Detecting the structural form of cast shadows patterns

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

Shazaan Khambiye
The School of Optometry and Vision Science,
The University of New South Wales, Sydney,
New South Wales, Australia

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

Cast shadows are ubiquitous in the visual scene and they inform greatly about the scene’s three-dimensional spatial layout. In the present study we investigated the ability of the visual system to detect the structural form of cast shadows using Glass patterns consisting of local dot-pairs or dipoles oriented appropriately to convey global structure. “Cast shadow” Glass patterns were constructed by superimposing two opposite polarity (light-increment – the object – and light-decrement – the shadow) concentric Glass patterns and then spatially displacing the decrement pattern along a particular orientation conforming to the pattern’s lighting direction. We determined the shadow detection threshold, which specifies the amount of structural difference between the object (increment Glass-pattern) and its respective shadow (decrement Glass-pattern), required for observers to detect the cast shadow. This was achieved by varying the ratio between local opposite-polarity dipole pairs that were appropriately aligned along the lighting direction (congruent dipole pairs), and dipole pairs that were randomly oriented (incongruent dipole pairs), and were therefore inconsistent with the pattern’s lighting direction. We reported that thresholds were comparatively lower (i.e., the visual system is able to tolerate greater local pattern inconsistencies) for light-from-above patterns than for light-from-below patterns (Experiment 1), and detection is optimal for highly coherent patterns (Experiment 2) small spatial separations between opposite polarity (Experiment 3). Our findings demonstrated that the visual system is more sensitive to light-from-above configurations when detecting the form of cast shadows, and this detection process largely ignores local inconsistencies between the object and its respective shadow.

Keywords: cast shadows, stereopsis, Glass patterns, figure ground configuration


Introduction

Ascertaining the depth structure of the visual scene—the depth order of textures and objects in the environment relative to the observer—is arguably one of the more important tasks that the visual system performs. This operation is foremost in formulating the three-dimensional (3D) spatial layout that provides the basis for object perception and visually guided behavior. However, the retinal image is coded in two-dimensions (2D); depth structure is not directly mapped in this image. Depth structure must instead be indirectly inferred from a host of primary (e.g., binocular disparity) and secondary (e.g., size, perspective and motion parallax) cues that correlate directly with the depth configuration of features in the retinal image. While experimentation has shown that the visual system is most reliant on binocular disparity as a metric cue to depth, it is also reliant on a number of monocular cues that provide an ordinal representation of the depth position of objects (see Howard & Rogers, 1995; Kleffner & Ramachandran, 1992). Shadows (see Figure 1A)—the dark shading patterns and texture boundaries that reflect the structured interplay between the properties of objects and properties of the source of illumination—are ubiquitous in illuminated scenes, and are a prime example of a monocular cue to ordinal depth (e.g., Mamassian, 2004; Ramachandran, 1988a, b; Yonas, 1979).

Experimentation has consistently revealed that shadows provide a strong cue to the 3D shape of objects as well as the spatial layout of visual scenes (e.g., Khuu & Khambiye, 2012; Khuu, Moreland, & Phu, 2011; Ramachandran, 1988a, b; Khuu, Moreland, & Phu, 2011). This is likely because important
scene statistics such as the shape of objects, the position of the light source, the depth order of objects, and the position/slant of surfaces can be inferred from their statistical properties. Yonas (1979) differentiated between two types of object shadows: the primary or attached shadow, which occurs when the shadow is visible on the same object; and the derived or cast shadow, which occurs when the shadow of the object is detached and is projected on a different object or background (e.g., leaves casting shadows on a corrugated wall as in Figure 1A). This distinction between shadow types relates also to the sort of depth information that can be inferred from them (Mamassian, Knill, & Kersten, 1998). Attached shadows are particularly informative about the 3D shape of objects. For example, Ramachandran (1988a, b) showed that 2D circular tokens appear 3D, and either convex or concave, depending on the surface shading-gradient. Assuming a ‘light-from-above’ default (e.g., Ramachandran, 1988a, b), convexity is interpreted from circular tokens with ‘light to dark’ shading which ‘pop-out’ from the background, while tokens with the opposite ‘dark to light’ shading appear concave. By contrast, cast shadows provide very little information about object shape, but inform a great deal about the surface onto which they are cast. Because of this, they are powerfully informative about the spatial layout of the scene in both a qualitative (i.e., the depth order of objects relative to the background) and quantitative
manner (i.e., how far away the object is from its background).

Critical to understanding the perception of shadows is explaining how the visual system effectively associates the shadow with a potential casting object. Before useful depth information can be extracted, the visual system must first identify the shadow and then match it with its casting object—the so-called ‘shadow-correspondence problem’ (e.g., Mamassian, 2004). Where attached shadows are concerned, the solution to the correspondence problem is usually straightforward: the shading pattern is intrinsically ‘attached’ to the object, and a simple proximity matching/labeling operation provides a successful solution. However, this operation is not simple for cast shadows. The visual scene is most commonly a spatial array of light (an object) and dark (potentially a shadow) shapes that may be similar or dissimilar in form and may be spatially separated (e.g., the shadows and leaves in Figure 1 have similar shape and size), and it is not immediately clear how the visual system is able to solve this shadow matching problem. Indeed, an unconstrained shape matching operation will frequently produce spurious object-shadow pairings implying erroneous depth relationships.

