Texture discrimination based on global feature alignments

Flip Phillips   Psychology and Neuroscience, Skidmore College, USA

James T. Todd   Department of Psychology, The Ohio State University, USA

Three experiments are reported that examined the abilities of human observers to discriminate textures with identical distributions of orientation and spatial frequency. In Experiment 1, the stimuli consisted of low-pass filtered noise that was uniformly distributed and spatially isotropic. Observers were able to discriminate textures with identical image statistics when their frequencies were 1 cpd or less, but performance dropped off sharply for textures with higher frequencies. In Experiment 2, a novel procedure was employed with which it is possible to increase the high-frequency energy in the amplitude spectrum of a texture, while preserving the macroscopic alignments of its phase spectrum. The results reveal that this has little effect on performance, thus indicating that it is not spatial frequency per se that limits the abilities of observers to discriminate macroscopic texture patterns. When the phase spectra of these textures were randomly scrambled in Experiment 3, the frequency thresholds for discriminating textures reverted back to 1 cpd as was obtained in Experiment 1. These results provide strong evidence that human observers make use of two distinct procedures for discriminating patches of texture: One based on image statistics that is possible for all textures; and another based on macroscopic phase alignments that define features that are larger than 1°.

Keywords: texture discrimination, global feature alignment, spatial frequency

Introduction

Statistical approaches to visual perception have become increasingly popular over the past two decades, but they have proven to be particularly useful for the problem of texture discrimination. The basic idea of this approach is quite simple. The spatial structure at each image location is encoded by a bank of filters that are tuned to a variety of orientations and spatial scales—though filters sensitive to other higher order features (e.g., textons) may also be included. In order to compare two patches of texture, histograms are computed within each patch to represent the relative activations of each filter type. Based on this statistical analysis, two texture patches should be perceptually distinct whenever the difference between their respective feature histograms exceeds some critical threshold (e.g., Bergen & Adelson, 1988; Chubb & Landy, 1991).

Within this context, it is useful to keep in mind that surfaces in the natural environment have structures at a continuum of scales, which can be partitioned into three perceptually distinct regimes (see Koenderink & van Doorn, 1998, 1996). One range of scales that is outside the limits of visual sensitivity constitutes a surface's microstructure. Although these microscopic variations cannot be seen by the naked eye, they can be perceived indirectly by the pattern of surface reflectance. For example, this is what distinguishes matte versus shiny surfaces. A second range of scales that is referred to as mesostructure involves surface variations that are large enough to be visible but too small for the texture elements to be perceptually individuated. Examples in this regime include patterns of stucco or the tufting of carpets. Finally, a third range of scales called macrostructure include variations on a surface that are sufficiently large to have perceptually distinct shapes, such as the patterns of reflectance changes on a Holstein cow.

Let us now consider how this partitioning of scales is relevant to models of texture discrimination. The lower panels of Figure 1 show two images that were generated with a 500 × 500 grid of pixels that were randomly colored black or white. From a brief examination of these images, it is clear that they are perceptually indistinguishable. Note that this is precisely what would be expected from a statistically based model of texture discrimination. Even though the correlation between corresponding pixels in these images is close to zero, the structure of their respective feature histograms is nearly identical. The upper panels of Figure 1 show a similar pair of images that were generated with a much coarser 4 × 4 grid of pixels that were randomly colored black or white in equal proportions. Models of texture discrimination based on first-order feature histograms would again predict that the two patterns should be perceptually identical. However, it is immediately obvious when examining these images that they are perceptually quite different.

These demonstrations suggest that statistical models of texture discrimination may provide a good account of the
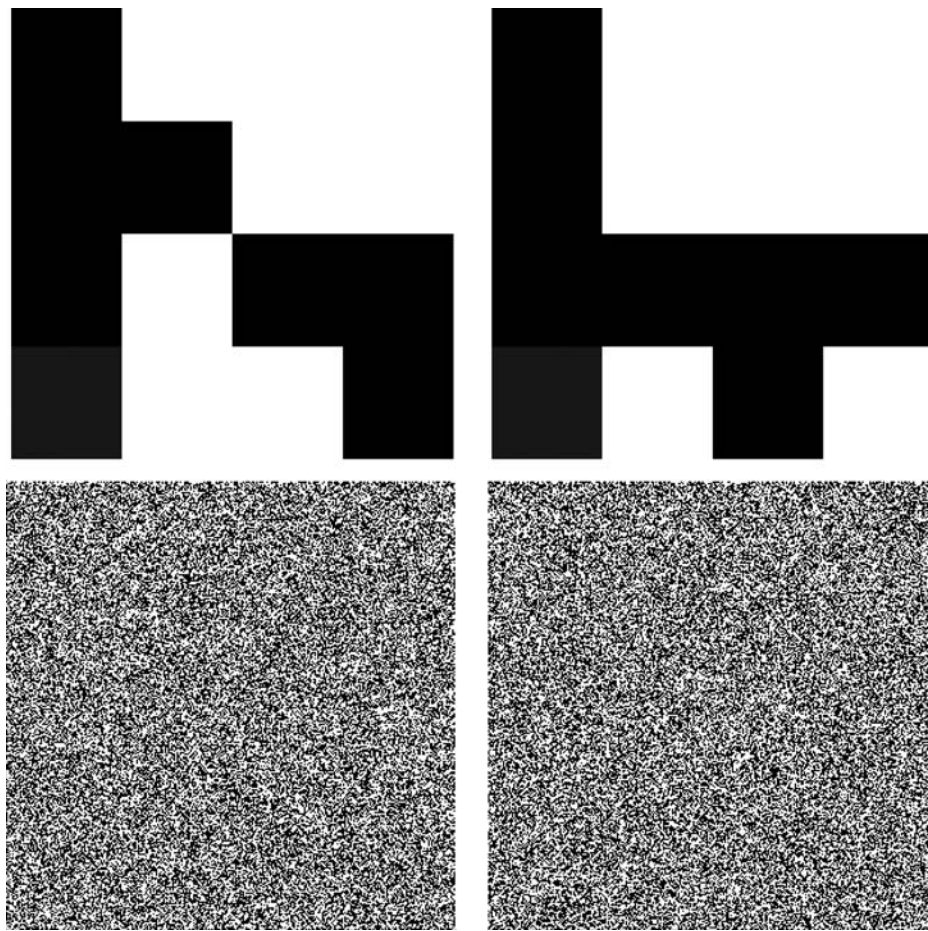
Figure 1. Two extreme scales of texture. The upper panels, consisting of 4 × 4 squares colored randomly black or white, are easily discriminable while the lower panels at 500 × 500 are far more difficult to discriminate.

ability of human observers to distinguish mesostructures, but that they are not so good at explaining observers' perceptions of macrostructures. It is interesting to note that within the Fourier domain, the distribution of orientation and spatial frequency used in most models of texture discrimination are represented entirely within the amplitude spectrum. However, it has long been known that most of the visually relevant information for identifying macroscopic shapes is contained within the phase spectrum. The classic procedure for demonstrating the importance of global phase is to combine the amplitude spectrum of one image with the phase spectrum of another (Huang, Burnett, & Deczky, 1975; Morgan, Ross, & Hayes, 1991; Oppenheim & Lim, 1981; Piotrowski & Campbell, 1982; Ramachandran & Srinivasan, 1961). When observers examine the resulting combination, it invariably appears more similar to the image that contributed the phase spectrum than the one that contributed the amplitude spectrum (e.g., see Figure 2). What this suggests is that macroscopic shapes are not distinguished by the relative activations of different filter types but rather by the pattern of alignments of the filter activations across different spatial locations. Thus, any model that can successfully explain the perceptual distinction between macroscopic textures must somehow take into account the relative alignments of local features in different regions.

It is important to recognize within this context that the pattern of alignments described above does not refer to Fourier phase congruence within local regions of an image. This has been proposed as a possible source of information for the detection of local edges (e.g., Kovesi, 1999; Morrone & Burr, 1988), though more recent psychophysical evidence suggests that this is not the mechanism by which the human visual system localizes edges (see Georgeson, May, Freeman, & Hesse, 2007; Hesse & Georgeson, 2005). In order to identify globally coherent shapes, the encoding of phase cannot be a purely local process. Rather, it must somehow represent how the orientations of features in one region of an image are related to the orientations of other features that may be spatially quite distant.

