The role of shape-from-shading information in the perception of local and global form in Glass patterns

Sieu K. Khuu
School of Optometry and Vision Science, The University of New South Wales, Sydney, New South Wales, Australia

Amy Moreland
School of Optometry and Vision Science, The University of New South Wales, Sydney, New South Wales, Australia

Jack Phu
School of Optometry and Vision Science, The University of New South Wales, Sydney, New South Wales, Australia

Three-dimensional (3D) shape can be inferred from the surface shading gradient of objects. Using Glass patterns, we investigated the importance of shape-from-shading information to the perception of global form. We examined whether different 3D shapes inferred from shading affect the extraction of local dipole orientations (local analysis) and the integration of dipoles in the perception of Glass structure (global analysis). In Experiment 1, we showed that incongruence in shading between partner dots prevents the recovery of the dipole orientation: partner dots with different 3D shapes are not paired to recover the dipole orientation. However, when incongruent “bipartite” partner dots (that have the same contrast polarity as shaded dots, but are two-dimensional) were used, the visual system was able to extract the local dipole orientation and detect the global pattern (Experiment 2). In Experiment 3, we showed that additional noise dipoles affect the detection of Glass structure regardless of the 3D shape difference between signal and additional noise dipoles. This demonstrates that the visual system combines different oriented 3D dipoles to detect Glass structure. Our findings show that shape-from-shading information impacts on the ability to detect form structure but in different ways at local and global stages of processing.

Keywords: spatial vision, 3D surface and shape perception, detection/discrimination, object recognition, perceptual organization


Introduction

A primary goal of the visual system is to construct an accurate representation of the three-dimensional (3D) visual world from the two-dimensional (2D) retinal image. The computation of depth structure is highly reliant on a number of binocular and monocular “pictorial” cues. One of the more effective monocular cues that aids in the computation of depth is the surface shading gradient of objects (Brewster, 1826). Shading information provides an indication of depth structure because there is a regular relationship between the surface reflectance pattern of an object and the inferred position of the light source. The visual system capitalizes on this relationship and infers depth structure based on the direction of the surface shading gradient of objects. In a classic demonstration, Ramachandran (1988a, 1988b) showed that 2D circular tokens appear 3D, and either convex or concave, depending on the surface shading gradient. Assuming the ecologically valid “light-from-above” default (e.g., Adams, 2007), circular tokens with “light to dark” (light–dark) shading are interpreted as being “convex” and appear to “pop out” from the background, while tokens with the opposite “dark to light” (dark–light) shading appear to be “concave” (Figures 1B and 1D; see also Kleffner & Ramachandran, 1992). Despite these powerful visual demonstrations, it remains unclear what the conditions are under which shading information cues depth structure and the computations employed by the visual system to derive 3D shape from shading.

The relationship between the perception of depth and the direction of shading has been extensively investigated using the visual search paradigm, which examines the computational rules underlying feature integration at earlier stages of visual processing (e.g., Adams, 2007; Kleffner & Ramachandran, 1992; Sun & Perona, 1998; Thornton & Gilden, 2007; Treisman & Gormican, 1988; Wolfe, 2007).
Typically, this paradigm requires observers to detect a shaded target embedded within an array of differently shaded distracter tokens and examines the impact of the number of distracters on search efficiency. Significantly, it has been noted that incongruence in the shading direction of the target and distracters produces a “pop-out” effect, with the time taken to identify the object not affected by the number of distracters (this is especially so when searching for a concave object in a field of convex objects; the opposite configuration produces a lesser effect; Aks & Enns, 1992; Kleffner & Ramachandran, 1992; Ramachandran, 1988a, 1988b). By contrast, visual search is inefficient when target and distracter tokens have similar shading patterns or when the shading gradient is horizontal in orientation (Kleffner & Ramachandran, 1992). These findings suggest that shading information may provide a pre-attentive cue that can facilitate search (Castiello, Lusher, Burton, & Disler, 2003; Elder Trithart, Pintilie, & Maclean, 2004; Mamassian, Jentzsch, Bacon, & Schweinberger, 2003; Rensink & Cavanagh, 2004; though see Wolfe, 2007 for an alternative explanation) and that the visual system is selective for the direction of the surface shading gradient, which is important for the perception of shape.

