Visual Perception of Texture

Michael S. Landy
Dept. of Psychology and Center for Neural Science
New York University

Norma Graham
Dept. of Psychology
Columbia University

Running head: Texture Perception

Corresponding author:

Michael S. Landy

Temporary address through August, 2002:

University of California
School of Optometry
Minor Hall
Berkeley, CA 94720-2020
(510) 642-7679
fax: (510) 643-5109
landy@nyu.edu

Permanent address (e.g., after September, 2002):

New York University
Dept. of Psychology
6 Washington Place, Rm. 961
New York, NY 10003
(212) 998-7857
fax: (212) 995-4349
landy@nyu.edu

1. Introduction

What is visual texture, and how might a study of the visual perception of texture help us to better understand human vision? In this chapter we will attempt to give the reader a feel for how the study of texture perception is useful both in understanding the impact of texture itself, as well as in providing a better understanding of basic visual mechanisms that respond not only to texture but to all visual stimuli. This review will be relatively brief and, of necessity, incomplete. We hope to give an overview of the different research areas concerned with texture perception and of the current issues. For a longer early review, we refer the reader to Bergen (1991).

Consider the scene in Fig. 1. The border between the sky and the trees/grass involves a difference in luminance, one that would easily be signaled by a linear mechanism such as a simple cell in primary visual cortex. The boundary between the zebras and the background also involves a change in chromaticity (although not visible in the black-and-white image in Fig. 1), which might be signaled by color-opponent mechanisms. But, the borders between pairs of zebras involve neither a difference in color nor in average luminance. These borders include stretches of boundary that are black on one side and white on the other, stretches where the colors are reversed, and stretches where there is no local visual information to signal the boundary (where black abuts black, or white abuts white). Nevertheless, we perceive a smooth, continuous occlusion boundary at the edge of each animal. It is as if the visual system possesses the capability of segmenting regions of the image based on a local textural property, such as separating “vertical stuff” from “horizontal stuff.”

Thus, texture is a property that is statistically defined. A uniformly textured region might be described as “predominantly vertically oriented”, “predominantly small in scale”, “wavy”, “stubbly”, “like wood grain” or “like water.” As Adelson and Bergen (1991) put it, texture is a property of “stuff” in the image, in contradistinction to visual features such as lines and edges, the “things” in the image (analogous to the linguistic difference between mass nouns like “water” and count nouns like “mouse”).

Another way of characterizing visual texture is in the uses to which it might be put. Texture is a property of an image region. Regions in the visual field can be characterized by differences in texture, brightness, color or other attributes. Relatively early processes in the visual system can use texture information to perform a tentative segmentation of the visual image into regions to ease the processing load on subsequent computational stages. The analysis of a single textured image region can lead to the perception of categorical labels for that region (“This looks like wood.” or “This surface looks slippery.”). The appearance of texture allows the observer to determine whether two textured regions appear to be made of the same or different stuff. If two abutting image regions have different surface texture, this may lead to the detection of the intervening texture border (like the border between adjacent zebras in Fig. 1). Such texture-defined boundaries may then be used to segment figure from ground and for 2-dimensional shape identification. Finally, continuous changes in texture properties may result in the percept of 3-dimensional shape (Gibson 1950). A focus of much research in this area is to characterize the mechanisms and representational schemes used to characterize texture, and thus to determine whether the same underlying mechanisms are responsible for each of the above perceptual capabilities.

2. Texture Segregation

Texture Features

Much of the work on the perception concerns the ability of observers to effortlessly discriminate certain texture pairs. For example, Fig. 2 shows rectangular regions of X's and of T's on a background of L's. Observers can perceive effortlessly that there is a region of X's different from the background, that this region has smooth, continuous borders, and that these borders form a rectangular shape. This is referred to as the segregation of figure from ground or segmentation of the image into multiple, homogenous regions. At the same time, none of these observations may be made about the region of T's without use of effortful scrutiny of the individual texture elements one by one.

This sort of observation led a number of investigators to consider what aspects of image structure led to pre-attentive segregation of textures. Beck and Attneave and their colleagues (Beck 1972, 1973; Olson and Attneave 1970) hypothesized that textural segmentation occurs on the basis of the distribution of simple properties of "texture elements" where the simple properties were things like the brightness, color, size, and the slopes of contours and other elemental descriptors of a texture. Marr (1976) added contour terminations as an important feature.

Julesz's early efforts were centered around image statistics. He first suggested (Julesz, Gilbert et al. 1973) that differences in dipole statistics were most important for texture pairs to segregate (these are the joint image statistics of the gray levels found at the opposite ends of a line segment of a particular length and orientation, as it is placed at all possible image locations, gathered for all possible pairs of gray levels, dipole lengths and orientations). But, counterexamples to this were found (e.g., Caelli and Julesz 1978). It was then suggested that textures with identical third-order statistics would prove indiscriminable (analogous to dipole statistics, these are joint image statistics of the gray levels found at the three corners of a triangle of a particular size, shape and orientation, as it is placed at all possible image locations, gathered for all possible triplets of gray levels, triangle shapes, sizes and orientations). Again, counterexamples to this hypothesis were found (Julesz et al. 1978).

Julesz noticed that the counterexamples were suggestive of an alternative explanation for texture segregation, similar to those of Beck and Marr. Julesz found that texture pairs that segregated easily but had identical 3rd-order statistics also differed in the amount of an easily discernible image feature (e.g., Caelli et al. 1978). The task then became one of identifying the list of image features, which Julesz (1981) dubbed “textons”, that were sufficient to explain segregation performance. The initial list of textons included such features as size, orientation, line terminations and line crossings.

It has been noted that the 3rd-order statistics used by Julesz were “population statistics.” That is, the counterexamples to Julesz’ various conjectures never had identical 2nd- or 3rd-order statistics within the actual, finite images observed. Rather, the identity was over all possible images that could have been generated by the process that generated the particular instantiation of texture currently in view. In fact, for continuous images, image pairs with identical 3rd-order statistics must be identical images, rendering that version of the conjecture trivial (Yellott 1993), and finite, discrete images are determined by their dipole statistics (Chubb and Yellott 2000). On the other hand, Victor (1994) makes the case for the appropriateness of the use of population statistics for theorizing about texture segregation.

The feature-based theories were echoed in research in the visual search field (Treisman 1985). Effortless “parallel” search for a target pattern in a field of distracter patterns was found whenever the target and distracters differed in a feature (e.g., size, orientation, etc.)

similar to the texton features that led to effortless texture segregation. For example, a target X was effortlessly and immediately located in a field of distracter L's. However, when the target was a T, the task became effortful and required serial scrutiny of the texture elements, requiring more time with every additional distracter added to the stimulus (Bergen and Julesz 1983). When the choice of target and distracters requires the observer to attend to a specific combination of two features, search becomes difficult and observers often perceive "illusory conjunctions" between features of neighboring objects (Treisman and Schmidt 1982). Somewhat analogous effects using texture elements having combinations of two features have been noted in texture segregation as well (Papathomas et al. 1999). However, Wolfe (1992) suggests that texture segregation and parallel visual search do not always follow the same rules.

A number of other observations have been made concerning when texture element stimuli do or do not segregate. Beck (1982) has pointed out that textures segregate based not only on the particular texture elements used, but also on their arrangement, reminiscent of the Gestalt laws of figural goodness. As in the search literature (Treisman and Gormican 1988), texture segregation may show asymmetries (Beck 1973; Gurnsey and Browse 1989). For example, a patch of incomplete circles will easily segregate from a background of circles, whereas the reverse pattern results in poor segregation. It has been suggested this is due to a difference in the variability of responses of underlying visual mechanisms to the two possible texture elements (Rubenstein and Sagi 1990).

Nothdurft (1985) suggested that finding an edge between two textures is analogous to finding a luminance-defined edge. To determine a luminance boundary involves locating large values of the derivative of luminance (the luminance gradient) across an image. Finding texture boundaries might involve the determination of other aspects of image structure (local scale, local orientation, etc.), and segregation would then result from large values of the "structure gradient."

