

# Veridicality of three-dimensional shape perception predicted from amplitude spectra of natural textures

Andrea Li and Qasim Zaidi

*SUNY State College of Optometry, 33 West 42nd Street, New York, New York 10036*

Received January 12, 2001; accepted March 6, 2001; revised manuscript received March 29, 2001

We show that the amplitude spectrum of a texture pattern, regardless of its phase spectrum, can be used to predict whether the pattern will convey the veridical three-dimensional (3-D) shape of the surface on which it lies. Patterns from the Brodatz collection of natural textures were overlaid on a flat surface that was then corrugated in depth and projected in perspective. Perceived ordinal shapes, reconstructed from a series of local relative depth judgments, showed that only about a third of the patterns conveyed veridical shape. The phase structure of each pattern was then randomized. Simulated concavities and convexities were presented for both the Brodatz and the phase-randomized patterns in a global shape identification task. The concordance between the shapes perceived from the Brodatz patterns and their phase-randomized versions was 80–88%, showing that the capacity for a pattern to correctly convey concavities and convexities is independent of phase information and that the amplitude spectrum contains all the information required to determine whether a pattern will convey veridical 3-D shape. A measure of the discrete oriented energy centered on the axis of maximum curvature was successful in identifying textures that convey veridical shape. © 2001 Optical Society of America

*OCIS codes:* 330.0330, 330.5020, 330.5510, 150.0150, 100.2960.

## 1. INTRODUCTION

The pattern of markings on a surface is a powerful visual cue to the three-dimensional (3-D) shape of the surface.<sup>1</sup> The two-dimensional (2-D) image in Fig. 1A (left) conveys the percept of a corrugated 3-D surface solely from the surface markings, i.e., without stereo, motion, silhouettes, or other cues to 3-D shape. Li and Zaidi<sup>2</sup> showed that observers perceived veridical 3-D shape only in the presence of surface markings oriented along lines of maximum curvature. In Fig. 1A these markings are the noisy sparse contours that appear to run along the curvature of the surface from left to right.

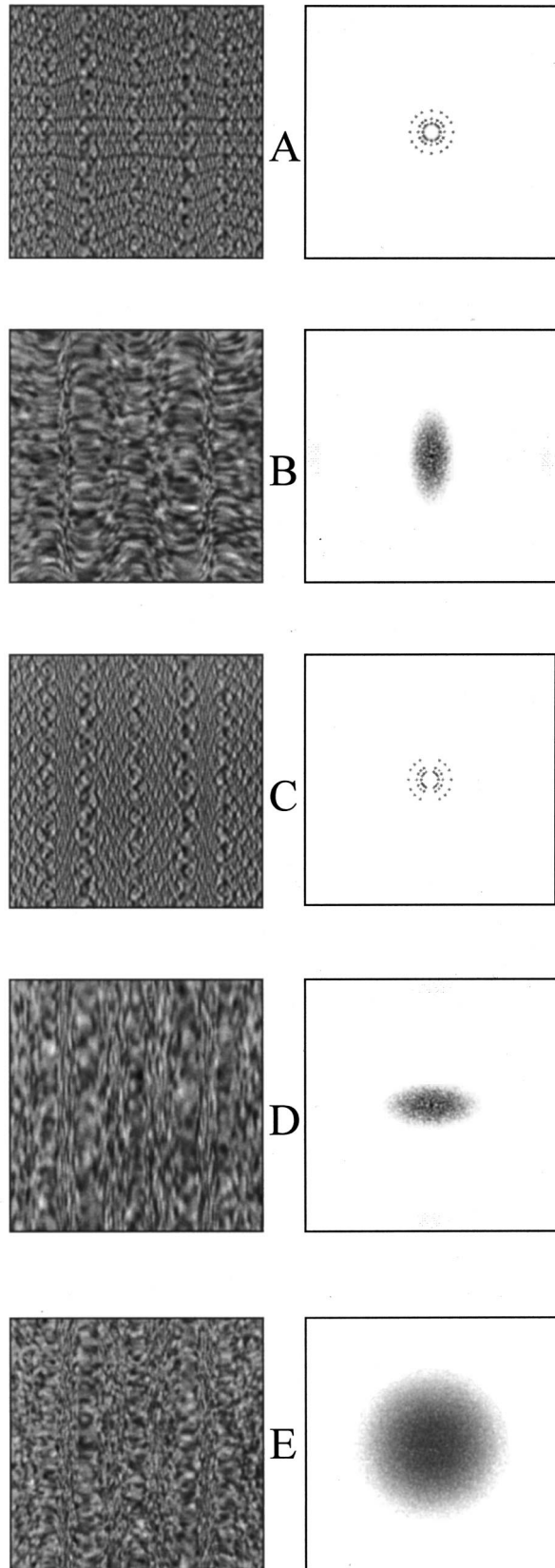
Before corrugation and projection, the texture pattern used in Fig. 1A consisted of eight different complex gratings, spaced 22.5 deg in orientation beginning at 0° (horizontal), each complex grating being the sum of three different spatial frequencies in random phase. The visible contours along projected lines of maximum curvature are formed by the horizontal components of the texture. The right panel of Fig. 1A is the 2-D amplitude spectrum of the pattern. (0, 0) in spatial frequency lies at the center of the panel, and higher energies are indicated by darker points. The vertical axis in frequency space, i.e., the horizontal axis in the image, is the axis of maximum curvature for this 3-D shape.

Figure 1 illustrates the main results of Li and Zaidi.<sup>2</sup> Besides patterns in which energy along the axis of maximum curvature is relatively discrete from energy at neighboring orientations (Fig. 1A), the other class of patterns that convey veridical shape are those whose amplitude spectra contain energy that is oriented predominantly along the critical axis, as in Fig. 1B. The critical

surface markings in this image are visible as short streaks lying along projected lines of maximum curvature. Synthetic patterns not belonging to the above classes (Fig. 1C–1E) did not convey veridical shape.<sup>2</sup> The amplitude spectra of these patterns either do not contain any components along the critical axis (Fig. 1C), are oriented along a noncritical axis (Fig. 1D), or are isotropic (Fig. 1E). The left panel of Fig. 1C contains similar texture gradients<sup>3–8</sup> and changes in frequency<sup>9</sup> as in Fig. 1A but does not convey veridical shape. Hence texture gradients and frequency modulations are neither necessary nor sufficient for conveying veridical 3-D shape.<sup>2</sup>

Li and Zaidi<sup>10</sup> showed mathematically that only energy components oriented within a few degrees of the axis of maximum curvature will distinguish concavities from convexities. This explains why patterns missing these necessary components will not convey the veridical shape of a corrugation. Components at other orientations provide information sufficient only to distinguish extrema of curvature from planar portions but not to distinguish between signs of curvatures or signs of slant. This mathematical result links the perception of shape from complex textures to the perception of shape from parallel rulings studied by Stevens<sup>11</sup> and may remove the need to incorporate Stevens's strong constraint that rulings are interpreted as following lines of maximum curvature on the surface. Knill<sup>12</sup> has recently derived a strong corollary to the above result—i.e., the texture flow on a developable surface follows parallel geodesics of the surface. However, since some patterns containing the critical component also do not convey veridical shape (Figs. 1D and 1E), sufficiency conditions may be more complicated.

The goal of this paper was to identify sufficiency conditions by using a set of naturally occurring textures. We first measured how well each of a set of naturally occur-



ring patterns conveyed 3-D shape when overlaid on a flat surface that was then corrugated and projected in perspective. Second, we determined whether a pattern's capacity for conveying veridical shape was contingent upon its amplitude spectrum regardless of its phase structure. This question was tested by generating a set of test patterns that had amplitude spectra identical to the original textures but randomized phase spectra. Our results show that the amplitude spectrum can be used to predict whether a pattern will convey veridical shape. Third, we devised objective measures based on the amplitude spectra of patterns to separate patterns that convey veridical shape from those that do not.

## 2. EXPERIMENT 1: PERCEIVED ORDINAL SHAPE FROM BRODATZ TEXTURES

The purpose of the first experiment was to quantitatively document the shapes conveyed by various everyday texture patterns. Local relative depth judgments were used to reconstruct perceived shape for corrugated surfaces in perspective projection.

### A. Texture Patterns

Texture patterns from the Brodatz<sup>13</sup> collection were used in this study. They consist of gray-level frontoparallel photographs of various man-made and natural materials, such as woven textiles and animal skins. The collection provides images containing textures at various scales and orientations and has been extensively analyzed.<sup>14</sup> Unlike many of the artificial texture patterns used in previous shape-from-texture studies,<sup>3-8</sup> most of these natural patterns cannot be easily parsed into individual texture elements. Patterns were eliminated if they contained shading caused by the 3-D texture of the material, as this would have resulted in projected images with shading cues that were inconsistent with the corrugated shape of the surface. They were also eliminated if the texture consisted of just a few large objects, as in the case of certain lace patterns that contained large floral patterns only once or twice across the entire pattern. Of the 56 patterns used, 40 were man-made materials (mainly textiles), and the remaining 16 fell into one of three naturally occurring categories: animal pelts (8), terrains such as pebbles or grass (4), and solid surfaces such as wood and stone (4).

Fig. 1. *Patterns that convey veridical shape*: A, Left: Octotroptic plaid pattern drawn on a surface corrugated in depth as a function of horizontal position and projected in perspective with the center of the image at eye height. The pattern consists of eight compound gratings, each oriented 22.5 deg from the next. Each compound grating is the sum of three frequencies at random phases. Right: The amplitude spectrum of the uncorrugated pattern. (0, 0) in spatial frequency lies at the center of the panel, and energy amplitude increases with increasing darkness. B, Pattern of white noise filtered with an elliptical filter oriented along the axis of maximum surface curvature. *Patterns that do not convey veridical shape*: C, Same complex plaid pattern as in A minus the horizontal compound grating. The pattern contains all the relevant texture gradients and frequency modulations consistent with the corrugated surface. D, Pattern of white noise filtered with an elliptical filter oriented along the axis of minimum curvature. E, Pattern of isotropic broadband noise.

Patterns were scanned in at 150 dpi on a Hewlett-Packard 4p Scanner and saved in PCX format. Each pattern was  $600 \times 600$  pixels. They were then ported into Matlab, which was used for all image computations. The amplitude spectrum of each pattern was computed. These spectra are shown in the leftmost columns of Fig. 2. The corresponding corrugated, projected images of the patterns are shown in the second column. The pattern number is shown above the two panels to the right.

### B. Corrugation and Projection

Shape-from-texture studies have conventionally used either flat or singly curved surfaces such as cylinders.<sup>3-9</sup> Results from Li and Zaidi<sup>2,10</sup> showed the necessity of using stimuli that contain both convex and concave curvatures. For example, although Sakai and Finkel<sup>9</sup> showed that one-dimensional frequency modulations were effective at conveying the shapes of convex cylindrical surfaces, Li and Zaidi<sup>2</sup> showed that they are not sufficient for distinguishing convexities from concavities, even under perspective projection.