Despite the complexity of resolving the shadow correspondence-problem for the perception of cast shadows, the visual system performs with relative efficiency (Mamassian & Goutcher, 2001). Rather than being ignored, cast shadows have been shown to contribute substantially to determining the 3D spatial layout of the visual scene (see Dee & Santos, 2011). For example, Hubona, Wheeler, Shirah, and Brandt (1998) demonstrated that cast shadows greatly influence the perceived depth of stereoscopically defined objects. Kersten, Knill, Mamassian, and Bülthoff (1996) demonstrated that the relative position of the cast shadow can govern the perceived trajectory of a moving object and that a moving cast shadow can elicit illusory motion in a stationary object. Moreover, cast shadows can be used as an effective cue to aid visual search and the segregation of local visual information (e.g., Cunningham, Beck, & Mingolla, 1996; Lovell, Gilchrist, Tolhurst, & Troschiano, 2009; Rensink & Cavanagh, 2004), and aids in the disambiguation of object and surface shape (see Cavanagh & Leclerc, 1989; Madison, Thompson, Kersten, Shirley & Smits, 2001), which has been shown to contribute to the recognition of objects (Braje, Legge, & Kersten, 2000; Tarr, Kersten, & Bülthoff, 1998). Recently, Mamassian (2004) has suggested that the visual system solves the shadow correspondence problem by implementing a coarse scale analysis. That is, the visual system is largely insensitive to local differences in the structural congruence (e.g., conforming to a particular lighting direction) between the shadow and the casting object, but instead image characteristics such as their “center of mass” is used as a basis for matching. Mamassian (2004) noted that cast shadow percepts are evident even when local shadow-object matches represent different lighting directions. These ‘impossible shadows’ are commonly observed in art and indicate that the visual system emphasizes global rather than local image properties in the shadow matching process (see Casati, 2008; Cavanagh, 2005). This coarse analysis is perhaps optimal as it provides quick means of discerning the 3D structure of the visual scene and ignoring fine detail that may hinder the detection process. However, this process is not well understood and remains the focus of much research.

While it has been suggested that coarse scale analysis is applied to solve the shadow correspondence problem, it is not clear by how much structural congruence may fluctuate before a cast shadow percept is no longer identifiable. Indeed, when the shape of the shadow and object are completely incongruent, the percept does not resemble a cast shadow, interfering with object recognition and detection (see Castiello, 2001; Dee & Santos, 2011; Lovell et al., 2009). What has yet to be well established is how sensitive the visual system is in detecting and identifying changes in the local structural form of cast shadow patterns. It is the goal of the present study to examine this issue. Quantifying sensitivity to the structural form of cast-shadow patterns was investigated in the present study using Glass patterns (see Figure 1A). Glass patterns (after Glass, 1969) are random dot stimuli consisting of dot-pairs of the same polarity (dipoles) whose orientation conforms to some common geometric rule. For example, a concentric Glass pattern (see Figure 1) 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 particularly useful because their analysis by the visual system is well understood (e.g., Gallant, Conner, Rakshit, Lewis, & Van Essen, 1996; Smith, Bair, & Movshon, 2002), and reflects both local and global levels of computation: the orientation of local dipoles is initially extracted, and then combined at a later stage by cells with large receptive fields (in areas along the ventral route of processing, e.g., V4 and Infero-temporal [IT] cortex) to determine the global form (e.g., Gallant et al., 1996; Khuu & Hayes, 2005; Wilson & Wilkinson, 1998). Additionally, such pathways are believed to also subserve the processing of cast shadows (see Castiello, Lusher, Burton, & Disler, 2003), which might suggest a common neural locus of global form and cast shadows might be located in areas along the ventral pathway. Given this, the use of cast shadow Glass patterns in the present study is appropriate as such stimuli are likely to activate cortical areas important for the processing of global form and shadow information.
Glass patterns can be constructed to convey compelling depictions of cast shadows (see Figure 1B). ‘Cast shadow Glass patterns’ are constructed by superimposing two identically formed, but opposite polarity, Glass patterns (one light-increment and another light-decrement of equal absolute contrast) and then offsetting the decrement pattern in a particular direction over a small spatial distance. As depicted in Figure 1B, the percept strongly resembles a single increment Glass pattern with an accompanying shadow appearing in 3D, rather than two distinct opposite polarity Glass patterns. The form relationship in these patterns can be independently manipulated to gauge how it affects the ability of the visual system to form shadow relationships. This type of pattern is ideal for the present purposes as it allows systematic manipulation of the level of congruence between the two patterns (perceptually, the object and shadow) required for the perception of a cast shadow.

The present study reports three experiments. In Experiment 1, we used cast-shadow Glass patterns to measure the amount of structural similarity/congruence between opposite polarity dipoles for correct identification of a cast shadow for different lighting directions. Previous studies have established that the visual system follows a light-from-above prior (e.g., Adams, 2007; Ramachandran, 1988a, b). One might predict, then, that the visual system may prefer light-from-above patterns over those in which light is implied to come from below. In Experiment 2, we examine whether the detection of cast shadows reflects global operations by comparing pattern detection performance between patterns with and without global form coherence. Finally in Experiment 3, we examined the impact of shadow separation—the spatial distance between the object and its shadow—on the ability to detect cast shadows, to clarify the rules applied by the visual system to formulate a match between object and shadow with regards to proximity. That is, over what separation range is the congruence optimal, and how far can the object and shadow be separated before the visual system is prevented from making a cast shadow judgment, denoting increment and decrement patterns as structurally distinct.

