Maybe they are all circles: Clues and cues

Chris M. P. Muller Research Institute MOVE, Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands

Eli Brenner Research Institute MOVE, Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands

Jeroen B. J. Smeets Research Institute MOVE, Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands

When several cues provide information about the same property of a visual scene, a weighted average of the singe-cue estimates can provide a more reliable estimate than that of any individual cue. Some cues rely on assumptions about the scene, such as that shapes are isotropic. Assuming that an elliptical image arises from viewing a circle at an angle allows one to extract the circle's angle from the aspect ratio in the image. This study investigates whether the weight given to image shape as a slant cue depends on the prevailing circumstances. Neither rotating an object to provide direct evidence that it is circular, nor surrounding an object with circles rather than ellipses increased the weight assigned to image shape relative to that assigned to binocular information. Thus the weight given to slant cues does not seem to rely on an elaborate analysis of the scene.

Keywords: 3D surface and shape perception, binocular vision, space and scene perception

Citation: Muller, C. M. P., Brenner, E., & Smeets, J. B. J. (2009). Maybe they are all circles: Clues and cues. *Journal of Vision, 9*(9):10, 1–5, http://journalofvision.org/9/9/10/, doi:10.1167/9.9.10.

Introduction

Knowing a surface's shape enables us to judge its three-dimensional slant from the two-dimensional image on the retina. For instance, if we know that the surface is circular we can use the aspect ratio of the elliptical image to estimate the slant. Although judging the slant from the aspect ratio is only justified if one knows that the surface is circular, observers use the aspect ratio even if evidence that the surface is circular is lacking. They presumably do so because they implicitly consider that real surfaces are more likely to be circles than ellipses (Knill, 2007). Considering how likely it is to encounter various objects is essential for choosing between all the possible interpretations of a given retinal image (Chan, Stevenson, Li, & Pizlo, 2006). Similarly, considering that motion is slow and smooth guides people's choice between the possible interpretations of ambiguous two-dimensional projections of rotating three-dimensional rotating objects (Rokers, Yuille, & Liu, 2006).

When we look at the wheel of a bicycle in daily life, we usually see it at some angle. The image of the wheel on our retina is therefore usually deformed into an ellipse. Because we know that the wheel is really circular, we can use this deformation to judge its orientation. Without that knowledge we cannot, because the same elliptical image on the retina could arise from viewing a circle at a certain angle, or from viewing any of a wide range of ellipses from other specific angles. Actually, the retinal image of a slanted circle is not a perfect ellipse, but in this paper we will consider retinal projections of circles to be perfect ellipses, because the additional deformations are usually small enough to be ignored (0.02 minutes of arc at most in this study).

If one looks at a large collection of bicycles, such as those that can be found outside a Dutch university, one typically sees bicycle wheels from every possible angle. If one of the bicycles happened to have an elliptical wheel, we would probably interpret its shape as belonging to a bicycle with normal circular wheels that was hastily put away in an awkward manner so that we view the wheel from an unusual angle, rather than interpreting it to be elliptical. On the other hand, if we looked at a large collection of deformed bicycles in a painting by Salvador Dali, we would probably be less confident that the wheels were depictions of circular wheels. Many studies have shown that the context, both past and present, can affect the way we interpret a scene (Jacobs & Fine, 1999; McKee, 1983; Mitchison & Westheimer, 1984; van Ee, Banks, & Backus, 1999).

The extent to which the visual system makes certain assumptions should obviously depend on how likely it is that the assumptions are correct (Knill, 2007; Mamassian & Landy, 2001). Given that the learned assumptions that our visual system relies on to disambiguate a scene can be context-specific (Knill, 2007), we can expect the contents of a scene to affect the visual system's confidence in various assumptions. If it does, we should be able to instantaneously influence the confidence in the assumption

doi: 10.1167/9.9.10 Received March 31, 2009; published August 24, 2009 ISSN 1534-7362 © ARVO

that surfaces are likely to be circular by simply changing the viewing circumstances. A change in such confidence should in turn give rise to a change in the weight that is assigned to image shape in a weighted average of all available cues for estimating slant, because the perceived slant is considered to be a weighted average of several cues (Hillis, Watt, Landy, & Banks, 2004; Knill & Saunders, 2003; Muller, Brenner, & Smeets, 2008). We tried to manipulate the confidence in the assumption of circularity in two ways: by rotating the surface so that the assumption that it is a rigid object would be violated if it were not a circle, and by surrounding the surface by either circles or by clearly non-circular ellipses.