While previous visual search investigations have focused extensively on the means by which shading information can be used as a cue to identify and perceptually segregate local features, comparatively very little is known about how shape-from-shading information facilitates the integration of local features in the perception of global form structure; such a process reflects computation beyond simple local feature detection. It is well recognized that complex global structures in the natural world comprise different 3D shapes (for example, a forest floor). The present study sought to establish whether, and how, different local structures are combined to extract the spatial layout of objects. The importance of this issue stems from the observation that while homogeneous/congruent shading of texture elements can lead to the perception of global form, as it may provide the basis for perceptual grouping (for example, the perception of radial structure in Figure 1A), it is entirely unclear whether perceived global form is affected when the shading of local elements is incongruent (i.e., in different or opposite directions). Given that the visual system may treat oppositely shaded textures separately as convex and concave in 3D shape, such elements may not be associated, reducing the perceptibility of the global pattern. This would accord with previous investigations showing that a cluster of similarly shaded elements within an array of oppositely shaded elements is grouped as a subset that then “pops out” from the array as a whole (see, e.g., Kawabe & Miura, 2004; Kleffner & Ramachandran, 1992). However, while differently shaded elements may produce a perceptually different 3D shape, this does not automatically mean that they are not combined. It has been well established that the mechanisms responsible for the computation of global form may integrate local elements regardless of whether or not they are perceptually different. For example, global form sensing mechanisms are able to integrate local features defined by different colors (Cardinal & Kiper, 2003), contrast (Or, Khuu, & Hayes, 2007; Wilson, Switkes, & De Valois, 2004), or across different depths (see, e.g., Earle, 1985; Khuu & Hayes, 2005) to detect global structure. If the visual system were able to integrate differently shaded/3D elements, global form would be unaffected regardless of the degree of incongruence between local elements. The present study investigated whether and how congruently and incongruently shaded elements are associated at both the local and global levels in the computation of global form.
Local and global form processing was investigated in this study using Glass patterns (after Glass, 1969; see Figure 1). Glass patterns are random dot stimuli consisting of dot pairs (dipoles) whose orientation conforms to some common geometric rule. For example, a concentric Glass pattern is produced by orientating dipoles 90 degrees to lines passing through the center of a region, while a radial pattern is constructed by placing dipoles along radii. Glass patterns are commonly used to investigate form processing because their analysis by the visual system must reflect both local and global levels of computation: the orientation of local dipoles is initially extracted, and then combined at a later stage where the global form can be determined (Dakin, 1997; Wilson & Wilkinson, 1998). It is believed that the extraction of local form information reflects computation in early cortical areas such as primary visual cortex (V1) and secondary visual cortex (V2), which contains cells capable of detecting the orientation of dipoles (see Smith, Bair, & Movshon, 2002; Smith, Kohn, & Movshon, 2007), while integration of local form information occurs at much higher cortical areas such as V4 where global detectors tuned to complex structures such as concentric and radial global forms exist (see, e.g., Chen, Chang, Liu, Chen, & Han, 2004; Gallant, Connor, Rakshit, Lewis, & Van Essen, 1996; Wilkinson et al., 2000). Glass patterns are additionally advantageous in that a measure of human sensitivity to global form (analogous to global dot motion stimuli: e.g., Badcock & Khuu, 2001; Khuu & Badcock, 2002; Newsome & Pare, 1988) can be obtained by changing the ratio between dipoles oriented in the pattern direction (signal dipoles) and those that have random orientations, until the global structure can be just detected. Typically, Glass detection thresholds are 15–30% signal dipoles (Badcock, Clifford, & Khuu, 2005; Khuu & Hayes, 2005; Wilson & Wilkinson, 1998). Using this method, previous studies have clarified the relative importance of stimulus attributes such as luminance polarity (Badcock et al., 2005; Or et al., 2007; Or, Khuu, & Hayes, 2010; Wilson et al., 2004), color (Cardinal & Kiper, 2003), and stereopsis (Khuu & Hayes, 2005) to the perception of global form. The present study used a similar experimental approach in seeking to examine the relative importance of shape-from-shading information to the detection of Glass pattern structure.

Since form information in Glass patterns is analyzed in two computational stages, we examined the role of shape-from-shading information in the extraction of the orientation of dipoles (local analysis) and in the detection of Glass structure embedded in noise (global analysis). In Experiment 1, we examined whether the visual system is able to group incongruently shaded partner dots (and, therefore, perceptually different 3D shapes) to extract the local dipole orientation—a local analysis. In Experiment 2, we examined the importance of the perception of 3D shape on grouping by comparing the results of Experiment 1 with Glass pattern stimuli consisting of “bipartite” partner dots that had a square-wave luminance profile. These stimuli have the same luminance polarity as shaded elements but do not convey a strong impression of 3D shape. We, therefore, expected that thresholds would not be dependent on the shading orientation of dipoles. Finally, in Experiment 3, we examined the importance of shape-from-shading information to the integration of dipoles at the global level by quantifying the extent to which Glass pattern detection thresholds are affected by noise dipoles that have the opposite 3D shape (i.e., a different shading direction) to signal dipoles.

**Experiment 1: The effect of direction of shading on the extraction of local dipoles in the perception of Glass structure**

To determine the orientation of dipoles, partner dots must be detected and paired by a common mechanism at the local stage of analysis. After the orientation of dipoles is extracted, they are later combined to detect the Glass structure. In Experiment 1, we investigated whether and how congruence in shading between partner dots affects their pairing and the consequent appearance of Glass structure. As mentioned, different shading gradients may lead to different interpretations of 3D shape. We, therefore, investigated whether the visual system is able to pair perceptually concave and convex partner dots to extract the orientation of dipoles.

