

Three-dimensional shape from non-homogeneous textures: Carved and stretched surfaces

Andrea Li

Department of Vision Sciences, SUNY College of Optometry,
New York, NY, USA



Qasim Zaidi

Department of Vision Sciences, SUNY College of Optometry,
New York, NY, USA



We examined the perception of 3D shape for surfaces folded, carved, or stretched out of textured materials. The textures were composed of sums of sinusoidal gratings or of circular dots, and were designed to differentiate between orientation and frequency information present in perspective images of the surfaces. Correct perception of concavities, convexities, saddles, and slants required the visibility of signature patterns of orientation modulations. These patterns were identical to those identified previously for developable surfaces (A. Li & Q. Zaidi, 2000; Q. Zaidi & L. Li, 2002), despite the fact that textures were statistically homogeneous on developable surfaces but not on carved or stretched surfaces. Frequency modulations in the image were interpreted as cues to distance from the observer, which led to weak but qualitatively correct percepts for some carved and stretched surfaces but to misperceptions for others, similar to the misperceptions for developable surfaces (A. Li & Q. Zaidi, 2003). Irrespective of whether texture on the surface is homogeneous or non-homogeneous, similar neural modules can be used to locate signature orientation modulations and thus extract shape from texture cues.

Keywords: 3D shape from texture, orientation modulations, frequency modulations, homogeneity, non-developable surfaces

Introduction

In the perspective image of a curved three-dimensional (3D) surface, the statistics of the texture pattern change with the curvature of the surface. (We follow convention in using the term texture for surface markings that form a repetitive pattern.) Even the most sophisticated shape-from-texture models assume that the texture on the surface is statistically homogeneous (i.e., stochastically stationary and invariant to translation on the surface), and inhomogeneities in the image arise from the projection of segments of the surface that depart from being fronto-parallel with respect to the observer (Clerc & Mallat, 2002; Garding, 1992; Malik & Rosenholtz, 1997). This assumption is true only under very restricted conditions. A widely studied case is that of developable surfaces that can be unfolded into a flat plane without stretching or cutting (e.g., cylinders, cones, and sinusoidal corrugations). For the subset of patterns that are statistically homogeneous over a flat sheet, developable surfaces can be formed from that sheet so that the texture is homogeneous over the whole surface. Developable surfaces can have very complex shapes, as shown by Huffman (Stix, 1991; "Geometric Paper Folding: Dr. David Huffman" [link]); however, they can only have local Gaussian curvatures equal to zero (maximum curvature times minimum curvature), so it requires other operations such as carving or stretching to make more general surfaces, which have local Gaussian curvatures that vary from greater than to less than zero. Whereas it is possible to carefully paint a carved

or stretched surface with homogeneous texture (Clerc & Mallat, 2002), under generic conditions, the texture on a carved or stretched surface is not homogeneous if the surface is like a saddle or an ellipsoid and has varying Gaussian curvature. In instances such as skin and clothing, the inhomogeneity may change as the surface deforms. Thus, for most complex shapes, texture inhomogeneities in an image are not caused solely by the projection so that estimating the projective transform and reversing it, as in Garding (1992), Malik and Rosenholtz (1997), and Clerc and Mallat (2002), is not sufficient to infer the 3D shape of the surface.

In this work, we examine the perception of 3D shape from texture cues for developable surfaces on which the texture is homogeneous, and carved and stretched surfaces on which it is not. We show that the assumption of homogeneity is not necessary for extracting 3D shape because observers correctly perceive 3D curvatures and slants when signature patterns of orientation modulations are visible, irrespective of whether the texture on the surface is homogeneous or not. We also show that, in the generic case, these orientation modulations will appear in perspective images of carved, stretched, and developable surfaces only at the locations of the correct curvatures or slants. Shape from texture can thus rely on neural modules that extract signature orientation modulations, irrespective of the homogeneity of the texture pattern and whether the surface is developable, carved, or stretched. When signature patterns of orientation modulations are not visible, observers infer

shape using spatial frequency modulations as cues to distance. This leads to correct percepts for images where spatial frequency varies with distance from the observer, but incorrect percepts where the spatial frequency varies with the slant of the surface. (Note: Throughout this work, when we refer to *correct* percepts, we explicitly mean that the perceived *signs* of curvatures and *directions* of slants are identical to those of the simulated 3D surface.)

In studying 3D shape, we have used sinusoidal corrugations and depth plaids as samples from a set of basis shapes (i.e., shapes that in combination could generate a wide variety of shapes) (Bracewell, 1995). Here we first compare the perception of sinusoidally carved surfaces to developable surfaces (Figure 1), and then generalize the results to depth plaids (sums of orthogonal sinusoidal corrugations) containing positive and negative Gaussian curvatures. Flat, foldable materials have only a single surface pattern, but 3D solids can exhibit different surface patterns depending on the direction of the cut (e.g., the veneer of wood is more varied if it is cut across the grain than if the wood is cut parallel to the grain). In addition, the surface pattern is statistically similar for certain parallel cuts, but not for others. In this study, surfaces were carved from the two classes of solids shown in Figure 2. The constant-z solid was formed by repeating identical planar patterns along the z-axis (i.e., the axis of carved depth). The constant-x solid was formed by repeating identical planar patterns along the x-axis, orthogonal to the axis of carved depth. The same planar patterns were also folded into developable corrugations, and stretched onto corrugated solids.

The perspective images of textured surfaces presented to observers in this study were computed by projecting 1.5 cycles of the sinusoidally curved developable and carved surfaces onto the image plane of a CRT monitor. When viewed monocularly at a viewing distance of 1 m, the retinal image coincided with that of a real 3D sinusoidally curved surface with an amplitude of 7 cm and a wavelength of 10 cm (Figure 3). To restrict the shape cues solely to texture variations in the image, all surfaces were presented in fronto-parallel view and without occluding contours. The effects described in this paper are robust enough to be seen in the perspective images in this study even in less than perfect viewing conditions (e.g., at reading distance).

We used two classes of texture patterns to separate the contributions of orientation and frequency modulations to shape perception. Shape perception by the visual system is a complex process that is likely to involve interactions across many areas of cortex. For example, Mumford (1992) proposed that neurons in higher areas of the cortex can function as deformable templates or matched-filters, and cells at lower levels transmit difference signals between feedback from activated higher level neurons and inputs from lower level neurons. Murray, Kersten, Olshausen, Schrater, and Woods (2002) have provided fMRI evidence compatible with such linkage between LOC and V1. In any neural model that involves feedback to V1 and/or extensive lateral interactions, it is not possible to treat V1 neurons as inde-

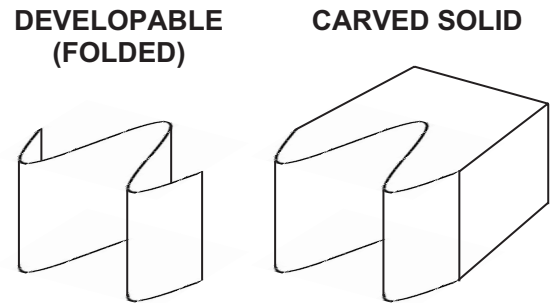


Figure 1. Developable (folded) and carved solid surfaces with identical sinusoidally corrugated shapes from the observer's viewpoint.

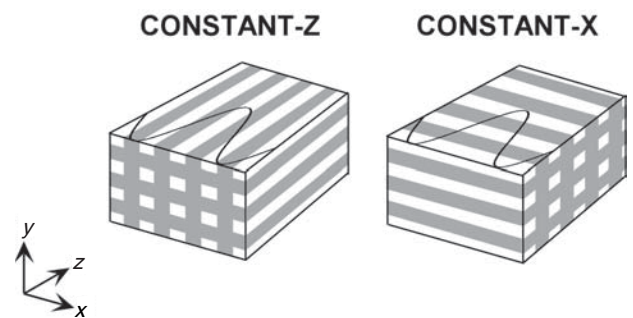


Figure 2. Solids used for carved surfaces: constant-z solids contain identical planar patterns repeated along the z-axis, and constant-x solids contain identical planar patterns repeated along the x-axis. The sinusoidal curves show the cuts that are made through the solids.

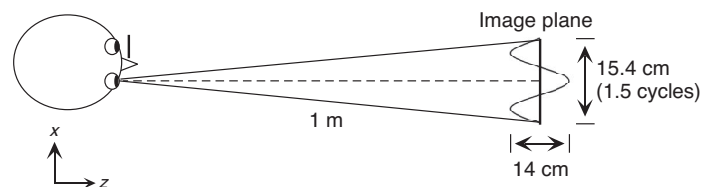


Figure 3. CRT displays formed retinal images identical to perspective projections of real 3D surfaces with the indicated shape and dimensions.

pendent filters, but the currencies of both the feed-forward and feed-back signals are the receptive field properties of V1 cells. Because V1 neurons are tuned for orientation and spatial frequency, it is useful to parse texture variations in an image into orientation and frequency modulations.

The first class of patterns we used was composed of oriented sinusoidal gratings, shown with their amplitude spectra in the left column of Figure 4: a horizontal-vertical plaid, an octotropic plaid consisting of eight gratings of the same spatial frequency equally spaced in orientation (components shown in Figure 5), and the octotropic plaid minus the horizontal grating. The second class, shown in the right

column of Figure 4, consisted of patterns made of circular dots: a pattern consisting of uniformly sized dots that were randomly positioned (with a minimal overlap constraint), a pattern in which the uniformly sized dots were horizontally and vertically aligned, and a pattern in which the size of the aligned dots was randomly varied. While the elements of all three of the patterns are isotropic, the first pattern is the only one that is also globally isotropic as shown by its amplitude spectrum. The other two patterns contain concentrations of energy at discrete orientations as shown by their amplitude spectra.

For scales greater than one-sixteenth of the patterns in Figure 4, the textures are statistically homogenous. When larger versions of these textures are folded into a developable surface, the texture on the surface remains statistically homogenous. The texture on the surface, however, is not homogeneous when solids containing these patterns are carved, or when elastic versions of these patterns are stretched over curved surfaces. However, the surface markings on carved or stretched surfaces are not randomly non-homogenous, but rather are locally affine transformations of these homogeneous patterns, where the affine transformation is a function of the local curvature. Texture distortions in perspective images are therefore due to a combination of the shape-caused and the projective transformations.

In the following sections, we will analyze perspective images in terms of local changes in orientation and frequency, and examine how each contributes to shape percepts. We will show that the orientation modulations of critical components are essentially immune to the particular carving or stretching process, but frequency modulations are not. As a result, correct curvature perception occurs whenever signature orientation modulations are visible, irrespective of whether the surface is carved, stretched, or folded. These results are illustrated in the images in this work, and were confirmed empirically by psychophysical experiments described in "Appendix A," in which observers were asked to judge the relative depth of two test locations at various phases along the surface. Quantitative measurements of shape percepts are presented in "Appendix A," and will be referred to within appropriate sections.

