



Limitations on shape information provided by texture cues

Qasim Zaidi *, Andrea Li

SUNY College of Optometry, 33 West 42nd Street, New York, NY 10036, USA

Received 20 July 2001; received in revised form 20 August 2001

Abstract

This paper uses visual, empirical and formal methods (Li & Zaidi, *Vision Research*, 40 (2000) 217; Li & Zaidi, *Vision Research*, 41 (22) (2001a) 2927) to examine the roles of oriented texture components in conveying veridical percepts of concave and convex surfaces that are pitched towards or away from the observer. The results show that pairs of components, oriented symmetrically around the axis of maximum curvature, combine to provide the geodesic orientation modulations that are critical for veridical shape perception. The degree of pitch determines the orientations of the critical pair of components. Perspective is crucial to the veridical perception of concavities and convexities, regardless of the degree of pitch. The results of this paper reconfirm that veridical shape perception depends on extracting critical patterns of oriented energy, but also show that the class of textures capable of conveying veridical percepts of developable shapes in general views is even more restricted than that identified by Li and Zaidi (*Journal of Optical Society of America A*, 18 (2001b), 2430). © 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Shape; Texture; Contour; Perspective; Orientation; Symmetry

1. Background

Explicit knowledge of the capacity of surface markings to convey shape and depth is at least as old as the Renaissance discovery of vanishing-point perspective. Paintings as early as Lorenzetti's Annunciation of 1344 used perspectival renderings of surface markings to enhance the three-dimensionality of objects and ground planes. Locally correct perspective cues, foreshortening, and the use of curved strokes and cross-hatchings to convey 3-D curvatures can be found in much older art, e.g. Greek vases, Hellenistic and Greco-Roman mosaics and wall paintings (Panofsky, 1927), and the 5th Century paintings at Ajanta, India (Spink, 1994; Behl, 1998).

Gibson's (1950) seminal contribution was to use the mathematical language of gradients to bring attention to the shape and depth cues in images that are provided by systematic variations in surface markings. After Gibson, the Visual Perception community exclusively followed the gradient model in studying shape-from-texture (e.g. Braunstein & Payne, 1969; Vickers, 1971; Rosinski & Levine, 1976; Cutting & Millard, 1984; Todd & Aker-

strom, 1987; Blake, Bulthoff, & Sheinberg, 1993; Cummings, Johnston, & Parker, 1993; Knill, 1998a,b,c). These studies culminated in sophisticated ideal observer analyses indicating the primacy over size and density cues of the foreshortening of texture elements (Blake et al., 1993; Knill, 1998c). In much of the perception literature, 'texture' has consisted of randomly-repeated discrete elements like polka-dots or squares. Since most natural textures do not consist of discrete elements, the computer vision literature generalized the concept of textures to statistically homogeneous surface markings (Garding, 1992; Malik & Rosenholtz, 1997; DeBonet, 1997; Zhu, Wu, & Mumford, 1997; Haralick, 1979; Rosenfeld, 1971), and explored models for extracting 3-D shape from cues such as texture flows (Hel & Zucker, 1989; Knill, 2001), spatial frequency modulations (Turner, Gerstein, & Bajcsy, 1991; Sakai & Finkel, 1993), and affine deformations of the texture pattern (Aloimonos, 1988; Garding, 1992; Malik & Rosenholtz, 1997).

Since Gibson's work, a lot has been learned about early neural processing of spatial patterns (Graham, 1989). In particular, we know that the first stage of cortical processing involves orientation and spatial frequency selective filters (Shapley & Lennie, 1985, Lennie, 1998). Our aim (Li & Zaidi, 2000, 2001a,b; Zaidi & Li, 2000) is to understand the extraction of 3-D shape from

* Corresponding author. Tel.: +1-212-780-5142; fax: +1-212-780-5137.

E-mail address: qz@sunyopt.edu (Q. Zaidi).

texture cues in neural terms. To this end, we have considered shape-from-texture in terms of the spatial primitives of frequency and orientation. Given the power of Fourier analysis and synthesis (Campbell & Robson, 1968; Bracewell, 1995), we have looked at texture patterns as composed of oriented sinusoidal gratings, and used sinusoidal corrugations as our simulated 3-D shapes. This has led us to some new findings and, more importantly, to new analytic methods that provide fresh insights into the issues. In particular, the unexpected discovery that certain classes of texture patterns do not convey veridical shape enabled us to go beyond gradient- or transform-based analyses of the complete texture pattern, to shape analyses based on patterns of oriented energy or contours created by specific components of the texture pattern.

The shape of an object can be regarded as variation in orientation across the surface of the object. This variation in surface orientation produces variations in surface markings across the projected image. Li and Zaidi (2000) showed empirically that, for upright corrugations, texture patterns that did not contain a discrete component parallel to the axis of maximum surface curvature did not qualify even minimally as being able to convey qualitatively veridical shapes because based on variations in these patterns simulated concavities looked similar to simulated convexities, and the zero-crossings in the slanted portions of the corrugation were seen as the extremum of concave segments. In other words the shapes looked qualitatively distorted from veridical. Fig. 1 provides illustrations. The texture in Fig. 1A (octotropic plaid) contains eight components, each consisting of compound sinusoidal gratings. The components are oriented 22.5° apart and include a horizontal component parallel to the axis of maximum curvature. These components form geodesics of the corrugation. The projected image conveys simulated concavities as well as convexities. The texture pattern in Fig. 1B is identical to Fig. 1A, except that it is missing the horizontal component, and though the simulated 3-D corrugation is identical to that in Fig. 1A, observers

cannot perceive both concavities and convexities from the information provided by the seven remaining geodesics. Note that Fig. 1B contains texture gradients, texture flows, frequency modulations, and affine deformations caused by the corrugation, yet does not convey a veridical percept. The secret to conveying veridical shape seems to be the existence of orientation modulations of the critical horizontal component which are visible in Fig. 1A as sparse wavy contours. Such orientation modulations could be carried by luminance defined contours (Fig. 1C), streaks (Fig. 1D), or contrast defined contours (Fig. 1E). Oriented energy could be extracted from contrast modulations through one of the class of non-linear second-order processes that have been proposed for extracting contrast modulations (e.g. Heeger, 1992; Sperling, Chubb, Solomon, & Lu, 1994; Mareshal & Baker, 1998; Schofield & Georgeson, 1999). Frequency modulations exhibited by other geodesics convey curvature or slants, but not the signs, i.e. concavity versus convexity or left versus right slant. These results translated shape from texture into the spatial primitives of orientation and frequency modulations.

Li and Zaidi (2001a) performed a mathematical analysis of shape from texture cues and showed that the results above were a direct consequence of the shape information contained in various oriented components of the texture pattern. In particular, for upright sinusoidal corrugations, only components oriented within a few degrees of the axis of maximum curvature exhibited orientation changes that could distinguish concavities from convexities and right from left slants. Other components exhibited orientation and frequency changes that were different for curvatures than for slants, but were similar for different signs of curvature and for different signs of slant. Together the two sorts of information led to stable 3-D shape percepts. An analysis of this sort can reveal if there is sufficient information to accomplish a visual task. Lack of such information precludes successful accomplishment of the task. Whether such information is used successfully by a particular

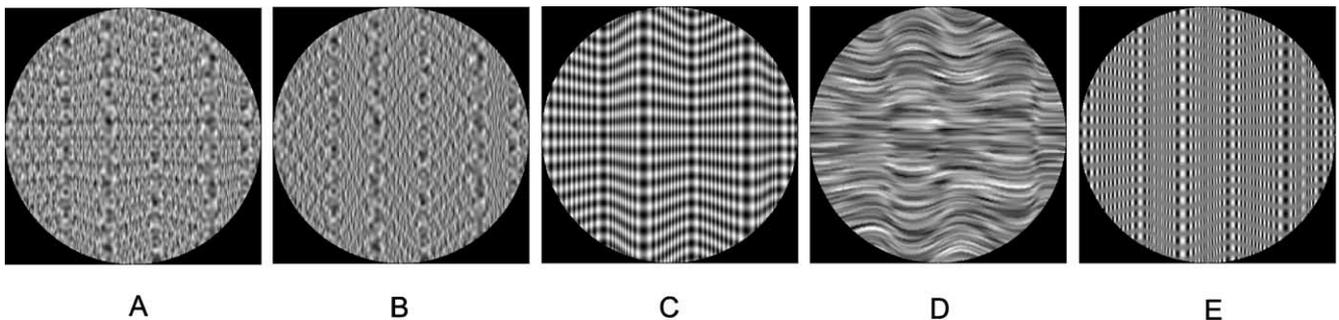


Fig. 1. Perspective projections of simulated sinusoidal corrugations with a zero crossing in the center, convexity to the left and concavity to the right. Pre-corrugation textures: (A) octotropic plaid; (B) octotropic plaid minus the component parallel to the axis of maximum curvature; (C) horizontal-vertical plaid; (D) random streaks oriented along the axis of maximum curvature; (E) vertical grating modulated in contrast along the horizontal axis. Patterns do not qualify as conveying veridical shape if the convex and concave portions are not perceived in the correct locations in the images.

visual system is a question to be settled by psychophysical experiments.

The purpose of this paper is to apply the methods of Li and Zaidi (2000, 2001a) to the case of corrugations that are pitched out of the fronto-parallel plane. We will show visually, empirically and formally that component pairs oriented symmetrically around the axis of maximum curvature combine to facilitate veridical shape perception as the surface is pitched from upright to horizontal, and thus in general no single geodesic is sufficient. For complex textures, this analysis can predict which textures convey concavities and convexities at each pitch and which do not. In Section 5, the results of this analytic approach enable us to refute the objections to our work raised by Todd and Oomes (2002), to show that the class of textures that convey veridical shape is much more restricted than claimed by Todd and Oomes, and to explain the importance of perspective and the issue of generic versus informative stimulus conditions in shape-from-texture.

2. Experiment 1: component contributions to percepts of pitched concavities and convexities

We begin by applying the visual methods of Li and Zaidi (2000) and the experimental methods of Li and Zaidi (2001a) to the case of corrugations that are pitched out of the fronto-parallel plane. Fig. 2 shows half-cycles of simulated 3-D corrugations. When presented correctly on a fronto-planar screen, with the center set at eye-height, each image simulates a sinusoidal concavity (CCV) or convexity (CVX) projected in perspective. The distance of the image plane d was 100 cm, the wavelength of the corrugation was 15.4 cm, the peak to trough amplitude was 20 cm, and each image spanned 4.5 by 9° of visual angle. The numbers along the bottom indicate the angle of rotation (pitch in degrees) of the surface around the horizontal axis at eye-level. (For clarity of expression, “deg” will indicate the pitch of the corrugation, while “ $^\circ$ ” will indicate the orientation of the texture component.) 0 deg is upright, a positive value

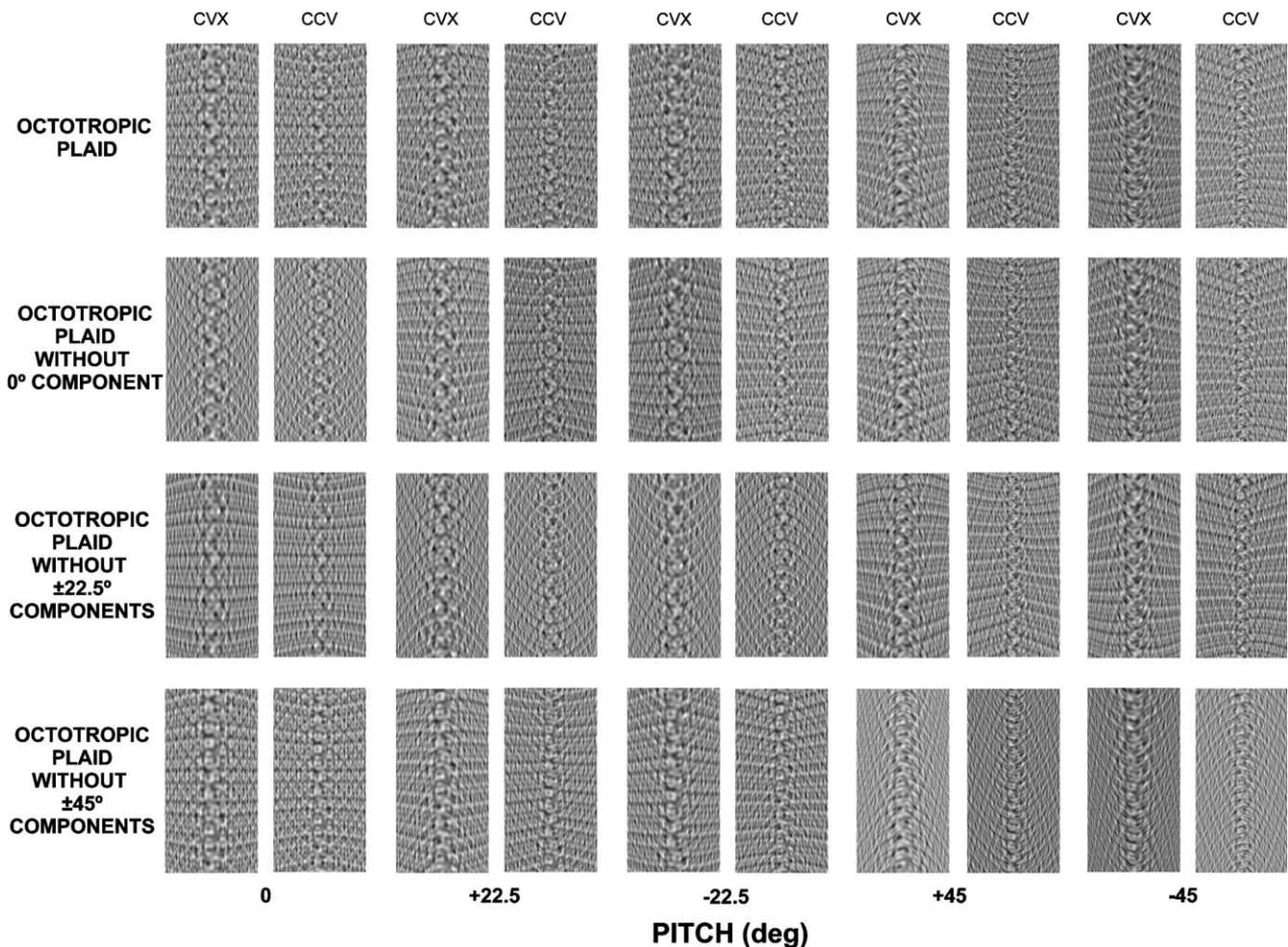


Fig. 2. Half-cycles of simulated 3-D corrugations (CVX = convex, CCV = concave), Columns show surfaces pitched at 0 deg, +22.5 deg, -22.5 deg, +45 deg, -45 deg from upright. First row shows octotropic plaid textures, while in the other rows selected components of the octotropic plaid have been removed. These images were used in Experiment 1.

