Human sensitivity to phase perturbations in natural images: a statistical framework

Mitchell G A Thomson¶, David H Foster#, Robert J Summers¶
Neurosciences Research Institute, Aston University, Birmingham B4 7ET, UK;
e-mail: thomsmga@aston.ac.uk; # Department of Optometry and Neuroscience, University
of Manchester Institute of Science and Technology, PO Box 88, Manchester M60 1QD, UK
Received 11 November 1998, in revised form 24 November 1999

Abstract. Fourier-phase information is important in determining the appearance of natural scenes, but the structure of natural-image phase spectra is highly complex and difficult to relate directly to human perceptual processes. This problem is addressed by extending previous investigations of human visual sensitivity to the randomisation and quantisation of Fourier phase in natural images. The salience of the image changes induced by these physical processes is shown to depend critically on the nature of the original phase spectrum of each image, and the processes of randomisation and quantisation are shown to be perceptually equivalent provided that they shift image phase components by the same average amount. These results are explained by assuming that the visual system is sensitive to those phase-domain image changes which also alter certain global higher-order image statistics. This assumption may be used to place constraints on the likely nature of cortical processing: mechanisms which correlate the outputs of a bank of relative-phase-sensitive units are found to be consistent with the patterns of sensitivity reported here.

1 Introduction

The ecological approach to visual coding is based upon the premise that physical characterisations of the natural environment can help place in context the aims of, and constraints upon, human visual processing. One aim of the ecological approach is to better understand the role of spatial image redundancies in visual coding; much of the research in this area has made use of the global Fourier transform as a tool for both analysing (eg Field 1987) and manipulating (eg Knill et al 1990; Tadmor and Tolhurst 1994; Thomson and Foster 1997) statistical image properties. In particular, those second-order scene statistics quantified by the image power spectrum appear broadly consistent from image to image, and this analytical result has been related to the properties of visual cortical cells (Field 1987). Simple Fourier-domain image-processing demonstrations (eg Oppenheim and Lim 1981; Piotrowski and Campbell 1982) imply, however, that it is image *phase* spectra which convey the majority of the information used by organisms to discriminate different scenes. This both motivates and complicates an extension of the ecological approach: those scene statistics captured by the phase spectrum do show some consistencies from image to image (Thomson 1999a, 1999b), and these consistencies might be expected to have greater import for sensory coding than the power-spectral consistencies (we must surely discriminate different scenes efficiently), yet if the phase spectrum is also the chief determinant of image-specific structure, it will not be so easy to relate those consistencies to the processing carried out by the human visual system.

Given these difficulties, a study which aims to provide a better understanding of human sensitivity to phase-related structure in natural scenes would be well advised to follow a rather parsimonious approach. The key goals of the present study are (i) to define the simplest global image statistics that are potentially sensitive to changes in Fourier phase spectra of images, and (ii) to assess the ability of these statistics

¶ Current address: Colour & Imaging Institute, University of Derby, Kingsway House East, Derby DE22 3HL, UK; e-mail: m.g.a.thomson@derby.ac.uk

to account for human visual sensitivity to phase perturbations in natural images. As in previous studies, the global Fourier transform is used to perform the phase perturbations (Oppenheim and Lim 1981; Piotrowski and Campbell 1982); the class of statistical image measures investigated here, however, is defined not in the Fourier domain but in higher-order statistical domains (Nikias and Petropulu 1993). Human visual sensitivity to image phase perturbations is assessed partly by means of some new suprathreshold image-processing demonstrations and partly through a psychophysical image-discrimination experiment. One measure, based on the normalised fourth image moment or *kurtosis*, accounts for human performance particularly well; when this finding is used to place constraints on the mechanisms underlying early cortical processing, it is shown to be consistent with a system which processes image structure by correlating the outputs of relative-phase-sensitive cells.

2 Previous work on phase manipulations

The techniques of phase randomisation and phase quantisation have been used in several studies to perturb localised scene structure across the whole region of support of an image. For example, it is well known that natural images whose phase spectra have been completely randomised convey very little local structural information, and human observers rapidly and consistently categorise such images as 'unnatural' (Field 1987). Julesz and Schumer (1981) claimed that the slightest scrambling of the Fourier phase spectrum renders an image unrecognisable, and Oppenheim and Lim (1981) even asserted that no variation in the amplitude spectrum of a pattern is needed to recognise it as long as the phase spectrum is adequately coded. More recent studies (Juwells et al 1991; Tadmor and Tolhurst 1993; Lohmann et al 1997), however, imply that the relative importance of Fourier amplitude and phase as determinants of perceived image structure appears to depend on the exact configuration of the phase spectrum of the image. Piotrowski and Campbell (1982) showed that natural images could still be recognised after severe quantisation of their Fourier phase spectra, but the differential effects of phase quantisation and phase randomisation are still unclear. In the most systematic study to date, Hübner et al (1988) determined detection thresholds for the effects of quantisation and randomisation on random grayscale textures; their results showed that the detectability varied monotonically with increased randomisation or quantisation, and that the thresholds obtained with the use of these two techniques were similar provided that both techniques shifted the phase components by the same average amount (cf Piotrowski and Campbell 1982).