Developable surfaces

When any of the patterns in Figure 1 are folded, the texture on the surface is unchanged, but texture distortions are visible in perspective images. Images of the developable

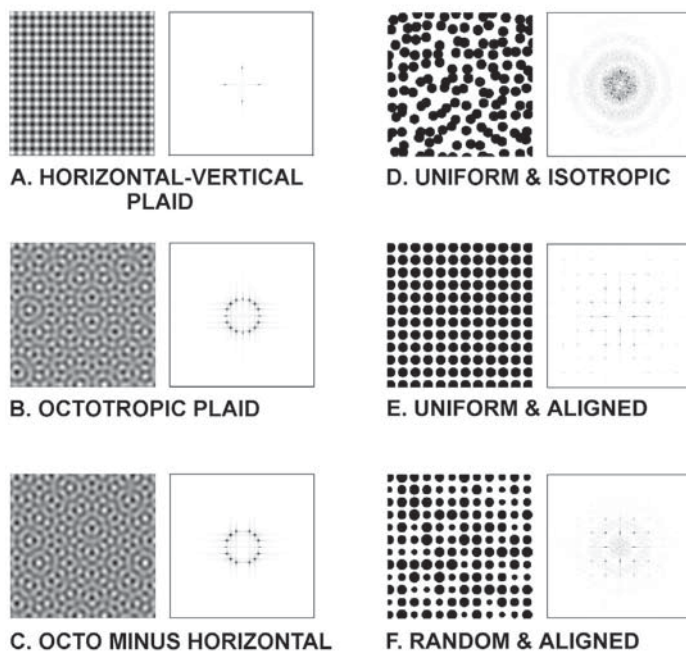


Figure 4. Two groups of texture patterns (shown here together with their amplitude spectra) were examined: composites of sinusoidal gratings (A-C) and locally isotropic dot patterns (E-F). A. Horizontal-vertical plaid. B. Octotropic plaid consisting of eight gratings of the same frequency as those in A, equally spaced in orientation (see Figure 5). C. Octotropic plaid without the horizontal component. D. Uniformly sized, randomly positioned dot pattern. E. Uniformly sized, horizontally and vertically aligned dot pattern. F. Randomly sized, horizontal and vertically aligned dot pattern.

corrugations overlaid with the sinusoidal grating patterns are shown in Figure 6. Surfaces with a central concavity are presented in the top row and surfaces with a central convexity in the bottom row. For both the horizontal-vertical plaid and the octotropic plaid, it is easy to identify right and left slants, and thus concavities and convexities. Observers correctly identified right and left slanting portions of the surface for these two patterns (upper left and middle panels of Figure A1). Slants are not easily distinguished, however, in the images in the third column of Figure 6 where the texture pattern is missing the horizontal grating. For this pattern, observers confused left and right slants, and often classified both as flat (upper right panel, Figure A1). This shows that the information supplied by the horizontal grating is crucial to correct shape perception for up-right corrugations (Li & Zaidi, 2000; Li & Zaidi, 2001a).

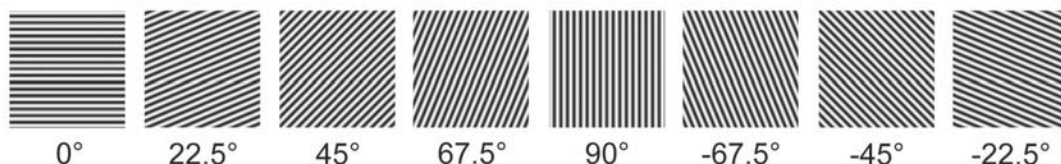


Figure 5. Eight oriented grating components of octotropic plaid (Figure 4B).

This information is visible as contours that bow inward toward the center of the image at local concavities, bow outward at local convexities, and converge rightward or leftward, respectively, at rightward and leftward slants.

It is easy to show why the horizontal grating uniquely carries the shape information in these images. For the surface in concave phase, Figure 7 shows the effects of corrugation and perspective projection on the eight oriented components of the octotropic plaid (see "Appendix B" for mathematical derivations of projected orientations and frequencies). The image of the horizontal component (0°) is the only one that shows patterns of orientation modulations that are different for different signs of curvature. The image of the vertical grating (90°) shows frequency modulation but no changes in orientation. For all the oblique components, the local orientation and frequency at fronto-parallel portions of the surface (i.e., at centers of concavities and convexities) equal the original orientation and frequency, and increase with increasing slant. When all eight components are added to form the image in Figure 6B (top), the horizontal component is visible because its orientation modulations vary only between $\pm 10^\circ$, whereas the minimum orientation of any other component is $\pm 22.5^\circ$, and at slanted portions of the surface, the frequency of the horizontal component is lower than that of the other components. The other seven components do not convey shape individually (Figure 7) or summed together (Figure 6C). Consequently, images of the octotropic pattern contain sufficient information for shape to be perceived correctly, but this is not true for images of the octotropic pattern minus the horizontal component. As will be shown in this work, for images of upright shapes, the pattern of orientation modulations of the horizontal component is universal for texture patterns containing discrete energy parallel to the axis of maximum curvature.

Frequency modulations in the image, however, will be shown to vary as a function of how the surface is formed. For example, frequency modulations in perspective images of developable surfaces are caused largely by changes in surface slant. Figure 8 shows an aerial view of a patterned surface slanted at two different angles with respect to the observer's eye. Because the frequency of the pattern on the surface is constant, as slant increases, the projected width of the pattern in the image plane decreases, and the frequency in the image increases. In images of textured objects

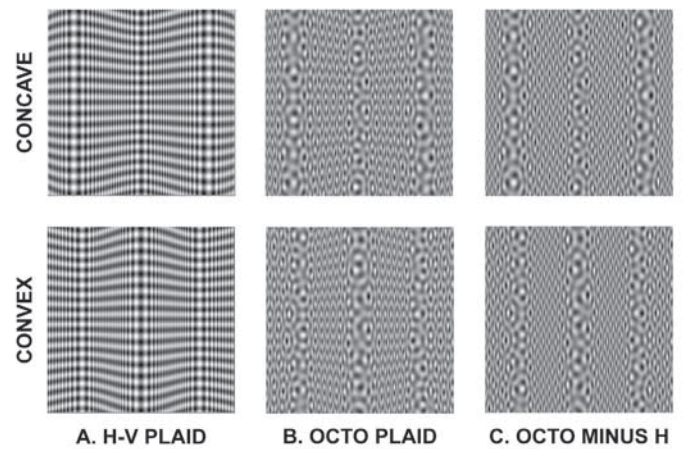


Figure 6. Perspective images of developable surfaces containing central concavities (top) and convexities (bottom) overlaid with the three grating composite patterns. Signature orientation modulations of the horizontal component in both plaid patterns (A-B) contain sufficient information to correctly convey concavities, convexities, right slants, and left slants. Subtracting the horizontal component (C) eliminates these orientation modulations, and the surface shapes are not correctly perceived.

whose internal depth is substantially less than their distance from the observer, spatial frequency modulations are due more to changes in slants than to changes in distances with respect to the observer. Consequently, images of rightward and leftward slants exhibit similarly increased frequency because of the slant, with little difference between them from changes in distance. As a result, images of concave and convex portions of the corrugation exhibit similar high-low-high frequency gradients. Observers cannot resolve this ambiguity and perceive convex and concave curvatures both as convexities. This percept is consistent with the frequency gradient functioning as a cue to relative distance from the eye because the effect of distance is to increase the spatial frequencies in the image of a pattern (Li & Zaidi, 2003). It is worth noting that in cases where the observer is navigating through a textured environment, there is a large range of distances to the observer. Consequently, the frequency modulations in the retinal image are mainly due to changes in distance, and thus provide veridical cues as to the shape of the environment.

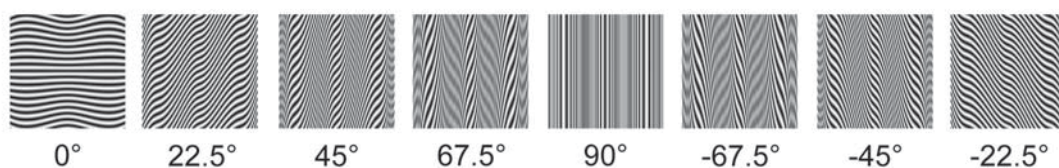


Figure 7. Perspective images of the developable surface (with a central concavity) overlaid with each of the eight grating components of the octotropic plaid. The horizontal component exhibits the signature orientation modulations. All other components exhibit low frequencies at concavities and convexities and high frequencies at left and right slants. Orientation modulations of these components are all steeper than those exhibited by the horizontal component.

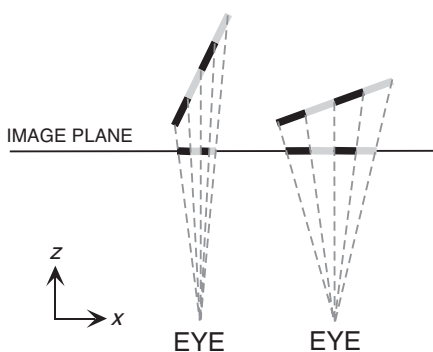


Figure 8. Frequency modulations in images of developable surfaces are largely slant-caused. Aerial view of a vertical grating on a flat surface at two different slants. As slant increases, frequency in the perspective image increases.

Images of the developable corrugations overlaid with the three dot patterns are shown in Figure 9. The images of the corrugations with the isotropic dot pattern (Figure 9A) exhibit slant-caused frequency modulations along the horizontal axis with high-low-high frequency gradients at both concavities and convexities. For this pattern, observers confused left and right slants (lower left panel of Figure A1). Li and Zaidi (2003) showed that for globally isotropic patterns, observers report that concavities and convexities both appear convex, indicating that, rather than attribute these modulations to changes in surface slant, observers attribute them to changes in distance. This is done despite the fact that frequency changes due solely to distance would be iso-

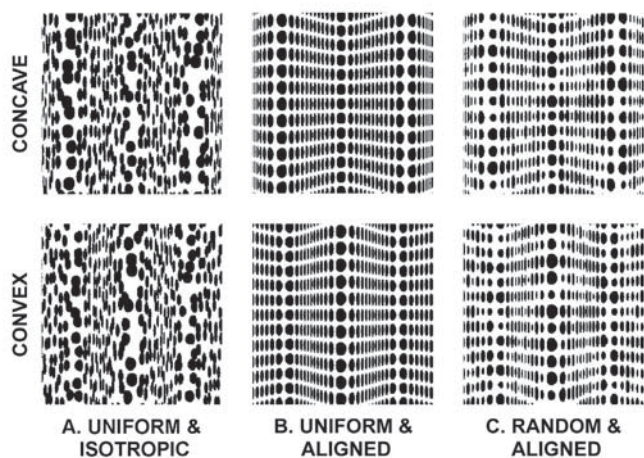


Figure 9. Perspective images of the developable surfaces overlaid with the three dot patterns. Slant-caused frequency modulations in the globally isotropic dot pattern (A) are misinterpreted as changes in distance, and as a result concavities are misperceived as convex. Horizontal and vertical alignment of the dots (B) adds the signature orientation modulations of the horizontal component (see Figure 6) and concavities become distinguishable from convexities. Randomizing the size of the aligned dots (C) makes little difference in the percepts.

tropic, whereas the frequency changes in Figure 9A are almost exclusively along the axis of maximum curvature, suggesting that the frequency modulations in the image are more potent cues to 3D shape than shape changes of texture elements. Orientation modulations are difficult to perceive for the isotropic dot pattern, but the modulations are apparent when the dots are horizontally and vertically aligned in the texture (Figure 9B). These modulations are similar to those of the horizontal component in Figure 7. Concavities, convexities, right slants, and left slants are all identifiable. Randomizing the size of the aligned dots, as in Figure 9C, may compromise the ability to extract frequency modulations but it does not affect the shape percepts much; the different surface shapes are easily distinguishable, and observers correctly identified left and right slants (lower middle and right panels of Figure A1).

