Measuring perceived differences in surface texture due to changes in higher order statistics

K. Emrith,[1,*] M. J. Chantler,[1] P. R. Green,[2] L. T. Maloney,[3] and A. D. F. Clarke[1]

[1]*School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK*
[2]*School of Life Sciences, Heriot-Watt University, Edinburgh, UK*
[3]*Department of Psychology, New York University, New York, New York 10003, USA*
**Corresponding author: k.emrith@hw.ac.uk*

Received August 10, 2009; revised January 7, 2010; accepted March 8, 2010;
posted March 24, 2010 (Doc. ID 115546); published April 30, 2010

We investigate the ability of humans to perceive changes in the appearance of images of surface texture caused by the variation of their higher order statistics. We incrementally randomize their phase spectra while holding their first and second order statistics constant in order to ensure that the change in the appearance is due solely to changes in third and other higher order statistics. Stimuli comprise both natural and synthetically generated naturalistic images, with the latter being used to prevent observers from making pixel-wise comparisons. A difference scaling method is used to derive the perceptual scales for each observer, which show a sigmoidal relationship with the degree of randomization. Observers were maximally sensitive to changes within the 20%–60% randomization range. In order to account for this behavior we propose a biologically plausible model that computes the variance of local measurements of phase congruency. © 2010 Optical Society of America

OCIS codes: 330.5000, 330.5020, 330.5510.

1. INTRODUCTION

The frequency channel model has provided many valuable insights into human spatial vision [1,2], and it has also provided a wealth of features for computer classification of the image texture [3,4]. Much of this research has been carried out on the first and second order statistics of images and it is now well understood as to how these statistics influence perception, and how models—typified by the filter-rectify-filter (FRF) structure—can be used to account for these effects [5–8]. Many studies have also exploited the frequency channel model to investigate "*n*th order statistics" particularly with respect to the preattentive segregation of patterned image regions or abutted images [9–19]. Although these studies provided strong psychophysical evidence of the ability of humans to perform discrimination and segregation tasks, they were, however, based on highly stylized binary textures constructed using geometric elements for which the control and computation of third and higher order statistics are reasonably straightforward.

Unfortunately standard methods for modeling higher order statistics *per se* in natural (gray-level) images are often extremely complex [20–22] and it is difficult to obtain an intuitive understanding of what the many parameters represent. However, many researchers have pointed out that most of the visually pertinent information in an image is encoded in its phase spectrum and there are numerous examples that show that phase randomizing an image (also referred to as "scrambling") leaves it largely unrecognizable [23–28]. Furthermore, it is well known that there is an intimate relationship between the higher order statistics and the phase spectra, and studies have demonstrated that visually salient features in images (such as edges and bars) correspond to the points where the various harmonics representing the image (as a two-dimensional signal) have the same phase alignment (known as phase congruency) [29,30].

While phase randomization almost certainly changes the higher order statistics of images, it is also likely to affect the first order statistics since phase randomizing a wide bandwidth image will (due to the central limit theorem) result in a normal distribution. Some studies have investigated the first order natural image statistics that could potentially account for the visual effects of phase perturbing natural images [8,31,32]. Additionally, it has also been reported that, while the first and second order statistics are independent in natural images, a correlation between them appears when the images are phase scrambled [8,33,34].

Researchers have recently exploited techniques for gradual phase randomization of natural images in order to investigate the amount that is required before recognition or segregation becomes impossible [32,35–40]. Of these studies only two [32,38] have related the observer's performance to metrics directly derived from the image data. Thomson *et al.* [32] measured higher order *moments* (skewness and kurtosis) derived from first order statistics of the image histogram rather than higher order statistics. Hansen and Hess [38] used phase scrambling over different frequency bands to investigate the degree of cross-scale phase alignment required for the identification of natural or naturalistic scenes. They proposed a "structural sparseness metric" (SSM) in order to aid the interpretation of their results and concluded that observers are less tolerant to phase scrambling when identifying more structured scenes (for example, a scene representing

 © 2010 Optical Society of America

a dense forest will contain significantly more structural information than one depicting the sea or a blue sky).

To our knowledge, all of the above mentioned papers employed phase spectra scrambling while seeking to maintain constant power spectra. However, while several studies do normalize the mean and variance of the image luminance they do not explicitly control other higher order moments (skewness and kurtosis) derived from the image histogram. Thus it is unclear as to whether the perceived image variations are due to changes in the higher order moments of first order statistics or due to higher order statistics. Furthermore, none of these studies directly investigate the correlation of an image metric with the perceived changes caused by the gradual perturbation of the phase spectrum.

A. Current Study

Given that the contribution of first and second order statistics for segregation or discrimination tasks is well researched, it is intriguing to investigate how well we can detect changes in natural images that are due solely to changes in higher order statistics (that is statistics higher than second order). The goals of this paper therefore are (1) to investigate the ability of observers to perceive differences in images caused solely by changes in higher order statistics and (2) to propose a biologically plausible image processing model that accounts for these perceptions.

We present two experiments in which we use natural images and synthesized naturalistic images to investigate the ability of humans to detect small changes in higher order statistics. We do this by keeping the image histogram and the power spectrum constant while gradually phase randomizing the image. In the first experiment we use a large number of phase randomization levels to derive perceptual scales for each observer, whereas in the second experiment we use a smaller number of randomization levels with, however, a more extensive set of images.

2. METHODS

A 2-alternative forced choice (2AFC) procedure was used to capture human judgments. This method was preferred over some other popular methods such as the method of adjustment [41] or ratio scaling since it reduces the burden of having to arbitrarily assign values to randomized textures being discriminated, is easy to implement, and also requires a few trials to fit the observers' judgments to a perceptual scale. The estimation of a perceptual scale that corresponds to the amount of phase randomization was performed using the technique of maximum likelihood difference scaling (MLDS) [42]. The MLDS is a method that has been used for estimating suprathreshold differences across a range of images that have undergone some physical changes, for example, in the quantification of color differences [42], the direct measurement of human perception of image compression [43], or the estimation of gloss scales [44]. The MLDS works by using a set of four stimuli (a quadruple) chosen randomly from a full set of textures with different degrees of randomization. The MLDS requires the use of non-overlapping quadruples for each reference texture. A set of N randomized images (for each reference texture) allows the generation of $N!/(4!(N-4)!)$ non-overlapping quadruples. For 11 degrees of randomization we obtain 330 non-overlapping quadruples [e.g., (2,3) and (2,8) is an overlapping quadruple and is not counted in the 330 quadruples]. A more detailed explanation of the MLDS is provided in Appendix A.

A. Observers

The observers asked to perform the experiments presented in this study had normal or corrected-to-normal vision. All of the observers were naïve to (1) perceptual texture characteristics and the nature of the stimuli, and (2) the purpose of the experiments.

B. Stimuli

A set of 12 stimuli comprising six natural textures and six computer synthesized naturalistic textures were used in the psychophysical experiment. The natural textures were captured under unknown illumination conditions, whereas the naturalistic textures were synthesized under controlled conditions. In the second case, a Lambertian model of reflectance was used to render surface height maps that were generated using a random or semi-random placement of texture elements. Where primitives overlapped we took the maximum of the height of any primitive at that position. Unlike the natural textures, each synthetic naturalistic one could be generated repeatedly using different seeds controlling the placement of the texture elements. A total of ten seeds was used for each synthetic texture. This procedure ensured that the pairs of synthetic textures could not be compared pixel-wise, and so the synthetic textures provide a control for the possibility that observers compared the natural texture pairs in this way. Figure 1 shows all the reference textures used in the experiment. The first two rows in Fig. 1 show the natural textures, and rows 3 and 4 show the naturalistic ones.

