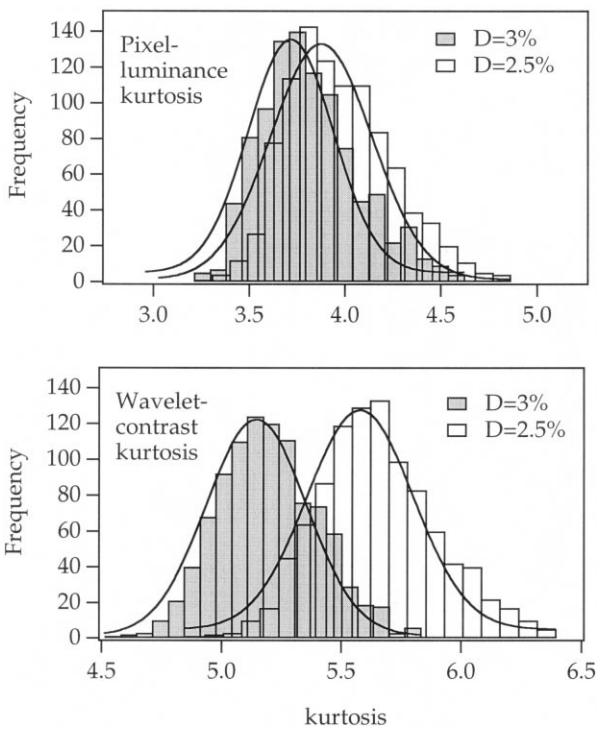

Fig. 3. Distributions of kurtosis for two wavelet textures whose densities D (3 and 2.5%) are just discriminable. Each distribution is based on 1000 sample textures. The kurtosis values in the top figure have been calculated from the pixel-luminance histograms of the raw textures. The kurtosis values in the bottom figure have been calculated from wavelet-contrast histograms (e.g. see Fig. 2), using a kernel matched to the vertically-oriented, highest-spatial-frequency wavelet used in synthesising the textures. Best fitting Gaussian functions have been fitted to each distribution, and show that the distributions are very close to normal. The performance of a model observer, d'_m, is given by the difference between the means of each pair of distributions divided by a measure of their average standard deviation (see Eq. (5)). Note how a bigger d'_m is provided by the wavelet-contrast histograms.

