Three-dimensional symmetric shapes are discriminated more efficiently than asymmetric ones

Zili Liu

Department of Psychology, University of California, Los Angeles, 1285 Franz Hall, Box 951563, Los Angeles, California 90095

Daniel Kersten

Department of Psychology, University of Minnesota, Elliott Hall, 75 East River Road, Minneapolis, Minnesota 55455

Received September 19, 2002; revised manuscript received February 10, 2003; accepted March 10, 2003

Objects with bilateral symmetry, such as faces, animal shapes, and many man-made objects, play an important role in everyday vision. Because they occur frequently, it is reasonable to conjecture that the brain may be specialized for symmetric objects. We investigated whether the human visual system processes three-dimensional (3D) symmetric objects more efficiently than asymmetric ones. Human subjects, having learned a symmetric wire object, discriminated which of two distorted copies of the learned object was more similar to the learned one. The distortion was achieved by adding 3D Gaussian positional perturbations at the vertices of the wire object. In the asymmetric condition, the perturbation was independent from one vertex to the next. In the symmetric condition, independent perturbations were added to only half of the object; perturbations on the other half retained the symmetry of the object. We found that subjects' thresholds were higher in the symmetric condition. However, since the perturbation in the symmetric condition was correlated, a stimulus image provided less information in the symmetric condition. Taking this into consideration, an ideal-observer analysis revealed that subjects were actually more efficient at discriminating symmetric objects. This reversal in interpretation underscores the importance of ideal-observer analysis. A completely opposite, and wrong, conclusion would have been drawn from analyzing only human discrimination thresholds. Given the same amount of information, the visual system is actually better able to discriminate symmetric objects than asymmetric ones.

OCIS codes: 330.4060, 330.5020, 330.5510, 330.6100.

1. INTRODUCTION

How is the shape of a three-dimensional (3D) object represented in the visual system, and how does its representation in the brain differ from the physical shape in the world? One way of formalizing the question of representation is to ask how the physical shape space is mapped to the perceptual shape space.[1,2] The visually perceived world is not a replica of the physical world but rather reflects ecological and functional requirements. We would not expect such a mapping to be uniform, but rather we expect that certain physical details may be exaggerated and others reduced. We can experimentally address the nature of this mapping by testing whether two similar objects can be discriminated and what determines such perceptual discrimination.

Because of their biological, adaptive, and social significance, objects with bilateral symmetry have been considered an important subclass of objects[3–15] (see Ref. 16 for a review). This raises the possibility that the human visual system is specialized in some way for the processing of bilaterally symmetric objects. If true, such specialization could be manifest in two ways (Fig. 1). First, the visual system may be particularly good at detecting departures from symmetry. In fact, most studies have shown that human observers are sensitive to small deviations from perfect symmetry, with the interpretation that vision has mechanisms specialized for symmetry. On a closer look, however, it appears that this good sensitivity is not appropriately compared.[17–20] That is to say, the evidence so far has shown that a small deviation from perfect symmetry can be detected; but whether the same amount of deviation can be better detected from an asymmetric object is unknown.

This paper addresses the second sense in which symmetry may be special. That is, special in the sense that deformations of a symmetric object may be detected more easily than deformations of an asymmetric object. In other words, if the same amount of deviation is applied to a symmetric object in such a way that the object remains symmetric after the deviation, can this deviation be better detected? For reasons that will become clear below, the question will be addressed in the context of discrimination. Two objects, both deformed from the same 3D symmetric object, one with more deformation than the other, are presented to a human subject. The subject decides which of the two is more similar to the original symmetric object that has been learned. There are two ways to accomplish the deformation. (1) Symmetric condition: The deformation is applied in such a way that the deformed objects remain symmetric. (2) Asymmetric condi-

1084-7529/2003/071331-10$15.00

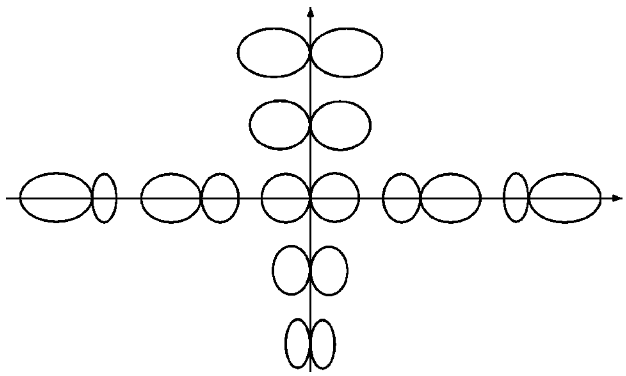

Fig. 1. Schematic illustration of a physical shape subspace, where along one dimension (horizontal) the degrees of asymmetry of an object varies and along an orthogonal dimension (vertical) shape varies while symmetry is unchanged. In this paper we ask whether two shapes are equally discriminable along the two axes, respectively, if their physical difference is kept constant (by an ideal observer).

tion: The two objects become asymmetric after the deformation. The question is which condition is easier to detect?