Carved constant-z corrugations

The surface texture was homogeneous for all the developable examples above, but that is not the case for the carved surfaces that follow. Figure 10 shows perspective images of corrugations carved from constant-z solids formed by repeating a single texture pattern along the z-axis (Figure 2). In the images of the solids patterned with the horizontal-vertical plaid (Figure 10A) concavities, convexities, right and left slants can all be correctly identified. The orientation modulations of the horizontal component appear identical to those of the horizontal component on the developable surface. Because the carved solid's axis of maximum curvature is horizontal, the horizontal component is not distorted on the surface and is identical to the

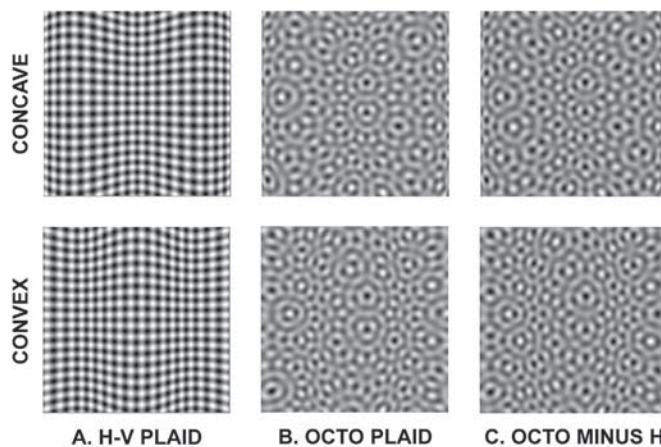


Figure 10. Perspective images of the sinusoidal surfaces carved from constant-z solids with the three grating component planar patterns. The horizontal component in the HV plaid exhibits the same signature orientation modulations that convey concavities and convexities; however, the surfaces appear more gradually curved than their developable counterparts (Figure 6). These orientation modulations are invisible in the octotonic plaid patterns (B-C), which both appear flat.

undistorted horizontal component on the developable surface. Projection thus results in patterns of orientation modulations of the horizontal component that are identical for both kinds of surfaces.

Despite identical orientation modulations, the shapes of the surfaces in Figure 10A appear more gradually curved than their developable counterparts in Figure 6A. (These percepts are quantified in "Appendix A2.") This is because the frequency of the vertical component modulates much less than for the developable surface. Figure 11 shows an aerial view of a constant- z solid formed by vertical grating planar patterns carved at two different angles (indicated by the thick dark grey lines). Unlike for the developable surface, the frequency on the surface of the cut decreases with increasing slant angle. However, as slant angle increases, the projected width of a unit length of solid decreases in the image. These two tendencies counteract each other, so that in the perspective image, the frequency is essentially unaffected by slant. Modulations in the image thus are mainly due to changes in distance from the observer. Consequently, the frequency gradients around concavities and convexities are distinct from one another: low-high-low for concavities and high-low-high for convexities. Variations in spatial frequency on the carved surface show that the texture is not homogeneous on a surface carved with multiple slants.

The images in Figures 10B and 10C appear flat. This is particularly surprising for Figure 10B, where the horizontal component of the octotropic plaid could be expected to contribute the signature orientation modulations. The reason is revealed by Figure 12, which shows the images of the eight components for the carved constant- z solid (see "Appendix C" for mathematical derivations of projected orientations and frequencies). As expected, the horizontal component exhibits the signature orientation modulations that observers use to perceive shape correctly for the horizontal-vertical plaid. However, the images of the $\pm 22.5^\circ$ components contain orientations and frequencies that are similar to those of the horizontal component and mask the orientation modulations of the horizontal component in the summed image. In Figure 13, these two components are subtracted from the octotropic plaid; the signature orientation modulations of the horizontal component become visible, and concavities, convexities, and right and left slants become distinguishable (upper middle and right pan-

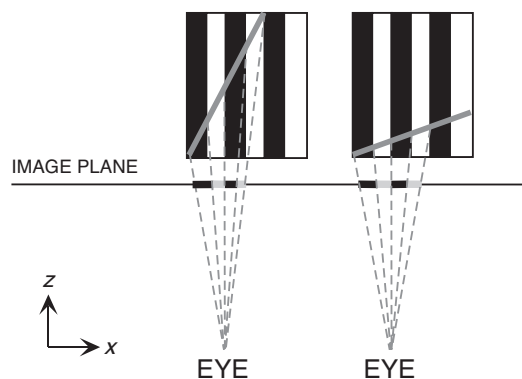


Figure 11. Frequency modulations for carved constant- z solids. Aerial view of a constant- z solid formed by vertical grating planar patterns. As the angle of the cut is increased, the frequency on the surface of the cut decreases; however, projection increases the frequency in the image plane. As a result there is little frequency modulation in the image.

els, Figure A2). It is interesting that the distance caused frequency modulations of the seven other components in Figure 12 are consistent with correct percepts of the central concavity, but the perceived shape is essentially flat when all seven components are combined in Figure 10C.

Images of the carved constant- z corrugations with the dot patterns are shown in Figure 14. All the images for the dot patterns in Figure 14 contain frequency modulations determined by distance. Orientation modulations are visible in the aligned dot patterns (Figure 14B and 14C), but not in the isotropic pattern (Figure 14A). In Figure 14A, concavities and convexities are discernible, but just barely, from the frequency cue to distance. While observers make some correct slant judgments for this pattern, a large proportion of the slants are classified as flat (lower left panel, Figure A2). Signs of curvature and slant are easily identifiable when signature orientation modulations are visible (Figure 14B and 14C). The addition of random frequency modulations in Figure 14C hardly affects the shape percepts (lower middle and right panels, Figure A2).

It is worth pointing out that all six of the patterns in Figure 10 and Figure 14 are inhomogeneous on the surface of the solid, but that frequency and orientation modulations signal correct locations and signs of curvature. The orientation modulations, in particular, are identical for the

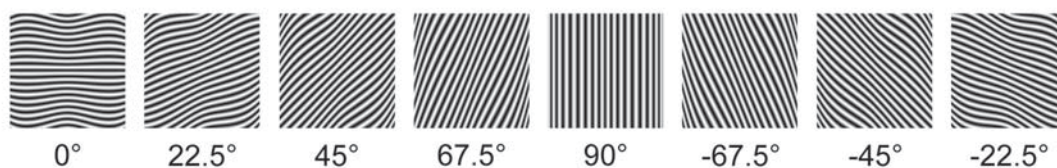


Figure 12. Perspective images of carved constant- z solids (with central concavity) with each of the eight grating patterns of the octotropic plaid. The orientation modulations of the horizontal component are the same as those for developable surfaces. The orientation modulations of the $\pm 22.5^\circ$ components overlap in range with those of the horizontal component.

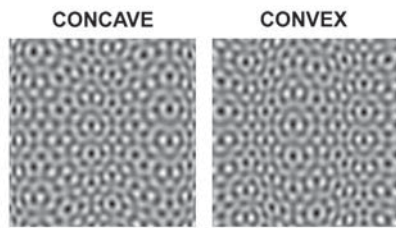


Figure 13. Octotropic plaid from Figure 10B without the $\pm 22.5^\circ$ components. The signature orientation modulations of the horizontal component are revealed and concavities and convexities become distinguishable.

developable and carved surfaces, and provide unambiguous cues to the signs of curvature and slant. Parsing the perspective image in terms of orientation and frequency modulations thus obviates a need to restrict shape-from-texture models to homogenous textures.

Carved constant-x corrugations

When the corrugation is carved from a constant-x solid formed by repeating a single texture pattern along the x-axis (Figure 2), the texture on the surface is inhomogeneous, but the inhomogeneities and hence the perspective images are quite different from the carved constant-z solid. Figure 15 shows images of the corrugations carved from constant-x solids formed by grating patterns. Concavities, convexities, right and left slants are all identifiable for the images of the horizontal-vertical plaid in Figure 15A and the octotropic plaid in Figure 15B, and observers identify slants correctly (upper left and middle panels, Figure A3). In the images, the horizontal component gives rise to the same signature

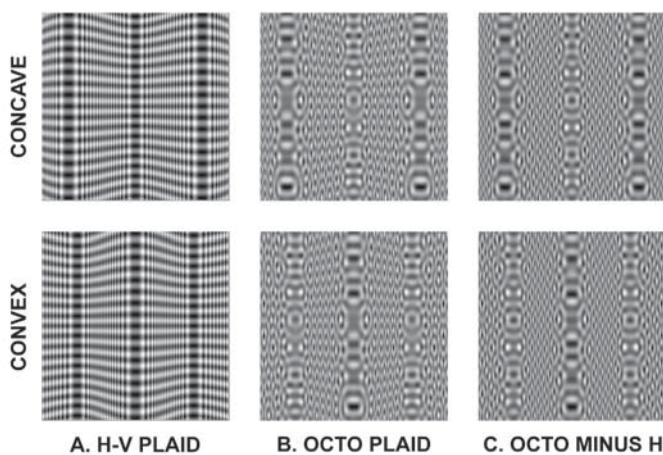


Figure 15. Perspective images of the carved constant-z solids with the three grating component patterns. Signature orientation modulations of the horizontal component in the plaid patterns (A-B) are different for concavities and convexities. Subtracting the horizontal component from the octotropic plaid (C) removes the orientation modulations.