All the reference textures were forced to follow normal intensity distributions before being phase randomized. While the naturalistic textures were generated with normal intensity distributions, the natural textures were mapped to normal distributions using the mean and standard deviation of their original distributions. Figure 2 shows the pea images before (left column) and after (right column) the mapping process. To obtain test stimuli with varying amounts of higher order statistics, the reference textures were subjected to different degrees of phase randomization. To ensure that the randomized images for each reference texture varied only in their higher order statistics, the textures were normalized to have the same first and second order statistics. The normalization process was performed at each randomization stage so that the resulting image was constructed using the partially randomized phase spectrum of each reference texture and its original power spectrum. This allowed the second order statistics to be kept constant for all partially randomized images. Additionally, each randomized image was subjected to a D'Agostino–Pearson normality test to verify whether its intensity distribution had deviated from the original normal distribution. If the null hypoth-

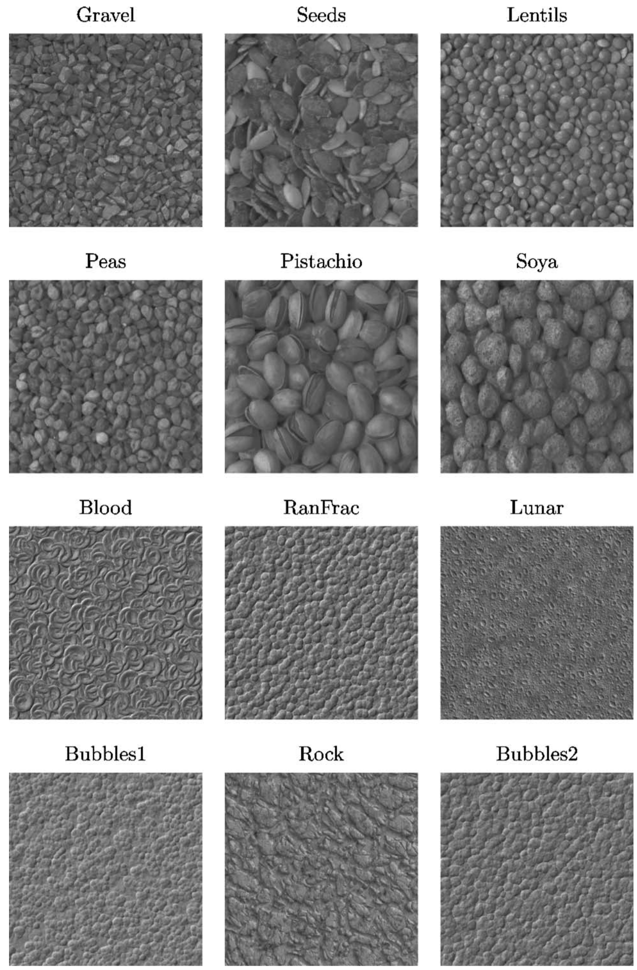
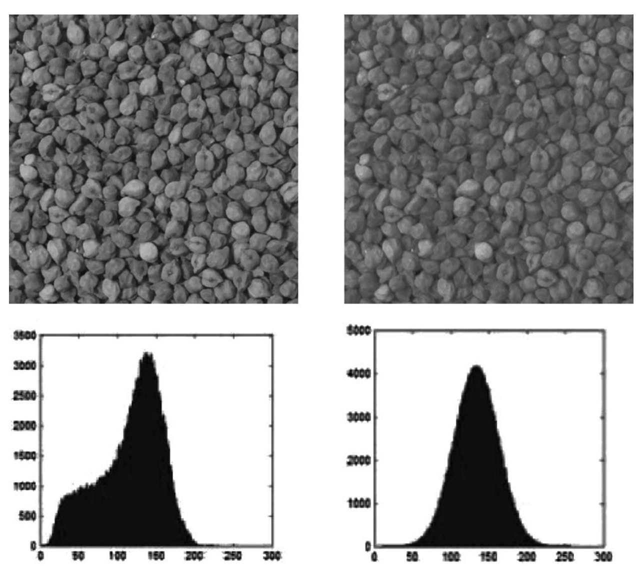


Fig. 1. Images used in the psychophysical experiments. Top two rows show the six natural images and the bottom two rows show the six computer synthesized textures.



(a) original (b) mapped to normal

Fig. 2. Example of a natural image whose intensity distribution is mapped to a normal one.

esis of a non-normal distribution was not rejected at the 0.05 level of significance, the processed image was forced to the reference texture's normal distribution and the second order statistics were again adjusted. This process is repeated until both first and second order statistics were the same as the original image.

Gradual phase randomization was performed by adding a random variable to the principal phase values of the original texture images. The random variable was drawn from the uniform distribution $[0,\sigma]$. Eleven degrees of phase randomization were used with degree 0 being $\sigma=0$ and degree 10 being $\sigma=2\pi$ for 100% phase randomization and with the other degrees representing linear increments in σ.

Complex conjugate symmetry was maintained in the randomized phase spectra in order to provide a zero power imaginary spatial domain image and to ensure that the second order statistics in the real spatial domain image remained constant. Figure 3 shows one synthetic (blood) texture and one natural (seeds) texture at four different levels of phase randomization (0%, 30%, 60%, and 100%). All the reference textures were non-periodic (and also tileable in the case of synthetic textures), and fully phase randomizing them results in visually continuous images. On the other hand, phase randomizing highly periodic textures does not lead to visually continuous images (see Fig. 4) and therefore such textures were not used in this study.

C. Experimental Setup

A 20 in. TFT (thin-film transistors) monitor (NEC LCD2090UXi) with a pixel pitch of 0.255 mm (100 dpi) was used to display a 2×2 array (quadruple) of images of size 512×512 pixels for each trial. The calibration of the gamma responses ($\gamma=2.2$) was performed using a Gareth Macbeth Eye One Pro spectrometer. The luminance of the monitor was fixed at 120 cd/m^2 with the color temperature set at 6500 K for a frame rate of 60 Hz. Observers fixed the screen from a distance of 70 cm, where it subtended a visual angle of 11°.

3. EXPERIMENT 1

The objective of experiment 1 was to investigate how well observers could discriminate between pairs of texture images that differ in their higher order statistics, expressed as the 11 degrees of randomization described in Section 2. While a larger number of trials (with more than 11 randomization levels) would provide higher confidence levels in estimating the perceptual scales using the MLDS, 330 trials per reference image is more realistic for the perceptual task considered. Experiment 1, however, considered only a subset of natural and synthetic images since it is not practical for observers to judge trials from all 12 reference textures (i.e., 3960 trials) at one go. Two natural (gravel and seeds) and two synthetic (blood and RanFrac) reference textures were used in this experiment.

A. Procedure

Six observers participated in this experiment. The observers were presented with two pairs (a quadruple) of stimuli (a,b) and (c,d) displayed one above the other and were

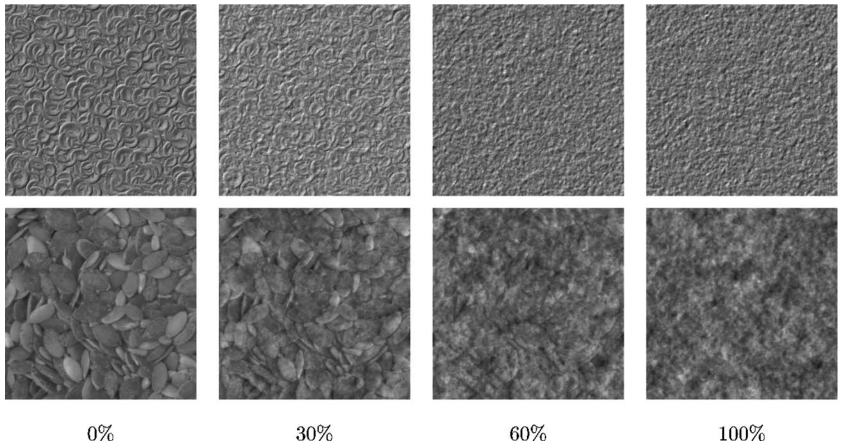


Fig. 3. Reference textures at different levels of randomization: Blood (top) and seeds (bottom).

asked to identify the pair that had the larger perceptual difference. The trials for each texture were presented sequentially with the observer having the option to take a break in between each set of trials presented. No time limit was imposed on observers in making their choice; however, the interface presented to them required that one of the two pairs was selected (forced choice mechanism). Observers did not have the option to return to a previous trial.

The MLDS technique requires that each trial is composed of textures having an ordered degree of randomization; however, there was no restriction in the way in which the images were presented to the observers. Thus the position (top or bottom) of each pair was randomized at each trial and also the position (left or right) within each pair was also randomized. Additionally, since the stimuli belonging to each reference texture were presented sequentially, the order in which the sets were presented to each observer was alternated. This was done in order to balance any effect of fatigue. The result R_i for each trial i was saved in a binary form ($R_i=0/1$) with a value of zero corresponding to the upper pair having the larger perceptual difference or 1 for the lower pair. The final results for each test texture were fed to the MLDS program to estimate the perceptual scales. A MLDS package implemented using the R programming language was used for the estimation process [45].

B. Results

Plots of the estimated perceptual scales for six observers are displayed in Figs. 5 and 6 for the chosen naturalistic (blood and RanFrac) and natural (gravel and seeds) images. Each plot shows how the difference scale values vary for the selected textures when their phase spectra were gradually randomized. The bootstrap procedure described by Maloney and Yang [42] was used to estimate the confidence intervals (± 1 SD) shown in the plots. We observe that all the plots in both Figs. 5 and 6 exhibit a sigmoidal behavior, monotonically increasing from 0 and saturating at 1. For both natural and synthesized textures, the plots clearly show that for an amount of phase randomization varying from 0% to 20%, the changes in the difference scale values are low for most of the observers. This indicates that observers encountered appreciable difficulty in discriminating texture pairs within the 0%–20% range.