To carefully examine this issue, Glass pattern detection thresholds were measured using patterns in which the number of congruently shaded dipoles was varied. We expected that, if shading incongruence impacts on the recovery of dipoles, Glass pattern thresholds would be optimal when all partner dots (forming a dipole) are congruent (and, therefore, have the same 3D shape), but thresholds would be negatively impacted when partner dots are incongruent. However, if shading congruence does not impact detection, thresholds would remain unchanged. Additionally, we measured and compared the detectability of global structure for patterns in which the direction of shading within dipoles was either light–dark or dark–light (from top to bottom). Given the “light from above” prior, these patterns appear to consist of either convex or concave dipoles (see Figures 1A and 1C). Note, however, that the Glass pattern depicted in Figure 1C can also be perceived as being convex, with the light source coming from below. This ambiguity in 3D form (for dark–light shaded elements) has been previously noted (see, e.g., Kleffner & Ramachandran, 1992; Ramachandran, 1988a, 1988b) and its implication for the detection of global form...
is further considered in the Results and discussion section of this experiment.

Methods

Observers

Six observers (age: 20–34) participated in this study. Two were authors, while the others were experienced observers but naive to the goals of the study. All had normal or corrected-to-normal visual acuity.

Stimuli

The stimulus consisted of 200 non-overlapping dipoles randomly positioned in a 25° × 25° stimulus area set to a background luminance of 70 cd/m². Dipoles consisted of two similar sized circular dots that were separated by a fixed (center to center) dot distance of 0.6°. Thus, the dipole density of the stimulus was uniformly 0.32 dipole/deg². Each dot forming a dipole had a smooth shading pattern that transitioned from one side of the dot to the other in a sinusoidal fashion along the 315/135° direction, making the axis of shading from top-left to bottom-right. This shading orientation was used because previous studies have noted that the visual system is more sensitive to 3D structures that are consistent with the light source above and to the left (Mamassian & Goutcher, 2001; McManus, Buckman, & Woolley, 2004; Sun & Perona, 1998). Shaded dots were created by windowing a section of sinewave grating (spatial frequency: 1 cpd, Michelson contrast: 0.96), with a circular aperture that was hard-edged (radius: 0.125°). Specifically, to produce this “black–white” shaded dot, the phase of the sinewave grating was shifted by 135° and the resulting dot pattern transitioned from increment to decrement with the luminance balance point (corresponding to the background luminance) coinciding with the center of the dot. Thus, each dot had absolute contrast of 43% and balanced regions of positive and negative polarities. This study used only Glass patterns conveying radial structure. Radial structures were used as they are ubiquitous in natural scenes, and neural mechanisms that are selective for radial form have been shown to exist (Chen et al., 2004; Gallant et al., 1996). These were produced by aligning signal dipoles with radii from the center of the stimulus. Additionally, signal dipoles were either convex (Figure 1A) in shape conveyed by light–dark shading or concave in shape (Figure 1C) conveyed by dark–light shading.

Observers viewed the stimulus binocularly in a dark room at a viewing distance of 60 cm. To ensure steady fixation, a black cross was presented in the middle of the stimulus area and observers were instructed to fixate on this mark throughout the duration of the experiment. Stimuli were generated using MATLAB version 7 and displayed on a linearized 24-inch Mitsubishi Diamond Pro monitor driven at a frame rate of 100 Hz.

Procedure

As mentioned, the goal of the present study was to examine the impact of local shading congruence on the perception of Glass structure. To measure sensitivity to Glass structure, we used a two-interval forced-choice paradigm (2IFC) in conjunction with a staircase procedure. One interval presented a radial Glass pattern, while the other presented a Glass pattern that had the same number of dipoles but was one in which dipoles were randomly oriented. The order of interval presentation was randomized between trials. Both intervals were shown for 1 s and were separated by a period of 1 s, during which time a blank screen was shown at the background luminance. The task of the observer was to judge the interval containing the radial pattern by pressing appropriate keys on a keyboard. A staircase procedure that converged on the 79% correct performance level (three correct responses to step down, one incorrect response to step up) was used to modify the number of signal dipoles required for observers to just reliably discriminate the interval containing the signal Glass pattern. The staircase began with an initial signal level of 80% (i.e., 160 dipoles) and the initial step size of the staircase was 8 dipoles. On the first and subsequent reversal, the step size was halved. After the third reversal, the step size was 1 dipole and remained at this value until the end of the trial. The staircase lasted for 8 reversals and the average of the last 4 reversals provided an indication of the Glass pattern detection threshold.

The staircase procedure was repeated for Glass patterns in which dipoles had congruent or incongruent shading gradients. Congruent dipoles consisted of partner dots in which shading orientation was identical. Incongruent dipoles consisted of partner dots that had opposite shading orientations such that one dot was shaded light–dark while its partner was dark–light. Across conditions, the percentage of dipoles that were congruent in both signal and noise intervals was varied. There were 5 congruence levels: 0, 25, 50, 75, and 100%. Thus, at 0% (e.g., Figures 1B, 1D, and 1F), all partner dots in both the signal and noise intervals had opposite shading gradients, while at 100% all partner dots had the same shading orientation (e.g., Figures 1A, 1C, and 1E).