Methods

Observers

Ten participants took part in this study. Two of them were authors. The other eight were experienced psychophysics observers who were naive as to the purpose of the study. All had normal or corrected to normal vision, and had normal stereo acuity. Each observer took part in both a main session and an additional session that were performed on separate days. Each session was preceded by 5 practice trials that were not recorded. The study was approved by the ethics committee of the Faculty of Human Movement Sciences.

Setup

Our setup consisted of an Apple G5 computer that generated the images and processed the responses, a 57 cm (diagonal) Sony Trinitron monitor (resolution 1096 × 686 pixels) on which the images were displayed, and Crystal Eyes stereo shutter spectacles that allowed us to present alternate images to the two eyes. The images were generated at a refresh rate of 160 Hz (80 Hz per eye). Observers sat 1 meter from the screen, so that the screen filled approximately 27° × 17° of visual angle.

Stimuli

The stimulus display was a simulation of an array of 7 × 7 elliptical or circular discs whose centers formed a plane that was slanted in depth (Figure 1). The central disc in the array served as a reference. It was always slanted around a horizontal axis in the frontal plane (at the center of the screen). Each of the other discs was constantly rotating around a random axis at 180°/s. This rotation axis passed through the center of the circle or ellipse in question, but was completely independent of the circle's or ellipse's initial slant. The centers of all the discs specified a plane that served as a probe. The observer could rotate this plane around its horizontal axis at the center of the screen. Changing the probe surface's slant did not influence the reference disc, or change the orientation of the probe discs or of their axes of rotation, it only changed the positions of the probe discs' centers.

The surface of the reference disc was rendered with a red-green gradient. The probe discs had a blue-purple gradient (see Figure 1). The different colors made it easy for observers to distinguish the reference disc from the other discs. The gradient made it possible to see the rotation even if a circle rotated around an axis orthogonal to the surface itself (i.e. around the surface normal). We used smooth gradients in order not to introduce additional information about slant from the rotation itself (motion parallax). Such information would be consistent with the slant from the aspect ratio, which could lead to a higher monocular cue weight. It is important to realize that rotating the reference disc does not increase the precision with which the slant can be derived from the aspect ratio, it only provides additional support for the assumption of circularity, thereby confirming that using the aspect ratio is justified.

The probe discs could either be circular, or elliptical. The constant rotation around random axes ensured that these shapes were clearly perceived as such. When the probe discs were circular, their (simulated) diameters were all 4 cm (giving a maximal width of 2.3° of visual angle at the 1 meter viewing distance). When the probe discs were elliptical, their major axes were always 4 cm but their minor axes varied randomly in length between 1.3 and 4 cm. The reference disc was a simulation of an elliptical object, and therefore contained a slant cue conflict if observers interpreted it as arising from a circular object; i.e. there was always a discrepancy between the slant indicated by its elliptical outline on the screen (assuming circularity) and the slant indicated by binocular disparity.

Paradigm

Observers set the slant of the probe surface to match the slant of the central (reference) disc by moving the computer mouse from left to right. They indicated that they were satisfied with their setting by clicking the computer mouse, which also started the next trial. On each trial we recorded the slant that observers set. Sixteen conditions were randomly interleaved in one session. There were 25 trials for each condition.