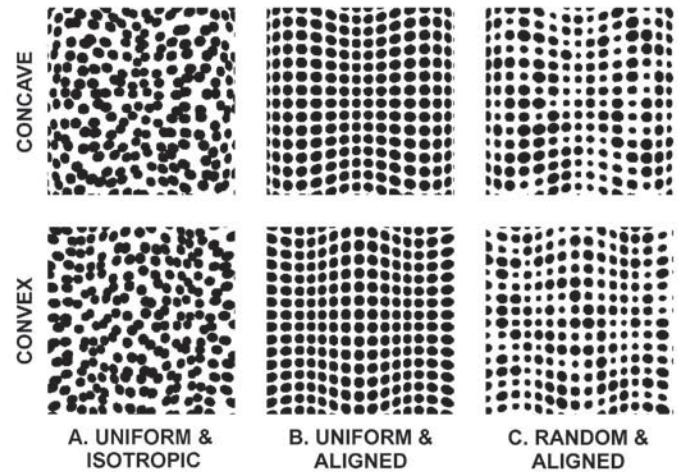


Figure 14. Perspective images of carved constant-z solids with the three dot planar patterns. Distance-caused frequency modulations in the random dot pattern (A) roughly convey concavities and convexities; however, they are much more compelling when the dots are aligned in the solid (B). Randomizing the size of the aligned dots (C) makes little difference in the percept.

orientation modulations as the developable surface because the horizontal component is not distorted by the carving along the horizontal axis. When the horizontal component is subtracted from the planar pattern of the constant-x solid in Figure 15C, the image no longer contains sufficient information to distinguish signs of curvatures and slants. As a result, observers confuse left and right slants (upper right panel, Figure A3).

In the images of the corrugation with the horizontal-vertical plaid (Figure 15A), frequency modulations are similar to but even more pronounced than those of the developable surfaces in Figure 6A. Figure 16 shows an aerial view of a constant-x solid formed by repeating vertical gratings along the x-axis. As the angle of the cut increases, the frequency on the surface of the cut increases. Because in-

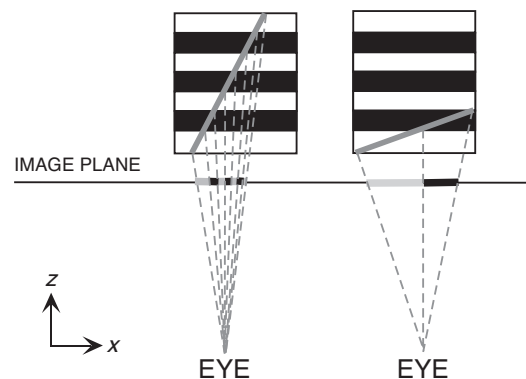


Figure 16. Frequency modulations for carved constant-x solids. Aerial view of a constant-x solid with a vertical grating planar pattern. As the angle of the cut increases, the frequency on the surface of the cut increases. Further, projection increases the frequency in the image plane. As a result, the frequency in the image increases with increasing slant.

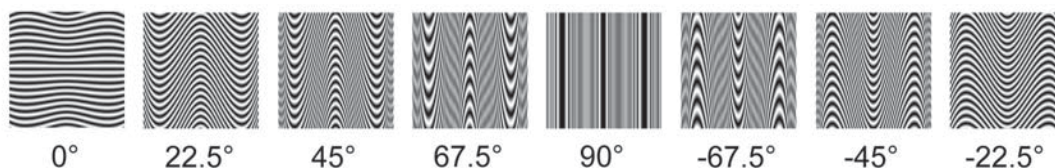


Figure 17. Perspective images of the carved constant-x solid with each of the eight grating components of the octotropic plaid. The horizontal component exhibits the same signature orientation modulations. All other components exhibit slant-caused frequency gradients similar to those for the developable surfaces, and steeper orientation modulations than those of the horizontal component.

creasing the slant also decreases the projected width of a unit surface length in the image, the projected frequency in the image increases much more with increasing slant than for the developable surface. The directions of the frequency gradients are similar for developable and constant-x solids, but the projected frequency for the constant-x solid will be zero when the slant of the cut is zero (i.e., where the surface is fronto-parallel). Concavities and convexities thus exhibit similar high-zero-high frequency gradients. In addition, portions of the surface that are at equal depths (e.g., the peaks of the convexities) cut through identical portions of the planar pattern along the x -axis. Because the surface is periodic and presented with either a central concavity or convexity, the images are symmetric about the vertical mid-line (e.g., Figure 15B-C).

Figure 17 shows the distortions of the eight components of the octotropic plaid within the constant-x solid (see “Appendix D” for mathematical derivations of projected orientations and frequencies). The orientations of the horizontal component are much shallower than the orientations of the other components. In addition, the frequency of the horizontal component remains nearly constant. All of the non-horizontal components exhibit high frequencies where the surface is slanted and low where it is fronto-parallel. Consequently, when the eight components are added together in Figure 15B, the orientation modulations of the horizontal component are visible, especially at slanted portions of the surface.

Figure 18 shows the carved constant-x corrugations with the dot patterns. The images exhibit symmetric distortions about the vertical mid-line. For all three patterns, the slant-caused frequency gradients are similar to but more pronounced than those for the developable surfaces in Figure 9. Frequency modulations are the only cue in the isotropic dot pattern in Figure 18A, and the concavities in the surface appear convex for most viewers, indicating again that frequency modulations are interpreted as distance rather than slant. This is quantitatively confirmed by the fact that observers confuse left and right slants for this pattern (lower left panel, Figure A3). For the aligned dot pattern in Figure 18B, the signature orientation modulations enable concavities, convexities, right and left slants to become distinguishable. Randomizing the size of the dots in Figure 18C does not significantly change the 3D percepts (lower middle and right panels, Figure A3).

The sources of texture inhomogeneities in images of constant-x surfaces are quite different from those for constant-z surfaces. For example, for the vertical component, increasing the slant of the carving decreases the frequency on the constant-z surface but increases the frequency on the constant-x surface. For both surfaces, frequency gradients locate local extrema of curvature, but only for the constant-z surface are the gradients different for concavities and convexities. However, both types of carvings leave the horizontal component undistorted on the surface. As a result, whenever the orientation modulations of this component are visible, observers perceive the correct signs and locations of curvatures and slants.

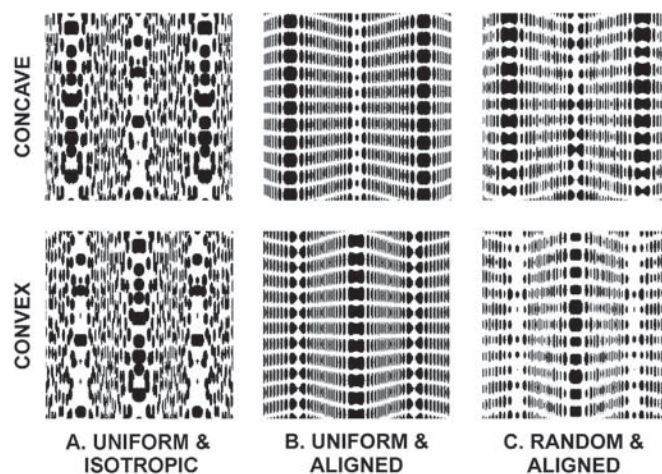


Figure 18. Perspective images of the carved constant-x solid with the three dot patterns. Slant-caused frequency modulations in the isotropic dot pattern (A) are misinterpreted as changes in distance and concavities appear convex. Aligning the dots horizontally and vertically in the solid (B) adds the signature orientation modulations to the image that are different for concavities and convexities. Randomizing the size of the aligned dots (C) makes little difference in the percepts.

Carved depth plaids (constant-z)

So far we have shown that the orientation modulations of the horizontal component are the same for developable and carved surfaces curved along a single axis, that these orientation modulations are different for concavities, convexities, left slants and right slants, and that whenever these

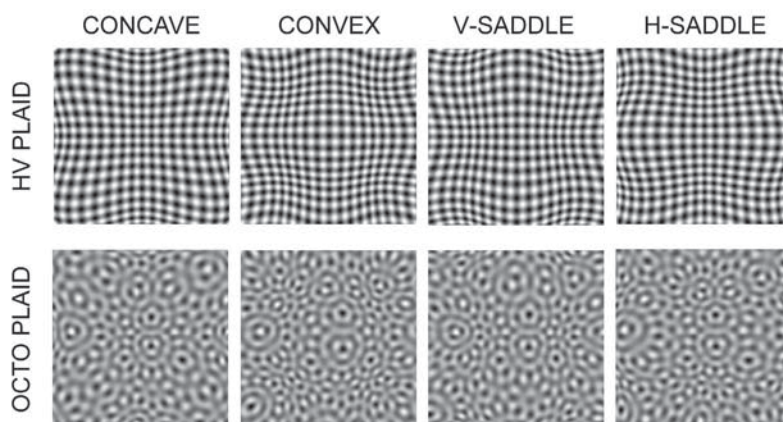


Figure 19. Perspective images of depth plaids curved sinusoidally along the horizontal and vertical axes. The surfaces are carved from constant- z solids with the horizontal-vertical plaid (top) and octotropic plaid (bottom) planar patterns. For each pattern, four different phases of the depth plaid are shown: concave in which curvature along both axes contain a central concavity, convex in which both contain a central convexity, a vertical saddle in which the surface is centrally concave along the vertical axis and convex along the horizontal axis, and a horizontal saddle that is centrally concave along the horizontal axis and convex along the vertical axis. Signature orientation modulations of the horizontal and vertical grating components along each of the two axes of curvature combine to convey the 2D locations of concavities, convexities, and saddles. These modulations are invisible for the octotropic plaid (bottom) and all the images appear flat.

orientation modulations are visible, observers perceive the correct signs of curvatures and slants of 3D surfaces. Do similar rules exist for doubly curved (i.e., inherently non-developable) carved solids? We examined depth plaids that were sums of orthogonal sinusoidal corrugations. The corrugations of these depth plaids had the same amplitude and wavelength as the surfaces above. The surfaces were simulated as carved from constant- z solids formed by each of the six texture patterns in Figure 4.

Figure 19 shows four different phases of the depth plaid for the horizontal-vertical and octotropic plaid patterns. In the leftmost column, the central curvatures along both axes are concave, and in the second column, both are centrally convex. The third column shows a vertical saddle where the curvature parallel to the vertical axis is concave while the curvature parallel to the horizontal axis is convex, and the fourth column shows a horizontal saddle where the curvature parallel to the vertical axis is convex and the curvature parallel to the horizontal axis is concave. For the horizontal-vertical plaid, all the curvatures described above are easy to identify. In the leftmost image, the orientation modulations of the horizontal component are identical to those in the image of the leftmost panel of Figure 12 and are distinct for signs of curvatures and slants along the

horizontal axis. Orientation modulations about the vertical axis are identical to a 90° rotated version of the horizontal modulations, and provide distinct information about curvatures and slants along that axis. Jointly, these two sets of orientation modulations enable correct localization of the concavities, convexities, and saddles in the surface. It appears that for the case where the curvature of a surface can be decomposed into a sum of curvatures along single axes, the shape of the surface can be extracted by simply combining the cues for the curvatures along each axis.