Each pattern was overlaid on a flat surface that was then corrugated sinusoidally in depth ( $z$ ) as a function of horizontal position ( $x$ ):

$$z = A_{3D} \cos(2\pi f_{3D}x + \phi) + d, \quad (1)$$

where  $A_{3D}$  and  $f_{3D}$  are the amplitude and frequency of the corrugation, respectively, and  $\phi$  is the phase of the corrugation at the center of the image. We then computed the perspective projection of each corrugated pattern onto the plane of the CRT display (see Appendix of Ref. 2). The amplitude of the depth modulation ( $A_{3D}$ ) was computed to be 8 cm from peak to trough for a viewing distance  $d$  of 44 cm. The projected image was 381 pixels square and subtended 21 deg visual angle. Each image contained three full cycles of the corrugation. The patterns in Fig. 2 have all been corrugated and projected with a central phase of zero.

### C. Equipment

Stimuli were presented using a Cambridge Research Systems Visual Stimulus Generator (CRS VSG2/3) on a BARCO 7651 color monitor with a  $736 \times 550$  pixel screen running with a refresh rate of 100 frames/s. Through the use of 12-bit DACs, after gamma correction, the VSG2/3 is able to generate 2861 linear levels for each gun.

### D. Psychophysical Method

Perceived ordinal shape was reconstructed for each pattern from a series of local relative depth judgments.<sup>2</sup> Each projected image contained a central square fixation flanked by two smaller test dots positioned to the left and right of the fixation along a  $-45^\circ$  or  $+45^\circ$  diagonal. The test dots were  $0.5^\circ$  visual angle away from the fixation and abutted it at its diagonal corners. Observers were asked to fixate the central dot and indicate whether the location corresponding to the left or right dot appeared closer or whether the two locations appeared at equal depths. To sample both convex and concave portions of the surface, the central fixation location was set to corre-

spond to one of 12 different phases along a single cycle of the corrugation. The surface was projected in perspective with the center of each image at eye height.

One texture pattern was used for each session. We randomly interleaved 5 trials for each of the 12 phases for test dots along the  $+45^\circ$  diagonal, and the same for test dots along the  $-45^\circ$  diagonal. Images were presented against a black background for 1.5 s followed by a gray mask of the same size and average luminance that remained on until the observer made a response.

Observers rested their chins in a headrest, and viewing was monocular in a dark room. Each session lasted approximately 5 min. One of the authors (AL) and a student (JR) served as observers. JR had participated in similar experiments on synthetic textures.

### E. Ordinal Shape Reconstruction

Each relative depth judgment provides the sign of the local slope of the surface. The perceived ordinal shape of the surface was reconstructed by integrating these local signs cumulatively from left to right in the following way. Initially, the estimate of perceived relative depth at each test location was set to zero (i.e., equivalent to the plane of the screen). Distances beyond the plane of the screen were indicated by positive numbers, distances in front of the screen by negative numbers. From the results of each trial, the estimated perceived depth of the right test location was incremented and decremented according to the observer's response: If the right test location appeared closer, the depth estimate at that location was decremented by 1; if the left location appeared closer (i.e., the right location appeared farther away), the depth estimate at the right test location was incremented by 1. Depth estimates were left unchanged if the two locations were reported at equal depths. Since there were five trials for each test dot configuration at each location, resulting relative depth estimates ranged from  $-5$  to  $5$ . The final array of relative depth estimates was then added from left to right, starting at zero. For example, the final array of depth estimates of a surface for which all local relative depth relationships were seen correctly was (5, 5, 5, 0,  $-5$ ,  $-5$ ,  $-5$ ,  $-5$ , 0, 5, 5) which yielded a reconstruction of (5, 10, 15, 15, 10, 5, 0,  $-5$ ,  $-10$ ,  $-10$ ,  $-5$ , 0) (Fig. 3, top). Results across the two diagonal test dot configurations did not differ systematically, and so depth estimates were averaged across the two.

This reconstruction provided an ordinal estimate of the perceived shape. Hence global percepts of a triangular waveform and a flattened sinusoid could yield the same reconstruction, as could percepts of different amplitudes of depth modulation. However, the method successfully reveals perceived convexities and concavities.<sup>2</sup>

### F. Classification of Veridicality

The reconstructed perceived shape corresponding to each texture pattern is shown for each of the two observers in the rightmost columns of Fig. 2. Each plot shows perceived shape for one cycle of corrugation as a function of the phase (in radians). Observers' initials are indicated at the top of each column. The dashed vertical line in the plots indicates the central phase of the corrugated pat-



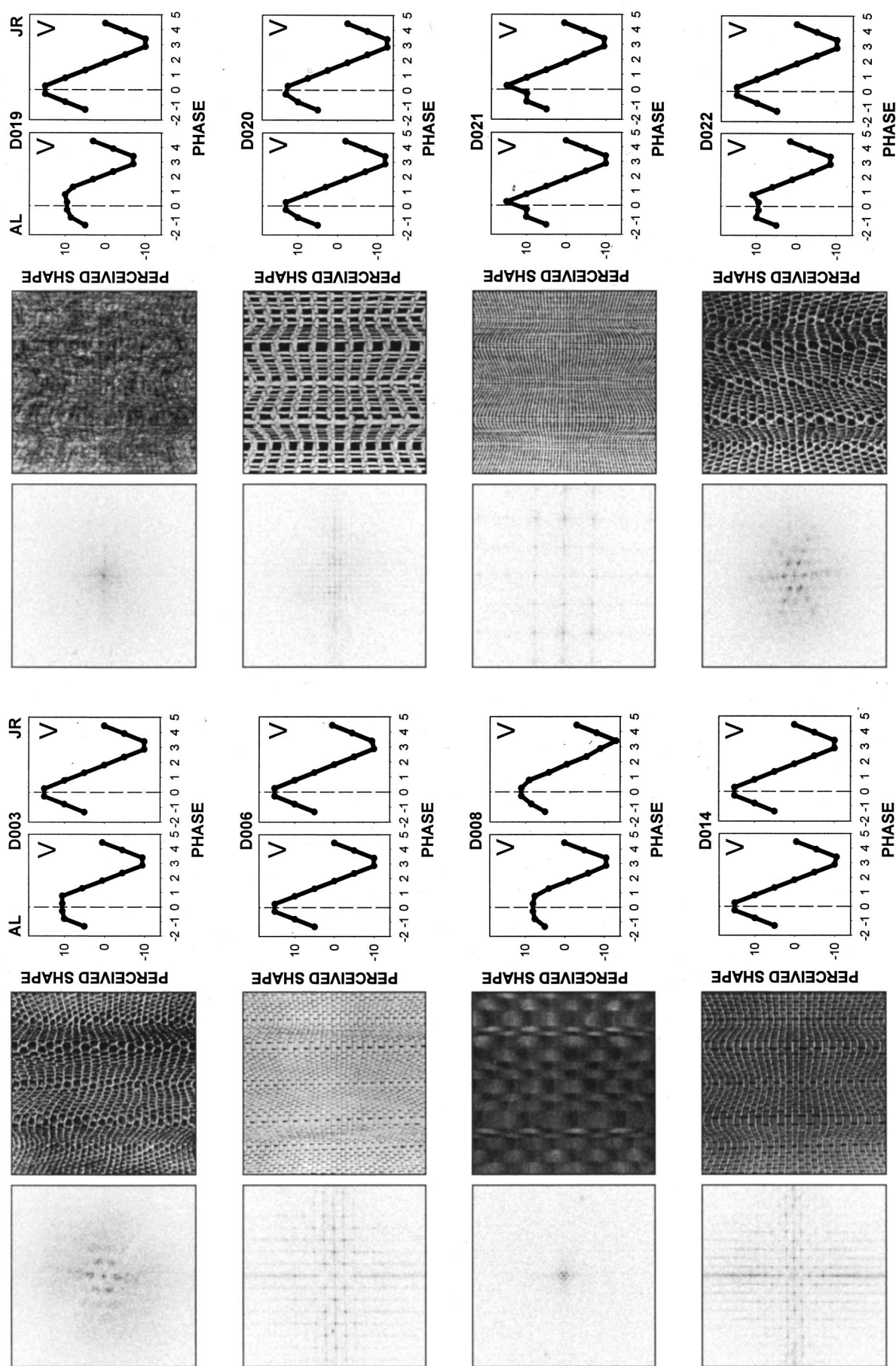


Fig. 2. Amplitude spectra, corrugated patterns, and data for 56 Brodatz patterns, arranged in groups of four panels per pattern. Amplitude spectra of uncorrugated patterns are shown in the first column, corrugated projected patterns in the next two columns to the right, and data for two observers in the next two columns to the right (AL on the left and JR on the right). The pattern number is indicated above the two data plots. The dashed vertical line in the plots indicates the central phase of the corrugated patterns shown to the left. The perceived shape is classified according to the best-fitting template (see Fig. 3), which appears in the upper right or upper left corner of each data panel. V, veridical; HR, half-rectified; FR, fully rectified; F, flat; O, other. Patterns that were perceived veridically by both observers are presented first, followed by those perceived veridically by only one observer, and finally by neither observer. (Continues on next six pages.)