To answer this question, we must ensure that the deformations in the two conditions are equal or, more precisely, the signal-to-noise ratio of the stimulus in the two conditions are identical. At issue is the correlated nature of the deformation applied to the symmetric condition. For a symmetric object to remain symmetric after deformation, the deformation itself has to be symmetric. In contrast, for a symmetric object to become asymmetric after deformation, the deformation has to be asymmetric. For this reason, we use the ideal observer to equalize the stimulus signal-to-noise ratio in both conditions.[21–24]

Calculating the performance of ideal observers is often theoretically difficult; however, when an appropriate task and object representation are chosen, ideal observer analysis can be made tractable. In Subsection 1.A we will elaborate on the ideal observer approach.

The second question we ask is the following. If perceiving a symmetric object is different from perceiving an asymmetric one, then to what extent does the assumption that an object is symmetric *per se* facilitate ascertaining the 3D structure of the object? We know that a shaded thick wire object, as compared with its silhouette or a thin wire-frame counterpart, offers the following additional information about its 3D shape: shading, self-occlusion that provides relative depth ordering between the occluding cylinder and the occluded, and the shape of the creases between two neighboring cylinders. Our question is if we replace the shaded objects with their silhouettes, does an assumption of bilateral symmetry provide sufficient information to make discriminations of accuracy comparable with the shaded case? In other words, can human subjects achieve similar discrimination performance if the only 3D shape information is the knowledge that the object has bilateral symmetry?

Previously, we have found that human subjects can take advantage of the information from shading, self-occlusion, and the shape of creases between two neighboring cylinders.[25] This is evident from their better discrimination performance for asymmetric thick wire objects than for the thin wire-frame counterparts, which lack the aforementioned shape information. We have also found that when an image of a wire object can be interpreted as projected from either a symmetric or an asymmetric object, human subjects prefer the symmetric interpretation.[26] In other words, when a 3D object can be interpreted as symmetric, it will be. However, it remains an open question whether such preference for a symmetric object interpretation translates into a more precise specification of the symmetric object's shape, just as shading, self-occlusion, and crease shapes do. The last of the three experiments in this paper will address this question.

A. Ideal-Observer Approach

For clarity, we use the same example of symmetry discrimination as above to motivate the ideal-observer approach. To answer under which of the conditions it is easier to discriminate, it is straightforward to set up a psychophysical task to measure the thresholds of a subject's performance in the two conditions. However, the threshold difference cannot tell us whether better performance in one condition is due to richer information in the stimulus or to superior specialization in the brain or to both. More specifically, in condition (1), the fact that both objects remain symmetric implies that stimulus information is redundant. When one half of an object is deformed, the other half has to be deformed accordingly to retain the symmetry. In theory, therefore, one half of the object and the plane of symmetry completely specify the entire object. In condition (2), on the other hand, since the deformation to one half of the object is completely independent of the other half, deformation at all vertices of an object is informative for the discrimination task.

To make a fair comparison, the stimulus information in the two conditions has to be equalized. The first step in achieving this is to quantify the input stimulus information in the following way. Given a specific task, an ideal observer is defined to have unlimited computational capacity and achieve the best possible performance.[21,27,28] The only limitation of the ideal observer's performance is the amount of stimulus input information, so the ideal observer can achieve as best as the input information allows. Consequently, the ideal observer's performance is a measure of the input stimulus information. The more input information, the better it performs.

Now if we give the same task to human subjects, they will also achieve a certain performance level, typically measured by the threshold at a fixed proportion correct. If we combine the thresholds of the ideal observer and human observers appropriately (see Appendix A), we have normalized the stimulus information and obtained a measure called statistical efficiency. It means the proportion of stimulus information that a human observer has effectively used in the given task. A simple example is detecting which of two patterns of a large number of random dots was drawn from a distribution with a higher mean number.[29] The ideal observer could count all the dots, whereas a human subject could not, and the human's performance would be lower than the ideal observer's. We can interpret the result as if the human subject can see only a fraction of the dots but otherwise use them optimally in the decision.

Coming back to the symmetry-discrimination task, we measure the threshold and statistical efficiency in the following way. The 3D symmetric object consists of a chain of four cylinders that have five vertices total. The deformation is done by adding 3D positional Gaussian perturbations at each vertex. In the symmetric condition, vertices 1–3 have independent perturbations, whereas vertices 4 and 5 are perturbed such that the resultant shape remains symmetric. In the asymmetric condition, perturbations at all five vertices are independent of each other. In any one trial, there will be two deformed objects, one drawn from a Gaussian distribution in which the variance is constant, whereas the other is drawn from a distribution with a larger variance that is adjustable in order to find the discrimination threshold. We call the first object the target and the second object the distractor. The experiment is to find out, with a staircase procedure, how much larger the second variance has to be for an observer to be 75% correct, and the resultant standard deviation will be the measure of the threshold. Defining the task in this way ensures the existence of a simple and provably optimal ideal observer in the signal-detection sense, which we will elaborate in the following.

