The use of subtle illumination cues for human judgement of spatial layout

Cindee Madison *Daniel J. Kersten*
Psychology Department
University of Minnesota

William B. Thompson *Peter Shirley* *Brian Smits*
Computer Science Department
University of Utah

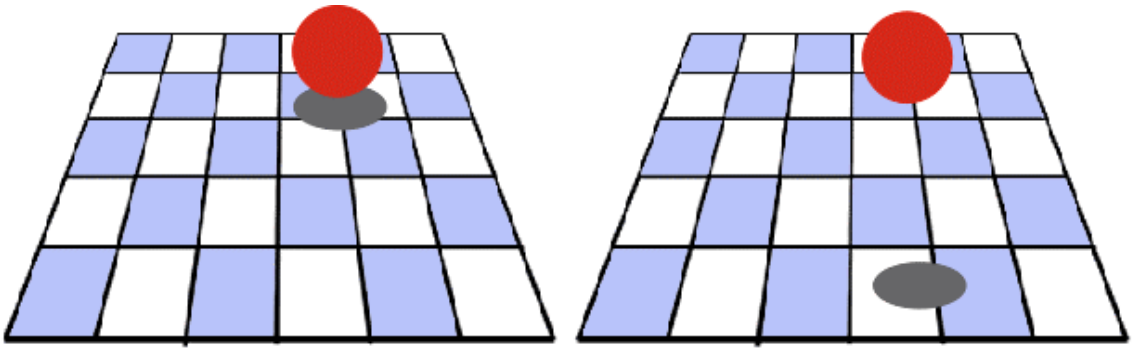
Abstract

The interaction of light with surfaces creates complex lighting effects that provide potentially useful information about the spatial relationships between objects. Previous research has shown cast shadows to be effective in determining the 3D layout of a scene. Interreflections are another source of information for spatial relationships; these, however, have been largely ignored in studies of human perception. The purpose of this study was to determine the effectiveness of interreflection for providing cues to contact, and to investigate how interreflection and shadow information combine in the perception of object contact.

Introduction

Determining the spatial relationships between objects in a scene is important for actions such as path planning, object avoidance, reaching, and grasping. In order to determine spatial layout, human vision takes advantage of the effects of surface occlusion of one object by another in the images received by the eyes. For example, if the surfaces are separated in depth, occlusion results in binocular disparity differences for the projections of corresponding feature points in the two eyes. Motion parallax is another well-known cue to surface depth relationships. There are also, so-called pictorial

cues to depth which require neither motion nor stereo vision. For example, in a monocular static image, the displacement between a cast shadow and the casting object in an image grows with the relative depth between two surfaces, and recent research has shown that shadows are sufficiently strong to override some other depth cues and determine the trajectory of an object in a scene [1,2]. Consider for example a ball on a homogenous plane. A shadow can perceptually drive the apparent location of the object.(figure 1) Even when the shadow is lightened, there can be an effect (figure 2), which has been shown to be strong for the case when the ball and shadow are both moving.

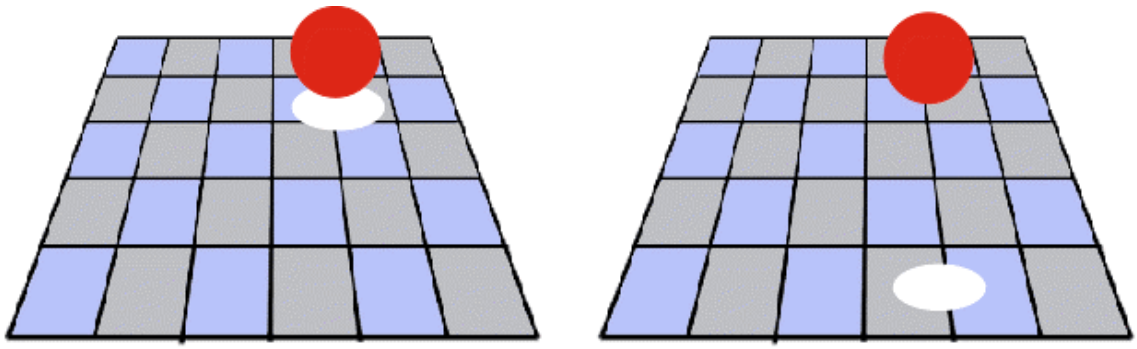


The shadow in this image makes the ball appear to be located close to floor in the back right corner of the floor.