This equivalence of randomisation and quantisation is particularly interesting, since one might intuit that quantising phase values could introduce image structure, whereas randomising phase values might be expected to destroy image structure. Is this finding true for natural images as well as for random textures? Do the visual effects of randomisation and quantisation depend critically on the nature of each individual image? The following section describes some new image-processing demonstrations designed to answer these questions and to test some of the findings reported in the studies cited above.

3 Demonstrations
3.1 *Methods*
The data set used to produce the demonstrations was an ensemble of eighty-five calibrated, digitised natural scenes of linear size $N = 512$ pixels; the images and calibration procedures have been described elsewhere (Thomson and Foster 1997). Phase perturbations, carried out on a Sun IPX SPARCstation (Sun Microsystems, USA), were applied by Fourier-transforming each image; computing its power and phase spectra as a function of Cartesian spatial frequencies (u, v); altering the values in the phase

spectrum $\phi(u, v)$; and inverse-Fourier-transforming the resulting spectra to produce a new image. A second, much larger image ensemble was generated by replacing the natural phase spectrum of each image with random deviates drawn from a uniform distribution over the interval $[-\pi, +\pi)$; this procedure was performed 100 times for each image, with a different random seed each time, producing a total of 8500 *random-phase* images. All random deviates were generated with a linear-congruential pseudo-random-number generator with a very long cycle (Press et al 1988).

Images in either ensemble could be phase-perturbed in one of two ways, as follows. *Phase quantisation* restricts the gamut of image phase states to a smaller number of permissible levels. An *n*-level phase quantisation divides the range of the principal phase spectrum $[-\pi, +\pi)$ into *n* equal intervals of size $2\pi/n$; all Fourier components whose phase values fall anywhere within a given interval are then shifted to a new phase value defined by the midpoint of that interval. For the purposes of the present study, *n* was restricted to the range 3–16; 2-level quantisation introduces symmetries into images, caused by a phase-orientation ambiguity (Piotrowski and Campbell 1982), and quantisations finer than 16-level appear to have no visual effect (Hübner et al 1988). Following Hübner et al (1988), *phase randomisation* is defined here by equivalence with phase quantisation; before *n*-level randomisation can be performed, one must first perform an *n*-level quantisation of the original image, then, using equation (1), compute the mean absolute phase difference (MAPD) of the phase values $\phi_o(u, v)$ present in the original image and the corresponding values $\phi_f(u, v)$ present in the phase-quantised image:

$$\text{MAPD} = \frac{1}{N^2} \sum_{u=1}^{N} \sum_{v=1}^{N} \left[|\phi_o(u, v) - \phi_f(u, v)| \bmod 2\pi \right]. \tag{1}$$

Randomisation then proceeds as follows: add to the value of each Fourier component in the original image a random deviate drawn from a uniform distribution over the interval $[-J, +J)$, with J chosen such that the MAPD of the original and randomised images is the same as that of the original and quantised images.

The demonstrations require a suprathreshold psychophysical paradigm capable of revealing which image phase perturbations have perceptual significance and which do not. The paradigm adopted here is similar to that used by Julesz and coworkers in their investigations of texture processing (eg Julesz et al 1973). If X is the original image, drawn from either the natural- or random-phase ensemble, and Y is the same image after phase perturbation, a composite image Z may be constructed by combining the pixel values of X with the corresponding pixel values of Y in proportions defined by a 2-D Gaussian function g centred over the image and with standard deviation of $N/4$, thus: $Z = gX + (1 - g)Y$. This partitioning process creates a central area, derived mainly from image X, and a surround area, derived mainly from image Y, without producing edge artifacts at the centre–surround interface. One may reasonably assume that if a composite can be segmented instantly and effortlessly by a human observer, this is because the underlying *statistics* of images X and Y are sufficiently different— the observer does not have access to both original images in their entirety, and so cannot perform a simple discrimination task that might be based on, for example, local luminance differences. Moreover, in the demonstrations shown here, images X and Y always had identical power spectra, so the statistical differences used by the observer to perform the segmentations must be due to differences in the phase spectra of those images.