The ideal observer is assumed to have knowledge of the 3D structure of each object, namely, the (x, y, z) coordinates of the five vertices and their ordering. From each input object image, the ideal observer is assumed to know the two-dimensional (2D) (x, y) coordinates of the vertices and their ordering except that it does not know which end of the cylinder chain is the head and which is the tail. The ideal observer also knows the value of the constant variance of the Gaussian added to the target. Consequently, the ideal observer's job is to consider all possible 2D images of the learned 3D object, calculate the probability that an input image is generated from each of the 2D images, and integrate the probability (as a function of viewing angle) over all 2D images to obtain the likelihood of an input image being generated from the learned object. Since there are two input images in any trial, the ideal observer obtains two likelihood measures and selects the one more likely to be more similar to the learned object. We have thus summarized the operations of the ideal observer; interested readers are referred to Liu *et al.*[25] for more technical details. As outlined in Appendix A of the present paper, the statistical efficiency E can be expressed as[23]

$$E = \frac{(\sigma_d{}^I)^2 - \sigma_t{}^2}{(\sigma_d{}^H)^2 - \sigma_t{}^2}, \tag{1}$$

where σ_t is the standard deviation for the target; $\sigma_d{}^I$ is the ideal observer's threshold, namely, the standard deviation for the distractor at 75% correct; and $\sigma_d{}^H$ is the threshold of human subjects.

For clarity, we will specify in more detail the computations of the ideal observer after specifying the experimental task in Subsection 1.B.

B. Rationale of the Experiments

We describe three experiments. Experiment 1 addresses the question of whether it is easier to discriminate one asymmetric object from another or one symmetric object from another. In the asymmetric condition, independent positional perturbations are added to all vertices of a symmetric wire object, so that the resultant object becomes asymmetric. In contrast, in the symmetric condition, independent perturbations are added to only half of an object, whereas the other half are adjusted to retain symmetry. Discrimination thresholds for human and ideal observers are measured, and statistical efficiencies are calculated.

Experiment 2 addresses the question of whether any effect in experiment 1 is due to the difference in the number of independent noise sources. That is to say, in the asymmetric condition in experiment 1, five vertices are added with independent perturbations, whereas in the symmetric condition, only three vertices are. So it might be that any difference in experiment 1 is due to subjects' limited ability in dealing with more noise sources in the asymmetric condition. Therefore, in this experiment, independent noise is added to only three vertices to make a symmetric object asymmetric. The prediction is that if the difference in experiment 1 is only due to this difference of the number of noise sources, then no difference should be found between the new condition and the symmetric condition in experiment 1.

Experiment 3 addresses the putative specialness of bilateral symmetry from a different perspective. We ask whether the information that an object is symmetric (or nearly so, with asymmetric perturbations) can also provide 3D shape information in the following sense. Since there is no depth information when a silhouette image is presented, the 3D pose of the corresponding 3D object, or, in other words, the object's 3D shape, is ambiguous. However, if the object is known or assumed to be symmetric, then the uncertainty of the 3D pose of the object and the object's 3D shape is reduced. In this experiment we compare symmetric wire objects that are either silhouettes or shaded. Our hypothesis is that if human subjects' performance does not deteriorate after we remove from a symmetric object shape information, such as shading, self-occlusion, and the shape of the creases between two neighboring cylinders, it will be evidence that the knowledge of symmetry alone is sufficient to reduce the uncertainty of the object's pose and shape, just as the removed information did for asymmetric objects.

2. EXPERIMENT 1

A. Stimuli

Each symmetric object consisted of five spheres linked by four cylinders in a chain in three dimensions. The diameter of each sphere and that of each cylinder were identical (0.22 cm). The length of each cylinder was 2.54 cm. Figure 2 shows one example object from multiple views. The objects were rendered with Lambertian shading, with a point light source at infinity. The illumination was directed toward the camera from the upper front, with an angle of 26.6° relative to the vertical direction. Ambient light with an intensity of 3% relative to the directed light was also present.

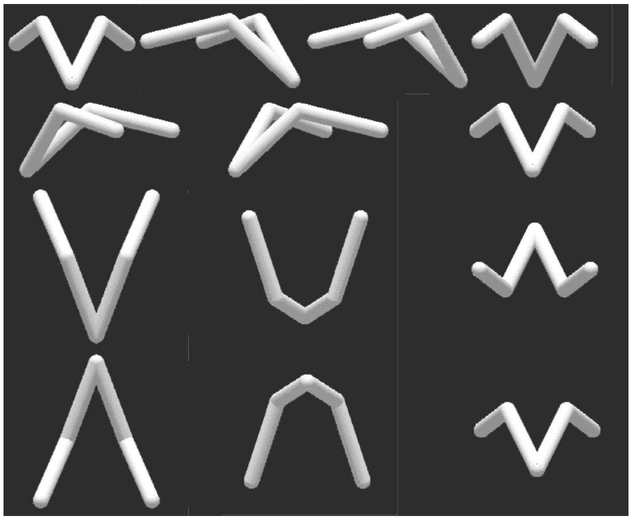

Fig. 2. Three-dimensional symmetric object from multiple views by rotation around the y axis (top two rows, left to right) and then around the x axis (bottom two rows). These 11 views (the top left, second row right, and bottom right are identical) are called the learned views.