of pitch indicates that the top of the surface is further from the observer, whereas a negative pitch indicates that the top is inclined towards the observer. All the pairs of images in Fig. 2 simulate a convexity (left) and a concavity (right). In this experiment, texture patterns for which the percepts of both of these shapes are not stable and veridical will be considered as not satisfying a minimal criterion for veridical shape perception. In Experiment 2, where slanted half-cycles of corrugation are used besides concave and convex half-cycles, a more general treatment of veridicality will be presented.

In the experiment, observers were shown a single simulated concavity or convexity for 0.5 s and asked to use switches to classify the percept as convex, concave, or other. All the surfaces from Fig. 2, the top row of Figs. 4 and 11 were randomly interleaved in one experiment. Viewing was monocular in a dark room, with the observer's head stabilized in a chin rest to set eye-height equal to the center of the screen. The results are tabulated in Table 1 for two naïve observers and one of the authors. Each entry shows the proportion correct (over 10 trials). Proportions greater than or equal to 0.8 are in bold to signify veridical reports. Starred values indicate which simulated concavities were reliably (≥ 0.8) but incorrectly perceived as convexities. Convexities were never reliably misperceived as concavities.

In all the images in the top row of Fig. 2, the texture pattern is an octotropic plaid and Table 1 shows that the

minimal criterion for conveying veridical shape is met for all the shown pitches. As noted in our earlier papers, the percept of the corrugation is more triangular than sinusoidal. The depth or shallowness of the percept changes with the nature of the surface markings and we intend to treat these perceptual changes in a later publication. An additional distortion in the percept that should be noted is the gross underestimation of pitch for both positive and negative rotations. An octotropic plaid thus seems to us to be capable of conveying only qualitatively veridical percepts of shape and pose of the surface.

We now strip away components of the octotropic plaid to show how different components contribute to shape percepts at different poses. In the second row of Fig. 2, the component parallel to the axis of maximum curvature (0°) has been removed from the octotropic plaid. This pattern conveyed veridical percepts at pitches of $\pm 22.5^\circ$ and $\pm 45^\circ$, but upright (pitch = 0°) concavities were reported as convexities (Table 1). In the third row, the $\pm 22.5^\circ$ components have been removed from the texture pattern. Veridical percepts are conveyed at pitches of 0° and $\pm 45^\circ$, but at $\pm 22.5^\circ$ pitches concavities are always misperceived as convexities. In the fourth row, the removal of the $\pm 45^\circ$ components leaves unchanged the shape carrying capacity at pitches of 0° and $\pm 22.5^\circ$, but makes concavities at pitches of $\pm 45^\circ$ appear convex. Taken together,

Table 1
Results of Experiment 1 for each observer

	0 deg		22.5 deg		-22.5 deg		45 deg		-45 deg	
	cvx	ccv								
<i>Observer LT</i>										
Octotropic plaid (Fig. 2)	1	0.7	1	1	1	0.9	1	0.9	1	0.9
Minus 0° (Fig. 2)	1	0*	1	0.9	1	1	1	1	1	0.6
Minus $\pm 22.5^\circ$ (Fig. 2)	1	1	1	0*	1	0.1*	1	1	1	1
Minus $\pm 45^\circ$ (Fig. 2)	1	1	1	0.8	1	0.9	1	0*	1	0*
$0^\circ + 90^\circ$ plaid (Fig. 11)	0.9	1	0.9	0*	1	0.4	1	0*	1	0.3
Parallel projection (Fig. 11)	0.7	0.3	0.8	0.6	0.7	0.4	0.6	0.4	0.5	0.2*
<i>Observer AC</i>										
Octotropic plaid (Fig. 2)	1	1	0.8	1	0.9	0.8	0.9	1	1	1
Minus 0° (Fig. 2)	0.9	0.2	1	1	0.9	0.7	0.9	1	1	0.9
Minus $\pm 22.5^\circ$ (Fig. 2)	1	0.9	1	0*	1	0.1*	1	1	1	1
Minus $\pm 45^\circ$ (Fig. 2)	0.9	0.9	1	1	1	1	1	0.3	1	0.2*
$0^\circ + 90^\circ$ plaids (Fig. 11)	1	1	1	0.3	1	0.2*	0.9	0.7	0.8	0.1*
Parallel projection (Fig. 11)	0.1	0	0.9	0.1	0.6	0	0.6	0.1	0.4	0*
<i>Observer AL</i>										
Octotropic plaid (Fig. 2)	1	0.9	1	0.9	1	1	1	0.9	1	0.9
Minus 0° (Fig. 2)	1	0*	1	0.9	1	0.9	1	0.7	1	0.9
Minus $\pm 22.5^\circ$ (Fig. 2)	1	1	1	0.1*	1	0*	1	0.9	1	0.7
Minus $\pm 45^\circ$ (Fig. 2)	1	1	1	0.7	1	0.7	1	0*	1	0.2*
$0^\circ + 90^\circ$ plaid (Fig. 11)	1	1	1	0*	1	0*	1	0*	1	0*
Parallel projection (Fig. 11)	0	0	1	0.2	0.7	0.1	0.5	0.2*	0.5	0.1*

For each texture pattern, proportion correct shape identification over ten trials is shown for concave and convex shapes at various degrees of pitch from upright. Proportions greater than or equal to 0.8 are in bold to signify veridical reports. Starred values indicate which simulated concavities were reliably (≥ 0.8) but incorrectly perceived as convexities.

Fig. 2 shows that different pairs of texture components are critical for conveying shape information at different pitches.

Fig. 2 demonstrates that no component of a texture provides sufficient information for veridical shape perception at all poses of a corrugated surface. For this corrugated surface, oriented components (of the uncorrugated pattern) seem to provide maximal information at a pitch equal to the orientation of the component. In Fig. 2, notice that all the images that convey veridical percepts have approximate mirror symmetry around the horizontal and vertical mid-lines, whereas the images that do not convey veridical percepts have translational symmetry around the horizontal mid-line.

3. Contributions of single and symmetric pairs of geodesics to percepts of pitched concavities and convexities

In this section we show how the critical orientation modulations in Fig. 2 are synthesized from geodesics of either a single direction parallel to the axis of maximum curvature or pairs of directions symmetric around this axis.

3.1. Single components

Fig. 3 illustrates simulated half-cycles of corrugations containing concavities (CCV) and convexities (CVX) in the upright and pitched poses indicated on the bottom, using texture patterns consisting of single gratings with pre-corrugated orientations of 0° , 90° , $\pm 22.5^\circ$ and $\pm 45^\circ$ (0° is parallel to the axis of maximum curvature). These images are also perspective projections of geodesics that are parallel on the surface of the corrugation.

The first row of image pairs in Fig. 3 illustrates simulated half-cycles for the 0° component. Each of the first pair of images (upright corrugations, pitch = 0) exhibits bilateral symmetry around the vertical mid-line. This bilateral symmetry conveys percepts of symmetric surface curvature. The pattern of orientations for convexities is qualitatively different from that for concavities, thus providing information for making shape distinctions. Both images are mirror symmetric around the horizontal mid-line, but contours in the concave image are bowed towards the mid-line whereas contours in the convex image are bowed away from the mid-line. As was noted with reference to Fig. 2, this second mirror symmetry seems to be crucial for distinguishing between concave and convex surfaces. The data in Li and Zaidi (2001a) showed that observers successfully extracted veridical shape from the information provided by the symmetries in the orientation modulations of the horizontal component, i.e. they not only distinguished slants from extrema of curvature, but also distinguished con-

cave from convex and right slanted from left slanted surfaces.

The second and third pair of images in the first row of Fig. 3 show the corrugations of the first pair pitched at $+22.5^\circ$ and -22.5° , respectively. The images exhibit more compelling depth at these pitches than in the upright case, but the perceived signs of the surface curvatures appear unstable, and the concavities appear convex more frequently than concave. All four images are mirror symmetric around the vertical mid-line, but have approximate translational symmetry around the horizontal mid-line as the contours are almost parallel and either all bowed upwards or all bowed downwards. In addition, the curves for a convexity at $+22.5^\circ$ pitch are similar in orientation pattern to the curves for a concavity at -22.5° pitch, and curves for a concavity at $+22.5^\circ$ pitch are similar in orientation pattern to the curves for a convexity at -22.5° pitch. This indicates that these geodesics do not provide sufficient information to disambiguate sign of pitch from sign of curvature. The same ambiguity characterizes the fourth and fifth pairs of images at pitches equal to $+45^\circ$ and -45° , respectively.

The bias for perceiving both convex and concave surfaces as convex is explainable in terms of the information in the images. As will be formally shown in Section 4, for surfaces of different slants at a constant distance from the observer, spatial frequency in the image is lowest at fronto-parallel slants, and increases monotonically with both positive and negative slants. Hence slant-caused spatial frequency modulations do not provide information about the sign of slant. In addition though, in perspective projection, for a constant slant, spatial frequency in images increases as a function of distance from the observer. For example, in these simulations, the spatial frequencies at the peaks of the convexities are lower than at the peaks of the concavities because the zero crossings were set at the same distance from the observer. In the absence of unambiguous information from orientation modulations, a bias for perceiving convexities would result if spatial frequency modulations were being used as cues to distance from the observer, i.e. the fronto-parallel segments of both convex and concave surfaces would be inferred as nearest to the observer based on the lowest spatial frequencies at fronto-parallel slant in the image of each surface.