The signature patterns of orientation modulations of the horizontal-vertical plaid are physically present in the images of the solid with the octotropic plaid pattern in the bottom row of Figure 19; however, they are not visible and the surfaces appear flat. Similar to the constant- z solid carved along a single axis, the signature orientation modulations along each axis of curvature are being masked by neighboring components of the planar pattern ($\pm 22.5^\circ$ mask the 0° component, and $\pm 67.5^\circ$ mask the 90° component). When these two sets of neighboring components are subtracted from the images in the bottom row, the signature orientation modulations about each of the two axes are revealed (Figure 20) and concavities, convexities, and saddles become distinguishable.

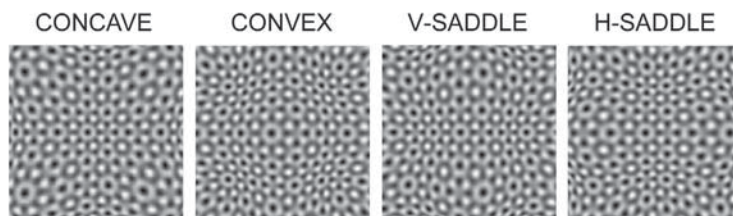


Figure 20. When the four components closest to the horizontal and vertical components in orientation are subtracted from the octotropic plaid in Figure 20B ($\pm 22.5^\circ$ for the horizontal component, $\pm 67.5^\circ$ for the vertical component), the signature orientation modulations along each axis are revealed and the images correctly convey the local surface shapes.

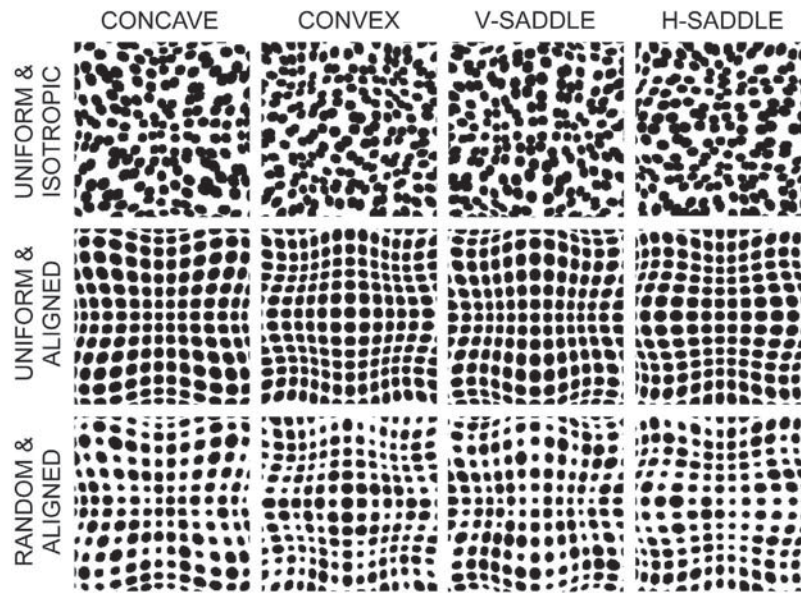


Figure 21. Perspective images of depth plaids carved from constant-z solids with the three dot patterns. Because frequency modulations are caused by distance and are interpreted as such, concavities, convexities, and saddles are correctly conveyed for the isotropic dot pattern (top); however, they are more compelling when the dots are horizontally and vertically aligned in the solid (middle) such that the signature orientation modulations are visible. Randomizing the aligned dots (bottom) makes little difference in the percepts.

Figure 21 shows the depth plaids carved with the three dot patterns. For the isotropic dot pattern, local concavities, convexities, and saddles are discernible, but just barely, from the frequency cues to distance. However, the surface shapes are much more compelling in the middle row where the horizontally and vertically aligned dots in the texture pattern add the signature orientation modulations about each axis of curvature. Randomizing the size of the dots in the bottom row affects the percepts very little.

For these depth plaids, inhomogeneity of texture on the surface is not an impediment to correct perception of curvatures for those cases where signature orientation modulations are visible. As in the case of curvature along a single axis, the orientation modulations occur naturally at the correct locations.

Textured deformable materials

Another class of 3D surfaces on which texture markings are generally non-homogenous is surfaces formed by deforming or stretching textured materials. Examples of deformable materials include animal skins and stretchable clothing.

The top row of Figure 22 shows fronto-parallel views of three unstretched materials patterned with the horizontal-vertical plaid, the octotropic plaid, and the isotropic dot pattern. If these materials are deformed by stretching so that they each have a sinusoidally corrugated shape, the images of the stretched surfaces (bottom row) are identical to those of the carved constant-z solid formed by the same

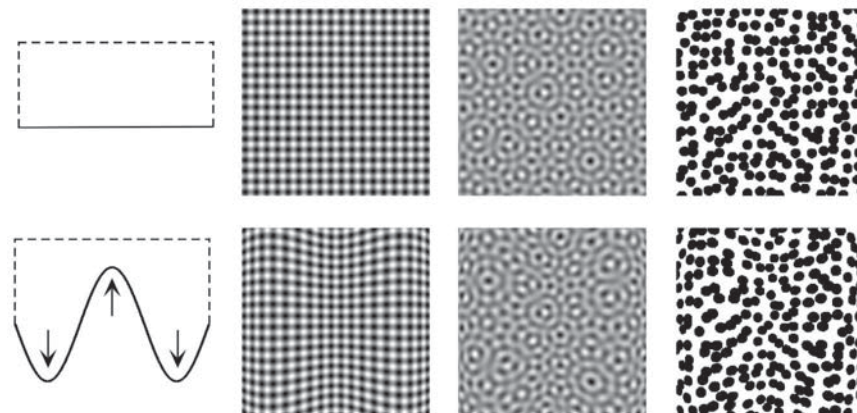


Figure 22. If the patterns in the upper row are stretched so that they are sinusoidally corrugated in depth, the perspective images of the stretched surfaces (bottom row) are identical to those of carved constant-z solids formed by these same planar patterns.

planar patterns. The stretched horizontal-vertical plaid material contains the critical orientation flows that are sufficient for identifying correct surface curvature. For the octotropic plaid, the flows are invisible because of masking and the surface appears flat. For the isotropic dot pattern, the stretching results in distance-based frequency modulations that yield weak but qualitatively correct shape percepts.

Forsyth (2002) has suggested that shape from texture may be the method with the most practical potential for recovering detailed deformation estimates for moving, deformable surfaces such as clothing and skin. Clothing that is not stretchable is like the class of developable surfaces, except for the discontinuities at the seams where even nearest neighbors are not preserved, whereas skin and stretchable clothing stretch in systematic ways with movements. It appears that for certain classes of surface textures, for both of these cases, orientation modulations will arise generically in perspective images and will be informative about curvatures.

Discussion

All the images that the reader has seen in this work are flat surfaces containing repeating but statistically non-homogeneous patterns. When these are viewed monocularly, even without access to stereo or motion, 3D shape percepts are extremely vivid if the signature orientation modulations are visible. This suggests that the visual system automatically creates percepts of curvature corresponding to signature orientation modulations. Given that signature orientation modulations automatically evoke corresponding shape percepts, the question of whether these percepts are correct reduces to whether these modulations occur in the correct locations in perspective images of real solids. This work shows that this is true for developable, carved, and stretched surfaces under many different conditions. This also suggests that the same neural mechanisms of extracting orientation modulations from images will suffice for all these different conditions. Similarly, a discrete number of mechanisms tuned to extract frequency modulations can provide information about distances to different parts of the surface. In other words, rather than perform the reverse optics operations of assuming texture properties, estimating texture distortions from the image, and then reversing the projection transform to infer the 3D shape, the visual system might instead signal the presence of 3D shape features automatically from the outputs of a discrete number of matched filters configured for particular orientation and frequency patterns.

Our work differs from other computational approaches in the way that we have characterized the information present in perspective images of texture surfaces. There are an infinite number of ways to parse this information. Some of the ways that have been shown to be useful are deformation gradients (Garding, 1992), local affine deformations of

the spectrum of a pattern (Malik & Rosenholtz, 1997), and deformations of wavelets (Clerc & Mallat, 2002). We have parsed the information in terms of orientation and frequency modulations. This has been useful because orientation modulations are generically different for concavities, convexities, right slants and left slants, whereas frequency modulations are not. The corollary is that unless the texture pattern contains discretely oriented energy that distorts into signature orientation flows, the textured image will not contain information that is different for different signs of curvatures and slants. Consequently, to identify 3D shapes from texture cues, the minimum requirement for a visual system, machine or natural, is that it be able to extract orientation modulations and be able to differentiate between orientation modulations that are signatures for distinct 3D features. Further, as shown by the octotropic plaid pattern, only those visual systems will identify 3D shapes correctly that can extract the signature orientation modulations in the presence of distractor orientations. Thus, correct shape perception relies both on the information contained in the image, and on the capacity of the visual system to extract the relevant information.

In this study, we have looked at only a limited number of texture patterns and at upright corrugated solids and plaids, so it is worth examining whether these results generalize to naturally occurring texture patterns, 3D solids, and 3D shapes. For the case of homogenous textures on upright developable shapes, we have previously examined the Brodatz (1966) set of natural and man-made textures (Li & Zaidi, 2001b). For these texture patterns, we found that similar to synthetic patterns, visibility of the signature orientation modulations and the perception of correct curvatures and slants can be predicted by the discreteness of energy in the critical Fourier component. For example, for certain natural textures, such as wood with fairly parallel grain, shapes are perceived correctly or incorrectly depending on whether the axis of 3D curvature is parallel or orthogonal to the grain. These results are likely to generalize to non-developable surfaces because the oriented components that distort into the signature orientation modulations are the same as for developable surfaces. We have also shown that whereas the Fourier component parallel to the axis of maximum curvature is critical for upright corrugations, other components provide the signature modulations for pitched corrugations (Zaidi & Li, 2002), and that this is the reason why texture patterns can convey more varied shapes than the parallel contours explored by Stevens (Stevens, 1981). In addition, however, shape percepts are also correct for two nongeneric but theoretically important classes of images: first, if signature orientation modulations are defined solely by contrast variations (i.e., without Fourier energy) (Li & Zaidi, 2000), and second if the orientation modulations are created by illusions (see the pattern “Primrose Hill” on the website of Akiyoshi Kitaoka [link]).

In the 3D solids we simulate in this study, a pattern is repeated exactly through the solid. It is more likely that in solids such as marble or wood, the pattern changes slightly

in parallel planes. However, if the global spectrum does not change appreciably across planes, the information contained in images of carved solids will be similar to that described in this work. The depth plaids we chose to explore as examples of doubly curved surfaces are extended and periodic, whereas most objects in the world are limited and not periodic.

However, linear combinations of depth plaids of different frequencies can be used to synthesize many different shapes, and given the range of slants in each depth map, it seems probable that signature orientation flows will be the critical information for correct perception of all 3D shapes.