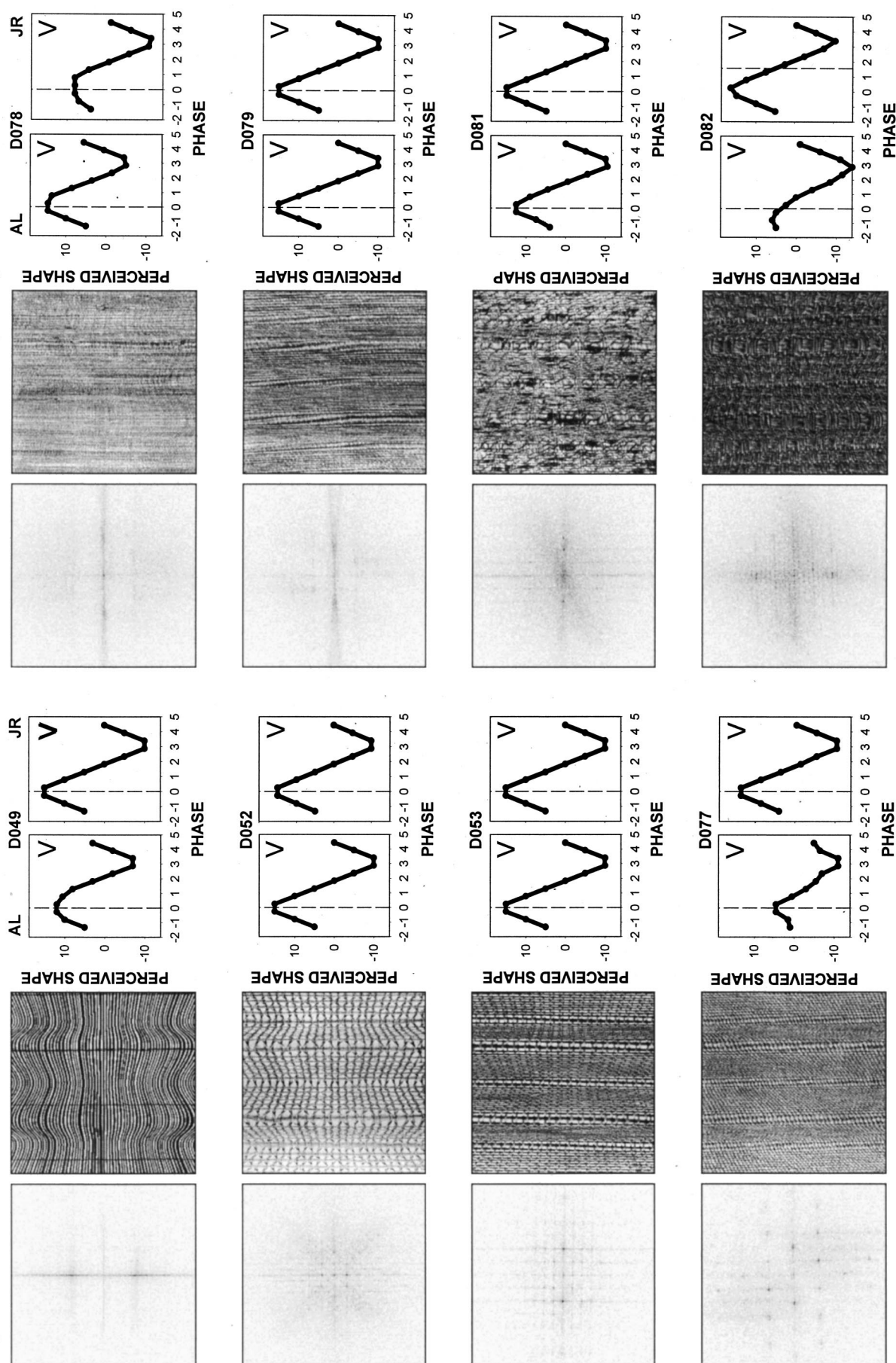


Fig. 2. (continued)



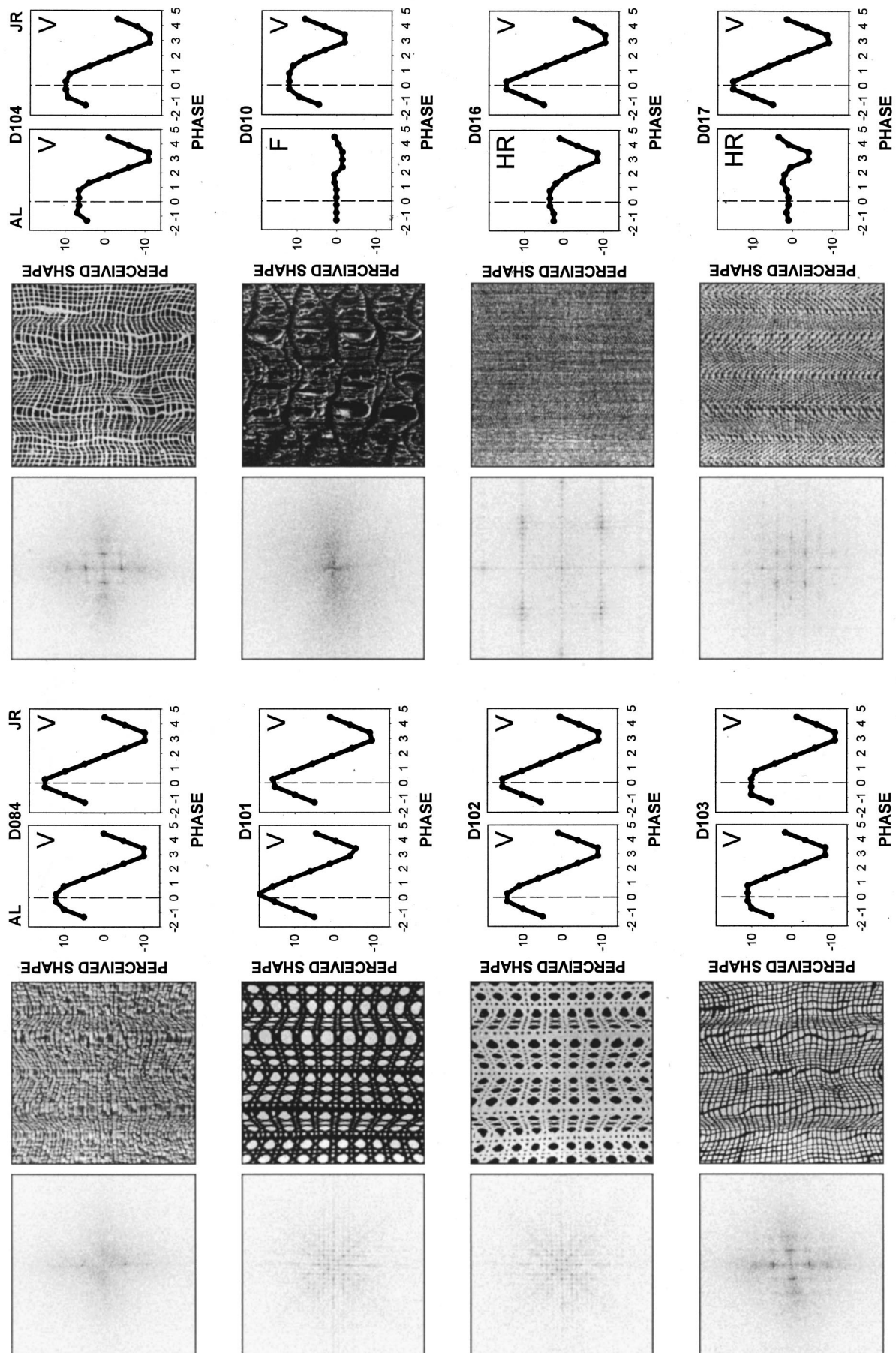


Fig. 2. (continued)

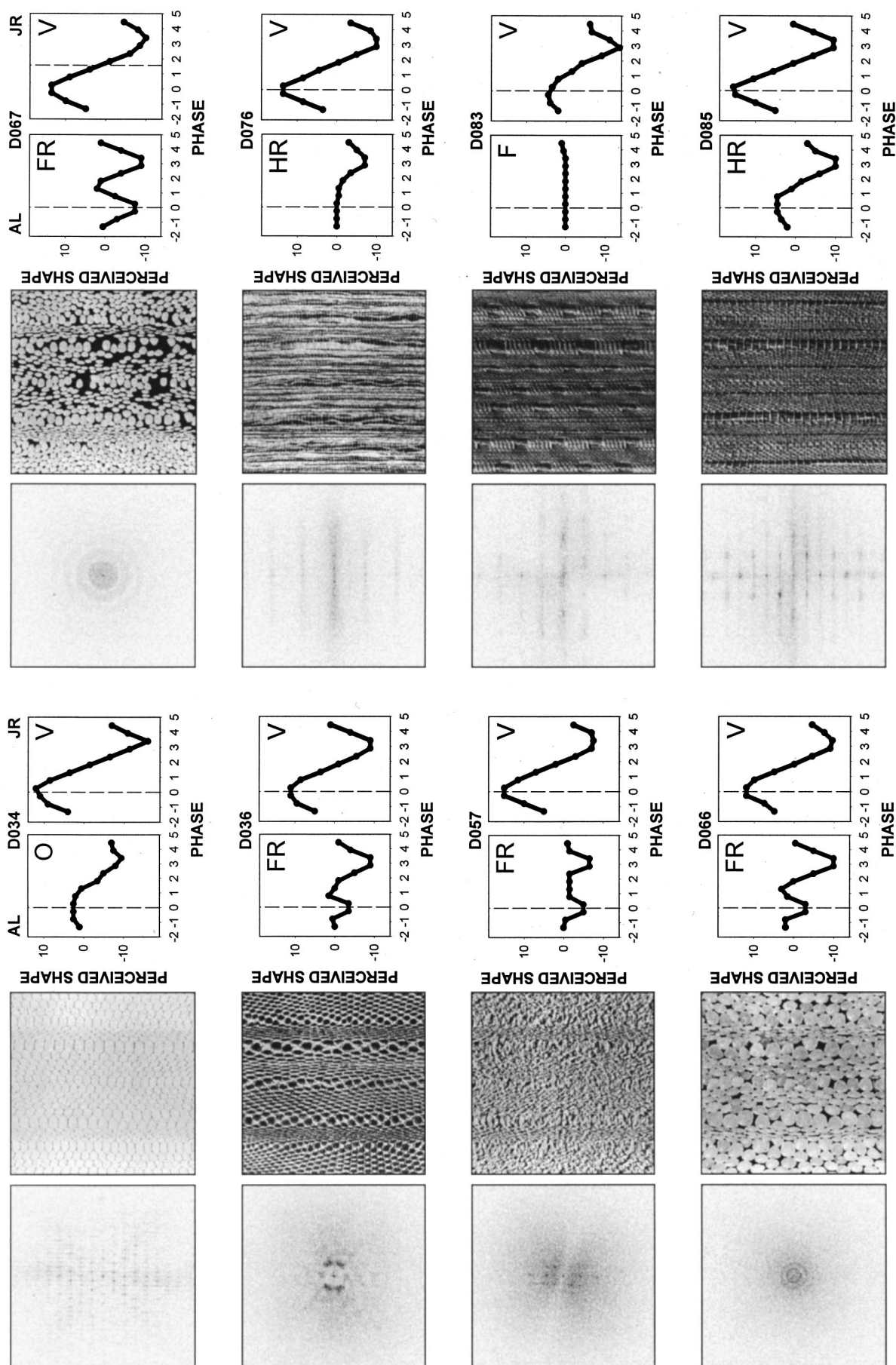


Fig. 2. (continued)