3.2 Results

Figure 1 shows nine composites generated with the paradigm described above. Composites A and B each show a natural image (surround) combined with its random-phase counterpart (centre); the salience of the statistical differences between natural-phase and random-phase image data gives a 'fuzzy' or 'noisy' percept in the centre of the stimulus. Composite C is, in effect, a control experiment, constructed by combining two different random-phase versions (generated with different random seeds) of the same natural scene: the composite is not segmentable. Composites D and E each show a natural image (surround) combined with a 4-level phase-randomised version of the same image (centre): Composite D is not segmentable, yet composite E is, so the effects of phase randomisation do indeed depend on the nature of the original image. Composite F, which is not segmentable, shows one of the random-phase images (surround) combined with the same image after 3-level phase quantisation (centre): even severe quantisation appears to introduce no new structure into random-phase images. None of the composites in the bottom row (G, H, I) is segmentable, yet all were created by combining a 3-level phase-randomised version (surround) with a 3-level phase-quantised version (centre) of the same image: when the processes of quantisation and randomisation are normalised for MAPD, the resulting visual effects do indeed appear perceptually equivalent.
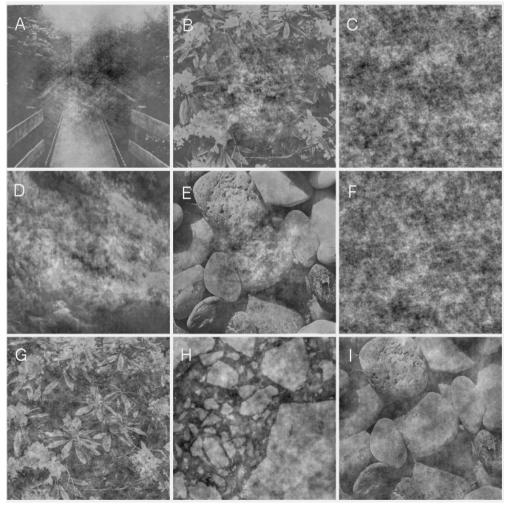

Figure 1. A montage of nine composite images; see text for description.

These observations may be combined with the results of the studies cited above in order to summarise the perceptual significance of statistical image changes induced by phase-perturbing both natural and random-phase images. For natural images, increasing the severity of phase perturbations generally increases the salience of the resulting visual effects, and natural images are always readily distinguished from their random-phase counterparts. Human visual sensitivity to intermediate levels of phase perturbation depends strongly on the nature of individual images, but not on the exact perturbation method (provided that the MAPD normalisation is used). For random-phase images, changes in the precise phase configuration induced by re-randomising the phase spectra of these images do not generate cues for image segmentation; this finding is not trivial, since one could easily prescribe an image *power* spectrum that would render such changes perceptually significant.

4 A statistical framework

This section is directed towards the first goal of the present study: to define the simplest global image statistics that are potentially sensitive to changes effected in image phase spectra. The statistics considered here are conventional nth-order statistics, not those studied by Julesz and colleagues [see Klein and Tyler (1986) for a discussion of the advantages of conventional over Julesz statistics]. Given this, global second-order statistics cannot possibly account for the visual effects of image phase perturbation, since the phase-perturbation procedures leave the power spectra of the images unchanged. How, then, should one proceed within the statistical framework? Perhaps the most obvious approach would be to first compute multiple, local first-order statistics, then consider how these are to be combined to produce global sensitivity; such an approach is in direct accord with existing anatomical and physiological evidence, which supports the notion that early visual processing is subserved by multiple local analysers that feed into 'collator units'. Although this agreement with experimental findings is encouraging, a number of theoretical issues make it difficult to pursue this 'bottom – up', local-analyser approach systematically outside the linear framework; for example, the well-known problem of disconfounding the precise sensitivities of the local analysers from the details of how their outputs are to be combined. An alternative strategy, then, is to compute simple, global higher-order statistics (that is, statistics of order 3 or higher) on the images themselves in an attempt to predict human sensitivity to phase-related image changes. Such a strategy is pursued here, not because we believe the visual system computes higher-order statistics, but because it allows constraints to be placed more directly on the way in which the outputs of visual analysers are combined. The two approaches are in fact complementary rather than mutually contradictory (indeed, there would be little point in following an approach which could not be related ultimately to the physiology), since many multiple local-analyser arrangements can be recast as a single mechanism whose sensitivity to the higher-order statistics of the visual input is easily defined.

