# Texture analysis and perception

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

Number of words (including references): 8790
Number of figures: 7
Short title: Texture Perception

Address correspondence to:
Michael S. Landy
Dept. of Psychology
New York University
6 Washington Place, Rm. 961
New York, NY 10003
(212) 998-7857
fax: (212) 995-4018
landy@nyu.edu

What is visual texture? The Oxford English Dictionary provides several definitions of the word. The first definition comes from the etymology; the word comes from Latin "textūra" meaning "a weaving". Later definitions elaborate and extend this theme, including "any natural structure having an appearance or consistence as if woven; a tissue; a web, e.g. of a spider" and, finally, "The constitution, structure, or substance of anything with regard to its constituents or formative elements." In other words, texture is "stuff" (Adelson & Bergen, 1991) that is composed of smaller elements. This latter definition thus encompasses all the sorts of images that we typically describe as visual texture, whether regular and clearly human-made (wallpaper, fabric, floor tiles), irregular deriving from natural variations in surface reflectance (wood grain), resulting from the interplay of illumination and a variegated 3-d surface (plaster) or from the separate elements in a scene (a field of wheat), or combinations of these factors. In each case, the image consists of repeating smaller elements (individual design elements in a tile, single small bumps in a plaster wall, single growth lines in wood) that repeat, possibly with variation, across the textured image. One can even extend this definition to encompass the idea of auditory "texture" (McDermott & Simoncelli, 2011).

In this chapter I will concentrate on the broad array of research on the visual coding and perception of texture that has appeared in the last decade or so since my previous review of this topic (Landy & Graham, 2004). For earlier reviews, see Bergen (1991) and Landy (1996). As we will see, there has been a tremendous amount of work in this area in the intervening years, requiring major revisions in our understanding of the processing of visual texture and its neural underpinnings. I begin by reviewing psychophysical work on the detection of texture-defined boundaries ("texture segregation") and its importance for the development of biologically grounded models of texture coding. I relate models of texture segregation to models of texture appearance based on image statistics. Then, I discuss recent work on the use of surface texture for visual estimation of object shape and surface properties in 3-d scenes. The chapter concludes with a survey of studies of the neural processing of texture. In recent years, neurophysiological work on texture coding has ranged from single-cell physiology and optical imaging to visual evoked potentials (VEP) and functional magnetic resonance imaging (fMRI).

## Models of texture segregation

### *The back-pocket model*

Consider the image of the boardwalk in Figure x.1. Simple cells in primary visual cortex have receptive fields that can be modeled as applying an orientation-selective, linear spatial filter to the retinal image. This computation emphasizes oriented lines in an image, such as those between neighboring boards in the walkway. But, linear filtering cannot specifically respond to the texture-defined borders in this image, such as the boundaries between differently oriented boardwalk sections. This is because, on a coarse scale, the average luminance is identical on either side of these boundaries. What differs across such a boundary is the average local orientation (i.e., a local average statistic of the image). What kind of computation can reveal such boundaries, which are clearly salient to human observers?

A number of researchers, first clearly described by Bergen and Adelson (1988), have suggested that texture-defined boundaries may be computed by the sort of mechanism we have called the

"back-pocket" model of texture segregation (Chubb & Landy, 1991). In its simplest form, the back-pocket model (Figure x.2) consists of a sequence of three feedforward stages: filter, rectify, filter (FRF; sometimes also called LNL for linear, nonlinear, linear). To make this concrete, Figure x.3 illustrates this sequence of operations on a section of the boardwalk image. The first linear spatial filter is tuned for orientation and spatially frequency so that it responds preferentially to one of the two constituent textures. In Figure x.3B, a vertically tuned filter has been applied. However, a linear filter will typically respond both with strong positive values (when the receptive field is located so that excitatory regions align with brighter parts of the image) and negative values (when aligned with darker image regions). Thus, a spatial average of these linear filter response will be identical on either side of the texture edge. In Figure x.3B the average on the left and right sides is identical (gray, indicating a filter output of zero). The nonlinearity (e.g., a threshold or point-wise squaring operation) converts these regions of high *variance* in response to regions producing higher *average* response (Figure x.3C, where now black represents a value of zero). After this stage a "second-order" linear filter, typically of a substantially larger scale, will respond strongly to the texture-defined edge (Figure x.3D). The kinds of image structure delineated by FRF models is quite prevalent in natural images and is not correlated with first-order, luminace-defined structure (Schofield, 2000).

Some recent work serves to support the FRF model and establish its basic properties. In much of this work, researchers have developed psychophysical techniques that seek to measure the variety and filtering properties of the second-order filters using experimental methods analogous to those classically used to measure the properties of first-order filters (i.e., of spatial frequency channels; Graham, 1989). In these studies, second-order stimuli involve modulations of texture that observers must detect or discriminate. Contrast sensitivity for the detection of orientation modulation (OM; see Figure x.4 for example stimuli) of is relatively flat over a large range of spatial frequency (Landy & Oruç, 2002). Discrimination of second-order modulation depth with OM textures results in subthreshold summation, i.e., thresholds decrease in the presence of pedestal contrast compared to no pedestal (the classic "dipper function" of first-order contrast discrimination). This improvement in threshold indicates either a further nonlinearity (i.e., FRFR) or a reduction in spatial uncertainty due to the presence of the pedestal (Landy & Oruç, 2002). Subthreshold summation (i.e., dipper functions) has been observed using OM stimuli as well as spatial-frequency-modulated (FM) and contrast-modulated (CM) textures with no facilitation across texture types (Kingdom, Prins, & Hayes, 2003). These results are reasonably consistent with FRF predictions, further supporting the FRF model. Kingdom and colleagues also argued that normalization (e.g., Heeger, 1992), sometimes referred to as surround suppression, was needed to explain their results (see following section for further discussion of normalization processes in texture perception). As in the first-order case, the second-order contrast sensitivity function is an envelope over multiple, narrower bandwidth second-order spatial frequency channels. This has been demonstrated, for example, using a second-order version of the summation technique originally used by Graham and Nachmias (1971) to infer the existence of first-order channels. Using this technique with OM stimuli, Landy and Oruç (2002) found that second-order channels had a similar bandwidth (ca. one octave) to first-order channels. Data from a critical-band masking experiment also suggest that second-order channels of modest bandwidth (ca. 1.5 octaves) are used to identify texture-defined letters (Oruç, Landy, & Pelli, 2006). Prins and Kingdom (2003) tested the FRF model using a "2×2 task" (2 tasks × 2 intervals) with OM, FM or CM stimuli. Subjects were required to both detect the interval in

which there was second-order modulation and identify the type of modulation in that stimulus (OM vs. CM, OM vs. FM, or CM vs. FM). For two of the three pairings (except CM vs. FM), the thresholds for detection and identification were identical. This is consistent with a model in which FRF channels are "labeled" with the first-order spatial filter they use, so that knowledge of the specific channel that detected the stimulus also identifies which type of modulation it contained. Ellemberg and colleagues (2006) also used the 2×2 task in the manner of Watson and Robson (1981). The two discriminanda differed in second-order spatial frequency or orientation, and they increased the difference in spatial frequency or orientation until detection and discrimination modulation thresholds were identical, indicating the two stimuli were detected by distinct channels. They found that the number of distinct second-order orientation and spatial-frequency channels was similar to the number of first-order channels (at least for lower spatial frequencies). Consistent with the FRF model, adaptation to first-order (grating) patterns is most effective at impeding texture segregation when the adapter spatial frequency matches the tuning of the first-order filter for the FRF channel that is the most informative for the task (Prins & Kingdom, 2006).

In the basic FRF model, the second-order linear filter uses the same input (i.e., from the same first-order filter) for both its excitatory and inhibitory regions. Both Prins and colleagues (2003) and Graham and Wolfson (2004) provide evidence against the existence of more complex orientation-opponent and double-opponent arrangements in which two orthogonal first-order channels, after rectification, contribute to the responses of a second-order channel.

***Extensions of the model***

Recent psychophysical work in texture discrimination has sought to flesh out and extend the basic FRF model. For example, an FRF model that computed local texture "energy" (with a square-law nonlinearity) should be insensitive to the local phase of the carrier patterns. But, textures with a preponderance of even phase (0 or 180 deg, i.e., positive or negative cosine phase) patterns segregate when placed on a random phase background (Hansen & Hess, 2006), suggesting that the sign of cosine phase *is* encoded (although a local luminance nonlinearity might provide an alternative explanation of these results). A number of studies have followed up the observation that orientation-defined texture edges are easiest to detect when one of the bordering textures shares a local (carrier) orientation with the second-order texture edge (Wolfson & Landy, 1995). This effect may be due to texture elements aligned with and abutting the texture edge, while texture elements interior to a texture-defined figure aid in segregation when orthogonal to the boundary (Giora & Casco, 2007). Segregation improves further when interior texture elements align, grouping to form long "chains" (Harrison & Keeble, 2008). Alberti and colleagues (2010) confirmed this finding, but also found that collinear texture elements are detrimental to motion-defined segregation.

