3D Shape Recognition with Stereo

S. C. de Vries, A. M. L. Kappers, and J. J. Koenderink
Utrecht Biophysics Research Institute, The Netherlands

Abstract. We report here on the use of two novel quantities, the shape index and the curvedness, in shape-from-stereo measurements. In a recognition task experiment subjects are able to categorize shapes which are shown as random dot stereograms on the basis of their shape index. We also find that shape perception in the investigated region is scale independent. Detection of hyperbolic surfaces is not as good as detection of elliptic surfaces.

Introduction

Research on stereo has been done for over a century. All kinds of aspects of stereo vision have received quite extensive attention. With the introduction of the random dot stereogram (RDS) by Julesz (Julesz, 1960) research on stereo was boosted. This paradigm openend whole new worlds for research. Most of the experiments done with RDS were concerned with flat surfaces in depth, whereas more complicated forms were mostly neglected perhaps due to the amount of computer time that is involved in generating this kind of surfaces. Whenever studies on 3D form perception were done, shapes were chosen rather arbitrarily (Uttal, 1987) or the set of shapes investigated was restricted (Johnston, 1988; Rogers, 1989). Here we present a more systematic approach to the perception of shapes with stereo.

First we need a more concise description of what the meaning of 'shape' is. We will restrict ourself to local surface patches. After a convenient choice of the base we can write this surface in the Monge patch representation $z = f(x,y)$. This can be rewritten as a Taylor series expansion:

$$z = a + bx + cy + dx^2 + exy + fy^2 \ldots$$

Another base transformation reduces this to

$$z = 1/2(K_1x^2 + K_2y^2) + \ldots$$

As a first approach to shape description we can use the second order terms, which are connected to the surface curvature. Although on a patch we can draw an infinite number of curved lines their curvature is constrained in a way already revealed by Gauss. He showed that for an arbitrary curvature K the following relation holds: $K(\phi) = K_+\sin^2\phi + K_-\cos^2\phi$, with K_+ and K_- the so-called *principal curvatures* having maximum curvature and minimum curvature on the patch respectively and ϕ the angle between K and the curved line (the abovementioned

K_1 and K_2 correspond to K_+ and K_-, respectively). Gauss also showed that K_+ and K_- are perpendicular to each other. Koenderink (Koenderink, 1990) uses K_+ and K_- to define the *shape index*, a quantity that describes form in second order. It is defined as

$$S = -\frac{2}{\pi}\arctan\frac{K_+ + K_-}{K_+ - K_-} \tag{1}$$

The shape index is scale independent, a feature we appreciate in a quantity that describes shapes, because, for instance, spheres of different radius have different curvatures but we regard their shape as being the same. In Figure 1 the shape index is illustrated. As can be seen in this figure the shape index scale can be divided in a number of different regions. The first division you can make is at the middle where at the left of 0 the shapes are concave and to the right they are convex. A second division at -0.5 and at 0.5 separates the elliptic shapes from the hyperbolic shapes.

From this description, it seems that the shape index approach to shape definition is powerful enough for the first fundamental steps into the 3D shape perception area. It can be used not only for shape from stereo, but for every shape from X research (where X can be motion, shading, texture, etc.). We report here its use in recognition tasks, but we have also used it for shape discrimination measurements, where it proved to be a very useful tool too.

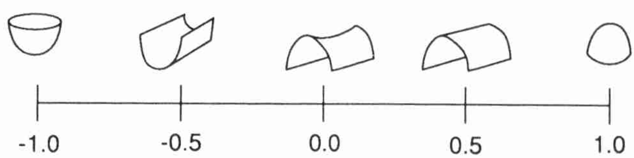

Figure 1. The shape index scale.

In addition to S Koenderink (Koenderink, 1990) proposes another shape quantity, the curvedness. It is defined as

$$C = \sqrt{\frac{K_+^2 + K_-^2}{2}} \tag{2}$$

and it weighs the *amount* of curvature of a surface.