Which global higher-order statistics should be considered? For any order n, the simplest possible nth-order statistic of a signal is the nth moment, ie the expected value of the nth power of the signal. The third and fourth moments of an image, then, warrant consideration on grounds of parsimony—they are the simplest global odd-order and even-order statistics that are potentially sensitive to image phase changes. Indeed, Klein and Tyler (1986) found that these measures could account satisfactorily for the results of a variety of psychophysical phase-discrimination experiments. The simplicity of these higher-order moments is, however, rather deceptive; they are not *pure* third-order and fourth-order statistical measures, since they generally also depend on second-order signal statistics. It is usually assumed that this dependence can be removed by normalising with the appropriate power of the variance; this produces from the third moment a measure termed *skewness*, and from the fourth moment a measure

termed *kurtosis*[1]. This normalisation does not guarantee, however, that the skewness and kurtosis measures will be independent of the *form* of the power spectrum: to do so requires that the power spectrum of the signal be completely flattened before these measures are computed, a process sometimes termed *whitening*. When this procedure is followed for natural-image data, the resulting measures—which may be termed phase-only skewness and phase-only kurtosis—do display certain consistencies across image ensembles; for example, the phase-only kurtosis of natural images is strictly positive, whereas the phase-only kurtosis of random-phase images is close to zero (Thomson 1999a).

5 Simulations

5.1 Methods

Can either of these higher-order statistics, with or without the modification of spectral whitening, account for the visual effects of phase-perturbing natural images? One simple and robust finding, reported in many of the phase-perturbation studies cited earlier, and illustrated here in figure 1, is that human observers do not confuse natural images with random-phase images. Figure 1 also shows that re-randomising random-phase images does not produce visually salient statistical effects. Taken together, these two observations effectively define bounds on human sensitivity to image-phase perturbations; an important preliminary test of the behaviour of the higher-order image statistics involves determining whether these statistics predict those bounds, as follows. Consider a difference measure which computes the magnitude of the changes induced in a given higher-order statistic as a result of complete phase randomisation of an image: if such a measure were computed for (a) the natural-image ensemble and (b) the random-phase ensemble, the distributions over (a) and (b) should not overlap—in fact one would expect the distribution over (b) to be centred around some very small value.[2]

These difference measures may now be defined formally as follows, given an original image I_o and a phase-perturbed image I_f, both constrained to be zero-mean, as functions of x and y:

- *Skewness difference*, Δs:

$$\frac{\frac{1}{N^2}\left| \sum_{x=1}^{N}\sum_{y=1}^{N}[I_f(x, y)^3] - \sum_{x=1}^{N}\sum_{y=1}^{N}[I_o(x, y)^3] \right|}{\left\{ \frac{1}{N^2}\sum_{x=1}^{N}\sum_{y=1}^{N}[I_o(x, y)^2] \right\}^{3/2}} ; \tag{2}$$

- *Phase-only skewness difference*, Δs_W: computed as in equation (2), but both original and phase-perturbed images have flat power spectra;
- *Kurtosis differences*, Δk:

$$\frac{\frac{1}{N^2}\left| \sum_{x=1}^{N}\sum_{y=1}^{N}[I_f(x, y)^4] - \sum_{x=1}^{N}\sum_{y=1}^{N}[I_o(x, y)^4] \right|}{\left\{ \frac{1}{N^2}\sum_{x=1}^{N}\sum_{y=1}^{N}[I_o(x, y)^2] \right\}^{2}} ; \tag{3}$$

- *Phase-only kurtosis differences*, Δk_W: computed as in equation (3), but both original and phase-perturbed images have flat power spectra.

---

[1] In the most common definition of kurtosis, the value 3 is subtracted from the normalised fourth moment in order to ensure that the kurtosis of a Gaussian random process is zero. This distinction is irrelevant to the measures described in this section, since they are all *difference* measures.

[2] In theory, phase-only kurtosis or skewness differences computed from a process with Gaussian phase statistics are zero, but the images described here are finite 2-D samples, and so some residual higher-order correlations would be expected.

For the sake of completeness, we also consider two first-order difference measures described by Hübner et al (1988), both of which were found to model human sensitivity to phase perturbations in grayscale textures rather well. These are the MAPD measure set out in equation (1), and a 'bit-metric' measure which quantifies local luminance differences between original and phase-perturbed images (for the purposes of computing this measure, I_f and I_o are constrained to be strictly positive):

$$\frac{1}{N^2} \sum_{x=1}^{N} \sum_{y=1}^{N} \frac{|I_f(x, y) - I_o(x, y)|}{I_o(x, y)} .$$

Like the skewness and kurtosis measures, this bit metric can also be made into a phase-only measure by computing it on spectrally whitened images.