In addition to linear spatial filtering and point nonlinearities, normalization, in which the response of one neuron is normalized by a local, pooled response of nearby neurons (e.g., Heeger, 1992), is likely a widespread computation used throughout cortical circuits (Carandini & Heeger, 2011). This leads to an extension of the FRF model to $F_1R_1N_1F_2R_2N_2$, in which "N" stands for a normalization stage. Normalization has been suggested as an important component of texture processing (e.g., Graham, 1991; Graham & Sutter, 2000). Several recent papers

suggest that first-order normalization ($N_1$) is needed to explain psychophysical performance using second-order stimuli (Kingdom et al., 2003; Li & Zaidi, 2009; Motoyoshi & Nishida, 2004). The Motoyoshi and Nishida (2004) result also suggests that there are FRF (or FRNF) channels in which the first-order filter is not orientation-tuned; several other studies are consistent with this observation (Motoyoshi & Kingdom, 2007; Prins, 2008). Schade and Meinecke (2009, 2011) demonstrate a kind of second-order metacontrast masking. They found that detection of a second-order pattern was impeded if it was followed by a masker containing a task-irrelevant texture-defined border. This might be modeled as normalization with a normalization pool that is extended in time. Finally, evidence is beginning to emerge for second-order normalization as well, both in terms of surround suppression's effects on sensitivity (Wang, Landy, & Heeger, 2011) and on perceived contrast (Ellemberg, Allen, & Hess, 2004).

Texture segregation experiments have used a wide variety of natural and computer-generated stimuli in which texture-defined edges are signaled by modulations of contrast, orientation, spatial frequency, line terminations, etc. Thus, modelers are motivated to find simple explanations for these complex phenomena that don't require positing a new texture "channel" for each newly defined type of stimulus (Prins & Kingdom, 2003). This efficiency might result from channels that are "carrier invariant", responding to several forms of texture modulation (e.g., CM, OM and FM) or, in the case of CM stimuli, responding to contrast modulation of a range of carrier patterns. Two studies found cross-adaptation between different kinds of contours including both first-order (luminance-modulated or LM) and second-order (CM, OM and illusory contour or IC), suggesting a common mechanism is used to encode all of these types (Filangieri & Li, 2009; Hawley & Keeble, 2006). However, other studies (e.g., Larsson, Landy, & Heeger, 2006) suggest that LM mechanisms are separate from those coding second-order patterns.

The FRF model is consistent with the observation that, for OM patterns, detectability of an edge is a function of the "orientation gradient" (Landy & Bergen, 1991; Nothdurft, 1985), i.e., the change in orientation across the edge divided by the distance over which it occurs. However, in an interesting series of papers (Ben-Shahar, 2006; Ben-Shahar & Zucker, 2004; Ben-Yosef & Ben-Shahar, 2008), Ben-Shahar showed that perception of boundaries in OM patterns is more complex than this. He examined orientation flow patterns in which orientation changes continuously across the image and found that contours are perceived in locations in which the orientation gradient is constant across the image (Figure x.5). He suggested a theory for the encoding of second-order spatial patterns in which two local measures of orientation change are computed: normal and tangential curvature of the orientation flow. This appears consistent with the finding that structure in texture made of random pairs of oriented texture elements is most salient when the pairs consist of co-circular elements, i.e., support a local orientation flow (Motoyoshi & F. A. A. Kingdom, 2010). It may also relate to the many behavioral results suggesting that texture-defined edges are most salient when a texture on one side of the edge has orientation content oriented parallel to the edge (Casco, Grieco, Campana, Corvino, & Caputo, 2005; Giora & Casco, 2007; Wolfson & Landy, 1995).

There has been substantial work on developing biologically inspired models of texture segregation. Two modeling efforts take distinct approaches to this task. In the model of Thielscher and Neumann (2003, 2005, 2007), a feed-forward chain of filtering and rectification, representing computations in cortical areas V1, V2 and V4, is modulated by feedback signals

from V4 back to V1. They show that the feedback signals are necessary to account for human performance. On the other hand, the models of Zhaoping Li (Jingling & Zhaoping, 2008; Zhaoping, 2003; Zhaoping, Guyader, & Lewis, 2009; Zhaoping & May, 2007) are primarily feedforward, with some interaction between model neurons within a single cortical area.

### *Texture and visual attention*

It is well-known that visual attention can improve performance in low-level visual tasks. A series of studies by Carrasco and colleagues (Barbot, Landy, & Carrasco, 2011; Carrasco & Yeshurun, 2009; Yeshurun & Carrasco, 2000, 2008; Yeshurun, Montagna, & Carrasco, 2008) indicates that both exogenous attention, induced involuntarily by sudden stimulus change in the periphery, and endogenous attention, the top-down allocation of spatial attention to a region of interest, have strong effects on second-order processing. For example, consider the central performance drop, in which performance at a texture-segregation task is best at a near-peripheral location and worse at the fovea (Gurnsey, Pearson, & Day, 1996; Kehrer, 1989). It has been suggested that this task requires a particular scale of processing for optimal performance. At each eccentricity, there is a scale or spatial frequency to which the visual system is most sensitive. This preferred spatial frequency is highest at the fovea and drops with increasing eccentricity, so that it is "just right" for this texture-segregation task in the near-periphery. With exogenous attention, performance improves in this task at peripheral locations and worsens at the fovea, consistent with the view that exogenous attention improves spatial resolution wherever it is directed (Yeshurun & Carrasco, 2000). In locations where the optimal scale is already too high for the task (near the fovea), the increase in resolution makes performance worse, but where it is too low (farther in the periphery), performance improves with attention. However, endogenous attention is more flexible and improves performance at all eccentricities in this task (Yeshurun et al., 2008). In addition to improving resolution, attention also improves second-order contrast sensitivity (Barbot, Landy, & Carrasco, 2011).

### *Development*

The developmental sequence for perception of texture-defined patterns lasts throughout childhood. Global orientation leads to significant visual evoked potentials (VEPS) in 2-5 month olds, and global orientation patterns (e.g., circles vs. pinwheels) lead to responses by 6-13 months (Norcia et al., 2005). On the other hand, behavioral performance in simple spatial discrimination tasks reaches adult levels by age 12 for luminance-defined patterns, but continues to improve significantly after age 12 for second-order, contrast-defined patterns (Bertone, Hanck, Cornish, & Faubert, 2008; Bertone, Hanck, Guy, & Cornish, 2010).

## Texture coding and statistics

Early work on texture segregation was concerned with what "order" of statistics were encoded by the visual system (Caelli & Julesz, 1978; Caelli, Julesz, & Gilbert, 1978; Julesz, Gilbert, Shepp, & Frish, 1973; Julesz, Gilbert, & Victor, 1978). In this work the term "second-order statistics" meant dipole statistics, in which one computes the distribution of gray values seen at the ends of a dipole (a line segment of a particular length and orientation) positioned at a random position on the texture, with one distribution per choice of dipole length and orientation. Third

order statistics are analogous, except that one randomly positions triangles (of a particular shape and size) on the texture. Counterexamples were found to the hypotheses that 2nd-order or 3rd-order statistics determined texture appearance (Caelli & Julesz, 1978; Julesz et al., 1978).

Understanding texture perception by looking at image statistics remains an important approach. Even the encoding of first-order statistics (graylevel distributions) is complex. Chubb and colleagues (1994, 2004) showed that graylevel distributions are encoded by human observers using only three statistics: mean (average luminance), variance (i.e., contrast), and a third statistic they term "blackshot", because it is primarily sensitive to distinctions among the darkest graylevels in the scene. Both 1st- and 4th-order statistics are encoded in a continuous, graded fashion (Victor & Conte, 2004). Differences in image statistics that are most salient are also those that are most informative in natural images (Tkačik, Prentice, Victor, & Balasubramanian, 2010), i.e., are principal components for the 1st-order statistics, or are maximum-entropy components for 4th-order statistics. Segregation of texture-defined borders between natural image textures requires 3rd- or higher-order statistics, because performance differs for phase-scrambled textures that preserve 2nd- but not higher-order statistics (Arsenault, Yoonessi, & Baker, 2011). The encoding of phase was also examined by Emrith and colleagues (2010). They varied the degree of phase scrambling and used the maximum-likelihood difference scaling method (Knoblauch & Maloney, 2008) to determine perceptual sensitivity to the degree of phase scrambling. Observers were most sensitive to intermediate degrees of phase scrambling, and the authors suggested that a model with a third stage of nonlinearity and linear filtering (i.e., FRFRF) would be required to account for this behavior.

Many natural textures can be characterized by the distribution of local orientation across the texture (cf. Figure x.1) , just as experimenters have often used OM textures in their experiments. When observers are asked to judge the mean orientation of textures consisting of collections of oriented local texture elements (oriented Gabor patches), sampling efficiency (the fraction of the texture elements observers used for the judgment) depends only on the number of texture elements in the stimulus (Dakin, 2001). It is independent of the density of texture elements or overall size of the texture image. Girshick and colleagues (2011) also asked observers to discriminate the mean orientation of patterns containing oriented texture elements. In their study, orientation uncertainty (the variance) was varied to measure biases in orientation estimation and infer the underlying orientation prior (in the Bayesian-estimation sense) used by observers. They found that human observers use a prior that peaks at the cardinal orientations (horizontal and vertical) similar to the distribution of orientations in natural images, and that a neural population code that emphasized the cardinal orientations (with more neurons tuned to these orientations possessing narrower orientation bandwidth) could account for their results. If the stimuli to be discriminated include both first-order (luminance-defined) and second-order (contrast-defined) texture elements, observers are not able to combine the two cues and use first-order luminance information even when it was less informative (Allen, Hess, Mansouri, & Dakin, 2003). However, combinations of spatial frequency and orientation cues do lead to improved psychophysical performance from cue integration [much as Landy and Kojima (2001) had found previously using a localization task] and stronger VEP signal during the task (Bach, Schmitt, Quenzer, Meigen, & Fahle, 2000).