Across all the images in the second row, the orientation for the pre-corrugated vertical component (90°) is almost vertical, and hence provides no information about 3-D shape. The minor convergence at the top or bottom solely indicates the pitch. On the other hand, the vertical component exhibits spatial frequency changes in the images. Convex and concave curvatures both exhibit the lowest spatial frequencies at fronto-parallel slants and show similar increases in spatial frequency as a

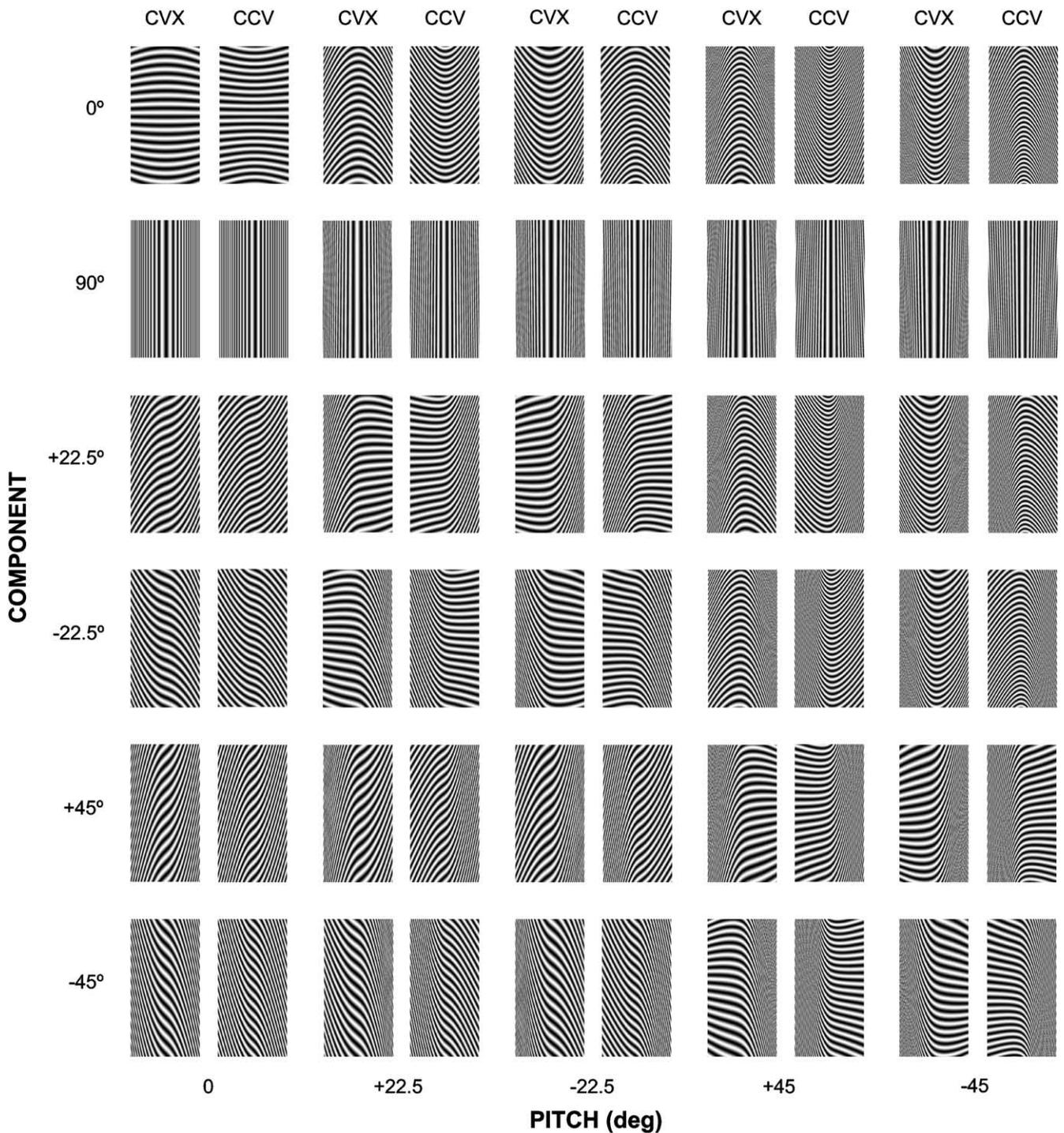


Fig. 3. Half-cycles of simulated 3-D corrugations pitched from the upright similar to Fig. 2. Texture patterns consisted of single gratings with pre-corrugation orientations of 0° , 90° , $\pm 22.5^\circ$, $\pm 45^\circ$ (0° = axis of maximum curvature).

function of the absolute value of slant. Therefore these frequency modulations provide information about the presence of local surface curvature versus slanted or flat fronto-parallel segments, but do not provide sufficient information to distinguish a concavity from a convexity. The use of a distance-based interpretation of frequency modulations would again lead to both concave and

convex surfaces being perceived as convex. The results of Li and Zaidi (2001a) showed that when observers extract curvature from the frequency modulations of the vertical component they almost always identify it as a convexity.

For the $+22.5^\circ$ and -22.5° components (rows 3 and 4 in Fig. 3), images for upright convex surfaces are similar

to images for upright concave surfaces, i.e. the orientation and/or frequency modulations exhibited by these components support inferences of curvature but cannot be used to distinguish concavities from convexities (Li & Zaidi, 2001a). In the second and third pairs of images at ± 22.5 deg pitch, the $+22.5^\circ$ and -22.5° components convey vague percepts of convexities and concavities, and the perceived surfaces seem to consist of ridges rather than symmetric curvatures. The contours do not have mirror symmetry around the vertical mid-lines. Stevens (1981) and Knill (2001) have remarked that rulings along geodesics not on the axis of maximum curvature convey distorted surface percepts. However, notice the marked similarity between the orientation patterns in the lower spatial frequency halves of these images and the corresponding halves of the images for the same curvatures for the 0° component at 0 deg pitch. The $\pm 22.5^\circ$ components at ± 45 deg pitches exhibit roughly parallel contours that again confound sign of pitch and curvature. The $+45^\circ$ and -45° components also do not by themselves convey stable percepts of symmetric convex and concave surfaces, but notice the similarity between the images at ± 45 deg pitches to the images for $\pm 22.5^\circ$ components at ± 22.5 deg pitches.

3.2. Pairs of components

In Fig. 4, we have used pairs of components of the octotropic plaid to demonstrate their contribution to 3-D shape percepts at different pitches. The top row shows pitched corrugations for a combination of 0° and 90° components. These images were used in Experiment 1, and the results are shown in Table 1. The percepts were veridical only at the upright 0 deg pitch. At pitches of ± 22.5 deg and ± 45 deg, the images show possibly even more compelling percepts of corrugations than in the upright case, but the data show that concavities are not distinguished from convexities. The contours contributed by the 0° component exhibit orientation modulations that are almost parallel through the whole image, and the images for convexities at positive pitch are similar to the images for concavities at equal negative pitch and vice versa. This indicates that it will be difficult to disambiguate sign of pitch from sign of curvature, unless there is ancillary information about one or both. For example, pitch information could potentially be provided by the convergence or divergence of the 90° component due to distance, but for the bas-relief regime studied in this paper, such information is obviously not adequate. For larger visual angles, disambiguating

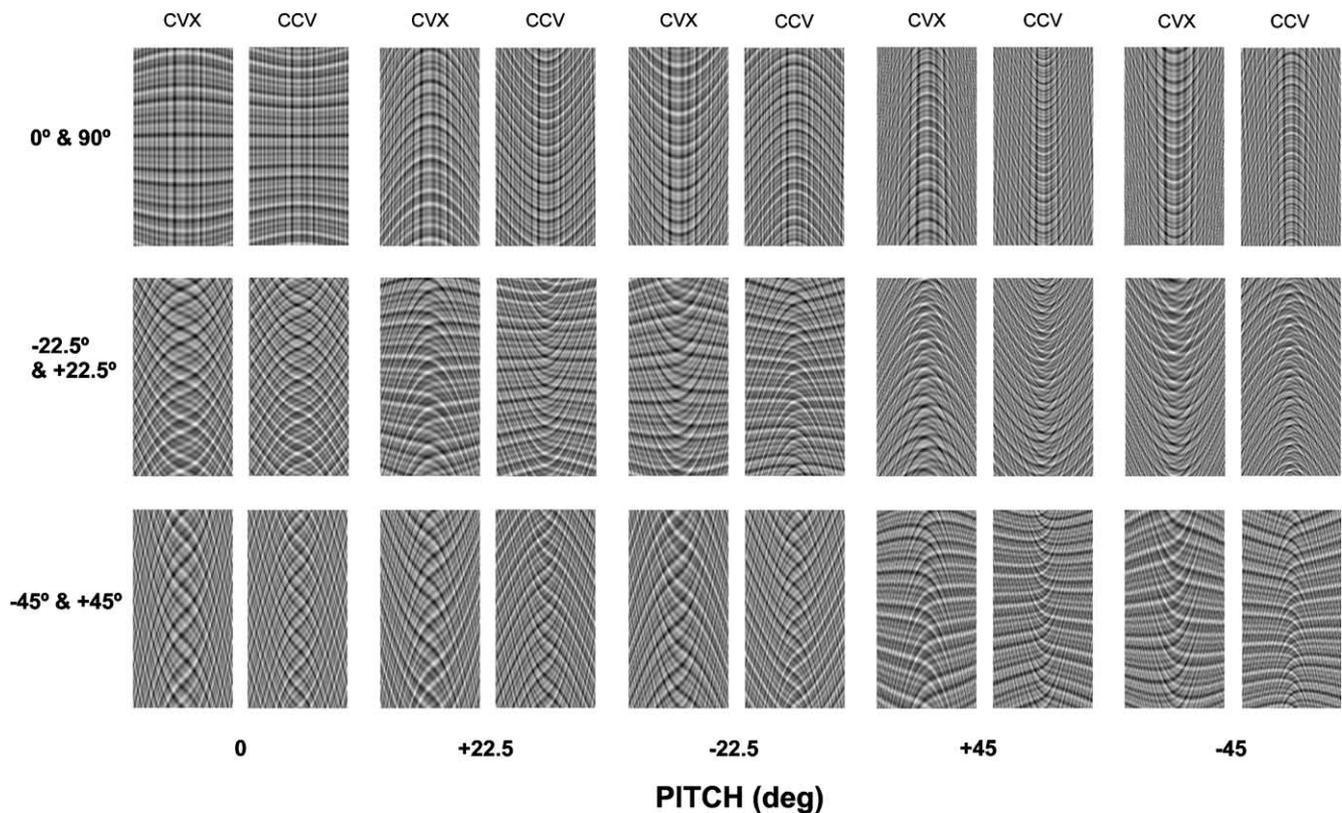


Fig. 4. Half-cycles of simulated 3-D corrugations pitched from the upright similar to Fig. 2. Texture patterns consisted of pairs of components of the octotropic plaid. Images in the top row were used in Experiment 1.

information could be provided by the 90° component due to the greater amount of perspective.

The second row of Fig. 4 shows corrugations using texture patterns made by combining components at +22.5° and –22.5°. Observers had difficulty making surface-based judgments on those images that contained discontinuities near the vertical mid-line, but at pitches of ±22.5 deg, notice the similar patterns of symmetry of the contours between these images and the upright images in the top row (ignore the center vertical sections of the images). Combining the ±22.5° components results in four distinct sets of orientation modulations that can potentially distinguish convexity from concavity, and positive pitch from negative. Fig. 2 showed that these components are critical for veridical percepts at pitches of +22.5° and –22.5°.

The texture pattern in the bottom row combines +45° and –45° components. These components are critical for veridical shape at ±45 deg pitches in Fig. 2, and in Fig. 4 at these and only these pitches, the patterns of symmetry of the contours of the ±45° combination are similar to the upright images in the top row. Fig. 4 shows that there are similar orientation modulations in the images of pairs of geodesics at the pitches at which they contribute to stable veridical percepts, despite the fact that these modulations are being conveyed by different components of the texture pattern.

4. Information analysis of orientation and frequency modulations caused by surface slant

In this section, we derive and graph expressions for frequency and orientation modulations in perspective images as a function of surface slants at constant distance from the observer. For different degrees of pitch, we will separately analyze patterns of orientation and frequency modulations contributed by texture components in isolation and in symmetric pairs. These derivations generalize the methods of Li and Zaidi (2001a) to pitches other than upright, and enable us to analyze the shape information provided by oriented texture components. This section provides a systematic mathematical counterpart of the visual analysis in Section 3. The visual analysis resumes in Section 5.

The values of local orientation and spatial frequency in each image were obtained from expressions derived in Appendix A. The image plane is assumed to be at distance d from the eye. The component oriented ω radians from the horizontal in the pre-corrugated pattern is overlaid on the vertical axis at a height y_0 with respect to eye-height onto a surface with a local slant of θ radians. The surface is rigidly pitched α radians around the horizontal axis. The local orientation in the image is given by:

$$\tan^{-1} \left(\frac{d \cos \alpha \sin \omega + d \sin \alpha \sin \theta \cos \omega - y_0 \sin \theta \cos \omega}{d \cos \theta \cos \omega - y_0 \sin \alpha \cos \theta \cos \omega} \right) \quad (1)$$

The expression for local spatial frequency is too long to include here, but the derivation is straightforward as shown in the Appendix A.

The graphs in Fig. 5 show local orientations and spatial frequencies of the 0° (open symbols) and 90° (closed symbols) components as a function of local corrugation slant around the vertical axis for perspective projections with the same distance parameter and range of slants as Figs. 2–4. The pitch of the corrugation in each row is indicated on the left. Curves are plotted for three horizontal slices through the corrugations corresponding to the following heights in the upright image: upward pointing triangles represent $y_0 = +4.4$, downward pointing triangles represent $y_0 = -4.4$, and circles represent $y_0 = 0^\circ$ of visual angle with respect to eye height.

For the horizontal component in the upright case, substituting $\alpha = 0$ and $\omega = 0$ in Eq. (1) gives:

$$\tan^{-1} \left(\frac{-y_0 \tan \theta}{d} \right) \quad (2)$$

For constant y_0 and d , orientation in the image is a monotonic function of slant of the surface, i.e. for equal positive and negative slants, orientations will be equal but of opposite signs. This results in rotational symmetry of order 2 which is shown by the open symbols in the top orientation graphs (pitch = 0). This function obviously also implies that a convexity, which is a traverse from negative to positive slants, will result in image orientations that are a mirror image of a traverse from positive to negative slants, i.e. a concavity, e.g. compare the curves formed by upward pointing open triangles in the convex and concave panels. This difference in the orientation patterns is potential information for distinguishing between concavities and convexities. In addition in Eq. (2), for a constant d and θ , orientations in the image will be equal and opposite for $+y_0$ and $-y_0$, i.e. for points equally above and below eye-height. Within each orientation panel, curves formed by upward pointing open triangles are mirror symmetric around the horizontal mid-line to the curves formed by downward pointing open triangles. As noted earlier, this second mirror symmetry provides the critical information for distinguishing between concave and convex surfaces.