Our results that patterns of orientation modulations obviate the need to calculate texture gradients or assume homogeneity have implications for neural and computational models of shape from texture. Our results suggest that a neural implementation of the extraction of 3D shape-from-texture would require only a small number of mechanisms, each receiving input from local orientation sensitive operators configured in signature patterns of orientation modulations that represent individual 3D shapes. Other mechanisms receiving input from frequency sensitive operators would contribute supplementary inferences about relative distance along the surface. In preliminary work (Zaidi & Li, 2002), we showed that such matched filters had reasonable success in locating and identifying concavities and convexities by extracting the orientation modulations of the critical component from multiple orientations at each point. There were no false alarms from these matched filters, indicating that signature orientation modulations almost never occur accidentally. There were, however, misses when the orientation information was noisy or changed contrast. The hard-wired inputs for the matched filters were provided by V1-like orientation-tuned filters, which may be inadequate. As a front-end to the matched filters, we are experimenting with a population coding model of extracting local orientation, similar to that of

Fleming, Torralba, and Adelson (2004). In addition, the illusions presented by Kitaoka indicate that the perceived orientation modulations are affected by local lateral interactions among neighboring orientations in the image. Because lateral interactions in V1 predominantly affect the gains of neurons (Cavanaugh, Bair, & Movshon, 2002; Muller, Metha, Krauskopf, & Lennie, 2003), each feed-forward connection must come from a cluster of neurons with similar orientation selectivities. The number of matched filters at higher levels can be kept manageable by implementing them as deformable templates (Yuille, 1991) in recurrent feedback schemes like that proposed by Mumford (1992) and Lee and Mumford (2003). It remains to be tested whether this scheme can enable automatic object identification in natural scenes, especially for deformable surfaces like animal skins (Forsyth, 2002).

Appendix A

A.0. Psychophysical methods

To measure perceived slant along the surface, we used a local relative depth task similar to that used in Li and Zaidi (2000). Perspective images of the textured surfaces were presented against a background of mean grey at 44 cd/m². Surfaces were presented in one of four different central phases as shown in Figure A0. For the two images on the left, the projection was centered, respectively, to the left and right of a concavity (phase = $-\pi/8$ and $\pi/8$), and for the two images on the right, the projection was centered, respectively, to the left and right of a convexity (phase = $7\pi/8$ and $9\pi/8$). Thus the images at phases $-\pi/8$ and $9\pi/8$ were centered on rightward slanting portions of the surface, and at $\pi/8$ and $7\pi/8$ they were centered on leftward slanting portions of the surface. Each image contained two thin, red, vertical lines, each of which subtended

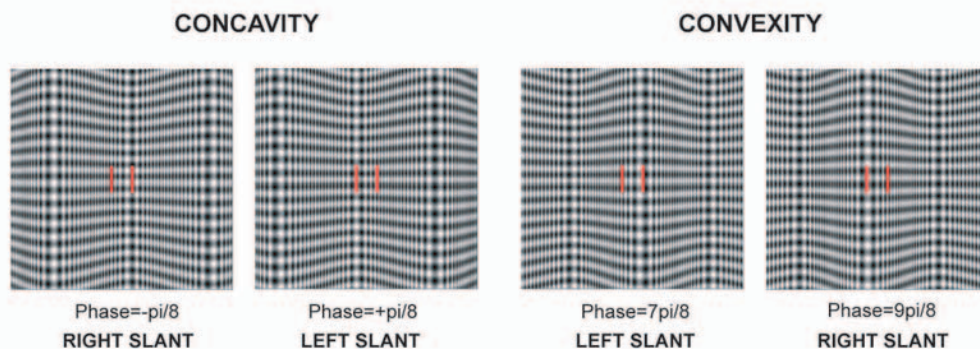


Figure A0. Example stimuli used in psychophysical experiments. For each surface type and texture pattern, the surface was presented in four different central phases: $-\pi/8$, $+\pi/8$, $7\pi/8$, and $9\pi/8$. The first two phases were centered slightly to the left and right of a concavity, and the latter two to the left and right of a convexity. Thin red vertical lines were placed 0.4 deg to the left and right of the vertical mid-line. One line was always at the center of the concavity or the convexity. For phases $-\pi/8$ and $9\pi/8$, the surface between the two lines was locally slanted to the right; for phases $+\pi/8$ and $7\pi/8$, it was slanted to the left. Observers judged which location on the surface as indicated by each of the two lines appeared closer to them in depth, or if they appeared at equal depths.

0.5 deg, displaced 0.4 deg to the left and to the right of the central vertical mid-line (0.8 deg apart). (In Figure A0, the lines have been thickened and lengthened for visibility.) One of the lines was always located at the center of either the concavity or the convexity. Observers were told that these lines indicated two locations directly behind them on the surface. The task was to indicate which of the two locations on the surface appeared closer to them, or if they appeared at equal depths. If the surface presented in phases of $-\pi/8$ or $9\pi/8$ (slanted to the right) was perceived correctly, observers should have responded that the left line appeared closer to them in depth. If the surface presented in phases of $\pi/8$ or $7\pi/8$ (slanted to the left) was perceived correctly, observers should have indicated that the right line appeared closer to them. If any surface appeared fronto-parallel, observers indicated that the two red lines appeared at equal depths.

Stimuli were generated using Matlab, and presented on a SONY GDM-F500 flat screen monitor with an 800 x 600 pixel screen running at a refresh rate of 80 frames/s via a Cambridge Research Systems Visual Stimulus Generator (CRS VSG 2/3) controlled through a 400-MHz Pentium II PC. Through the use of 12-bit DACs, after gamma correction, the VSG was able to generate 2861 linear levels per gun.

There were a total of 72 different images (6 texture patterns x 4 surface phases x 3 surface types). We divided the images by surface type, so in the first session, observers viewed images of developable surfaces, in the second, im-

ages of carved constant-z surfaces, and in the third, images of carved constant-x surfaces. Each image was presented 8 times for a total of 256 trials, presented in random order. Each session thus contained 16 presentations of a rightward slant for a particular pattern, and 16 presentations of a leftward slant for the same pattern. Viewing was monocular with the head position fixed in a chinrest at a distance of 1 m. At this distance, the retinal image coincided with that of a simulated 3D surface with the physical parameters shown in Figure 3. Each session began with 1 min of adaptation to a screen of mid-grey. After adaptation, each image was presented onscreen until the observer made a response via a response box. There was no feedback.

Data will be presented from three paid observers. All were naive about the purposes of the experiment, but had previously served as observers in similar psychophysical experiments. All had normal or corrected-to-normal acuity.

Figure A1 shows data averaged across the three observers for developable surfaces. Each panel represents data for one of the six texture patterns. The frequency with which each simulated slant (left or right) was reported as each of the perceived slants (left, right, or fronto-parallel) is indicated by the size of the dot in the graph, with the areas adding to unity along the vertical axis for each simulated slant. In each panel, data for the two phases representing right slants were collapsed, as were data for the two phases representing left slants. If all slants were perceived correctly, the graph should show two large dots along the diagonal.

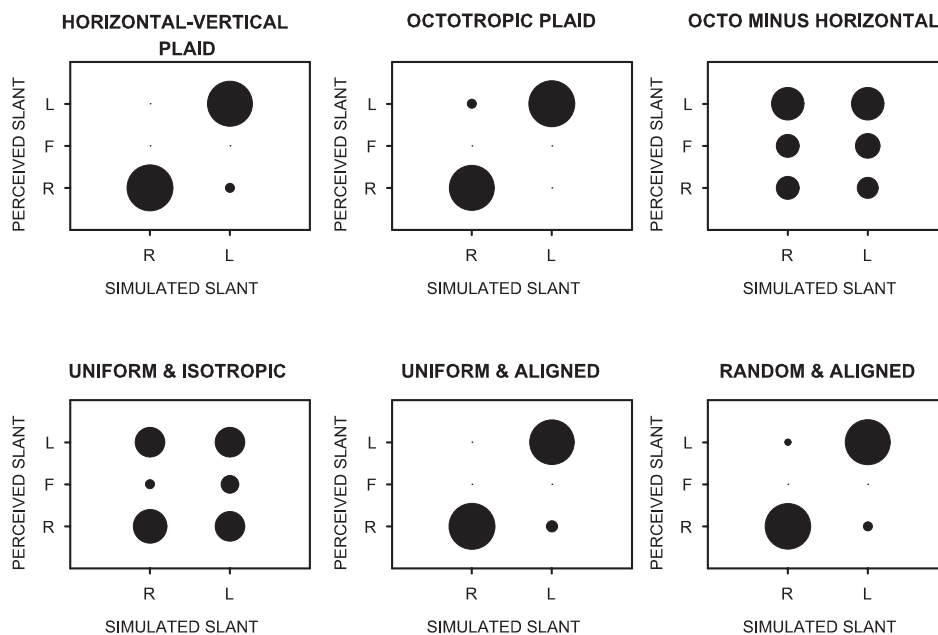


Figure A1. Data for developable surfaces. Frequency with which right and left slants are reported as each of the perceived slants is represented as the size of the dot. Observers made correct slant judgments for both plaids and aligned dot patterns. In the absence of the critical orientation flows, observers interpreted slant-caused frequency modulations as cues to distance, and as a result left slants and right slants were confused (isotropic pattern), and were sometimes reported as flat (octo minus horizontal).

A.1. Developable surfaces

For the horizontal-vertical plaid, the octotropic plaid, and the two aligned dot patterns, observers perceived right and left slants correctly. For the octotropic plaid minus the horizontal and the isotropic dot pattern, observers confused right slants for left and vice versa, and sometimes perceived both as fronto-parallel. Thus observers made correct slant judgments only when the orientation flows were visible. The confusion of right and left slants explains why concavities for the octotropic plaid minus the horizontal and the isotropic dot pattern appear convex (see Figure 6 and Figure 9). Slant-caused frequency modulations in the image (see Figure 8) are consistent with and interpreted as changes in distance, with low frequencies marking closer portions of the surface and high frequencies marking farther portions.

A.2. Carved constant-z solids

Results for carved constant-z solids are presented in the same format in Figure A2. Although observers judged slants correctly for the carved horizontal-vertical plaid solid, some slants were perceived as fronto-parallel. This is consistent with the slightly flattened percept of this solid conveyed in Figure 10. Results for the carved octotropic plaid solid show that the surface was perceived as flattened overall. However, when the $\pm 22.5^\circ$ components were subtracted from the planar pattern making up the solid, the orientation flows of the horizontal component were unmasked, and observers perceived right and left slants correctly. The carved aligned dot pattern solids also exhibited the critical

orientation flows and observers made correct slant judgments. For the isotropic dot pattern solid, these flows are absent. Observers interpreted distance-caused frequency modulations in the image (see Figure 11) correctly by making correct slant judgments in a small proportion of trials; however, most slants were perceived as fronto-parallel.