How are all these studies of image statistics related to the FRF and similar models? The FRF computation is an image transformation. A pixel in the output image is a linear combination of the pixels in the intermediate image that results from the first linear filter and rectification. The pixels of this intermediate image are a nonlinear combination of input image pixels. For example, if the nonlinearity is a square law, then output pixels are a quadratic polynomial function of input pixels. For more severe nonlinearities, this function will have higher polynomial order. Thus, the task of determining the number and types of FRF channels in the visual system is tantamount to determining the number and types of image statistics we encode. This connection is clearest in the work on texture-synthesis models that take an input image, compute a collection of statistics on that image, and then try to synthesize new images that match those statistics. If the statistics correspond to those human observers use to determine texture appearance, then the synthesized images should look like they are made of the same "stuff" as the original image. Two such texture-synthesis models have been developed that begin with a multi-scale, oriented image pyramid, in analogy to the multiple spatial frequency and orientation channels in human vision. The first of these (Heeger & Bergen, 1995) uses only first-order statistics (the histogram of values in each spatial channel and the pixel histogram). A more recent such model (Portilla and Simoncelli, 2000) also includes various correlational statistics between each pyramid value and its spatial, scale and orientation neighbors. Figure x.6 shows example syntheses by this model, which works quite well on a wide variety of textures.

The Portilla and Simoncelli (2000) model (the "PS" model) has been examined as a potential model for human visual coding in several psychophysical studies. Balas (2006) asked observers to discriminate the odd texture in a group of three synthesized textures, asking whether various subsets of the statistics in the PS model are necessary for describing human performance. Humans can detect differences in textures when any subclass of statistics are left out of the model, but the salience of the difference depends on the class of statistics that are omitted (the marginal statistics, similar to those in the Heeger/Bergen model, "HB", are the most salient) and on the particular natural texture being synthesized. Surprisingly, the power spectrum is a better predictor of human texture classification than either the HB or PS model (Balas, 2008). This shows that the ability of a model to synthesize textures that humans perceive as matching an original texture does not imply that a model can predict human texture classification. In a particularly interesting paper, Balas and colleagues (2009) suggested that the sorts of statistics in the PS model are all that is encoded in peripheral vision. They did this by showing that performance in a peripheral visual crowding task was similar to performance in which the peripheral search patterns were replaced by "mongrel" patterns (i.e., resynthesized images using the PS model, Figure x.7) but observed foveally. Following this logic a step further, Freeman and Simoncelli (2011) created image "metamers" by replacing peripheral portions of an image with PS-model resyntheses. To do this, they had to choose an optimal image patch size for measuring local image statistics in each region of the image to drive the texture-synthesis process so that the full synthetic image appeared identical to original image. By using a cortical magnification model in which the pooling region for statistics grows with eccentricity, they found that the cortical scaling that could be supported with this manipulation matched that of cortical area V2, suggesting that texture statistics might be encoded in that brain area.

## Shape from texture

Surface texture is one of the many pictorial cues to 3-dimensional shape, and work in this area has often centered on the assumptions observers make about surface textures to compute shape from texture, and what computation is used (e.g., Knill, 1998). Li and Zaidi (Li & Zaidi, 2000, 2001a,b, 2003, 2004; Zaidi & Li, 2002) have developed a simple idea concerning the aspects of texture that support 3-d perception. Originally working with developable (i.e., singly curved) surfaces and textures synthesized as sums of sine wave gratings, they found that veridical perception of qualitative aspects of 3-d shape (e.g., concave vs. convex, etc.) required that the texture include contours along the orientation of maximum curvature, and that those spectral components need to be relatively isolated (without additional components at nearby orientations). In later work, this idea was confirmed with natural textures and with doubly curved 3-d shapes. Todd and colleagues have produced a series of counterexamples and experimental data they see as contradicting this theory (Thaler, Todd, & Dijkstra, 2007; Todd & Oomes, 2002; Todd & Thaler, 2010; Todd, Oomes, Koenderink, & Kappers, 2004; Todd, Thaler, Dijkstra, Koenderink, & Kappers, 2007), resulting in an extended argument between the two groups played out in the journals that doesn't seem to have been resolved to either group's satisfaction. Todd and colleagues have been concerned with the precise shapes observers perceive and, in particular, in the circumstances under which observers misperceive shape. They have suggested several heuristic computations that accord with these patterns of nonveridical shape perception, though these heuristics mainly apply to textures that are collections of individual, nonoverlapping texture elements. This class does not include the sorts of noise textures examined by Li and Zaidi. On the other hand, Li and Zaidi's work has been concerned mainly with qualitative shape perception (e.g., the sign of perceived slant or curvature) and not on perceived depth per se.

## Surface properties from texture

The definition of texture suggests that texture should be regarded as a surface property. Indeed, one can speak of 3-d texture as a property of a surface akin to shape but at a far smaller scale. Perception of roughness depends on viewing conditions. It seems that observers use the amount of cast shadow in an image of a rough surface as a cue to surface roughness, even though cast shadow can vary with such things as the direction of illumination for a constant amount of surface roughness (Ho, Landy, & Maloney, 2006; Ho, Maloney, & Landy, 2007; Landy, Ho, Serwe, Trommershäuser, & Maloney, 2011). Surface roughness (or bumpiness) and gloss are not coded independently, in the sense that varying one affects perception of the other surface property (Ho et al., 2008). Padilla and colleagues (2008) looked at perceived surface roughness of 3-d surfaces with a height function that was fractal (i.e., had a $1/f^{\beta}$ spectrum), and found that roughness increased with decreasing value of spectral slope $\beta$.

## Neural coding of texture

### *Physiological studies in animals*

There has been a modest effort to determine how second-order patterns are encoded in cortex and in which brain areas the computation takes place. In the cat, Baker and colleagues have

suggested in a series of papers (Song & Baker, 2006; 2007; Zhan & Baker, 2006; 2008) that second-order patterns (both CM and IC) are encoded in area 18. In single-cell experiments, they find that a subpopulation of area 18 cells are tuned for the orientation of the modulator and that many behave in a carrier-invariant fashion. Using optical imaging, they find that there are similar first- and second-order orientation maps in area 18. These findings only occur for second-order patterns in which the carrier pattern has a spatial frequency above the linear (first-order) passband of the cell or cortical area. They suggest these results are consistent with the behavior of an FRF model. However, these results remain controversial. In rhesus macaque area V2 (perhaps a region homologous to cat area 18), cue-invariant second-order responses are nearly absent (El-Shamayleh & Movshon, 2011), and in a small set of cat area 18 cells these authors were unable to replicate Baker's results. They suggested that tuning for second-order patterns might emerge gradually across several cortical regions rather than all at once in area V2. In area 18 of the cat cells have similar disparity tuning to the contrast envelope of second-order patterns as they have to luminance disparity (Tanaka & Ohzawa, 2006). Many cells in cat area 17 are tuned for the second-order orientation and spatial frequency of CM gratings and this second-order tuning arises simply by the cell having a nonclassical surround that is asymmetrically placed or organized relative to the classical receptive field (Tanaka & Ohzawa , 2009). Finally, in awake macaque, population responses from area V4 appear to be capable of representing 3-d textures (Arcizet, Jouffrais, & Girard, 2008). In this study, macaques viewed images from a database of 3-d surface textures (the CURET database) while responses of V4 cells were recorded. Many cells were tuned to subsets of these images, and a clustering algorithm on the population responses was able to cluster individual textures independent of the illumination direction as well as a similar algorithm applied to a pyramid Gabor representation of these images.

***Visual evoked potential studies***

Several groups have studied the visual evoked potential (VEP) response in humans corresponding to the segregation of textures (the "tsVEP"). The tsVEP has two components: The early (100 ms) component is task-independent while the later (230 ms) component is abolished if the subject performs an attention-distracting task (Heinrich, Andrés, & Bach, 2007). The tsVEP is stronger for more salient textures, e.g., when the texture elements are parallel to the texture edge (Casco, Grieco, Campana, Corvino, & Caputo, 2005). The tsVEP really is a response to texture *per se* and is not just due to the first-order content of the image, because a change in orientation content leads to a VEP signal that cuts off at 17 Hz, whereas the tsVEP cuts off at 12 Hz (Lachappelle, McKerral, Jauffret, & Bach, 2008). A series of studies by Appelbaum and colleagues (Appelbaum, Wade, Pettet, Vildavski, & Norcia, 2008; Appelbaum, Wade, Vildavski, Pettet, & Norcia, 2006) used a variety of second-order patterns in which they tagged responses to the figure and ground textures by cycling different versions of each at different frequencies ("frequency tagging"). Responses to the figure begin in area V1 and continue on to the lateral occipital complex (LOC) independent of the texture cue, whereas responses to the background appear in more medial regions. Nonlinear interaction responses (e.g., at frequency f1+f2, where f1 and f2 are the figure and ground frequencies, respectively) signal the texture-defined boundary, and disappear when a luminance-defined gap is introduced between figure and ground.