The plots for the synthetic textures (Fig. 5) show sharp slopes for the range 20%–60% of phase randomization. This shows a greater ability of observers in discriminating between the synthetic texture pairs presented, thus indicating that observers were more sensitive to smaller differences in randomization within this range. Beyond the 60% mark, all observers perceived a little change in the appearance in the partially randomized synthetic textures.

Inspection of Fig. 3 suggests the basis for the sigmoidal relationship; at 30% randomization, the texture elements remain as visible as in the original image, whereas from 60% onward they are not. It should be noted that the observers were not explicitly asked to judge the texture pairs based on the visibility of texture elements.

While the plots for natural textures show similar shapes (see Fig. 6), we observe that the steep slopes extend to 70%–80% phase randomization. The greater ability of humans to judge perceptual difference between natural textures within a range up to 80% phase randomization may be due to the pixel-wise comparisons that observers were able to make for natural textures.

Additionally, the plots in Fig. 5 show very similar behavior for the two synthesized textures, while this is not the case for the two natural textures. We observe that the behavior for the seed texture is more linear within the range 20%–60% than for the gravel texture. A possible explanation may be that while the synthetic textures are

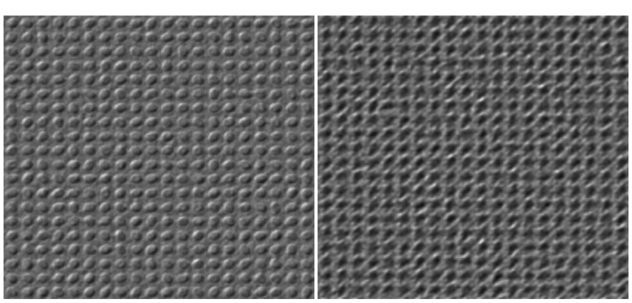


Fig. 4. Effect of full phase randomization on the appearance of a highly periodic texture (left, original; right, randomized).

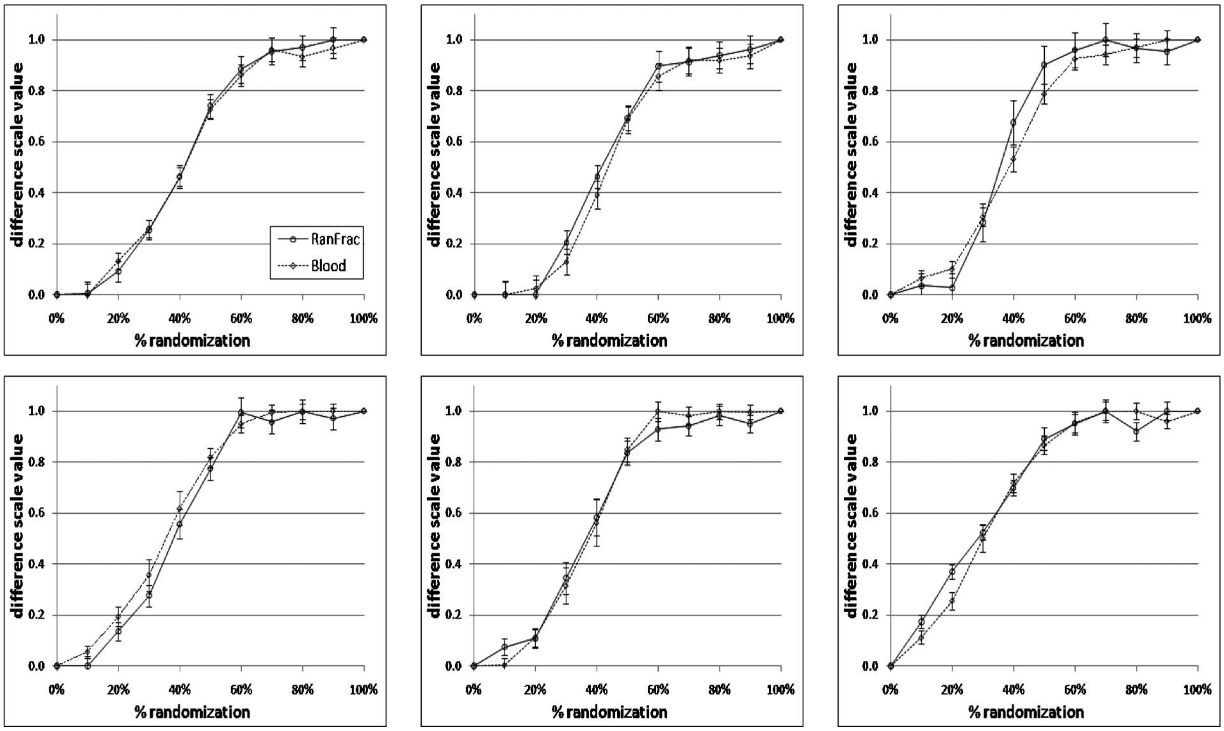


Fig. 5. Plots showing the behavior of individual observers' difference scales with changing amount of phase randomization for synthetic textures blood and RanFrac.

made up of a single texture element, the texture elements from the natural textures vary in size, shape, and contrast.

4. EXPERIMENT 2

Experiment 2 was carried out to investigate whether the behavior of the perceptual scales for the natural and synthetic textures is maintained for a larger set of textures. In this experiment a set of eight reference textures, comprising four synthetic and four natural textures, was used. To allow a larger set to be tested, the number of randomization levels was decreased leading to a fewer trials per reference image. Figures 5 and 6 from experiment 1 showed that beyond 80% randomization, observers were

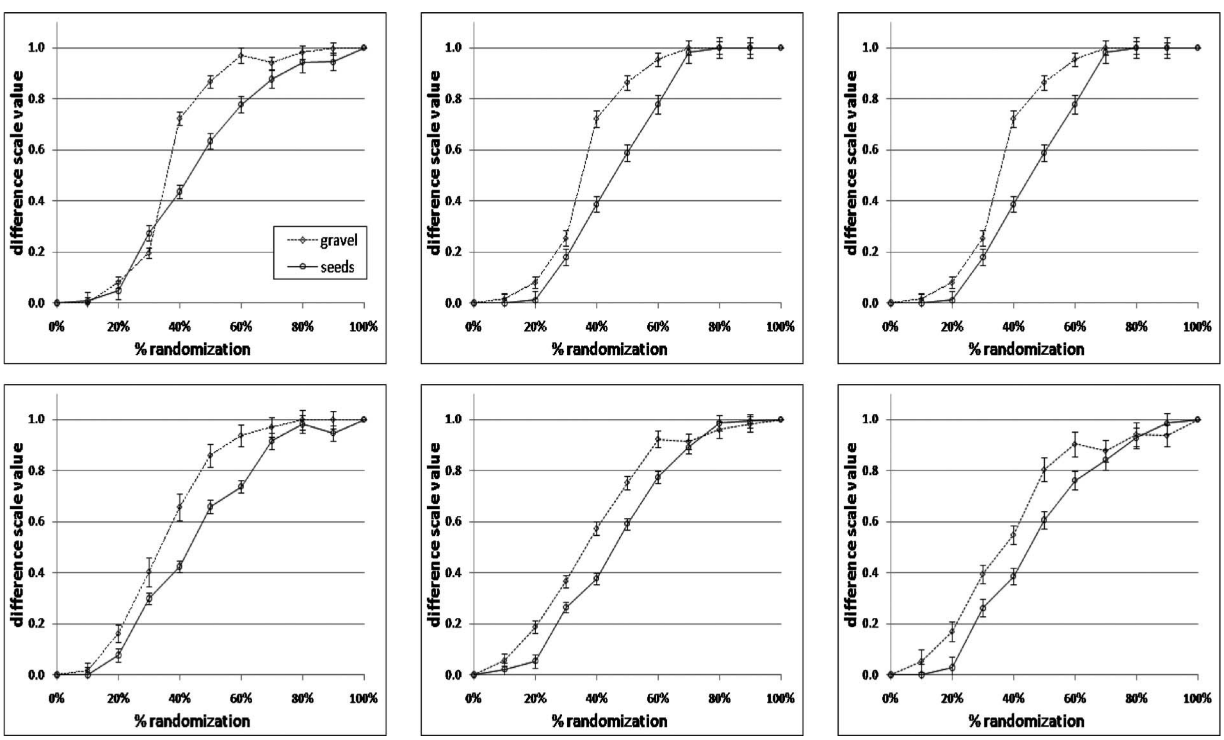


Fig. 6. Six plots showing the behavior of difference scales with changing amount of phase randomization for natural textures gravel and seeds.

unable to perceive any changes in the appearance of the randomized textures. Thus, only the first nine (i.e., 0%–80%) degrees of randomization were presented in the current experiment.