All procedures were repeated for stimuli in which dots were shaded but were increment—smoothly transitioning from white to background (see Figure 1D)—or decrement—smoothly transitioning from black to background. These dots were created by selectively windowing the appropriate increment and decrement regions of the sinusoidal shading pattern. This was achieved by shifting the sinewave grating used to create the dots’ shading profile by 90° to produce increment shading or by 180° to create decrement shading. These procedures ensured that increment and decrement dots had the same absolute relative contrast of 43%, the same as the previously described black–white shaded dots that appear perceptually
3D. As increment and decrement Glass dots do not convey a strong sense of 3D shape (see, e.g., Elder et al., 2004), the visual system may be able to extract the dipole orientation and detect the global form regardless of the shading congruence of partner dots. Thus, sensitivity to these patterns provides a good comparison to the 3D shape from shading stimuli previously described.

A block comprised 30 staircase runs: five levels of shading congruence (0, 25, 50, 75, and 100%) for three different shaded dot types (increment, decrement, and black–white), for two dot shading directions (either light–dark or dark–light). Observers each completed 5 blocks and results were averaged over the 5 thresholds for each condition. Order of stimulus presentation was randomized within and between each block. Observers were given at least 2 practice trials to familiarize themselves with the task prior to data collection. No feedback was given to indicate the correctness of response.

Results and discussion

The results of Experiment 1 are shown in Figure 2. Glass pattern detection thresholds in terms of the minimum percentage of signal dipoles required to detect the global structure are plotted as a function of the percentage of congruently shaded dipoles for black–white (triangles), increment (square), and decrement (circle) dot patterns. In Figures 2A and 2B, the results for light–dark shaded and dark–light shaded patterns are shown, respectively. Results from the five observers were similar and were, therefore, averaged. Error bars represent one standard error of the mean.

The results indicated a number of interesting findings. First, regardless of the percentage of congruent dipoles and the direction of shading (i.e., light–dark or dark–light), Glass pattern detection thresholds for increment (squares) and decrement (circles) patterns were the same and remained unchanged at approximately 25–30% signal dipoles. These results show that simply changing the shading direction of partner dots has no impact on Glass pattern detection. Arguably, this is because, irrespective of the shading direction, partner dots do not convey a strong sense of 3D and the visual system is, therefore, able to appropriately pair them to detect the global pattern. Verification of this effect is obtained by comparing Figures 1D and 1E, which show 100% and 0% congruent increment patterns, respectively. Consider Figure 1F: though the shading pattern of partner dots is incongruent, radial structure is still perceptible and appears equal in salience to the congruent pattern depicted in Figure 1E.

Second, the results obtained for increment and decrement Glass patterns differed markedly from those for patterns created with black–white shaded dots that strongly convey 3D shape from shading (triangles). For those patterns, detection thresholds depend on the percentage of congruently shaded dipoles. For conditions in which there was no, or a small percentage of, congruent dipoles, Glass pattern thresholds were 75% signal dipoles. This represents an approximate three-fold increase in detection threshold compared to increment and decrement patterns at that level. Note that global structure is not lost entirely from these patterns and remains perceptible at signal levels above this point. However, as the percentage of congruent dipoles increases, the Glass pattern structure becomes more perceptible, and at a congruence level of 100%, thresholds are comparable to those of increment and decrement patterns. Additionally, for 3D patterns there was no difference in threshold between light–dark and dark–light shaded Glass patterns. For these stimuli, the shading direction infers different 3D shape, such that dots were either convex (light–dark elements) or concave (dark–light elements). While previous research has noted that convex structures are processed much faster than concave structures in visual search (e.g., Ramachandran, 1988a, 1988b), the present study shows that this anisotropy does not
not exist when detecting global form. Given that the stimulus duration was sufficiently long for the visual system to accurately compute the 3D shape from shading of objects, this is perhaps unsurprising.

The finding that shape-from-shading information impacts on the recovery of dipole orientation and global form is analogous to previous reports that Glass pattern structure is largely lost if partner dots are of opposite polarity (see, e.g., Badcock et al., 2005; Glass & Switkes, 1976; Kovacs & Julesz, 1992) but recovers visibility if the pattern is in motion (see Or et al., 2007, 2010). For example, Badcock et al. (2005) reported that the form for Glass patterns consisting of opposite polarity partner dots was only detectable at very high signal levels, a finding that mirrors observations from incongruent dipoles in the present study. The results of Badcock et al. (2005) can be attributed to the fact that information from “on” and “off” channels at the local stage of analysis remains separate and does not combine to recover the orientation of dipoles. However, careful inspection of our perceptually 3D stimulus shows that this explanation cannot account for the results of the present study. The black–white shaded dots (those that induce 3D shape from shading) contain a balanced amount of increment and decrement regions, and changing the direction of shading does not affect this balance; such dots will equally activate both “on” and “off” channels regardless of their shading direction. If polarity-selective channels accounted for our findings, one would expect Glass pattern detection thresholds to remain the same irrespective of the congruence of dipoles. This was not observed in the present study, rather the opposite effect. Thus, the basis for the segregation of partner dots must be due to a difference in perceived 3D shape. In the same way, the results of the present study cannot also be accounted for by the selective operations of “second-order” detectors. While black–white shaded dots will drive second-order detectors, such detectors are equally sensitive to both congruent and incongruent partner dots and, therefore, will not be able to discriminate between them.