Finally, much of the literature assumes that effortless texture segregation and parallel visual search are truly effortless. That is, they require no selective attention to operate (demonstrated by, e.g., Braun and Sagi 1990). However, Joseph, Chun and Nakayama (1997) had observers perform an effortful, secondary task and noticed a large decrement in search performance in a search task that typically yields performance independent of the number of distracters. Thus, it is possible that even parallel search and, by extension, effortless texture segregation still require selective visual attention. Alternatively, texture segregation may not require focal visual attention, but attention may be used to alter the characteristics of visual mechanisms responsible for texture segregation (e.g., Yeshurun and Carrasco, 2000). Early literature also assumed that texture segregation was effortless in the sense of being "immediate". However at least some textures take substantial time to process (e.g. Sutter and Graham 1995) thus undermining the notion that preattentive texture segregation is always immediate and effortless.

We have treated texture as if it is somehow an isolated cue that can signal the presence, location and shape of an edge. However, texture can co-occur in a stimulus with other cues to edge presence such as luminance, color, depth or motion. Rivest and Cavanagh (1996) showed that perceived edge location was a compromise between the position signaled by texture and by other cues (motion, luminance, color). In addition, localization accuracy was better for 2-cue than for single-cue stimuli. Landy and Kojima (2001) found that different textural cues to edge location were combined using a weighted average, with greater weight given to the more reliable cues. This is analogous to the cue combination scheme that has been seen with multiple cues to depth (including depth-from-texture) by Landy, Maloney, Johnston and Young (1995), among others.

Current Models of Texture Segregation

How might one model the aspects of texture segregation performance we have just surveyed? If an edge is defined by a difference in luminance (a typical light/dark edge), then a band pass linear spatial filter similar to a cortical simple cell can detect the edge by producing a peak response at the location of the edge. But, a typical texture-defined edge (e.g., Figs. 2 and 4A) has the same average luminance on either side of the edge and thus will not be detected by any purely linear mechanism.

Several early investigators (e.g. Beck 1972; Julesz 1981) suggested that observers calculate the local density of various image features, and that differences in these texton or feature statistics on either side of a texture-defined edge results in effortless texture segregation. However, it was never clearly described exactly what an image feature was and how it would be computed from the retinal image. The image features discussed (e.g. lines of different slopes, line terminations and crossings) were clearly tied to the kinds of stimuli employed in most texture studies of the period (basically, pen-and-ink drawings), and would not be applied easily to natural, gray-scale images.

An alternative line of modeling suggests that we need look no further than the orientation- and spatial frequency-tuned channels already discovered in the spatial vision literature through summation, identification, adaptation and masking experiments using sine wave grating stimuli (De Valois and De Valois 1988; Graham 1989, 1992). For example, Knutsson and Granlund (1983) suggested that the distribution of power in different spatial frequency bands might be used to segregate natural textures and ran such a computational model on patchworks of textures drawn from the Brodatz (1966) collection (a standard collection of texture images often used in the computational literature).

Bergen and Adelson (1988) pointed out that even the example of X’s, L’s and T’s (Fig. 2) could be accounted for by the distribution of power in isotropic channels similar in form to cells found in the LGN and layer 4 of primary visual cortex. Further, they showed that if the size of the X’s was increased to effectively equate the dominant spatial frequency or “scale” of the different texture elements, the segregation of X’s from a background of L’s could be made difficult. This was strong evidence against the texton or feature theories.

A plethora of similar models based on filters selective for spatial frequency and orientation have been investigated (Bovik et al. 1990; Caelli 1985; Fogel and Sagi 1989; Graham 1991; Landy and Bergen 1991; Malik and Perona 1990; Sutter et al. 1989; Turner 1986; for an alternative view, see Victor 1988). These models are so similar in basic design that Chubb and Landy (1991) referred to this class as the “back pocket model of texture segregation,” as texture perception researchers pull this model from their back pocket to explain new phenomena texture segregation.

The basic back pocket model consists of three stages (Fig. 3). First, a set of linear spatial filters, akin to the simple cells of primary visual cortex, is applied to the retinal image. Second, the outputs of the first-stage linear filters are transformed in a nonlinear manner (by half- or full-wave rectification, squaring and/or gain control). Finally, another stage of linear filtering is used to enhance texture-defined contours. If this third stage consisted only of spatial pooling, the resulting outputs would resemble those of cortical simple cells. But, often this linear filter is modeled as bandpass and orientation-tuned, so that it enhances texture-defined edges much as an orientation-tuned linear spatial filter enhances luminance-defined edges.

This process is illustrated in Fig. 4. Fig. 4A shows an orientation-defined texture border (Wolfson and Landy 1995). In Fig. 4B a vertically-oriented spatial filter has been applied. The responses are larger to the vertically-oriented portion of the image, but these responses are both strongly positive (when the filter is centered on a texture element) and negative (when the filter is positioned off to the side of a texture element). As a result, the average value of the output is identical on either side of the texture border, but on the left the response variability is greater. In Fig. 4C the responses of Fig. 4B have been rectified, resulting in larger responses in the area of vertically-oriented texture. Finally, in Fig. 4D a 2nd-order, larger scale, vertically-oriented spatial filter has been applied, resulting in a peak response at the location of the texture-defined edge. For a detection experiment ("Was there a texture-defined edge in this briefly-flashed stimulus" or "Were there two different texture regions or only one?"), a model would try to predict human performance by the strength of the peak response in Fig. 4D as compared to peaks in responses to background noise in stimuli not containing texture-defined edges. For further examples, see Bergen (1991) and Bergen and Landy (1991).

A wide variety of terminology has been used to describe the basic model outlined in Fig. 3, making the literature difficult to approach for the neophyte. The basic sequence of a spatial filter, a nonlinearity and a second spatial filter has been called the back pocket model (Chubb and Landy 1991), a LNL (linear, nonlinear, linear) model, a FRF (filter, rectify, filter) model (e.g. Dakin et al. 1999), second-order processing (e.g. Chubb et al. 2001) or a simple or linear channel (the first L in LNL) followed by a comparison-and-decision stage (e.g. Graham, Beck, and Sutter 1992),

About the term "2nd-order". The phrase "2nd-order" can be particularly troublesome. In some hands, and as we will use it here, it merely refers to the 2nd-stage of linear filtering following the nonlinearity in a model like that of Fig. 3. As such, it has been applied to models in a wide variety of visual tasks (Chubb et al. 2001). But, “2nd-order” has another technical definition that has also been used in similar contexts. If the nonlinearity in Fig. 3 is a squaring operation, then the pixels in the output image (after the second stage of linear filtering) are all computed as 2nd-order (i.e., quadratic) polynomials of the pixels in the model input.

In this chapter, we will refer to the model of Fig. 3 as a 2nd-order model, meaning that it contains a 2nd-order linear spatial filter. Of necessity, this 2nd-order linear filter must follow an intervening nonlinearity. Otherwise, there would simply be two sequential linear filters which is indistinguishable from a single, lumped linear spatial filter. We will use this term regardless of the polynomial order of the intervening nonlinearity.

There is also a more general use of “2nd-order”. In this usage a 2nd-order entity (e.g., a neuron) pools, after some intervening nonlinearity, the responses from a number of other entities (called 1st-order) but, in this more general usage, the 1st -order entities do not form a linear filter characterized by a single spatial weighting function as they do in Fig. 3. Rather, the 1st-order entities can be an assortment of neurons sensitive to various different things (e.g. different orientations or different spatial frequencies). See the introduction to Graham and Sutter (1998) for a brief review of such general suggestions.

3rd-order models. 2nd-order models are not the end of the story. For example, Graham, Sutter, and Venkatesan (1993) used an element-arrangement texture stimulus consisting of two types of elements, arranged in stripes in one region and in a checkerboard in another region. Consider the case where each texture element is a high frequency Gabor pattern (a windowed sine wave grating) and the two types of elements differ only in spatial frequency. Consider a 2nd-order model like that just described with the first linear filter tuned to one of the two types of Gabor patches, and the second linear filter tuned to the width and orientation of stripes of elements. This 2nd-order model would yield a response to these element-arrangement textures that was of the same average level although of high contrast in the striped region and low contrast in the checked region. To reveal the texture-defined edge between the checkerboard and striped regions, therefore, requires another stage of processing, which could be a pointwise nonlinearity followed by an even larger-scale linear spatial filter (another NL) thus producing a sequence LNLNL. For an illustration of such a model's responses see Graham et al. (1993), Fig. 4.