The shadow in this image makes the ball appear to be well above the right front corner of the floor.

Figure 1: Ball-in-Box Images

In both images the red ball is in the same position on the checkerboard floor.



The white ellipse in this image makes the ball appear to be located close to floor in the back right corner of the floor.

The white ellipse in this image makes the ball appear to be well above the right front corner of the floor.

Figure 2: Ball-in-Box Images

In this image white shadows are used instead of dark shadows.
In both images the red ball is in the same position on the checkerboard floor.

Perceptual inferences about depth can be quantitative estimates, or discrete decisions. For example, binocular disparity, motion parallax, and cast shadow displacement are sources of quantitative information about relative depth. Human vision, however, is also effective at utilizing image information to determine qualitative depth relations, such as "in front of" or "behind". For example, surface occlusion produces unmatched feature pairs in the two eyes, so-called "da Vinci stereopsis", which, given knowledge of the eye-of-origin, determines which surface is in front [3]. An example of a monocular static cue for qualitative depth is a "T-junction" which is created in the image when one surface edge crosses another.

Human vision also makes other types of categorical perceptual judgments. In addition to perceiving whether one object is in front or behind another, the visual system can determine whether one surface is touching another. Such contact information is

particularly critical for reach and grasp. Contact information can also be effectively used if a large-scale frame-of-reference, such as the ground plane, has a known depth map. Then, knowledge of contact of an object with the ground plane determines the viewer's distance to the object. For example, by determining the contact between an object and a homogeneously textured ground plane observers can quite accurately judge the distance to the object [4]. Research on children has shown shadows to be effective in influencing depth and height judgements of an object on a plane, an ability that was found to improve with age.[5].

Although motion and binocular parallax are potentially useful for determining surface contact through the detection of surface orientation discontinuities, it is worth considering whether there might be effective monocular static sources of contact information. To be used reliably, cues for relative depth require image features that correspond to surface points, rather than to highly variable image features due to the effects of illumination. For example, specularities change with eye and light position. Shadows and interreflections between surfaces change with variations in relative object and light source positions. Because these types of image intensity changes are not tightly tied to the locations of surface points, illumination has historically been considered something to be discounted in studies of both human and computer vision.

Relationship between shadows and interreflections. Though historically ignored, the interaction of light with objects results in a number of subtle illumination effects which may be useful cues for surface attributes and relations. As noted above, cast shadows are effective in determining the 3D layout of a scene, and other studies are finding shadows play a role in object recognition [1,2,6] Interreflection is another effect that is closely related to shadows [7]. Consider the intersection of two surfaces (figure 3).