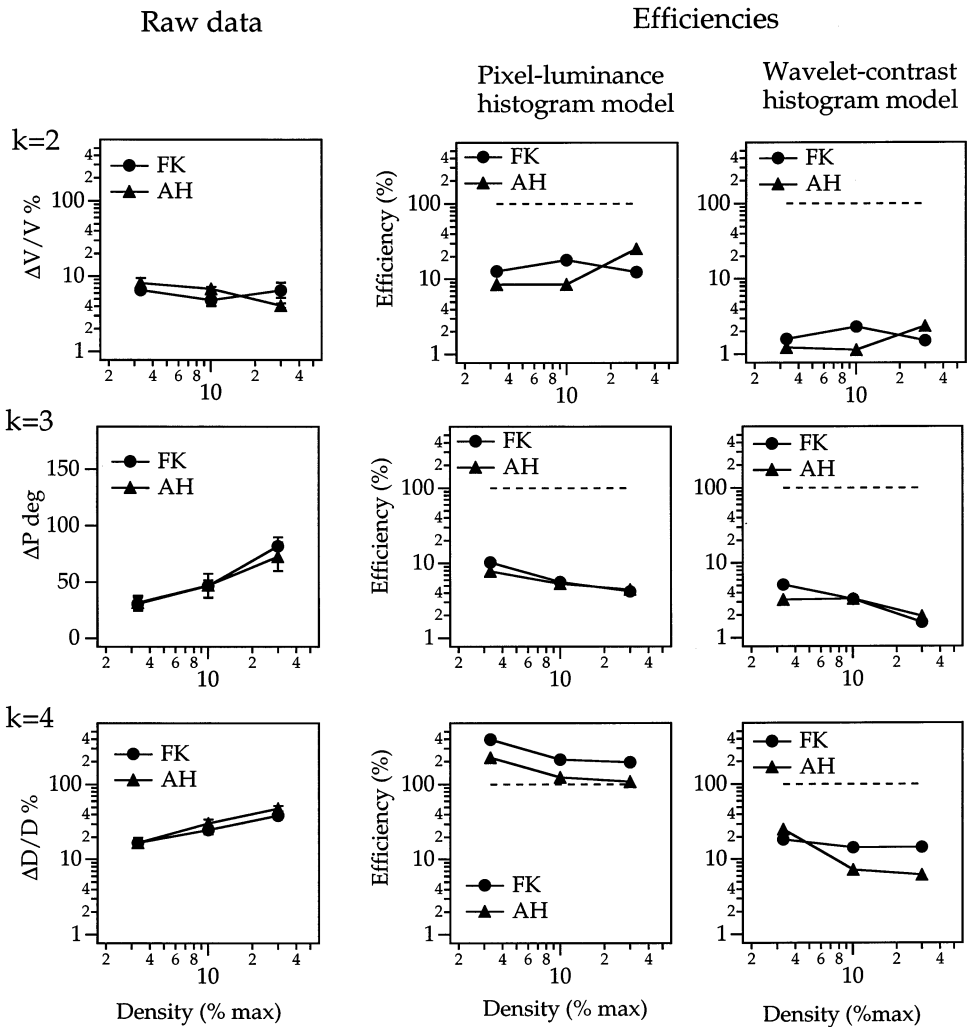

Fig. 4. Results of experiment. Raw data are shown in the left hand column, and efficiencies with respect to both pixel-luminance histogram, and wavelet-contrast histogram model observers, are shown on the right. Top, contrast, or variance ($k = 2$), discrimination; middle, phase, or skew ($k = 3$), discrimination; bottom wavelet density, or kurtosis ($k = 4$), discrimination. For further details see text.

the wavelet-contrast histograms have the greater separation.

Efficiency of human performance with respect to this model observer, F_m, is given by:

$$F = 100(d'_e/d'_m)^2\ \%, \tag{6}$$

where d'_e is the empirically measured value of human sensitivity, equal to 0.78, the value of d' for a 2AFC task at the 70.7% correct level (Macmillan & Creelman, 1991).

3. Results

The results for two observers are shown in Fig. 4. In all figures the data are shown for three wavelet densities D (3.0, 10 and 30%), where density is defined as the number of wavelets expressed as a percentage of the total number of pixels (36 864). Both observers performed similarly on all tasks. The left hand column in Fig. 4 shows the raw data. The top, $k = 2$ (variance), and bottom, $k = 4$ (kurtosis), discrimination results are expressed as Weber fractions (the difference between test and comparison levels at threshold, divided by the comparison level), while the middle, $k = 3$ (skew), data are expressed as the difference in Gabor-sinusoid phase between the test and the comparison textures at threshold.

As the figure shows, variance (contrast) discrimination thresholds were more-or-less constant at around 6% across the nearly 10-fold change in density; skew (phase) discrimination thresholds increased from about 30 to 75°; kurtosis (density) discrimination thresholds increased from about 17 to 43%. Although our textures are made from specified numbers of wavelets, the high degree of wavelet overlap introduces extrinsic noise into any internal representation of the texture's properties. This extrinsic noise makes the absolute discrimination results for variance ($\Delta V/V$), skew (ΔP), and kurtosis ($\Delta D/D$), not very useful for comparison of performance

between these three statistics. However, a dimensionless measure of efficiency is useful. The middle column of Fig. 4 plots the efficiency, F (see Section 2), with which observers performed the task, compared to a model observer who computed the variance (Fig. 4, top), skew (Fig. 4, middle), and kurtosis (Fig. 4, bottom), of the pixel luminance histogram. While efficiency is less than 100% for discriminating variance and skew in the pixel luminance histograms of wavelet textures, it is greater than 100% for the discrimination of kurtosis — the moment that was manipulated by varying D. This result shows that a pixel-luminance-histogram model is invalid as a description of the basis for human performance for the density discrimination data; efficiencies, by definition, cannot exceed 100%. Although we cannot rule out the possibility that variance and skew in the pixel-luminance histogram underlies the basis for discriminating the top and middle texture pairs shown in Fig. 1, we can rule out this possibility for the bottom pair, which differ in kurtosis. This $>100\%$ efficiency result shows that the visual system is sensitive to some property of our wavelet textures that the pixel histogram statistics fail to capture.

Consider now the right hand column in Fig. 4, where efficiencies have been normalised to a model that computes the statistics of a wavelet-filtered set of stimuli, the wavelet-contrast-histogram model. The filter's convolution kernel was a vertical Gabor matched to that of the highest spatial-frequency Gabor micropattern used in making the textures ($s = 8$). Fig. 2 shows the result of convolving two textures with $D = 100\%$, and $D = 2\%$, and the resulting pixel histograms (see also Fig. 3). As can be seen in Fig. 4, efficiency is now less than 100% for all conditions, though the ordering of efficiency, kurtosis > variance > skew, is preserved. Although our highest-frequency kernel captures more of the information in the texture than any of the lower-frequency kernels, because there are more highest-frequency wavelet micropatterns in the texture than any other, it obviously does not capture all the information. The plotted efficiencies are therefore greater than they would be if a model were used that performed a complete wavelet decomposition of the textures. Such a model would calculate the statistics from the response histograms of a set of quasi-orthogonal wavelet filters spanning the full range of orientations and spatial frequencies of the textures. A model based on a full wavelet transform would however, at least for the images we used, result in efficiencies that preserved the order found for the single-scale wavelet transform. Our single-scale exercise is instructive, therefore, only in so far as it demonstrates that a model-observer based on a wavelet transform is inherently more plausible than one based on the pixel histogram (as only the former produces efficiencies consistently less than 100%), and that the wavelet-histogram kurtosis of our textures appears to be more efficiently discriminated than is either the variance or the skew.