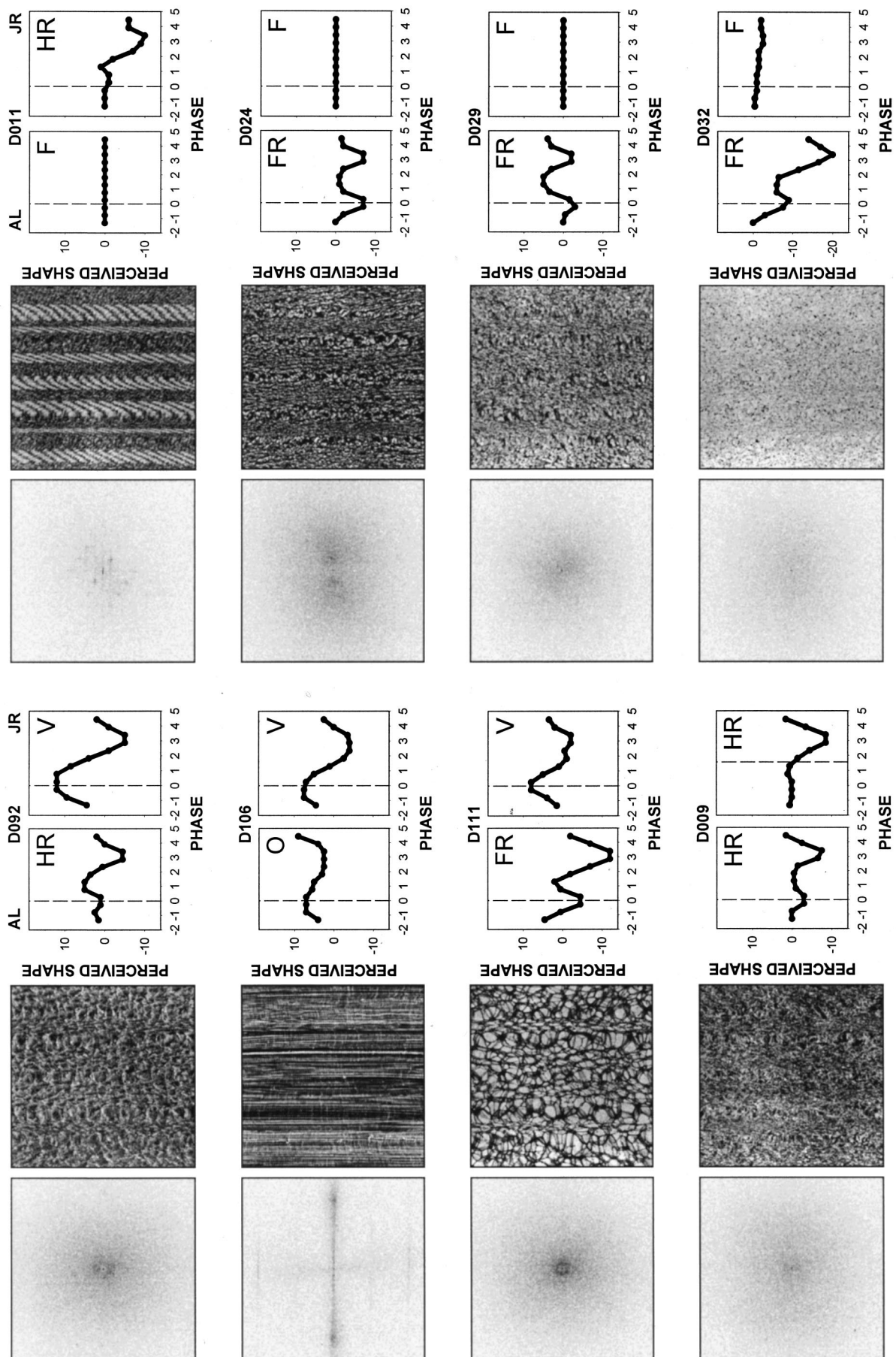


Fig. 2. (continued)



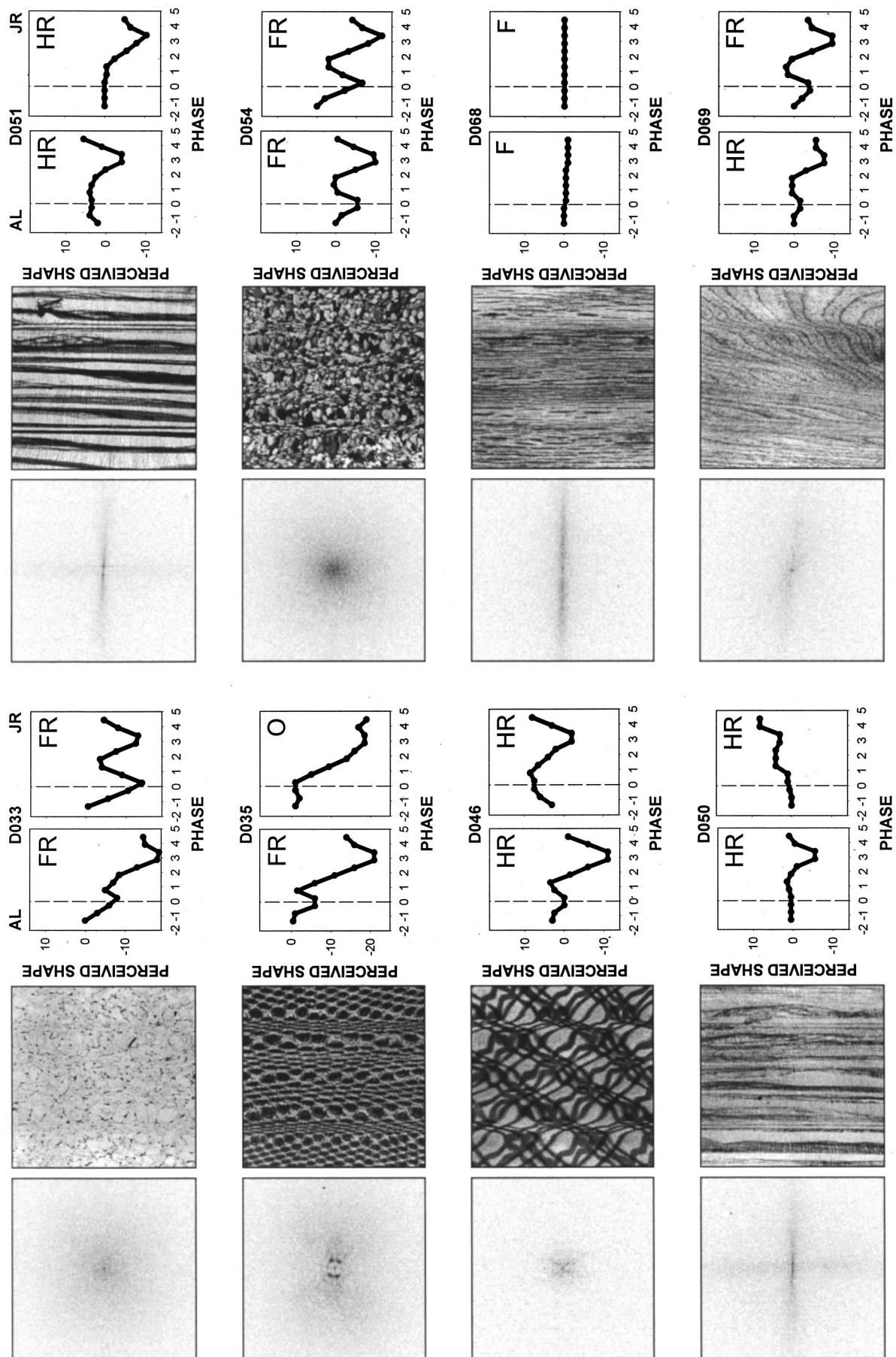


Fig. 2. (continued)

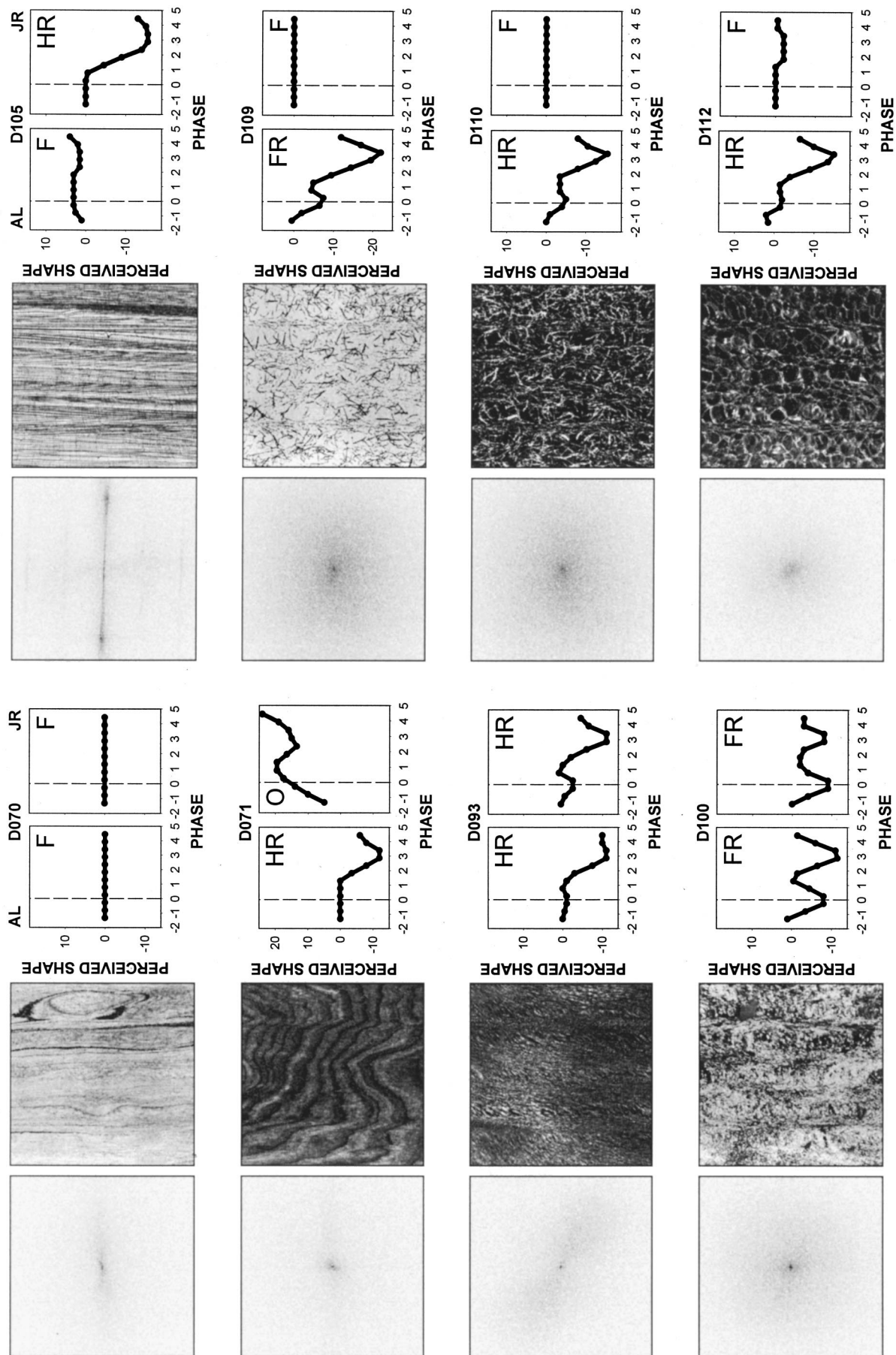


Fig. 2. (continued)



terns shown on the left. A number of different types of shapes resulted from the relative depth judgments, including half-rectified, fully rectified, and flat surfaces. There was one notable tendency in the data: In almost all cases, if any portion of the surface was seen incorrectly, it was the concave portion of the corrugation.