For the 0° component at arbitrary pitch, Eq. (1) simplifies to:

$$\tan^{-1} \left(\frac{\tan \theta (d \sin \alpha - y_0)}{d - y_0 \sin \alpha} \right) \quad (3)$$

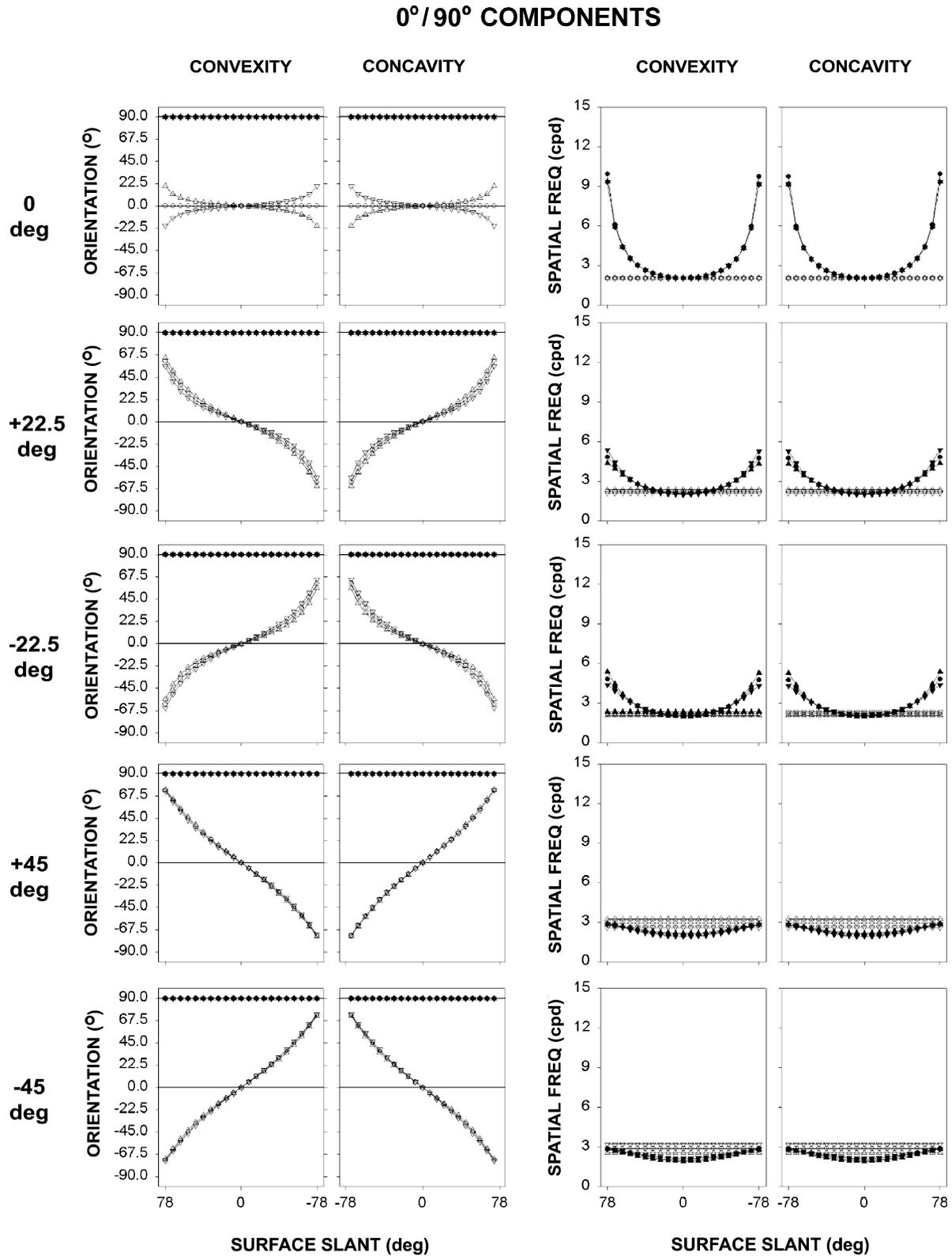


Fig. 5. Local orientations and spatial frequencies at each pitch as a linear function of corrugation slant. See text for derivations. Components are 90° (closed symbols) and 0° (open symbols).

When slant is upright, Eq. (3) reduces to Eq. (2), and since for our conditions $y_0 \ll d$, orientation is a damped

monotonic function of slant θ . In Eq. (3), as α tends to $\pi/2$, orientation tends to the same value as slant, and

height in the image (y_0) becomes less of a factor. In Fig. 5, for pitches of +22.5 deg, -22.5 deg, 45 deg and -45 deg, each of the orientation curves is approximately rotationally symmetric of order 2, and the orientation modulations are of larger amplitude for the pitched than for the upright surfaces, but the curves vary little by height in the image. Even though at any particular pitch, the curves for concavity are approximately mirror images of the curves for convexity, the curves for convexity at positive pitch are similar to curves for concavity at negative pitch, and the curves for concavity at positive pitch are similar to curves for convexity at negative pitch, thus there is a confounding of information for sign of curvature with information for sign of pitch. The 0° component thus cannot provide sufficient information to distinguish concave from convex curvatures for pitched corrugations.

From Eq. (1) the orientations for the pre-corrugated 90° component are derived as vertical across images for all pitches, and hence provide no information about 3-D shape. These values appear as overlapping filled symbols for the three image heights in Fig. 5. On the other hand, as shown in the graphs on the right, the vertical component exhibits spatial frequency changes in the image. Convex and concave curvatures both exhibit the lowest spatial frequencies at fronto-parallel slants (slant=0 deg) and show similar increases in spatial frequency for positive and negative slants. Therefore these frequency modulations provide information about the presence of surface curvature but do not provide sufficient information to distinguish a concavity from a convexity. As stated in Section 3, if these slant-caused frequency modulations are interpreted as distance-caused frequency modulations, both concave and convex surfaces will be perceived as convex.

Fig. 6 shows the derived local orientations and spatial frequencies for the $\pm 22.5^\circ$ components in the same format as Fig. 5. For the upright case ($\alpha=0$ deg), the local orientation for a component oriented at ω is given by:

$$\tan^{-1} \left(\frac{d \sin \omega - y_0 \sin \theta \cos \omega}{d \cos \theta \cos \omega} \right) \quad (4)$$

For fronto-parallel slants, when $\theta=0$ is substituted into Eq. (4), the local orientation in the image is derived as equal to ω , the pre-corrugated orientation of the texture component. This can help locate the curves for each component. The graphs for upright pitch show that the local orientation and frequency curves show systematic effects of surface slants that are similar for convex and concave surfaces and across heights in the image, thus the orientation and/or frequency modulations exhibited by these components can support inferences of curvature, but cannot be used to distin-

guish concavities from convexities. In the +22.5 deg pitch concavity panel, the lower spatial frequency section of the +22.5° component corresponds to negative slants (-78 to 0) and for these slants the upward and downward pointing triangles (filled) form curves that have similar shapes to the upward and downward pointing triangles (open) for the same slants for the concave 0° component at 0 deg pitch. Similarly for the low frequency slants of the -22.5° component (0 to 78), the upward and downward pointing triangles (open) form curves that have similar shapes to the upward and downward pointing triangles (open) for the same slants for the concave 0° component at 0 deg pitch. For each of the four curvature/pitch pairs, i.e. convexity/+22.5, concavity/+22.5, convexity/-22.5, concavity/-22.5, combining the low frequency portions of the $\pm 22.5^\circ$ component images results in four distinct sets of orientation modulations that are similar to the orientation modulations of the 0° component at 0 deg pitch, and thus provide sufficient information to distinguish convexity from concavity, and positive pitch from negative.

Fig. 7 shows the derived local orientations and spatial frequencies for the $\pm 45^\circ$ components in the same format as Fig. 5. In rows 4 and 5 representing pitches of ± 45 deg, the orientation graphs for the $\pm 45^\circ$ components are similar to the graphs for the $\pm 22.5^\circ$ components at pitches of ± 22.5 deg in Fig. 6. Therefore, combining the low frequency portions of the $\pm 45^\circ$ components for each of the four curvature/pitch pairs (i.e. convexity/+45, concavity/+45, convexity/-45, concavity/-45) also results in four distinct sets of orientation modulations that provide sufficient information to distinguish convexity from concavity, and positive pitch from negative.

The analyses of this section show that the pair of geodesics that are perspective projections of the components distributed symmetrically at $\pm p^\circ$ around the axis of maximum curvature will provide sufficient information for qualitatively veridical shape perception at pitches of $\pm p$ deg, whereas singly these geodesics will not be sufficient. This analysis explains why texture patterns that contain particular discrete components convey veridical shape for particular pitches in Fig. 2, and why patterns missing these components do not. We cannot conceive of how any gradient-based model could make this prediction.

5. New and old ideas in shape from texture

In this section, we build on the results of the previous sections to bring into sharp relief the contrasts between our approach to shape-from-texture and the traditional approaches exemplified by Todd and Oomes

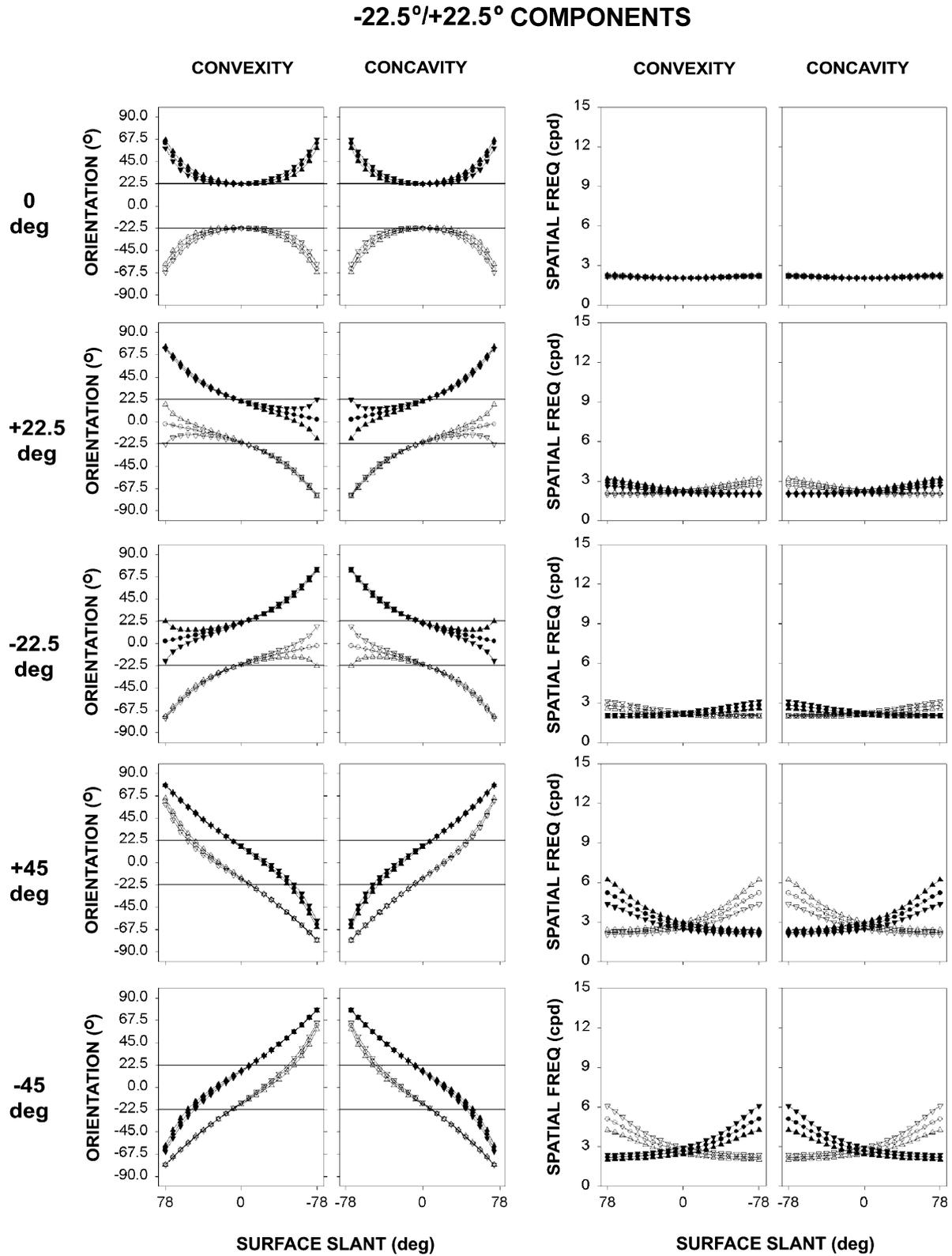


Fig. 6. Same as Fig. 5 for components 22.5° (closed symbols) and -22.5° (open symbols).