4. Discussion

The goal of the study was less to understand the mechanisms for discriminating textures, than to examine the case for the importance of higher-order statistics in texture discrimination. We showed that for a particular class of textures, human observers were at least as sensitive, and in some cases more sensitive, to the higher histogram moments compared to the lower moments. In addition, we showed that whereas a model of human performance based on the pixel-luminance histograms of our textures was insufficient (at least for discriminating their density, or kurtosis), a model based on the wavelet-contrast histogram was sufficient. This in no way implies that other basis functions besides wavelets would not allow equally valid models. However, it does demonstrate that the visual system is sensitive to the local pairwise correlations in our textures that are present by virtue of their wavelet composition, correlations that a wavelet transform captures but that a pixel description ignores.

Could our textures, differing ostensibly in only skew or kurtosis, be discriminated on the basis of their lower-order statistics? Chubb and Yellott (2000) have recently claimed that any discreet image has a unique dipole histogram; that is it has unique Julesz 2nd-order statistics. However, as Chubb and Yellott are at pains to point out, this may not be relevant to perception. At the very least it seems implausible to us that the visual system would go to the considerable effort of extracting the higher moments from histograms of 2nd-order combinatorial relationships, when in principle those higher moments were available in the 1st-order combinatorial histograms.

We are not, of course, the first to demonstrate sensitivity to higher-order statistics in textures. We have already said in the Introduction that Julesz and his co-workers published examples of textures discriminable purely on the basis of their higher-order (Julesz) statistics. More recently, Malik and Perona (1990) showed that for textures composed of even-symmetric Gabor elements, texture pairs made from opposite-phase Gabors were easily discriminable, as is the case for the middle texture pair in Fig. 1. Graham (1991) and Graham, Beck, and Sutter (1992) have shown that in non-stochastic textures consisting of regularly-spaced, identically-oriented Gabor elements, segregation occurs when one part of the texture consisted of Gabor elements arranged in vertical columns, the other in diagonal columns. Presumably, the two parts of these 'second-order' textures have similar lower-order statistics, and therefore segregation occurs on the basis

of higher-order statistics. Of particular relevance to the present study are also the findings of Durgin and coworkers (Durgin, 1995; Durgin & Huk, 1997), who have demonstrated that the visual system can be selectively adapted to different texture densities, which implies the existence of mechanisms that are specialised for encoding texture density.

What is new in our study is that we have demonstrated comparable, or even higher, sensitivities to higher-order statistics in a novel class of textures that embody many of the properties of naturally occurring textures. Moreover, because the statistical properties of our wavelet textures can be manipulated along a continuum (wavelet-contrast, -phase, and -density, are all continuous variables), our stimuli are potentially useful for addressing a wide range of issues in texture discrimination.

4.1. Current texture models

Which of our wavelet textures would be discriminated by current models of texture segregation? The most influential current class of texture model is the filter-rectify-filter model (see Wilson, 1999 for a recent review). Fig. 5 shows the principle by which a filter-rectify-filter model distinguishes two natural textures (for details of particular model implementations see, for example, Fogel & Sagi, 1989; Malik & Perona, 1990; Bergen & Landy, 1991; Chubb & Landy, 1991; Landy & Bergen, 1991; Rubenstein & Sagi, 1993). Each texture is initially convolved with a '1st-stage' Gabor-kernel; full-wave rectification, then low-pass filtering to smooth out the high-frequency ripples in energy, follows. Full-wave rectification would presumably be achieved physiologically by carrying the excitatory and inhibitory responses in separate on- and off-centre pathways. The final intensity maps reveal that the bottom texture in Fig. 5 has more energy at the chosen filter scale/orientation than the top texture. The process is repeated (in parallel) for other filter scales/orientations. A '2nd-stage' filter is typically employed to detect any difference in energy between each intensity map when the two textures abut. This stage is not shown here, since our textures are non-abutting. Although this model, at least in its basic form in Fig. 5, would distinguish between our wavelet textures which differed in variance, it would not distinguish between those differing only in either skew or kurtosis, because any differences in these statistics are not represented in the mean filter-energy response.