To classify the perceived ordinal shapes, we determined the similarity between the reconstruction for each pattern and each of the templates in Fig. 3. The veridical template for the simulated surface is shown in the top panel. Three nonveridical templates, half-rectified, fully rectified, and flat (frontoparallel), were also tested against the reconstructions (see Appendix A for details). If the data did not fit any of the templates, then the shape was classified as "other nonveridical."

A veridical reconstruction results from the percept of a surface whose convexities and concavities are seen correctly with a high degree of certainty; that is, this reconstruction is achieved only if the observer makes correct relative depth judgments at all locations on every trial. By simple statistical considerations, if the difference in perceived relative depth between two locations in the reconstruction is either 4 or 5, we can reject the hypothesis that the slant between the locations is perceived as flat and that the value results from random response. Perceived relative depths of 3 or less between two locations in the reconstruction can result from a surface that is perceived with a high degree of certainty as frontoparallel or from an uncertain percept resulting in random relative depth responses. Regardless of the certainty with which the different nonveridical shapes are seen, it is the veridical versus nonveridical classification that is the main focus of this paper.

### G. Results

The classification of the perceived ordinal shape is indicated in the upper right or left corner of each reconstruction panel in Fig. 2 (V, veridical; FR, fully rectified; HR, half-rectified; F, flat; O, other). The patterns are grouped by frequency of veridicality. That is, patterns that were perceived as veridical by both observers are shown first, followed by those that were perceived as veridical by one observer, and finally by neither observer. The results show that not all natural textures convey veridical ordinal shape. Observer AL and JR, respectively, perceived 37.5% and 45% of the patterns veridically, with 29% of the 56 patterns seen as veridical by both.

We informally tested the prediction that the amplitude spectra of the patterns that convey veridical ordinal shape contain dominant or discrete energy along the critical axis, whereas the spectra of patterns that convey nonveridical shapes contain dominant or discrete energy along noncritical axes, or are isotropic.<sup>2</sup> A classification of the spectra by eye by one of the authors (performed without knowledge of which spectrum belonged to which pattern to minimize bias) yielded 28 patterns that were predicted to convey veridical shape and 28 patterns predicted to not convey veridical shape. Approximately 80% of the classifications were consistent with the perceived shapes obtained for each observer. Therefore classification by eye of the amplitude spectrum might be reasonably predictive of whether or not a pattern will convey

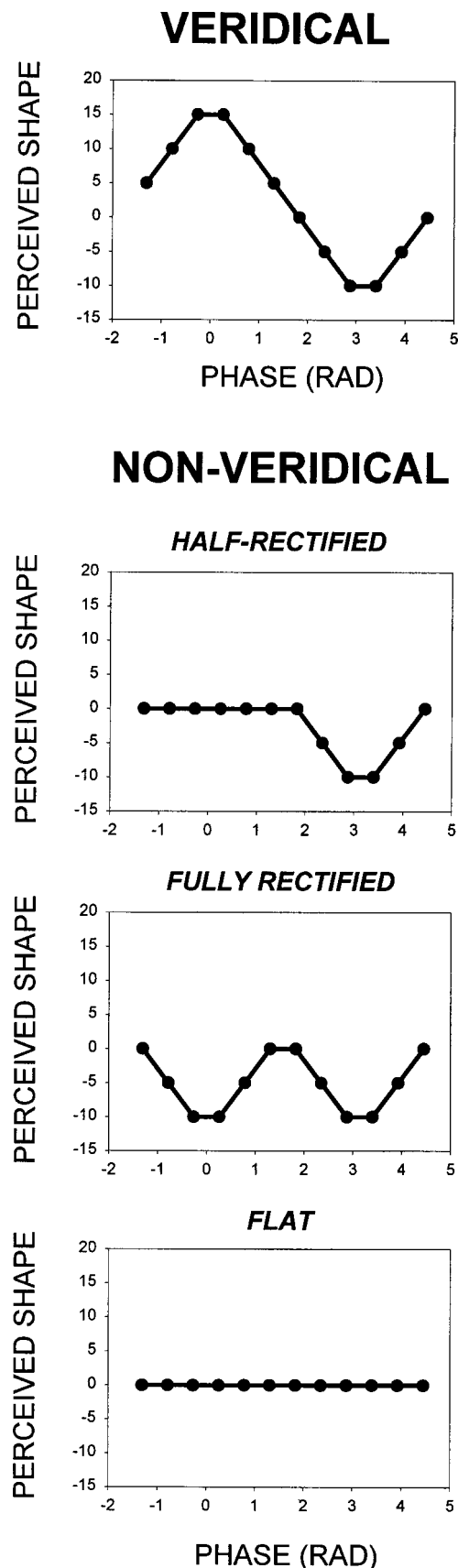
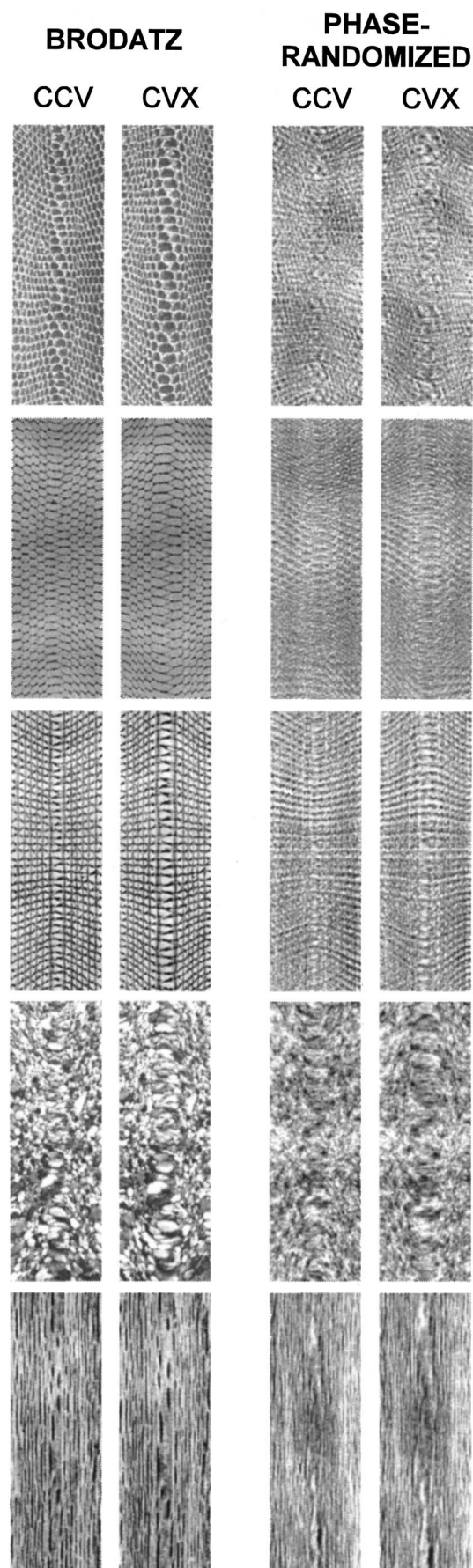


Fig. 3. Template shapes to which individual data were compared. Top, veridical; bottom, nonveridical: half-rectified, fully rectified, and flat.



veridical ordinal shape. In a later section we examine whether an objective classification can do better, but before doing so, we test whether the capacity for a pattern to convey veridical shape is also contingent upon its phase structure.

### 3. EXPERIMENT 2: RANDOMIZING PHASE STRUCTURE

#### A. Generation of Phase-Randomized Patterns

The amplitude spectrum of each pattern quantifies the power at each spatial frequency in the pattern. The phase spectrum quantifies the phase relations between the various frequencies in the pattern. To test whether these phase relations play a role in determining whether a pattern will convey veridical shape, for each of the 56 Brodatz patterns we generated a pattern that had the same amplitude spectrum but whose phase spectrum was randomized.

The 2-D Fourier transform of a pattern  $T$  can be expressed in terms of its amplitude-spectrum matrix  $A$  and phase spectrum matrix  $P$ :  $F(T) = Ae^{iP}$ . To generate the phase-randomized pattern ( $T_r$ ), we used the amplitude spectrum  $A$  with a random-phase-spectrum matrix  $P_r$  to create  $F(T_r) = Ae^{iP_r}$ . The new pattern was the inverse Fourier transform of  $F(T_r)$ . All images were normalized to have the same mean luminance. Each of the Brodatz and the phase-randomized patterns was overlaid on a flat surface, corrugated, and projected in perspective as either a central concavity or a central convexity. The central  $7^\circ \times 21^\circ$  vertical strip containing 80% of a full cycle of the sinusoidal corrugation was used as the stimulus. The two left columns of Fig. 4 show some examples of concavities and convexities of normalized Brodatz patterns and the two right columns show their phase-randomized equivalents. Randomizing the phase structure removes many of the sharp contours that are created by precise phase relations between harmonics.

#### B. Psychophysical Method

The number of stimulus patterns was large (56 Brodatz patterns plus 56 phase-randomized patterns); therefore we used a global shape identification paradigm in which observers were asked to identify the 3-D surface shape in a three-alternative forced-choice task (concave, convex, or other). We have used a similar paradigm previously<sup>10</sup> and found it to yield reliable results.

The experiment consisted of 7 sessions, each repeated 5 times. Each session contained 320 trials: 32 stimuli (8 Brodatz patterns and their 8 phase-randomized analogs, each at two simulated curvatures), presented for 10 trials each randomly interleaved. Each session lasted approximately 10 min. Each of the 224 stimuli were thus presented 50 times. There was no feedback. Observers'

Fig. 4. Left: examples of Brodatz patterns corrugated and projected in concavity and convexity phase. From top to bottom, Brodatz pattern numbers are D003, D034, D052, D054, and D068. Right: phase-randomized versions of the patterns in the left column. Patterns have all been normalized to have the same mean luminance.



heads were fixed, and viewing was monocular in a dark room.

The same observers from Experiment 1 also participated in this experiment. Observers fixated a central square in a mid-gray screen for 1 min at the beginning of each session. The stimuli were presented for 1 s each, followed by the mid-gray screen. A three-button response box was used to indicate whether the surface appeared concave, convex, or other. The third choice of "other" was included to minimize random biases toward one or the other curvature on trials where the observer was unsure of either. Stimuli were presented by using the VSG2/3 on a SONY GDM-F500 flat-surface color monitor with a  $800 \times 600$  pixel screen running with a refresh rate of 100 frames/s.