**Experiment 1: Quantifying human sensitivity to cast-shadow patterns**

Glass patterns are typically used to measure human sensitivity to global form 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 thresholds are 15%–25% signal, e.g., Badcock, Clifford, & Khuu, 2005). In Experiment 1, we used a variant of this method to measure human detection of cast shadows patterns. Rather than changing the form coherence of Glass patterns (as in the aforementioned studies), the structural congruence between the shadow and casting object was systematically altered. ‘Structural congruence’ refers to the extent to which local increment and decrement dipoles (belonging to the two separate patterns) in the pattern are aligned along a particular orientation conforming to a particular lighting direction. For example in Figure 1B, increment and decrement dipoles are separated over a small distance and aligned vertically to produce a congruent cast shadow pattern inferring a single light source overhead. However, such patterns are not perceived as cast shadows if increment and decrement dipoles are not locally aligned, but randomly positioned, signifying many different lighting directions. For example, in Figure 1C decrement dipoles are randomly placed relative to increment dipoles. For such incongruent patterns a cast shadow is not perceived but rather two opposite polarity concentric Glass patterns that appear distinct in form. Note that regardless of the shadow congruence level, the global form of both increment and decrement Glass patterns remains the same (appearing concentric) because the local orientation of dipoles does not change, only their relative positions. Changing the shadow congruence therefore interferes with the ability to match local information to detect cast shadows, not the ability to discriminate a difference in the global form/shape of the two Glass patterns.

Changing the level of object-shadow congruence in the pattern enables a measure of the tolerance of the visual system to local (but not global) structural changes in the shadow. This ‘shadow detection threshold’ denotes the minimum number of congruently aligned increment and decrement dipole pairings required for correct identification of a cast shadow above chance level. Employing Glass patterns in this manner affords precision in measuring the sensitivity of mechanisms responsible for the processing of cast shadow information and provides a direct indication of the tolerance of such mechanisms to local inconsistencies when detecting cast shadow patterns. This approach largely differs from previous investigations that have used reaction time (e.g., Rensink & Cavanagh, 2004), qualitative stimuli (Mamassian et al., 1998; Yonas, 1979), or stimuli in which the shadow bares no resemblance to the object or allows for systematic change in shape congruence (Castiello et al., 2003) to describe the perception of cast shadows.

As mentioned, previous studies have established that the detection of shadows is governed by assumptions regarding the direction of lighting. It is clear that the visual system is selective for shadows that imply a light from above to a greater extent than those that imply...
light from below (Adams, 2007; Lovell et al., 2009; Rensink & Cavanagh, 2004; Sun & Perona, 1998). Previous investigations have also shown selectivity to light-from-above patterns when searching for cast shadow targets (e.g., Lovell et al., 2009; Rensink & Cavanagh, 2004). In Experiment 1, we examined whether lighting direction affects the detection of the global form of cast-shadow patterns.

**Methods**

**Observers**

Data were obtained from five participants aged 22-35 years. All were naïve to the purpose of the experiment but were experienced in psychophysical experimentation. All had normal or corrected-to-normal visual acuity.

**Stimuli**

The stimulus consisted of two opposite polarity, one light increment and the other light decrement, Glass patterns placed within a circular stimulus area (radius: 12.4°) that was set to a background luminance of 67 cd/m². Dipoles from both patterns consisted of two similar sized circular dots (radius: 0.125°) that were separated by a fixed distance of 0.25°. Each Glass pattern consisted of 800 dipoles and was either light-increment (120 cd/m², Weber Contrast: 0.8) or light-decrement (14 cd/m², Weber Contrast: –0.8) in appearance. The overall dipole density of the stimulus was 3.26 dipoles/deg². At the local level, increment and decrement (I-D) dipoles were separated over a small spatial distance, and in Experiment 1, this “shadow separation” was set to 0.125°. Decrement dipoles were either placed in alignment consistent with a particular light direction (“congruent dipole pairs”), or were randomly aligned relative to its increment dipole partner (“incongruent dipole pairs”). As depicted in Figure 1A, these construction methods produced an orthographically presented cast shadow which resembled a vertically aligned object projecting its separated shadow on a vertical background surface. This study used only Glass patterns conveying concentric structure; a pilot study revealed no difference in sensitivity between different types of global form when detecting cast shadows. Concentric structures are ubiquitous in natural scenes, and neural mechanisms that are selective for concentric form have been shown to exist (e.g., Chen & Foley, 2004; Gallant et al., 1996).

Observers viewed the stimulus binocularly in a dark room at a viewing distance of 70 cm. Stimuli were generated using MATLAB version 7 and displayed on a linearized 24-inch Mitsubishi Diamond Pro monitor (Mitsubishi, Dong Guan, China) driven at a frame rate of 100 Hz.

**Procedure**

To measure human sensitivity to the structural form of cast shadows, we systematically altered the number of congruent dipole pairs in the stimulus required for correct identification of the cast shadow pattern above chance level. To measure this cast shadow detection threshold, we used a two-interval forced-choice paradigm (2IFC) in conjunction with a staircase procedure. In one interval, a congruent cast shadow pattern was presented (briefly for 150 ms to minimize eye movements), while in the other interval a “non-shadow” pattern was presented in which all increment and decrement dipole pairs were incongruent in position (e.g., Figure 1C). Observers were required to fixate on a black cross at the center of the stimulus and to judge the interval containing the cast shadow pattern. The presentation order was randomized between trials. A staircase procedure that converged on the 79% correct performance level was used to modify the shadow congruence level until observers could reliably discriminate the interval containing the cast shadow pattern. The staircase began at an initial shadow congruence level of 80% (i.e., 640 congruent dipole pairs) and the initial step size of the staircase was 8%. On the first and subsequent reversal the step size was halved. After the third reversal the step size was 1% and remained at this value until the end of the trial. The staircase lasted for eight reversals and the average of the last four reversals provided an indication of the shadow detection threshold. No feedback was given to indicate the correctness of response.

The above procedures were repeated for cast shadow patterns in which the orientation of congruent dipole pairs was systematically varied over 360° of orientation at regular steps. Beginning at 0°, which denotes light directly from the left, and increasing by 45° in a clockwise direction, these orientations were: 45° (above and to the left), 90° (directly above), 135° (above and to the right), 180° (directly from the right), 225° (below and to the right), 270° (directly below), and 315° (below and to the left). A block comprised eight staircase trials: the eight different lighting directions repeated once. In total there were 10 blocks with the order in which the different lighting-directions were presented to the observer randomized within and between blocks. The data were averaged across the blocks to provide an estimate of the detection threshold for each of the eight lighting directions.