Methods

We use an ATARI 1040 ST computer to display the RDS. The frame rate of the monitor of the ATARI is 70 Hz. Between frames the display is switched from depicting the left eye's view to the rigth eye's view. A pair of LCD-shutter spectacles running synchronously to the monitor is used to select the right image for each eye. Subjects sit at a 60 cm distance from the display with their head in a chin rest. They have normal or corrected-to-normal vision and their stereo vision abilities are checked by means of a series of test images.

The shapes we use in this experiment belong to the family of the quadratic surfaces given by $z = (K_+x^2+K_-y^2)/2$. K_+ and K_- can be calculated from (1) and (2) when the shape index and curvedness are given. The quadrics are calculated on a 6 x 6 cm square domain centered on the screen. The rest of the screen is filled with random dots to serve as a background with the same disparity as the screen. The quadrics (or better: the cyclopian images) are translated 7.5 cm in front of the screen so that no parts of the shape will lie behind the screen, which can serve as an unwanted shape cue. Starting at 2.5 cm from the center the quadratic shapes are blurred to hide the form of the boundary which can also be an unwanted cue. Within the range of 2.5 cm only foreground pixels are visible. Outside of the range of 3 cm from the center only the random dot background is visible. In the blurred region the ratio of foreground to background pixels decreases with distance from the center.

The subjects are shown quadratic RDSs drawn randomly from a pool of 501 precalculated RDSs which are equally distributed along the shape index scale. Display time is 3 s. The subjects' task is to categorize each image into one of eight possible categories. The categories are also equally spaced along the scale. Each subject gets three sessions of 200 stimuli. Experiments are done with three values of the curvedness, 0.3, 0.5, and 0.8 cm^{-1}. In a final experiment the curvedness has random values ranging from 0.3 to 0.8 cm^{-1}. The value of 0.3 cm^{-1} is about the lowest that can be represented given the resolution constraints of the screen. The maximum value of 0.8 cm^{-1} is not the highest possible, but it has a reasonable high curvature. It is about the curvature of a thumb.

Results

In Figure 2 the results for one subject and a curvedness of 0.5 cm^{-1} are shown. Depicted in this figure are the results of the subject for each of the eigth categories (which were, of course, offered to the subject in a random order). So if we take, for instance, the upper left graph, we see the distribution of answers whenever the subject was offered a shape somewhere in category 1 (shape index -1 – -0.75). A peak at category 1 means that it is recognized as such. It can be seen from the figure that scores for the elliptic regions (cat. 1, 2, 7, and 8) are higher and with smaller distributions than in the saddle shaped regions (cat. 3, 4, 5, and 6). Results of other subjects are similar. The scores are always centered around the category that is offered to the subjects, with a peak in the middle. So confusion of shapes decreases with distance on the shape index scale. This is an important conclusion because it indicates that we can use this distance as a measure of discrimination performance in discrimination experiments.

The results for the two other curvednesses were qualitatively the same. This means that curvedness and (which is almost the same) scale does not influence shape perception, at least not in the range of curvednesses we have investigated.

This can be concluded also from Figure 3. In this figure the results of the categorization experiment with the randomized curvedness are shown.

In this experiment subjects could not use absolute curvature measurements of one of the principle curvatures as could be done in the previous experiments. Here, the subjects had to estimate the ratios of the principle curvatures before being able to categorize the shape. Again, the results are qualitatively the same as in Figure 2. In Figure 4 results are plotted differently. We have divided the curvedness range into 15 categories and we have counted the number of correct responses in each category. The figure shows that there is no correlation between scores and curvedness, again this indicates that shape perception is scale independent.

Discussion

Whether the poor results in the hyperbolic area are due to the human visual system or are inherent to the task is not clear. To categorize elliptic shapes, a qualitative comparison of the two principle curvatures and an estimate of one of them is sufficient for categorization. For hyperbolic shapes one needs to compare the principle curvatures more carefully. This could explain the results in the experiments with constant curvedness. On the other hand, the similarity of these results and those from the experiments with random curvedness (where the task becomes more difficult in the case of elliptics) suggests that it is not the task itself that is the cause of the difference in perception between elliptic and hyperbolic shapes. In a pilot experiment not described here, subjects had to respond with values of shape index instead of category number. In this experiment accuracy was needed in both the elliptic and hyperbolic cases. Again