As Figures 2A and 2B show, perceptually 3D patterns (triangles) consisting of light–dark (convex) and dark–light (concave) shaded dots give similar results indicating that they are equally affected by the number of congruently shaded dipoles. At first, these findings seem inconsistent with a well-noted asymmetry in the visual search of shaded elements: search for a concave element among convex distracters is efficient, while search for a convex target among concave is less so, with reaction time increasing markedly with set size (see Aks & Enns, 1992; Kleffner & Ramachandran, 1992; Sun & Perona, 1998). This asymmetry is thought to stem from the fact that the visual system naturally assumes an overhead light source (i.e., a “light-from-above” prior; see, e.g., Adams, 2007; Ramachandran, 1988a, 1988b; Thomas, Nardini, & Mareschal, 2010), making a light–dark shaded element unambiguously convex in shape. By contrast, the 3D shape of a dark–light dot is more ambiguous, and though a “light-from-above” prior would bias the 3D shape of this dot so that it appears concave, it can also appear convex if the light source is assumed to come from below. The implication for this observation is that a similar asymmetry may exist for the detection of Glass patterns. For example, making the dipoles for light–dark shaded convex patterns incongruent (by reversing the shading direction of one partner dot from light–dark to dark–light) may not greatly impact the perception of Glass structure because the changed dot may remain perceptually convex like its partner. If this were the case, the visual system would be able to recover the local orientation of dipoles. However, our data do not suggest this outcome, nor is it consistent with perception of a pattern such as that illustrated in Figure 1B—the dark–light shaded dots in that image are clearly concave in shape. This can be accounted for simply by the observations of Kleffner and Ramachandran (1992). They noted that when light–dark and dark–light shaded elements are viewed in conjunction, their 3D shape is not ambiguous nor can they be seen as simultaneously convex and concave; they appear unequivocally as 3D opposites. If light is assumed to come from above and uniformly illuminate the entire visual scene (Ramachandran, 1988a, 1988b), light–dark shaded elements must be convex, biasing oppositely shaded dark–light elements to appear concave. Certainly in the absence of light–dark convex dots, the 3D form of dark–light dots can be ambiguous depending on the assumed position of the light source.

Another major difference between our study and those using visual search is the relative proportions of concave and convex elements in the stimulus. A typical visual search task presents a single target of a particular shading direction in an array of distracters with the opposite shading direction. In the present study, patterns consisted of many of “concave” and “convex” elements. Under these conditions, perceptual grouping between similarly 3D-shaped elements may increase their salience. This accords with the findings of Kawabe and Miura (2004) who showed that the noted asymmetry in visual search between concave and convex elements is not observed when the target stimulus consists of a group of identically shaded elements.

In summary, Experiment 1 showed that the shading congruence of partner dots affects grouping and the subsequent recovery of local orientation. At the local stage of analysis, 3D shape from shading is used to segregate partner dots, preventing their pairing and negatively impacting on the ability to detect global structure. This effect can be informally confirmed by comparing the perceptibility of Glass structure in congruent and incongruent patterns shown in Figures 1A–1D. In Figures 1A and 1C, all dipoles are signal and shaded in the same direction. For those patterns, the radial structure is readily
Experiment 2: The effect of shading congruence on the perception of bipartite Glass patterns

In Experiment 1, the perceived 3D shape of black–white shaded dots was given exclusively by their smooth luminance profile that conformed to a sinusoidal modulation. If this cue were removed, dots would appear 2D or “flat.” Given the lack of perceived 3D structure, it might be predicted that the visual system would pair such dots to detect the local orientation of dipoles (despite incongruence in shading direction) in much the same way as it did the increment and decrement dipoles used in Experiment 1. One way to examine this possibility is to employ “luminance-step” or “bipartite” dots in which the luminance profile changes abruptly. In effect, such dots are half-white and half-black (see Figure 3). Bipartite dots have been frequently used in past studies as a natural comparison to perceptually 3D dots (see, e.g., Kawabe & Miura, 2004; Ramachandran, 1988a, 1988b) because they have the same contrast polarity but do not convey depth structure. In Experiment 2, we examine the extent to which the visual system is able to detect Glass structure conveyed by bipartite dots and compare these results with those obtained with 3D patterns in Experiment 1. The same observers as in Experiment 1 participated in Experiment 2. Glass patterns were generated using the same procedures as in Experiment 1, but bipartite dots were created using a square-wave luminance profile. Only Glass patterns containing light–dark shaded dots were examined. As in Experiment 1, we measured Glass pattern detection thresholds for different percentages of congruently shaded dipoles.