Conditions

There were four main conditions, arising from the four combinations of two different manipulations. The

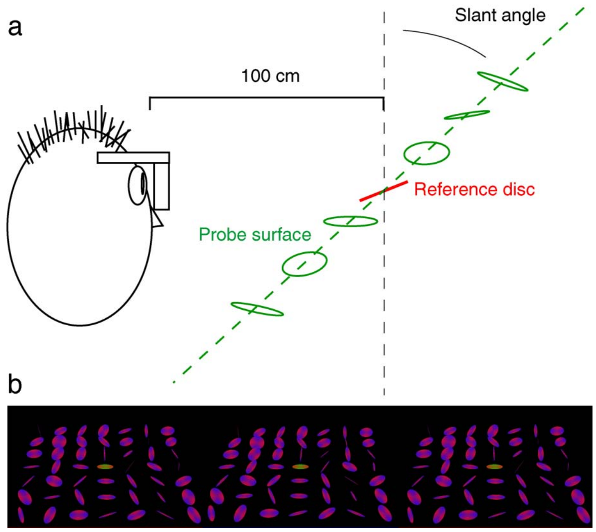

Figure 1. The task. (a) Diagram showing a side view of the task (not to scale). The orientation of the red line indicates the perceived slant of the reference disc. The observer's task was to rotate the probe surface so that its slant matched that of the reference disc. The orientation of the probe surface had to be inferred from the positions of the circles or ellipses. (b) Stereogram of one stimulus frame for either uncrossed (two leftmost images) or crossed (two rightmost images) fusion.

first manipulation was that the reference disc could rotate around its surface normal, rather than being static. This does not affect the instantaneous information about the disc's slant, but the fact that the outline does not change when the disc rotates provides evidence that the reference disc is circular. The other manipulation involved the shapes of the probe discs. These were either all circular or all non-circular ellipses. Presumably one would be biased toward assuming that the reference disc had the same shape as the surrounding discs.

Each of these four main conditions was performed using four different conflicts of the reference disc, leading to the total of 16 conditions. The discrepancy between the slants from image shape and binocular disparity was always 10 degrees, which is small enough for us to expect weighted averaging of the cues rather than the percept flipping between the values of the two cues (Muller et al., 2008; van Ee, van Dam, & Erkelens, 2002). The actual slants were the two possible combinations of a 63 degree slant and a 73 degree slant, and the two possible combinations of a 58 degrees slant and a 68 degrees slant. Slant was defined as how far backward the top was oriented from the fronto-parallel plane (see Figure 1a). The slants were chosen on the basis of the known influence of slant on the relative weights of monocular and binocular cues (Knill & Saunders, 2003).

Additional session

Beside the main session described above, we also subjected participants to an additional session that was identical to the original one except that there was no conflict between the cues for the reference disc (assuming that it is circular). Again there were 16 conditions, but rather than the four pairs of cue conflicts of the original session there were simply four reference angles: 58, 63, 68 and 73 degrees. The additional session was conducted in order to correct for possible influences of other, uncontrolled cues (image blur, lens accommodation, motion parallax if the observer does not hold his or her head completely still) and of possible response biases. How this correction was achieved is explained in the next section.

Analysis

In order to evaluate the effect of our manipulations we need to determine the weight given to the reference disc's image shape as a slant cue. Since this cue always indicated a different slant than was indicated by the binocular disparity, the extent to which observers relied on each type of information can be determined from the set slant. If the perceived slant is a weighted average of that indicated by the two cues, the set slant provides direct information about their relative weights. For instance, setting a slant exactly half way between that specified by the two cues indicates that both cues are given equal weight.

However, as we already pointed out, the perceived slant may also be influenced by other factors. This may be differently so for the reference disc than for the probe plane, which could result in response biases. To account for this we did not relate the set slants in the main session to the simulated slants, but to the slants set in the additional session without cue conflicts. We first determined each subject's average set slant for each of the 16 conditions in each session. We then used the averaged slants to determine the weight attributed to the monocular cue (image shape) in the main, cue-conflict session by comparing the set slant for the reference disc in each condition of that session with the two set slants for the reference discs when both cues had one of the two values within that conflict in the additional, (cue-consistent) session. Note that this is only really a correct procedure if the individual cue estimates are unbiased or have the same bias. Having exactly the right correction is not essential, because we are mainly interested in comparing the values across conditions, but it makes the weights we find more meaningful and hopefully removes response bias differences between observers. We performed a repeated measures ANOVA on the calculated monocular slant cue weights with the factors *conflict type* (63–73, 73–63, 58–68, 68–58), *reference motion* (rotating