A. Procedure

The same procedure as that for experiment 1 was used. A total of four observers participated in this experiment. For each reference texture a set of 126 trials (quadruples) was presented to observers (i.e., a total of 1008 trials for eight textures). While the trials for each reference texture were presented in sequence, the presentation order for the natural and synthetic textures was randomized.

B. Results

Figure 7 shows the plots (left column) for the natural and synthesized textures tested and also the mean behavior (right column) for four observers who participated in this

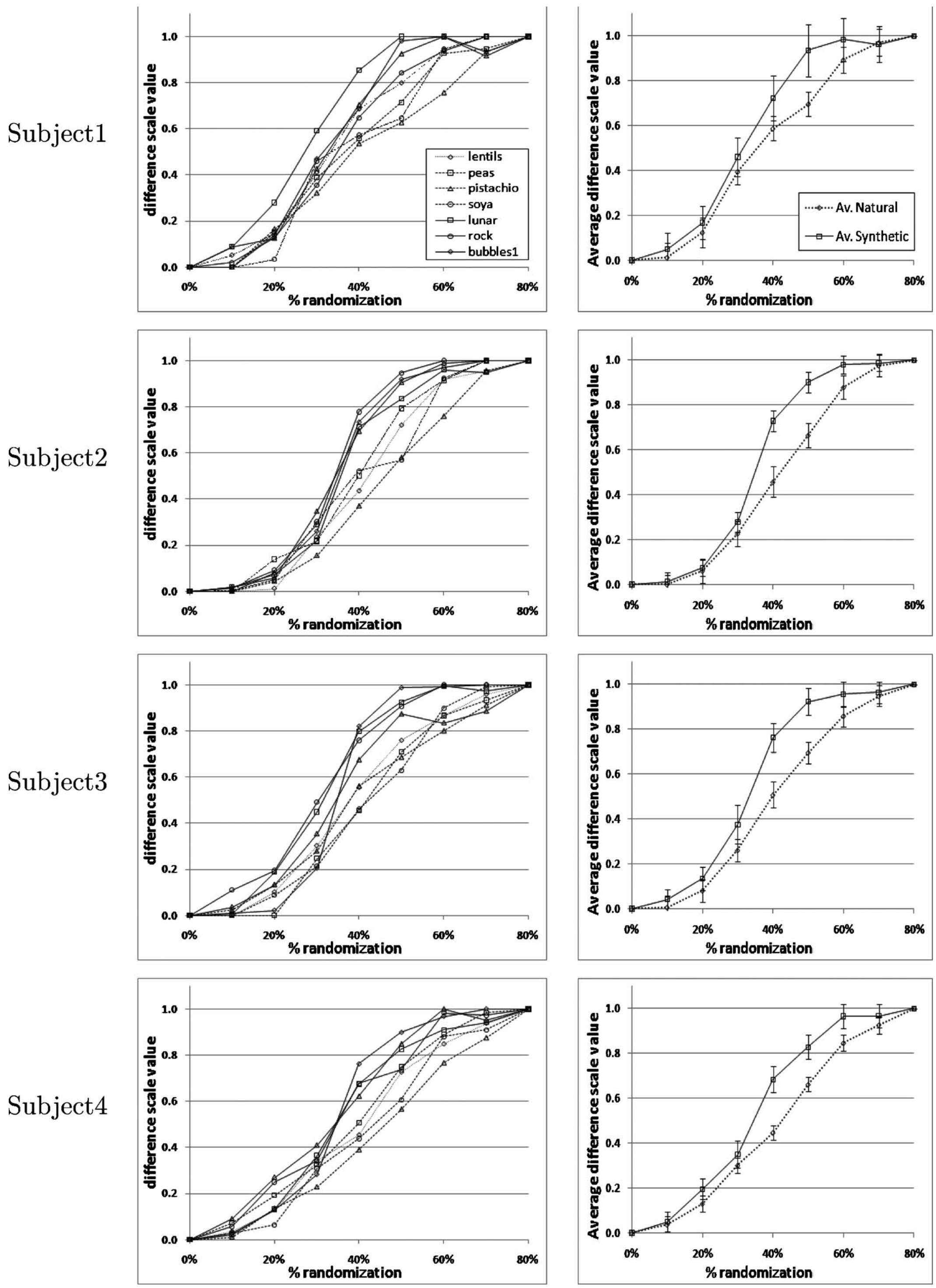


Fig. 7. Plots showing the behavior of difference scales for a set of four synthetic and four natural textures using only nine randomization levels (0%–80%). Column 1 shows the plots for all textures and column 2 shows the average behavior for four subjects.

experiment. The plots for the eight different textures confirm the general behavior (i.e., monotonic and sigmoidal) of the perceptual scales derived in experiment 1, and also provide additional evidence that observers had a greater ability to discriminate smaller changes in the appearance of synthetic textures within the 20%–60% range of randomization while being more sensitive to a greater range (up to 70%–80%) for natural textures. Additionally, the perceptual scales for the natural textures also appear more linear, with a rather constant slope, as compared to the scales for the synthetic textures. These observations may be due to the fact that observers may have used strong localized information as characterized by highlights and shadow information to make pixel-wise comparison for natural textures, although most structural information was destroyed beyond the 60% randomization level as illustrated in Fig. 3.

5. MODELING OF PERCEPTUAL DIFFERENCES

In this section we seek a biologically plausible model that can generate a perceptual measure with the following characteristics: (1) it has a monotonic sigmoidal relationship with increasing phase randomization and (2) it shows a steep change in the range 20%–60% of phase randomization.

It is well known that higher order statistics (i.e., higher than second order) are affected by the phase relationship of patterns [26]; however, a simple model to represent phase information is difficult to achieve due to its complex representation (mainly due to phase wrapping in the range $[-\pi,+\pi]$). Since it is known that natural images contain structure that is aligned locally in phase space [29,30], we have investigated and applied Kovesi's phase congruency model [30] to represent the change in the appearance of textured surfaces at different levels of phase randomization. We observed in Fig. 3 that randomizing the phase spectra of the texture images destroys the spatial arrangements of local features and changes the appearance of the texture surfaces. In this study, therefore, we provide what we believe to be a novel feature derived from Kovesi's phase congruency model [30], which characterizes the change in the appearance of the surface textures, and we show that this feature meets the criteria described above.

Kovesi's phase congruency model [30] was inspired by the *local energy model* presented by Morrone and Burr [29], which models the way in which the human visual system uses odd and even symmetric receptors in the visual cortex to decode local features such as edges and lines. Morrone and Burr's model [29] uses phase congruence information to detect these local features and it has been applied successfully both to the segmentation of visual scenes and to predict the perceptual appearance of these scenes [30,46,47]. Their model consists of two stages. In the first stage quadrature filter pairs (odd and even symmetric) are applied at different spatial frequencies and orientations. A function based on the sum of squares of the responses from these filters is computed, and peaks corresponding to salient features are identified. The second stage classifies these peaks in terms of different perceptual features (such as edges or bars). Morrone and Burr's model [29], however, does not provide good localization of local features due to its dependence on the local contrast. Kovesi [30] improved Morrone and Burr's model [29] to provide better localization of features by computing the phase congruence information that is invariant to changes in the image contrast and also by identifying and compensating for noise.

The appearance of surface texture is primarily characterized by the presence of local perceptual features such as edges, lines, or corners. Thus, any change in those perceptual features would contribute to changing the appearance of surfaces. Kovesi's model [30] has been successfully utilized for the detection of edges and localized features in images and has also been shown to perform better than other detectors such as Canny or Prewitt [30,48]. We have therefore employed it in order to investigate how the edge information changes with the varying amount of phase randomization. Figure 8 shows phase congruency maps of the blood texture at different levels of phase randomization. These maps were generated using a MATLAB implementation of phase congruency available at [49] and described in [30].