### C. Results

In Fig. 5, each panel shows the number of correct responses out of 50 for each Brodatz pattern plotted against the number of correct responses for the phase-randomized

analog. Each point thus corresponds to one of the 56 texture patterns. The data for simulated concavities are plotted in the left column, and data for simulated convexities are plotted in the right column. The solid horizontal and vertical lines in each panel indicate 35 out of 50 correct responses. In 50 repetitions, the probability that an observer has made 35 or more correct responses by chance is less than 0.05; therefore stimuli that were identified correctly in more than 35 out of 50 trials were considered veridical identifications. Stimuli that were perceived veridically for both the Brodatz and the phase-randomized patterns fall in the small square at the upper right of each panel; those that were perceived nonveridically for both patterns fall in the larger square at the lower left of each panel. Most of the data fall in these two squares, suggesting that there is a high degree of agreement between the veridicality of percepts for the Brodatz patterns and their phase-randomized analogs. In all four panels, the cluster at the upper right corner indicates that a number of surfaces were reported as veridi-

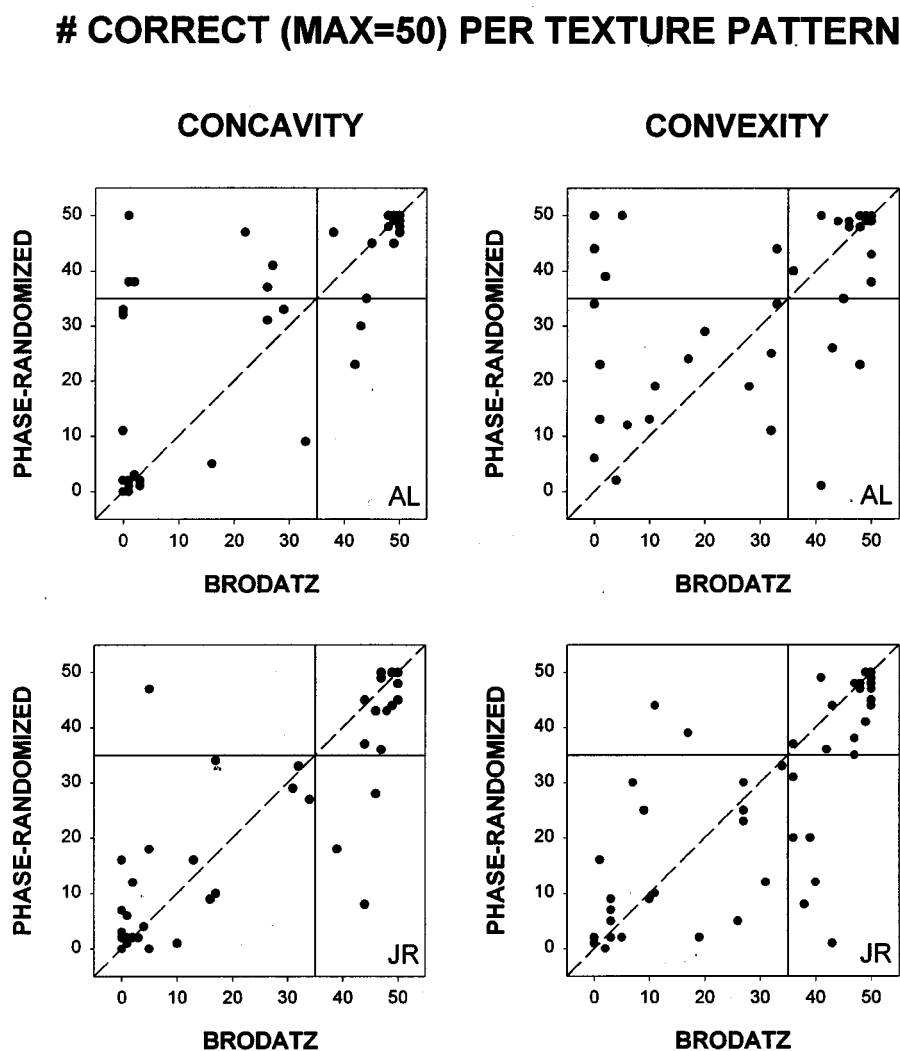


Fig. 5. Number of correct responses out of 50 for each pattern plotted for the Brodatz patterns along the abscissa and for the phase-randomized patterns along the ordinate. Data are plotted by simulated curvature in the two columns and by observer in the two rows. Each panel contains 56 points, each corresponding to one of the Brodatz/phase-randomized patterns. The solid vertical and horizontal lines indicate the 35/50 correct response boundaries. Patterns for which observers perceived veridical shape in both the Brodatz and the phase-randomized conditions fall in the square at the upper right of each panel; those for which observers perceived nonveridical shapes in both conditions fall in the square at the lower left of each panel.

**Table 1. Number of Patterns Seen Veridically for Simulated Curvatures**

Pattern	Type of Curvature			
	Concave		Convex	
	AL	JR	AL	JR
Brodatz and phase-randomized	26	17	35	28
Brodatz only	4	4	3	6
Phase-randomized only	7	3	6	2
Neither	19	32	12	20
Total	56	56	56	56

cal with a high degree of certainty. The clustering at the lower left corner suggests that nonveridical shapes were also reported with a high degree of certainty. Very few

$$\alpha = \tan^{-1} \left( \frac{d \sin \omega - y \cos \omega \sin \theta}{d \cos \theta \cos \omega} \right), \quad (2)$$

$$\sigma = \frac{f[d + \cos(\omega + \pi/2)\cos \theta]}{\{\cos^2 \theta [\cos^2(\omega + \pi/2)(d^2 + y^2) + d^2 \sin^2(\omega + \pi/2) - 2yd \sin(\omega + \pi/2)\cos(\omega + \pi/2)\sin \theta]\}^{1/2}}. \quad (3)$$

points fall around 25/50, which would have indicated random responding. Stimuli that were perceived veridically for only the Brodatz patterns fall in the rectangle at the lower right. Stimuli that were perceived veridically for only the phase-randomized patterns fall in the upper left rectangle.

Table 1 shows the numbers of patterns in each subdivision of each panel in Fig. 5. The number of patterns for which the correct curvature was reported for both the Brodatz and the corresponding phase-randomized pattern is tabulated in the first row, the number of patterns for which the correct curvature was reported for only the Brodatz pattern in the second row, for only the phase-randomized pattern in the third row, and for neither pattern in the fourth row. The perceived shapes for the Brodatz and the phase-randomized patterns were in close agreement. For observers JR and AL, 87.5 and 80.4% of the simulated concave patterns and 85.7 and 83.9% of the simulated convex patterns fall in the first and fourth rows. The high concordance between the shapes conveyed by the Brodatz and the phase-randomized patterns indicates that the amplitude spectrum contains the critical shape-conveying information about a pattern; randomization of the phase structure had little or no effect on the veridicality of the perceived shape.

For observer AL, there were four patterns whose shapes were perceived veridically for the Brodatz pattern and nonveridically for the phase-randomized pattern (D035, D051, D054, and D066). These four patterns were, however, perceived as nonveridical in Experiment 1. For observer JR, there were also four patterns perceived as veridical for the Brodatz pattern and as nonveridical for the phase-randomized pattern (D008, D016,

D081, and D082). These patterns were perceived as veridical in Experiment 1. It is possible that for these cases, for this observer, phase randomization reduced the perceived contrast of critical contours in the image that led to the nonveridical percepts.

#### 4. ANALYSIS OF THE AMPLITUDE SPECTRUM

To develop an objective metric that can be used directly to indicate whether a pattern will convey veridical 3-D shape, we consider how the amplitude spectra for patterns that convey veridical shape differ quantitatively from those of patterns that do not convey veridical shape.

Li and Zaidi<sup>10</sup> showed that for a developable surface of local slant  $\theta$  at a distance  $d$ , a texture component of spatial frequency  $f$ , oriented at  $\omega$  with respect to the horizontal, will have an image orientation  $\alpha$  and spatial frequency  $\sigma$  given by

The slants for the concavity and convexity simulations in Experiment 2 range from  $-84$  to  $+84$  deg. For these slants, only components oriented within 3 deg of the axis of maximum curvature were found to provide sufficient differences in  $\alpha$  to distinguish concavities from convexities.<sup>10</sup> We therefore derived measures to distinguish patterns that convey veridical shape based on amount and discreteness within a 6-deg wedge centered around the axis of maximum curvature.

The first pair of measures was based on the energy in the critical orientations. The energy in the amplitude spectrum of each Brodatz pattern within the 6-deg wedge centered on the axis of maximum curvature (gray wedge in Fig. 6) was measured in two ways. At each spatial frequency  $f$  within the wedge, we took either the maximum or the mean energy across the 6 deg ( $\theta_{\text{crit}}$ ) and weighted the resulting profile by the human contrast-sensitivity function<sup>15</sup>:

$$\mu(f) = \text{mean}(E_{f, \theta_{\text{crit}}}) * \text{CSF}_f, \quad (4)$$

$$m(f) = \text{max}(E_{f, \theta_{\text{crit}}}) * \text{CSF}_f. \quad (5)$$

The filtered energy was then summed over frequencies to yield a single value, in units of canonical contrast threshold, for each texture pattern. The two values were computed:

$$C_{\text{mean}} = \sum_f \mu(f), \quad (6)$$

$$C_{\text{max}} = \sum_f m(f). \quad (7)$$

In Fig. 7 the cumulative numbers of veridical (solid curves) and nonveridical (dashed curves) patterns from



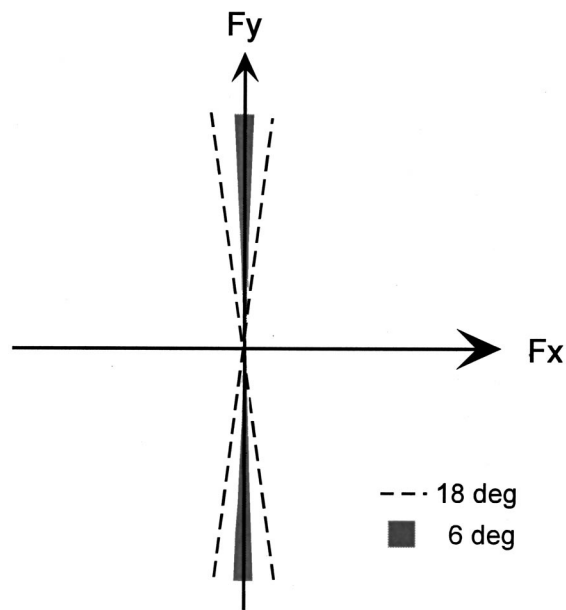


Fig. 6. Discreteness index defined as the ratio of the energy within the critical 6-deg wedge of orientations of the amplitude spectrum (gray) and the surrounding 18-deg wedge (dashed lines).