The present article describes three experiments that were designed to explore these issues in greater detail. In all three studies, observers were asked to make same–different judgments for pairs of random noise images that
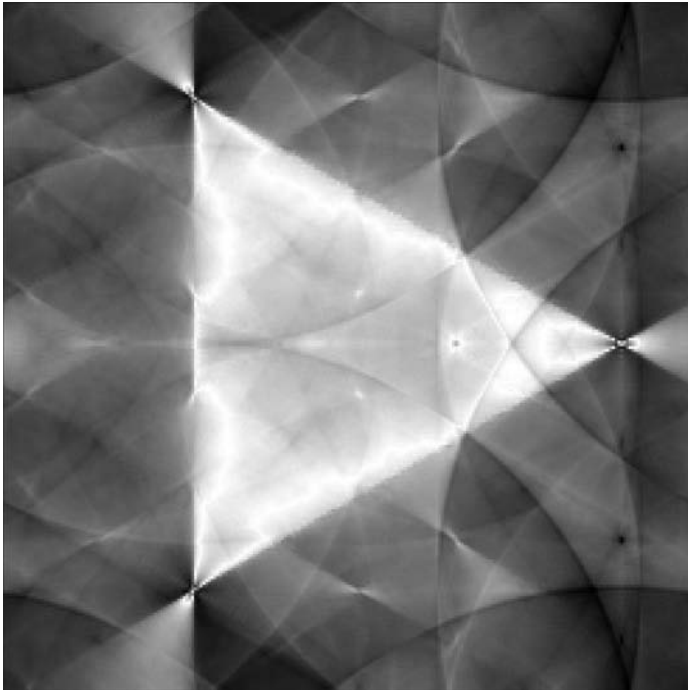
Figure 2. This image was constructed by combining the amplitude spectrum of a black circle on a white background with the phase spectrum of a black triangle on a white background and subjecting the combination to the inverse Fourier transform. It is clear that the resulting image is dominated by the figure whose phase information is present.

were generated from the same underlying distribution. In Experiment 1, we use low-pass filtered random noise in an effort to determine the boundary between mesostructure and macrostructure by measuring the threshold cutoff frequency for which observers can no longer discriminate different noise patterns. In Experiment 2, we attempt to demonstrate that it is not low spatial frequency per se that allows observers to discriminate these patterns, but rather it is the patterns of phase alignments in their respective macroscopic features. This is achieved by increasing the spatial frequency of our low-frequency stimuli from Experiment 1 in a manner that preserves their macroscopic phase alignments. Finally, in Experiment 3 we test whether performance with these latter stimuli would be impaired if their phase spectra were randomly scrambled.

Experiment 1

We first establish baseline performance using unstructured noise patterns of varying spatial frequency. This will establish a limit for the comparison task that we can compare against performance in structured and partially structured conditions.

Methods

Subjects

The observers were eight adults, all with normal or corrected-to-normal vision. All were members of the laboratories conducting the research and aware of the purpose and scope of the experiments.

Apparatus

The apparatus consisted of an Apple Power Mac computer with an ATI 1950 series graphics card (Advanced Micro Devices), driving an Apple 30″ Cinema HD LCD display, calibrated to a luminance of 150 cd/m^2, running at a spatial resolution of 2560×1600 pixels with 8-bit components.[1] Subjects were seated at a viewing distance of 57 cm ($1° = 1$ cm) from the LCD monitor. The experiment control software was written using an in-house, Python-based, experiment package—*eelpy*. A chin rest was used to maintain a constant viewing distance. All observations were made monocularly with dim background lighting. Responses were recorded using an X-Keys programmable USB button box (PI Engineering).

Stimuli

The stimuli consisted of patches of low-pass filtered, spatially isotropic, uniformly distributed noise. The filters ranged from 0.14 to 12.7 cycles/deg in spatial frequency over 14 steps, spaced at intervals scaled by $1/\sqrt{2}$. Each stimulus was histogram equalized after filtering to normalize the mean luminance across different stimuli. A sample set of stimuli is shown in Figure 3. For a detailed description of the stimulus generation process, see Appendix A.

Procedure

On a given trial, the stimulus presentation proceeded as follows: First, a black screen with a randomly positioned red fixation cross was displayed for 1.0 s. After fixation, the stimulus patch was displayed at a second random location, centered within ±5° of the initial fixation. The patch remained on screen for 0.5 s. After which, a second random fixation and stimulus appeared, again for 1.0 s and 0.5 s, respectively. This second stimulus was either the same as in the first interval, or a patch of different "base" noise that was filtered at the same spatial frequency. Thus, it appeared phenomenally *similar* but did not always have the same features in the same configurations. This was followed by a black screen until the subject responded "same" or "different" via the button box.

The fixations and random stimulus placement served to prevent the subject from concentrating their attention on one particular feature or location on the stimulus patch. Furthermore, in order to prevent the use of features near
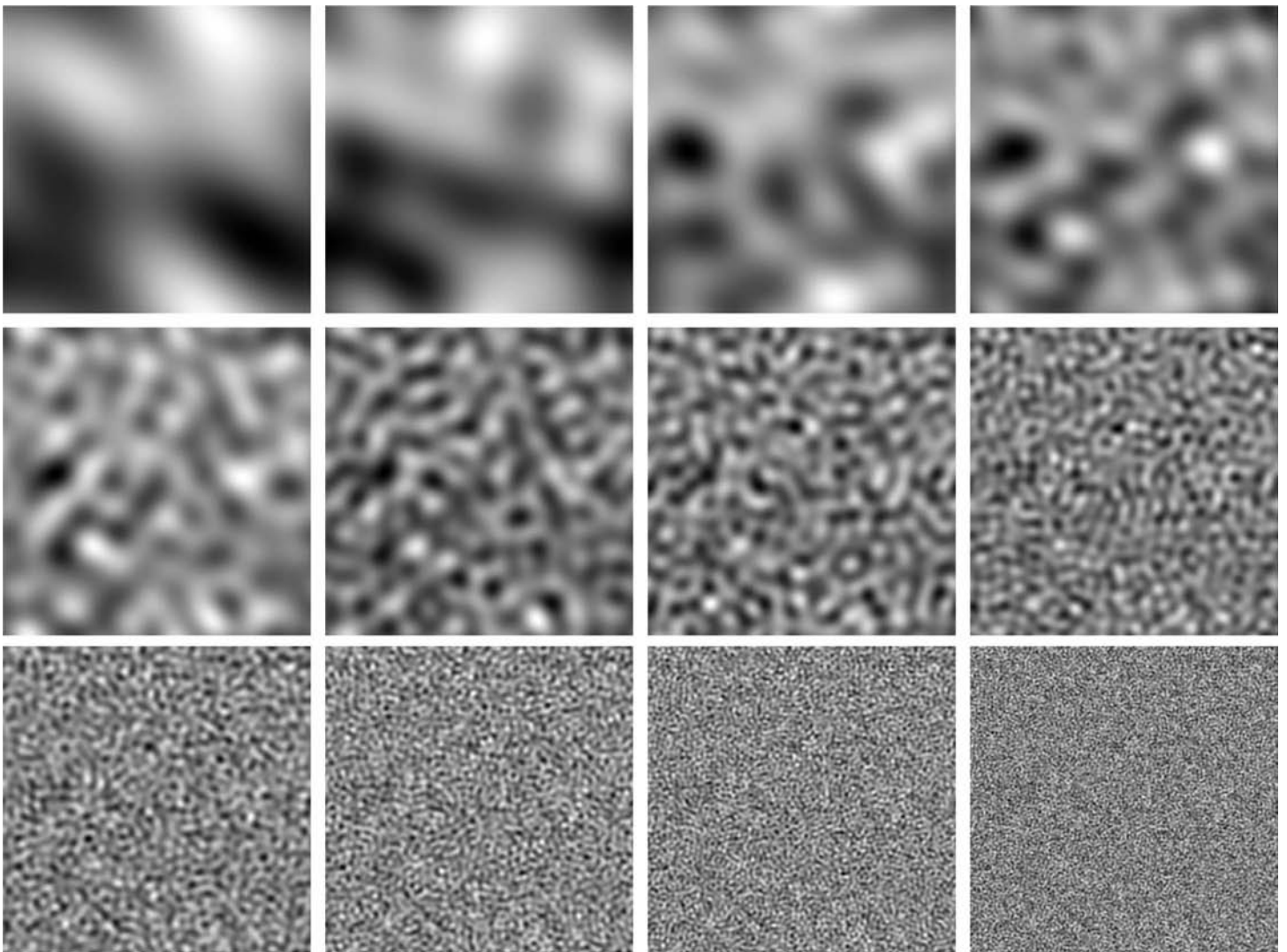
Figure 3. Sample stimulus set from Experiment 1. Each stimulus consists of a patch of low-pass filtered uniformly distributed noise. Filters ranged from 0.14 to 12.7 cycles/deg, spaced at intervals scaled by $1/\sqrt{2}$.

the edges and corners of the patch for matching,[2] the noise patch was randomly shifted within its own 0.25° window. Each subject participated in two blocks of 24 repetitions of each of the 14 stimulus levels, for a total of 672 trials/subject (Figure 4).