(a)

(b)

Fig. 3. Example stimuli in the test stage when objects were distorted versions of the learned symmetric object. (a) Both objects were symmetric after the distortion. (b) Both were asymmetric.

The images were rendered with parallel projection on a Silicon Graphics computer with the Open Inventor software (Silicon Graphics, Inc.). The viewing distance was 57 cm. An object was approximately 5 cm in size, which was approximately 5° visual angle. Subjects viewed the stimuli binocularly in a darkened room.

B. Method

Each subject was tested with six symmetric objects; three became asymmetric after structural distortion was added, and the other three remained symmetric after structural distortion. Each object was tested in a single block. The order of the test and which objects were symmetric and which asymmetric were counterbalanced among the subjects. Within each block the experimental procedure was as follows.

1. Learning

A symmetric image of the object was shown with $x = 0$ being the symmetric plane. The subject rotated the object by pressing the middle button of the computer mouse. Each rotational step was 60°. The object was rotated first around the y axis and then around the x axis (when all the images were symmetric). Hence the object was shown from 11 viewing angles. Figure 2 shows the sequence of these images for one example object. A subject looked through this image sequence twice.

2. Practice

Two images were presented side by side to the subject, who decided which was the learned object. Feedback was provided by a computer beep for correct responses. One image was chosen randomly from one of the 11 learned views. The other image was generated either from a different symmetric object or from the same learned object by adding distortions. The distortion was created by adding Gaussian positional noise $N(0, 0.254 \text{ cm})$ in three dimensions to one symmetric half of the object; the other half of the object was distorted in such a way that symmetry was preserved.

At any time during the practice, the subject could press the middle mouse button to review the learned object. The image of the learned object would change into the first image of the object and rotate around the y axis and then around the x axis under the subject's mouse button control. Meanwhile, the image of the distractor object remained unchanged.

The subject needed 100 trials to finish this practice, in addition to the review trials. This practice was the same no matter if the object would be distorted symmetrically or asymmetrically during the next stage, the test stage, so that the subject was trained to learn both the symmetry and the geometric details of the object. Figure 3 shows an example of the object when the subject was tested under the symmetric and asymmetric conditions, respectively.

3. Test

Two object images were presented side by side to the subject, who decided which was more similar to the learned object. No feedback was provided. Both objects were distorted versions of the same learned symmetric object and rendered from the same viewing angle. Both objects were generated by adding Gaussian positional distortions in three dimensions to the vertices. For one object, the standard deviation was always 0.254 cm. For the other, the standard deviation was larger and could be adjusted to find a threshold. We used a staircase procedure to find the threshold value of the larger standard deviation that gave rise to 75% correct performance.[30] At any time during the test, the subject could press the middle mouse button and review the learned object from the 11 viewpoints. Discounting these review trials, there were 500 test trials.

The total number of 500 test trials was divided into three conditions, which were randomly interleaved. In 200 of the trials, the viewing position was from one of the 11 learned positions. We call this the learned view condition. In another 200 trials, the viewing position was randomly chosen from the viewing sphere, with a random rotation around the viewing direction. We call this the novel view condition. The remaining 100 trials served as refresh trials that were identical to those in the practice stage, and feedback was provided.

C. Subjects

Twelve subjects, naïve to the purpose of the experiment, participated; each was with six prototype objects, three for the symmetric condition and three for the asymmetric one, in a counterbalanced manner.

D. Results

1. *Thresholds*

Within-subjects analysis of variance (ANOVA) revealed a significant main effect of learned versus novel viewing angles [$F(1, 11) = 7.92, p < 0.02$]. As expected, the threshold for the learned views was smaller than for the novel views (0.71 cm versus 0.84 cm). ANOVA also revealed a significant main effect of symmetric versus asymmetric conditions [$F(1, 11) = 4.72, p = 0.05$]. The threshold for the asymmetric condition was smaller than for the symmetric condition (0.71 cm versus 0.85 cm); this means that the performance was better for the asymmetric than for the symmetric objects. The interaction was not significant ($F < 1$). Figure 4 shows the threshold performance when the subjects were correct 75% of the time.

2. *Ideal Observer Computation*

We start by describing the computation for the asymmetric condition (see Liu *et al.*[25] for more details). The question is, given two input images [i.e., for each image, the (x, y) coordinates of each of the five vertices and their order] of two deformed objects, which is more likely to come from the learned object? The 3D structure of the learned object [i.e., the (x, y, z) coordinates of each of the five vertices and their order] is known. Unknown are the projection angle from which both images are obtained and which end vertex is the head and which is the tail.

The ideal observer tries out all possible projection angles and, for each angle, tries all possible angles of rotation (in 360°). For each particular projection angle and rotation angle, a Euclidean distance between vertices of an input image and the newly projected image is computed [the origin (0, 0, 0) is assumed known to the ideal observer]. The likelihood is then computed by using the Gaussian with the known variance, which is the constant variance of the target object. Because of the head–tail ambiguity, the same computation is repeated with a flipped head–tail correspondence between the two images above. The likelihood values for all the projection angles and rotation angles are integrated with equal weighting. Finally, the input image with the larger total likelihood will be chosen as the one more similar to the learned object.