Here we will call this LNLNL sequence a 3rd-order model. But, to avoid confusion, let us note that Graham and her colleagues refer to the first LNL as a "complex channel" or "2nd-order channel" and the final NL is an instance of what they call the comparison-and-decision stage.

About the terms "Fourier" and "non-Fourier". There is also possible confusion about the terms Fourier and non-Fourier. A stimulus like that in Fig. 4A in which the edge can be found by the model in Fig. 3 has been referred to as "non-Fourier" (first applied to motion stimuli by Chubb and Sperling 1988). The term was used because the Fourier spectrum of this stimulus does not contain components that correspond directly to the texture-defined edge. But some others (e.g., Graham and Sutter 2000) have used the term "Fourier" channels for the first linear filters (the simple channels) in Fig. 3 and reserved the term "non-Fourier" for the complex channels (the initial LNL) in what we called third-order models above (LNLNL).

This confusing terminology is the result of a difference in emphasis. In this chapter we concentrate on models that localize (i.e., produce a peak response at) edges between two abutting textures.. But, others (e.g. Graham and Sutter 2000; Lin and Wilson 1996) have emphasized response measures that can be used to discriminate between pairs of textures (whether simultaneously present and abutting or not) by any later, nonlinear decision process. Thus, finding the edge in an orientation-defined texture like that of Fig. 3 is, in Graham and Sutter's terms, "Fourier-based" as the power spectra of the two constituent textures differ, whereas finding the edge in a Gabor-patch element-arrangement texture like that of Graham, Sutter, and Venkatesan (1993) is "non-Fourier-based" as the power spectra of the two constituent textures do not differ.

Model Specification

The models of texture segregation just described are complicated, with many details that require elucidation. Are the initial linear filters of a 2nd-order pathway the same spatial filters as the spatial frequency channels that have been described using grating experiments? What is the nature of the following nonlinearity? Are there fixed, 2nd-order linear filters and what is their form? This is an area of current, active research and most of these issues have not been convincingly decided.

Graham, Sutter and Venkatesan (1993) and Dakin and Mareschal (2000) provide evidence that the initial spatial filters in a 2nd-order pathway used to detect contrast modulations of texture are themselves tuned for spatial frequency and orientation. In the same article, Graham and colleagues also demonstrated that the initial spatial filters in a 3rd-order pathway (their "complex channels") were orientation and spatial-frequency-tuned as well.

The back pocket model includes a nonlinearity between the two stages of linear spatial filtering that is required to demodulate the input stimuli. For small 1st-order spatial filters, Chubb, Econopouly and Landy (1994) provided a technique called histogram contrast analysis that allowed them to measure aspects of the static nonlinearity, showing that it included components of higher order than merely squaring the input luminances. Graham and Sutter (1998) found that this nonlinearity must be expansive. They also (Graham and Sutter 2000) suggest that a gain control mechanism acts as an inhibitory influence among multiple pathways both of the type called 2nd-order and 3rd-order here.

First-order spatial frequency channels were first measured using sine wave grating stimuli and various experimental paradigms including adaptation, masking and summation experiments (reviewed in Graham 1989). Recently, researchers have used analogous experiments to examine the 2nd-order linear filters. To do so, researchers hope to deliver to the 2nd-order filter something like the sine wave grating stimuli of classic spatial frequency channel studies. The usual ploy is to use a stimulus that has a sine wave (or Gabor) pattern that is used to modulate some aspect of textural content across the stimulus. The assumed 1st-order filter and subsequent nonlinearity demodulate this stimulus, providing as input to the 2nd-order linear filter a noisy version of the intended grating or Gabor pattern.

Studies of texture modulation detection have revealed a very broad-band 2nd-order texture contrast sensitivity function (CSF) using a variety of texture modulations including contrast (Schofield and Georgeson 1999, 2000; Sutter et al. 1995), local orientation content (Kingdom et al. 1995) and modulation between vertically- and horizontally-oriented filtered noise (Landy and Ternes 1995). This function is far more broad-band than the corresponding luminance CSF. A demonstration of this effect is shown in Fig. 5A. A modulator pattern is used to additively combine a vertical and horizontal noise texture. The modulator increases in spatial frequency from left to right, and in contrast from bottom to top. As you can see, the texture modulation becomes impossible to discern at approximately the same level for all spatial frequencies. The sample data in Fig. 5B confirm this observation.

Evidence for multiple, 2nd-order filters underlying this broad 2nd-order CSF has been equivocal, with evidence both pro (Arsenault et al. 1999; Schofield and Georgeson 1999) and con (Kingdom and Keeble 1996). Many studies have found texture discrimination to be scale-invariant, suggesting the existence of a link between the scale of the corresponding 1st- and 2nd-order spatial filters (Kingdom and Keeble 1999, Landy and Bergen 1991, Sutter et al. 1995). It has also been suggested that the orientation preferences of the 1st- and 2nd-order filters tend to be aligned (Dakin and Mareschal 2000; Wolfson and Landy 1995). This alignment of 1st- and 2nd-order filters has also been supported for element-arrangement stimuli that require a 3rd-order model to detect the texture-defined edges (Graham and Wolfson 2001).

If there is an obligatory link between the scales of the 1st- and 2nd-order filters, this suggests that the preferred 2nd-order scale should depend on eccentricity. This was first demonstrated by Kehrer (1989) who noted that performance on an orientation-defined texture-segregation task at first improves as the target texture moves into the periphery, and then falls as the eccentricity further increases. The poor foveal performance was dubbed the central performance drop (CPD). This argument that the CPD is due to the relation between the scale of the 2nd-order pattern and the local scale of the 2nd-order filter has been made by Yeshurun and Carrasco (2000) who, in addition, suggested that the 2nd-order spatial filters are narrowed as a consequence of the allocation of selective attention.

The temporal properties of the 1st- and 2nd-order filters are not well understood although some information is available (Lin and Wilson 1996; Motoyoshi and Nishida 2001; Schofield and Georgeson 2000; Sutter and Graham 1995, Sutter and Hwang 1999).

The possibility that the wiring between 1st-order and 2nd order filters is more complicated than that shown in Fig. 3 remains open as well (e.g. appendix in Graham and

Sutter 1998; Mussap 2001) with some particular interest in possible lateral excitatory and inhibitory interactions among different positions within the same filter (Motoyoshi 1999; Wolfson and Landy 1999).

Early filters are not the only visual processes that play an important role in determining the conscious perception of textured stimuli. Consider He and Nakayama (1994), who constructed a series of binocular demonstration stimuli, involving both texture and disparity. The foreground surface consisted of a set of textured squares. The background stimuli consisted of a region of I shapes surrounded by L shapes that, monocularly, segregated quite easily. However, when seen in depth with the squares (that abutted the L’s and I’s) in front, both the L’s and I’s were perceived as occluded by the squares. They underwent surface completion; that is, they were both perceived as larger rectangles occluded by the squares, and texture segregation became effortful. This suggests that higher-level, surface-based representations are involved in judgments about the objects perceived on the basis of textured regions in the stimulus.

3. Texture Appearance

The previous section concentrated on research concerning observers’ abilities to detect borders between differing textures. Here, we consider research more directly measuring the appearance of textures. If two images both appear to be a “grassy field” then, at some level of analysis, the representations of the two images must be similar. To understand the appearance of texture might involve developing such a representation as well as a metric within that representation space so that textures are perceived as similar if their representations are close, and dissimilar if far. Indeed, there is even evidence that texture appearance (or, at least, region-based) mechanisms can be responsible for texture segregation in some cases (Wolfson and Landy 1998) as certain texture pairs can be discriminated just as well when they are separated as when they abut (forming an edge). Using region-based as well as edge-based mechanisms may be optimal for segregation processes (Lee 1995).