Results and discussion

Results for Experiment 1 are shown in Figure 5. Fitting these data to a Log-Weibull distribution determined that the 60%, 75%, and 90% thresholds for these stimuli were located at 1, 2, and 3 cycles/deg, respectively. This establishes a transition region from macro- to mesostructure in the texture images. When the noise varies at approximately 1 cpd or less, coherent structures emerge that the subject can take advantage of for discrimination. For example, note in the upper right panel of Figure 3 that there are two vertically aligned black blobs that perceptually

pop out from the overall pattern, and in the upper left panel, there is a diagonally oriented black blob that is surrounded on the left by diagonally oriented white blobs. All of the observers reported during debriefing that they used this type of macroscopic feature whenever they were available in order to determine whether or not two texture patches were the same. However, at approximately 3 cpd, the blobs in the texture become too small to discern any specific configural alignments among them, and as shown in Figure 5, the discrimination of these higher frequency textures is much more difficult.

Experiment 2

The anecdotal reports from the subjects in Experiment 1 confirmed that they use configural alignments of features
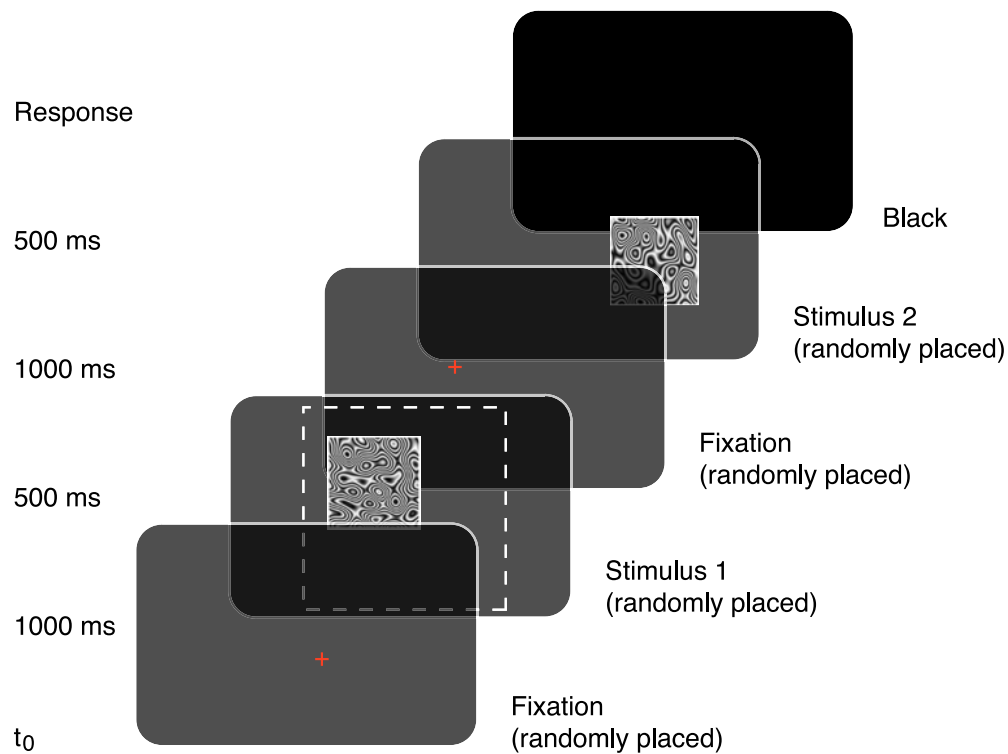
Figure 4. The task was a "same–different" paradigm with the two stimuli offset spatially and temporally. The trial starts with a 1-s presentation of randomly placed fixation cross, followed by a randomly positioned stimulus visible for 500 ms. This sequence is repeated a second time with a second stimulus—same or different from the first. After which the subject responded "same" or "different".

as the primary cue for discriminating textures with equivalent histograms, and the psychophysical results reveal that this strategy breaks down for textures with frequencies above 3 cpd, in which it is difficult to identify individual feature elements. Having established a boundary between the macro- and mesostructures, we will now introduce a new manipulation that increases the high-frequency energy in the amplitude spectrum of a texture, while preserving the macroscopic alignments of its phase spectrum. Our goal is to demonstrate that it is the phase spectrum that defines macroscopic structure of a texture, rather than its amplitude spectrum.

Methods

Subjects

The observers were seven adults, all with normal or corrected-to-normal vision. All were members of the laboratories conducting the research and aware of the purpose and scope of the experiments.

Apparatus

The apparatus was the same as in Experiment 1.

Stimuli

Similar to Experiment 1, our stimuli started with patches of low-pass filtered, spatially isotropic, uniformly distributed noise. Having established the baseline performance thresholds in Experiment 1, we narrow the basic

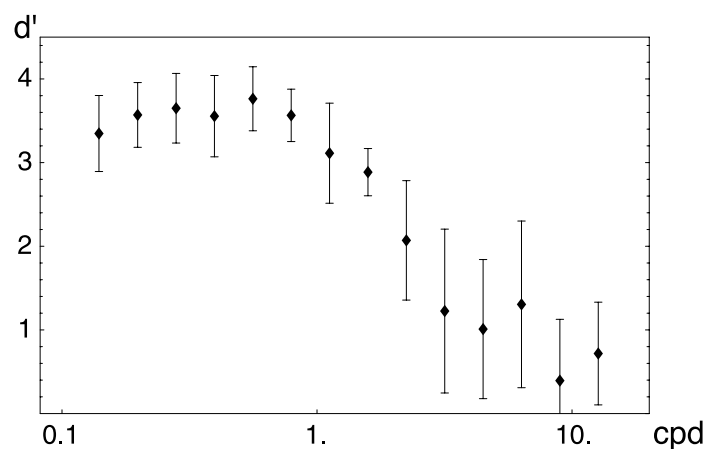Figure 5. Results from Experiment 1. Error bars are $\pm\hat{\sigma}$. Fitting these results to a Log-Weibull distribution shows that the 60%, 75%, and 90% thresholds for these stimuli were located at 1, 2, and 3 cycles/deg, respectively, establishing a transition from coherent macrostructure to the more ambiguous mesostructure.
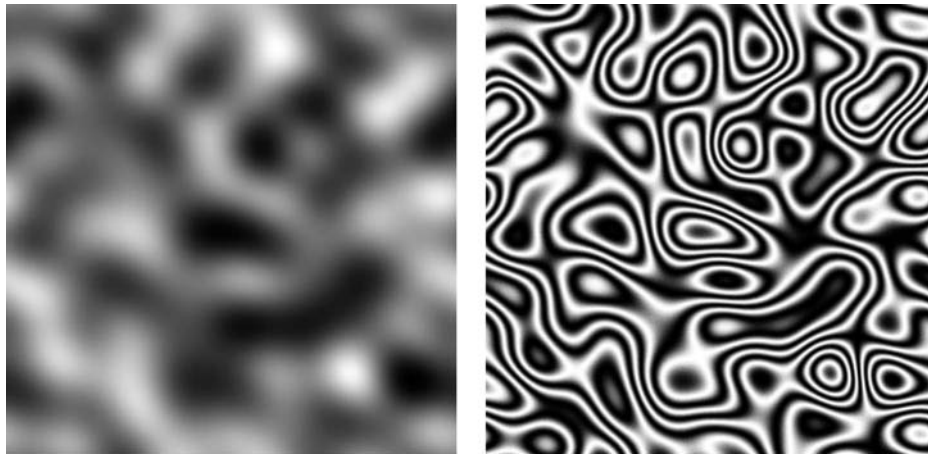
Figure 6. An example of the introduction of high-frequency components to the carrier stimuli. On the left is a low-pass filtered stimulus as used in Experiment 1 and on the right is a stimulus that has had high-frequency sinusoids imposed to provide linear structure.

frequencies to range on the interval 1 to 3 cycles/deg over five steps, spaced at intervals scaled by $1/\sqrt{2}$. We refer to these as the "carrier" frequencies. The resulting images are subjected to a further transformation that remaps its grayscale ramp through 3 to 9 sinusoidal modulations. This creates linear structures at varying scales. This process can be seen in Figure 6. In all, five carrier frequencies were used and each was subjected to three levels of linear structuring, for a total of fifteen spatial frequency conditions. For a detailed description of the stimulus generation process, see Appendix A.

A sample set of stimuli is shown in Figure 7. In each row, the carrier frequency increases, right to left, and the frequency of the superimposed linear structure increases from top to bottom.

Procedure

The procedure was the same as Experiment 1. A temporally separated pair of images was presented and the subject indicated if they were "same" or "different". Each subject participated in two blocks of 24 repetitions of each of the 15 stimulus levels, for a total of 720 trials/subject.