Our results, showing relatively good efficiency at kurtosis discrimination, argue for an approach to modelling texture segregation which incorporates sensitivity to skew, and kurtosis. Further support comes from the

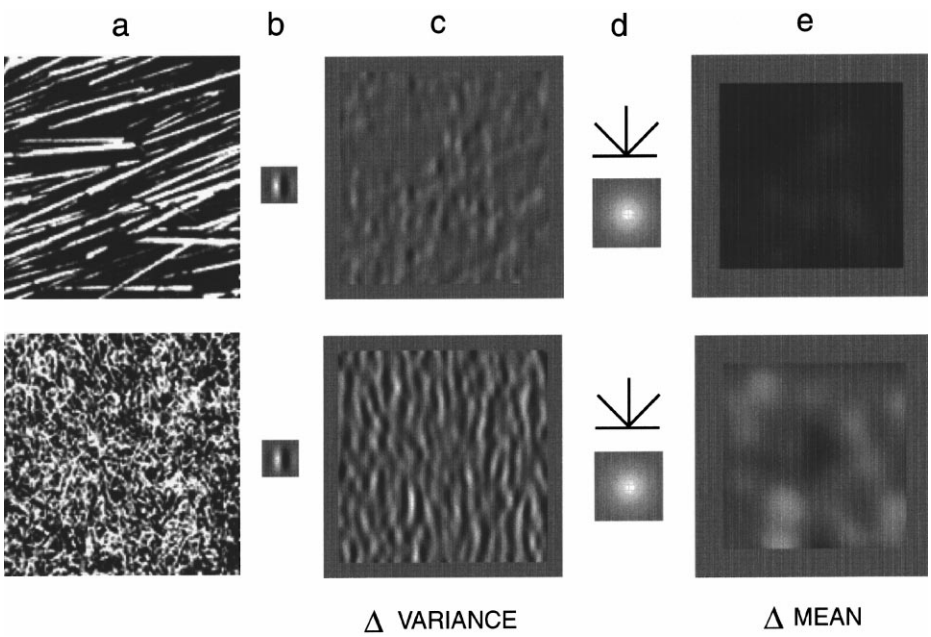

Fig. 5. Conventional filter-rectify-filter model of texture discrimination. The two textures in (a) are filtered by a '1st-stage' vertically-oriented Gabor kernel (b) to produce (c); (c) is full-wave rectified and then smoothed by a Gaussian filter (d), producing the image in (e) whose average intensity corresponds to the amount of contrast energy in the texture at the 1st-stage filter's scale/orientation. Note that the lowpass filter is not the '2nd-stage' filter of conventional filter-rectify-filter models, which employ 2nd-stage filters to detect the borders between intensity maps. Note also that the effect of rectification and smoothing in (d) is to transform the difference in variance in (c) to the difference in mean intensity in (e). The procedure shown here is repeated for other 1st-stage scales/orientations, and the gamut of energy-intensity maps (or the signals elicited by the borders between abutting texture-pairs) used in the decision process. This model however would not distinguish between our wavelet textures that differ in either skew or kurtosis, since they would have near-identical energy maps at all orientations and scales. The natural textures here and in other figures are taken from the Brodatz album (Brodatz, 1966).