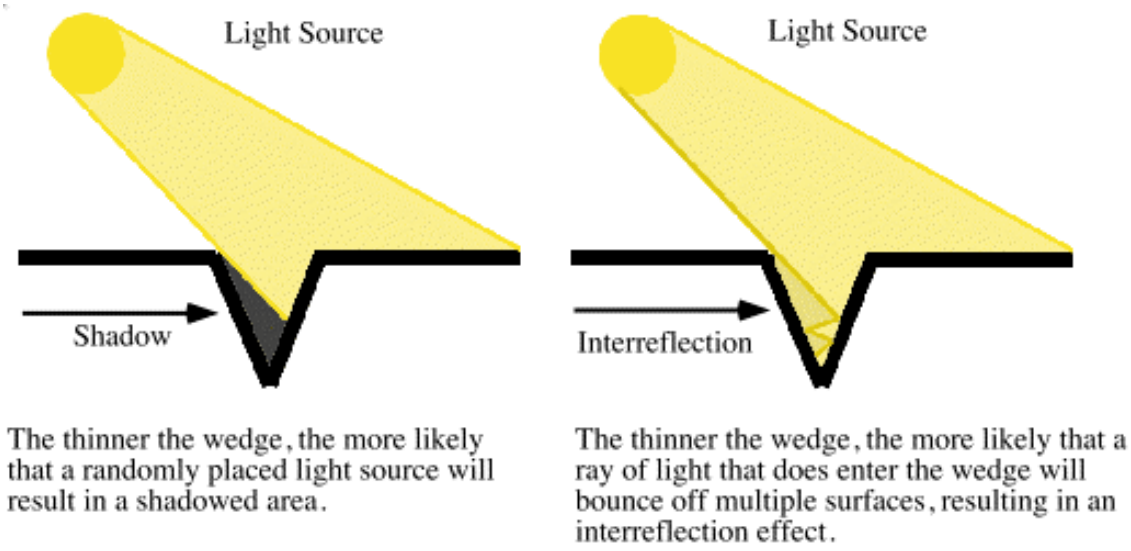


Figure 3: Relationship Between Shadows and Interreflections

Decreasing the angle between the surfaces also decreases the likelihood for a randomly placed light source to strike either of the surfaces; hence the more likely for there to be a shadow. However, as the angle between the two surfaces decreases, any light source that does enter will result in more bounces between the surfaces, and hence the more likely for there to be interreflection. Since interreflection is closely tied to shadows, it may also affect visual processes that incorporate shadows. Recent computational studies have shown how interreflection can be used to extract color or shape from image intensity values [8-11]. However, even though interreflections have image properties similar to shadows, little is known about the effectiveness of interreflections for the human perception of spatial layout, or other perceptual tasks.

Surface albedo and interreflections. The subtlety of interreflection raises questions about its importance to the human visual system. There is evidence, however, that indicates perception can use interreflection information. A study by Gilchrist and Jacobsen [12] studied human responses to "a world of one reflectance." Two small rooms with objects were painted either all matte white or all matte black. Participants could accurately distinguish the white room from the black room, even when the rooms were illuminated to reflect the same average luminance. However, measurement of the luminance profiles of the two rooms found greater variation in the black room than in the white room. Because interreflection is greater between the white surfaces, the white room will have a more uniform luminance profile than the black room. These results suggest human sensitivity to intensity variance, a function of indirect illumination, to determine surface albedo.

Kersten, Hurlbert and Bloj reported results in which participants viewed a concave folded card consisting of two opposing surfaces [13,14]. One surface of the card was white paper, the other red paper. For a concave card, interreflections naturally result in the white side becoming tinged with pink. (Note that if the card was folded to be convex, there would be no interreflection, so a white surface would remain white, unaffected by the adjacent red card.) Although the physical shape of the card remained constant, the *perceived* shape of the card, concave versus convex, was driven by a stereo cue. They found that participants accurately discounted the interreflection in the concave condition, thus seeing the white paper as white. But when the same image was seen as convex, observers interpreted the white surface as being a pinkish paper instead of white. So given cues to shape, the viewer can discount effects of illumination, again showing sensitivity to interreflection.

Spatial layout, interreflections and shadows. Both of the above experiments showed that human vision can use interreflection information, but both involved judgements of albedo. If an observer is unsure of the geometry between two surfaces, can interreflections be used to judge spatial relationship? The purpose of this study is to investigate whether human vision uses interreflections to determine object contact and if so, how this information combines with shadows.

Methods

Computer graphics was used to simulate images of a rectangular box which was either in contact or slightly above a floor. The floor was a continuous, textured ground plane (Figure 4). Shadows and interreflections were generated based on a model of the

physics of light transport [15]. Research participants viewed these images and made decisions about the contact of the box with the ground plane.