Results and discussion

The results for Experiment 2 are shown in Figure 8. The left panel of this figure shows the thresholds obtained for the individual carrier $\times$ modulation combinations. Note that there is relatively little decline with increasing frequency in comparison to Experiment 1. In other words, the increase of high-frequency energy caused by the sinusoidal modulations of the grayscale had relatively little effect on observers' performance. The right panel of Figure 8 shows performance only as a function of the carrier frequency. Note that when we plot the data in a

way that ignores the frequencies introduced by the sinusoidal modulations, the results are quite similar to Experiment 1. That is to say, the increase of high-frequency energy caused by the carrier textures had a large effect on observers' performance. These findings show clearly that it is not spatial frequency per se that limits the abilities of observers to discriminate macroscopic texture patterns. The important characteristic of our sinusoidal modulations of the grayscale is that they increase the high-frequency energy in the amplitude spectrum, while preserving the phase alignments of the carrier texture in the phase spectrum. This suggests that the perceptually relevant information for discriminating these textures is contained primarily in their phase spectra. If this hypothesis is correct, then we should be able to attenuate the performance for textures with sinusoidal modulations by randomly scrambling their phase spectra. Experiment 3 was designed to test this prediction.

Experiment 3

Methods

Subjects

The observers were four adults, all with normal or corrected-to-normal vision. All were members of the laboratories conducting the research and aware of the purpose and scope of the experiments.

Apparatus

The apparatus was the same as in Experiments 1 and 2.

Stimuli

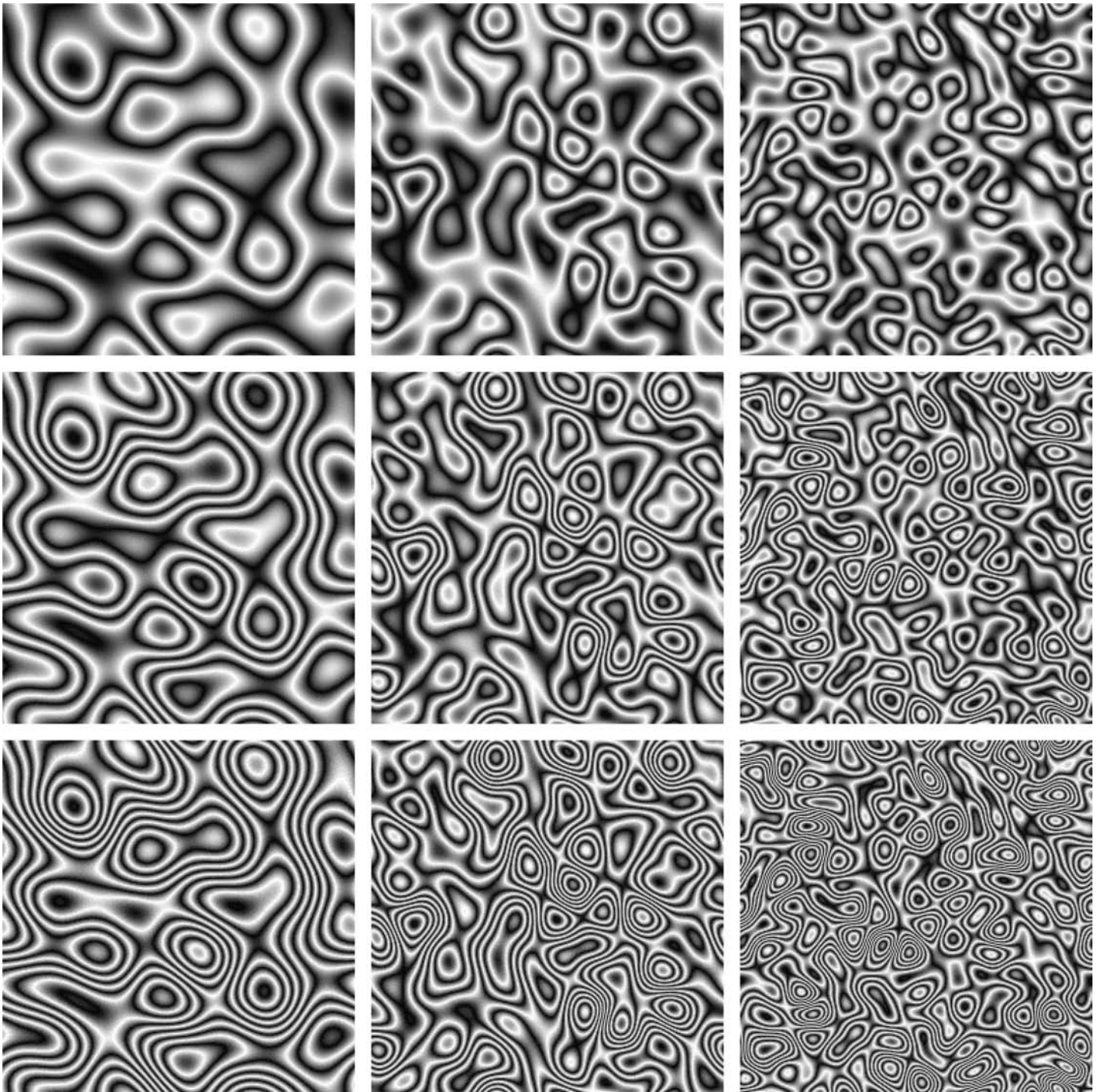The stimuli consisted of patches of low-pass filtered, spatially isotropic, uniformly distributed noise, subjected

Figure 7. Sample stimulus set for Experiment 2. In each row, the carrier frequency increases, right to left, and the frequency of the superimposed linear structure increases from top to bottom.

to modulation as in Experiment 2. Based on the results of Experiment 2, three levels of overall spatial frequency (e.g., carrier × modulation) were selected to cover the transition range from meso- to macrostructure as established in Experiment 1—2, 4, and 6 cycles/deg. This was accomplished by using three carrier frequencies of 0.33, 0.66,

and 1.0 cycles/deg, respectively, and modulating each by six cycles. In order to disrupt the macroscopic alignments in these images, we adopted a technique used previously by Hansen and Hess (2007), Thomson, Foster, and Summers (2000), and Wichmann, Braun, and Gegenfurtner (2006) to randomize the global phase spectra. In this
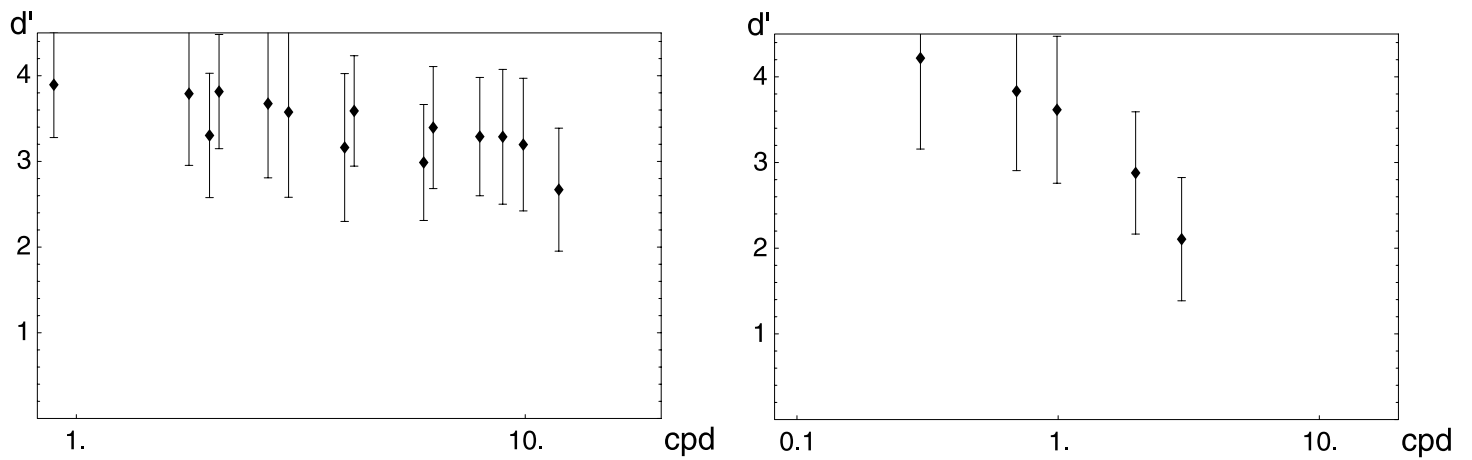
Figure 8. Results from Experiment 2. d' is a function of cycles per degree. On the left are the results for the frequency of the linear structure. On the right, the results are collapsed onto the five carrier frequencies. Error bars indicate $\pm\hat{\sigma}$.

experiment, each stimuli's global phase, ϕ, was jittered at one of five levels, ranging from $\pm\pi/5$ to $\pm\pi$, normally distributed. A sample set of stimuli is shown in Figure 9. For a detailed description of the stimulus generation process, see Appendix A. The effect of the phase scrambling is to transition the stimulus from containing structured information (for example, using a 1 cpd carrier in the lower row of Figure 9) to unstructured, higher frequency information (6 cpd in the lower row of Figure 9).