5.2 Results

The histograms shown in figures 2a (Δs), 2b (Δk), 3a (MAPD), and 3b (bit-metric) represent the distributions of these difference measures across the two image ensembles
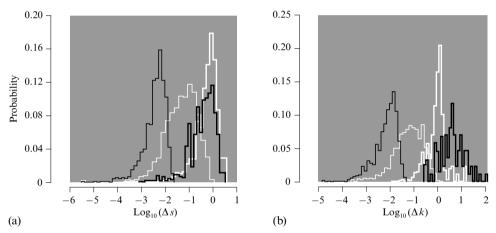

Figure 2. Histograms showing the distribution of (a) the skewness difference and (b) the kurtosis difference measures across the natural-image ensemble (thick lines) and the random-phase ensemble (thin lines). Dark-coloured lines indicate the phase-only measures, Δs_W in (a) and Δk_W in (b), and light-coloured lines indicate the non-phase-only measures, Δs in (a) and Δk in (b).
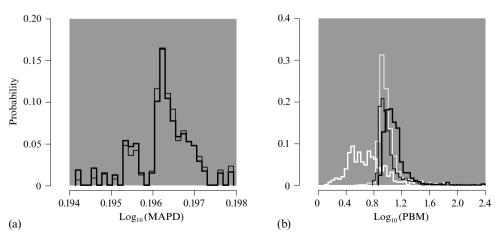

Figure 3. Histograms showing the distribution of (a) the MAPD and (b) the bit-metric measure (PBM) across the natural-image ensemble (thick lines) and the random-phase ensemble (thin lines). In (b) dark-coloured lines indicate the phase-only bit-metric measures and light-coloured lines indicate the non-phase-only bit-metric measure.

described above; the abscissas are plotted on a logarithmic scale, purely because the range of values is generally very large. Dark-coloured lines indicate the phase-only measures and light-coloured lines indicate the non-phase-only measures; thick lines show the distribution for the natural-phase images, and thin lines show the distribution over the random-phase images.

Only two of the difference measures, Δs_{W} and Δk_{W}, satisfy the requirement that there should be no overlap between the distribution over the natural images (thick lines) and the distribution over the random-phase images (thin lines); indeed, the peaks of these two distributions are several orders of magnitude apart in both cases.

These findings prompted a more qualitative investigation of the behaviour of the Δs_{W} and Δk_{W} measures under increasingly severe phase perturbation, as follows. Each image in the natural and random-phase image libraries was subjected to n-level quantisation with $n = 3, 4, 5, 6, 7, 8, 10, 12, 14, 16$. The Δs_{W} and Δk_{W} measures were computed from equations (2) and (3), respectively, this time comparing the phase-only skewness and phase-only kurtosis of the original images with those of the phase-perturbed images; the results are plotted as a function of n in figures 4a (Δs_{W}) and 4b (Δk_{W}). For the purposes of comparison, two more simulations were performed. First, the random-phase ensemble was used instead of the natural-image ensemble; the corresponding data are also shown in figures 4a and 4b. Second, the technique of phase randomisation, instead of phase quantisation, was used to phase-perturb images in the natural-image ensemble: the top part of figure 4a shows the ratio (Δs_{W} under randomisation)/(Δs_{W} under quantisation) as a function of n, and the top part of figure 4b shows the corresponding data for Δk_{W}.

Figures 4a and 4b are rather similar (although notice in figure 4a that the error bars for the natural-image data are rather large, and that the randomisation/quantisation ratio is not always exactly one): as n decreases, randomisation and quantisation
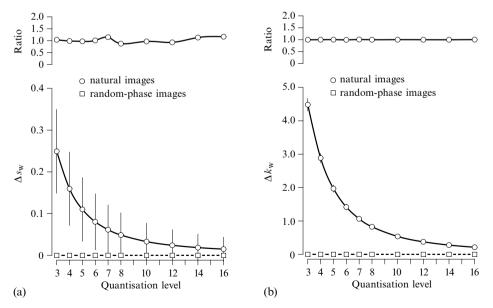

Figure 4. (Bottom) (a) Δs_{W} and (b) Δk_{W} plotted as functions of n under phase-quantisation of natural images (solid lines) and random-phase images (dotted lines): vertical bars show ± 1 SEM of the distribution of Δs_{W} and Δk_{W} over all images in the ensemble. (Top) the ratio (Δs_{W} under randomisation)/(Δs_{W} under quantisation) plotted as a function of n in (a) and the corresponding ratio for Δk_{W} plotted in (b).