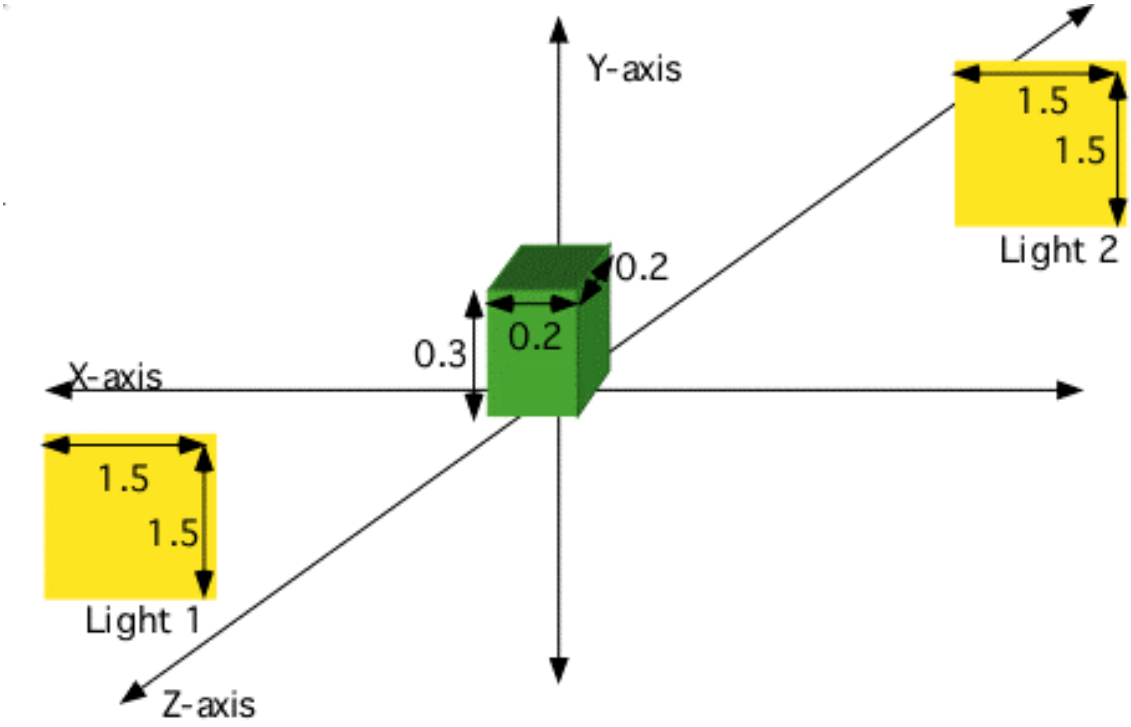


Figure 4: Geometry of the Scene:
The green Box is centered at the origin. Light 1 is centered at (-4.0, 2.25, 12.0). Light 2 is centered at (1.0, 2.25, -12.0). In the touching condition, the Green Box lies on the floor plane, in the above condition it is 0.010 units above the floor. All measurements are in absolute units.

The objects were modeled with simple geometry and the material of both the box and the floor were ideal lambertian reflectors. The box has the albedo of the green square in the top row of the Macbeth ColorChecker Chart [16], and the floor is grey with albedo 0.7. The light sources were modeled as diffusely illuminating rectangles (1.5 X 1.5 units) with a grey (flat) intensity curve. The computed images were tone mapped to an image format with 256 levels per channel, where zero intensity was mapped to the lowest image level,

the largest RGB component was mapped to the largest image level, and the rest of the pixels were mapped linearly between those extremes The scene was generated in four (one light source, Figure 5) conditions, and four (two light source, Figure 6) conditions: 1) No Shadow + No Interreflection; 2) Shadow + No Interreflection; 3) No Shadow + Interreflection; 4) Shadow + Interreflection.