### *fMRI studies*

There have been a number of studies that attempt to link particular brain areas to texture processing using fMRI. In the simplest of these, subjects are shown textured vs "equated" non-textured objects, looking for the areas that respond more strongly to texture or adapt to recently viewed texture (Cant, Arnott, & Goodale, 2009; Cavina-Pratesi, Kentridge, Heywood, & Milner, 2010a; 2010b; Stylianou-Korsnes, Reiner, Magnussen, & Feldman, 2010) or to areas that respond more strongly to shape defined by texture (Georgieva, Todd, Peeters, & Orban, 2008). A variety of brain areas are implicated for texture processing by these studies, including the posterior collateral sulcus and right inferior parietal lobule, consistent with some neurophysiological findings. For shape-from-texture, several brain regions respond more strongly to stimuli that appear 3-d due to texture cues, including LOC, lateral occipital sulcus and the inferior temporal gyrus.

In studies such as these, especially those using subtraction contrasts, it is hard to determine what function is carried out by the revealed brain areas, i.e., what aspects of the textured stimuli lead to stronger responses. Single-unit studies suggest that only a small subset of neurons in early visual cortical regions encode second-order patterns, so fMRI methods are required that can demonstrate the function of a potentially small subset of the neurons in a voxel. fMRI adaptation is an experimental technique suited to providing evidence for tuning of a subset of neurons in a voxel and this technique has demonstrated second-order orientation and spatial frequency tuning in several visual cortical areas (Hallum, Landy, & Heeger, 2011; Larsson et al., 2006; Montaser-Kouhsari, Landy, Heeger, & Larsson, 2007). In all of these studies, the adaptation index (the reduction of response due to adaptation to the same orientation or frequency of modulation compared to adaptation to a different modulation) increases steadily from V1 through downstream visual areas. Consistent orientation-selective adaptation to LM, CM and OM patterns was found in V1, V2, V3, V3A/B, LO1, hV4 and VO1 (Larsson et al., 2006). For first-order (LM) patterns, adaptation was no stronger in downstream areas than in V1, suggesting that the adaptation effect derived from V1 processing and was inherited by downstream areas. In contrast, stronger adaptation was found in VO1 for OM patterns and additionally in V3A/B and LO1 for CM patterns, suggesting further downstream processing, consistent with the "gradualist encoding" suggestion of El-Shamayleh and Movshon (2011). A similar gradual increase in adaptation was found for illusory contours, with the addition of significant adaptation in areas V7 and LO2 (Montaser-Kouhsari, Landy, Heeger, & Larsson, 2007). Significant tuning for second-order spatial frequency was found in V1, V2, V3 and V4 (Hallum, Landy, & Heeger, 2011). However, modeling these responses suggested that first-order normalization was involved in second-order spatial frequency tuning and that the bulk of the adaptation may have occurred in V1. Finally, fMRI responses increase in several cortical areas with increasing salience of the texture-defined border (achieved by increasing the orientation gradient), especially in areas hV4, V3 and LOC, (Thielscher, Kölle, Neumann, Spitzer, & Grön, 2008).

**References**


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

Alberti, C. F., Pavan, A., Campana, G., & Casco, C. (2010). Segmentation by single and combined features involves different contextual influences. *Vision Research*, *50*, 1065–1073.

Allen, H. A., Hess, R. F., Mansouri, B., & Dakin, S. C. (2003). Integration of first- and second-order orientation. *Journal of the Optical Society of America A*, *20*, 974–986.

Appelbaum, L. G., Wade, A. R., Pettet, M. W., Vildavski, V. Y., & Norcia, A. M. (2008). Figure-ground interaction in the human visual cortex. *Journal of Vision*, *8*(9):8, 1–19.

Appelbaum, L. G., Wade, A. R., Vildavski, V. Y., Pettet, M. W., & Norcia, A. M. (2006). Cue-invariant networks for figure and background processing in human visual cortex. *Journal of Neuroscience*, *26*, 11695–11708.

Arcizet, F., Jouffrais, C., & Girard, P. (2008). Natural textures classification in area V4 of the macaque monkey. *Experimental Brain Research*, *189*, 109–120.

Arsenault, E., Yoonessi, A., & Baker, C. (2011). Higher order texture statistics impair contrast boundary segmentation. *Journal of Vision*, *11*(10):14, 1–15.

Bach, M., Schmitt, C., Quenzer, T., Meigen, T., & Fahle, M. (2000). Summation of texture segregation across orientation and spatial frequency: electrophysiological and psychophysical findings. *Vision Research*, *40*, 3559–3566.

Balas, B. (2008). Attentive texture similarity as a categorization task: Comparing texture synthesis models. *Pattern Recognition*, *41*, 972–982.

Balas, B. J. (2006). Texture synthesis and perception: using computational models to study texture representations in the human visual system. *Vision Research*, *46*, 299–309.

Balas, B., Nakano, L., & Rosenholtz, R. (2009). A summary-statistic representation in peripheral vision explains visual crowding. *Journal of Vision*, *9*(12):13, 1–18.

Barbot, A., Landy, M. S., & Carrasco, M. (2011). Exogenous attention enhances 2nd-order contrast sensitivity. *Vision Research*, *51*, 1086–1098.

Ben-Shahar, O. (2006). Visual saliency and texture segregation without feature gradient. *Proceedings of the National Academy of Sciences USA*, *103*, 15704–15709.

Ben-Shahar, O., & Zucker, S. W. (2004). Sensitivity to curvatures in orientation-based texture segmentation. *Vision Research*, *44*, 257–277.

Ben-Yosef, G., & Ben-Shahar, O. (2008). Curvature-based perceptual singularities and texture saliency with early vision mechanisms. *Journal of the Optical Society of America A*, *25*, 1974–1993.

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

Bergen, J. R., & Adelson, E. H. (1988). Early vision and texture perception. *Nature*, *333*, 363–364.

Bertone, A., Hanck, J., Cornish, K. M., & Faubert, J. (2008). Development of static and dynamic perception for luminance-defined and texture-defined information. *Neuroreport*, *19*, 225–228.

Bertone, A., Hanck, J., Guy, J., & Cornish, K. (2010). The development of luminance- and texture-defined form perception during the school-aged years. *Neuropsychologia*, *48*, 3080–3085.

Cant, J. S., Arnott, S. R., & Goodale, M. A. (2009). fMR-adaptation reveals separate processing regions for the perception of form and texture in the human ventral stream. *Experimental Brain Research*, *192*, 391–405.

Caelli, T., & Julesz, B. (1978). On perceptual analyzers underlying visual texture discrimination: Part I. *Biological Cybernetics*, *28*. 167–175.

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

Carandini, M., & Heeger, D. J. (2011). Normalization as a canonical neural computation. *Nature Reviews Neuroscience*, *13*, 51–62.

Carrasco, M., & Yeshurun, Y. (2009). Covert attention effects on spatial resolution. In N. Srinivasan (Ed.), *Progress in Brain Research: Attention* (Vol. 176, pp. 65–86). Amsterdam: Elsevier.

Casco, C., Grieco, A., Campana, G., Corvino, M. P., & Caputo, G. (2005). Attention modulates psychophysical and electrophysiological response to visual texture segmentation in humans. *Vision Research*, *45*, 2384–2396.

Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., & Milner, A. D. (2010a). Separate processing of texture and form in the ventral stream: evidence from FMRI and visual agnosia. *Cerebral Cortex*, *20*, 433–446.

Cavina-Pratesi, C., Kentridge, R. W., Heywood, C. A., & Milner, A. D. (2010b). Separate channels for processing form, texture, and color: evidence from FMRI adaptation and visual object agnosia. *Cerebral Cortex*, *20*, 2319–2332.

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.

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 (Eds.), *Computational Models of Visual Processing* (pp. 291–301). Cambridge, MA: MIT Press.

Chubb, C., Landy, M. S., & Econopouly, J. (2004). A visual mechanism tuned to black. *Vision Research*, *44*, 3223–3232.

Dakin, S. C. (2001). Information limit on the spatial integration of local orientation signals *Journal of the Optical Society of America A*, *18*, 1016–1026.

El-Shamayleh, Y., & Movshon, J. A. (2011). Neuronal responses to texture-defined form in macaque visual area V2. *Journal of Neuroscience*, *31*, 8543–8555.

Ellemberg, D., Allen, H. A., & Hess, R. F. (2004). Investigating local network interactions underlying first- and second-order processing. *Vision Research*, *44*, 1787–1797.

Ellemberg, D., Allen, H. A., & Hess, R. F. (2006). Second-order spatial frequency and orientation channels in human vision. *Vision Research*, *46*, 2798–2803.