produce equivalent monotonic increases in Δs_{W} and Δk_{W} measured from natural images, but essentially no change in Δs_{W} and Δk_{W} measured from random-phase images. Thus far, then, the effects of phase perturbation on the behaviour of Δs_{W} and Δk_{W} is consistent with the associated visual effects. To test the performance of these simple statistical measures still further, a psychophysical image-discrimination experiment was devised, and this is described in the following section.
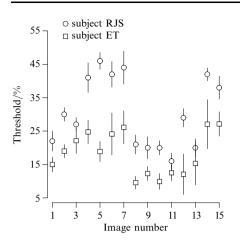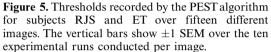
6 Experiment
6.1 *Methods*
The experiment was designed to investigate further the image-dependence of human visual sensitivity to phase perturbations. One obvious threshold psychophysical paradigm would require observers to discriminate a slightly phase-perturbed image from a natural image, but under these circumstances observers' sensitivities might be determined by one particular 'feature' in the natural images, ie they may not perform the tasks statistically. To avoid this, observers were instead required to discriminate a partly phase-randomised image from two fully phase-randomised images with the same power spectrum. Partly phase-randomised images were generated by adding a random deviate drawn from a uniform distribution over the interval $[-J, +J)$ to each phase value in the original image; for the fully randomised images, the value of J was π. The stimulus parameter r was the *difference* in randomisation between the partly and fully randomised images, expressed as a percentage of full phase randomisation: $r = 100(\pi - J)/\pi$ %. A 3-alternative forced-choice paradigm was used, and image presentation was temporal-sequential: on each trial an observer was presented with a stimulus; an interstimulus field; a second stimulus; an interstimulus field; a third stimulus; and a poststimulus field. Each stimulus consisted of a single image presented in the centre of the visual field on a uniform gray background. The observer was instructed to choose which one of the three stimuli was the odd one out. One of the three stimuli contained the partly phase-randomised image and the other two contained fully phase-randomised images; the (temporal) order of these three stimuli was random. All three images had identical power spectra but were phase-randomised with the use of different random seeds. The interstimulus and poststimulus fields consisted of uniform gray backgrounds. The presentation time for all stimuli and all intervals was 300 ms each; observers used a three-button mouse to signal their responses. Stimuli were presented on a high-resolution monitor (Sony Japan Trinitron model) driven by a Silicon Graphics O2 computer system (SGI, UK). The mid-gray of the monitor (gray-level 128) had a luminance of 30 cd m^{-2}; with the viewing distance set at 1.5 m, the available spatial-frequency bandwidth was $0.125 - 45$ cycles deg^{-1}. In a single experimental run, the PEST algorithm (Taylor and Creelman 1967) was used to determine a single threshold. The algorithm was initialised with a stimulus parameter of 50% and a step size of 10%; it terminated when the step size was less than 2.5%, and the magnitude of the step size could never exceed 20%. Ten estimates of observer threshold were determined for each of fifteen different images; these fifteen images were selected at random from the natural-image ensemble. There were thus 150 runs in an entire experiment, and these runs were conducted in random order. There were two observers, RJS and ET, both of whom had corrected-to-normal visual acuity (6/6 Snellen acuity). Observer ET was unaware of the purposes of the experiment; RJS is a coauthor.

6.2 *Results*
The results of the experiment are plotted in figure 5, which shows the thresholds recorded by the PEST algorithm for both subjects over the 15 different images.

Notice that the observers' thresholds vary greatly from image to image, from around 10% to 45%. The two subjects' threshold levels are quite different—observer ET

Figure 5. Thresholds recorded by the PEST algorithm for subjects RJS and ET over fifteen different images. The vertical bars show ± 1 SEM over the ten experimental runs conducted per image.

is around $10\% - 15\%$ more sensitive than observer RJS—but the trends as a function of image number do show similarities.

In an attempt to determine whether the s_W and k_W measures discussed above could account for human performance, these measures were used to construct model observers, as follows. First, each image used in the experiment was subjected to progressive phase randomisation, ranging from no randomisation at all ($r = 100\%$) to complete phase randomisation ($r = 0\%$) in steps of 5%. When Δs_W and Δk_W were plotted as functions of r, the data were found to be almost perfectly fitted by logistic functions, ie curves of the form Δs_W or $\Delta k_W = c/[1 + \exp(a - br)]$, where a, b, c are the parameters determined by the fitting algorithm. Fitting these data in this way allowed us to compute the value of r for any given value of Δs_W and Δk_W; in fact, by repeating the entire process 100 times for 100 different random seeds, we were also able to estimate errors on r due to the variation in random seed. All that remained, then, was to set a threshold level of Δs_W or Δk_W at which the model observers would judge a partly phase-perturbed image to be discriminable from fully-phase-randomised images. The human observers' threshold data shown in figure 5 were used to do this, as follows. The fitted curves described above were used to reparameterise the human observers' thresholds in Δs_W and Δk_W rather than in r; these thresholds were then averaged across all images to yield a single threshold value; this threshold was then used as the model observer's threshold. This procedure was performed separately for each of the two human observers because of the significant difference in their overall threshold levels. The thresholds set by the two human observers are plotted in figures 6a (Δs_W) and 6b (Δk_W) against the corresponding thresholds predicted by the model observers.