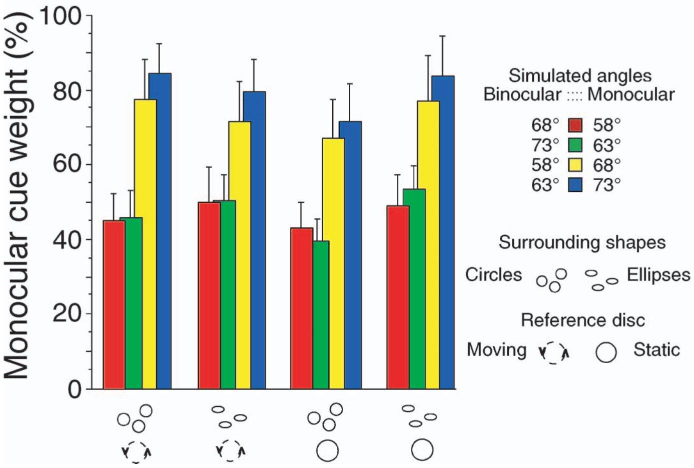

Figure 2. The average weight given to the monocular cue in each of the 16 conditions. The four groups of columns represent the 4 manipulations. The bar color indicates the cue conflict. Error bars are between-subject standard errors.

or static) and *surrounding shapes* (circular or non-circular ellipses).

Results

Figure 2 shows the weights that observers gave to the monocular slant cue (image shape). The four groups of columns represent the four main conditions. The four bar colors indicate the slant values for each cue. The clearest finding is that more weight is given to the monocular cue when the angle indicated by the monocular cue is larger than that of the binocular cue than when the angle indicated by the binocular cue is larger. This is consistent with earlier findings, and originates in the fact that the relative resolution of monocular slant information increases with increasing slant (Knill & Saunders, 2003). Apart from this effect of slant (F_3 = 4.0, P = 0.02), the ANOVA only revealed a significant interaction between target motion and surrounding shapes (F_1 = 19.9, P = 0.002). When surrounded by circles, rotating the reference disc increased the weight given to monocular cues, but when surrounded by ellipses it decreased this weight. The main effects were not significant. Overall the average monocular weight was slightly larger for the rotating reference disc than for the static one, as predicted. It was slightly smaller when surrounded by circles than when surrounded by ellipses, which is the opposite of what one would predict.

Discussion

It has been shown that observers gradually change the strength of their assumption of circularity after extensive exposure to non-circular ellipses (Knill, 2007). It is not known whether such changes can also occur instantaneously as a consequence of details of the context within which judgments are made. The fact that the influence of exposure in Knill's study carried over across sessions (on separate days) suggests that it is context specific. In that case the weights were changed within the context of the experiment despite experiencing normal circumstances between sessions. Here, we examined the context within individual trials.

We reasoned that placing a reference disc within an array of objects that are obviously not circular may make observers less inclined to assume that the reference disc is circular, in which case they should reduce the weight given to the monocular (image shape) cue for slant. Our results do not support this prediction. Thus people do not appear to consider the shapes of surrounding objects when determining the likelihood of the reference being circular in order to assign a weight to the monocular slant cue. This may be because observers were instructed to consider the array of shapes as a whole, although the probe itself was made up of the circles or ellipses, making it hard to ignore them. It may also be because the observers consider the shapes they saw on previous trials to be just as important as those on the specific trial in question, in accordance with the slow learning shown by Knill (2007). They may also not relate the target to the other items because they consider it to be a different kind of object (it had a different color, it did not always rotate, and it held the only position at which there was never a clearly non-circular object). Whatever the reason, apparently our observers did not consider the simultaneously visible context (the shapes that are visible during the trial) to be particularly important when evaluating the validity of the assumption of circularity (and using it to determine the weights assigned to the monocular and binocular cues).

On individual trials we also either did or did not provide direct evidence that the reference disc was circular, by having it either rotate around its axis or remain static. Observers did not assign a higher weight to the monocular cue for slant in the presence of evidence that the reference disc was circular than when such evidence was absent. Perhaps they judge shape and slant independently, just as they do shape and size (Brenner & van Damme, 1999).