(2002). The differences begin at a fundamental level. Todd and Oomes present the following statement as a ‘theoretical explanation’: ‘What makes these images

appear 3-D is that the curvature of the image contours is perceptually attributed to the depicted surface’. In our view, this statement is at best a description of two

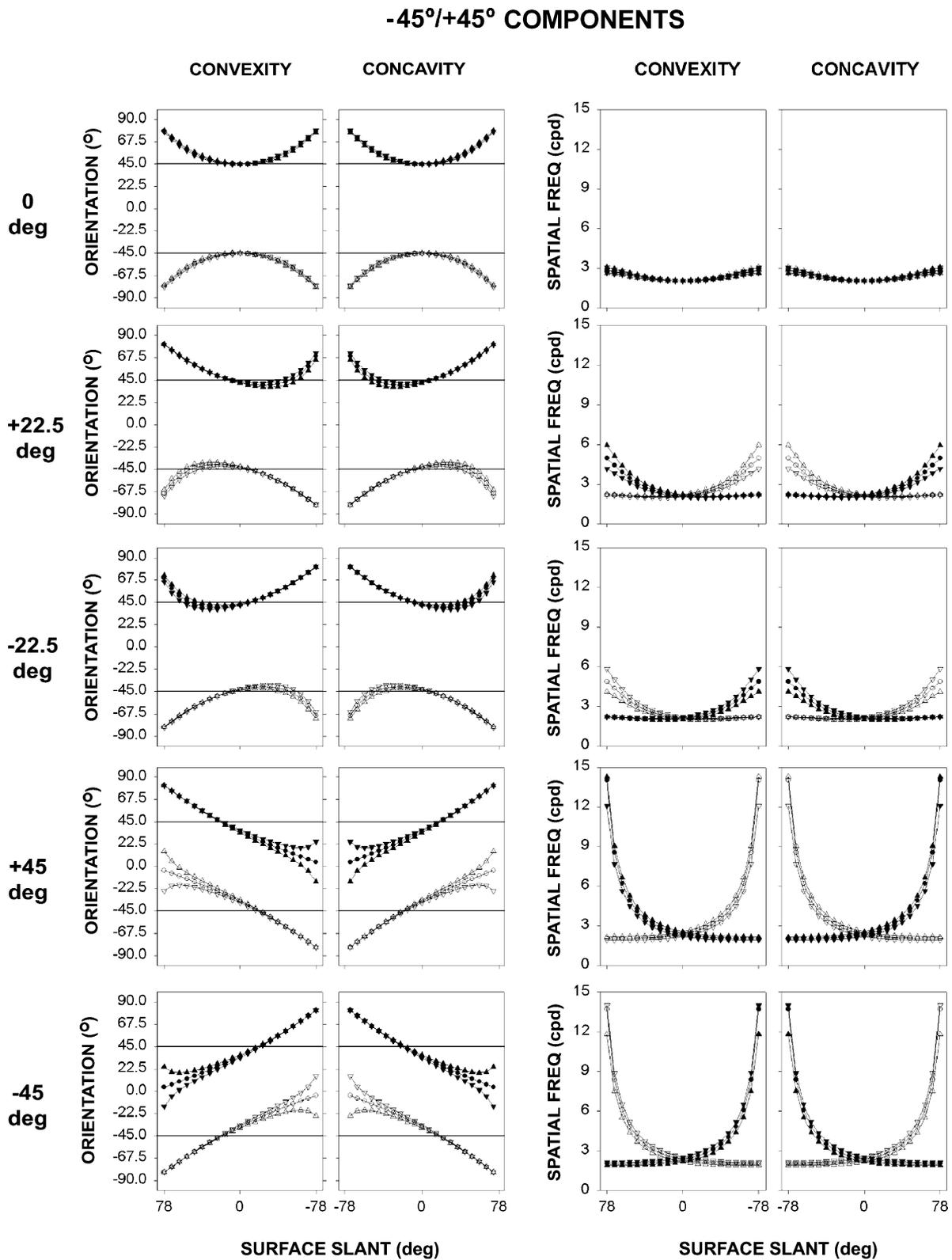


Fig. 7. Same as Fig. 5 for components 45° (closed symbols) and -45° (open symbols).

inseparable phenomenal experiences that arise from the same neural processes. Our aim is to understand how neurons extract 3-D shape from texture cues, and to

build a mechanistic model that can identify sign, depth, location, orientation, etc. of surface curvatures and slants.

5.1. Experiment 2: Developable surfaces and critical orientations of discrete energy

Developable surfaces are shapes that can be formed by bending a thin, flat material without stretching or tearing. Therefore, the texture pattern on the surface remains unchanged, and any image variations in the pattern are due to the projection. Developable surfaces thus provide a good starting place for studying shape-from-texture, and we have used non-twisted developable surfaces. Li and Zaidi (2000) pointed out that in studying shape from texture, it is imperative that both concavities and convexities be considered. In Figs. 7 and 10 of their paper, Todd and Oomes present ‘perceptually compelling’ convex cylinders made out of polka-dots

and lines oriented 45° to the axis of maximum curvature. From this demonstration they claim that these texture patterns support 3-D shape extraction despite not containing discrete energy along the axis of maximum curvature. Li and Zaidi (2000, 2001b) had already shown that many patterns that do not contain discrete energy along the axis of maximum curvature will convey convexities but will not convey concavities for upright corrugations.

In Fig. 8, we present perspective projections of four half-cycles of simulated corrugations: concave, convex, right and left slanted, and a fifth flat fronto-parallel surface for patterns consisting of polka-dots, a 45° grating, and a horizontal–vertical plaid. Images were computed for a sinusoidal corrugation with a wavelength

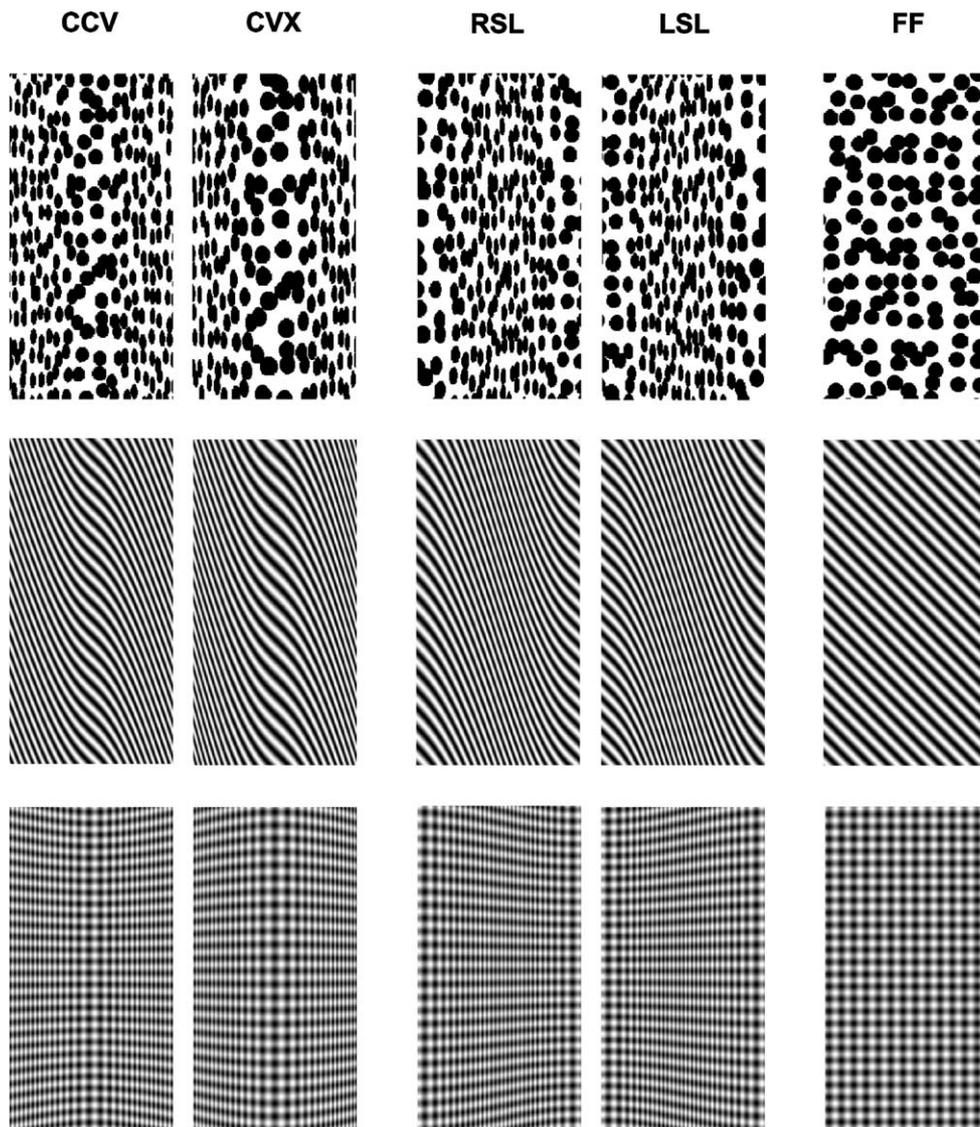


Fig. 8. Perspective projections of four half-cycles of simulated upright 3-D corrugations: concave (CCV), convex (CVX), right slanted (RSL), left slanted (LSL), and a flat fronto-parallel (FF) surface, for patterns consisting of polka-dots (top), a 45° grating (middle) and a horizontal–vertical plaid (bottom). The plaid conveys veridical percepts for all five surfaces, but the other two patterns do not. In particular, notice that the concave surfaces appear convex in the top two rows. These images were used in Experiment 2.

of 15.4 cm, peak to trough amplitude of 14 cm, and a viewing distance of 100 cm (similar width to depth ratio as the cylinders in Todd and Oomes' Figs. 7 and 10). These stimuli were presented randomly interleaved 20 times each for 1 s in a 5AFC experiment. Results for three naive observers and one of the authors are shown in Fig. 9. For each simulated shape, the area of the dot represents the frequency with which each perceived shape was reported. The largest dots in this figure represent 100%. For the polka-dots and the 45° grating patterns, observers perceived simulated concavities and convexities both to be convex, and also confused

right and left slants with concavities or flat fronto-parallel surfaces. In other words, simulated convexities for these two patterns may be perceptually compelling as shown by Todd and Oomes, but simulated concavities also look like convexities. These results can be contrasted in the same figures with the veridical results and percepts for the horizontal-vertical plaid pattern, which does contain discrete energy parallel to the axis of maximum curvature. It should be pointed out that Li and Zaidi (2001a) showed formally and empirically that for upright corrugations, a 45° component does not provide sufficient information to distin-

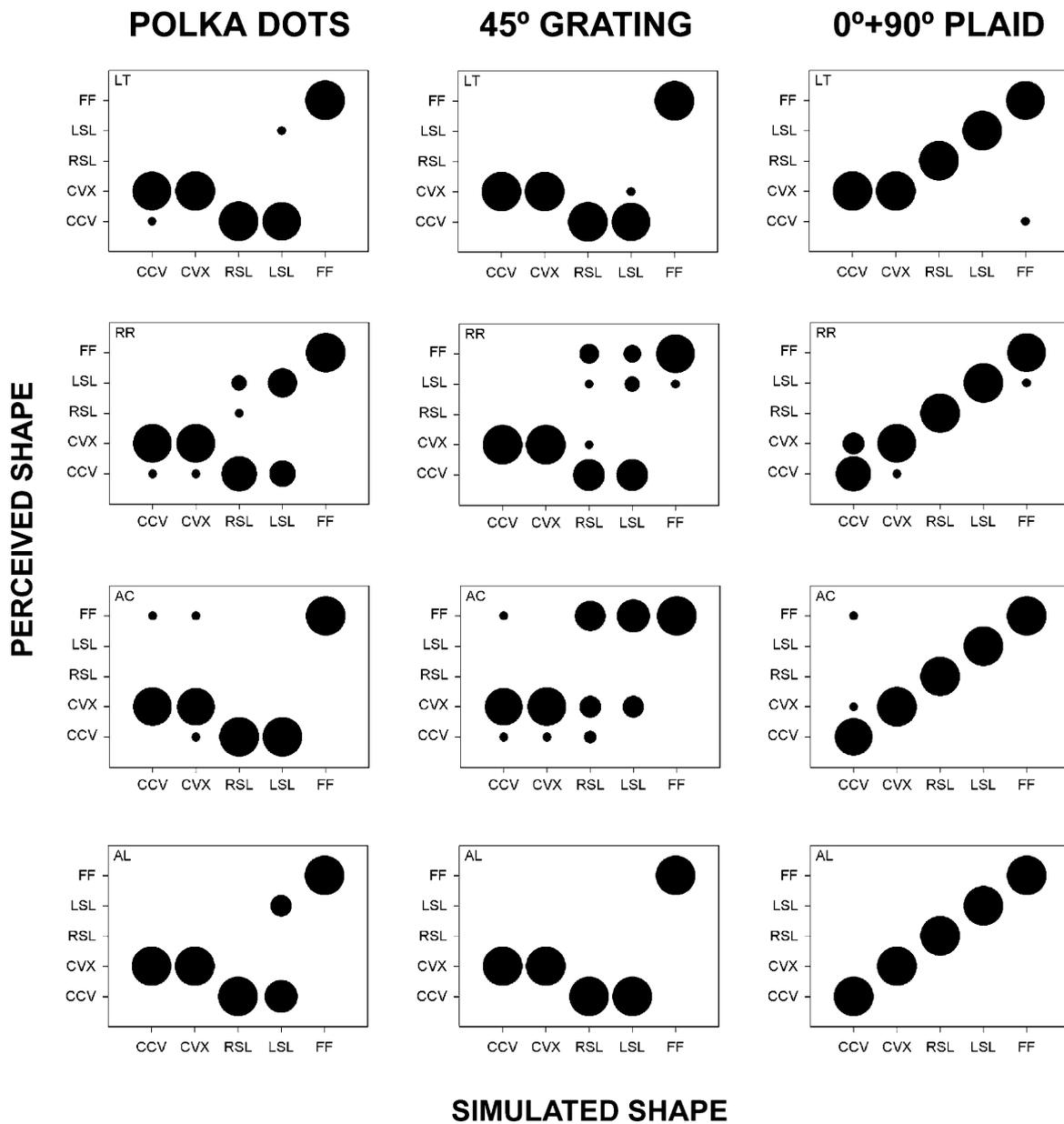


Fig. 9. Results of Experiment 2, separated by texture patterns and observers. Areas of dots represent frequencies with which each perceived shape was reported for each simulated shape. Dot areas sum to 100% frequency for each simulated shape.

guish concavities from convexities, or left slants from right.