The findings of Experiment 2 with bipartite patterns (circles) are shown in Figure 4 along with the results for perceptually 3D patterns (consisting of black–white shaded dots) from Experiment 1 (triangles). As predicted, thresholds for detecting Glass pattern structure from bipartite dots were unaffected by the percentage of incongruently shaded dipoles: thresholds for these patterns remained at approximately 30–40% signal dipoles across all congruence levels. This contrasts markedly with the results for the perceptually 3D patterns used in Experiment 1, where a dependence on shading congruence was noted (see Experiment 1).

The results of Experiment 2 can be accounted for by the fact that bipartite dots do not imply 3D structure and are perceptually 2D regardless of their shading direction. Thus, the visual system is able to pair partner dots and, subsequently, perceive the global pattern. Indeed, the Glass structure is perceptible for both congruent and incongruent bipartite patterns shown in Figure 3. This result conclusively shows that the results of Experiment 1 arise from an...
Experiment 3: The effect of shading congruence on the integration of dipoles in the perception of Glass patterns

Experiments 1 and 2 showed that shading congruence affects the extraction of local signal dipoles only if elements convey 3D structure. In Experiment 3, we examined whether shading congruence impacts on the ability to detect form at the global stage of processing. This was achieved by measuring the degree to which the ability to detect Glass structure is affected by the addition of noise dipoles to the pattern—randomly oriented dipoles that are incongruently shaded relative to signal dipoles. It has been well established that Glass pattern detection thresholds increase with the addition of noise dipoles, provided that all dipoles are processed by a common mechanism (see, e.g., Badcock et al., 2005; Khuu & Hayes, 2005). Therefore, by systematically changing the percentage of additive noise, it is possible to establish the dependence of global form mechanisms on 3D shape-from-shading information. If thresholds are unaffected by the addition of incongruently shaded noise dipoles (i.e., thresholds are at the same level as when no noise dipoles are present), separate global form processing mechanisms for the analysis of 3D shapes would be implied. However, if the visual system uses a common mechanism to combine dipoles of different 3D shape, additional noise dipoles would interfere with the integration of signal dipoles regardless of their shading congruence.

Methods

Observers were the same as those in the previous experiments. Stimuli were Glass patterns constructed in a manner similar to those described previously but had half the dipole density (at 100 dipoles, 0.16 dipole/deg²; see Figure 5A). Partner dots forming a dipole always had the same shading direction and were either light–dark producing a “convex” pattern or dark–light producing a “concave” pattern. To this pattern, additional 100 noise dipoles were added. Additional noise dipoles had either identical (e.g., Figure 5B) or opposite (e.g., Figure 5C) shading gradients as signal dipoles. In different conditions, we varied the percentage of congruent additional noise dipoles at the levels: 0, 25, 50, 75, and 100%. Thus, at 0% all additional noise dipoles had the opposite shading gradient to signal dipoles, while at 100% they were the same. As in Experiment 1, we examined sensitivity to Glass patterns consisting of black–white dipoles and increment or decrement shaded dipoles. Conditions were repeated for signal dipoles shaded light–dark and dark–light. A baseline condition was conducted for each type of Glass pattern in which detection of Glass structure was measured in the absence of any additional noise dipoles. This condition allowed us to examine the extent to which additional noise dipoles interfere with the perception of Glass structure.

A block comprised 33 staircase runs: five levels of shading congruence, repeated for the three Glass pattern types, for the two shading directions (light–dark and dark–light), plus three baselines for each of three Glass pattern types. Observers each completed 5 blocks and results were averaged over the 5 runs for each condition. Stimulus conditions were randomized within and between each block. Observers were given two practice runs to familiarize themselves with the task. No feedback was given to indicate the correctness of their response.

Results and discussion

The results of Experiment 3 are shown in Figures 6A (light–dark shaded Glass pattern) and 6B (dark–light shaded Glass pattern). The threshold in terms of the percentage of signal dipoles (averaged over five observers)
required to detect the Glass structure is plotted as a function of the percentage of additional noise dipoles that had shading gradients congruent with signal dipoles. Different symbols represent different Glass pattern types (increment: square, decrement: circle, black–white shaded dipoles: triangles). Open symbols indicate thresholds for the baseline conditions.

A number of trends are evident in these data. First, for increment and decrement Glass patterns, additional noise dipoles did interfere with the ability to detect signal dipoles regardless of the percentage of incongruently shaded additional noise dipoles. Glass pattern detection thresholds were between approximately 22 and 27% signal dipoles and approximately 1.7 times the threshold of the corresponding baseline condition (open circle, approximately 12.5% signal dipoles).