Since the mean threshold of a given model observer over all images has effectively been fixed to agree with the mean threshold of the corresponding human observer, what is of interest here is whether the model can predict the significant image-by-image variation in observer thresholds. Of the two model observers, the one based on phase-only kurtosis (figure 6b) predicts the observers' thresholds rather well; the data lie close to a straight line of unit gradient and the correlation coefficient is 0.93. Notice also that, in both figures, the human-observer and model-observer error estimates are of similar magnitude, implying that the uncertainty in the human observers' thresholds is explained mainly by the variations in the random seeds.
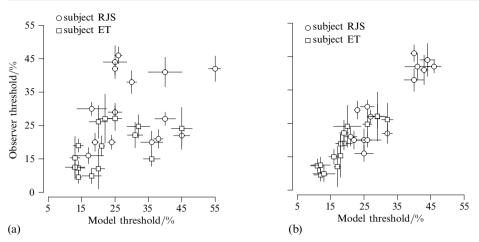
Figure 6. Thresholds set by the two human observers plotted against the corresponding thresholds predicted by a model observer that computes (a) Δs_W and (b) Δk_W. The human-observer data are those plotted in figure 5; the horizontal bars show ± 1 SEM of the distribution of (a) Δs_W and (b) Δk_W over the 100 random seeds.

7 Discussion

The present work has introduced global statistics that are—under the conventional definitions—the simplest global statistics that could potentially account for image changes induced by phase perturbation. It has shown that measures based on image skewness and kurtosis can, when appropriately normalised to remove the effects of second-order image structure, account for the observed patterns of human visual sensitivity to the phase-domain changes described here; in addition, the phase-only kurtosis measure can even account for the strong image dependence of the visual effects of phase perturbation.

Of course, such simple measures could never be expected to provide a complete description of human visual sensitivity to phase perturbations—we know there are ways of processing images which will change phase-only skewness or kurtosis radically without inducing visually salient image changes (and the converse is probably also true)—and it may be that the behaviour of these statistical measures in this respect could be improved easily through, for example, the introduction of spatial-frequency selectivity. As stated earlier, however, it is not the aim of the present study to argue that the visual system computes higher-order statistics. The rationale for the higher-order-statistical approach is instead as follows: now that we know the human visual system shows similar patterns of sensitivity under phase-related image changes to the trends observed in analytically defined higher-order statistics, this similarity can be used to place constraints simultaneously on (a) what type of analysers might be involved, and (b) how the outputs of those analysers might be combined. To discover what those constraints are, we need to define the class of visual mechanisms that are sensitive to specifically those phase-domain image changes that also affect the phase-only image kurtosis.

For simplicity's sake, the following part of the discussion is restricted to the phase-only kurtosis statistic, but arguments analogous to that set out below could be constructed for any phase-only higher-order statistic. Fourth-order statistics are the lowest even-higher-order statistics, and even-higher-order statistics are easier to relate to the properties of neural mechanisms (Thomson 1999a) than odd-higher-order statistics.

Recall that natural images always have positive phase-only kurtosis, whereas random-phase images have phase-only kurtosis values close to zero. What sort of consistencies in the phase spectra of natural images could give rise to this positive phase-only

kurtosis? The answer cannot lie in the *absolute* Fourier phases; these would be altered by a simple translation, under which all higher-order moments are (by definition) invariant. The key is to consider the properties of the phase-only kurtosis measure not in the space domain, but in the Fourier domain. Since image kurtosis is a fourth-order statistic, it must be formed from the sum of fourfold combinations of image data; this is true not only in the spatial domain, but also in the Fourier domain, where kurtosis is defined as the integral of fourfold combinations of frequency components (Nikias and Petropulu 1993). The space of all these fourfold combinations is very large; one way of simplifying it is to consider each combination of four frequency components as two pairs of frequency components. This allows one to define the *relative phase* of each pair as the difference in the absolute phases of its two members; and it follows from the (frequency-domain) definition of kurtosis that positive phase-only kurtosis will arise if the relative phase of one pair is, on average, positively correlated with the relative phase of the other pair (Thomson 1999a). This phenomenon—positive correlations between relative spatial phases—is indeed a property of natural images (Field 1994; Summers and Thomson 1999), and this simple model of fourth-order image structure suggests one way of constructing a simple encoder that would be sensitive to changes in phase-only image kurtosis: operate a mechanism which correlates the outputs of a bank of relative-phase-sensitive analysers.