The polka-dot images show variations in size along the horizontal axis. These variations are similar to the variations in the width of the bars of the vertical component in Fig. 3. Since spatial frequency is inversely proportional to the width of the bars, the spatial frequency graphs in Fig. 5 also represent the inverse of the size gradients. Spatial frequency modulations (or size changes) due to the slant of the surface do not provide sufficient information about the direction of slant, but for a constant slant, size decreases and spatial frequency increases in images as a function of distance from the observer. A bias for perceiving convexities would result if spatial frequency modulations were being used as cues to distance from the observer, i.e. the fronto-parallel segment of a convexity or a concavity would be inferred as being nearest to the observer based on the lowest spatial frequencies or largest element sizes at fronto-parallel slants. The right slanted and left slanted polka-dot surfaces in Fig. 8 exhibit spatial frequency variations that are opposite to the concave and convex surfaces, i.e. the highest spatial frequencies are in the center. If our suggestion about spatial frequency as an absolute cue to distance is correct, these surfaces should be perceived as concavities, and the results in Fig. 9 show that the percepts are in accord with our suggestion.

Todd and Oomes claim that the cylinders in their Figs. 7 and 10 look perceptually compelling because of gradual orientation changes relative to the viewing direction. Their figures actually show shapes that have gradual orientation changes in the middle but have extremely sharp slants at the edges of the cylinders. If just the edges of these cylinders are masked off, the remaining shape percept is almost flat. This indicates that the compelling part of the shape percept is actually supplied by the sharp slants (or discontinuities) at the edges, not the gradual orientation changes in the middle. In any case, compelling convexities can be displayed using polka-dots or 45° gratings even with sinusoidal corrugations as shown in Fig. 8. The point Li and Zaidi (2000, 2001b) made was that such patterns could not convey distinct upright concavities and convexities. Fig. 4 showed that by itself the 45° grating does not convey the veridical shape of the corrugation at any pitch. In Fig. 10 we show that even for perspective projections, at pitches other than upright, concavities made from polka-dot textures appear to be 'perceptually compelling' non-veridical convexities. Texture patterns with isotropic amplitude spectra are thus generally incapable of conveying veridical shape for surfaces that have both concavities and convexities, whereas patterns with discrete orientations in the amplitude spectrum, like the octotropic plaid, are generally capable of conveying qualitatively veridical shape percepts.

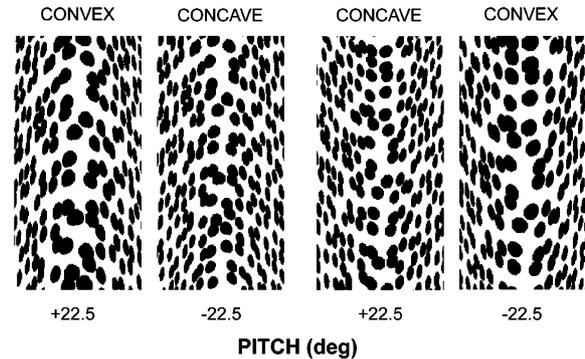


Fig. 10. Perspective projections of the simulated polka-dot concave and convex surfaces from Fig. 8, presented at pitches of ± 22.5 deg. Notice that at both $+22.5$ deg and -22.5 deg pitch, simulated concave surfaces appear perceptually compelling as convex surfaces.

5.2. Generic versus informative conditions

Todd and Oomes (2002) use the terms 'generic' and 'nongeneric' in their title and use the term 'degenerate' for the shapes and poses of our surfaces, but they neither define their terms nor provide any empirical evidence on the statistical frequency of stimuli and views that could bear on the issue of genericity. The term generic is not absolute, but relative to the context. For example, in scale space theory (Koenderink, 1984a), the term generic is applied to surfaces for which the features occur and change in a stable way, i.e. if the surface is slightly perturbed, the pattern in which the features evolve does not change under smoothing flows (Lu, Cao, & Mumford, 2001). Sinusoidal corrugations are generic surfaces under this definition. In an analogous fashion, in the computer vision literature, a nongeneric or accidental viewing position is defined as one which, if perturbed slightly, would reveal a different topological or differential structure of the scene (Binford, 1981). The genericity principle has been used to justify various rules for interpreting images (Lowe, 1985). For a visual system, these rules can be considered perceptual biases or priors that develop because of repeated exposure to generic views. There is certainly frequent exposure to upright shapes as these abound in the natural and man-made environment, e.g. cliffs and buildings. In addition, there is empirical evidence that in the absence of stereo information, observers tend to see pitched faces of irregular tetrahedra and hexahedra as upright, revealing a bias for assuming planes and edges to be fronto-parallel (Griffiths & Zaidi, 2000a).

For real objects, an accidental view is one that conceals features that are revealed by a slight perturbation of viewing position. Since textured, upright corrugations are veridically perceived for a large class of textures (Li & Zaidi, 2001b), it is difficult to conceive of these conditions as 'degenerate' for the domain of shape from texture. More importantly, Fig. 2 shows that at different

poses of the corrugation, different texture patterns convey veridical shape, and this depends systematically on the constituent components. Instead of viewing each pitch of the corrugation as ‘degenerate’ for the texture components that do not convey veridical shape, we think that it is more profitable to assess conditions in terms of informativeness in ruling out accidentally efficacious texture cues to shape. For example, the convex cylinders used by Todd and Oomes are very low on informativeness because they do not even reveal that polka-dot textures cannot convey veridical shape in general.

5.3. Critical role of perspective

We address two issues concerning perspective and shape-from-texture. The first issue concerns the role of perspective cues in shape-from-texture. Since the amount of perspective information is a function of visual angle, for any object, perspective cues are reduced at large distances from the observer. In Fig. 11, the first row of Fig. 2 is depicted for $d = \infty$, which yields an orthographic projection in which the projecting rays are parallel, and the image plane is perpendicular to the projecting rays (Foley, van Dam, Feiner, & Hughes, 1987; Mundy & Zisserman, 1992). The bottom row of each observer’s data set in Table 1 shows the empirical results for the perception of the orthographic projections depicted in Fig. 11. For the vast majority of conditions, observers do not report veridical percepts. Only in a few cases, and only for convexities, are the percentage correct greater than or equal to 0.8. This should be contrasted with the top row of the table for each observer, where percepts were almost universally veridical for the same textured shapes presented in perspective projection.

This issue can be separately analyzed for orientation and frequency modulations. In Eq. (1), if we substitute $d = \infty$, then the expression for local orientation reduces to:

$$\tan^{-1} \left(\frac{\cos \alpha \sin \omega + \sin \alpha \sin \theta \cos \omega}{\cos \theta \cos \omega} \right) \quad (5)$$

i.e. there is no y_0 component and the contours will be parallel across different heights in the image. As noted in Section 3, and as can be seen in Fig. 11, such parallel contours confound sign of curvature and sign of pitch: note the similarity of the convex image at +22.5 deg pitch to the concave image at –22.5 deg pitch, and the concave at +22.5 deg pitch to the convex at –22.5 deg pitch. In terms of frequency modulations, even in orthographic projection, local frequencies increase as slant increases on either side of fronto-parallel. However, the frequency changes are similar for concavities and convexities and thus cannot be used to distinguish between the two. In addition though, for constant slants, perspective causes increases in spatial frequency as a function of distance from the observer. Hence, if frequency modulations due to slant are interpreted in terms of the distance cue, concave surfaces would give illusions of convexity. This would be true for orthographically as well as perspective projected surfaces.

The secondary issue concerns perspective versus orthographic projection of stimuli for experiments. There is ample evidence that visual inferences are consistent with perspective projections, e.g. Shepard’s table illusion (Shepard, 1990), indicating that orthographic projections are likely to be interpreted as perspective projections of surfaces distorted from the veridical. In fact, Todd and Oomes complain that the perspective image of a corrugated horizontal grating (Fig. 5 of Li and Zaidi (2000); see also the pair of half-cycles in the top left of Fig. 3) appears bistable to them. Images that simulate perspective projections of a scene are calculated for a fixed eye position. At correct eye-positions, our observers reported veridical concavities and convexities in both local-depth and global-shape tasks. If visual inferences are based on perspective cues, percepts of the image of the 0° component corrugation will predictably change signs of curvature if the image is presented at incorrect distances, eye-heights, or slants. The

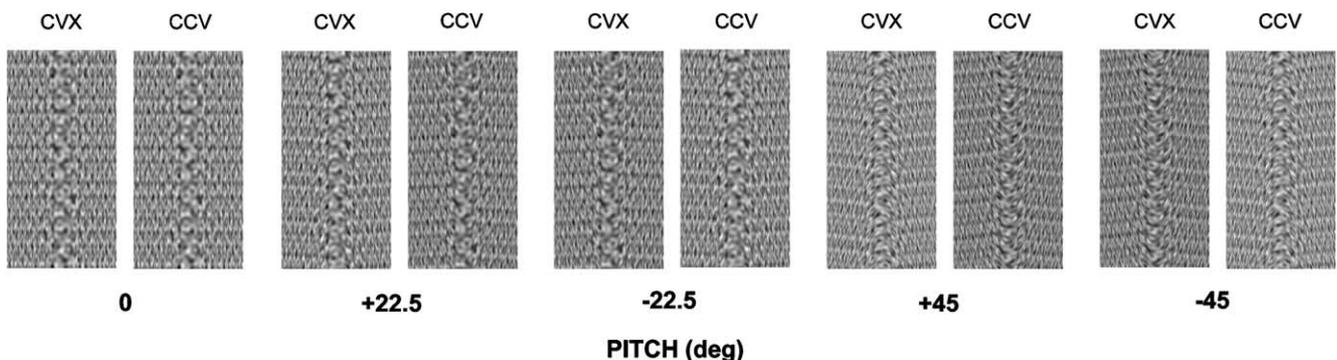


Fig. 11. Simulated surfaces from the top row of Fig. 2 presented in orthographic projection and used in Experiment 1. Notice that perspective projection (Fig. 2) conveys stable veridical percepts of concave and convex surfaces, but that orthographic projection (Fig. 11) does not.

distorting effects of incorrect eye-position for near perspective have been well known since Leonardo da Vinci's writings in the 15th–16th centuries (Panofsky, 1927).

Calculating orthographic projections used to be the computationally feasible compromise, but increases in computing power have made it very easy to calculate perspective projections, so there is no reason not to do so in studying the efficacy of texture cues to convey veridical shape.

5.4. Multiple curvatures in perspective and orthographic projection

The empirical results shown in this paper are for isolated concavities and convexities. The results of Experiment 1 showed that for simulated concavities and convexities for the $0^\circ + 90^\circ$ plaid, where the orientation modulations are contributed entirely by the 0° component, concavities are not seen reliably as concavities for the ± 22.5 deg pitch images. In fact, the percept is bistable for most observers. Is it possible that shape percepts are more likely to be veridical when multiple curvatures are present in an image, like in the figure from Stevens (1981) that Todd and Oomes used as their Fig. 11? Consider the top row of Fig. 12, where using the 0° component of the octotropic plaid, we have simulated perspective projections of corrugations with zero crossings in the center. Images in the left column simulate a convexity to the left and a concavity to the right; images in the right column simulate a concavity to the left and a convexity to the right. The left and right surfaces are pitched at $+30$ deg and -30 deg respectively and then rolled around the vertical axis by 15 deg. Veridical perception requires seeing one concavity to the right of one convexity in the left column, and seeing one convexity to the right of one concavity in the right column. To most observers the surfaces look like corrugations, but observers sometimes report a convexity and a concavity, and sometime two convexities. Even when a concavity and a convexity are reported, the perceived signs are either bistable or the opposite of veridical in the negative pitch case. The simulated concavities are especially perceptually bistable in sign. The orientation modulations in the images are almost parallel, and thus the bistability is due to the ambiguity between sign of curvature and sign of pitch that was discussed in Sections 3 and 4.