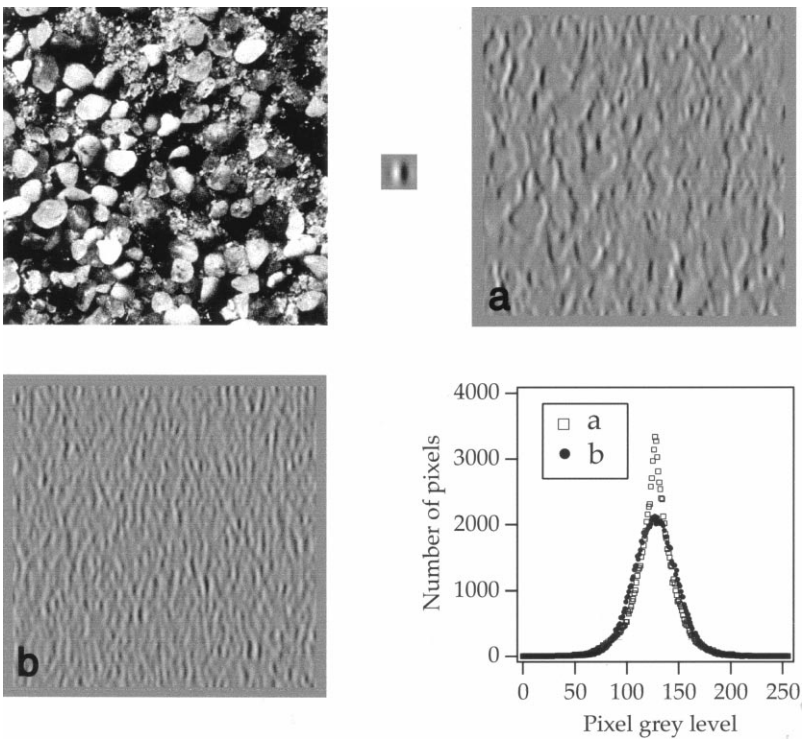

Fig. 6. The importance of kurtosis. The natural texture (top left) is filtered by a vertical Gabor kernel to produce (a). (b) is produced by randomly-positioning thousands of the same kernel, and then equalizing the variance to that of (a). The histograms of (a) and (b) are shown on the bottom right, and show a marked difference in kurtosis due to the abundance of local structure in the natural texture. The conventional filter-rectify-filter model in Fig. 5 would fail to distinguish between (a) and (b).

simple demonstration shown in Fig. 6, which illustrates how relatively high kurtosis (relative to the Gaussian) is a ubiquitous property of wavelet-filtered natural textures. High kurtosis in natural textures reflects the presence of higher-order statistical structure primarily due to the concentration of energy at edges (in Fig. 6 stones), with less energy in between (Morrone & Burr, 1988; Field, 1994; Thomson, 1999).

4.2. How might skew and kurtosis be discriminated?

Although our wavelet textures that differ in skew and kurtosis cannot be discriminated using the basic filter-rectify-filter model shown in Fig. 5, the introduction of a simple thresholding non-linearity would make discrimination possible. For skew, a threshold imposed on the response of an even-symmetric, say 'on-centre', filter could convert two images that differ in skew into two images that differ in contrast energy, thus allowing a simple filter-rectify-filter model to discriminate the textures. In the filter-rectify-filter model of Malik and Perona (1990), thresholding was combined with the use of half-wave rectification as the intermediate non-linearity; the positive and negative components of the filtered image were separately represented as two positive images, and then each was thresholded. Using this model, textures with positive and negative skew would be differentially represented in the two half-wave rectified images, prior to any 2nd stage of filtering. A simple threshold could also be used to discriminate images with different kurtosis. For an image with high kurtosis, the high magnitude local structure will get through the threshold; for an image with low kurtosis, and the right choice of the threshold, little or no image structure would pass the threshold. It would be necessary to normalise the filtered images by their average amount of energy to avoid the confounding effects of overall image contrast, but with such a proviso, a simple filter-rectify-filter model could easily discriminate between low and high kurtosis.

On the other hand, rather than manipulate the form of the intermediate non-linearity, an extension of the filter-rectify-filter model might be employed. Fig. 7 shows the idea. The two equal-in-variance wavelet-filter response patterns in Fig. 7a are full-wave rectified, and then subjected to a 2nd-stage of bandpass filtering, using a filter about twice the size of the 1st stage filter. The 2nd stage filter output is then subjected to the same steps of full-wave-rectification-plus-lowpass-filtering as was the 1st stage response in the filter-rectify-filter model shown in Fig. 5. The more kurtotic texture in Fig. 7 produces a higher-in-luminance 2nd stage energy-intensity map.

If one wished to instantiate a segmentation process based on a difference in kurtosis, one could use a 3rd stage filter to detect the border between the two inten-

sity maps shown in Fig. 7e. One could then term this a 'filter-rectify-filter-rectify-filter' model. Essential for this model to succeed would be the removal by normalisation of any differences in overall energy, or variance, resulting from the 1st stage, as we have done in constructing the images shown in Fig. 7a.

This extension of the basic filter-rectify-filter model has much in keeping with those models where the purpose of the 2nd-stage filter is to capture the pattern of contrast energy within a texture region. Such models have been advanced to account for the segmentation of so-called 'second-order' textures (Graham, 1991; Sutter & Graham, 1995; Graham & Sutter, 1998), and the detection of textures that are sinusoidally modulated in the critical texture variable (e.g. for contrast modulation, Sutter, Sperling, & Chubb (1995); for orientation modulation, Kingdom & Keeble (1996) and Gray & Regan (1998); for spatial frequency modulation, Arsenault, Wilkinson, & Kingdom (1999)).