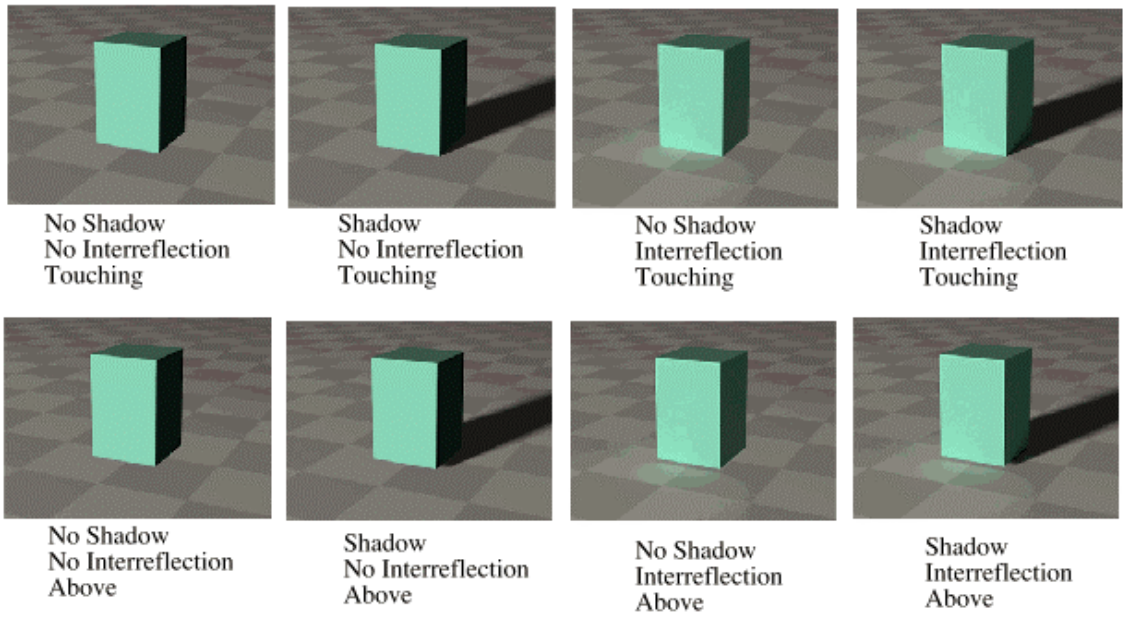


Figure 5: One Light Source Images

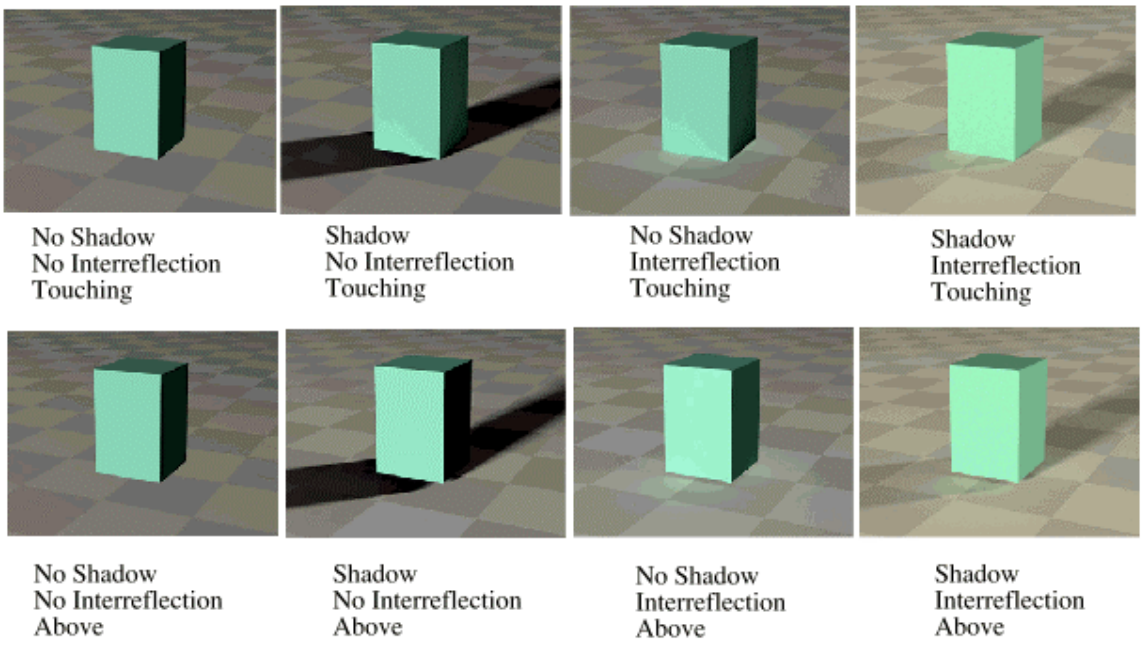
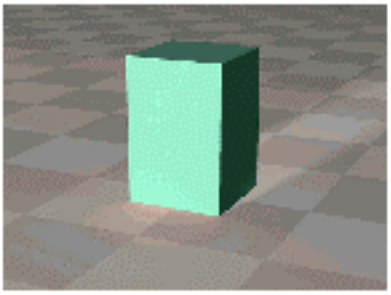
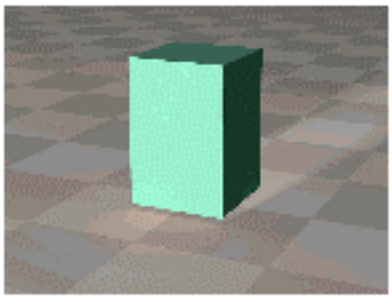


Figure 6: Two Source Light Images

A final condition was used in which the image rendering was not based on a realistic model of the physics of light. This condition had one light source, a white shadow, and a photometrically incorrect (red instead of green) interreflection (Figure 6). In each of the above conditions the scene was rendered twice, with the box slightly above the floor, and with the box touching the floor.