One approach to this problem of measuring texture appearance is a classical one: elicit similarity judgments from observers and try to build a representation. Having done so, one can then ask whether the underlying dimensions have any semantic basis, or whether dimensions satisfy any of the properties of other perceptual dimensions (such as the additivity and metamerism of color space). Three dimensions appeared to suffice for sets of natural textures (Rao and Lohse 1996) as well as artificial ones (Gurnsey and Fleet 2001; Harvey and Gervais 1978). A texture analogy to color matching experiments with artificial, one-dimensional textures provide satisfactory appearance matches with four texture primaries (Richards and Polit 1974). As with color matching, this technique shows that one can account for texture matches with the four primaries, but does not explain texture appearance. Color appearance depends on the particular metameric match as well as on color context. Similarly, texture appearance can depend on context. For example, Durgin (2001) shows that perceived texture density of a texture patch depends on the density of the surrounding texture.

An alternative approach is to analyze an instance of texture so as to estimate its representation, then to use that representation to generate new instances of texture. The proposed representational scheme is considered successful if the newly generated textures are classified as "made of the same stuff as the original" by observers. The first such model, by Heeger and Bergen (1995), represented the input texture image as the histograms of values in each level of an oriented pyramid representation of the image, that is, as the statistics of the responses from a collection of orientation- and spatial frequency-tuned spatial filters. The resulting newly-generated texture images were occasionally striking in their similarity to the original. But, in other instances, especially those involving correlations between different image areas at long distances, the results were quite poor. More recent models incorporate higher-order statistics including correlations between pairs of filter responses across space, spatial frequency and orientation (De Bonet and Viola 1998; Portilla and Simoncelli 2000; Zhu, Wu and Mumford 1998). Fig. 6 shows two sample textures (inset squares) that were extrapolated by the technique of Portilla and Simoncelli (2000). Clearly, the technique has captured a good deal of that which defines the appearance of these textures. The technique is somewhat less successful with purely periodic textures (tiles), binary or pen-and-ink textures, or with pseudo-textures that are, for example, collections of small objects (e.g., a pile of jellybeans). It remains to be seen whether a metric (Euclidean, Minkowski, or other) applied to one of these texture representation spaces will correlate well with observers' judgments of the perceptual similarity of textures.

Few psychophysical tests of these new statistical characterizations of texture have been carried out. Kingdom, Hayes and Field (2001), in an analogy to the work of Chubb and colleagues in the luminance domain (1994), found that observers were most sensitive to kurtosis in the histograms of wavelet (that is, multi-scale, orientation-tuned) coefficients in artificial textures. Durgin (2001) has suggested that texture density is a separate dimension from either mean (luminance) or variance (RMS contrast).

The texture representation schemes just discussed are image-based. That is, all content of the representation is based on simple statistics based on responses of filters to the texture. A complete theory of texture perception might involve recognition that natural textures are associated with real-world materials, and the appearance of texture may well relate to perception of the particular material from which the image derived (wood, plastic, water, grassland, etc.) or properties of the real-world material that might relate to actions the observer might wish to take. This is the concept of an “affordance” (Gibson 1979). Is this material sticky? Will it crumble in my hand? Will I be able to walk on it in bare feet? There has been a great deal of work, notably in the computer graphics world, to understand image properties of natural materials to be able to simulate these materials in virtual displays. On the other hand, very little research has been done on the perception of real-world textural properties. Recently, there has been some effort to understand the variety of images one can find of natural textures as viewpoint and lighting conditions are varied (Dana, van Ginneken, Nayar and Koenderink 1999).

4. Shape from Texture

Gibson (1950) pointed out that the perspective distortion of surface texture is a cue to surface layout. For example, consider a ground plane that is painted with randomly placed circles. As the surface recedes into the distance, three different “texture gradients” may be distinguished: size (further texture elements are smaller in the retinal image), density (further texture elements are closer together in the retinal image) and compression (further elements are more slanted relative to the line of sight, and hence form more eccentric ellipses in the retinal image).

The computational literature is replete with suggested algorithms for the computation of shape from texture. These algorithms vary in how restrictive an assumption is made about the surface texture. The earliest algorithms (e.g., Witkin 1981) assumed an isotropic texture (all orientations were equally represented on the surface, which is true of the above example). More recent algorithms (e.g., Aloimonos 1988) only assume texture homogeneity (i.e., the texture is statistically the same at all positions on the surface). A particularly interesting algorithm is that of Malik and Rosenholtz (1997). This algorithm makes weak assumptions about the underlying surface texture. It looks for affine distortions in image statistics from one location to another as seen in the responses of a bank of spatial filters varying in orientation and spatial frequency preference, much like the first stage in the current models of texture segregation.

Psychophysical research on the perception of shape from texture has followed a similar history. Cutting and Millard (1984) discuss the three possible texture gradients and manipulated them independently in their stimuli. They found that perception of slant for planar stimuli depended mainly on the size gradient, whereas perception of curved stimuli was almost completely determined by the compression gradient. Rosenholtz and Malik (1997) found texture isotropy to be unnecessary for human observers to estimate surface orientation, consistent with their computational theory. Li and Zaidi (2000) examined the types of surface texture that would give a veridical percept of shape when mapped onto a corrugated surface in perspective, and found that several aspects of the Fourier power spectrum were predictive of observer accuracy, corresponding to the availability of oriented energy along lines of maximum and minimum curvature in the surface.

A second line of psychophysical research has been to derive ideal (maximum a posteriori) observers and to compare the reliability of human observer estimates of surface layout with those of the ideal observer. Blake, Buelthoff and Sheinberg (1993) derived such a model with the assumption of isotropic, homogeneous surface texture, and demonstrated that observer estimates of surface curvature must use the compression gradient. Buckley, Frisby and Blake (1996) applied the same strategy to the estimation of surface slant, and found that texture compression dominates observer judgments even for fields of view large enough that, for the ideal, texture density should dominate. Finally, in a series of three papers, Knill (1998a-c) derived ideal observers for slant from texture that use the three texture gradient cues and derived the reliability of each cue as a function of slant and field of view. He found that human observers became more reliable with increasing slant and field of view just as did the ideal observers. Again, performance was so good that observers must have used texture compression and, at least in part, an assumption of isotropy.

5. Neurophysiology

The physiological substrate for the first-stage linear filters in texture segregation models is likely to be the spatial frequency and orientation selective cells in cortical area V1. Further, V1 is sufficiently complicated that other attributes of the current models, such as the normalization or other nonlinearities and subsequent spatial pooling, could certainly also occur in V1. There are also lateral interactions between neurons in V1 (both excitatory and inhibitory) that have been reported that go beyond the classical receptive field. There has been some controversy over the function of these has lateral interactions in V1. Some have suggested lateral interactions enhance responses to pop-out stimuli (Kastner et al. 1997, 1999; Nothdurft et al. 1999), to texture elements near texture borders (Nothdurft et al. 2000), to orientation contrast (Knierim and van Essen 1992; Sillito et al. 1995), and to figure rather than ground (Lamme, 1995; Zipser et al. 1996). Li (2000) even describes a neural network model of segmentation that includes such processes.

However, the responses to orientation contrast stimuli are a complex function of the contrasts of the figure and ground (Levitt and Lund, 1997) suggesting that these V1 responses are primarily the result of a gain control mechanism that is only an initial stage of the computation of texture borders and figure-ground. Several groups find that input from outside the classical receptive field is mainly suppressive, consistent with this view (Freeman et al. 2001; Rossi et al., 2001; Sceniak et al. 2001; Walker et al. 2000). An in-depth review of a large range of results from areas V1 up through MT and V4 (Lennie 1998) concludes that it may be too much to attribute such functions as pop-out and figure-ground segregation to area V1, and that these functions probably occur in V2 through V4 or even at higher levels. Lennie suggests that “Spatial interactions in V1 probably have a less exotic role; they provide lateral inhibition in the domain of local structure so that, by analogy with lateral inhibition in the luminance domain, signals from regions of common structure are suppressed and contrasts in structure are made salient.” In this view, it is not until area V4 that the system has even grouped regions of similar structure to find contours, regions, surfaces and, perhaps, compute surface slant. And thus, in this view, many of the processes called into play by texture stimuli (e.g. the conscious perception of a surface as having a particular texture) would be determined predominantly by still higher-level cortical areas. A recent fMRI study of static texture segregation (Kastner et al. 2000) concurs, finding little response to texture borders in V1 or V2/VP, and increasing responses as one proceeds downstream from V3 to V4 and TEO.