**Results and discussion**

The results of Experiment 1 are shown in Figure 2, which plots the minimum shadow detection threshold (the average proportion congruent dipole-pairs) required to detect cast shadow Glass patterns for
different lighting directions (circles). Error bars signify 95% confidence intervals. There were a number of findings. First, detection thresholds for cast shadow Glass patterns (gray circles) were affected by lighting direction. A repeated-measures one-way ANOVA revealed a significant main effect of lighting direction on detection thresholds, $F(7, 4) = 33.64, p < 0.0001$. As the lighting direction is systematically changed (from 0°), there is an initial decrease in threshold with the lowest value (~0.25) corresponding to a lighting direction of 90° (i.e., light directly above). Thresholds increase thereafter with the highest detection threshold observed for patterns in which light was from below (i.e., ~0.5 for 270°). These results demonstrate the predicted asymmetry between detecting cast shadow patterns in which the lighting direction is above and below. Post-hoc Bonferroni comparisons (Neter, Wasserman, & Kutner, 1990) revealed that thresholds for light-from-above patterns were lower and significantly different from patterns in which light was from below when compared along the same orientation axis: 90° and 270° (mean difference: 0.37, $p < 0.01$); 45° and 225° (mean difference: 0.24, $p < 0.01$); 135° and 315° (mean difference: 0.23, $p < 0.01$). This finding confirms that when detecting the structural form of cast shadow patterns, the visual system is selective for ecologically valid cast shadows in which light is implied to be overhead (see General discussion). This observation is in agreement with previous studies examining both cast and attached shadow perception that demonstrate a preference to patterns in which light is from above compared to other lighting directions (e.g., see Lovell et al., 2009; Ramachandran, 1988a, b; Sun & Perona, 1998). In addition to these, the present study demonstrates that lighting direction is also an important factor when detecting the structural form of cast shadows. Second, congruence for light-from-above (90°) was significantly different and lower than for patterns in which light is implied to come in from the sides (light from the right: 180°, mean difference: 0.24, $p < 0.01$) and 0° (light from the left, mean difference: 0.29, $p < 0.01$), however, both patterns are not significantly different from patterns in which light is from below (i.e., 270°). These findings additionally demonstrate that the visual system is selective for light-from-above patterns over those in which the light source is implied to come from below and to the sides (see below for discussion).

The results of Experiment 1 demonstrate a clear asymmetry in the ability to detect cast shadow patterns depending on direction of lighting. What might account for this? It is likely that this difference reflects the ecological validity of these two pattern types. Given that light comes to natural scenes from a single overhead light source, the visual system is likely to consider light-from-above patterns as cast shadows, but light-from-below patterns are unnatural, and therefore not perceived as shadows (see Lovell et al., 2009; Rensink & Cavanagh, 2004, for further discussion). Given this distinction, different mechanisms might be expected to underlie their detection. We suggest that our data might be consistent with the functioning of two separate mechanisms, one that implements a coarse scale analysis and detects cast shadows conveyed by light-from-above patterns, and another “non-shadow” mechanism capable of detecting finer pattern detail that is exclusively responsible for detecting light-from-below patterns. As shown in Figure 2, reflecting the operations of a coarse scale detector, detection thresholds (proportion congruent dipole pairs) for light-from-above patterns were approximately 0.25–0.35, which indicates that a cast shadow is still perceptible even when well over half of the pattern consists of incongruently aligned dipole-pairs. However, for light-from-below patterns, detection thresholds were much higher at approximately 0.45–0.5. Higher detection thresholds arise because the mechanism underlying the detection of such patterns is sensitive to the local features of the stimulus. Thus, when the structural congruence of the pattern is decreased (by increasing the proportion of incongruently aligned dipole pairs), it is unable to ignore the local inconsistencies between the shadow and object, which interferes with the ability of this mechanism to detect the form (see below for a more detailed description) of the stimulus.
If light-from-below patterns are not considered as shadows, what might be mediating their perception? It is important to note that, to convey a shadow percept, increment and decrement dipole pairs in our stimulus are placed in physical alignment along a particular lighting direction. It might be suggested that for light-from-below patterns, observers are not detecting a 3D cast shadow, but rather by judging whether a majority of increment and decrement dipole pairs are aligned along a particular direction. Note that this is an alignment judgment because the visual system cannot effectively pair increment and decrement dots to extract a local orientation signal (see Badcock et al., 2005; Glass & Switkes, 1976; Smith et al., 2002). Such a strategy must reflect a fine scale matching operation, in which the visual system makes multiple comparisons of the alignment between increment and decrement dipole pairs over a large spatial area before integration and form detection. However, for the dense random dot stimulus used in the present study, decreasing the structural congruence of the pattern will mean more incongruently aligned dipole pairs in the pattern, and this will disrupt the ability of the visual system to perceive the pattern’s overall alignment. Thus, judging spatial alignment is optimal only at high congruence levels. This would explain why higher detection thresholds are observed light-from-below patterns as shown in Figure 2, but not for light-from-above patterns whose perception is mediated by a coarse scale detector (selective for shadow configurations implying light from above) capable of tolerating local inconsistencies in the pattern.