Numerically, 32,776 points on the viewing sphere were chosen in such a way that their distribution was maximally uniform on the sphere (the enclosed convex hull had the maximum volume). Each point represented a projection angle. For each projection angle, 360 rotational angles were chosen (1° each).

For the symmetric condition, the computation was similar except for the following two differences. The first was that, since independent noise was added only to three vertices on one side of an object, the Euclidean distance was calculated between two images accordingly, as well (because of the head–tail ambiguity, the head–tail flip was still needed). The second difference was that, since an object after deformation remained symmetric, the information regarding the plane of symmetry was not completely lost, so the ideal observer used the information as follows.

For any bilaterally symmetric object, a line connecting any two corresponding points is perpendicular to the plane of symmetry. Therefore all such lines are parallel to each other. In our case, the lines connecting vertices 1 and 5 and vertices 2 and 4 were parallel to each other. Since the object remained symmetric after deformation, these two lines remained parallel to each other as well. Now, since the 3D symmetric object was projected to its 2D image orthographically, these two lines were also parallel to each other. Therefore, although the projection angle was still not completely known, we do know that the rotational angle around the projection direction had to be such that the two new lines had to be parallel with the old lines. More specifically, there were only two such rotational angles (differing by 180°). Again, the head–tail ambiguity remained.

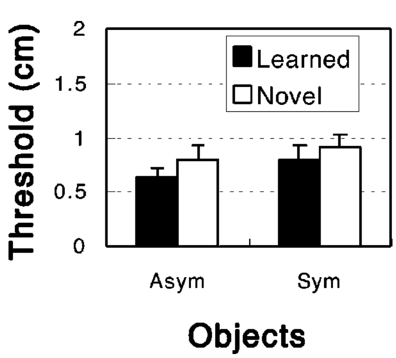

Fig. 4. Discrimination threshold of 12 subjects when they were correct 75% of the time. The error bars are standard errors between subjects. Note that these error bars cannot directly reflect the statistical significance, since within-subjects analysis is used (and the same for Figs. 5–7 and 9).

3. *Statistical Efficiencies*

We calculated the statistical efficiencies for the symmetric and asymmetric conditions separately. For the symmetric condition, we simulated 20 objects, each with the learned view condition and novel view condition, and each condition was simulated with the staircase procedure for 2000 trials. Because the ideal observer had knowledge of the 3D structure of each object, its performance for the learned view and the novel view conditions should be the same. We simulated both conditions to self-check that the thresholds of both conditions for each object were identical, and they indeed were. We obtained an average threshold of 0.379 cm with a standard deviation of 0.013 cm.

For the asymmetric condition, we needed to simulate only nine objects with just the novel condition to obtain an average threshold with a much smaller standard deviation than the symmetric ones: 0.321 ± 0.007 cm.

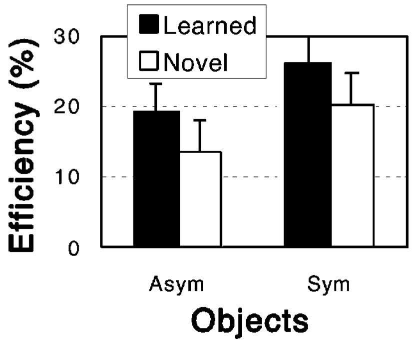

Fig. 5. Statistical efficiencies for the symmetric versus asymmetric conditions and learned versus novel views. Note that the pattern of symmetric versus asymmetric comparison is completely reversed here in contrast to the threshold performance.

We then calculated the statistical efficiencies using Eq. (1). ANOVA demonstrated, as expected, a better performance for the learned view than for the novel view [22.50% versus 16.90%, $F(1, 11) = 12.82$, $p < 0.01$]. The ANOVA also indicated better performance for the symmetric condition than for the asymmetric condition [23.04% versus 16.36%, $F(1, 11) = 15.00$, $p < 0.01$], which was unexpected from the discrimination-threshold results. The interaction was not significant ($F < 1$). Figure 5 shows the result.

3. EXPERIMENT 2

There was a potential confound in experiment 1. Namely, subjects' threshold difference between the symmetric and the asymmetric conditions might have had nothing to do with symmetry but rather might be due to the different number of noise sources. That is, there were only three independent noise sources for the symmetric condition and five for the asymmetric condition.

To test this possibility, we conducted experiment 2 under the same conditions as experiment 1, except that independent noise was added to only three vertices, 1, 3, and 5, and no noise was added at vertices 2 and 4. The resultant object therefore would be asymmetric.

If the result in experiment 1 is completely due to the difference of the number of noise sources, then we expect no difference between the results of experiment 2 and the symmetric condition of experiment 1. On the other hand, if symmetry versus asymmetry has a genuine effect, then we expect a difference in thresholds and statistical efficiencies between this experiment and the symmetric condition in experiment 1.

A. Subjects

Twelve fresh subjects, naïve to the purpose of the experiment, participated, each with three prototype objects.