In the second row of Fig. 12, the above surfaces are depicted in orthographic projection. Notice that the shift from perspective to orthographic projection has a barely discernible effect on the orientation modulations – they are now perfectly parallel. Not surprisingly, the percepts have not changed much from the perspective projections, i.e. a corrugation can be perceived but it is unstable and most often reversed in phase for

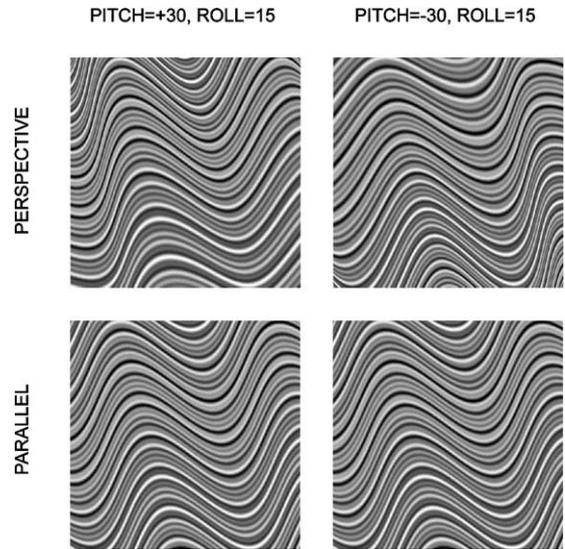


Fig. 12. Simulated corrugated surfaces for the 0° component of the octotropic plaid with zero crossings in the center. Left column images simulate a convexity to the left and a concavity to the right, right column images simulate a concavity to the left and a convexity to the right. Pitch and roll of surfaces are given at the top of columns. First row, perspective projections; second row, orthographic projections. Veridical perception requires seeing one concavity to the right of one convexity in the left column, and seeing one convexity to the right of one concavity in the right column.

the negative pitch. In contrast, where perspective projected images conveyed veridical percepts of concave and convex surfaces (Fig. 2 top row), orthographic projection removed the critical cues (Fig. 11). In other words, orthographic projection will not change the perception of images that contain roughly parallel orientation modulations in perspective projection, but these are images that do not reliably convey veridical percepts of concave and convex surfaces.

The figure from Stevens (1981) seems to consist of a 0° component bent into a sinusoidal corrugation, but with ambiguity as to sign of pitch and curvatures since it is similar to both of the pitched and rolled surfaces in the bottom row of Fig. 12. Todd and Oomes use Stevens' figure to argue that perspective projection is not necessary for texture cues to convey compelling percepts of shape. In Stevens' figure, if the curvature perceived as concave is isolated by masking the rest of the figure, it frequently is perceived as convex. In fact, the concavity can be made to switch signs of curvature just by attending to different portions of the figure. The cause of this bistability is explained in the previous paragraphs. Parenthetically, while trying to approximate the pitch and roll parameters used by Stevens, we discovered that concavities are considerably more perceptually bistable for many other pitch and roll parameters, and that texture cues are even less informative if a corrugation is first rolled and then pitched.

5.5. Silhouettes, occlusions and image junctions

Silhouettes, occlusions, and image junctions are very powerful cues to 3-D shape (e.g. Clowes, 1971; Koenderink & van Doorn, 1979; Sugihara, 1984; Koenderink, 1984b; Willats, 1992; Van Effeltherre, 1994; Koenderink, van Doorn, Christou, & Lappin, 1996; Albert, 1999; Tse, 1999; Albert & Tse, 2000; Rubin, 2001; Tse, 2002; Griffiths & Zaidi, 1998; Griffiths & Zaidi, 2000b; Zaidi, Spehar, & Shy, 1997), and might even be the primary cues in Todd and Oomes' Fig. 12. It is worth pointing out that Fig. 12 of Li and Zaidi (2001a) shows that volumetric (non-developable) concavities, convexities and saddles can all be depicted solely through orientation modulations without the aid of occlusions or junctions.

In our published work, we have purposely used stimuli that do not contain occlusions, silhouettes or image junctions so as to concentrate on texture cues to shape. However, Griffiths and Zaidi (2001) have examined the simple case of trapezoidal silhouettes of planar texture patterns. As Ames (1951) showed, fronto-parallel trapezoidal silhouettes are interpreted as slanted rectangles in monocular viewing. Inside the silhouettes, the texture patterns simulated slants that were either consistent with the perceived slant of the trapezoid, or with the opposite slant. Texture patterns that contained discrete energy oriented along the axis of slant, could convey signs of slants in isolation. Such patterns combined with consistent trapezoids to give vivid impressions of depth, but dissociated from inconsistent trapezoids to create the impression of a slanted surface viewed through an oppositely slanted window. Texture patterns that did not contain discrete energy oriented along the axis of slant could not convey the sign of slant by themselves. These patterns were captured by the silhouette and seen as slanted with the silhouette. These results show that silhouette and texture cues to shape combine or not combine, and when they combine the combination seems to be weighted by the informativeness of each cue (Landy & Maloney, 1995). Clearly, this analysis needs to be systematically extended to curved silhouettes, occluding contours, and texture discontinuities, before general cases can be understood.

5.6. Neural extraction of 3-D shape from texture cues

In our view, there is a basic and an applied reason to study shape-from-texture. The basic question concerns the extraction of 3-D shape from texture cues by neurons in the brain. Towards this aim, Zaidi and Li (2000) proposed a neural model consisting of filters matched to the orientation modulations of the component parallel to the axis of maximum curvature for upright corrugations. Fig. 13 shows a set of matched filters for developable concavities and convexities of a single depth

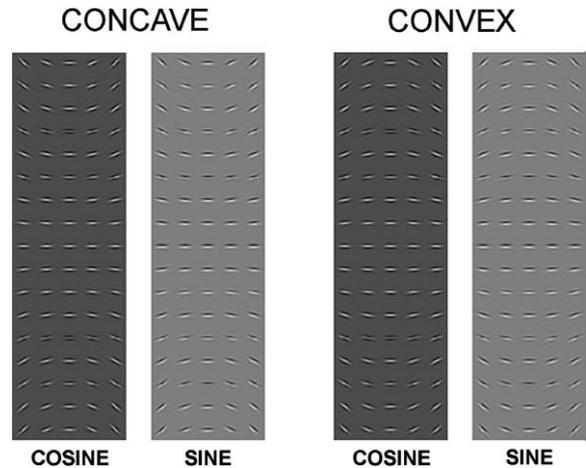


Fig. 13. Matched filters for identifying concave and convex surfaces. Filters are made up of feed-forward connections from V1-like orientation- and spatial frequency-tuned neurons simulated as Gabors in sine and cosine phase.

amplitude and orientation. Matched filters were made out of simple feed-forward connections from independent V1-like orientation-tuned neurons selective for the same peak spatial frequency. For graphical clarity, the filters in Fig. 13 show a sparse spatial sampling. The model had reasonable successes in identifying depths, locations and orientations of concavities and convexities, but was not perfect for simulated shapes that contained noisy versions of the critical information. We have continued to refine this model using lateral connections between the V1-like neural elements (Das & Gilbert, 1999; Li, 1998; Hess & Field, 1999), and will soon prepare it for publication. It is interesting that in Fig. 2, all the images that convey veridical concavities and convexities exhibit orientation modulations similar to the upright corrugation, irrespective of the components of the texture pattern that are supplying the critical modulations. This result suggests that even though our matched-filter model was designed for upright corrugations, it will also perform well on pitched corrugations. It also appears possible that a neural filter, matched to frequency modulations based on distance from the observer, will extract convex curvature for both convex and concave surfaces textured with patterns like polka-dots that have isotropic amplitude spectra, and thus will be compatible with psychophysical data (Figs. 8–10 and Li & Zaidi, 2001b).

5.7. Rendition of shapes using texture cues

An important application of shape-from-texture results is in the rendition of shapes in computer graphics (e.g. Peachey, 1985; Hanrahan & Hauberli, 1990; Arad & Elber, 1997; Interrante, 1997; Interrante & Kim, 2001). In medical imaging applications, e.g. computer

rendering of critical concavities like the insides of arteries, veridicality of the percept can be crucial, and a perceptually compelling non-veridical shape could cause problems. In choosing textures for these purposes it is important to use our strategy for ruling out classes of textures that do not convey veridical percepts for the complete range of required shapes and poses. For a texture pattern to be generally useful in conveying veridical shape, at the very least it should enable observers to distinguish between concavities and convexities. For developable surfaces, at a minimum, this rules out all textures not containing a discrete component parallel to the axis of maximum curvature (Li & Zaidi, 2000), notwithstanding the ability of such textures to convey compelling percepts of a few shapes in a few poses (Todd & Oomes, 2002). In addition, the results of this paper show that the class of generally useful texture patterns is even more restricted than that identified by Li and Zaidi (2000), and consists of textures that contain energy in a number of discrete orientations. Li and Zaidi (2001b) provided procedures to measure the magnitude and discreteness of oriented energy in the amplitude spectra of natural textures. For upright corrugations, we were concerned only with the orientation parallel to the axis of maximum curvature. For pitched concavities, similar procedures could be used to measure discrete energy along the orientations critical for the required pitches.

6. Summary

We have tried to introduce new analytic approaches to the study of shape-from-texture, and have progressively learned more about the issues with each succeeding paper. In our first paper (Li & Zaidi, 2000), we made the unexpected discovery that certain classes of texture patterns do not convey even ordinal veridical shape for upright corrugations, and that the critical information for shape is supplied only by discrete components parallel to the axis of maximum curvature. This paper showed the necessity of simulating concavities as well as convexities, and linked shape-from-texture to the shape-from-contour work of Stevens (1981). In fact, we thought that our results supported Steven's contention that observers assume that parallel surface markings follow lines of maximum curvature. In our second paper (Li & Zaidi, 2001a), we discovered that this assumption was not needed, since for upright corrugations only the orientation modulations of the component parallel to the axis of maximum curvature contain sufficient information to distinguish signs of curvatures or slants. In our third paper (Li & Zaidi, 2001b) we went considerably beyond the textures used in the shape-from-texture literature, and tested our ideas on a large set of everyday textures (Brodatz, 1966). We showed that the abilities of

textures to convey shape can essentially be predicted from their amplitude spectra, irrespective of their phase spectra, and that for upright corrugations, veridical shape is conveyed only by those textures that contain discrete energy within a narrow angle around the axis of maximum curvature.

In this paper, we have applied our methods to the case of corrugations pitched towards or away from the observer. This has led to the new result that different pairs of geodesics are necessary to convey veridical shape at different pitches of the corrugation, each pair arising from components symmetrically oriented around the axis of maximum curvature. This result supplements Stevens' (1981) result that shape is misperceived if the surface markings consist of only one component oriented along a geodesic other than the axis of maximum curvature. Whereas Li and Zaidi (2000) and Knill (2001) linked shape-from-texture to Stevens' work on shape from parallel geodesics, this paper goes beyond that by showing why textures with multiple discretely oriented components will in general convey veridical 3-D shape better than geodesics of any single orientation. However, the results of this paper also show that, contrary to the claims of Todd and Oomes (2002), the class of texture patterns capable of conveying veridical percepts of developable shapes in general views is even more restricted than that identified by Li and Zaidi (2000, 2001b). The analysis in the paper explains why perspective information is critical in the extraction of veridical shape-from-texture cues. We feel that in our papers we have barely scratched the surface of this area of inquiry. This paper shows some examples where shapes are perceived as distorted, and others where the curvatures are perceived as reversed or ambiguous. Future explorations of more general shapes and poses with our methods should lead to a hierarchy of shape and pose conveying properties of different texture types.

Acknowledgements

The authors wish to thank A. Fuzz Griffiths, Hal Sedgwick, Molly Aitken, David Knill and an anonymous reviewer for their constructive comments, and Long Tran, Rocco Robilotto and Andrew Chepaitis for serving as observers. This work was supported by NEI grant EY13312 to QZ.