To provide a direct measure (in the absence of shadow percept) of the ability of the visual system to make an alignment judgment, a supplementary experiment was conducted in which we repeated Experiment 1 with Glass pattern stimuli consisting of red (CIE \( x = 0.62, y = 0.33 \)) and green (CIE \( x = 0.28, y = 0.59 \)) dipoles of approximately equal perceptual luminance. Such patterns (see Figure 1D) do not appear 3D because there is no/little systematic luminance difference between colored dipoles to induce a cast shadow effect. Note that it was not possible to ensure complete isoluminance given the size and density of our stimulus. However, it has been shown that isoluminant point does not appreciatively change with stimulus eccentricities up to approximately 20° of visual angle (see Bilodeau & Faubert, 1997). This critical area is slightly smaller than our stimulus size (25°), and any potential luminance difference between colored dots is likely to be small, and unlikely to lead to a cast shadow percept (see Figure 1D). Importantly, Red and Green (R-G) Glass patterns were structurally identical to the I-D patterns employed previously, but appeared as two color distinct patterns aligned along a particular direction. Thus, the detection of the structural congruence of these patterns must rely on judging the alignment of different colored dipoles because there is no cast shadow cue. If it were the case that light-from-below patterns are not considered shadows, but instead judged on the physical alignment of dipole pairs, one would predict that detection thresholds for I-D patterns would be similar to R-G Glass patterns.

While the R-G and I-D patterns used in this experiment are perceptually different, previous studies have well established that R-G Glass patterns are equally detectable as I-D patterns at high contrasts (with thresholds approximately 20% signal dipoles; compare Cardinal & Kiper, 2003 and Wilson & Wilkinson, 1998). More importantly, the mechanisms responsible for the detection of Glass pattern form are not selective for differences in color and luminance contrast (see Rentzeperis & Kiper, 2010), which suggests that R-G and I-D patterns are equally perceptible and likely to be processed by a common form mechanism. This allows for their comparison.

The supplementary experiment used a two-interval forced-choice procedure. Observers (the same as in the main experiment) were asked to judge the interval containing a pattern in which R-G dipole pairs were congruently aligned along a particular direction. The other interval contained a similar pattern, but one in which the R-G Glass pattern was incongruent and consisted of local dipoles pairs that were randomly aligned. This procedure was repeated for alignment directions of 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°, which, in Experiment 1, corresponded to the implied direction of lighting with cast shadow patterns.

The results of this supplementary experiment are shown in Figure 2 as detection thresholds (proportion congruent dipole-pairs, black squares, error bars signify 95% confidence intervals). In Figure 2, detection thresholds for R-G Glass patterns were approximately 0.5–0.6, and importantly, changing lighting direction did not affect detection thresholds (one-way repeated-measures ANOVA: \( F(7, 4) = 0.84, p = 0.5634 \)). These results suggest that the visual system is particularly poor at detecting the alignment of dense random dot stimuli, and requiring congruence levels to be over 0.5 for correct identification of the aligned pattern.

When detection thresholds for different lighting directions were compared (two-way repeated-measures ANOVA) for R-G and I-D Glass patterns, significant main effects were observed for the type of pattern, \( F(1, 32) = 283, p < 0.0001 \), and lighting direction, \( F(7, 32) = 15.87, p < 0.0001 \). Additionally, their interaction was also significant, \( F(7, 32) = 13.05, p < 0.001 \), indicating that detection thresholds were dependent on both the lighting direction and the type of pattern. When
thresholds were compared (using Bonferroni multiple comparisons test) between R-G and I-D Glass patterns for the same lighting direction/dipole pair-orientation, significant differences were observed for light-from-above patterns (45°: mean difference: 0.32, $p < 0.0001$; 90°: mean difference: 0.37, $p < 0.0001$; 135°: mean difference: 0.26, $p < 0.0001$). These findings indicate that when detecting patterns in which light is from above (and also to the sides), the visual system is not simply judging the alignment of dipoles, but rather, is mostly to be sensitive to, and reliant on, the 3D percept conveyed by such patterns. Consistent with the observations made by Mamassian (2004), this process is tolerant to local inconsistencies in the structural congruence of the pattern. We also note that judging the detection thresholds for patterns in which light was horizontal and from the sides were lower than for R-G patterns (0°: mean difference: 0.14, $p < 0.0001$ and 180°: mean difference: 0.14, $p < 0.0001$). This suggests that for these pattern types the visual system might interpret them as cast shadows, but one that is less compelling than for those in which light is from above (see Figure 2). This observation is very much consistent with the findings of a number of previous studies showing the 3D shape of horizontally shaded objects are more ambiguous than for vertically shaded elements (Khuu & Khambiye, 2012; Kleffner & Ramachandran, 1992; Sun & Perona, 1998).

When light-from-below patterns were compared to the results obtained with R-G Glass patterns, they were not significantly different at the $p = 0.05$ level (225°: mean difference: 0.086; 270°: mean difference: 0.052, 315°: mean difference: 0.073). Note, though, that those thresholds were consistently lower than for R-G patterns at each orientation (see Figure 2). These findings suggest that light-from-below patterns might not be interpreted as shadows because their detection is comparable to judging the physical alignment of elements, reflecting an operation that is not tolerant to local inconsistencies in the structural congruence of the pattern.

In summary, the results of Experiment 1 demonstrate that, in the detection of light-from-above patterns, this operation is very tolerant to local pattern inconsistencies (between the object and its shadow), reflecting a coarse scale operation when detecting cast shadows. We have quantified this process in terms of degree of structural congruence required for the correct identification of the cast shadow. However, the detection of light-from-below patterns is likely to reflect a judgment of the pattern’s 2D alignment, rather than them being perceived as cast-shadows. Indeed for these patterns detection thresholds were not significantly different to those when judging the 2D alignment of R-G Glass patterns.