B. Results

1. Thresholds

Two-way ANOVA was used to compare the thresholds from this experiment and those of the symmetric condition in the last experiment, with the main factors of view (learned versus novel) and condition (asymmetric versus symmetric). As expected, the main effect of view was significant; it was easier to discriminate the learned views than the novel views [0.84 cm versus 1.13 cm, $F(1, 22) = 14.01$, $p < 0.001$]. The main effect of condition was not significant [symmetric: 0.85 cm; asymmetric: 1.12 cm, $F(1, 22) = 1.17$, $p = 0.29$]. Of particular interest was the interaction, which was significant, meaning that discrimination dropped more quickly from the learned to novel views for the asymmetric condition in the current experiment (0.87 to 1.36 cm) than for the symmetric condition (0.80 to 0.90 cm) [$F(1, 22) = 5.96$, $p < 0.02$, Fig. 6]. This means that, although noise had been added to the same number of vertices, discrimination is different for symmetric versus asymmetric objects.

2. Ideal Observer Computation

We describe here the ideal observer for the task in experiment 2. Since no noise was added to vertices 2 and 4 and since the origin (0, 0, 0) was always known, the 2D coordinates of vertices 2 and 4 (four knowns) completely determine the projection angle (3 unknowns), except that we still have to consider the head–tail ambiguity. We used 20 objects and each object again with 2000 trials. The average threshold and standard deviation were 0.339 ± 0.004 cm.

3. Statistical Efficiencies

We computed the statistical efficiencies using Eq. (1) for the learned and novel views. We then performed a two-way between-subjects ANOVA, comparing efficiencies in

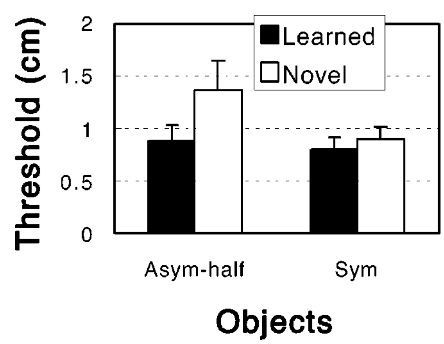

Fig. 6. Discrimination thresholds for the learned and novel views of the symmetric condition in experiment 1 and of the asymmetric condition in experiment 2.

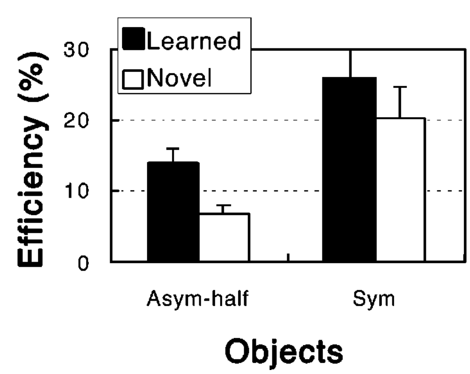

Fig. 7. Statistical efficiencies of the symmetric condition in experiment 1 and of the asymmetric condition in experiment 2 for the learned and novel views.

this experiment and those in experiment 1 of the symmetric condition. Both main effects were significant. Efficiencies for the learned views were higher than for the novel ones (19.95% versus 13.50%), $F(1, 22) = 12.93$, $p < 0.002$. Efficiencies for the symmetric objects were higher than for asymmetric ones (23.04% versus 10.41%), $F(1, 22) = 9.21$, $p < 0.006$. The interaction was not significant [$F(1, 22) < 1$]. Figure 7 shows the result.

4. EXPERIMENT 3

So far, the computations of the ideal observer have assumed that only 2D information, namely, the (x, y) coordinates of an object's vertices, is available from an input image. However, shading, self-occlusion, and the shape of creases formed at the joints between neighboring cylinders may provide potential 3D information about the length and depth orientation of each of an object's cylinders.[25] Although the computations of the ideal observer, so far, are all valid, since no bias is introduced in the comparisons, we would still like to know to what extent such 3D information is used by human subjects.

To address this question, we changed the objects into silhouettes so that an object had the same luminance everywhere and all three potential sources of 3D information were absent (Fig. 8). We predict that if subjects' performance became worse, then it would be evidence that the 3D information was used. Otherwise, it means that these sources of 3D information play no significant role when the prototype objects are symmetric in three dimensions.

A. Subjects

Four subjects from experiment 1 plus eight fresh subjects participated; each was with six prototype objects, three in the symmetric condition and three in the asymmetric, in a counterbalanced manner.

B. Results

We compared the results with those in experiment 1 in a three-way ANOVA (silhouette versus shaded, learned view versus novel view, and symmetric versus asymmetric). We first report here the within-subjects analysis with the four subjects. Only two main effects were significant, both confirming the results from experiment 1 [Fig. 9(a)]. The thresholds for the symmetric condition were higher than for the asymmetric condition (0.765 cm versus 0.667 cm): $F(1, 3) = 15.30$, $p < 0.03$. Thresholds for the learned views were lower than for the novel views (0.622 cm versus 0.810 cm): $F(1, 3) = 13.03$, $p = 0.037$. Importantly, the difference between the silhouette and the shaded conditions was not significant (0.767 cm versus 0.665 cm): $F(1, 3) = 2.53$, $p = 0.21$.