Neither the main effect of target motion nor that of surrounding shape was significant, but we did find a significant interaction between the two. On average, the monocular cue was given most weight in the condition in which there was most support for using that cue (rotating target surrounded by circles). However, the weight was not lowest in the condition with the least support for using the monocular cue (static target surrounded by ellipses). The differences between the conditions were very small (see Figure 2), so although we cannot be certain that the two manipulations had no effect, if they did it is quite modest. If there is a small effect, one may be able to reveal it by presenting each condition separately, with the same environment on many trials, so that it accumulates. Regardless of whether this does or does not happen, our study demonstrates that the instantaneous context does not have much influence on the use of the monocular cue to judge slant.

It is well established that cues that provide different estimates for the same property of a visual scene are combined by weighted averaging, and that the more precisely the visual system can estimate a property from a certain cue, the more weight that estimate is assigned in the weighted average (Hillis et al., 2004; Knill & Saunders, 2003; Landy, Maloney, Johnston, & Young, 1995; van Beers, Sittig, & Denier van der Gon, 1999). It is also known that humans use assumptions to disambiguate retinal images and that these assumptions can be altered by prolonged viewing (Knill, 2007). In this study we attempted to alter the strength of the assumption of (radial) symmetry by direct visual information, and by doing so to manipulate the weight that observers attribute to a monocular cue for slant. The information that we provided did not affect the weights.

Acknowledgments

This research was supported by the Netherlands Organization for Scientific Research (NWO; MaGW grant 452-02-007).

Commercial relationships: none.
Corresponding author: Chris M. P. Muller.
Email: christiaan.muller@ugent.be.
Address: Research Institute MOVE, Faculty of Human Movement Sciences, Vrije Universiteit, Amsterdam, The Netherlands.

References

Brenner, E., & van Damme, W. J. (1999). Perceived distance, shape and size. *Vision Research, 39,* 975–986. [PubMed]

Chan, M. W., Stevenson, A. K., Li, Y., & Pizlo, Z. (2006). Binocular shape constancy from novel views: The role of a priori constraints. *Perceptions & Psychophysics, 68,* 1124–1139. [PubMed] [Article]

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision, 4*(12):1, 967–992, http://journalofvision.org/4/12/1/, doi:10.1167/4.12.1. [PubMed] [Article]

Jacobs, R. A., & Fine, I. (1999). Experience-dependent integration of texture and motion cues to depth. *Vision Research, 39,* 4062–4075. [PubMed]

Knill, D. C. (2007). Learning Bayesian priors for depth perception. *Journal of Vision, 7*(8):13, 1–20, http://journalofvision.org/7/8/13/, doi:10.1167/7.8.13. [PubMed] [Article]

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43,* 2539–2558. [PubMed]

Landy, M. S., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: In defense of weak fusion. *Vision Research, 35,* 389–412. [PubMed]

Mamassian, P., & Landy, M. S. (2001). Interaction of visual prior constraints. *Vision Research, 41,* 2653–2668. [PubMed]

McKee, S. P. (1983). The spatial requirements for fine stereoacuity. *Vision Research, 23,* 191–198. [PubMed]

Mitchison, G. J., & Westheimer, G. (1984). The perception of depth in simple figures. *Vision Research, 24,* 1063–1073. [PubMed]

Muller, C. M. P., Brenner, E., & Smeets, J. B. J. (2008). Testing a counter-intuitive prediction of optimal cue combination. *Vision Research, 49,* 134–139. [PubMed]

Rokers, B., Yuille, A., & Liu, Z. (2006). The perceived motion of a stereokinetic stimulus. *Vision Research, 46,* 2375–2387. [PubMed]

van Beers, R. J., Sittig, A. C., & Denier van der Gon, J. J. (1999). Integration of proprioceptive and visual position-information: An experimentally supported model. *Journal of Neurophysiology, 81,* 1355–1364. [PubMed] [Article]

van Ee, R., Banks, M. S., & Backus, B. T. (1999). An analysis of binocular slant contrast. *Perception, 28,* 1121–1145. [PubMed]

van Ee, R., van Dam, L. C., & Erkelens, C. J. (2002). Bi-stability in perceived slant when binocular disparity and monocular perspective specify different slants. *Journal of Vision, 2*(9):2, 597–607, http://journalofvision.org/2/9/2/, doi:10.1167/2.9.2. [PubMed] [Article]