As a property of visual cortical cells, sensitivity to relative rather than absolute spatial phase is not a new concept; in fact, it accords well with existing physiological evidence (eg Victor and Conte 1996). These relative-phase analysers are almost certainly not the 'front end' of the visual cortex; anatomical and physiological studies lead us to believe that they are themselves formed from the outputs of translation-sensitive units. Two different stages of the mechanism predicted here could thus be reconciled with the notions of local-contrast (Tadmor and Tolhurst 1994) detectors (the translation-sensitive units) and local-energy (Morrone and Burr 1988) detectors (the relative-phase analysers), but it is the combinatorial aspects of the mechanism which are of greater interest here—not just the implied existence of correlator units, an idea which has received some attention in the literature (see, for example, Field 1994), but also the notion that some sort of spectral whitening might be necessary for the correlators to function efficiently. A recent computational study (Summers and Thomson 1999) suggests that the limited spectral whitening achieved by the retina (Atick and Redlich 1992) is probably inadequate in this respect; far more important, probably, is the 'effective whitening' achieved by having spatial-frequency-selective cortical cells with approximately octave-wide spatial-frequency bandwidths (Field 1987).

Finally, notice that an analyser sensitive to relative spatial phase, although translation-insensitive, has a fundamentally local sensitivity (Victor and Conte 1996); the fact that local sensitivity is predicted by a global statistical approach refutes the notion that global image models are only useful for describing global image-processing phenomena. We expect that the higher-order statistical framework will continue to be useful in computational and psychophysical studies of natural-image perception; indeed, one advantage of the approach described here is that it may be taken one step further: to ask, for example, out of the set of all possible relative-phase-sensitive units, which ones will process the sort of relative-spatial-phase relationships found in natural images most efficiently?

8 Conclusions
The visual effects of quantising and randomising the Fourier phases of natural and random-phase images have been illustrated by way of demonstrations and psychophysical experimentation. The resulting percepts depend qualitatively not on the nature of the phase perturbations themselves, but on the nature of the phase spectrum of each individual image. These results are predictable from simple global higher-order statistical

measures—in particular, the phase-only image kurtosis—but not from simple first-order measures. This finding implies that the visual cortical mechanisms involved in processing those structural image changes induced by phase perturbations could operate by correlating the outputs of banks of relative-phase-sensitive units.

References

Atick J J, Redlich N A, 1992 "What does the retina know about natural scenes?" *Neural Computation* **4** 196 – 210

Field D J, 1987 "Relations between the statistics of natural images and the response properties of cortical cells" *Journal of the Optical Society of America A* **4** 2379 – 2394

Field D J, 1989 "What the statistics of natural images tell us about visual coding" *SPIE—Human Vision, Visual Processing and Digital Display* **1077** 269 – 276

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

Hübner M, Caelli M, Rentschler I, 1988 "Visual phase resolution for grayscale textures" *Perception & Psychophysics* **43** 319 – 325

Julesz B, Schumer R A, 1981 "Early visual perception" *Annual Review of Psychology* **32** 575 – 627

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

Juwells I, Vallmitjana S, Carnicer A, Campos J, 1991 "The role of amplitude and phase of the Fourier transform in digital image processing" *American Journal of Physics* **59** 744 – 748

Klein S A, Tyler C W, 1986 "Phase discrimination of compound gratings: generalized autocorrelation analysis" *Journal of the Optical Society of America A* **3** 868 – 879

Knill D C, Field D J, Kersten D, 1990 "Human discrimination of fractal images" *Journal of the Optical Society of America A* **7** 1113 – 1123

Lohmann A W, Mendlovic D, Shabtay G, 1997 "Significance of phase and amplitude in the Fourier domain" *Journal of the Optical Society of America A* **14** 2901 – 2904

Morrone M C, Burr D C, 1988 "Feature detection in human vision: a phase-dependent energy model" *Proceedings of the Royal Society of London, Series B* **235** 221 – 224

Nikias C, Petropulu A, 1993 *Higher-order Spectra Analysis* (New York: Prentice-Hall)

Oppenheim A V, Lim J S, 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

Press W H, Flannery B P, Teukolsky S A, Vetterling W T, 1988 *Numerical Recipes in C* (Cambridge, UK: Cambridge University Press)

Summers R J, Thomson M G A, 1999 "Nonuniform phase perturbations in natural images" *Perception* **28** Supplement, 127

Tadmor Y, Tolhurst D J, 1993 "Both the phase and amplitude spectrum may determine the appearance of natural images" *Vision Research* **33** 141 – 145

Tadmor Y, Tolhurst D J, 1994 "Discrimination of changes in the second-order statistics of natural and synthetic images" *Vision Research* **34** 541 – 554

Taylor M M, Creelman C D, 1967 "PEST: efficient estimates on probability functions" *Journal of the Acoustical Society of America* **41** 782 – 787

Thomson M G A, 1999a "Visual coding and the phase structure of natural scenes" *Network: Computation in Neural Systems* **10** 123 – 132

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

Thomson M G A, Foster D H, 1997 "The role of second- and third-order statistics in the discriminability of natural images" *Journal of the Optical Society of America A* **14** 2081 – 2090

Victor J D, Conte M M, 1996 "The role of high-order phase correlations in texture processing" *Vision Research* **36** 1615 – 1631

p