We now look at the between-subjects ANOVA between the eight subjects in the current experiment and the remaining eight subjects in experiment 1. The results were consistent with the within-subjects analysis above [Fig. 9(b)]. Perhaps owing to between-subjects variations, the thresholds for the symmetric condition were now only marginally significantly worse than for the asymmetric condition (0.926 cm versus 0.784 cm): $F(1, 14) = 3.16$, $p = 0.097$. Thresholds for the learned

Fig. 8. Experimental stimuli in the test phase when all the objects were silhouettes.

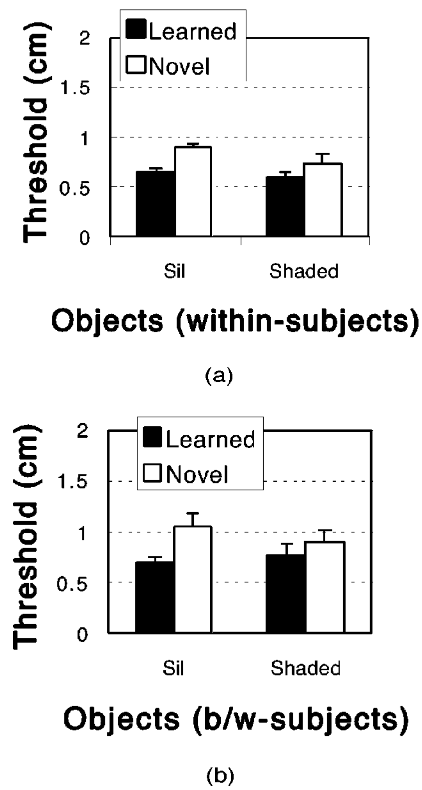

Fig. 9. Threshold comparisons between experiment 1 and the silhouette condition of experiment 3. (a) Within-subjects comparison (four subjects). (b) Between-subjects comparison (eight subjects each). There was no significant difference between the silhouette and the shaded conditions.

views were better than for the novel views (0.735 cm versus 0.975 cm): $F(1, 14) = 11.24$, $p < 0.005$. Thresholds between the silhouette and the shaded conditions were not significantly different (mean $\pm$ standard error: 0.875 ± 0.075 cm versus 0.835 ± 0.080 cm): $F(1, 14) = 0.05 < 1$, $p = 0.828$.

5. DISCUSSION

We have demonstrated that it is harder, in terms of threshold, to discriminate between two deformed objects when the deformation retains the symmetry. However, when we take into consideration that a symmetric deformation has only half of the degrees of freedom as an asymmetric deformation, we found that human observers are better able to discriminate symmetric deformations when the input image information is equalized. We stress that without the ideal-observer analysis such a conclusion cannot be reached. This suggests that the visual system is indeed better equipped to process symmetric objects.[3,16,20,31–33]

In addition, these results suggest that symmetry within a 2D image does not account for all of the advantages for symmetric objects. Otherwise, we should not expect any difference for the novel views between the symmetric and the asymmetric conditions, since a novel view was almost always an asymmetric image. Within-image symmetry, on the other hand, does suggest an advantage. This is because, for the learned views, the symmetric condition is more efficient than the asymmetric condition. Note that in half of the trials of the symmetric condition, the images were symmetric. Whereas for the asymmetric condition, practically no image was symmetric. We remark that, within the learned views of the symmetric condition, we cannot directly compare the symmetric images with asymmetric ones because we used a single staircase procedure to obtain a single threshold. Therefore we cannot rule out the possibility that this advantage is completely due to the 3D symmetry and has nothing to do with the 2D symmetry in an image.

Although our results show that object discrimination is more efficient for symmetric than for asymmetric objects, we emphasize that this does not imply distinct physiological mechanisms for symmetric versus asymmetric objects. In Bayesian terms, symmetric objects may be handled more efficiently as a consequence of being more common or more important. Yet, somehow priors and task relevance have to be instantiated in the brain. This kind of specialization may not be completely based on familiarity with many near-symmetric objects in the world, natural and artificial alike, because the objects used in the current study are novel. High efficiencies may have to do with the way the visual system codes for near-symmetric objects, where symmetry is a type of regularity that the visual system exploits to form economic object representations.[12,34]

The results in the third experiment indicated that, for symmetric prototype objects, shading, self-occlusion, and the shape of creases between two neighboring cylinders do not provide more information about the 3D shape of an object than the silhouette of the object does already. This is, in fact, consistent with subjects' impressions that a silhouette of a symmetric object already appears to be a 3D symmetric object. This is in stark contrast with the previous finding in Liu *et al.*[25] In that study, shaded (with self-occlusion and creases at the junctions), but asymmetric, objects gave rise to better (i.e., smaller) thresholds than objects (called tinker toys) of the same 3D configurations but with all the cylinders shrunk to a line, so that no shading, self-occlusion, and creases were available. Moreover, in that study the statistical efficiency for the shaded wire objects was also higher than for the tinker toys when the ideal observer, for the former, was assumed to be able to recover the full 3D structure of an object [the (x, y, z) coordinates of the vertices] from its single 2D image and, for the latter, the ideal observer had only the 2D (x, y) coordinates of the vertices. It was a strong result because the ideal observer for the shaded objects was a super-ideal observer, who was able to recover a 3D shape from shading from a single image. Such a super-ideal observer gave rise to the lower bound of statistical efficiency for the shaded wire objects, which was still higher than for the tinker toys. The fact that no statistically significant difference has been found here between the shaded and the silhouette symmetric objects is, we believe, a strong testament that the knowledge of symmetry alone provides information about the 3D shape of the objects.