A.3. Carved constant-x solids

Figure A3 shows data for the carved constant-x solids. Observers made correct slant judgments for both plaids and both aligned dot patterns for which the critical orientation flows were visible. In the absence of the critical orientation flows (octotropic plaid minus the horizontal, isotropic dot pattern), observers confused right slants and left slants. This explains why in Figure 15 and Figure 18, concavities for these two patterns appear convex. As for developable surfaces, slant-caused frequency modulations in the image (see Figure 16) are misinterpreted as distance.

Appendix B

Orientation and frequency in perspective images of developable surfaces

In this appendix, we derive local orientation and spatial frequency in the perspective projections of oriented texture components for developable surfaces. The derivation incorporates offsets to the equations from the appendix of Zaidi and Li (2002) that enable computations of orientation and frequency at locations on the surface horizontally

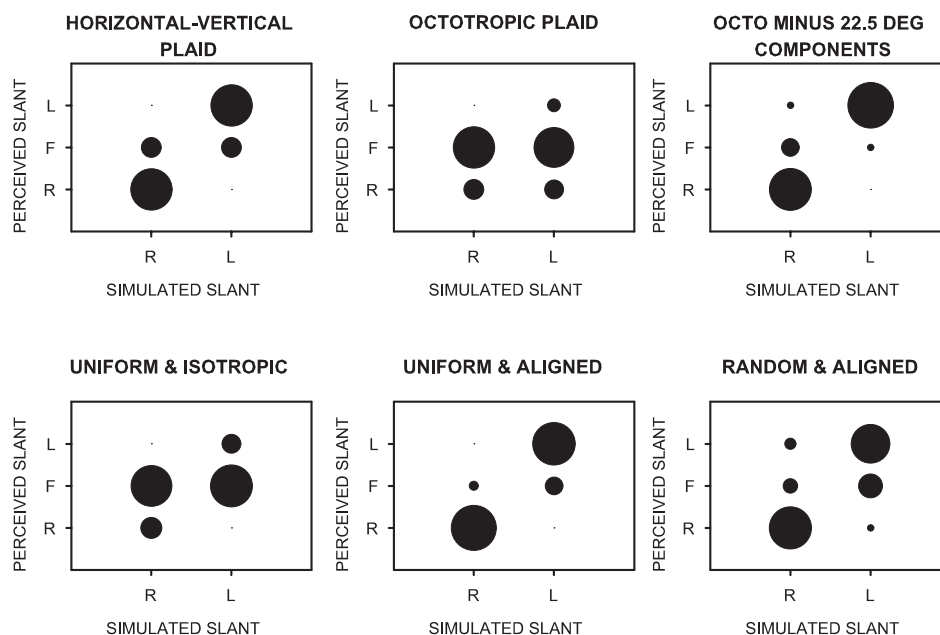


Figure A2. Data for carved constant-z solids. Observers made correct slant judgments for the horizontal-vertical plaid, octo plaid minus the 22.5° components, and aligned dot patterns. In the absence of the critical orientation flows (octo plaid, isotropic dot pattern), distance-caused frequency modulations led to percepts of flattened surfaces.

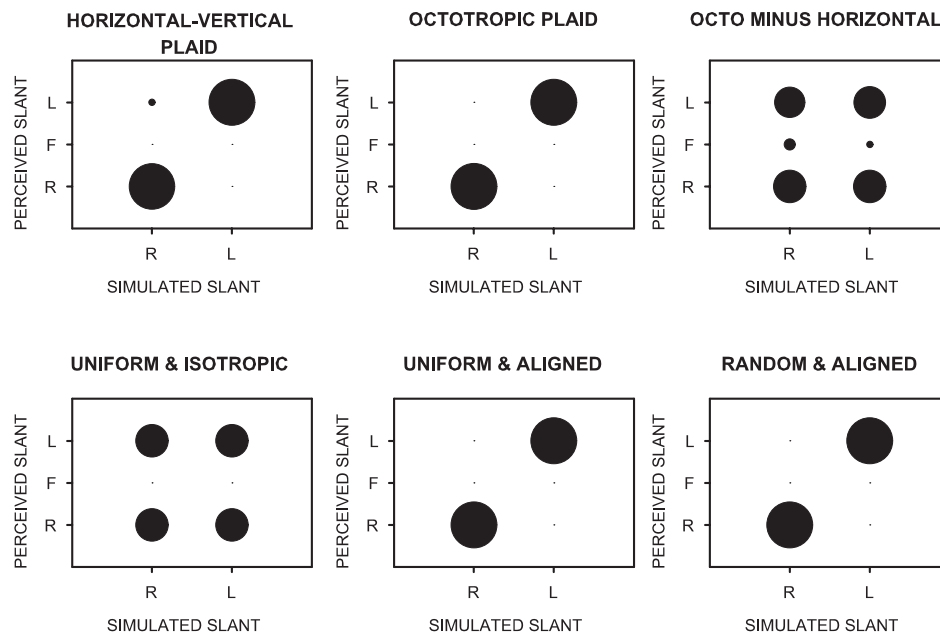


Figure A3. Data for carved constant-x surfaces. Observers made correct slant judgments for plaids and aligned dot patterns. In the absence of the critical orientation flows, observers interpreted slant-caused frequency modulations as cues to distance, and as a result left slants and right slants were confused (octo minus horizontal, isotropic dot pattern).

displaced from the line of sight (thus incorporating the effects of perspective), and locations on the surface that are displaced in depth from the image plane.

We start with a line of unit length oriented in the xy -plane at the angle of the texture component. The line is then slanted out of the xy -plane about a vertical axis at an angle equal to the local slant of the surface and its perspective projection in the xy -plane is computed. We also compute the perspective projection if the slanted line is additionally pitched about a horizontal axis at an angle equal to the pitch of the surface. The perspective coordinates of the

slanted line in the xy -plane then provide the projected orientation, and the projected frequency is equivalent to the inverse of the projected length of the line in the xy -plane.

The center of the image plane is defined as $(0, 0, 0)$ in 3D 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 start with the following parameters in radians:

- ω = orientation of the texture component from the horizontal axis in the xy -plane
- θ = local slant of the surface through the vertical axis
- α = pitch of the surface backwards through the horizontal eye-height line.

We consider a line of unit length, with one point at (x, y, z) (i.e., y units above eye-height), where z is the difference in depth between the surface and the image plane. Figure B shows views of this line in both the xy - (frontal) and xz - (aerial) planes. If the line were lying in the xy -plane at an angle of ω radians from the horizontal, the coordinates of the rightmost end point would be

$$(x + \cos\omega, y + \sin\omega, z).$$

If this line is slanted θ radians from the frontal plane through the vertical axis (i.e., slanted to the left or to the right), the coordinates of the end point would become

$$(x + \cos\theta \cos\omega, y + \sin\omega, z + \cos\omega \sin\theta).$$

The perspective image (u, v) of any point (x, y, z) is calculated as

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

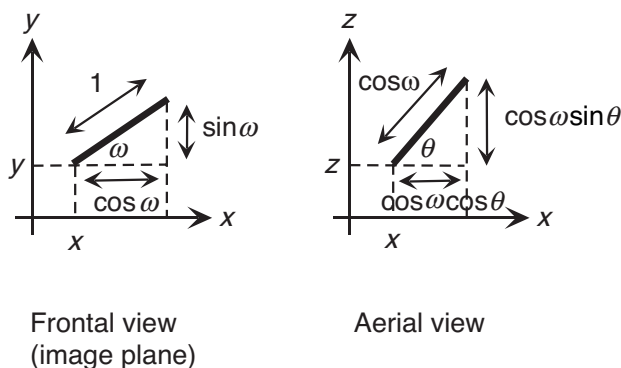


Figure B. Local orientation and frequency in the perspective image of a component oriented at ω on a developable surface are derived by taking a line of unit length in the image plane at the orientation of the component (ω), slanting it out of the fronto-parallel plane by an angle equal to the local slant of the surface (θ). Local orientation is computed as the orientation of the projected line, and local frequency of the component oriented at $(\omega + \pi/2)$ is the inverse of the length of the projected line.

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

In the perspective image, the line would extend from

$$(u_0, v_0) = [xd/(z+d), yd/(z+d)]$$

to

$$(u_1, v_1) = [d(x + \cos\theta \cos\omega) / (\cos\omega \sin\theta + z + d), \\ d(y + \sin\omega) / (\cos\omega \sin\theta + z + d)].$$

If the corrugation is pitched backwards α radians through the horizontal eye-height line, the 3D coordinates of the end-points of the line change to

$$(x, y\cos\alpha + z\sin\alpha, -y\sin\alpha + z\cos\alpha)$$

and

$$(x + \cos\theta \cos\omega, \cos\alpha(y + \sin\omega) + \sin\alpha(z + \cos\omega \sin\theta), \\ -\sin\alpha(y + \sin\omega) + \cos\alpha(z + \cos\omega \sin\theta)).$$

In the perspective image, the line would extend from

$$(u_0, v_0) = [xd/(-y\sin\alpha + z\cos\alpha + d),$$

$$d(y\cos\alpha + z\sin\alpha) / (-y\sin\alpha + z\cos\alpha + d)]$$

to

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

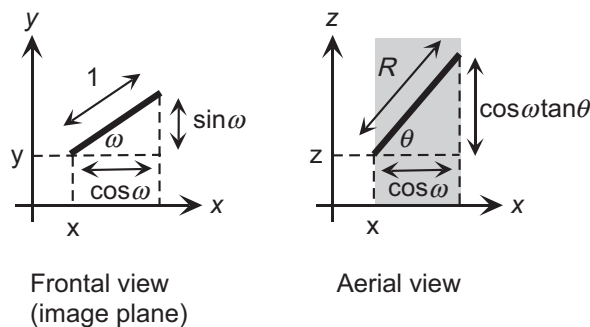


Figure C. Local orientation and frequency in the perspective image of a surface carved from a constant-z solid with a planar pattern of a component oriented at ω are derived by taking a line of unit length in the image plane at the orientation of the component (ω), repeating this line in depth along the z-axis, and carving the subsequently formed plane (shaded region is aerial view) at an angle of θ . Local orientation is computed as the orientation of the line on the planar cut (R), and local frequency of the component oriented at $(\omega + \pi/2)$ is the inverse of the length of the projected line.

The slope of the line in the perspective image is calculated as

$$\frac{v_1 - v_0}{u_1 - u_0},$$

and its length as

$$\sqrt{(u_1 - u_0)^2 + (v_1 - v_0)^2}.$$

The slope of this line provides the local projected orientation of a texture component at angle ω from the horizontal. Changes in the length of this line as a function of θ and α provide changes in local spatial frequency of the texture component oriented at $\omega + \pi/2$ radians.