### Experiment 2: Is global form necessary for the detection of cast shadows?

In Experiment 1 we concluded that light-from-below patterns are selectively analyzed by a shadow detector that is largely insensitive to local lighting-direction consistencies in the stimulus (see Casati, 2008; Mamassian, 2004). Implicit in this explanation is that such detectors must operate by first extracting the local dipole pairings, and then integrating local features to extract global information about the object (i.e., the Glass pattern) and its cast shadow. However, it is important to note that in Experiment 1, cast-shadow detection might occur without the need to integrate local features to detect the global form of the pattern. In particular, on a given trial the cast shadow might be determined by examining the congruence of a small subset of dipole pairs, without the need detect the global form of the pattern. This strategy would imply that the Glass patterns used in the present study are judged based on their local rather than their global properties.

In Experiment 2, we sought to resolve the above-mentioned issue by measuring and comparing detection thresholds between patterns with and without global form. As in Experiment 1, a “coherent” cast-shadow concentric pattern is constructed by orienting all dipoles tangent to radii from the centre of the stimulus, however, setting the form coherence to 0%, such that all dipoles were randomly oriented, produces a “random” shadow pattern (see Figure 3A, B). If the perception of global form was unnecessary for the detection of cast shadows, it is expected that thresholds for random shadow patterns will be the same as for coherent patterns. However, if perceiving cast shadows was dependent on global form, detection thresholds will be dependent on both the lighting direction as well the level of form coherence. In particular, an expected outcome is that the thresholds for coherent patterns will change with lighting direction (showing a sensitivity bias for light-from-above directions), but not to the same extent for random patterns.

In Experiment 2, shadow detection thresholds were measured (using the methods and procedures of Experiment 1) for random and coherent shadow patterns, and this was repeated for lighting directions of 0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°. Five observers who had normal or corrected-to-normal acuity acted as observers. One was an author of the study, while the others were new observers who were experienced in psychophysical experimentation, but were naïve to the goals of the study.

The results of Experiment 2 are shown in Figure 4, which plots the average cast-shadow detection thresh-
old as a function of the lighting direction angle. Results for coherent shadow patterns are indicated by the black squares (solid line), while data for random shadow patterns are indicated by the gray circles (dashed lines). A number of findings present themselves. First, thresholds required to detect random shadow patterns were approximately 0.4–0.5, and did not systematically change with the lighting direction, $F(6, 28) = 2.039, p = 0.0935$. Note that for this pattern, detection thresholds at 90° (implying light directly from above) appear to be lower than for other directions, however, post-hoc comparisons (Bonferroni’s multile comparison test) confirmed that this difference was not significant (at the $p = 0.05$ level). Second, in contrast, the detection of coherent shadow patterns was significantly affected by the lighting direction, $F(6, 28) = 20.54, p < 0.0001$, with significantly lower thresholds for light-from-above patterns than light-from-below when compared along the same direction axis (45° and 225°: mean difference = 0.19, $p < 0.001$; 90° and 270°: mean difference = 0.28, $p < 0.001$; 135° and 315°: mean difference = 0.20, $p < 0.001$) This outcome replicates the findings of Experiment 1, and is consistent with the assumption that light-from-above patterns are selectively processed by cast shadow detectors that are sensitive to coarse pattern detail.

A repeated-measures two-way ANOVA was conducted to compare the results between random and coherent shadow-patterns for the different lighting directions. This analysis revealed main effects for both coherence level (0% vs. 100%, $F(1, 32) = 16.14, p = 0.0003$) and lighting direction, $F(7, 32) = 13.06, p < 0.0001$, and a significant interaction effect, $F(7, 32) = 16.14, p = 0.0008$, between the two factors. The significant interaction effect confirmed the aforementioned observation that changing lighting direction produced different effects on the detection of random and coherent shadow patterns. Particularly, differences in detection thresholds were largest for light-from-above directions, however, for light-from-below directions the detection of both pattern types was the same. Indeed, Bonferroni multiple comparisons post-hoc tests confirmed that random and coherent shadow-patterns were significantly different at directions of 45° (mean difference = 0.194, $p < 0.001$), 90° (mean difference = 0.148, $p < 0.01$) and 135° (mean difference = 0.146, $p < 0.01$), but not significantly different ($p > 0.05$) for light-from-below directions of 225°, 270°, and 315°.
The findings of Experiment 2 demonstrate that global form is a requirement for the perception of cast shadows, and supports the notion that their detection is dependent on global properties of the stimulus. This dependency on global form suggests that the detection of cast shadows is largely driven by a global object recognition process that is selective for coarse detail in the stimulus. Without a coherent percept, the visual system is less able to attribute a shadow to the stimulus (see Figure 3). In the absence of global form, the pattern detection is likely to be based on other judgments, such as the alignment of dipole pairs, much like the detection of the R-G patterns and light-from-below I-D patterns used in Experiment 1. Indeed note that in Figure 4, detection thresholds for both pattern types were similar for light-from-below patterns which are considered to be unnatural (see Lovell et al., 2009; Rensink & Cavanagh, 2004) and unlikely to be treated as cast shadows.

**Experiment 3: Sensitivity to cast-shadow Glass patterns: The effect of shadow separation**

The spatial distance between an object and its shadow, or ‘shadow separation,’ is likely to affect the matching operation that enables this monocular cue to depth. As shown in Figure 1B, I-D dipole pairs that are near to each other produce compelling cast shadow percepts. Intuitively, large separations will likely reduce the percept of a cast shadow because I-D dipole pairs will be segregated, making them perceptually distinct and preventing their association. Thus, the visual system might place greater emphasis on close proximity matches when detecting the structural form of cast shadows (see also Ni, Braunstein, & Andersen, 2004). However, the spatial extent over which this association is optimal for our stimulus remains to be established. Experiment 2 examined the degree to which shadow separation affects shadow detection thresholds using the previously outlined procedures. The stimulus and procedures of Experiment 1 was employed, with the separation between I-D dipoles set to 0.25°, 0.5°, 0.75°, and 1.00°. This manipulation was repeated for two lighting directions of 90° (light-from-above) in which a cast shadow percept is evident, and 270° (light-from-below) in which the visual system might not treat as a cast shadow, and therefore, the results for this condition provides a baseline comparison. The same observers as in the previous experiment acted as observers.