Emrith, K., Chantler, M. J., Green, P. R., Maloney, L. T., & Clarke, A. D. F. (2010). Measuring perceived differences in surface texture due to changes in higher order statistics. *Journal of the Optical Society of America A*, *27*, 1232–1244.

Filangieri, C., & Li, A. (2009). Three-dimensional shape from second-order orientation flows. *Vision Research*, *49*, 1465–1471.

Freeman, J., & Simoncelli, E. P. (2011). Metamers of the ventral stream. *Nature Neuroscience*, *14*, 1195–1201.

Georgieva, S. S., Todd, J. T., Peeters, R., & Orban, G. A. (2008). The extraction of 3D shape from texture and shading in the human brain. *Cerebral cortex*, *18*, 2416–2438.

Giora, E., & Casco, C. (2007). Region- and edge-based configurational effects in texture segmentation. *Vision Research*, *47*, 879–886.

Girshick, A. R., Landy, M. S., & Simoncelli, E. P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nature Neuroscience*, *14*, 926–932.

Graham, N. V. S. (1989). *Visual Pattern Analyzers*. New York: Oxford University Press.

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

Graham, N., & Nachmias, J. (1971). Detection of grating patterns containing two spatial frequencies: a comparison of single-channel and multiple-channels models. *Vision Research*, *11*, 251–259.

Graham, N. & 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., & Wolfson, S. S. (2004). Is there opponent-orientation coding in the second-order channels of pattern vision. *Vision Research*, *44*, 3145–3175.

Gurnsey, R., Pearson, P., & Day, D. (1996). Texture segmentation along the horizontal meridian: nonmonotonic changes in performance with eccentricity. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 738–757.

Hallum, L. E., Landy, M. S., & Heeger, D. J. (2011). Human primary visual cortex (V1) is selective for second-order spatial frequency. *Journal of Neurophysiology*, *105*, 2121–2131.

Hansen, B. C., & Hess, R. F. (2006). The role of spatial phase in texture segmentation and contour integration. *Journal of Vision*, *6*, 594–615.

Harrison, S. J., & Keeble, D. R. T. (2008). Within-texture collinearity improves human texture segmentation. *Vision Research*, *48*, 1955–1964.

Hawley, S. J., & Keeble, D. R. T. (2006). Tilt aftereffect for texture edges is larger than in matched subjective edges, but both are strong adaptors of luminance edges. *Journal of Vision*, *6*, 37–52.

Heeger, D. J. (1992). Normalization of cell responses in cat striate cortex. *Visual Neuroscience*, *9*, 181–197.

Heeger, D. J., & Bergen, J. R. (1995). Pyramid-based texture analysis/synthesis. In *Proceedings of the 22nd Annual Conference on Computer Graphics & Interactive Techniques*, *30*, 229–238.

Heinrich, S. P., Andrés, M., & Bach, M. (2007). Attention and visual texture segregation. *Journal of Vision*, *7*(6):6, 1–10.

Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2006). How direction of illumination affects visually perceived surface roughness. *Journal of Vision*, *6*, 634–648.

Ho, Y.-X., Landy, M. S., & Maloney, L. T. (2008). Conjoint measurement of gloss and surface texture. *Psychological Science*, *19*, 196–204.

Ho, Y.-X., Maloney, L. T., & Landy, M. S. (2007). The effect of viewpoint on perceived visual roughness *Journal of Vision*, *7*(1):1, 1–16.

Jingling, L., & Zhaoping, L. (2008). Change detection is easier at texture border bars when they are parallel to the border: Evidence for V1 mechanisms of bottom-up salience. *Perception*, *37*, 197–206.

Julesz, B., Gilbert, E. N., Shepp, L. A., & 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., & Victor, J. D. (1978). Visual discrimination of textures with identical third-order statistics. *Biological Cybernetics*, *31*, 137–140.

Kehrer, L. (1989). Central performance drop on perceptual segregation tasks. *Spatial Vision*, *4*, 45–62.

Kingdom, F. A. A., Prins, N., & Hayes, A. (2003). Mechanism independence for texture-modulation detection is consistent with a filter-rectify-filter mechanism. *Visual Neuroscience*, *20*, 65–76.

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

Knoblauch, K., & Maloney, L. T. (2008). MLDS: Maximum likelihood difference scaling in R. *Journal of Statistical Software, 25*, 1–26.

Lachapelle, J., McKerral, M., Jauffret, C., & Bach, M. (2008). Temporal resolution of orientation-defined texture segregation: a VEP study. *Documenta Ophthalmologica. Advances in Ophthalmology*, *117*, 155–162.

Landy, M. S. (1996). Texture Perception. In G. Adelmen (Ed.), *Encyclopedia of Neuroscience*. Amsterdam: Elsevier.

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

Landy, M. S., & Graham, N. V. G. (2004). Visual perception of texture. In L. M. Chalupa & J. S. Werner (Eds.), *The Visual Neurosciences* (pp. 1106–1118). Cambridge, Mass.: MIT Press.

Landy, M. S., & 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., & Oruç, İ. (2002). Properties of second-order spatial frequency channels. *Vision Research*, *42*, 2311–2329.

Landy, M. S., Ho, Y.-X., Serwe, S., Trommershäuser, J., & Maloney, L. T. (2011). Cues and pseudocues in texture and shape perception. In J. Trommershäuser, K. Körding, & M. S. Landy (Eds.), *Sensory Cue Integration* (pp. 263–278). New York: Oxford University Press.

Larsson, J., Landy, M. S., & Heeger, D. J. (2006). Orientation-selective adaptation to first- and second-order patterns in human visual cortex. *Journal of Neurophysiology*, *95*, 862–881.

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

Li, A., & Zaidi, Q. (2001a). Information limitations in perception of shape from texture. *Vision Research*, *41*, 1519–1533.

Li, A., & Zaidi, Q. (2001b). Veridicality of three-dimensional shape perception predicted from amplitude spectra of natural textures. *Journal of the Optical Society of America A*, *18*, 2430–2447.

Li, A., & Zaidi, Q. (2003). Observer strategies in perception of 3-D shape from isotropic textures: developable surfaces. *Vision Research*, *43*, 2741–2758.

Li, A., & Zaidi, Q. (2004). Three-dimensional shape from non-homogeneous textures: carved and stretched surfaces *Journal of Vision*, *4*, 860–878.

Li, A., & Zaidi, Q. (2009). Release from cross-orientation suppression facilitates 3D shape perception *PloS One*, *4*, e8333.

McDermott, J. H., & Simoncelli, E. P. (2011). Sound texture perception via statistics of the auditory periphery: evidence from sound synthesis. *Neuron*, *71*, 926–940.

Montaser-Kouhsari, L., Landy, M. S., Heeger, D. J., & Larsson, J. (2007). Orientation-selective adaptation to illusory contours in human visual cortex. *Journal of Neuroscience*, *27*, 2186–2195.

Motoyoshi, I., & Kingdom, F. A. A. (2007). Differential roles of contrast polarity reveal two streams of second-order visual processing. *Vision Research*, *47*, 2047–2054.

Motoyoshi, I., & Kingdom, F. A. A. (2010). The role of co-circularity of local elements in texture perception. *Journal of Vision*, *10*(1):3, 1–8.

Motoyoshi, I., & Nishida, S. (2004). Cross-orientation summation in texture segregation. *Vision Research*, *44*, 2567–2576.

Norcia, A. M., Pei, F., Bonneh, Y., Hou, C., Sampath, V., & Pettet, M. W. (2005). Development of sensitivity to texture and contour information in the human infant. *Journal of Cognitive Neuroscience*, *17*, 569–579.

Nothdurft, H. C. (1985). Sensitivity for structure gradient in texture discrimination tasks. *Vision Research*, *25*, 551–560.

Oruç, İ., Landy, M. S., & Pelli, D. G. (2006). Noise masking reveals channels for second-order letters. *Vision Research*, *46*, 1493–1506.

Padilla, S., Drbohlav, O., Green, P. R., Spence, A., & Chantler, M. J. (2008). Perceived roughness of $1/f^{\beta}$ noise surfaces. *Vision Research*, *48*, 1791–1797.

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

Prins, N. (2008). Texture modulation detection by probability summation among orientation-selective and isotropic mechanisms. *Vision Research*, *48*, 2751–2766.

Prins, N., & Kingdom, F. A. A. (2003). Detection and discrimination of texture modulations defined by orientation, spatial frequency, and contrast. *Journal of the Optical Society of America A*, *20*, 401–410.

Prins, N., & Kingdom, F. A. A. (2006). Direct evidence for the existence of energy-based texture mechanisms. *Perception*, *35*, 1035–1046.

Prins, N., Nottingham, N. K., & Mussap, A. J. (2003). The role of local grouping and global orientation contrast in perception of orientation-modulated textures. *Vision Research*, *43*, 2315–2331.

Schade, U., & Meinecke, C. (2009). Spatial distance between target and irrelevant patch modulates detection in a texture segmentation task. *Spatial Vision*, *22*, 511–527.

Schade, U., & Meinecke, C. (2011). Texture segmentation: do the processing units on the saliency map increase with eccentricity. *Vision Research*, *51*, 1–12.