Appendix C

Orientation and frequency in perspective images of surfaces carved from constant-z solids

For the developable surface derivation in “Appendix B,” we computed the perspective projection of a slanted, pitched line of unit length (representing the oriented texture component). In this derivation, each oriented component in the xy -plane is repeated along the z -axis, and we compute the perspective projection of the slanted carving through this solid. We also compute the perspective projection of the carving if it is pitched about a horizontal axis.

The center of the image plane is defined as $(0, 0, 0)$ in 3D 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 start with the following parameters in radians:

- ω = orientation of the texture component from the horizontal axis in the xy -plane
- θ = angle at which the surface is carved with respect to the xy -plane
- α = pitch of the surface backwards through the horizontal eye-height line, after it has been carved.

We consider a line of unit length, with one point at (x, y, z) (i.e., y units above eye-height), where z is the difference in depth between the surface and the image plane. Figure C shows views of this line in both the xy - (frontal) and xz - (aerial) planes. If the line were lying in the constant- z plane at an angle of ω radians from the horizontal, the coordinates of the rightmost end point would be

$$(x + \cos\omega, y + \sin\omega, z).$$

Given that all xy -planes of the solid material to be carved are identical, the x -projection of this line will be identical for all values of z (indicated by the shaded area in the aerial view). The surface is carved at θ away from the xy -plane.

The length of the cut in the xz -plane is R and its x -projection will equal $\cos\omega$. The projected orientation and frequency are computed from the endpoints of the cut $R = \cos\omega/\cos\theta$. The coordinates of the endpoints of the cut will be (x, y, z) and

$$(x + \cos\omega, y + \sin\omega, z + \tan\theta \cos\omega).$$

If the cut is pitched backward through the horizontal eye-height line by α , the coordinates then become

$$(x, y\cos\alpha + z\sin\alpha, -y\sin\alpha + z\cos\alpha)$$

and

$$(x + \cos\omega, \cos\alpha(y + \sin\omega) + \sin\alpha(z + \tan\theta \cos\omega),$$

$$-\sin\alpha(y + \sin\omega) + \cos\alpha(z + \tan\theta \cos\omega)).$$

In the perspective image (see “Appendix B” for conversion of 3D spatial coordinates to perspective image coordinates), the cut would extend from

$$(u_0, v_0) = [xd/(-y\sin\alpha + z\cos\alpha + d),$$

$$d(y\cos\alpha + z\sin\alpha)/(-y\sin\alpha + z\cos\alpha + d)]$$

to

$$(u_1, v_1) = \left(\frac{d(x + \cos\omega)}{-\sin\alpha(y + \sin\omega) + \cos\alpha(z + \tan\theta \cos\omega) + d}, \right. \\ \left. \frac{d\cos\alpha(y + \sin\omega) + d\sin\alpha(z + \tan\theta \cos\omega)}{-\sin\alpha(y + \sin\omega) + \cos\alpha(z + \tan\theta \cos\omega)} \right).$$

The slope of the cut in the perspective image is calculated as

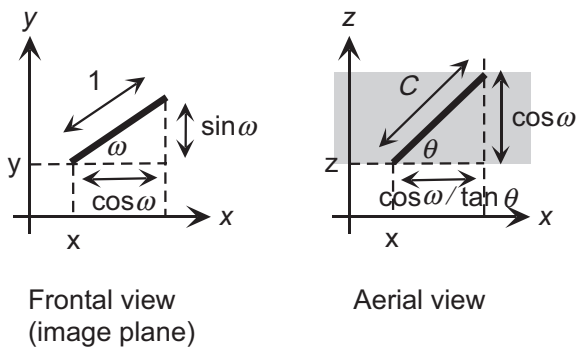


Figure D. Local orientation and frequency in the perspective image of a surface carved from a constant- x solid with a planar pattern of a component oriented at ω are derived by taking a line of unit length in the image plane at the orientation of the component (ω), repeating this line along the x -axis, and carving the subsequently formed plane (shaded region is aerial view) at an angle of θ . Local orientation is computed as the orientation of the line on the planar cut (C), and local frequency of the component oriented at $(\omega + \pi/2)$ is the inverse of the length of the projected line.

$$\frac{v_1 - v_0}{u_1 - u_0},$$

and its length as

$$\sqrt{(u_1 - u_0)^2 + (v_1 - v_0)^2}.$$

The slope of this cut provides the local projected orientation of a texture component at angle ω from the horizontal. Changes in the length of this cut as a function of θ and α provide changes in local spatial frequency of the texture component oriented at $\omega + \pi/2$ radians.

Appendix D

Orientation and frequency in perspective images of surfaces carved from constant- x solids

Every yz -plane of the constant- x solid contains the same oriented texture component that is repeated along the x -axis. This derivation computes the perspective projection of the slanted carving through this solid. We also compute the perspective projection if the carved solid is pitched about a horizontal axis.

The center of the image plane is defined as $(0, 0, 0)$ in 3D 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 start with the following parameters in radians:

- ω = orientation of the texture component from the horizontal axis in the xy -plane
- θ = angle at which the surface is carved with respect to the xy -plane
- α = pitch of the surface backwards through the horizontal eye-height line, after it has been carved.

We consider a line of unit length, with one point at (x, y, z) (i.e., y units above eye-height), where z is the difference in depth between the surface and the image plane. Figure D shows views of this line in both the xy - (frontal) and xz - (aerial) planes. If the line were lying in the constant- z plane at an angle of ω radians from the horizontal, the coordinates of the rightmost end point would be

$$(x + \cos\omega, y + \sin\omega, z).$$

Given that all yz -planes of the solid material to be carved are identical, the x -projection of this line will be identical for all values of x (indicated by the shaded area in the aerial view). The surface is carved at θ away from the xy -plane. The length of the cut in the xz -plane is C and its z -projection will equal $\cos\omega$. The projected orientation and frequency are computed from the endpoints of the cut $C = \cos\omega/\tan\theta$. The coordinates of the endpoints of the cut will be (x, y, z) and

$$(x + \cos\omega/\tan\theta, y + \sin\omega, z + \cos\omega).$$

If the cut is pitched backward through the horizontal eye-height line by α , the coordinates then become

$$(x, y\cos\alpha + z\sin\alpha, -y\sin\alpha + z\cos\alpha)$$

and

$$(x + \cos\omega/\tan\theta, \cos\alpha(y + \sin\omega) + \sin\alpha(z + \cos\omega), \\ -\sin\alpha(y + \sin\omega) + \cos\alpha(z + \cos\omega)).$$

In the perspective image (see "Appendix B" for conversion of 3D spatial coordinates to perspective image coordinates), the cut would extend from

$$(u_0, v_0) = [xd/(-y\sin\alpha + z\cos\alpha + d),$$

$$d(y\cos\alpha + z\sin\alpha)/(-y\sin\alpha + z\cos\alpha + d)]$$

to

$$(u_1, v_1) = \left(\frac{d(x + \cos\omega)}{-\sin\alpha(y + \sin\omega) + \cos\alpha(z + \tan\theta\cos\omega) + d}, \right. \\ \left. \frac{d\cos\alpha(y + \sin\omega) + d\sin\alpha(z + \tan\theta\cos\omega)}{-\sin\alpha(y + \sin\omega) + \cos\alpha(z + \tan\theta\cos\omega)} \right).$$

The slope of the cut in the perspective image is calculated as

$$\frac{v_1 - v_0}{u_1 - u_0},$$

and its length as

$$\sqrt{(u_1 - u_0)^2 + (v_1 - v_0)^2}.$$

The slope of this cut provides the local projected orientation of a texture component at angle ω from the horizontal. Changes in the length of this cut as a function of θ and α provide changes in local spatial frequency of the texture component oriented at $\omega + \pi/2$ radians.

Acknowledgments

This work was supported by National Eye Institute Grant EY13312 to QZ, and was presented in part at the Visual Sciences Society Meeting in Sarasota, FL, May, 2003, the European Conference on Visual Perception in Paris, France, August, 2003, and the Human Vision and Electronic Imaging Conference (SPIE) in San Jose, CA, January, 2004.

Commercial relationships: None.

Corresponding author: Andrea Li.

Email: ali@sunyopt.edu.

Address: Psychology Department, Queens College, 65-30 Kissena Boulevard, Flushing, NY 11367.

References

- Bracewell, R. N. (1995). *Two-dimensional imaging*. Englewood Cliffs, NJ: Prentice Hall.
- Brodatz, P. (1966). *Textures: A photographic album for artists and designers*. New York: Dover.
- Cavanaugh, J. R., Bair, W., & Movshon, J. A. (2002). Selectivity and spatial distribution of signals from the receptive field surround in macaque V1 neurons. *Journal of Neurophysiology*, 88(5), 2547-2556. [PubMed]
- Clerc, M., & Mallat, S. (2002). The texture gradient equation for recovering shape from texture. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(4), 536-549.
- Fleming, R. W., Torralba, A., & Adelson, E. H. (2004). Specular reflections and the perception of shape. Manuscript submitted for publication. *Journal of Vision*, 4(9), 798-820, <http://journalofvision.org/4/9/10/>, doi:10.1167/4.9.10. [PubMed][Article]
- Forsyth, D. A. (2002). Shape from texture without boundaries. *Proceedings of European Conference on Computer Vision*, 3, 225-239.
- Garding, J. (1992). Shape from texture for smooth curved surfaces in perspective projection. *Journal of Mathematical Imaging and Vision*, 2, 327-350.
- Lee, T. S., & Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *Journal of the Optical Society of America A*, 20(7), 1434-1448. [PubMed]
- Li, A., & Zaidi, Q. (2000). Perception of three-dimensional shape from texture is based on patterns of oriented energy. *Vision Research*, 40(2), 217-242. [PubMed]
- Li, A., & Zaidi, Q. (2001a). Erratum to "Information limitations in perception of shape from texture." *Vision Research*, 41(22), 2927-2942. [PubMed]
- 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. [PubMed]
- Li, A., & Zaidi, Q. (2003). Observer strategies in perception of 3-D shape from isotropic textures: developable surfaces. *Vision Research*, 43, 2741-2758. [PubMed]
- 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.
- Muller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (2003). Local signals from beyond the receptive fields of striate cortical neurons. *Journal of Neurophysiology*, 90(2), 822-831. [PubMed]

- Mumford, D. (1992). On the computational architecture of the neocortex II. The role of cortico-cortical loops. *Biological Cybernetics*, 66, 241-251. [[PubMed](#)]
- Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences*, 99(23), 15164-15169. [[PubMed](#)][[Article](#)]
- Stevens, K. A. (1981). The visual interpretation of surface contours. *Artificial Intelligence*, 17, 47-73.
- Stix, G. (1991). Profile: David Huffman. *Scientific American*, September, 54-58.
- Yuille, A. (1991). Deformable templates for face recognition. *Journal of Cognitive Neuroscience*, 3(1), 59-70.
- Zaidi, Q., & Li, A. (2002). Limitations on shape information provided by texture cues. *Vision Research*, 42(7), 815-835. [[PubMed](#)]