6. Conclusions

The perception of texture is a rich and varied area of study. In the early coding of texture borders, there is some common ground between current psychophysical data and models and the physiology of primary visual cortex, such as the suggestion that texture border coding involves a succession of linear spatial filters and nonlinearities that include both static nonlinearities as well as contrast gain control mechanisms. Less well understood, however are such higher-level computations involving texture as the calculation of figure-ground, the coding of texture appearance, and the determination of depth and 3-D shape from texture cues.

Acknowledgments

Michael Landy was supported by National Eye Institute grant EY08266 and Human Frontier Science Program grant RG0109/1999-B. Norma Graham was supported by National Eye Institute grant EY08459 We would like to acknowledge the helpful comments of Sabina Wolfson over a period of many years.

REFERENCES

*Adelson, E. H., and Bergen, J. R., 1991, The plenoptic function and the elements of early vision, in *Computational Models of Visual Processing*, (M. S. Landy and J. A. Movshon, eds.), Cambridge, MA: MIT Press, pp. 3-20.

Aloimonos, J., 1988, Shape from texture, *Artificial Intelligence*, 38:345-360.

Arsenault, A. S., Wilkinson, F., and 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., 1972, Similarity grouping and peripheral discriminability under uncertainty, *American Journal of Psychology*, 85:1-19.

Beck, J., 1973, Similarity grouping of curves, *Perceptual and Motor Skills*, 36:1331-1341.

Beck, J., 1982, Textural segmentation, in *Organization and Representation in Perception*, (J. Beck, ed.), Hillsdale, NJ: Erlbaum, pp. 285-317.

*Bergen, J. R. 1991, Theories of visual texture perception, in *Vision and Visual Dysfunction, vol. 10B*, (D. Regan, ed.), New York: Macmillan, pp. 114-134.

Bergen, J. R., and Adelson, E. H., 1988, Early vision and texture perception, *Nature*, 333:363-364.

Bergen, J. R., and Julesz, B., 1983, Parallel versus serial processing in rapid pattern discrimination, *Nature*, 303:696-698.

*Bergen, J. R., and Landy, M. S., 1991, Computational modeling of visual texture segregation, in *Computational Models of Visual Processing Landy*, (M. S. Landy and J. A. Movshon, eds.), Cambridge, MA: MIT Press, pp. 253-271.

Blake, A., Buelthoff, H. H., and Sheinberg, D., 1993, Shape from texture: Ideal observers and human psychophysics, *Vision Research*, 33:1723-1737.

Bovik, A. C., Clark, M., and Geisler, W. S., 1990, Multichannel texture analysis using localized spatial filters, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:55-73.

Braun, J., and Sagi, D., 1990, Vision outside the focus of attention, *Perception & Psychophysics*, 48:45-58.

Brodatz, P., 1966, *Textures*, New York: Dover.

Buckley, D., Frisby, J. P., and Blake, A., 1996, Does the human visual system implement an ideal observer theory of slant from texture?, *Vision Research*, 36:1163-1176.

Caelli, T., 1985, Three processing characteristics of visual texture segmentation, *Spatial Vision*, 1:19-30.

Caelli, T., and Julesz, B., 1978, On perceptual analyzers underlying visual texture discrimination: Part I, *Biological Cybernetics*, 28:167-175.

Caelli, T., Julesz, B., and Gilbert, E. N., 1978, On perceptual analyzers underlying visual texture discrimination: Part II, *Biological Cybernetics*, 29:201-214.

Chubb, C., Econopouly, J., and 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.

Chubb, C., and Landy, M. S., 1991, Orthogonal distribution analysis: A new approach to the study of texture perception, in *Computational Models of Visual Processing*, (M. S. Landy and J. A. Movshon, eds.), Cambridge, MA: MIT Press, pp. 291-301.

Chubb, C., Olzak, L., and Derrington, A., 2001, Second-order processes in vision: introduction, *Journal of the Optical Society of America A*, 18:2175-2178.

Chubb, C., and Sperling, G., 1988, Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception, *Journal of the Optical Society of America A*, 5:1986-2007.

Chubb, C., and Yellott, J. I., Jr., 2000, Every discrete, finite image is uniquely determined by its dipole histogram, *Vision Research*, 40:485-492.

Cutting, J. E., and Millard, R. T., 1984, Three gradients and the perception of flat and curved surfaces, *Journal of Experimental Psychology: General*, 113:198-216.

Dakin, S. C., and Mareschal, I., 2000, Sensitivity to contrast modulation depends on carrier spatial frequency and orientation, *Vision Research*, 40:311-329.

Dakin, S. C., Williams, C. B., and Hess, R. F., 1999, The interaction of first- and second-order cues to orientation, *Vision Research*, 39:2867-2884.

Dana, K. J., van Ginneken, B., Nayar, S. K., and Koenderink, J. J., 1999, Reflectance and texture of real-world surfaces, *ACM Transactions on Graphics*, 18:1-34.

De Bonet, J. S., and Viola, P., 1998, A non-parametric multi-scale statistical model for natural images, in *Advances in Neural Information Processing Systems 9*, (M. I. Jordan, M. J. Kearns and S. A. Solla, eds.), Cambridge, MA: MIT Press, pp. 773-779.

*De Valois, R. L., and De Valois, K. K., 1988, *Spatial Vision*, New York: Oxford University Press.

Durgin, F. H., 2001, Texture contrast aftereffects are monocular; texture density aftereffects are binocular, *Vision Research*, 41:2619-2630.

Freeman R. D., Ohzawa, I., and Walker, G., 2001, Beyond the classical receptive field in the visual cortex, *Progress in Brain Research*, 134:157-170.

Fogel, I., and Sagi, D., 1989, Gabor filters as texture discriminator, *Biological Cybernetics*, 61:103-113.

*Gibson, J. J., 1950, *The Perception of the Visual World*, Boston: Houghton Mifflin.

*Gibson, J. J., 1979, *The Ecological Approach to Visual Perception*, Boston: Houghton Mifflin.

*Graham, N., 1989, *Visual Pattern Analyzers*, New York: Oxford University Press.

Graham, N., 1991, Complex channels, early local nonlinearities, and normalization in perceived texture segregation, in Computational Models of Visual Processing, (M. S. Landy and J. A. Movshon, eds.), Cambridge, MA: MIT Press, pp. 291-301.

*Graham, N., 1992, Breaking the visual stimulus into parts, *Current Directions in Psychological Science*, 1:55-61.

Graham, N., Beck, J., and 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.

Graham, N., and Sutter, A., 1998, Spatial summation in simple (Fourier) and complex (non-Fourier) texture channels, *Vision Research*, 38:231-257.

Graham, N., and Sutter, A., 2000, Normalization: contrast-gain control in simple (Fourier) and complex (non-Fourier) pathways of pattern vision, *Vision Research*, 40:2737-2761.

Graham, N., Sutter, A., and Venkatesan, C., 1993, Spatial-frequency- and orientation-selectivity of simple and complex channels in region segmentation, *Vision Research*, 33:1893-1911.

Graham, N., and Wolfson, S. S., 2001, A note about preferred orientations at the first and second stages of complex (second-order) texture channels, *Journal of the Optical Society of America A*, 18:2273-2281.

Gurnsey, R., and Browse, R. A., 1989, Asymmetries in visual texture discrimination, *Spatial Vision*, 4:31-44.

Gurnsey, R., and Fleet, D. J., 2001, Texture space, *Vision Research*, 41:745-757.

Harvey, L. O., Jr., and Gervais, M. J., 1978, Visual texture perception and Fourier analysis, *Perception & Psychophysics*, 24:534-542.

He, Z. J., and Nakayama, K., 1994, Perceiving textures: beyond filtering, *Vision Research*, 34:151-162.

Heeger, D., and Bergen, J. R., 1995, Pyramid-based texture analysis/synthesis, *Proceedings of ACM SIGGRAPH 1995*, New York: Association for Computing Machinery, pp. 229-238.

Joseph, J. S., Chun, M. M., and Nakayama, K., 1997, Attentional requirements in a ‘preattentive’ feature search task, *Nature*, 387:805-807.

Julesz, B., 1981, Textons, the elements of texture perception, and their interactions, *Nature*, 290:91-97.