Procedure

The procedure was the same as Experiments 1 and 2. A temporally separated pair of images was presented and the
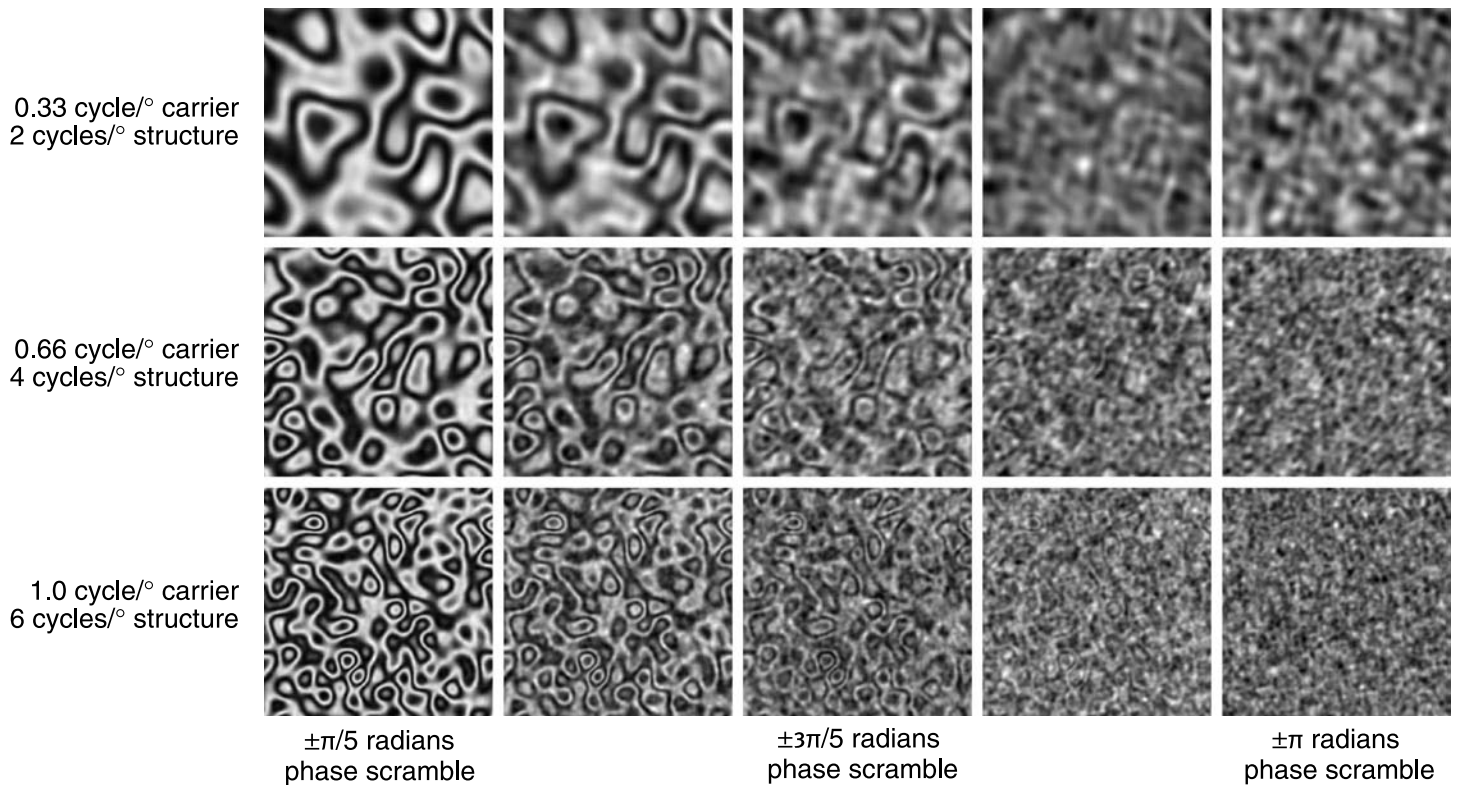

Figure 9. Sample stimulus set for Experiment 3. Three levels of spatial frequency were chosen from the stimulus set used in Experiment 2, using carrier frequencies of 0.33, 0.66, and 1.0 cycles/deg with modulations resulting in 2, 4, and 6 cycles/deg of structured information. These are shown in each row. The Fourier domain phase was scrambled in magnitudes ranging from $\pm\pi/5$ to $\pm\pi$ radians in five steps, seen here in each column. As the scrambling increases, the higher frequency modulated structure disappears and the images revert to the lower frequency carrier-based features.

Original texture   Mean of phase scrambled textures

0.33 cycle/° carrier
2 cycles/° structure

0.66 cycle/° carrier
4 cycles/° structure
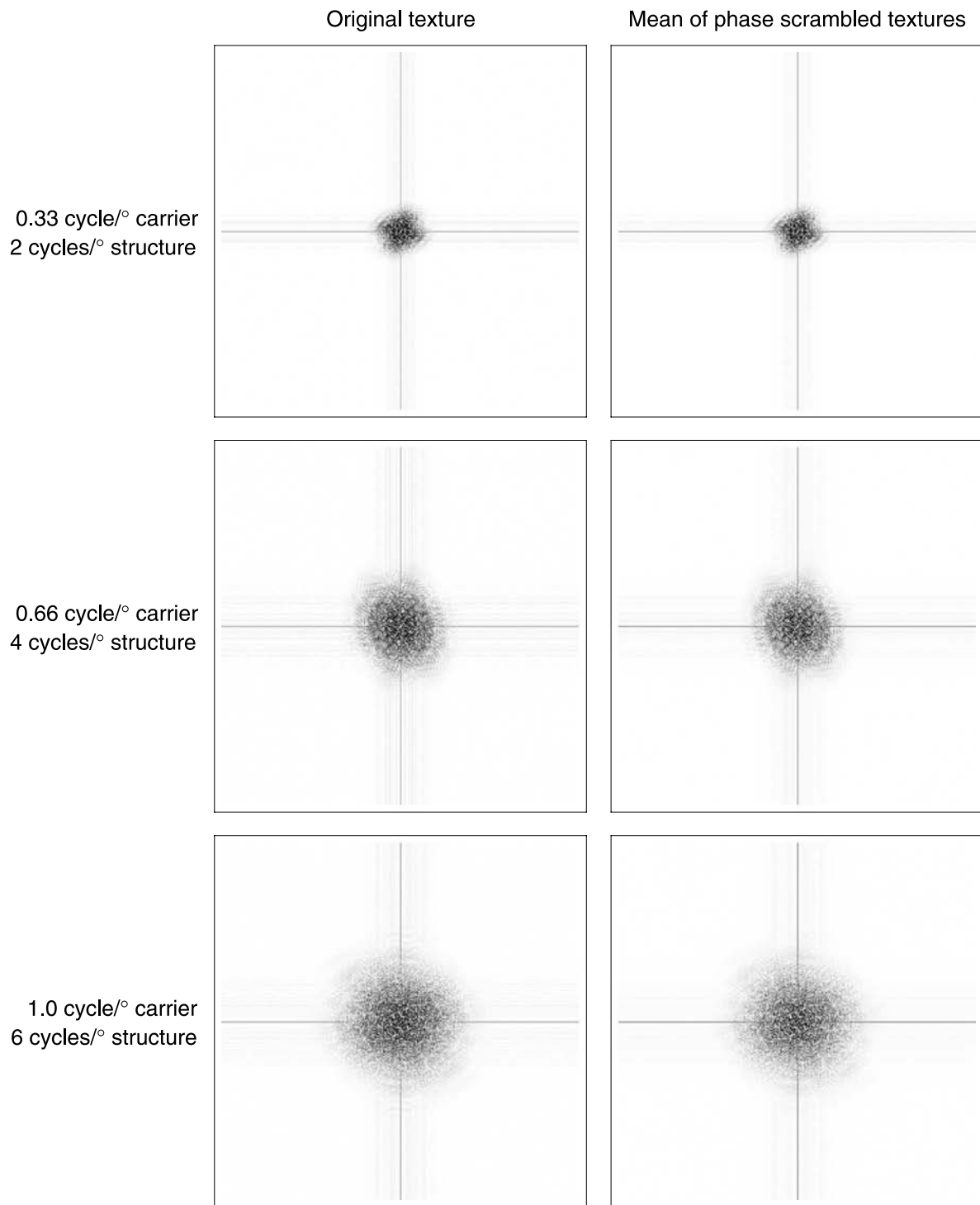
1.0 cycle/° carrier
6 cycles/° structure

Figure 10. Power spectra for the stimulus set shown in Figure 9. The three rows correspond to the three base frequencies shown in each row of Figure 9—0.33, 0.66, and 1.0 cycles/deg. On the left is the power spectra of the unscrambled stimulus and on the right is the mean of the phase-scrambled stimuli ($\pm\pi/5$ to $\pm\pi$ radians).

subject indicated if they were "same" or "different". Each subject participated in two blocks of 24 repetitions of each of the 15 stimulus levels, for a total of 720 trials/subject (Figure 10).