4.3. An alternative approach to discriminating wavelet textures

The above approach to discriminating the variance, skew, and kurtosis, in our wavelet textures is underpinned by the idea that texture discrimination is a relatively primitive process, employing mainly feedforward, and relatively simple, filtering mechanisms. An alternative approach is that some explication of image features precedes texture analysis, an old idea central to the thinking of Julesz (1981) and Beck (1982), and implicit in the notion of Marr (1982) of the primal sketch. Kingdom and Keeble (2000) and Kingdom and Hayes (2000) have recently argued for such an approach to account for the detection of sinusoidally modulated textures. In this approach, the feature content of the elements comprising textures — their orientation, contrast, size, blur, and colour, etc. — is explicitly derived prior to texture analysis. Feature analysis may well involve combining information from a number of 1st-stage filters tuned to various spatial frequencies and orientations in the same receptive field location. Texture analysis would then proceed with the extraction of information about the pattern, or statistics, of each attribute within the texture, possibly requiring selective attention to the relevant attribute. Inspection of the textures in Fig. 1 may provide clues as to which feature attributes are salient. For variance, the contrasts of features would presumably be the critical attribute. For skew it might be the relative incidences of small-dark versus small-bright features in the texture. For kurtosis, a number of features might be used. As the textures become more sparse, the contrast of individual features increases if the overall r.m.s. is held constant. In addition, there is an increased amount of featureless space between the elements. Both these changes in feature content could, in principle, be represented in a primal sketch description of wavelet tex-

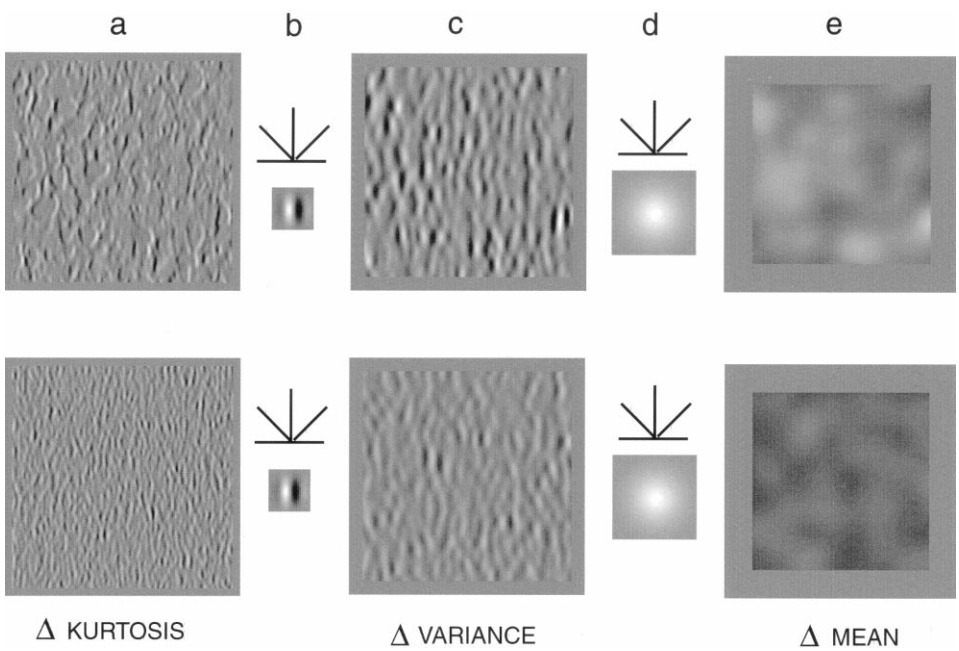

Fig. 7. How a difference in kurtosis in the filtered image might be recovered using an extension of the conventional filter-rectify-filter model. The filtered image in (a) is first rectified, then re-filtered by a '2nd-stage' bandpass filter, (b), twice the scale of the 1st-stage filter (not shown) which produced the images in (a). The result is an image that differs in variance rather than kurtosis. This difference in variance could then be detected by an additional Filter-Rectify-Filter procedure, that is an additional stage of rectification-plus-smoothing (d), to produce the difference-in-mean-intensity images in (e). However, this feed-forward energy model is only one possible way of discriminating textures with equal variance but different kurtosis, and in the text we argue for a more feature-based approach.

tures, and this description used by the visual system to discriminate the texture pairs.

4.4. Do we need to be sensitive to higher-order statistics?

It could be argued that while we are sensitive to the higher-order statistics in textures, it is sufficient to use lower-order statistic differences for texture discrimination, because the gamut of naturally occurring textures invariably differ in at least their 2nd-order statistics. First, in some naturally occurring textures, the higher-order statistics may be the most salient for discrimination; our data showing that higher-order statistics in our wavelet textures are more efficiently processed than power-spectral statistics (2nd order) is consistent with this possibility. Second, variations in density in the retinal image projection of a textured surface typically accompany changes in the depth and orientation of the surface; density is potentially a 'shape-from-texture' cue. Third, the study of texture perception is not reducible to the question of the discriminability of naturally occurring texture pairs. There is also the other side of the coin — the issue of identicality. Under what conditions are two textures perceived to be one and the same? Consider for example grass viewed through mist as opposed to viewed on a clear day; or viewed from two different angles. In both situations, the different views of the grass would result in different lower-order image statistics (the one viewed through mist would be of lower contrast, the one viewed from one angle would have different orientation statistics; both differences would be revealed as differences in the power spectrum). Yet we may want to identify the grass as the same in all situations, and for this purpose the similarity in higher-order statistics could be used. This, as was discussed in the Introduction, is the question which models of texture synthesis are primarily concerned.