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

Julesz, B., Gilbert, E. N., and Victor, J. D., 1978, Visual discrimination of textures with identical third-order statistics, *Biological Cybernetics*, 31:137-140.

Kastner, S., de Weerd, P., and Ungerleider, L. G., 2000, Texture segregation in the human visual cortex: a function MRI study, *Journal of Neurophysiology*, 83:2453-2457.

Kastner, S., Nothdurft, H. C., and Pigarev, I. N., 1997, Neuronal correlates of pop-out in cat striate cortex, *Vision Research*, 37:371-376.

Kastner, S., Nothdurft, H. C., and Pigarev, I. N., 1999, Neuronal responses to orientation and motion contrast in cat striate cortex, *Visual Neuroscience*, 16:587-600.

Kehrer, L., 1989, Central performance drop on perceptual segregation tasks, *Spatial Vision*, 4:45-62.

Kingdom, F. A. A., Hayes, A., and Field, D. J., 2001, Sensitivity to contrast histogram differences in synthetic wavelet-textures, *Vision Research*, 41:585-598.

Kingdom, F. A. A., and 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., and Keeble, D. R. T., 1999, On the mechanism for scale invariance in orientation-defined textures, *Vision Research*, 39:1477-1489.

Kingdom, F. A. A., Keeble, D. R. T., and Moulden, B., 1995, Sensitivity to orientation modulation in micropattern-based textures, *Vision Research*, 35:79-91.

Knierim, J. J., and Van Essen, D. C., 1992, Neuronal responses to static texture patterns in area V1 of the alert macaque monkey, *Journal of Neurophysiology*, 67:961-980.

Knill, D. C., 1998a, Surface orientation from texture: ideal observers, generic observers and the information content of texture cues, *Vision Research*, 38:1655-1682.

Knill, D. C., 1998b, Discrimination of planar surface slant from texture: human and ideal observers compared, *Vision Research*, 38:1683-1711.

Knill, D. C., 1998c, Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture, *Vision Research*, 38:2635-2656.

Knutsson, H., and Granlund, G. H., 1983, Texture analysis using two-dimensional quadrature filters, *Proceedings of the IEEE Computer Society Workshop on Computer Architecture for Pattern Analysis and Image Database Management*, Silver Spring, MD: IEEE Computer Society, pp. 206-213.

Lamme, V. A. F., 1995, The neurophysiology of figure-ground segregation in primary visual cortex, *Journal of Neuroscience*, 15:1605-1615.

Landy, M. S., and Bergen, J. R., 1991, Texture segregation and orientation gradient, *Vision Research*, 31:679-691.

Landy, M. S., and Kojima, H., 2001, Ideal cue combination for localizing texture-defined edges, *Journal of the Optical Society of America A*, 18:2307-2320.

Landy, M. S., Maloney, L. T., Johnston, E. B., and Young, M. J., 1995, Measurement and modeling of depth cue combination: In defense of weak fusion, *Vision Research*, 35:389-412.

Landy, M. S., and Ternes, C. M., 1995, 2nd-order spatial contrast sensitivity, *1995 OSA Annual Meeting/ILS-XI Program*, Washington, DC: Optical Society of America, p. 64.

Lee, T. S., 1995, A Bayesian framework for understanding texture segmentation in the primary visual cortex, *Vision Research*, 35:2643-2657.

*Lennie, P., 1998, Single units and cortical organization, *Perception*, 27:889-935.

Levitt, J. B., and Lund, J. S., 1997, Contrast dependence of contextual effects in primate visual cortex, *Nature*, 387:73-76.

Li, A., and Zaidi, Q., 2000, Perception of three-dimensional shape from texture is based on patterns of oriented energy, *Vision Research*, 40:217-242.

Li, Z., 2000, Pre-attentive segmentation in the primary visual cortex, *Spatial Vision*, 13:25-50.

Lin, L. M., and Wilson, H. R., 1996, Fourier and non-Fourier pattern discrimination compared, *Vision Research*, 36:1907-1918.

Malik, J., and Perona, P., 1990, Preattentive texture discrimination with early vision mechanisms, *Journal of the Optical Society of America A*, 7:923-932.

Malik, J., and Rosenholtz, R., 1997, Computing local surface orientation and shape from texture for curved surfaces, *International Journal of Computer Vision*, 23:149-168.

Marr, D., 1976, Early processing of visual information, *Philosophical Transactions of the Royal Society London B*, 275:483-519.

Motoyoshi, I., 1999, Texture filling-in and texture segregation revealed by transient masking, *Vision Research*, 39:1285-1291.

Motoyoshi, I., and Nishida, S., 2001, Temporal resolution of orientation-based texture segregation, *Vision Research*, 41:2089-2105.

Mussap, A. J., 2001, Orientation integration in detection and discrimination of contrast-modulated patterns, *Vision Research*, 41:295-311.

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

Nothdurft, H. C., Gallant, J. L., and Van Essen, D. C., 1999, Response modulation by texture surround in primate area V1: correlates of “popout” under anesthesia, *Visual Neuroscience*, 16:15-34.

Nothdurft, H. C., Gallant, J. L., and Van Essen, D. C., 2000, Response profiles to texture border patterns in area V1, *Visual Neuroscience*, 17:421-36.

Olson, R. K., and Attneave, F., 1970, What variables produce similarity grouping?, *The American Journal of Psychology*, 83:1-21.

Papathomas, T. V., Gorea, A., Feher, A., and Conway, T. E., 1999, Attention-based texture segregation, *Perception & Psychophysics*, 61:1399-1410.

Portilla, J., and Simoncelli, E. P., 2000, A parametric texture model based on joint statistics of complex wavelet coefficients, *International Journal of Computer Vision*, 40:49-71.

Rao, A. R., and Lohse, G. L., 1996, Towards a texture naming system: identifying relevant dimensions of texture, *Vision Research*, 36:1649-1669.

Richards, W., and Polit, A., 1974, Texture matching, *Kybernetik*, 16:155-162.

Rivest, J., and Cavanagh, P., 1996, Localizing contours defined by more than one attribute, *Vision Research*, 36:53-66.

Rosenholtz, R., and Malik, J., 1997, Surface orientation from texture: isotropy or homogeneity (or both)?, *Vision Research*, 16:2283-2293.

Rossi, A. F., Desimone, R., and Ungerleider, L. G., 2001, Contextual modulation in primary visual cortex of macaques, *Journal of Neuroscience*, 21:1698-1709.

Rubenstein, B. S., and Sagi, D., 1990, Spatial variability as a limiting factor in texture-discrimination tasks: implications for performance asymmetries, *Journal of the Optical Society of America A*, 7:1623-1643.

Sceniak, M. P., Hawken, M. J., and Shapley, R., 2001, Visual spatial characterization of macaque V1 neurons, *Journal of Neurophysiology*, 85:1873-1887.

Schofield, A. J. and Georgeson, M. A., 1999, Sensitivity to modulations of luminance and contrast in visual white noise: separate mechanisms with similar behavior, *Vision Research*, 39:2697-2716.

Schofield, A. J., Georgeson, M. A., 2000, The temporal properties of first- and second-order vision, *Vision Research*, 40:2475-2487.

Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., and Davis, J., 1995, Visual cortical mechanisms detecting focal orientation discontinuities, *Nature*, 378:492-496.

Sutter, A., Beck, J., and Graham, N., 1989, Contrast and spatial variables in texture segregation: Testing a simple spatial-frequency channels model, *Perception & Psychophysics*, 46:312-332.

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

Sutter, A., and Hwang, D., 1999, A comparison of the dynamics of simple (Fourier) and complex (non-Fourier) mechanisms in texture segregation, *Vision Research*, 39:1943-1962.

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

Treisman, A. M., 1985, Preattentive processes in vision, *Computer Vision, Graphics, and Image Processing*, 31:156-177.

Treisman, A. M., and Gormican, S., 1988, Feature analysis in early vision: Evidence from search asymmetries, *Psychological Review*, 95:15-48.

Treisman, A. M., and Schmidt, H., 1982, Illusory conjunctions in the perception of objects, *Cognitive Psychology*, 14:107-141.