White Shadow
Red Interreflection
Touching



White Shadow
Red Interreflection
Above

Figure 7: The "Faked" Images

Twenty-two participants viewed each image ten times in a random order. Participants made a judgement in which they ranked their confidence regarding the contact between the floor and the box. The degree of contact choices were: definitely touching, maybe touching, unsure, maybe above, and definitely above. The images were displayed at a resolution of 1024 X 768 on an Applevision 1710AV display at a distance of 64cm. The images were displayed at a gamma of 1.8, the monitor was calibrated using the following gamma correction applet (http://www.cs.berkeley.edu/~efros/java/gamma/gamma.html), and the images had a mean luminance of 120 cd/m^2, with shadow areas averaging around 15 cd/m^2 and the highly illuminated areas averaging around 200 cd/m^2. All participants had normal or corrected to normal vision, gave their informed consent, and received extra credit for their participation in a research experience program connected to a first year psychology class.

Data Analysis

Participant responses were analyzed independently. Because each image can be considered as having a “correct answer”, either touching or above as determined by the geometry of the image, Receiver Operating Characteristic (ROC) curves could be generated for each individual, for each condition. The ROC curve is generated by determining the correct number of responses for each criterion level (definitely touching, maybe touching, unsure, maybe above, and definitely above), and plotting hits (object is touching and participant says it is touching) against false alarms (object is not touching and participant says it is touching.)[17] The area under the ROC curve gives a measurement of sensitivity ranging from 0 to 1 with an area of one showing the highest sensitivity. Figure 8 shows the average area for the participants for each condition with standard error bars.

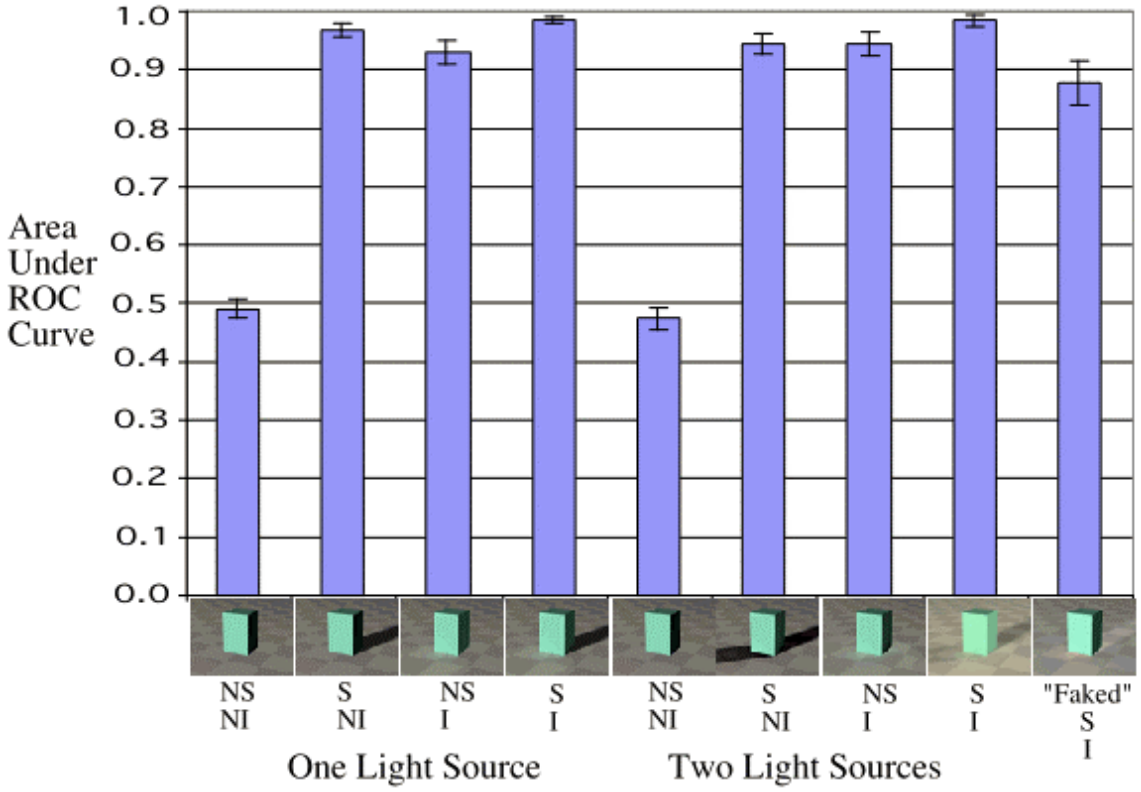


Figure 8: Participant results (grouped by image condition):
NS = no Shadow **S = Shadow**
NI = no Interreflection **I = Interreflection**

Results

It is important to note that in the case of no shadows or interreflections, performance for determining object contact was at chance. This suggests that participants did not have biased judgements resulting from the effect of other cues, such as coplanarity with the floor plane.