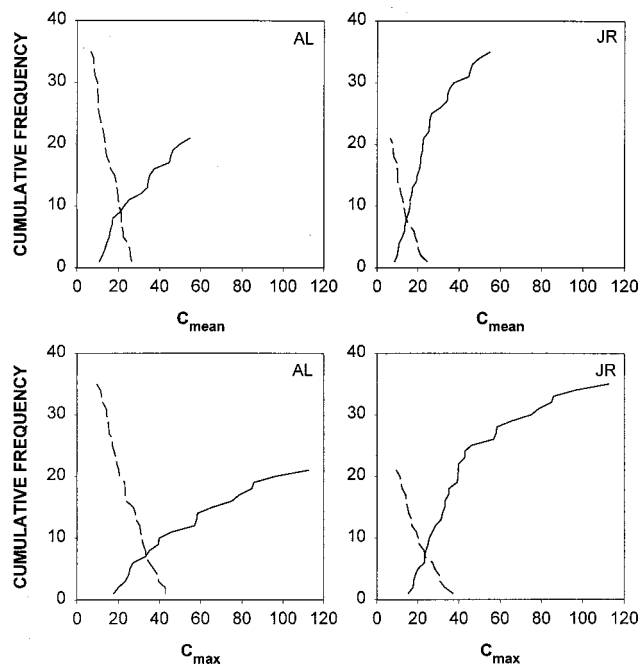


Fig. 7. Cumulative distributions for  $C_{\text{mean}}$  (top row) and  $C_{\text{max}}$  (bottom row) based on the classifications from Experiment 1. Distributions for observer AL are plotted on the left, for JR on the right.  $C_{\text{mean}}$  and  $C_{\text{max}}$  are in units of contrast threshold. For  $C_{\text{mean}}$ , the mean energy was computed at each spatial frequency within the slice. The resulting profile was filtered by the human CSF before summing. For  $C_{\text{max}}$ , the maximum energy was computed at each spatial frequency within the slice, before filtering with the human CSF and summing. For patterns that conveyed veridical shape (solid curves) values were accumulated from left to right. For patterns that did not convey veridical shape (dashed curves) values were accumulated from right to left.

Table 2. Minimum Number of Misclassified Patterns

Criterion Value	Experiment 1		Experiment 2	
	AL	JR	AL	JR
$C_{\text{mean}}$	11	15	18	11
$C_{\text{max}}$	10	13	18	9
$R_{\text{mean}}$	11	11	16	6
$R_{\text{max}}$	10	10	14	4

Experiment 1 are plotted versus  $C_{\text{mean}}$  and  $C_{\text{max}}$ . The cumulative frequency at each point on the solid curve represents the number of veridically perceived stimuli for which  $C_{\text{mean}}$  or  $C_{\text{max}}$  was less than or equal to the corresponding value along the ordinate.  $C_{\text{mean}}$  and  $C_{\text{max}}$  values for patterns that did not convey veridical shape were accumulated in decreasing order. The cumulative frequency for each point on the dashed curve represents the number of nonveridically perceived stimuli for which  $C_{\text{mean}}$  or  $C_{\text{max}}$  was greater than or equal to the corresponding value along the ordinate. These curves show the costs and benefits of using any value of  $C_{\text{mean}}$  or  $C_{\text{max}}$  to classify patterns as capable of conveying veridical shape or not. The abscissa of the solid curve at the criterion value indicates the number of patterns that will be incorrectly classified as veridical, and the abscissa value of the dashed curve indicates the number of patterns that will be incorrectly classified as nonveridical. The derived criterion value should be set to minimize the total costs associated with each type of minimization.

In the absence of different cost considerations for misclassification of veridical or nonveridical patterns, we chose the criterion that minimized the total number of misclassified patterns. The minimum number of misclassified patterns are tabulated by experiment and observer in Table 2. Criterion values of  $C_{\text{mean}}$  yielded 11 and 15 misclassified patterns for observers AL and JR, respectively, while criterion values of  $C_{\text{max}}$  yielded slightly fewer, 10 and 13 patterns respectively.

Similar cumulatives for Experiment 2 are shown in Fig. 8. Even though the psychophysical task and shape classification methods were different from those used in Experiment 1, and even though the perceived shapes differed for some patterns, the cumulatives for both  $C_{\text{max}}$  and  $C_{\text{mean}}$  for Experiment 2 are similar to those from Experiment 1. The criterion values of  $C_{\text{mean}}$  yielded 18 and 11 misclassified patterns for observers AL and JR, respectively, and criterion values of  $C_{\text{max}}$  yielded 18 and 9 patterns, respectively.

The difference in shapes conveyed by the isotropic pattern in Fig. 1E versus the patterns in Figs. 1A and 1B indicates that the discreteness of the energy in the critical orientation with respect to a local neighborhood of orientations may also be a factor. In two additional measures, we used a discreteness index to weight the critical energy. The discreteness index  $D$  is defined for each spatial frequency  $f$  as the ratio of the energy in the critical 6-deg wedge to the energy in a surrounding 18-deg wedge (Fig. 6).  $D$  for a particular spatial frequency  $f$  is thus the ratio of summed energies at that spatial frequency for the

6-deg wedge ( $\theta_{\text{crit}}$ ) with respect to the summed energies for the surrounding 18-deg wedge ( $\theta_{\text{sur}}$ ):

$$D_f = \frac{\sum_{\theta_{\text{crit}}} E_{f,\theta}}{\sum_{\theta_{\text{sur}}} E_{f,\theta}}. \quad (8)$$

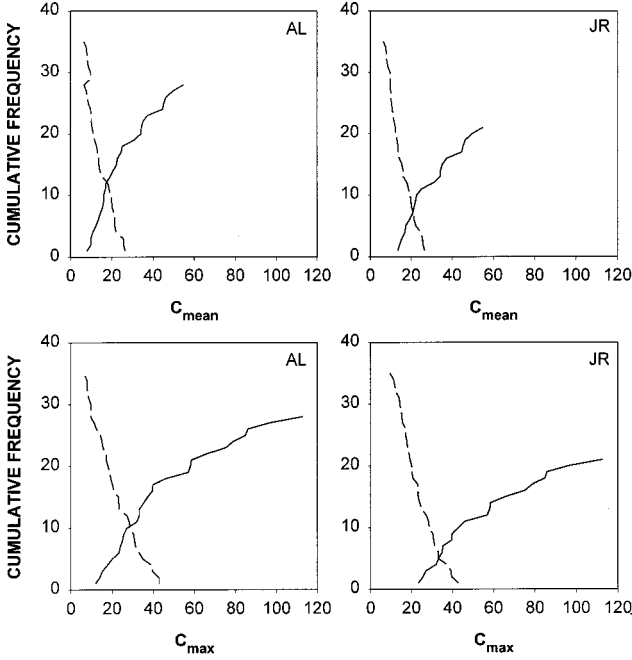


Fig. 8. Cumulative distributions of  $C_{\text{max}}$  and  $C_{\text{mean}}$  based on the classifications from Experiment 2. See Fig. 7 for details.

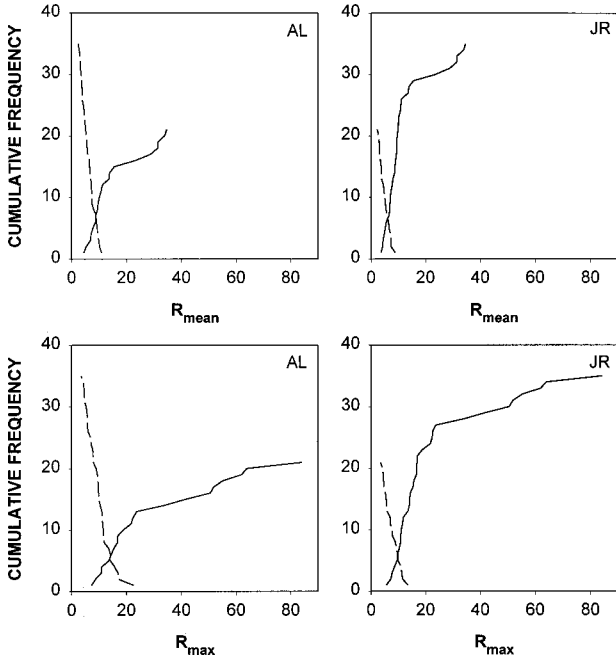


Fig. 9. Cumulative distributions for  $R_{\text{max}}$  and  $R_{\text{mean}}$  for Experiment 1. The maximum or the mean at each frequency was weighted by a discreteness index (the ratio of energy in the critical 6-deg wedge to the energy in the larger 18-deg wedge) before being filtered with the human CSF and summed.

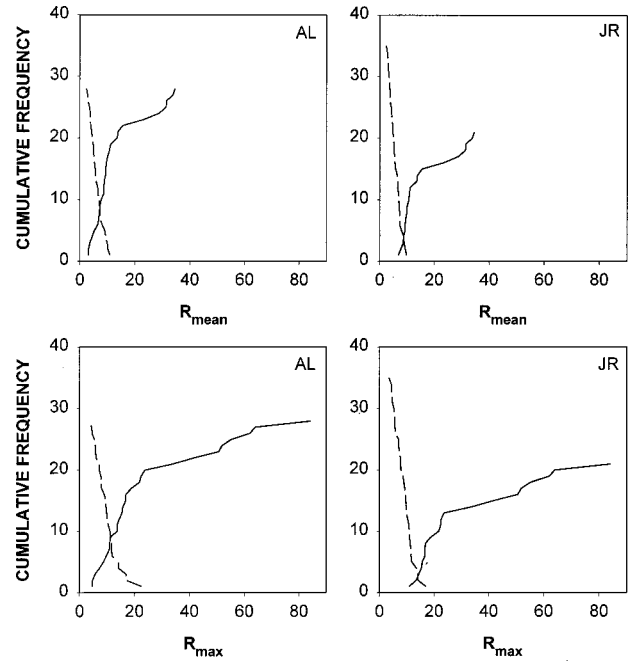


Fig. 10. Cumulative distributions of  $R_{\text{max}}$  and  $R_{\text{mean}}$  based on the classifications from Experiment 2. See Fig. 9 for details.

For an isotropic pattern like Fig. 1E,  $D_f$  will be equal to 0.33 at each frequency. For discrete and dominant patterns like Figs. 1A and 1B,  $D_f$  will be greater than 0.33. For discrete or dominant noncritical patterns like Figs. 1C and 1D,  $D_f$  will be less than 0.33. The new measures were computed as

$$R_{\text{mean}} = \sum_f D_f \mu(f), \quad (9)$$

$$R_{\text{max}} = \sum_f D_f m(f). \quad (10)$$

Figure 9 shows the cumulative curves based on  $R_{\text{max}}$  and  $R_{\text{mean}}$  for Experiment 1.  $R_{\text{max}}$  and  $R_{\text{mean}}$  yielded the same number of misclassifications as  $C_{\text{max}}$  and  $C_{\text{mean}}$  for observer AL (11 patterns for  $R_{\text{mean}}$  and 10 patterns for  $R_{\text{max}}$ ) and fewer respective misclassifications for observer JR (11 patterns for  $R_{\text{mean}}$  and 10 patterns for  $R_{\text{max}}$ ). For Experiment 2 (Fig. 10),  $R_{\text{max}}$  and  $R_{\text{mean}}$  yielded fewer misclassifications than  $C_{\text{max}}$  and  $C_{\text{mean}}$  for both observers (16 patterns for  $R_{\text{mean}}$  and 14 patterns for  $R_{\text{max}}$  for observer AL, 6 patterns for  $R_{\text{mean}}$  and 4 patterns for  $R_{\text{max}}$  for observer JR). Thus  $R_{\text{max}}$  consistently yielded the fewest number of misclassifications. If it is necessary to choose patterns that convey veridical shape, judging from our observers' percepts, any pattern for which  $R_{\text{max}}$  exceeds 15 units will work.