Turner, M. R., 1986, Texture discrimination by Gabor functions, *Biological Cybernetics*, 55:71-82.

Victor, J. D., 1988, Models for preattentive texture discrimination: Fourier analysis and local feature processing in a unified framework, *Spatial Vision*, 3:263-280.

Victor, J. D., 1994, Images, statistics, and textures: implications of triple correlation uniqueness for texture statistics and the Julesz conjecture: comment, *Journal of the Optical Society of America A*, 11:1680-1684.

Walker, G. A., Ohzawa, I., and Freeman, R. D., 2000, Suppression outside the classical cortical receptive field, *Visual Neuroscience*, 17:369-379.

Witkin, A. P., 1981, Recovering surface shape and orientation from texture, *Artificial Intelligence*, 17:17-45.

Wolfe, J. M., 1992, "Effortless" texture segmentation and "parallel" visual search are *not* the same thing, *Vision Research*, 32:757-763.

Wolfson, S. S., and Landy, M. S., 1995, Discrimination of orientation-defined texture edges, *Vision Research*, 35:2863-2877.

Wolfson, S. S., and Landy, M. S., 1998, Examining edge- and region-based texture mechanisms, *Vision Research*, 38:439-446.

Wolfson, S. S., and Landy, M. S., 1999, Long range interactions between oriented texture elements, *Vision Research,* 39:933-945.

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

Yeshurun, Y., and Carrasco, M., 2000, The locus of attentional effects in texture segmentation, *Nature Neuroscience*, 3:622-627.

Zipser, K., Lamme, V. A. F., and Schiller, P. H., 1996, Contextual modulation in primary visual cortex, *Journal of Neuroscience*, 16:7376-7389.

Zhu, S. C., Wu, Y., and Mumford, D., 1998, Filters, random fields and maximum entropy (FRAME) – Towards a unified theory for texture modeling, *International Journal of Computer Vision*, 27:107-126.

FIGURE LEGENDS

1. Types of image borders. A natural image containing borders signaled by differences in luminance, color and/or textural content.

2. Texture segregation. Notice that the region of X's on the left is easily segregated from the background of L's. One immediately perceives the borders between the two regions, and the shape of the region containing the X's. On the other hand, the border between the T's and L's is difficult to see, and the shape of the region of T's can only be discerned slowly, effortfully, and with item-by-item scrutiny.

3. The back pocket model of texture segregation. The retinal image is first processed by a bank of linear spatial filters. Then, some form of nonlinearity is applied. Here, a pointwise fullwave rectification is indicated. Next, a second stage of linear spatial filtering is applied to enhance the texture-defined edge. Subsequent decision processes are dependent on the particular psychophysical task under study.

4. Back pocket model. A. An orientation-defined edge. B. The result of the application of a linear, vertically-oriented, spatial filter. C. The result of a pointwise nonlinearity (squaring). D. A second, large-scale, vertically-oriented, spatial filter yields a peak response at the location of the texture-defined border in A.

5. The 2nd-order contrast sensitivity function. A. This figure is constructed using a modulator image to additively combine vertical and horizontal noise images (Landy and Ternes 1995). The modulator, shown as a function above the texture, has a spatial frequency that increases from left to right, and its contrast increases from bottom to top. Large

modulator values result in a local texture dominated by vertically-oriented noise, and small values by horizontally-oriented noise. Note that threshold modulation contrast is nearly independent of spatial frequency. B. Example data from a forced-choice modulation contrast detection experiment using sine wave modulators of noise patterns.

6. Texture appearance, representation and extrapolation. In the technique of Portilla and Simoncelli (2000), a texture is first analyzed using a bank of linear spatial filters varying in preferred spatial frequency and orientation. A set of statistics, both 1st-order and correlational, on that set of filter responses becomes the representation of the given texture. This representation may be used to generate new instances of the texture. In each panel, the inset square is the original texture, and the rest of the image is new texture generated using the technique.

Figure 1

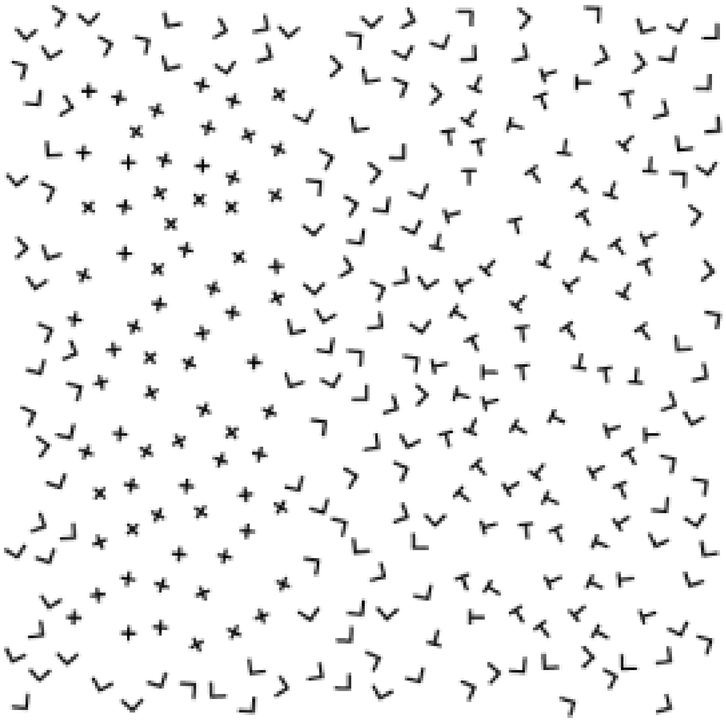


Figure 2

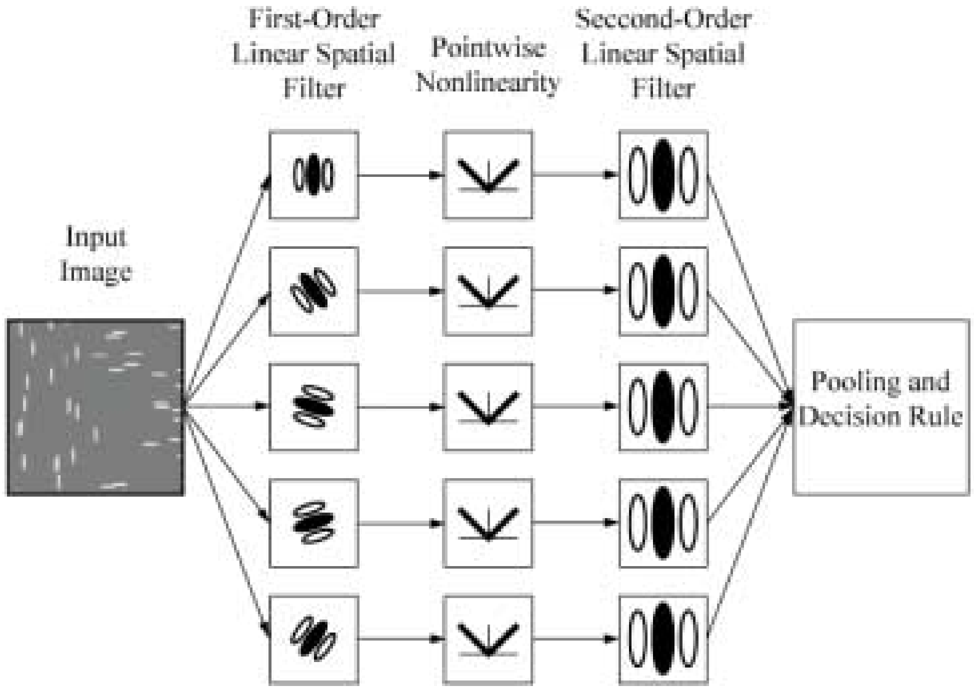

First-Order
Linear Spatial
Filter
Pointwise
Nonlinearity
Seccond-Order
Linear Spatial
Filter
Input
Image
Pooling and
Decision Rule