Schofield, A. J. (2000). What does second-order vision see in an image? *Perception*, *29*, 1071–1086.

Song, Y., & Baker, C. L. (2006). Neural mechanisms mediating responses to abutting gratings: luminance edges vs. illusory contours. *Visual Neuroscience*, *23*, 181–199.

Song, Y., & Baker, C. L. (2007). Neuronal response to texture- and contrast-defined boundaries in early visual cortex. *Visual Neuroscience*, *24*, 65–77.

Stylianou-Korsnes, M., Reiner, M., Magnussen, S. J., & Feldman, M. W. (2010). Visual recognition of shapes and textures: an fMRi study. *Brain Structure & Function*, *214*, 355–359.

Tanaka, H., & Ohzawa, I. (2006). Neural basis for stereopsis from second-order contrast cues. *Journal of Neuroscience*, *26*, 4370–4382.

Tanaka, H., & Ohzawa, I. (2009). Surround suppression of V1 neurons mediates orientation-based representation of high-order visual features. *Journal of Neurophysiology*, *101*, 1444–1462.

Thaler, L., Todd, J. T., & Dijkstra, T. M. H. (2007). The effects of phase on the perception of 3D shape from texture: psychophysics and modeling. *Vision Research*, *47*, 411–427.

Thielscher, A., & Neumann, H. (2003). Neural mechanisms of cortico-cortical interaction in texture boundary detection: a modeling approach. *Neuroscience*, *122*, 921–939.

Thielscher, A., & Neumann, H. (2005). Neural mechanisms of human texture processing: texture boundary detection and visual search. *Spatial Vision*, *18*, 227–257.

Thielscher, A., & Neumann, H. (2007). A computational model to link psychophysics and cortical cell activation patterns in human texture processing. *Journal of Computational Neuroscience*, *22*, 255–282.

Thielscher, A., Kölle, M., Neumann, H., Spitzer, M., & Grön, G. (2008). Texture segmentation in human perception: a combined modeling and fMRI study. *Neuroscience*, *151*, 730–736.

Tkačik, G., Prentice, J. S., Victor, J. D., & Balasubramanian, V. (2010). Local statistics in natural scenes predict the saliency of synthetic textures. *Proceedings of the National Academy of Sciences USA*, *107*, 18149–18154.

Todd, J. T., & Oomes, A. H. J. (2002). Generic and non-generic conditions for the perception of surface shape from texture. *Vision Research*, *42*, 837–850.

Todd, J. T., & Thaler, L. (2010). The perception of 3D shape from texture based on directional width gradients. *Journal of Vision*, *10*(5):17, 1–13.

Todd, J. T., Oomes, A. H. J., Koenderink, J. J., & Kappers, A. M. L. (2004). The perception of doubly curved surfaces from anisotropic textures. *Psychological Science*, *15*, 40–46.

Todd, J. T., Thaler, L., Dijkstra, T. M. H., Koenderink, J. J., & Kappers, A. M. L. (2007). The effects of viewing angle, camera angle, and sign of surface curvature on the perception of three-dimensional shape from texture. *Journal of Vision*, *7*(12):9, 1–16.

Victor, J. D., & Conte, M. M. (2004). Visual working memory for image statistics. *Vision Research*, *44*, 541–556.

Wang, H. X., Landy, M. S., & Heeger, D. J. (2011). Psychophysical evidence for normalization in second-order mechanisms. *Journal of Vision*, *11(11):1173*.

Watson, A. B., & Robson, J. G. (1981). Discrimination at threshold: labelled detectors in human vision. *Vision Research*, *21*, 1115–1122.

Wolfson, S. S., & Landy, M. S. (1995). Discrimination of orientation-defined texture edges. *Vision Research*, *35*, 2863–2877.

Yeshurun, Y., & Carrasco, M. (2008). The effects of transient attention on spatial resolution and the size of the attentional cue. *Perception & Psychophysics*, *70*, 104–113.

Yeshurun, Y., & Carrasco, M. (2000). The locus of attentional effects in texture segmentation. *Nature Neuroscience*, *3*, 622–627.

Yeshurun, Y., Montagna, B., & Carrasco, M. (2008). On the flexibility of sustained attention and its effects on a texture segmentation task. *Vision Research*, *48*, 80–95.

Zaidi, Q., & Li, A. (2002). Limitations on shape information provided by texture cues. *Vision Research*, *42*, 815–835.

Zhan, C. A., & Baker, C. L. (2006). Boundary cue invariance in cortical orientation maps. *Cerebral Cortex*, *16*, 896–906.

Zhan, C. A., & Baker, C. L. (2008). Critical spatial frequencies for illusory contour processing in early visual cortex. *Cerebral Cortex*, *18*, 1029–1041.

Zhaoping, L. (2003). V1 mechanisms and some figure-ground and border effects. *Journal of Physiology, Paris*, *97*, 503–515.

Zhaoping, L., Guyader, N., & Lewis, A. (2009). Relative contributions of 2D and 3D cues in a texture segmentation task, implications for the roles of striate and extrastriate cortex in attentional selection. *Journal of Vision*, *9(11):20*, 1–22.

Zhaoping, L., & May, K. A. (2007). Psychophysical tests of the hypothesis of a bottom-up saliency map in primary visual cortex. *PLoS Computational Biology*, 3(4):e62.

**Figure Legends**

Figure x.1. The boardwalk at Coney Island, Brooklyn, NY. Note the salient boundaries between boardwalk sections each consisting of parallel boards, with the sections' boards in different orientations.

Figure x.2. The back-pocket model of texture segregation, consisting of a linear filter, a point-wise nonlinearity, and a second-order linear filter at a coarser scale.

Figure x.3. The back-pocket model of Figure x.2 applied to the image of Figure x.1. (A) Portion of Figure x.1 with an orientation-defined edge. (B) Convolved with a vertically-oriented filter. (C) Full-wave rectified by squaring. (D) Convolved with a second, coarser-scale, vertically-oriented filter. In (B) and (D), a filter response of zero is represented as mid-gray; in (D) zero is represented as black.

Figure x.4. Typical second-order stimulus modulations used in texture experiments in which a noise carrier is modulated. (A) Contrast modulation – CM. (B) Orientation modulation – OM. (C) Frequency modulation - FM. (D) Illusory contour – IC. (E) First-order luminance modulation (LM). In A–C, a modulator, here a vertical sine wave, controls the contrast of one (A) or two (B–C) carrier patterns, here consisting of filtered noise.

Figure x.5. Example of a texture stimulus that has constant orientation gradient throughout the pattern and yet has salient contours, inconsistent with typical FRF models (modeled after Ben-Yosef & Ben-Shahar, 2008).

Figure x.6. Texture representation based on image statistics. In each panel, the inset square is an original texture and the rest of the image is new texture using the technique of Portilla & Simoncelli (2000).

Figure x.7. (A) Sample stimulus for visual crowding study. (B) "Mongrel" image based on the Portilla & Simoncelli (2000) model as described by Balas et al. (2009).

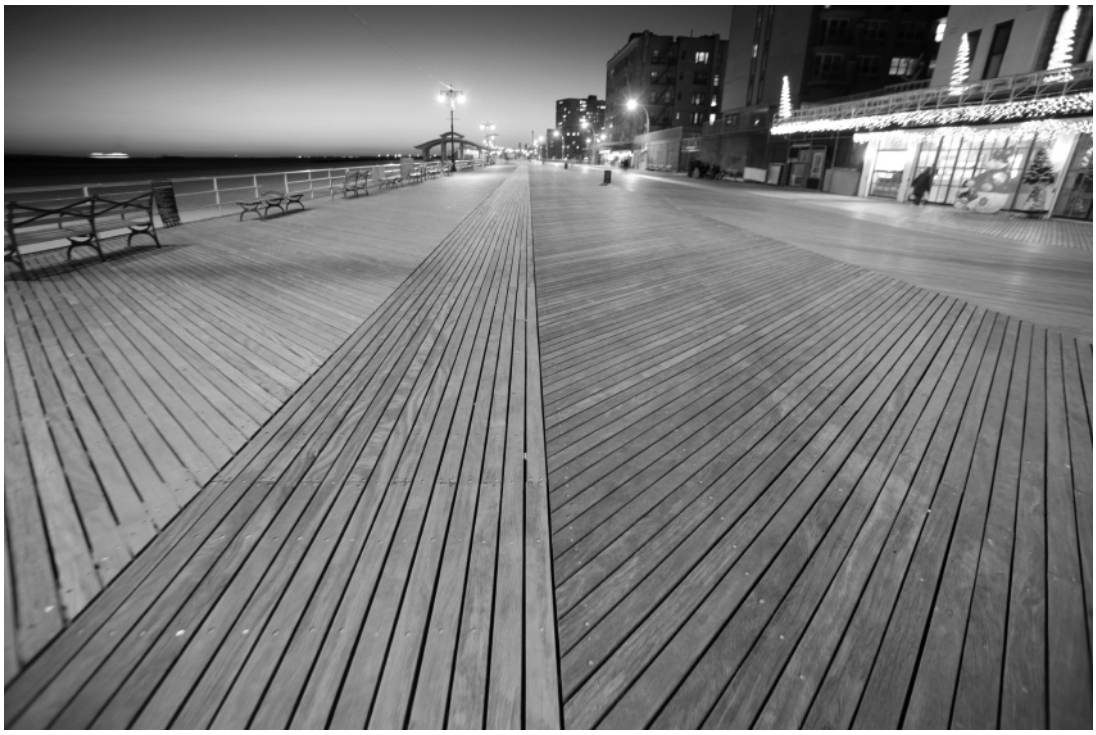


Figure x.1. The boardwalk at Coney Island, Brooklyn, NY. Note the salient boundaries between boardwalk sections each consisting of parallel boards, with the sections' boards in different orientations.