Adding just shadows or just interreflections resulted in a higher sensitivity to object contact. A combination of shadows with interreflections resulted in the highest sensitivity and was significantly greater than in images that had only shadows or only interreflections. Also note that the "faked" shadow and interreflection condition also resulted in a relatively high sensitivity. Though the illumination effects were not physically correct, subjects could still make an accurate response. This suggests that some property other than the color of the shadow and interreflection help play a role in signifying object contact. For instance, the alignment of the shadow contour with the corner of the box may be an important indicator, with the coincidental alignment of the contours at the corner point indicating contact. However, the most accurate judgement did require both correct shadow and correct interreflection, suggesting that agreement of all the cue information led to the best judgement.

Conclusion

The results of this research show that shadows and interreflections are important cues for object contact. Though interreflections appear more subtle than shadows in the image, under the conditions of this experiment, they convey equally strong information about object contact. However, the greatest sensitivity results from a combination of

both shadows and interreflections. When taken together with shadows, interreflection light effects may prove to be valuable cues to spatial layout and in general to a visual system that uses light information to make sense of the outside world.

References

[1] D. Kersten, D.C. Knill, P. Mamassian, and I. Bulthoff, “Illusory motion form shadows,” *Nature*, vol. 379, p.31, 1996.

[2] D.Kersten, P. Mamassian, and D.C. Knill, “Moving cast shadows induce apparent motion in depth,” *Perception*, vol. 26, pp. 171-192, 1997.

[3] K. Nakayama, and S. Shimojo, " da Vinci stereopsis: depth and subjective occluding contours from unpaired image points," *Vision Res* , vol. 30, pp. 1811-25, 1990.

[4] M.J. Sinai, T.L. Ooi, and Z.J. He, “Terrain influences the accurate judgement of distance,” *Nature* Vol. 395, pp 497-500, 1998.

[5] A. Yonas, L.T. Goldsmith, & J. L. Halstrom, "Development of sensitivity to information provided by cast shadows in pictures," *Perception*, vol. 7, pp. 333-341, 1978.

[6] M.J. Tarr, D. Kersten, and H. Bulthoff, “Why the visual system might encode the effects of illumination,” *Vision Research*, vol. 38, pp.2259-2275, 1998.

[7] M. Langer, "When do shadows become interreflections?" *International Journal of Computer Vision* (in press).

[8] S.K. Nayar, K.Ikeuchi, & T. Kanade, "Shape from interreflections" *International Journal of Computer Vision* , vol. 6, pp. 173-195, 1991.

[9] B.V. Funt, M.S. Drew, & J. Ho, "Color constancy from mutual reflection" *International Journal of Computer Vision* , vol. 6, pp. 5-24, 1991.

[10] D. Forsyth, & A. Zisserman, "Reflections on shading," in *Physics Based vision: Principles and Practice: Shape Recovery* (Wolff, LB., Shafer, S.A., & Healey, G.E. eds.), Jones and Bartlett Publishers, Boston, 1992.

[11] C.G.Christou, & J.J. Koenderink, "Light source dependence in shape from shading," *Vision Research* , vol. 37, pp. 1441-1449, 1997.

[12] A. Gilchrist and A. Jacobsen, “Perception of lightness and illumination in a world of one reflectance,” *Perception* , vol. 13, pp.5-19, 1984.

[13] D. Kersten, and A. Hurlbert, “Discounting the color of mutual illumination: A 3D shape-induced color phenomenon,” *Investigative Ophthalmology and Visual Science,* vol. 37, no.3, 1996.

[14] M. Bloj, A. Hurllbert, and D. Kersten. (personal communication).

[15] W.B. Thompson, P Shirley, B. Smits, D.J. Kersten, and C. Madison, “Visual Glue,” Technical Report UUCS-98-007,University of Utah, March 1998.

[16] C.S. McCamy, H. Marcus, and J.G. Davidson, "A Color-Rendition Chart," *Journal of Applied Photographic Engineering,* vol. 2, pp 95-99, 1976.

[17] D.M. Green, and John A. Swets, *Signal Detection Theory and Psychophysics.* Huntington, New York: Robert E. Krieger Publishing Company, 1974.