A. Single Feature Representation

Rather than a phase congruency map, we require a single measure per image as the basis of the perceptual scale. A visual inspection of the maps shown in Fig. 8 suggests that as the level of phase randomization is increased, the edge information is gradually degraded, leading to a noise image when the texture is fully randomized. This appears as a change in the distribution of the edge intensity values as shown in the bottom row in Fig. 8, which suggests that the histogram statistics of the phase congruency maps may provide useful perceptual scales. Figure 9 shows how the mean, variance, skewness, and kurtosis of the phase congruency histograms vary with increasing levels of phase randomization for the blood, RanFrac, and seed textures. While both skewness and variance change considerably with the change in randomization level, the variance is the only feature that behaves in a monotonic way for the three textures investigated. We observe that the behavior of skewness is not monotonic within the 0%–20% range for two of the textures considered (it is monotonic only for the RanFrac texture).

The phase congruency variance showed no significant difference in the behavior across the randomization levels when extracted from naturalistic textures generated using different placement seeds [see Fig. 10(a)]. This measure also converges at 100% randomization for both natural and synthetic textures as shown in Figs. 10(b) and 10(c). Additionally, changes in the variance are greatest in the range 20%–60% of phase randomization. All these observations make the phase congruency variance a suitable measure to represent the psychophysical data. In Figs. 10(b) and 10(c) we observe that although the variance converges for the selected textures, it is different at 0% randomization. This suggests that the current measure of higher order statistics may also account for the amount of structural information across different textures in addition to existing mechanisms that employ first and second order statistics [1,8,11,12,17,50] to do so.

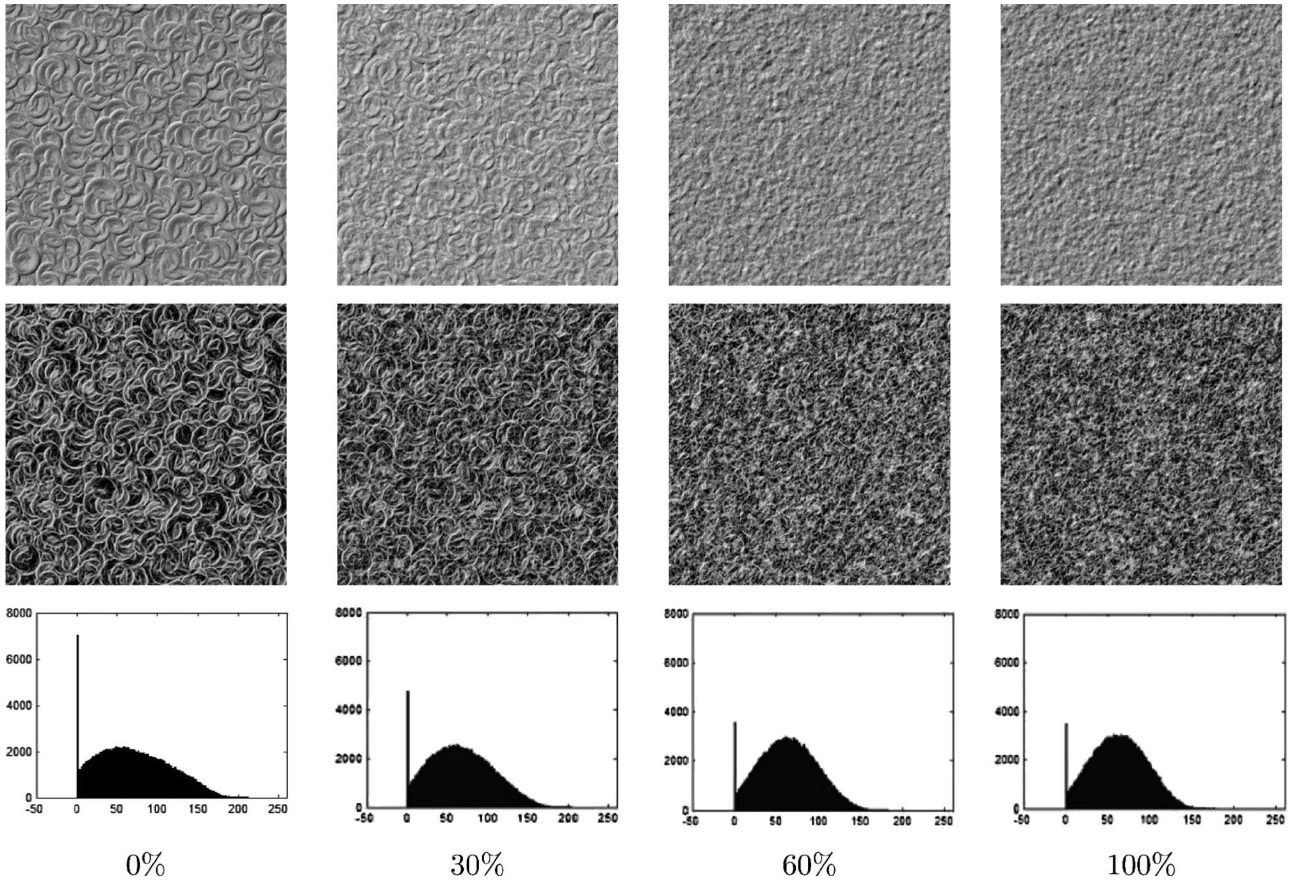


Fig. 8. Phase congruency maps (middle row) for different levels (0%, 30%, 60%, and 100%) of phase randomized blood images (top row) obtained after applying Kovesi's phase congruency model [30]. Bottom row shows how the edge intensity histogram of each map changes shape when the image is randomized.

B. Model

As a model for the computation of the phase congruency variance, we propose a two stage process, with the first stage corresponding to the computation of phase congruency, while in the second stage point-wise nonlinearity and pooling operations are used to estimate the variance of the phase congruency map for each randomized image. Figure 11 illustrates the different steps involved in this two stage model.

The first stage is specified in [30]. It uses a bank of

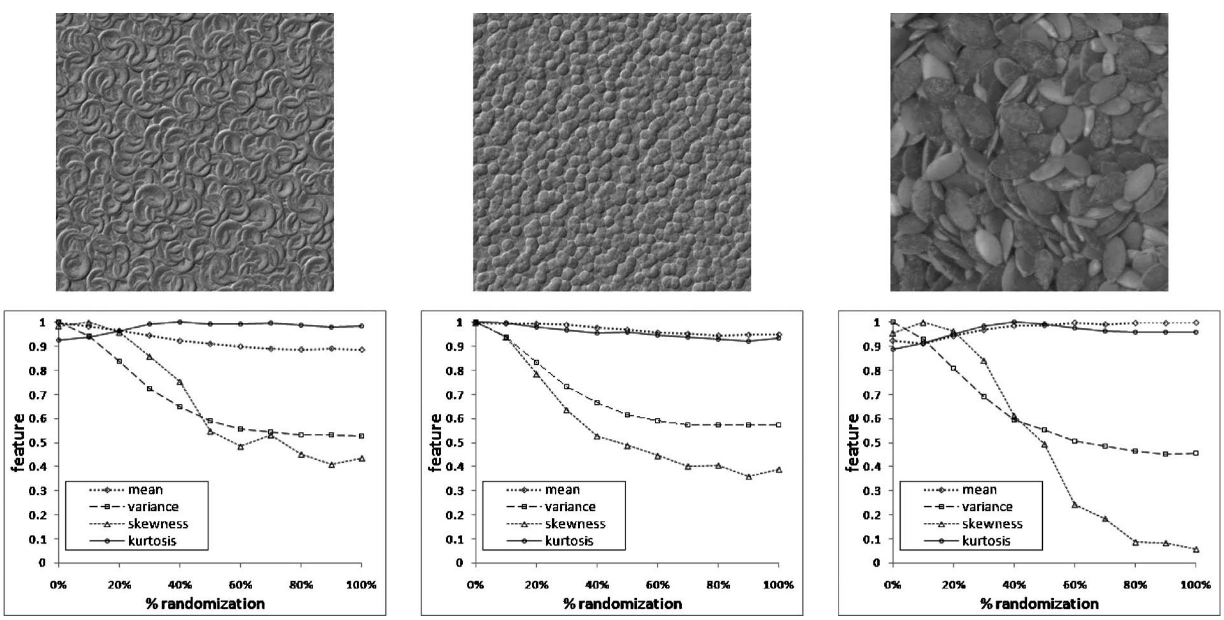
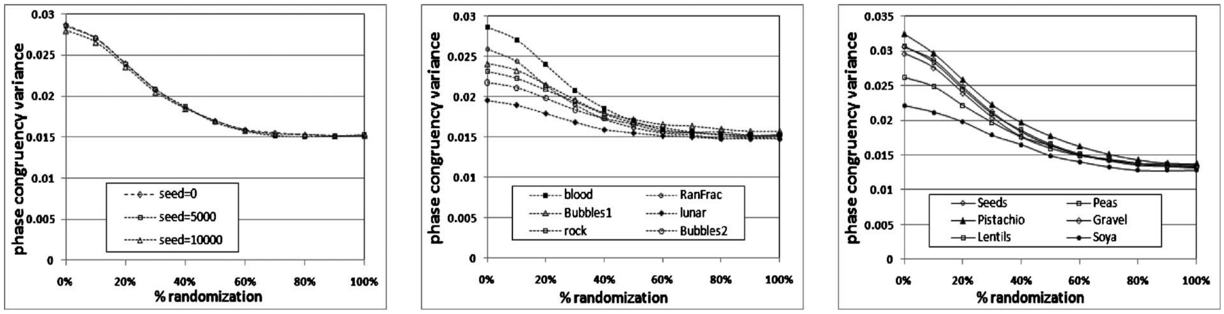


Fig. 9. Behavior of the mean, variance, skewness, and kurtosis of the phase congruency distribution for textures blood, RanFrac, and seeds across the different levels of randomization.