The results of Experiment 3 are shown in Figure 5, which plots shadow detection thresholds for patterns in which the lighting direction was 0° (gray circles) and 270° (black squares) against the shadow separation. The average observer data is shown and error bars signify 95% confidence intervals. A repeated-measures two-way ANOVA examining the effect of shadow separation and pattern type on detection thresholds revealed no significant interaction effects between the two factors, but significant main effects for both shadow separation, \( F(3, 32) = 39, p < 0.0001 \), and lighting direction, \( F(3, 32) = 36.03, p < 0.0001 \). Replicating the findings of Experiment 1, lower detection thresholds are observed for light-from-above patterns than for light-from below patterns. Post-hoc Bonferroni multiple comparisons indicated significant differences in detection thresholds between light-from-above and light-from-below patterns for separations of 0.25° (mean difference: 0.18, \( p < 0.05 \)), 0.5° (mean difference: 0.22, \( p < 0.05 \)) and 0.75° (mean difference: 0.13, \( p < 0.05 \)), but not at the largest separation of 1.00° (mean difference: 0.04, \( p > 0.05 \)). Additionally, as shown in Figure 5, detection thresholds for light-from-above patterns (gray circles) were dependent on the shadow separation. Increasing the separation between aligned I-D dipole pairs systematically reduced the perceptibility of the cast shadow stimulus. This is because at large separations opposite polarity dipole pairs are not associated, which negatively affects the detection of the global cast shadow. Spontaneous separation also affected the form of light-from-below patterns (black squares) in a similar manner to light-from-above patterns. As these patterns might not be perceived as cast shadows, higher detection thresholds at large

Figure 5. Average shadow detection thresholds (proportion congruently aligned dipole-pairs) plotted as a function of the shadow separation (in degrees). Results are depicted for cast shadow patterns in which light was implied to be above (gray circles) and below (black squares). Error bars signify 95% confidence intervals.
separations might be due to an inherent increase in correspondence noise, which would affect the perceived pattern alignment. Additionally, for the largest separation (1°) detection thresholds for both pattern types were not significantly different (p = 0.05). This perhaps suggests that at this separation, the visual system may not perceive a cast shadow for light-from-above patterns as their detectability is similar to no different from light-from-below patterns.

In conclusion, our data demonstrate that shadow separation (indicated by the local I-D dipole pair separation), directly affects the perception of cast shadows. Moreover, this finding suggest that while, global mechanisms are responsible for the detection of cast shadows (see Experiment 2), such mechanisms operate by first integrating local matches between I-D features of the stimulus (representing local matches between local features of the object and its shadow), and then integrating local shadow matches to detect the global shadow form. As demonstrated by this experiment, changing the shadow separation prevents the matching of local features (as a result of an increase in correspondence noise at larger separations), which, in turn, impairs the detection of cast shadows.

**General discussion**

The goal of the present study was to quantify human sensitivity to the structural form of cast shadows. We examined the amount of structural congruence between closely paired I-D Glass patterns required for observers to perceive a cast shadow pattern. The results of the present study indicate that the visual system is tolerant to changes in the local structural congruence, but this tolerance is dependent on stimulus properties such as the implied direction of lighting (see Experiment 1), pattern coherence (Experiment 2), and the shadow separation (see Experiment 3). It was found that the visual system is more sensitive to global form patterns in which light is implied as from above; light-from-below is unnatural and not perceived as a cast shadow. These findings confirm two known characteristics of cast shadows. First, the visual system assumes a light-from-above prior when detecting cast shadows (e.g., Jacobson & Werner, 2004; Lovell et al., 2009; Mamassian, 2004), and second, this shadow processing mechanism implements an analysis that is largely tolerant to local structural inconsistencies between the object and its shadow.

Previously, Mamassian (2004) noted that the visual system implements a coarse scale analysis when solving the shadow correspondence problem. The findings of the present study are consistent with this observation. As demonstrated in Figure 2, the detection of cast shadow patterns is optimal when light is implied from above, and for such patterns, incongruence in the structural alignment of dipole pairs can be as much as 75% of the total percent of elements before the pattern can no longer be detected. Significantly though, the present study reveals for the first time how tolerant this operation is in terms of shadow congruence. Additionally, our results agree with Mamassian’s (2004) proposal that the visual system solves the shadow correspondence problem by matching an object with its shadow based on its center of mass. Changing the structural congruence (by randomly aligning dipole-pairs) of the Glass pattern stimulus used in the present study disrupts the center of mass relationship between the object and its shadow, which is derived from the aggregate of local shadow pairings, and therefore interferes with the ability of the visual system to derive a cast shadow percept. Though note that in addition, as demonstrated in Experiment 2, form coherence is an additional requirement for the computation of cast shadows.

The findings of the present study confirm that the light-from-above prior influences the processes by which shadow matching is achieved, and therefore how the visual system solves the shadow correspondence problem. Our findings are consistent with visual search studies that show a difference in reaction time to upright and inverted cast shadow stimuli (e.g., Lovell et al., 2009; Rensink & Cavanagh, 2004). For example, Lovell et al. (2009) reported that search for upright cast-shadow patterns was more efficient than for inverted cast shadow patterns. Moreover, this selectivity to upright cast-shadow patterns is tolerant to large angular deviations between the object and its shadow, in a manner consistent with a coarse scale analysis. These findings contrast with those observed with inverted images in which small angular deviations have an immediate impact on search efficiency. The findings of the present study are entirely consistent with the visual search observations made by Lovell et al. (2009). When detecting light-from-above patterns the visual system is able to tolerate large-scale local change in the structural congruence (reflecting the operations of a coarse-scale pattern detector), but not for light-from-below patterns. We additionally suggest the likelihood that the detection of such patterns is perhaps mediated by an alignment judgment which is highly dependent on the structural congruence in the pattern. Indeed, detection thresholds for light-from-below patterns were no different for detecting 2D R-G patterns, which suggests that light-from-below patterns are not perceived as being 3D.