5. Conclusion

We have shown that the variance of the wavelet histogram of a texture is insufficient for discrimination between textures — we also need to consider the 3rd and 4th moments. It is possible that we also need to consider moments higher than the 4th. However, positional variance and the lack of perfect matching between filters and elements will tend to blur out higher-order moments, making it unlikely that moments above the 4th are useful.

In this paper we limited ourselves to 1st-order Julesz statistics (i.e. 1st-order wavelet histograms). However, we believe there are conditions of 2nd-order relationships — over position (e.g. Field et al., 1993; Kingdom and Keeble, 1996, 1998), and through scale (e.g. Morgan, Ross, & Hayes, 1991; Hayes, 1994, 1998) — that are important. For example many textures comprise contours (e.g. hair), and these are presumably salient features for discrimination. The demonstrations of De Bonet and Viola (1997) also illustrate the importance of relationships between vectors for texture discrimination.

Thirty years ago Julesz proposed that 1st- and 2nd-order statistics of relationships among pixels may provide a complete description of what is pertinent to texture discrimination. This proposal failed. However, we may find that 1st- and 2nd-order combinatorial relationships within wavelet histograms may be all that is needed for texture discrimination.

Acknowledgements

This research was funded by a Canadian NSERC (ref: OGP 01217130) and MRC # MT11554 grants to FAAK, by Australian Research Council grant # A79531386 to AH, and by NIH grant # MH50588 to DJF.

References

Arsenault, S. A., Wilkinson, F., & Kingdom, F. A. A. (1999). Modulation frequency and orientation tuning of second-order texture mechanisms. *Journal of the Optical Society of America A*, *16*, 427–435.

Beck, J. (1982). Textural segmentation. In J. Beck, *Organization and representation in perception*. Hillside, NJ: Erlbaum.

Bergen, J. R. (1991). Theories of visual texture perception. In D. Regan, *Vision and visual dysfunction*, vol. 10B. New York: Macmillan.

Bergen, J. R., & Landy, M. S. (1991). Computational modelling of visual texture segregation. In M. S. Landy, & J. A. Movshon, *Computational models of visual processing* (pp. 253–271). Cambridge, MA: The MIT Press.

Brodatz, 1966. Textures. Dover Publications, New York.

Chubb, C., & Landy, M. S. (1991). Orthogonal distribution analysis: a new approach to the study of texture perception. In M. S. Landy, & J. A. Movshon, *Computational models of visual processing* (pp. 291–301). Cambridge, MA: MIT Press.

Chubb, C., & Yellott, J. I. (2000). Every discrete, finite image is uniquely determined by its dipole histogram. *Vision Research*, *40*, 485–492.

Chubb, C., Econopouly, J., & Landy, M. S. (1994). Histogram contrast analysis and the visual segregation of IID textures. *Journal of the Optical Society of America A*, *11*, 2350–2374.

De Bonet, J.S., 1997. Multiresolution sampling procedure for analysis and synthesis of texture images. Computer Graphics, 361–368 (ACM SIGGRAPH).

De Bonet, J. S., & Viola, P. (1997). A nonparametric multi-scale statistical model for natural images. In M. Jordan, & M. Perrone, *Advances in neural information processing*, vol. 10. Cambridge, MA: MIT Press.

Durgin, F. H. (1995). Texture density adaptation and the perceived numerosity and distribution of texture. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 149–169.

Durgin, F.H., & Huk, 1997. Texture density after-effects in the perception of artificial and natural textures. *Vision Research*, *37*, 3273–3282.

Field, D. J. (1994). What is the goal of sensory coding? *Neural Computation*, *6*, 559–601.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local 'association field'. *Vision Research*, *33*, 173–193.

Field, D. J., Hayes, A., & Hess, R. F. (2000). The roles of polarity and symmetry in the perceptual grouping of contour fragments. *Spatial Vision*, *13*, 51–66.

Fogel, I., & Sagi, D. (1989). Gabor filters as texture discriminators. *Biological Cybernetics*, *61*, 103–113.

Gagalowicz, A. (1981). A new method for texture fields synthesis: some applications to the study of human vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *3*(5), 520–533.

Graham, N. (1991). Complex channels, early local nonlinearities, and normalization in perceived texture segregation. In M. S. Landy, & J. A. Movshon, *Computational models of visual processing* (pp. 291–301). Cambridge, MA: MIT Press.

Graham, N., & Sutter, A. (1998). Spatial summation in simple (Fourier) and complex (no-Fourier) texture channels. *Vision Research*, *38*, 231–257.