Appendix A

In this appendix, we derive local orientation and spatial frequency in the perspective projections of oriented texture components for pitched corrugations. The

center of the image plane is defined as (0, 0, 0) in 3-D space coordinates, the surface normal to the image plane at that point intersects the observing eye at a distance d , i.e. (0, 0, 0) is at eye-height. We consider a line of unit length, with one point on the image plane at (0, y_0 , 0), i.e. y_0 units above eye-height. If the line were lying in the image plane at an angle of ω radians from the horizontal, the coordinates of the end point would be ($\cos \omega$, $y_0 + \sin \omega$, 0). If this line is slanted θ radians from the frontal plane along an upright corrugation, the coordinates of the end point would become ($\cos \theta \cos \omega$, $y_0 + \sin \omega \cos \theta \sin \alpha$, $\cos \omega \sin \theta \sin \alpha$). The perspective image (u , v) of any point (x , y , z) is calculated as:

$$u = \frac{ud}{z+d}$$

$$v = \frac{yd}{z+d}$$

In the perspective image, the line would extend from (0, y_0) to [$d \cos \theta \cos \omega / (\cos \omega \sin \theta + d)$, $(y_0 + \sin \omega) d / (\cos \omega \sin \theta + d)$]. If the corrugation is pitched backwards α radians through the horizontal eye-height line, the 3-D coordinates of the end-points of the line change to (0, $y_0 \cos \alpha - y_0 \sin \alpha$, 0) and ($\cos \theta \cos \omega$, $(y_0 + \sin \omega) \cos \alpha - \cos \omega \sin \theta \sin \alpha$, $\cos \omega \sin \theta \cos \alpha - (y_0 + \sin \omega) \sin \alpha$). In the perspective image, the line would extend from (u_0 , v_0) = [0, $d y_0 \cos \alpha / (-y_0 \sin \alpha + d)$] to

$$(u_1, v_1) = \left(\frac{d \cos \theta \cos \omega}{\cos \omega \sin \theta \cos \alpha + d - (y_0 + \sin \omega) \sin \alpha}, \right. \\ \left. \times \frac{d((y_0 + \sin \omega) \cos \alpha - \cos \omega \sin \theta \sin \alpha)}{\cos \omega \sin \theta \cos \alpha + d - (y_0 + \sin \omega) \sin \alpha} \right).$$

The slope of the line in the perspective image is calculated as: $(v_1 - v_0) / (u_1 - u_0)$, and its length as $\sqrt{(u_1 - u_0)^2 + (v_1 - v_0)^2}$. The slope of this line is used as the estimate for the local projected orientation of a texture component at angle ω from the horizontal, and is plotted versus θ in Figs. 5–7 for various fixed values of α . Changes in the length of this line as a function of θ and α provide an estimate of changes in local spatial frequency of the texture component oriented at $\omega + \pi/2$ radians, and is plotted versus θ in Figs. 5–7 for various fixed values of α . Since lines for all slants are assumed to start at (0, y_0 , 0), these estimates show the differential effects of surface slant in perspective projection isolated from the differential effects of distance from the observer.

References

- Albert, M. K. (1999). Surface formation and depth in monocular scene perception. *Perception*, 28(11), 1347–1360.
- Albert, M. K., & Tse, P. U. (2000). The role of surface attraction in perceiving volumetric shape. *Perception*, 29(4), 409–420.
- Aloimonos, J. (1988). Shape from texture. *Biological Cybernetics*, 58(5), 345–360.
- Ames, A. (1951). Visual perception and the rotating trapezoidal window. *Psychological Monographs* 65(7).
- Arad, N., & Elber, G. (1997). Isomeric texture mapping for free-form surfaces. *Computer Graphics Forum*, 16(5), 247–256.
- Behl, B. K. (1998). *The Ajanta Caves: artistic wonder of ancient Buddhist India*. New York: Harry N. Abrams.
- Binford, T. O. (1981). Inferring surfaces from images. *Artificial Intelligence*, 17, 205–244.
- Blake, A., Bulthoff, H. H., & Sheinberg, D. (1993). Shape from texture: ideal observers and human psychophysics. *Vision Research*, 33, 1723–1737.
- Bracewell, R. N. (1995). *Two-dimensional imaging*. Englewood Cliffs, NJ: Prentice Hall.
- Braunstein, M. L., & Payne, J. W. (1969). Perspective and form ratio as determinants of relative slant judgments. *Journal of Experimental Psychology*, 81(3), 584–590.
- Brodatz, P. (1966). *Textures: a photographic album for artists and designers*. New York: Dover.
- Campbell, F. W., & Robson, J. G. (1968). Application of Fourier analysis to the visibility of gratings. *Journal of Physiology*, 197(3), 551–566.
- Clowes, M. B. (1971). On seeing things. *Artificial Intelligence*, 2, 79–116.
- Cummings, B. G., Johnston, E. B., & Parker, A. J. (1993). Effects of different texture cues on curved surfaces viewed stereoscopically. *Vision Research*, 33, 827–838.
- Cutting, J. E., & Millard, R. T. (1984). Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology*, 113, 196–216.
- Das, A., & Gilbert, C. D. (1999). Topography of contextual modulations mediated by short-range interactions in primary visual cortex. *Nature*, 399(6737), 643–644.
- DeBonet, J. S. (1997). Multiresolution sampling procedure for analysis and synthesis of texture images. Proceedings of AMC SIGGRAPH (pp. 361–368). August 3–8, 1997, Los Angeles, CA.
- Foley, J. D., van Dam, A., Feiner, S. K., & Hughes, J. F. (1987). *Computer graphics: principles and practice* (2nd ed.). New York: Addison-Wesley Publishing Company.
- Garding, J. (1992). Shape from texture for smooth curved surfaces. *Journal of Mathematical Imaging and Vision*, 2, 327–350.
- Gibson, J. J. (1950). The perception of visual surfaces. *American Journal of Psychology*, 63, 367–384.
- Graham, N. (1989). *Visual pattern analyzers*. New York: Oxford University Press.
- Griffiths, A. F., & Zaidi, Q. (1998). Rigid objects that appear to bend. *Perception*, 27(7), 799–802.
- Griffiths, A. F., & Zaidi, Q. (2000a). Perceived shapes and poses of isoperspectival solids. *Investigative Ophthalmology and Visual Science*, 41(4), S722.
- Griffiths, A. F., & Zaidi, Q. (2000b). Perceptual assumptions and projective distortions in a three-dimensional shape illusion. *Perception*, 29(2), 171–200.
- Griffiths, A. F., & Zaidi, Q. (2001). Looking through Ames' window. Vision Sciences Society Meeting, May 4–8, Sarasota, FL.
- Hanrahan, P., & Haerberli, P. (1990). Direct WYSIWYG painting and texturing on 3D shapes. Proceedings of SIGGRAPH, pp. 215–238 Dallas, TX.
- Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of IEEE*, 67 (5), 786–804.
- Heeger, D. J. (1992). Half-squaring in responses of cat striate cells. *Journal of Neuroscience*, 9, 427–443.
- Hel, O. Y., & Zucker, S. W. (1989). Texture fields and texture flows: sensitivity to differences. *Spatial Vision*, 4(2–3), 131–139.

- Hess, R., & Field, D. (1999). Integration of contours: new insights. *Trends in Cognitive Science*, 3(12), 480–486.
- Interrante, V. (1997). Illustrating surface shape in volume data via principal direction-driven 3D line integral convolution. Paper presented at the SIGGRAPH, Los Angeles, CA.
- Interrante, V., & Kim, S. (2001). Investigating the effect of texture orientation on the perception of 3D shape. *Proceedings of SPIE*, 4299, 330–339.
- Knill, D. C. (1998a). Surface orientation from texture: ideal observers, generic observers and the information content of texture cues. *Vision Research*, 38(11), 1655–1682.
- Knill, D. C. (1998b). Discrimination of planar surface slant from texture: human and ideal observers compared. *Vision Research*, 38(11), 1683–1711.
- Knill, D. C. (1998c). Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture. *Vision Research*, 38, 2635–2656.
- Knill, D. C. (2001). Contour into texture: information content of surface contours and texture flow. *Journal of the Optical Society of America A*, 18(1), 12–35.
- Koenderink, J. J. (1984a). The structure of images. *Biological Cybernetics*, 50, 363–373.
- Koenderink, J. J. (1984b). What does the occluding contour tell us about solid shape? *Perception*, 13(3), 321–330.
- Koenderink, J. J., & van Doorn, A. J. (1979). The internal representation of solid shape with respect to vision. *Biological Cybernetics*, 32, 211–216.
- Koenderink, J. J., van Doorn, A. J., Christou, C., & Lappin, J. S. (1996). Shape constancy in pictorial relief. *Perception*, 25(2), 155–164.
- Landy, M., & Maloney, L. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, 35(3), 389–412.
- Lennie, P. (1998). Single units and visual cortical organization. *Perception*, 27(8), 889–935.
- Li, A., & Zaidi, Q. (2000). Perception of three-dimensional shape from texture is based on patterns of oriented energy. *Vision Research*, 40, 217–242.
- Li, A., & Zaidi, Q. (2001a). Information limitations in perception of shape from texture. *Vision Research*, 41(22), 2927–2942.
- Li, A., & Zaidi, Q. (2001b). Veridicality of three-dimensional shape perception predicted from amplitude spectra of natural textures. *Journal of the Optical Society of America A*, 18(10), 2430–2447.
- Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, 10(4), 903–940.
- Lowe, D. (1985). *Perceptual organization and visual recognition*. Boston, MA: Kluwer.
- Lu, C., Cao, Y., & Mumford, D. (2001). Surface evolution under curvature flows. Image Processing, Computer Vision, and Computer Graphics (Special issue on Partial Differential Equations (PDE's)), submitted for publication.
- Malik, J., & Rosenholtz, R. (1997). Computing local surface orientation and shape from texture for curved surfaces. *International Journal of Computer Vision*, 23(2), 149–168.
- Mareshal, I., & Baker, C. (1998). A cortical locus for the processing of contrast-defined contours. *Nature Neuroscience*, 1(2), 150–154.
- Mundy, J. L., & Zisserman, A. (1992). Projective Geometry for machine vision. In J. L. Mundy, & A. Zisserman (Eds.), *Geometric Invariance in computer vision* (pp. 463–519). Cambridge, MA: MIT Press.
- Panofsky, E. (1927). Die perspektive als 'symbolische form', vortage der bibliothek warburg 1924–1925. New York: Zone Books, pp. 258–330.
- Peachey, D. R. (1985). Solid texturing of complex surfaces. *Proceedings of SIGGRAPH*, pp. 279–286 San Francisco, CA.
- Rosenfeld, A. (1971). Isotonic grammars, parallel grammars, and picture grammars. In B. M. D. Michie (Ed.), *Machine intelligence* (6) (pp. 281–294). Edinburgh: Edinburgh University Press.
- Rosinski, R. R., & Levine, N. P. (1976). Texture gradient effectiveness in the perception of surface slant. *Journal of Experimental Child Psychology*, 22, 261–271.
- Rubin, N. (2001). The role of junctions in surface completion and contour matching. *Perception*, 30(3), 339–366.
- Sakai, K., & Finkel, L. H. (1993). Characterization of the spatial-frequency spectrum in the perception of shape from texture. *Journal of Optical Society of America A*, 12(6), 1208–1224.
- Schofield, A. J., & Georgeson, M. A. (1999). Sensitivity to modulations of luminance and contrast in visual white noise: separate mechanisms with similar behavior. *Vision Research*, 39(16), 2697–2716.
- Shapley, R., & Lennie, P. (1985). Spatial frequency analysis in the visual system. *Annual Review of Neuroscience*, 8, 547–583.
- Shepard, R. (1990). *Mind sights*. New York: W.H. Freeman.
- Sperling, G., Chubb, C., Solomon, J. A., & Lu, Z. L. (1994). Full-wave and half-wave processes in second-order motion and texture. *Ciba Foundation Symposium*, 184, 287–303.
- Spink, W. (1994). *Ajanta: a brief history and guide*. Ann Arbor: Asian Art Archives of the University of Michigan.
- Stevens, K. A. (1981). The visual interpretation of surface contours. *Artificial Intelligence*, 17, 47–73.
- Sugihara, K. (1984). An algebraic approach to shape-from-image problems. *Artificial Intelligence*, 23, 59–95.
- Todd, J. T., & Akerstrom, R. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance*, 113, 221–224.
- Todd, J. T. & Oomes, A. H. J. (2002). Generic and nongeneric conditions for the perception of surface shape from texture. *Vision Research*, 42.
- Tse, P. U. (1999). Volume completion. *Cognitive Psychology*, 39(1), 37–68.
- Tse, P. U. (2002). A planar cut approach to volume recovery from silhouettes. *Psychological Review*, in press.
- Turner, M. R., Gerstein, G. L., & Bajcsy, R. (1991). Underestimation of visual texture slant by human observers: a model. *Biological Cybernetics*, 54, 215–226.
- Van Effeltherre, T. (1994). Aspect graphs for visual recognition of three-dimensional objects. *Perception*, 23(5), 563–582.
- Vickers, D. (1971). Perceptual economy and the impression of visual depth. *Perception and Psychophysics*, 10(1), 23–27.
- Willats, J. (1992). Seeing lumps, sticks, and slabs in silhouettes. *Perception*, 21(4), 481–496.
- Zaidi, Q., & Li, A. (2000). Neural model of shape from texture: developable surfaces. *Investigative Ophthalmology and Visual Science (Supplement)*, 41, S219.
- Zaidi, Q., Spehar, B., & Shy, M. (1997). Induced effects of backgrounds and foregrounds in three-dimensional configurations: the role of T-junctions. *Perception*, 26(4), 395–408.
- Zhu, S. C., Wu, Y., & Mumford, D. (1997). Minimax entropy principle and its applications to texture modeling. *Neural Computation*, 9(8), 1627–1660.