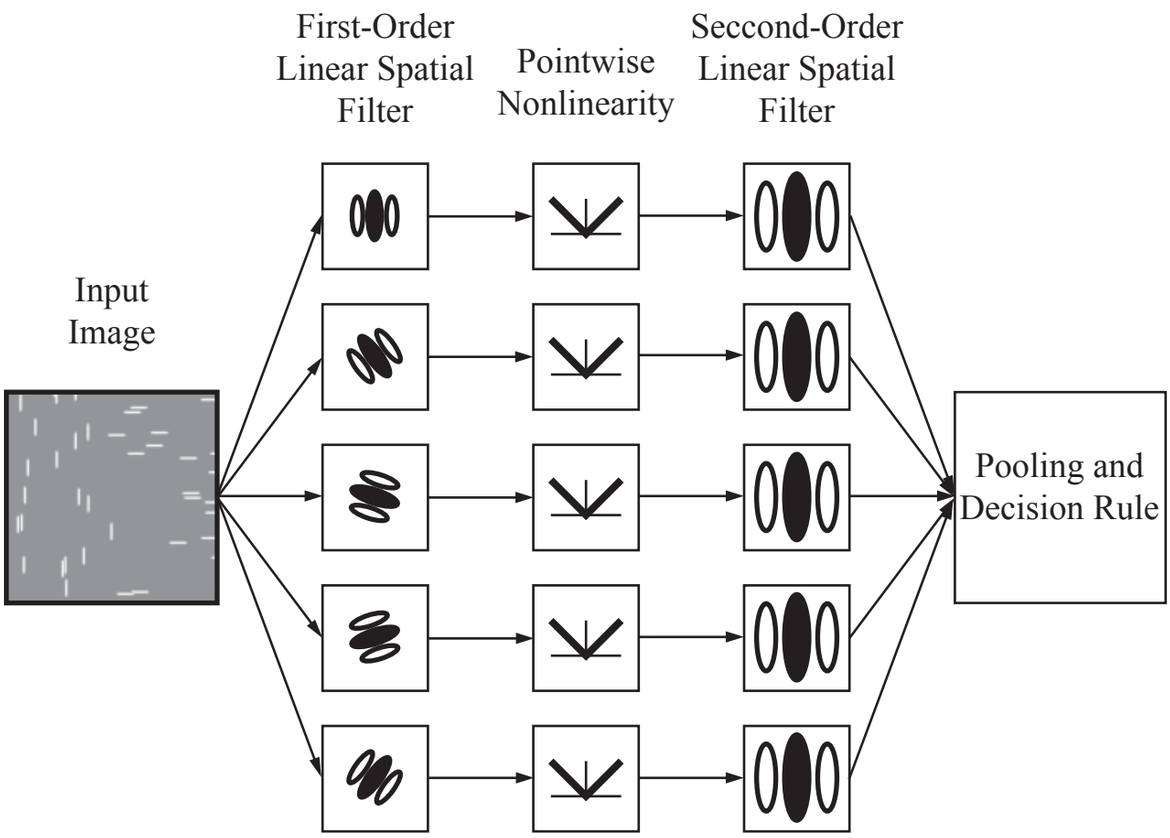



Figure x.2. The back-pocket model of texture segregation, consisting of a linear filter, a point-wise nonlinearity, and a second-order linear filter at a coarser scale.

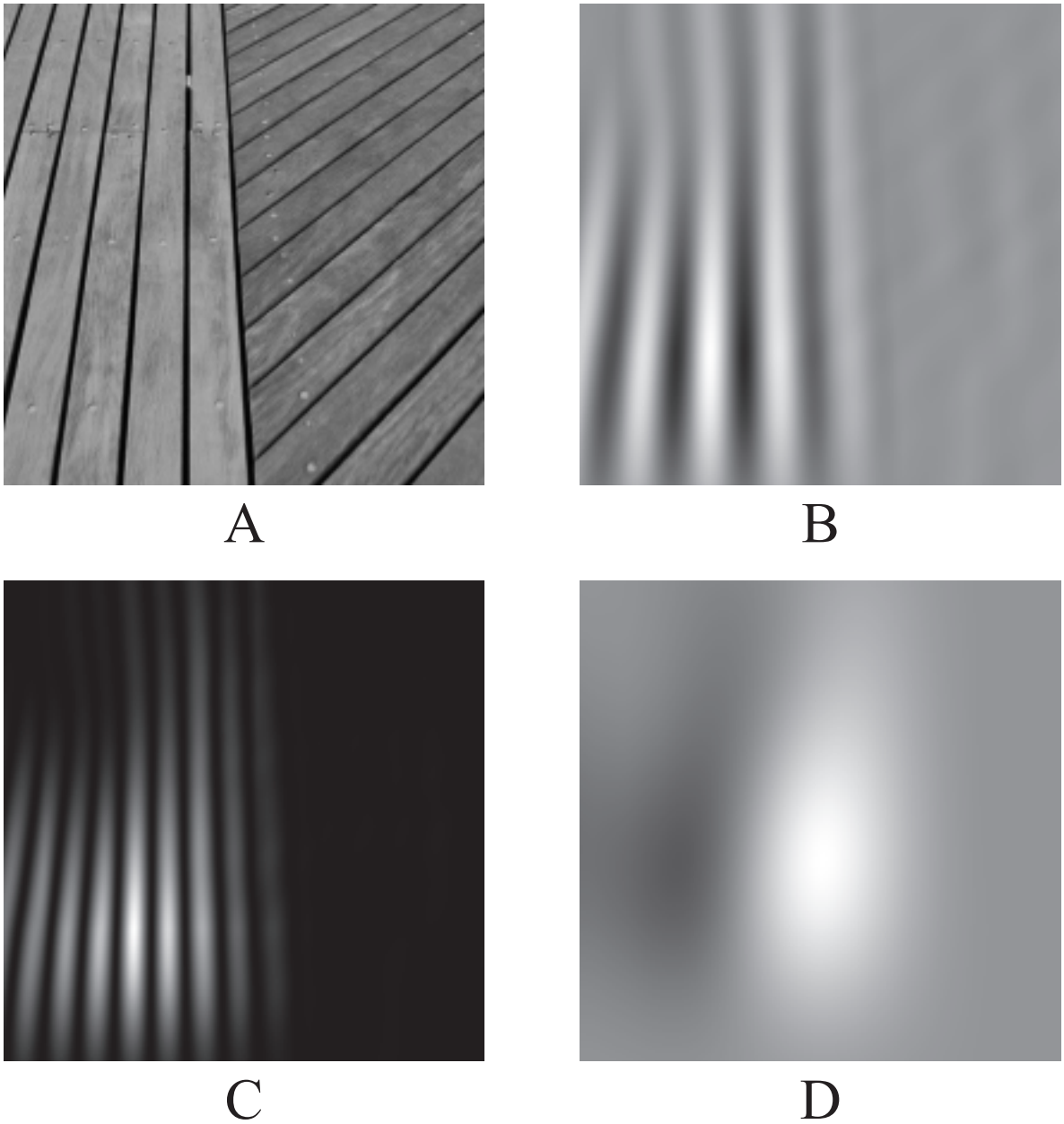



Figure x.3. The back-pocket model of Figure x.2 applied to the image of Figure x.1. (A) Portion of Figure x.1 with an orientation-defined edge. (B) Convolved with a vertically-oriented filter. (C) Full-wave rectified by squaring. (D) Convolved with a second, coarser-scale, vertically-oriented filter. In (B) and (D), a filter response of zero is represented as mid-gray; in (D) zero is represented as black.