## 5. DISCUSSION

The results of this paper show that texture patterns convey veridical 3-D shape if the magnitude of energy in a narrow wedge of orientations about the axis of maximum curvature is large relative to the energy in an encompass-

ing wider wedge. The span of the critical wedge will depend to some extent on the 3-D curvature of the surface and the viewing distance.

### A. Comparison between the Experimental Methods

Experiments 1 and 2 used local and global shape perception tasks, respectively. Experiment 1 involved the reconstruction of ordinal shape based on a series of local slant judgments, whereas Experiment 2 involved the direct judgment of concavity versus convexity. Comparisons of the numbers of patterns perceived veridically and nonveridically showed that a majority of the stimuli were perceived similarly in the two experiments: 87.5% of the patterns for AL and 75% for JR. Observer AL perceived more patterns as veridical in Experiment 2 and observer JR in Experiment 1. It is possible that some observers are better at extracting local slant while others are better at extracting global curvature.

### B. Materials That Convey Veridical Three-Dimensional Shape

To get an idea of how well different classes of materials convey veridical shape, we grouped the patterns as man-made versus natural (animal pelts, terrains, or solid surfaces). In Table 3 the total number of patterns in each category is listed for each category heading. The numbers of patterns in each category that conveyed veridical shape are listed for Experiment 1 and Experiment 2. Between 16 and 27 of the 40 man-made patterns conveyed veridical shape. Many man-made materials such as textiles contain gridlike patterns that can fall along lines of maximum and minimum surface curvature. When the pattern is tilted so that the seams follow nonprincipal lines of surface curvature, the projected image no longer conveys veridical shape. Only two to four of the eight animal pelt patterns conveyed veridical shape. The ones that did (e.g., D003, D022) contained visible contours or flows along projected lines of maximum curvature; those that did not (e.g., D036, D092, D093) contained either strong contours along nonprincipal lines of curvature or were isotropic. The four terrain and four solid-surface patterns were either isotropic or dominant along noncritical axes. Only one of these patterns conveyed veridical shape for only one observer in Experiment 2.

### C. Contrast Reversals and Changes in Scale

Contrast reversals had little or no effect on perceived shape. For the five pairs of patterns that were approximate contrast reversals of one another—D101/D102,

D103/D104, D109/D110, D111/D112, and D105/D106—the veridicality of the perceived shape agreed within each pair. This is not surprising since the spectral properties of contrast-reversed patterns are identical. Small changes in scale in relation to the shape of the surface also had little effect. For the three pairs of patterns that differed in magnifications—D016/D017 ( $4\times$ ), D032/D033 ( $1.5\times$ ), D066/D067 ( $0.33\times$ ), and D035/D036 ( $0.5\times$ )—a change in scale brought about only minor changes in the perceived shape without changing the veridicality. In only one case (D035/D036 for observer JR), the percept changed from nonveridical to veridical when the scale of the texture was made finer.

### D. Shape from Natural Textures

In our earlier work we used a small number of synthetic patterns to show that the amplitude spectra of patterns that conveyed veridical shape were either “discrete” or “dominant” along lines of maximum curvature. The study reported in the current paper extends this work to a large number of naturally occurring textures with complex amplitude and phase spectra. The results show that the shape information conveyed by a texture pattern is determined by its amplitude spectrum. In fact, the success of the  $R_{\max}$  measure confirms that the information is conveyed by a narrow range of orientations in the patterns. Our results suggest that the critical shape information can be directly processed through independent, band-limited, orientation-selective neurons and suggest that matched-filter paradigms<sup>16</sup> may provide a good model for extracting 3-D shape from 2-D texture cues.

This paper also provides a quantitative method for predicting whether a texture pattern will convey veridical 3-D shape when markings corresponding to lines of maximum curvature are visible. These results can be useful in the 3-D rendering of objects on computer screens. As an alternative to computing the lines of maximum curvature at every point on the surface,<sup>17</sup> our results suggest that a pattern containing a critical amount of discrete energy at a small number of different orientations will effectively convey the veridical 3-D shape of most surfaces.

*Note added in proof:* The results of this paper were obtained for upright surfaces. For corrugated surfaces pitched toward or away from the observer, Zaidi and Li<sup>18</sup> have recently shown that pairs of components oriented symmetrically away from the axis of maximum curvature will produce patterns of orientation modulations in the image that are critical for distinguishing concavities and convexities. As a result, for pitched corrugated surfaces, single oriented components, including those parallel to the axis of maximum curvature, will not be sufficient for veridical shape perception. The results of Zaidi and Li<sup>18</sup> showed that the critical pattern of orientation modulations that allows distinction of concavities from convexities matches the pattern that arises from a single component parallel to the axis of maximum curvature for upright corrugations. In addition, the shape and pitch of the surface can be used to predict which component pairs will combine to provide the critical orientation modulations. Thus for pitched surfaces, one can apply methods similar to those reported in this paper to determine whether the amplitude spectrum of a texture pat-

**Table 3. Numbers of Patterns in Each Material Category Conveying Veridical Shape**

Material Category	Experiment 1		Experiment 2	
	AL	JR	AL	JR
Man-made (40)	16	27	23	19
Natural				
Animal pelts (8)	2	4	4	7
Terrains (4)	0	0	1	0
Solid surfaces (4)	0	0	0	0



tern contains sufficient discrete energy in the relevant pairs of orientations to convey veridical shape.

## APPENDIX A

To determine the template that was most similar to the data, we first computed the difference between the array of local perceived slopes and the array of local slopes corresponding to each of the five possible templates. These error values were then normalized by the maximum possible error for that particular template. To simplify the explanation of maximum error, taking the veridical template as an example, the fixation locations in Fig. 3 were numbered from 1 through 12 from left to right. For each of locations 1–3 and 11–12, observers should respond for all five trials at each of these locations that the left test location is closer in depth than the right test location. That is, the slope of the surface at each of these locations should be +5. Measured perceived slopes at these locations can vary between –5 and +5, which yields a maximum error of 10 at each location. Similarly, the slope for locations 5–9 should be –5; that is, observers should respond for all five trials at each of these locations that the right test location is closer in depth than the left test location. This again yields a maximum error of 10 at each of these locations. At locations 4 and 10 the slope should be 0; that is, observers should always see the two flanking test locations as appearing at the same depth. Since the difference between veridical and observed responses can vary only between 0 and +5 or 0 and –5, the maximum error will be only 5 at these two locations. This yields a total maximum error of 110 for all locations for the veridical template. Similarly, the maximum error was 60 for the half-rectified and the flat templates and 100 for the fully rectified template.

The template that yielded the smallest normalized error was the template shape under which we classified the percept. If the data did not fit any of the templates (normalized error >0.25 for all templates), then the shape was classified as “other nonveridical.”

## ACKNOWLEDGMENTS

The authors thank Joseph Rappon for his time and diligence as an observer and Byung-Geun Khang and Fuzz Griffiths for helpful discussions of the manuscript. Parts of this work were presented at the ARVO '99 meeting in Ft. Lauderdale and at the ECVP '99 meeting in Trieste,

Italy. This work was supported by National Eye Institute grants NRSA EY06828 to A. Li and EY07556 and EY13312 to Q. Zaidi.

Address correspondence to Andrea Li, ali@sunyopt.edu.

## REFERENCES

1. J. J. Gibson, “The perception of visual surfaces,” *Am. J. Psychol.* **63**, 367–384 (1950).
2. A. Li and Q. Zaidi, “Perception of three-dimensional shape from texture is based on patterns of oriented energy,” *Vision Res.* **40**, 217–242 (2000).
3. B. G. Cummings, E. B. Johnston, and A. J. Parker, “Effects of different texture cues on curved surfaces viewed stereoscopically,” *Vision Res.* **33**, 827–838 (1993).
4. J. E. Cutting and R. T. Millard, “Three gradients and the perception of flat and curved surfaces,” *J. Exp. Psychol.* **113**, 196–216 (1984).
5. D. C. Knill, “Discrimination of planar surface slant from texture: human and ideal observers compared,” *Vision Res.* **38**, 1683–1711 (1998b).
6. D. C. Knill, “Surface orientation from texture: ideal observers, generic observers and the information content of texture cues,” *Vision Res.* **38**, 1655–1682 (1998a).
7. D. C. Knill, “Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture,” *Vision Res.* **38**, 2635–2656 (1998c).
8. J. T. Todd and R. A. Akerstrom, “Perception of three-dimensional form from patterns of optical texture,” *J. Exp. Psychol.* **113**, 221–224 (1987).
9. K. Sakai and L. H. Finkel, “Characterization of the spatial-frequency spectrum in the perception of shape from texture,” *J. Opt. Soc. Am. A* **12**, 1208–1224 (1993).
10. A. Li and Q. Zaidi, “Information limitations in perception of shape from texture,” *Vision Res.* **41**, 1519–1534 (2001).
11. K. A. Stevens, “The visual interpretation of surface contours,” *Artif. Intel.* **17**, 47–73 (1981).
12. D. C. Knill, “Contour into texture: information content of surface contours and texture flow,” *J. Opt. Soc. Am. A* **18**, 12–35 (2001).
13. P. Brodatz, *Textures: A Photographic Album for Artists and Designers* (Dover, New York, 1966).
14. A. R. Rao and G. L. Lohse, “Towards a texture naming system: identifying relevant dimensions of texture,” *Vision Res.* **36**, 1649–1669 (1996).
15. J. G. Robson, “Spatial and temporal contrast-sensitivity functions of the visual system,” *J. Opt. Soc. Am.* **56**, 1141–1142 (1966).
16. Q. Zaidi and A. Li, “Neural model of shape from texture: developable surfaces,” *Invest. Ophthalmol. Visual Sci.* **41**, S219 (2000).
17. V. Interrante, “Illustrating surface shape in volume data via principal direction-driven 3D line integral convolution,” presented at the SIGGRAPH conference, Los Angeles, California, August 3–7, 1997.
18. A. Li and Q. Zaidi, “Limitations on shape information provided by texture cues,” *Vision Res.* (to be published).