(a) (b) (c)

Fig. 10. Variation in phase congruency variance with changing levels of phase randomization for (a) a texture image generated using random placement of textons at different seeds, (b) six different synthetic textures generated using the same seed, and (c) six different natural textures.

logarithmic-Gabor quadrature filters tuned to different frequencies and orientations so as to capture localized feature information in the texture images. The filters are applied in the Fourier domain and the resulting spatial domain outputs are used to generate the phase congruency maps. The channels for encoding the edges correspond to a sequence of FRFs in the form of filter-rectify-filter-rectify-filter. With FRF layers able to detect only changes in second order statistics, at least one additional nonlinear layer is needed to capture changes in higher order statistics. The first FRF layer allows the generation of the phase congruency map PC, where the phase congruency at each orientation, PC_o, is computed as follows:

$$PC_o(x,y) = \frac{W_o(x)\lfloor E_o(x,y) - T \rfloor}{\sum_n A_{no}(x,y) + \epsilon}, \tag{1}$$

where $W_o(x)$ is a weighting function for the frequency spread at a given orientation o and $A_{no}(x,y)$ is the amplitude information derived using the response of the quadrature filter at each scale n and orientation o. $E_o(x,y)$ is the energy accumulated by the quadrature filters at N scales for a given orientation o and is given by $E_o(x,y)=\Sigma_n A_{no}\Delta\Phi_{no}(x,y)$, where $\Phi_{no}(x,y)$ is the weighted mean phase angle computed at each scale and orientation. T is used for noise compensation and ϵ is a small positive constant that is used when the sum of response vectors is very small leading to an ill-conditioned computation of phase congruency. The phase congruency map PC is obtained by summing the noise compensated energies at all orientations and then normalizing by the sum of amplitudes of the individual quadrature pairs applied at all the scales and orientations.

In the second stage, the phase congruency variance η is computed as follows: $\eta=\frac{1}{D}\Sigma(PC-\mu)^2$. It represents the second nonlinearity (rectify-filter) layer that the visual system uses to perceive differences in the appearance. μ is the mean phase congruency and D represents the number of pixels in the image. Note that although the variance is computed over the whole phase congruency map, it could also be computed over a local window. However, for the purpose of this paper we only require a single measure per image. Fitting the psychophysical data with the computed measure leads to a linear relationship in the

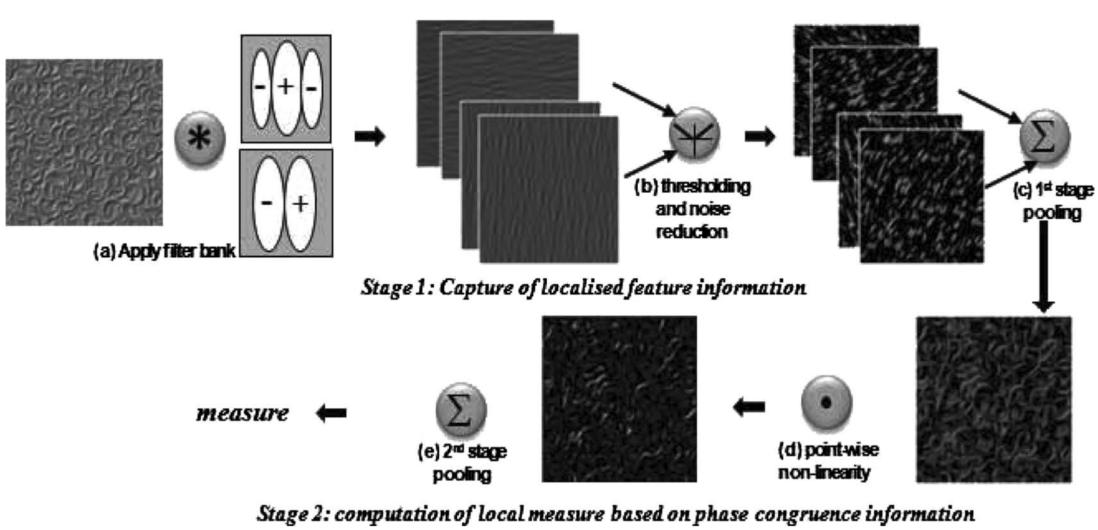


Fig. 11. Model: A two stage process for computing the higher order statistics measure to account for change in appearance.

log-log space. Figure 12 illustrates this relationship. The high correlation (R^2) values indicate excellent fits for both natural [Fig. 12(a)] and synthetic [Fig. 12(b)] textures.

6. DISCUSSION AND CONCLUSION

While it is well documented in the literature that most of the structural information within an image is characterized by its higher order statistics, no studies have so far investigated how well humans perceive small changes in the appearance as a result of changing such statistics. The current paper has addressed this issue by using both naturalistic and natural textures. We randomly perturbed their phase spectra by differing degrees while forcing all randomized images to have identical first and second order statistics as the originals.

Although several studies have investigated the effect of partially randomizing phase spectra on perception, the focus of those studies was on the ability of observers to perform recognition tasks. No quantitative measurements were made of the perceptual differences in the images that resulted. Thomson *et al.* [32] and Hansen and Hess [38] are the only authors to have proposed image metrics that change with the varying amount of phase randomization in their respective studies. However, neither of these studies controlled the first order statistics of their stimuli during the partial phase randomization process.

The experiment presented in this study captured observers' perceptions of changes in the appearance using a set of natural textures and synthetic textures (with naturalistic appearance). A perceptual scale, derived from the resulting psychophysical data, was shown to have sigmoidal monotonically increasing behavior for all images tested. We observed, from the perceptual scales for both natural and synthetic textures, that observers had considerable difficulty in perceiving differences in texture pairs which were phase randomized by less than 20%. For synthetic images, observers had a greater ability to discriminate small changes in the appearance within the 20%–60% range and encountered appreciable difficulty beyond the 60% mark. However, while the perceptual scales for natural images had the same shape, they indicated that observers had the ability to perceive changes in the appearance over a wider range of phase randomization (20%–70%). This may be due to the fact that observers could directly compare the gray levels in one region of a natural image with the same region in its paired image. The use of a randomized placement of texture elements prevented the observers from using the same strategy for the synthetic images.

Although the conditions in which the images were randomized suggest that a change in the appearance of the images may correspond to a change in the visibility of the texture elements, we cannot assume that the observers based their judgments on the perception of the structure. However, we showed (see Fig. 3) that the behavior in the range 20%–60% of randomization corresponds to considerable change in the visibility of the image structure.

We have also proposed an image-based metric that correlates well with the perceived changes. Kovesi's algorithm [30] was used to generate phase congruency maps that reflect the degree of phase congruence in local regions of the image. The algorithm is based on Morrone and Burr's model [29] for the detection of visually salient features in images, which is motivated both by the psychophysical data and by the properties of single cells in the visual cortex. In addition to being biologically motivated, the proposed metric satisfies two additional conditions necessary to account for the perceptual scale derived: (1) it has monotonic sigmoidal behavior and (2) it has the greatest change in gradient when extracted from images that are randomized in the range 20%–60%. It was also shown to correlate linearly with the perceptual difference measurement.