Graham, N., Beck, J., & Sutter, A. (1992). Nonlinear processes in spatial-frequency channel models of perceived texture segregation: effects of sign and amount of contrast. *Vision Research*, *32*, 719–743.

Gray, R., & Regan, D. (1998). Spatial frequency discrimination and detection characteristics for gratings defined by orientation texture. *Vision Research*, *38*, 2601–2617.

Hayes, A. (1994). Multi-resolution local-phase coding of image structure: an explanation of the 'blocked Lincoln' effect. *The Australian Journal of Psychology*, *46*, 13.

Hayes, A. (1998). Phase filtering' disrupts continuity through spatial scale of local image-structure. *Perception*, *27*, 81.

Heeger, D. J., & Bergen, J. R., 1995. Pyramid-based texture analysis/synthesis. Computer Graphics Proceedings, 229–238.

Julesz, B. (1981). Textons, the elements of texture perception and their interactions. *Nature*, *290*, 91–97.

Julesz, B., Frisch, H. L., Gilbert, E. N., & Shepp, L. A. (1973). Inability of humans to discriminate between textures that agree in second-order statistics-revisited. *Perception*, *2*, 391–405.

Kingdom, F. A. A., & Hayes, A. (2000). Mechanism independence for texture-property modulation. *Investigative Ophthalmology and Visual Science Supplement*, *40*, 317.

Kingdom, F. A. A., & Keeble, D. R. T. (1996). A linear systems approach to the detection of both abrupt and smooth spatial variations in orientation-defined textures. *Vision Research*, *36*, 409–420.

Kingdom, F. A. A., & Keeble, D. R. T. (1998). On the mechanism for scale invariance in orientation-defined textures. *Vision Research*, *39*, 1477–1489.

Kingdom, F. A. A., & Keeble, D. R. T., 2000. Luminance spatial frequency differences facilitate the segmentation of superimposed textures. *Vision Research*, *40*, 1077–1087.

Kingdom, F. A. A., & Whittle, P. (1996). Contrast discrimination at high contrasts reveals the influence of local light adaptation on contrast processing. *Vision Research*, *36*, 817–829.

Landy, M. S., & Bergen, J. R. (1991). Texture segregation and orientation gradient. *Vision Research*, *31*, 679–691.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: a user's guide*. Cambridge University Press: Cambridge.

Malik, J., & Perona, P. (1990). Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America A*, *7*, 923–932.

Marr, D. (1982). *Vision: a computational investigation into the human representation and processing of visual information*. San Francisco, CA: W.H. Freeman.

Morgan, M. J., Ross, J., & Hayes, A. (1991). The relative importance of phase and amplitude in patchwise image reconstruction. *Biological Cybernetics*, *65*, 113–119.

Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: a phase-dependent energy model. *Proceedings of the Royal Society of London Series B*, *235*, 221–245.

Moulden, B., Kingdom, F., & Gatley, L. (1990). The standard deviation of luminance as a metric of contrast random-dot images. *Perception*, *19*, 79–101.

Nothdurft, H. C. (1985). Sensitivity for structure gradient in texture discrimination tasks. *Vision Research*, *25*, 1957–1968.

Nothdurft, H. C. (1991). Texture segmentation and pop-out from orientation contrast. *Vision Research*, *31*, 1073–1078.

Portilla, J., & Simoncelli, E. P., 1999. Texture modeling and synthesis using joint statistics of complex wavelet coefficients. IEEE Workshop on Statistical and Computational Theories of Vision, pp. 1–32.

Rubenstein, B. S., & Sagi, D. (1993). Effects of foreground scale in texture discrimination task: Performance is size, shape and content specific. *Spatial Vision*, *7*, 293–310.

Simoncelli, E. P., & Portilla, J. (1998). Texture characterization via joint statistics of wavelet coefficient magnitudes. *Proceeding of the Fifth International Conference on Image Processing*, *1*, 1–5.

Sutter, A., & Graham, N. (1995). Investigating simple and complex mechanisms in texture segregation using speed-accuracy tradeoff method. *Vision Research*, *35*, 2825–2843.

Sutter, A., Sperling, G., & Chubb, C. (1995). Measuring the spatial frequency selectivity of second-order texture mechanisms. *Vision Research*, *35*, 915–924.

Thomson, M. G. A. (1999). Higher-order structure in natural scenes. *Journal of the Optical Society of America A*, *16*, 1549–1553.

Wilson, H. (1999). Non-Fourier cortical processes in texture, form, and motion perception. *Cerebral Cortex*, *13*, 445–477.

Yellott, J. I. (1993). Implications of triple correlation uniqueness for texture statistics and the Julesz conjecture. *Journal of the Optical Society of America A*, *10*, 777–793.