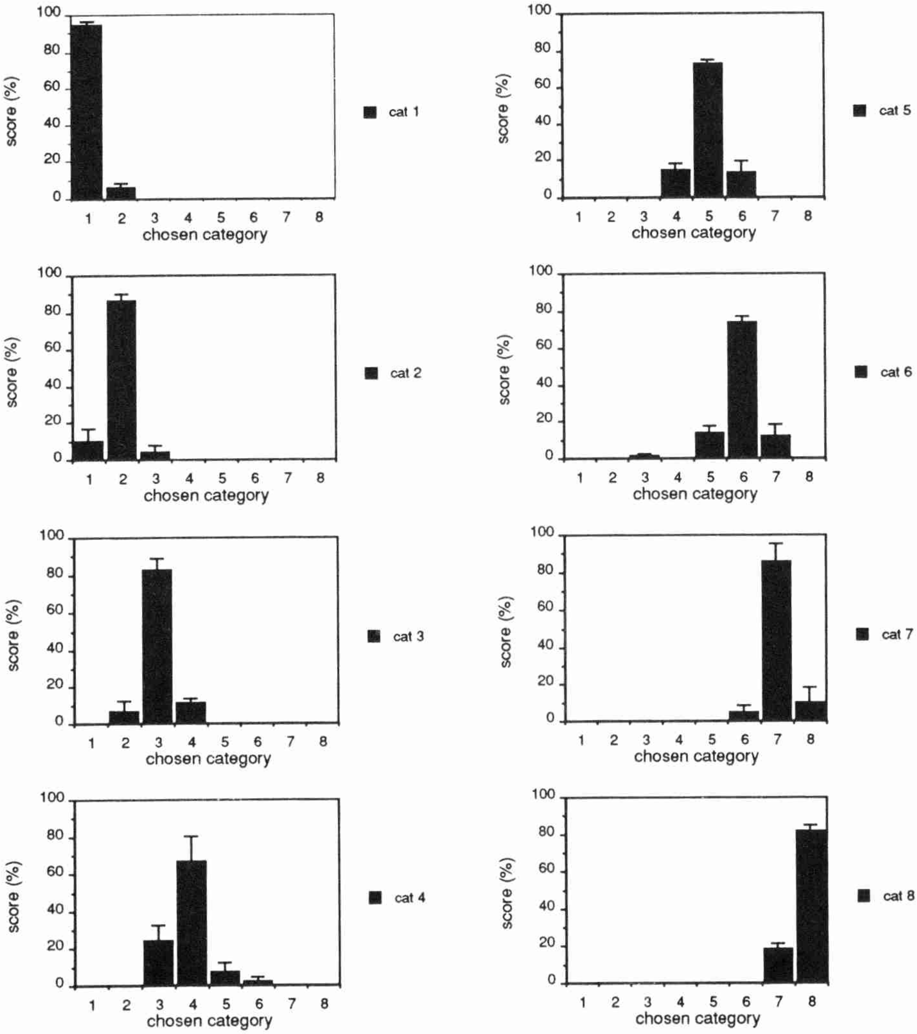

Figure 2. Results of the categorization experiment with $C = 0.5\ \text{cm}^{-1}$.

answers were more precise in the elliptic regions than in the hyperbolic regions, indicating that the difference is a perception matter and does not result from the task.