The measure proposed shares a common concept, the frequency channel model, with the SSM proposed by Hansen and Hess [38]. While the SSM exploits the distribution of a set of bandpass filter outputs, the measure we

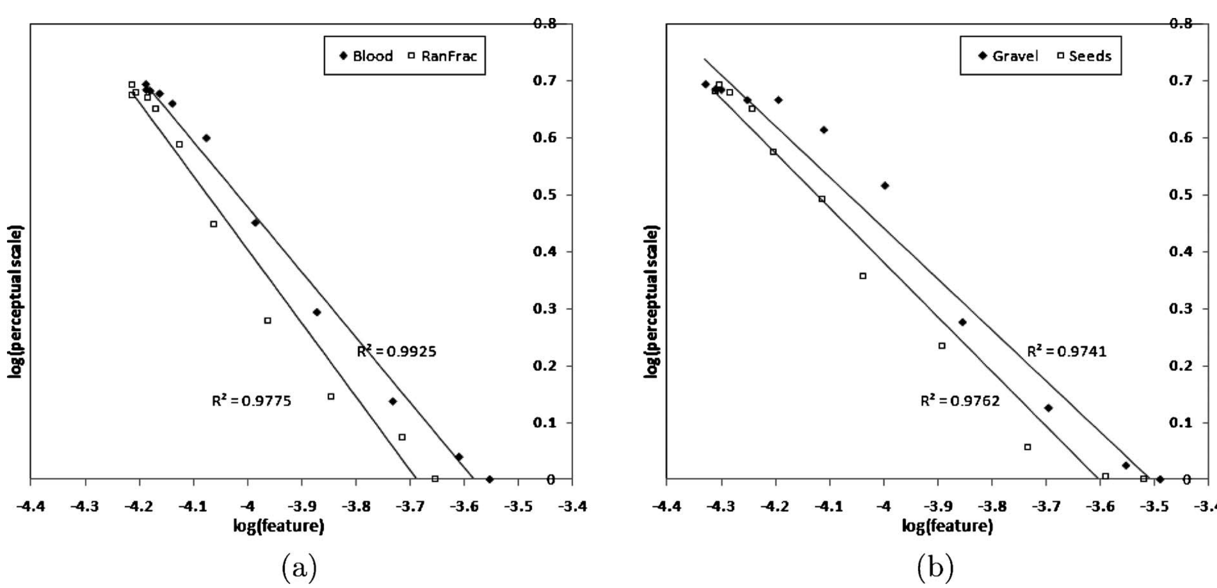


Fig. 12. Linear relationship between perceptual difference and phase congruency variance in a log-log space for (a) synthesized textures blood and RanFrac, and (b) gravel and seeds.

use employs a set of quadrature bandpass filters to compute the phase congruency in the image. We have chosen the latter because spatial phase congruence information has been widely used in the extraction of visually salient features (e.g., edges, bars, and lines) in images [29,30,48,51,52]. The SSM measure, in contrast, has only been applied in the study performed by Hansen and Hess [38].

While previous studies in texture discrimination have focused on the ability of humans to discriminate different categories of textures, the current study has investigated the ability of humans to perceive small changes in the appearance of the same texture. By using textures that differ *solely* in higher order statistics we have demonstrated that humans are very sensitive to the third and higher order statistics that contribute to the change in the appearance. Additionally, we provide a measure to characterize the change in the appearance and show that this measure correlates well with the perceived perceptual difference in textures.

APPENDIX A: MAXIMUM LIKELIHOOD DIFFERENCE SCALING

The MLDS is based on a model of the observer's perception of differences in psychophysical stimuli ordered on a physical scale. Initially, the experimenter selects a set of p stimuli, $\{I_1, I_2, \ldots, I_p\}$, with corresponding values $\{\phi_1 < \phi_2 < \cdots < \phi_p\}$ on the physical scale. On each trial the experimenter presents an observer with quadruples $(I_a, I_b; I_c, I_d)$ and asks him to judge which pair, I_a, I_b or I_c, I_d, exhibits the larger perceptual difference. We replace the notation $(I_a, I_b; I_c, I_d)$ with the simpler notation $(a, b; c, d)$ for convenience. Over the course of the experiment, the observer sees many different quadruples. In past work, experiments have used the set of all possible non-overlapping quadruples $a < b < c < d$ for p stimuli and the resulting scales have proven to be readily interpretable. Moreover, Maloney and Yang [42] reported extensive evaluations of this subset of all possible quadruples.

The data consist of a list of all quadruples presented and the observer's judgments. The goal of the MLDS is to assign values to $\{\psi_1 \leq \psi_2 \leq \cdots \leq \psi_p\}$ that best account for the observer's judgments. Maloney and Yang [42] proposed a stochastic model of difference judgment that allows the observer to exhibit some variation in judgment. Let $L_{ab} = |\psi_b - \psi_a|$ denote the unsigned perceived length of the interval I_a, I_b. The proposed decision model is an equal-variance Gaussian signal detection model [53] where the signal is the difference in the lengths of the intervals,

$$\delta(a,b;c,d) = |\psi_d - \psi_c| - |\psi_b - \psi_a|. \qquad \text{(A1)}$$

If δ is positive, the observer should judge the second interval larger; when negative, the first. We assume that the decision variable employed by the observer is

$$\Delta(a,b;c,d) = \delta(a,b;c,d) + \epsilon = L_{cd} - L_{ab} + \epsilon, \qquad \text{(A2)}$$

where $\epsilon \sim \mathcal{N}(0, \sigma^2)$: given the quadruple, $(a, b; c, d)$, the observer selects the pair I_c, I_d if and only if

$$\Delta(a,b;c,d) > 0. \qquad \text{(A3)}$$

In each experimental condition the observer completes n trials, each based on a quadruple $\mathbf{q}^k = (a^k, b^k; c^k, d^k)$, with $k = 1, n$. The observer's response is coded as $R^k = 0$ (the difference of the first pair is judged larger) or $R^k = 1$ (the second pair is judged larger). We fit the parameters $\mathbf{\Psi} = (\psi_1, \psi_2, \ldots, \psi_p)$ and σ by maximizing the likelihood of the observer's responses,

$$L(\mathbf{\Psi}, \sigma) = \prod_{k=1}^{n} \Phi\left(\frac{\delta(\mathbf{q}^k)}{\sigma}\right)^{1-R_k} \left(1 - \Phi\left(\frac{\delta(\mathbf{q}^k)}{\sigma}\right)\right)^{R_k}, \qquad \text{(A4)}$$

where $\Phi(x)$ denotes the cumulative standard normal distribution and $\delta(\mathbf{q}^k) = \delta(a^k, b^k; c^k, d^k)$ as defined in Eq. (A2).

At first glance, it would appear that the stochastic difference scaling model just presented has $p+1$ free parameters, $\psi_1, \ldots, \psi_p$, together with the standard deviation of the error term, σ. However, any linear transformation of $\psi_1, \ldots, \psi_p$ together with a corresponding scaling by σ^{-1} results in a set of parameters that predict exactly the same performance as the original parameters. Without any loss of generality, we can set $\psi_1 = 0$ and $\psi_p = 1$, leaving us with the $p-1$ free parameters, $\psi_2, \ldots, \psi_{p-1}$, and σ.

Equation (A4) is the likelihood for a Bernoulli random variable. Taking the negative logarithm allows the parameters to be estimated simply with a minimization procedure. We used the package *MLDS* described in [45] to estimate difference scales.

The fitted values $\psi_1, \ldots, \psi_p$ form the difference scale intended to capture human performance. These values can be plotted against the physical values $\{\phi_1 < \phi_2 < \cdots < \phi_p\}$ as a convenient summary of performance. We note that the choice of physical scale is arbitrary and any increasing transformation of the physical scale is a valid physical scale. The difference scale, however, is not arbitrary once its limits are fixed to be 0 and 1. This opens up the possibility of redefining physical scales of roughness or other attributes so that physical spacing better approximates the perceived difference as has been done for loudness by coding physical scale units in decibels.

Maloney and Yang [42] evaluated the distributional robustness of the MLDS. They varied the distributions of the error term ϵ while continuing to fit the data with the constant variance Gaussian error assumption. They found that the MLDS was remarkably resistant to failures of the distributional assumptions. Knoblauch and Maloney [45] also considered the possibility that the observer cannot judge differences in any consistent manner. Such a failure would likely result in a large value of σ relative to the scale limits of 0,1. They also proposed several diagnostic procedures intended to detect failures of the judgment model underlying the MLDS. Such procedures are analogous to testing for the pattern of residual values in the linear regression.