We report in Experiment 3 that changing spatial separation (shadow separation) between I-D dipoles forming cast shadow Glass patterns disrupted the perceptibility of the cast shadow percept. For the
largest separation of 1.00°, there was no difference between detection thresholds when detecting light-from-above and light-from-below cast shadow patterns. This suggests that the visual system is dependent on shadow separation and that spatial proximity is an important factor in shadow matching, which is likely to be restricted in its spatial extent of analysis. This finding at first hand appears to be at odds with the classic ball-and-shadow illusion (Kersten et al., 1996) in which an object’s cast shadow contributes to its perception of motion direction even over very large separations. However, a number of key differences exist between the stimulus used in the present study and those used by Kersten et al. (1996) that can account for this difference in findings. First, the ball-and-shadow illusion notably has two other cues that help to show an increasing depth effect—a scaled background of lines, and motion. The presence of a grid on the background grounds the percept of the stimulus and causes it to appear moving through 3D space. Second, motion cues provide continuity for the object. Because the original ball was matched with its shadow, the visual system assumes that the same objects moving through space are also separating due to the new distance between them. Our stimuli had no motion cue and were presented on a uniform background. Therefore, there was no perceived effect of increasing depth with increasing separation. Third, the shadow separation is likely to be dependent on the size and scale of the stimulus. In the present study, the dots used to construct Glass patterns were very small, and under natural conditions, are unlikely to cast a visible shadow at large separations. The visual system might be sensitive to this property, and considering also that the shadow size and contrast does not expectedly change with separation, it is unlikely to associate both I-D Glass patterns as a 3D object at large separations. Finally, because we used dense random dot displays (and vary structural congruence to measure pattern sensitivity), in an extension of the shadow correspondence problem, at large separations there is an increased chance that the visual system will make incorrect pairings between the same polarity but non-partner dots. Accordingly, this correspondence noise will interfere with pattern detection limiting pattern detection to small spatial separations.

While the present study and others have shown that cast-shadows can provide a powerful cue to determining the depth and 3D structure of objects, how and where they are processed in the visual system remains to be determined. One hypothesis is that a significant amount of shadow processing in the visual system is done without conscious awareness (see Dee & Santos, 2011). In a study of visual neglect (the defective ability of patients with unilateral damage to the brain to attend to visual space contralateral to the lesion—‘contralesional’), Castiello et al. (2003) describe shadow processing further, arriving at two conclusions: first that shadow processing is implicit, occurring at a location subsequent to the detection of luminance patterns, and second that shadows are thought to be processed in the temporal areas of the brain. In their first experiment, they investigated whether shadows were detected by the visual system when presented to the contralesional field. Their results showed that individuals with visual neglect were similar to controls in their processing of shadows, despite being unable to acknowledge their presence. These data support the hypothesis that shadow processing is implicit, occurring at a cortical stage subsequent to object shape analysis and perception. After depth information is extracted from cast shadows, it is believed that they are subsequently discounted by the visual system so that they do not interfere with the detection and recognition of objects (see Casati, 2008; Rensink & Cavanagh, 2004).

Castiello et al. (2003) also postulated on the location of shadow information processing within the brain by comparing the perception of shadows by: subjects with temporal lesions and neglect, subjects with frontal lesions and neglect, and subjects with lesions with no neglect, and controls. Their findings showed that subjects with temporal lesions and neglect performed much worse when the shadow was presented to the contralesional side in comparison to the unaffected field. This difference was not apparent in other subjects. Their conclusion from this supports their temporal prediction that integration of a shadow and its object occurs within the temporal areas of the right hemisphere. These findings are supported by previous observations of homologous data in experimental lesions in monkeys (Hietanen, Perrett, Oram, Benson, & Dittrich, 1992; Vogels & Biederman, 2002). More specifically, it is thought that these areas are located within the superior temporal sulcus (STS) and other higher-order processing areas of the ventral stream such as inferotemporal cortex (IT). Though solid, these findings might be limited if temporal brain lesions in the subjects was extensive across many visual areas. Speculation regarding the neural locus for shadow processing is therefore difficult. Mamassian et al. (1998) has speculated that because cast shadows inform about the 3D structural form of objects, they would be effective in acting neural areas along the dorsal pathway which has been shown to be involved in determining the spatial layout of the visual scene.

In conclusion, the present study reports that the ability of the visual system to detect the structural form of cast shadow Glass patterns is dependent on the implied lighting direction. Lower detection thresholds are observed for light-from-above patterns compared to light-from-below patterns. This finding is consistent
with the existence of a coarse scale detector that detects cast shadows conveyed by light-from-above configuration, while a separate mechanism (one that is more sensitive to the alignment of local information in the stimulus) may underlie the extraction of the global form of light-from-below patterns. Additionally, shadow separation was also shown to facilitate detection, such that the visual system is less sensitive to cast shadows at larger separations.

Acknowledgments

The authors 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) to S. Khuu.

Commercial relationships: none.

Corresponding author: Sieu Khuu.
Email: s.khuu@unsw.edu.au.
Address: The School of Optometry and Vision Science, The University of New South Wales, Sydney, New South Wales, Australia.

References


Perception of scene layout from optic contact, shadows, and motion. Perception, 33, 1305–1318.