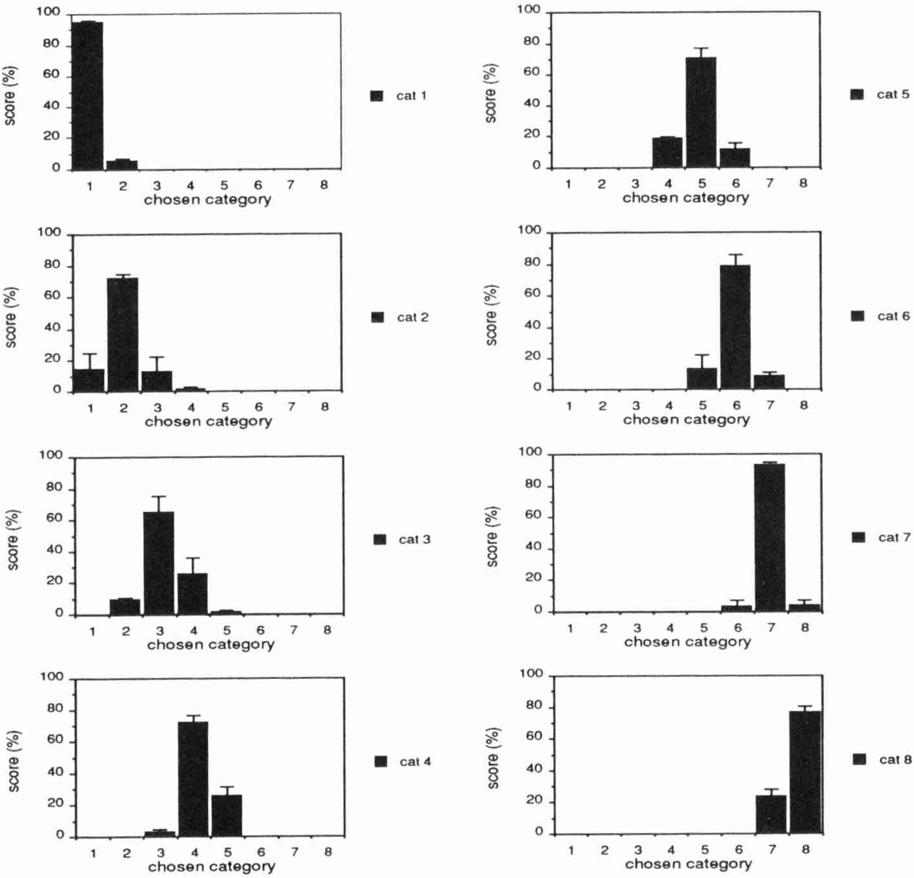

Figure 3. Results of the categorization experiment with curvedness in the range 0.3—0.8 cm^{-1}.

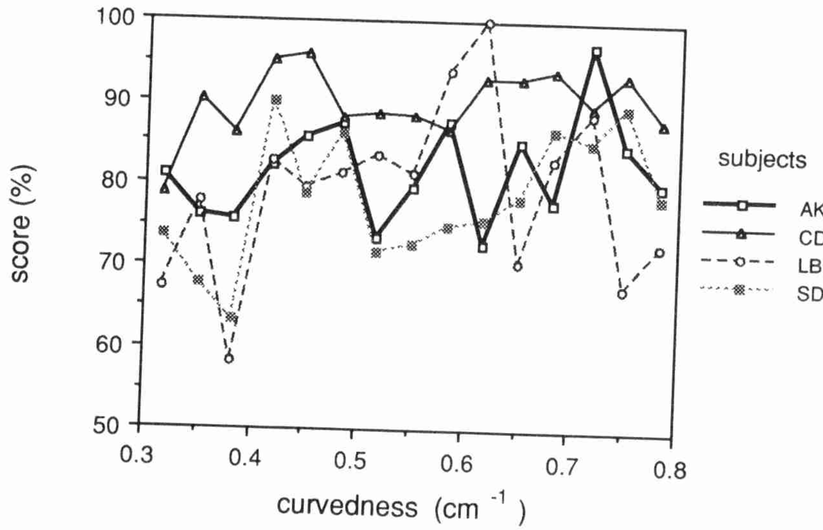

Figure 4. Percentage of correct answers for each of 15 curvedness categories.

Acknowledgment. This research is sponsored by the SPIN project '3D-Computer Vision' of the Dutch Ministry of Economic Affairs.

References

Johnston, E. B. (1988). Systematic distortion of shape from stereo. *Inv. Opthalm. Vis. Sci.*, *291*, 349.
Julesz, B. (1960). Binocular depth perception of computer generated patterns. *Bell Systems Technical Journal, 39*, 1125—1162.
Koenderink, J. J. (1990). *Solid shape.* Cambridge, MA: MIT Press.
Rogers, B. (1989). Disparity curvature and the perception of three-dimensional surfaces. *Nature, 339*, 135—137.
Uttal, W. P. (1987). *The perception of dotted forms.* Hillsdale, NJ: Lawrence Erlbaum Associates.