REFERENCES

1. M. S. Landy and N. Graham, "Visual perception of texture," in *The Visual Neurosciences*, L. M. Chalupa and J. S. Werner, eds. (MIT, 2004), pp. 1106–1118.
2. J. Malik and P. Perona, "Preattentive texture discrimination with early vision mechanisms," J. Opt. Soc. Am. A **7**, 923–932 (1990).
3. T. Randen and J. H. Husøy, "Filtering for texture classification: a comparative study," IEEE Trans. Pattern Anal. Mach. Intell. **21**, 291–310 (1999).
4. M. J. Chantler, M. Petrou, A. Penirschke, M. Schmidt, and G. McGunnigle, "Classifying surface texture while simultaneously estimating illumination," Int. J. Comput. Vis. **62**, 83–96 (2005).
5. P. Perona and J. Malik, "Detecting and localizing edges composed of steps, peaks and roofs," in *Proceedings of the ICCV* (1990), pp. 52–57.
6. I. Mareschal and C. L. Baker, "A cortical locus for the processing of contrast-defined contours," Nat. Neurosci. **1**, 150–154 (1998).
7. I. Mareschal and C. L. Baker, "Cortical processing of second-order motion," Visual Neurosci. **16**, 527–540 (1999).
8. A. P. Johnson and C. L. Baker, Jr., "First- and second-order information in natural images: a filter-based approach to image statistics," J. Opt. Soc. Am. A **21**, 913–925 (2004).
9. B. Julesz, "Visual discrimination of textures with identical third-order statistics," Biol. Cybern. **31**, 137–140 (1978).
10. T. Caelli and B. Julesz, "Psychophysical evidence for global feature processing in visual texture discrimination," J. Opt. Soc. Am. **69**, 675–678 (1979).
11. B. Julesz, "Spatial nonlinearities in the instantaneous perception of textures with identical power spectra," Philos. Trans. R. Soc. London, Ser. B **290**, 83–94 (1980).
12. B. Julesz, "A theory of preattentive texture discrimination based on first-order statistics of textons," Biol. Cybern. **41**, 131–138 (1981).
13. J. I. Yellott, Jr., "Implications of triple correlation uniqueness for texture statistics and the Julesz conjecture," J. Opt. Soc. Am. A **10**, 777–793 (1993).
14. C. Chubb, J. Econopouly, and M. S. Landy, "Histogram contrast analysis and the visual segregation of IID textures," J. Opt. Soc. Am. A **11**, 2350–2374 (1994).
15. J. D. Victor, "Images, statistics and textures: implications of triple correlation uniqueness for texture statistics and the Julesz conjecture: comment," J. Opt. Soc. Am. A **11**, 1680–1684 (1994).
16. J. D. Victor and M. M. Conte, "The role of high-order phase correlations in texture processing," Vision Res. **36**, 1615–1631 (1996).
17. C. Chubb and J. I. Yellott, "Every discrete, finite image is uniquely determined by its dipole histogram," Vision Res. **40**, 485–492 (2000).
18. C. W. Tyler, "Theory of texture discrimination of based on higher-order perturbations in individual texture samples," Vision Res. **44**, 2179–2186 (2004).
19. C. W. Tyler, "Beyond fourth-order texture discrimination: generation of extreme-order and statistically-balanced textures," Vision Res. **44**, 2187–2199 (2004).
20. M. O. Franz and B. Schőlkopf, "Implicit Wiener series for higher-order image analysis," in *Advances in Neural Information Processing Systems*, L. K. Saul, Y. Weiss, and L. Bottou, eds. (MIT, 2005), Vol. 17, pp. 465–472.
21. T. E. Hall and G. B. Giannakis, "Bispectral analysis and model validation of texture images," IEEE Trans. Image Process. **4**, 996–1009 (1995).
22. M. K. Tsatsanis and G. B. Giannakis, "Object and texture classification using higher order statistics," IEEE Trans. Pattern Anal. Mach. Intell. **14**, 733–750 (1992).
23. A. V. Oppenheim and J. Lim, "The importance of phase in signals," Proc. IEEE **69**, 529–541(1981).
24. L. N. Piotrowski and F. Campbell, "A demonstration of the visual importance and flexibility of spatial-frequency amplitude and phase," Perception **11**, 337–346 (1982).
25. D. J. Field, "Relations between the statistics of natural images and the response properties of cortical cells," J. Opt. Soc. Am. A **4**, 2379–2394 (1987).
26. D. J. Field, "What the statistics of natural images tell us about visual coding," in *Human Vision, Visual Processing and Digital Display*, Proc. SPIE **1077**, 269–276 (1989).
27. D. C. Knill, D. Field, and D. Kersten, "Human discrimination of fractal images," J. Opt. Soc. Am. A **7**, 1113–1123 (1990).
28. Y. Tadmor and D. J. Tolhurst, "Both the phase and amplitude spectrum may determine the appearance of natural images," Vision Res. **33**, 141–145 (1993).
29. M. C. Morrone and D. C. Burr, "Feature detection in human vision: a phase-dependent energy model," Proc. R. Soc. London, Ser. B **235**, 221–245 (1988).
30. P. Kovesi, "Phase congruency: a low-level image invariant," Psychol. Res. **64**, 136–148 (2000).
31. M. A. Thomson and D. H. Foster, "Role of second- and third-order statistics in the discriminability of natural images," J. Opt. Soc. Am. A **14**, 2081–2090 (1997).
32. M. A. Thomson, D. H. Foster, and R. J. Summers, "Human sensitivity to phase perturbations in natural images: a statistical framework," Perception **29**, 1057–1069 (2000).
33. V. Mante, R. A. Frazor, V. Bonin, W. S. Geisler, and M. Carandini, "Independence of luminance and contrast in natural scenes and in the early visual system," Nat. Neurosci. **8**, 1690–1697 (2005).
34. J. T. Lindgren, J. Hurri, and A. Hyvärinen, "Spatial dependencies between local luminance and contrast in natural images," J. Vision **8**, 1–13 (2008).
35. C. Baker, A. Yoonessi, and E. O. Arsenault, "Texture segmentation in natural images: contribution of higher-order image statistics to psychophysical performance," J. Vision **8**, 350 (2008).
36. W. Einhäuser, U. Rutishauser, E. P. Frady, S. Nadler, P. Kőnig, and C. Koch, "The relation of phase noise and luminance contrast to overt attention in complex visual stimuli," J. Vision **6**, 1148–1158 (2006).
37. F. A. Wichmann, D. I. Braun, and K. R. Gegenfurtner, "Phase noise and the classification of natural images," Vision Res. **46**, 1520–1529 (2006).
38. B. C. Hansen and R. F. Hess, "Structural sparseness and spatial phase alignment in natural scenes," J. Opt. Soc. Am. A **24**, 1873–1885 (2007).
39. A. Yoonessi and F. A. A. Kingdom, "Comparison of sensitivity to color changes in natural and phase-scrambled scenes," Vision Res. **25**, 676–684 (2008).
40. O. R. Joubert, G. A. Rousselet, M. Fabre-Thorpe, and D. Fize, "Rapid visual categorization of natural scene contexts with equalized amplitude spectrum and increasing phase noise," J. Vision **9**, 1–16 (2009).
41. S. Padilla, O. Drbohlav, P. R. Green, A. D. Spence, and M. J. Chantler, "Perceived roughness of $1/f^{beta}$ noise surfaces," Vision Res. **48**, 1791–1797 (2008).
42. L. T. Maloney and J. N. Yang, "Maximum likelihood difference scaling," J. Vision **3**, 573–585 (2003).
43. C. Charrier, L. Maloney, H. Cherifi, and K. Knoblauch, "Maximum likelihood difference scaling of image quality in compression-degraded images," J. Opt. Soc. Am. A **24**, 3418–3426 (2007).
44. G. Obein, K. Knoblauch, and F. Viénot, "Difference scaling of gloss: nonlinearity, binocularity, and constancy," J. Vision **4**, 711–720 (2004).
45. K. Knoblauch and L. T. Maloney, "MLDS: maximum likelihood difference scaling in *R* ," J. Stat. Software **25**, 1–26 (2008).
46. M. C. Morrone and D. C. Burr, "Capture and transparency in coarse quantized images," Vision Res. **37**, 2609–2629 (1997).
47. A. Perna and M. Morrone, "The lowest spatial frequency channel determines brightness perception," Vision Res. **47**, 1282–1291 (2007).
48. P. Kovesi, "Edges are not just steps," in *Proceedings of the ACCV2002* (2002), pp. 822–827.

49. http://www.csse.uwa.edu.au/~pk/Research/MatlabFns/.
50. C. Chubb, M. S. Landy, and J. Econopouly, "A visual mechanism tuned to black," Vision Res. **44**, 3223–3232 (2004).
51. G. S. Hesse and M. A. Georgeson, "Edges and bars: where do people see features in 1-D images?" Vision Res. **45**, 507–525 (2005).
52. E. J. Kelman, R. J. Baddeley, A. J. Shohet, and D. Osorio, "Perception of visual texture and the expression of disruptive camouflage by the cuttlefish, sepia officinalis," Proc. R. Soc. London, Ser. B **274**, 1369–1375 (2007).
53. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Wiley, 1974).