Results and discussion

Results from Experiment 3 are shown in Figure 11. In the low-frequency condition, the task is trivial regardless
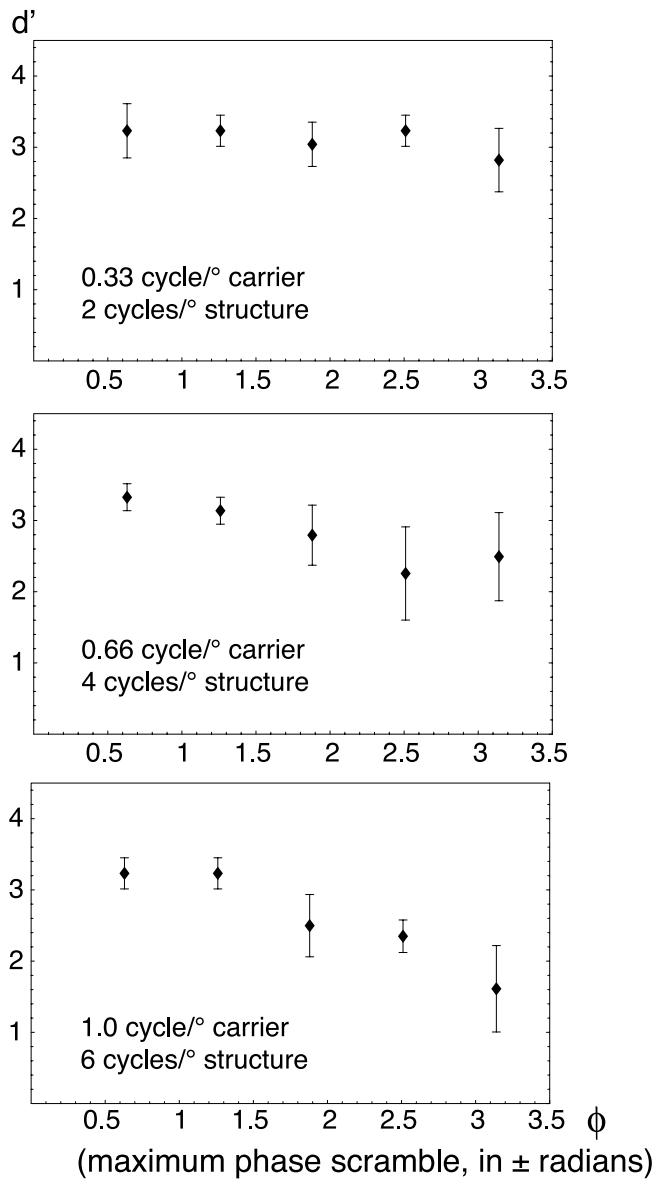
Figure 11. Results from Experiment 3. d' is a function of phase scrambling. Error bars indicate $\pm\hat{\sigma}$. Each graph shows the result from one of the three spatial frequency conditions, respectively. Recall that, as the phase scrambling increases, the appearance of the stimuli changes from predominantly that of the carrier to that of the total modulated frequency. Note that discrimination performance is very good, regardless of phase scrambling, in the 2 cycles/deg total condition, but in the 6 cycles/deg total condition, performance falls off as the amount of scrambling increases.

of the phase disturbance. However—as the spatial frequency of the scrambled information increases, performance begins to drop off as a function of phase scrambling. Recall that the stimuli were generated by modulating a given carrier frequency (0.33, 0.66, and 1 cpd)—all of which should be easily discriminable according to the results of Experiments 1 and 2 (a 75% threshold of 2 cpd). Indeed, in all three conditions, when the phase scrambling was relatively small, performance was consistent with the results from Experiment 2 showing that performance was dependent on the carrier frequency. However, as the scrambling increased, performance varied across the three conditions. In the low-frequency carrier condition, when the phase spectra was totally scrambled, the resulting stimuli contained features that were still above threshold, so the task was easily completed. However, in the high-frequency condition, as the disruption of the phase spectra yielded features below threshold, the task became more difficult.

General discussion

The research described in the present article provides strong empirical evidence that human observers make use of two distinct procedures for discriminating patches of texture. One possible approach that has been widely discussed in the literature is to compare the feature histograms for each patch. This would make it possible, for example, to easily distinguish a texture that produces uniform filter responses at all orientations from another that activates only a narrow range of orientations. This is the approach that observers most likely employ when making judgments of surface mesostructure. However, when the surface macrostructure contains perceptually distinct shapes or configurations, then it is also possible to discriminate textures based on the macroscopic alignments of filter responses at different spatial locations. During the debriefing sessions in the present experiments, observers spontaneously noted that they consciously searched for some salient feature in the first texture of each trial that they could attempt to match with a corresponding feature in the second texture. Examples they reported included such things as an "S" shaped contour or a collinear alignment of blobs.

Experiments 2 and 3 provide additional evidence that the relevant information for these macroscopic judgments is entirely contained within the Fourier phase spectra of the texture patterns. Experiment 2 demonstrates that texture patterns with frequencies in the mesostructure range can still be discriminated if the pattern of phase alignments is in the macrostructure range. Moreover, discrimination performance for these stimuli is reduced to chance when their phase spectra are randomly scrambled, as shown in Experiment 3.

In principle, it might be possible to account for these results using a statistical approach that incorporates higher order moments. For example, Thomson and colleagues (Thomson, 1999a, 1999b; Thomson & Foster, 1997; Thomson et al., 2000) have shown that phase perturbations of natural images can produce significant changes in higher order image statistics of degree 3 and above (e.g., skewness or kurtosis). Thus, if these higher order measures are important for the perceptual analysis of
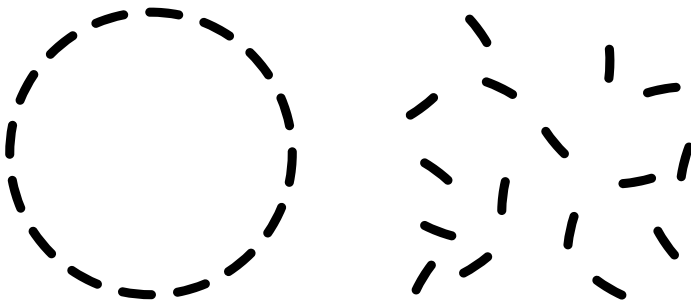
Figure 12. The pattern on the left shows a set of line segments whose orientations change as a function of position in a smoothly continuous manner so that the pattern is perceived as a circle. The pattern on the right contains the same set of line segments with the same orientations but their spatial positions have been randomly scrambled. If these patterns were visually encoded as histograms of local luminance values, or histograms of Gabor filter responses at different local scales and orientations, there would not be sufficient information in either of those data structures for determining whether or not the elements of a pattern are arranged in a circular configuration.

natural images, this could explain why recognition is impaired by phase scrambling. However, it cannot explain performance in the present experiments, because there were no systematic differences in skewness or kurtosis in the distributions of texture patterns that observers were required to discriminate.

An important limitation of this type of statistical approach is that it is insensitive to the long-range alignments of features in an image. Consider, for example, the pair of patterns presented in Figure 12. The pattern on the left shows a set of line segments whose orientations change as a function of position in a smoothly continuous manner so that the pattern is perceived as a circle. The pattern on the right contains the same set of line segments with the same orientations but their spatial positions have been randomly scrambled. If these patterns were visually encoded as histograms of local luminance values, or histograms of Gabor filter responses at different local scales and orientations, there would not be sufficient information in either of those data structures for determining whether or not the elements of a pattern are arranged in a circular configuration.

A similar distinction between higher order statistics and the alignment of features has recently arisen in theoretical analyses of the perception of gloss. Motoyoshi, Nishida, Sharan, and Adelson (2007) have argued that the apparent glossiness of a surface is determined statistically by the skewness of its luminance histogram or the skewness of the histogram of sub-band filter outputs. Anderson and Kim (2009) have shown, however, that this purely statistical analysis cannot adequately account for the perception of gloss because specular highlights only appear shiny when they are appropriately aligned with

the direction of surface curvature (see also Beck & Prazdny, 1981; Todd, Norman, & Mingolla, 2004).

Perhaps it might be possible to develop a statistical measure of contour alignments that incorporates some type of high order feature that explicitly encodes relative orientations in different spatial locations (e.g., see Geisler, 2008), but we suspect this is unlikely. Having a specific filter for each possible global shape or configuration is obviously not a feasible approach. A more sensible solution would be to somehow encode shapes using a more limited set of components (e.g., Biederman, 1987; Hoffman & Richards, 1984; Hoffman & Singh, 1997). However, a mere counting of these components is clearly insufficient to adequately characterize globally distinct shapes. It is also necessary to specify how they are arranged with respect to one another. Thus, we suspect that the most likely data structure for the perceptual representation of shapes is a graph of their components rather than a distribution.