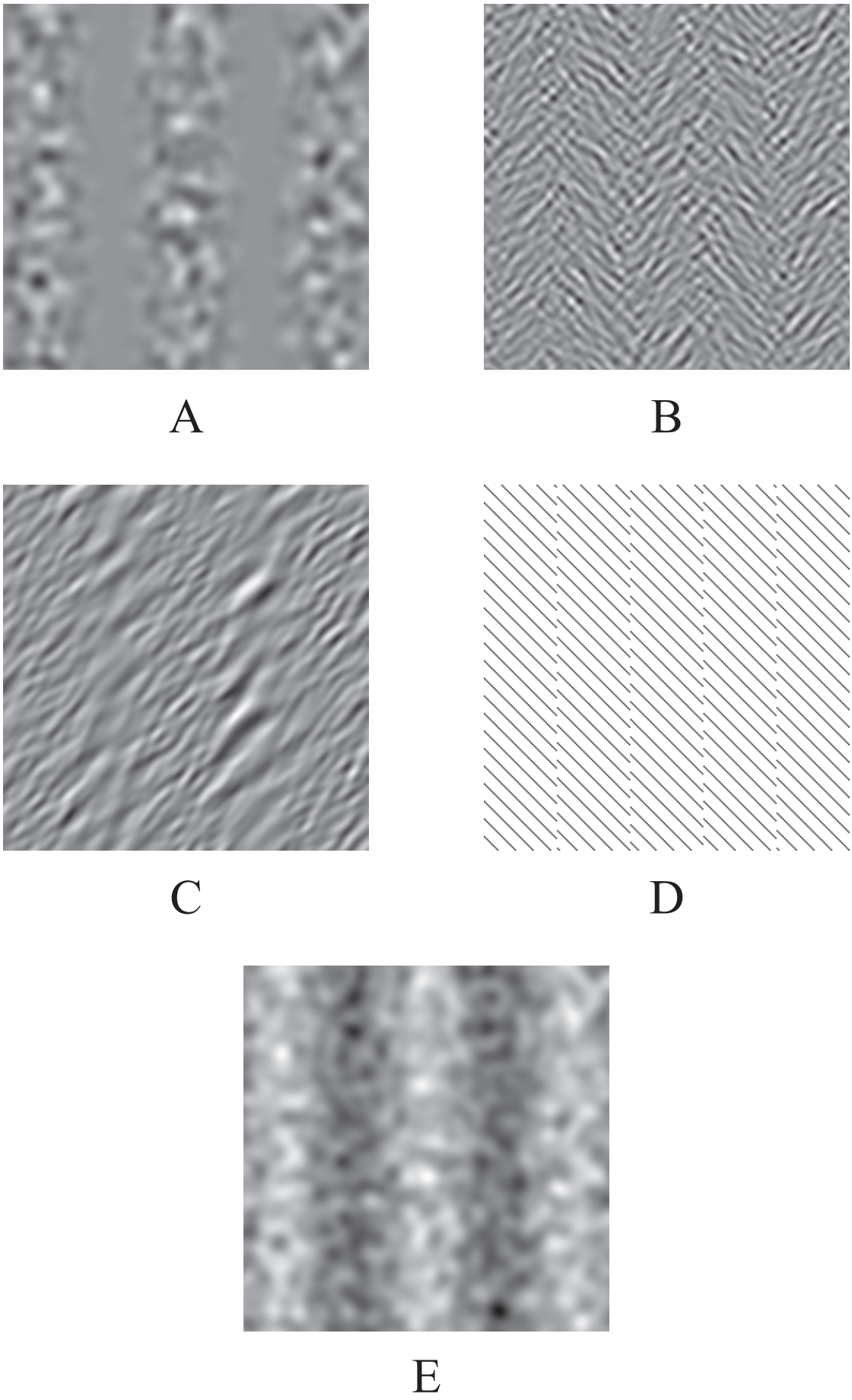



Figure x.4. Typical second-order stimulus modulations used in texture experiments in which a noise carrier is modulated. (A) Contrast modulation – CM. (B) Orientation modulation – OM. (C) Frequency modulation - FM. (D) Illusory contour – IC. (E) First-order luminance modulation (LM). In A–C, a modulator, here a vertical sine wave, controls the contrast of one (A) or two (B–C) carrier patterns, here consisting of filtered noise.

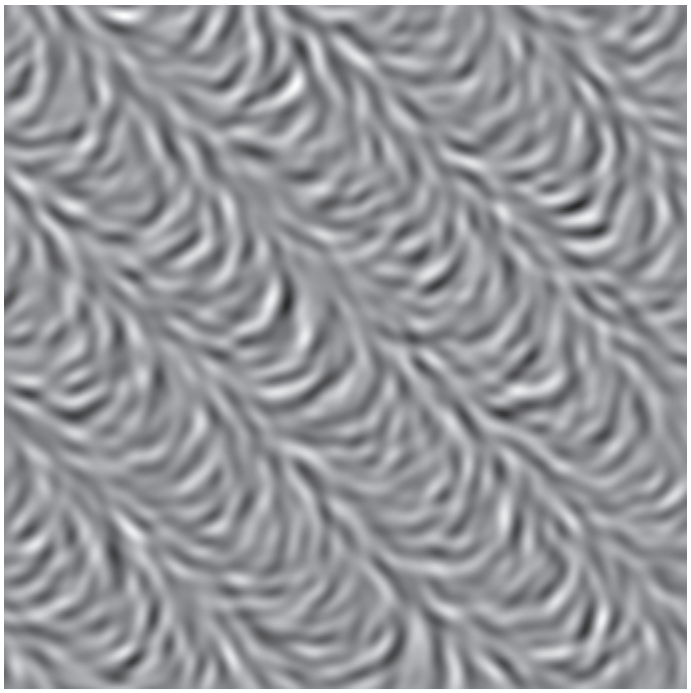
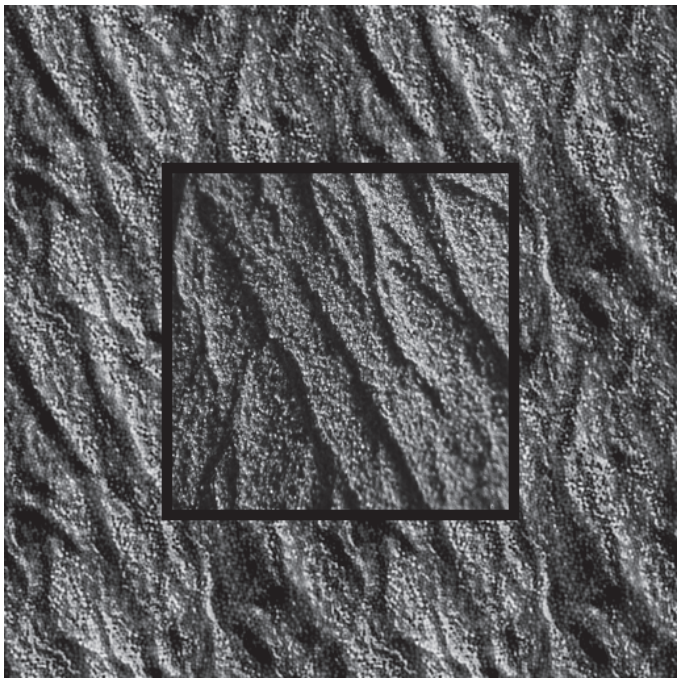
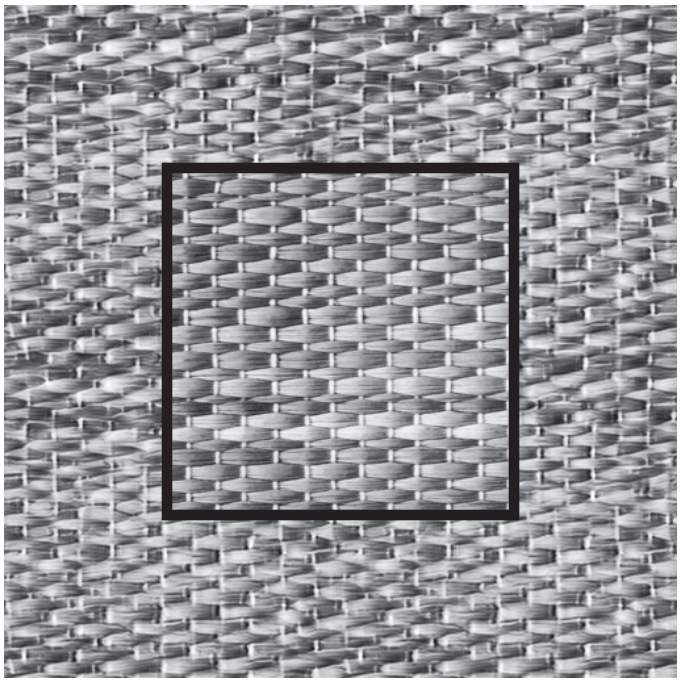


Figure x.5. Example of a texture stimulus that has constant orientation gradient throughout the pattern and yet has salient contours, inconsistent with typical FRF models (modeled after Ben-Yosef & Ben-Shahar, 2008).



A



B

Figure x.6. Texture representation based on image statistics. In each panel, the inset square is an original texture and the rest of the image is new texture using the technique of Portilla & Simoncelli (2000).

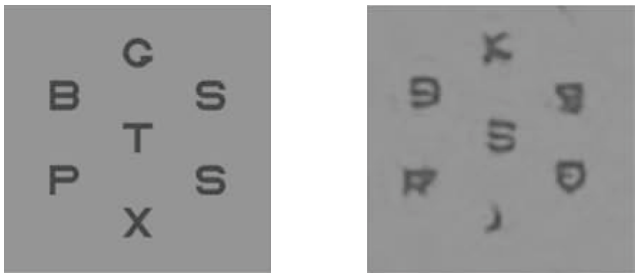



Figure x.7. (A) Sample stimulus for visual crowding study. (B) “Mongrel” image based on the Portilla & Simoncelli (2000) model as described by Balas et al. (2009).