Second, the pattern of results was equivalent for light–dark and dark–light shaded patterns. To illustrate, consider the increment radial pattern shown in Figure 5D. This pattern is equally masked by congruent (Figure 5E) and incongruent additional noise dipoles (Figure 5F). This result might have been expected because increment and decrement shaded dipoles are not perceived as being 3D, making their perceived shape unaffected by shading direction. Therefore, additional noise dipoles interfere with the detection of the signal dipoles because both dipole types are similar in 3D shape. Different global form detection mechanisms seem not to be required to process these two types of shading patterns because they are locally similar.

Third, and perhaps most interestingly, results from perceptually 3D patterns (triangles) consisting of “black–white” shaded dipoles showed a similar trend to those observed with increment and decrement patterns: regardless of the percentage of additional noise dipoles with incongruent shading, Glass pattern detection thresholds were approximately 1.5 to 2 times the corresponding baseline condition (open triangle). Results to this effect are observed for both “light–dark” (“convex”) and dark–light (“concave”) shaded patterns. That thresholds were so markedly affected suggests that, at the global stage of form processing, the visual system combines local dipoles to extract the global pattern irrespective of their perceived 3D shape from shading. This contrasts with results from Experiment 1, which showed that shape from shading prevents the grouping of partner dots to extract the local dipole orientation. Visual inspection of Figures 5A–5C provides conformation of this effect. Figure 5A shows a Glass pattern consisting of 100 perceptually convex dipoles. Figure 5B shows the same Glass pattern with additional 100 noise dipoles that have the same shading direction as signal dipoles. In this case, noise dipoles affect the detection of signal dipoles because both have the same 3D shape. In Figure 5C, additional noise dipoles have incongruent shading, and though they appear to be distinctly concave in shape and perceptually segregated from
“convex” signal dipoles, they nevertheless impair the perceptibility of Glass structure. Indeed, the effect of additional noise dipoles on Glass structure is comparable for both congruent (Figure 5B) and incongruent (Figure 5C) shading conditions.

General discussion

The goal of the present study was to examine the influence of shape-from-shading information on the perception of image structure. Previous research has well established that form analysis occurs in at least two computational stages: local and global. The present study finds that shape-from-shading information facilitates the detection of form at both computational levels.

In Experiment 1, we investigated the role of shape-from-shading information in detecting the orientation of dipoles. Sensitivity to global form was measured for patterns in which partner dots forming a dipole had congruent and incongruent shading gradients. Glass pattern detection thresholds remained unchanged for increment and decrement patterns regardless of the percentage of incongruently shaded dipoles. However, for black–white shaded Glass patterns (which are perceptually 3D in shape), global form detection was dependent on the percentage of congruently shaded dipoles, with thresholds negatively impacted at low levels. Thus, when these partner dots are incongruent they are perceived as 3D opposites, and the visual system is unable to pair them to extract the dipole orientation. However, Experiment 2 showed, using balanced bipartite dots that do not convey 3D shape but that have the same contrast polarity as black–white shaded dots, that the visual system is able to detect the local dipole orientation irrespective shading congruence. These results collectively show that, at the local stage of analysis, shape from shading can be used as a cue to segregate different 3D structures that are processed independently by the visual system. This concordance in results is not surprising since the techniques used in Experiment 1 and visual search paradigms both characterize processing at lower levels where local image features are first identified.

In Experiment 3, we examined whether shape-from-shading information influences the integration of dipoles. Specifically, we examined whether the shading congruence of additional noise dipoles (relative to signal dipoles) affected the perception of Glass structure. The results indicated that additional noise dipoles interfered with Glass pattern detection regardless of their shading direction. Accordingly, we concluded that the visual system combines additional noise dipoles to the detriment of global form detection, irrespective of their perceived 3D shape. Visual search models that can account for form processing at the local stage of analysis cannot, therefore, account for the results of Experiment 3.

Our results raise an important question: why are different 3D shapes combined in the detection of global structure? It is possible that, to achieve efficient processing, the visual system must be sensitive to regular properties of natural scenes. Natural scenes are highly textured (or matte) containing local elements and features at different depths and with different 3D shapes. A good example of this would be a forest floor, which consists of leaves and fallen branches that are randomly distributed in space. Illumination from above would produce local pockets of dark–light or light–dark shaded regions that may be interpreted by the visual system as being “convex” or “concave” in shape. While shape-from-shading information
will reveal the individual 3D structure of local features, in order to detect the spatial layout of the scene, the visual system must combine these features irrespective of their 3D shape. This suggestion would be consistent with the work of Mamassian et al. (2003), who argue that the computation of 3D shape from shading is predominantly “bottom-up” occurring at local areas of visual analysis in V1 and V2, before a global 3D representation of objects is interpreted in higher cortical areas. Additionally, Yamane, Carlson, Bowman, Wang, and Connor (2008) have found evidence for a neural network through which complex 3D shapes can be coded in the higher cortical areas along the ventral route of processing. This latter study found that a network of neurons codes 3D shape in terms of different local surface curvature properties as implied by shading information.