One possible mechanism for extracting the global phase alignments in an image is the boundary contour system that was originally proposed by Grossberg and Mingolla (1985a, 1985b). This system consists of two stages: The first stage is characterized by short-range competition between cells with different orientational tuning. Incoming signals that survive this competition are then passed to a second stage in which there is long-range cooperation between cells whose orientational tunings are approximately aligned with one another. Activity at this stage is then fed back to the first level in order to fill in gaps along a boundary. One of the most interesting properties of this system is that high-frequency filters can cooperate over large spatial extents. In other words, the scale of the phase alignments can be much larger than the scale of the local mechanisms that are aligned just as we observed in Experiment 2.

Although the model of Grossberg and Mingolla provides an elegant explanation of how a pattern of phase alignments could potentially be extracted from an image, it does not offer any suggestions about how boundary contours might be parameterized to encode their specific shapes. This is an important issue that remains for future research.

Appendix A

Stimulus construction

The stimuli in each experiment start as a patch of uniformly distributed noise. The noise is low-pass filtered at various spatial frequencies to create the base textures used in Experiment 1. These are further transformed by remapping image intensity to introduce linear coherent
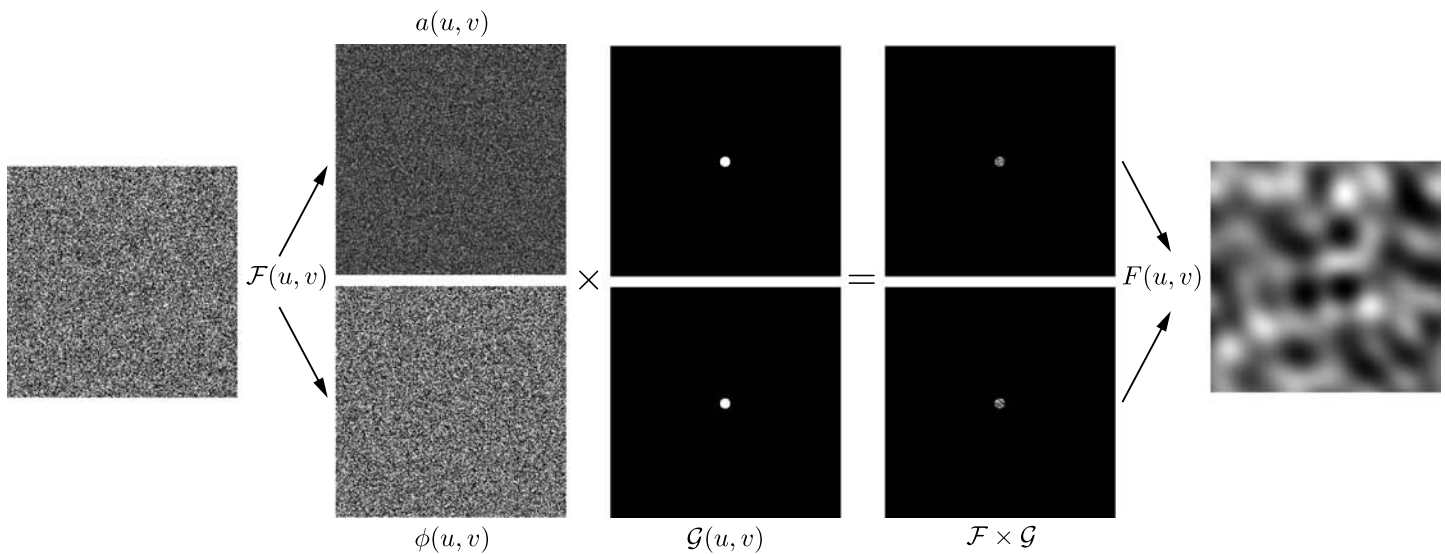
Figure A1. On the left is the base discrete uniformly distributed noise, *n*. It is decomposed using a Fourier transform, $\mathcal{F}(u, v)$. The log-scaled amplitude spectrum, *a,* is shown at the top and the phase spectrum, ϕ, at the bottom. The decomposed image is multiplied by the appropriately scaled infinite response low-pass filter, $\mathcal{G}$, and recomposed via the inverse Fourier transform, *F(u, v)*.
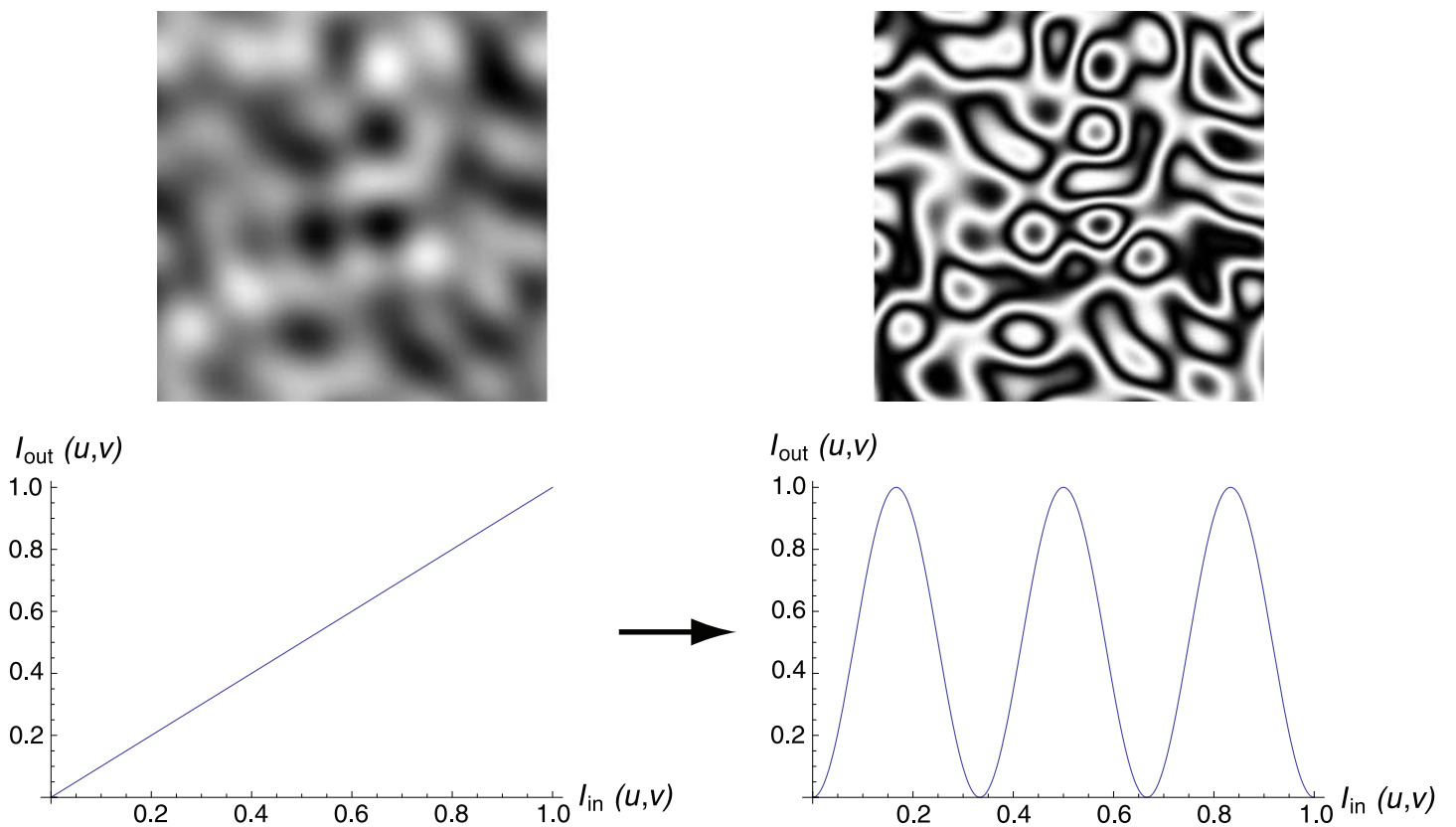