Figure 3

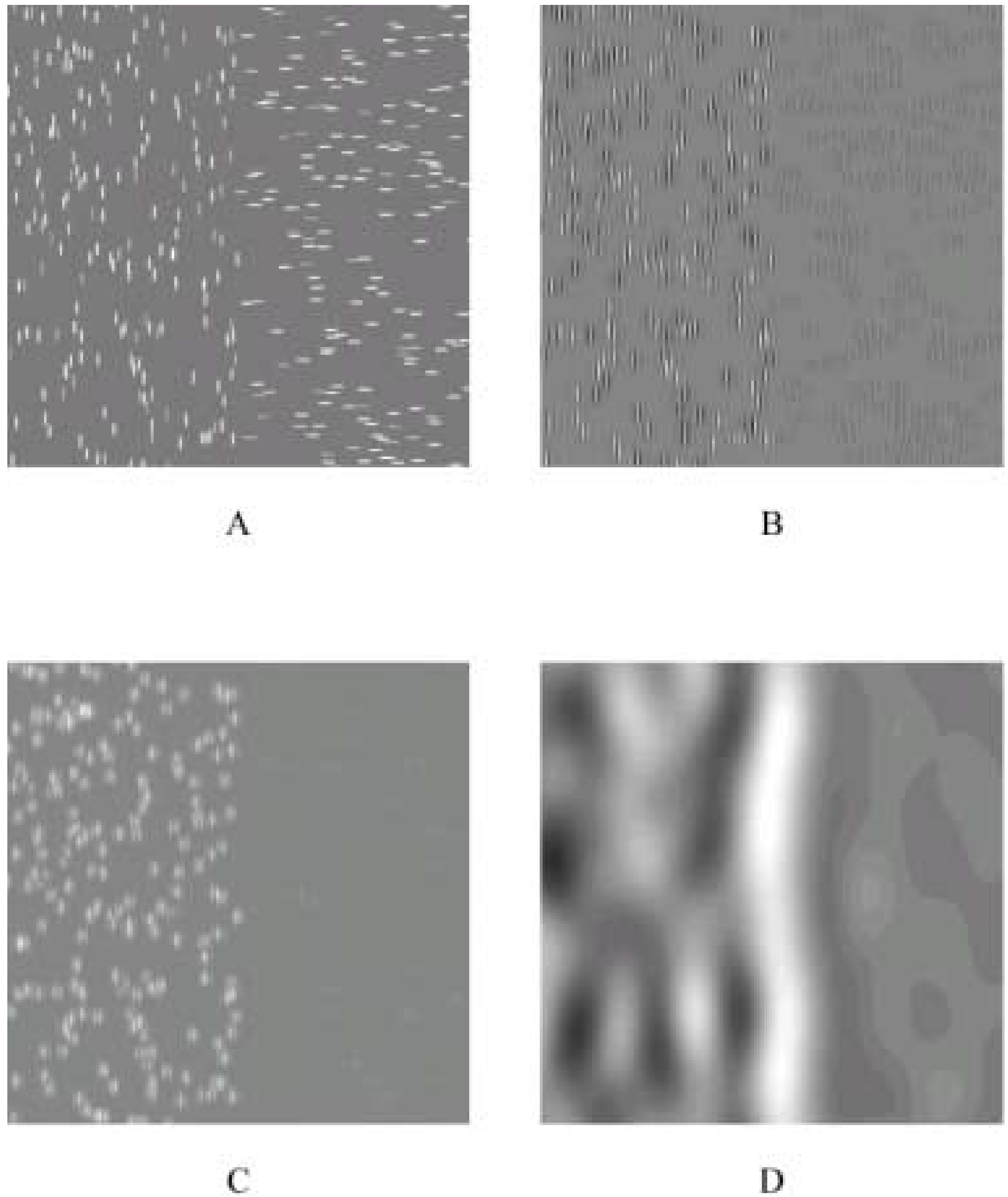
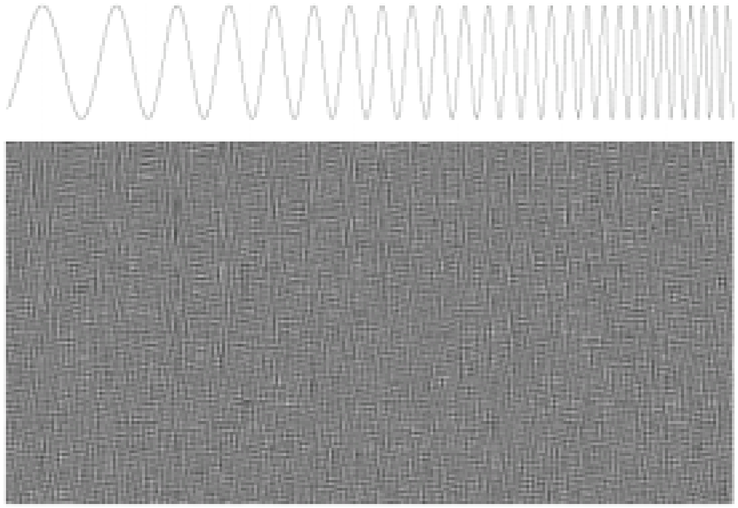
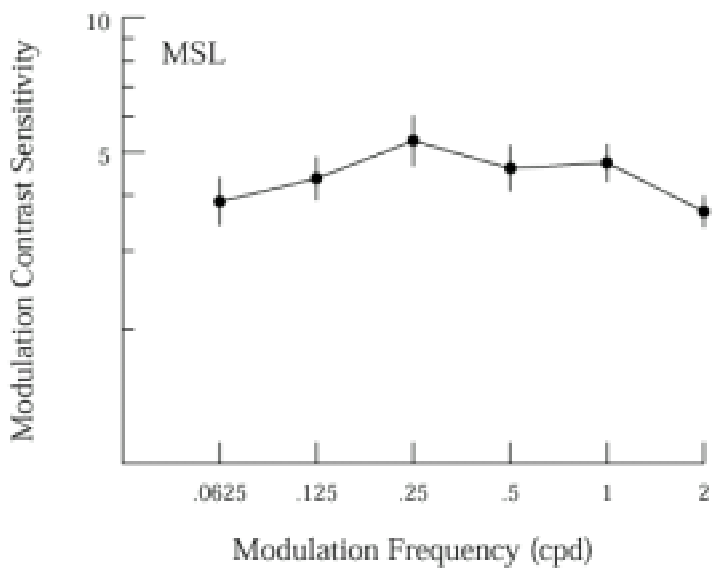


Figure 4

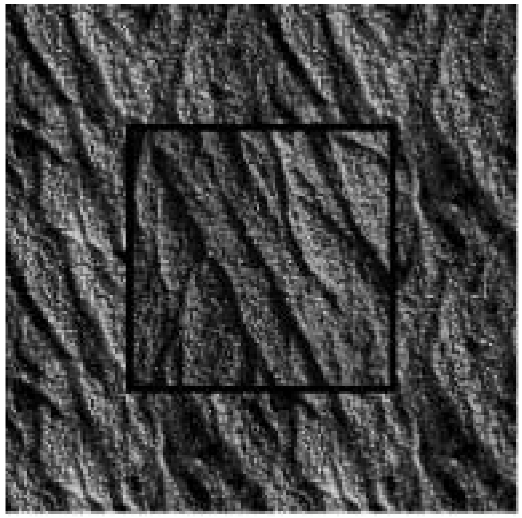


A

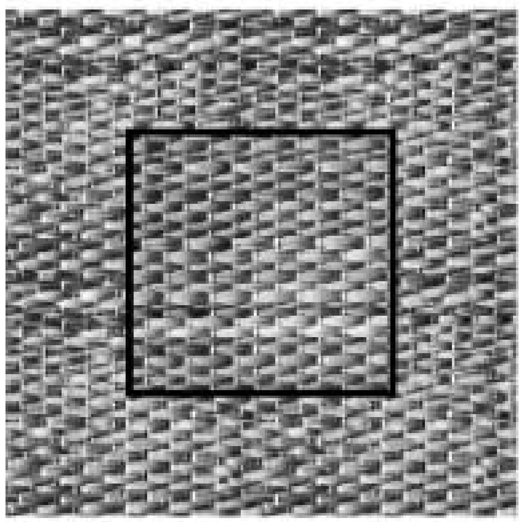

10
MSL
5
Modulation Contrast Sensitivity
.0625
.125
.25
.5
1
2
Modulation Frequency (cpd)

B

Figure 5



A



B

Figure 6