The results of this study also accord with those of Khuu and Hayes (2005) who investigated the role of stereopsis in the perception of Glass structure. They reported that Glass pattern structure becomes undetectable when partner dots forming dipoles are sufficiently separated in depth, signaled by a difference in binocular disparity. If one considers that different depth structures were conveyed by the shading gradient of dots in the present study, these findings accord well. Moreover, Khuu and Hayes demonstrated that additional noise dipoles interfered with the ability to detect Glass structure even if they occupied different positions in depth, though additional noise dots do not affect detection at disparities beyond ±1 degree of visual angle. These results also agree with the findings and conclusions of our present Experiment 3—that the visual system combines different 3D forms to detect global structure. Together, the findings of the present study and those of Or et al. (2007) clarify the importance of depth information to the perception of global form; though different cues can be used to signal depth, they have comparable effects.

The findings of Experiment 3 superficially appear at odds with Kleffner and Ramachandran (1992), who demonstrated that the appearance of a contour shape consisting of shaded tokens is not affected when embedded in a field of distracters with the opposite shading gradient. In that study, Kleffner and Ramachandran required observers to determine whether the spatial configuration of a group of shaded tokens was an “O” or a “C” shape. Observers could do this task with a high degree of accuracy when target and distracters had the opposite shading gradients and, therefore, the opposite 3D shape. They concluded that shape from shading reveals the shape of the stimulus, which segregates from the background. One might suggest that this indicates that oppositely shaded elements are not combined to detect global form. However, it is important to note that the experimental paradigm employed by Kleffner and Ramachandran is essentially a shape discrimination task between stimuli that are highly coherent. In the present study, we did not examine the perception of shape but rather sought to determine whether the visual system is able to integrate different 3D shapes to detect or see global form. Note that for our stimulus, Glass structure can still be perceived at signal levels above threshold. Under these circumstances, the global form of the stimulus is perceptible and allows observers to easily detect the shape of the stimulus from oppositely shaded background dipoles. This would, therefore, agree with the observations of Kleffner and Ramachandran (1992).

As mentioned in the Introduction section, the detection of form in Glass patterns is likely to involve a hierarchical network of cortical areas in the ventral route of processing. Candidate areas are V1 and V2, which are involved in the local analysis of form, and V4, which contains cells sensitive to complex structure and global shapes. It has been well established that these cortical areas are also important in the computation of shape from shading (see Hanazawa & Komatsu, 2001; Lee et al., 2002; Mamassian et al., 2003). We, therefore, need to ask whether there is agreement between the findings of the present study and physiological investigations examining cortical areas in which shape from shading is derived. Lee et al. (2002) reported that V1 and V2 cells are highly selective to different shading gradients and, therefore, code different 3D shapes. These cells are also involved in the detection of local orientation in Glass patterns (Smith et al., 2002, 2007), but given their selectivity for 3D form, they may be unable to recover local orientation signaled by pairings between different 3D shapes. However, such cells would most definitely be able to extract the local orientation of bipartite dipoles, as these dipoles do not have a 3D profile. This observation agrees with the findings of Experiments 1 and 2 of the present study. Additionally, previous studies have also shown that V4 cells respond to shape-from-shading information. For example, Hanazawa and Komatsu (2001) reported that, while a small proportion of V4 cells are tuned to the direction of shading, the majority are either tuned to opposite directions or are not tuned to shading at all. Given that V4 cells are also responsible for the coding of global form, the findings of Hanazawa and Komatsu would mean that they are likely to be sensitive to, and capable of integrating, different 3D shapes. This observation is entirely in agreement with the results of Experiment 3. As our study was behavioral in nature, connections between these two studies are necessarily speculative. However, the concordance between previous studies revealing the cortical areas responsible for the computation of shape from shading and the findings of the present study deserve to be highlighted.

Conclusions

The significance of the present study is that shape-from-shading information is important to the perception of global form. At the local stage of form analysis,
shape-from-shading information is used to define the 3D shape of individual elements: elements that are inferred as 3D based on their shading patterns are not combined with incongruent elements as local orientation is recovered. However, a different strategy is employed by the visual system at the global stage of processing: local orientations are combined irrespective of their 3D shape from shading. This suggests a global computational algorithm in which different local 3D forms are integrated to detect the structural layout of complex scenes or images.

Acknowledgments

We thank the observers who participated in the study and Joanna Kidd for her helpful comments. This research was supported by an Australian Research Council (ARC) Discovery Project Grant (DP110104713), an Early Career Researcher Grant, and a Science Faculty Research Grant from the University of New South Wales to S. Khuu.

Commercial relationships: none.
Corresponding author: Sieu K. Khuu.
Email: s.khuu@unsw.edu.au.
Address: School of Optometry and Vision Science, The University of New South Wales, Sydney, New South Wales 2052, Australia.

References