Figure A2. On the left is a base texture as derived above, whose depicted gray level is a linear function. For the image on the right, that linear ramp was remapped to a sinusoidal function, so that each blob in the original texture is transformed into a set of concentric dark and light bands.
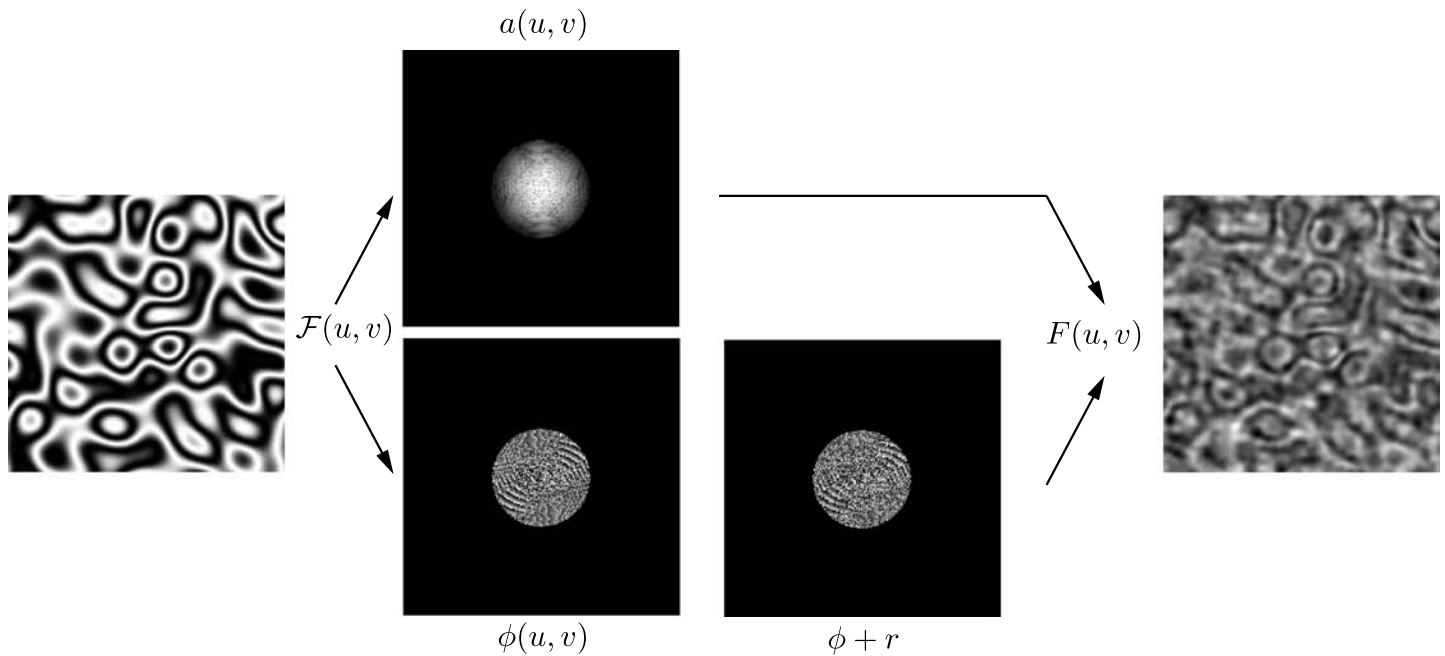
Figure A3. To create the phase-scrambled structured textures, the structured texture is first decomposed via Fourier transform. The phase is then subjected to a given amount jitter by the addition of a field of normally distributed random noise on the interval $\pm(0, \pi]$. This jittered phase is recombined with the amplitude spectra for the original texture and subjected to an inverse Fourier transform.

structure for Experiment 2. Finally, this structure is disrupted by jittering the phase spectrum for Experiment 3.

Base textures. First, define $n(x, y)$ as a two-dimensional field of uniformly distributed random discrete noise with dimension $N \times N$. This base noise is subjected to low-pass filtering at a specified frequency, f_{base}, by decomposing n using a Fourier transform, $\mathcal{F}(u, v)$. The decomposed image is multiplied by an infinite response low-pass filter, $\mathcal{G}$, and recomposed via the inverse Fourier transform, $F(u, v)$. An illustration of this process is shown in Figure A1.

Structured textures. To create the structured textures, a simple image intensity remapping is employed. Figure A2 shows an example of this process. The intensity of each pixel of the input image, $I_{in}(x, y)$, is mapped to a sinusoidal function, given by

$$I_{out}(x,y) = \frac{1}{2}\left(-\cos(f_{struct}2\pi I_{in}(x,y))-1\right)+1, \quad \text{(A1)}$$

where f_{struct} is the frequency of the desired structure. The highest spatial frequencies of the resulting textures are now roughly equivalent to the frequency of the base texture, f_{base}, multiplied by f_{struct}.

Phase-scrambled structured textures. Finally, these structured textures are subjected to phase scrambling to "de-cohere" the structure while maintaining the same amplitude spectra. To create these textures, the structured texture images are subjected to Fourier decomposition.

The phase spectrum is then combined with a map of normally distributed noise, r. The range of r varies on the interval $\pm(0, \pi]$. The amplitude spectrum is then combined, unmolested, with the jittered phase and subject to an inverse Fourier transform, which yields the final stimulus. This process is illustrated in Figure A3.

Footnotes

[1]The ATI 1950 series uses 10-bit DACs, but the Apple Pro displays are only 8-bit capable.
[2]For example, a large black blob that is situated on the upper right corner could be easily diagnostic.

References

Anderson, B. L., & Kim, J. (2009). Image statistics do not explain the perception of gloss and lightness. *Journal of Vision, 9*(11):10, 1–17, http://www.journalofvision.org/content/9/11/10, doi:10.1167/9.11.10. [PubMed] [Article]

Beck, J., & Prazdny, S. (1981). Highlights and the perception of glossiness. *Perception & Psychophysics, 30,* 407–410. [PubMed]

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

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94,* 115–147. [PubMed]

Chubb, C., & Landy, M. (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.

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59,* 167–192. [PubMed]

Georgeson, M. A., May, K. A., Freeman, T. C. A., & Hesse, G. S. (2007). From filters to features: Scale-space analysis of edge and blur coding in human vision. *Journal of Vision, 7*(13):7, 1–21, http://www.journalofvision.org/content/7/13/7, doi:10.1167/7.13.7. [PubMed] [Article]

Grossberg, S., & Mingolla, E. (1985a). Neural dynamics of form perception: Boundary completion, illusory figures, and neon color spreading. *Psychological Review, 92,* 173–211. [PubMed]

Grossberg, S., & Mingolla, E. (1985b). Neural dynamics of perceptual grouping: Textures, boundaries, and emergent segmentations. *Perception & Psychophysics, 38,* 141–171. [PubMed]

Hansen, B. C., & Hess, R. F. (2007). Structural sparseness and spatial phase alignment in natural scenes. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 24,* 1873–1885. [PubMed]

Hesse, G. S., & Georgeson, M. A. (2005). Edges and bars: Where do people see features in 1-D images? *Vision Research, 45,* 507–525. [PubMed]

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition, 18,* 65–96.

Hoffman, D. D., & Singh, M. (1997). Salience of visual parts. *Cognition, 63,* 29–78. [PubMed]

Huang, T., Burnett, J., & Deczky, A. (1975). The importance of phase in image processing filters. *IEEE Transactions on Acoustics, Speech, and Signal Processing, 23,* 529–542.

Koenderink, J. J., & van Doorn, A. (1998). Phenomenological description of bidirectional surface reflection. *Journal of the Optical Society of America A, 15,* 2903–2912.

Koenderink, J. J., & van Doorn, A. J. (1996). Illuminance texture due to surface mesostructure. *Journal of the Optical Society of America A, 13,* 452–463.

Kovesi, P. (1999). Image features from phase congruency. *Videre, 1,* 1–27.

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

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

Motoyoshi, I., Nishida, S., Sharan, L., & Adelson, E. H. (2007). Image statistics and the perception of surface qualities. *Nature, 447,* 206–209. [PubMed]

Oppenheim, A., & Lim, J. (1981). The importance of phase in signals. *Proceedings of the IEEE, 69,* 529–541.

Piotrowski, L. N., & Campbell, F. W. (1982). A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase. *Perception, 11,* 337–346. [PubMed]

Ramachandran, G., & Srinivasan, R. (1961). An apparent paradox in crystal structure analysis. *Nature, 190,* 159–161.

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

Thomson, M. G. (1999b). Visual coding and the phase structure of natural scenes. *Network, 10,* 123–132. [PubMed]

Thomson, M. G., & Foster, D. H. (1997). Role of second- and third-order statistics in the discriminability of natural images. *Journal of the Optical Society of America A, 14,* 2081–2090.

Thomson, M. G., Foster, D. H., & Summers, R. J. (2000). Human sensitivity to phase perturbations in natural images: A statistical framework. *Perception, 29,* 1057–1069. [PubMed]

Todd, J. T., Norman, J. F., & Mingolla, E. (2004). Lightness constancy in the presence of specular highlights. *Psychological Science, 15,* 33–39. [PubMed]

Wichmann, F. A., Braun, D. I., & Gegenfurtner, K. R. (2006). Phase noise and the classification of natural images. *Vision Research, 46,* 1520–1529. [PubMed]