So how may the knowledge of 3D symmetry be obtained from a 2D image? The following theoretical result may offer a clue. A necessary condition for a 2D image to be the parallel projection of a 3D symmetric object is this: All lines, each of which connects two corresponding points of a 3D symmetric object, have to be parallel in the image plane.[35] By extension, in a perspective projection, all these lines have to converge to a single point. (The parallel lines in the orthographic projection case are, of course, a special case, since they converge at infinity.)

There is also a theoretical possibility that our result of the symmetry advantage may be explained, at least in part, by the following. When symmetry is detected, it may produce a compelling percept to the extent that symmetry makes other differences less distinguishable.[19,20,36,37] We cannot rule out or confirm this possibility but speculate that, unless the image itself is symmetric, the symmetry percept is not as strong.

Human efficiencies are less than 100%. What are the possible sources of the inefficiencies? Specifically, such inefficiencies can be attributed to three categories: input stimulus encoding, representation of an object model, and the matching process (please see Section 4.1 in Ref. 25 and Refs. 23 and 38 for a discussion in a broader context). First, unlike the ideal observer who codes the (x, y) positions of the vertices precisely, humans are bound to be inaccurate, encoding the shape of an input stimulus, e.g., either by the vertex positions or by the lengths of the cylinders and their relative angles. We remark that assuming that the ideal observer always knows the origin of the coordinate system, which humans may not, results in a lower efficiency, since the uncertainty for the ideal observer is reduced. Second, even if humans can represent an input stimulus perfectly, they may not be able to represent the 3D structure of the object precisely, in contrast to the ideal observer. By precisely representing the 3D structure, we mean that the information contained in the representation can accurately describe the 3D shape of the object, no matter what the format of this representation is. Third, even if the human visual system can encode an input stimulus and represent the 3D structure of the object perfectly, it may still be imperfect in matching an internal object model with an input stimulus. For instance, humans may not be able to consider and integrate over all possible relative orientations of the model relative to the input. The visual system may not be able to effectively rotate the 3D model with a large angle to match with a 2D input.

APPENDIX A

We now outline the assumptions used in deriving the statistical efficiency, while leaving the derivation in Ref. 23. Statistical efficiency E is defined as the squared ratio of discrimination index d' for the human observer d'^H versus for the ideal observer d'^I, where $d'^I = x/\sigma$, with x as the strength of the signal and σ as the standard deviation of the Gaussian noise present,

$$E = \left(\frac{d'^H}{d'^I}\right)^2. \tag{A1}$$

As in the standard analysis, we assume that the human observer is ideal except that there is Gaussian internal noise σ_H added to the input signal and noise. Therefore $d'^H = x/\sqrt{\sigma^2 + \sigma_H{}^2}$. We now need another mathematical result that we proved in Ref. 23 in order to complete the derivation. This result says that the percentage correct for the ideal observer is determined by the ratio of standard deviations of target and of distractor noise, σ_t and σ_d. In our experimental setting, $P_{\text{thres}} = f(\sigma_t/\sigma_d)$. [In fact, $P_{\text{thres}}(\sigma_t/\sigma_d, n)$ is also a function of n, the number of object vertices with independent noise. This is relevant to the issue of different number of noise sources that is discussed in Subsection 1.B.] We assumed, in deriving this result, that an ideal observer knows about a prototype object in terms of a vector corresponding to the object's vertices. It also knows the vectors of an input target object [prototype plus noise $N(0, \sigma_t)$] and of an input distractor object [prototype plus noise $N(0, \sigma_d)$]. This ideal observer also knows the relative position and orientation between the prototype object and an input object. Note that this is an approximation to the ideal observer used in this paper. The ideal observer in this paper does not know the relative orientation of the prototype object relative to an input object. It instead needs to integrate over all possible orientation angles. Here the optimal strategy is to choose between the two input vectors the one that is closer in Euclidean distance to the prototype vector. The input target and distractor each have a probability distribution of this Euclidean distance, from which the percentage correct for the ideal observer can be derived as $P_{\text{thres}} = f(\sigma_t/\sigma_d)$. Having obtained this relationship, we then know that when the human and ideal observers have the same percentage correct,

$$\frac{\sigma_t{}^2}{(\sigma_d{}^I)^2} = \frac{\sigma_t{}^2 + \sigma_H{}^2}{(\sigma_d{}^H)^2 + \sigma_H{}^2}, \tag{A2}$$

which leads to Eq. (1).

ACKNOWLEDGMENTS

We thank the two reviewers for helpful comments. This research was previously reported at the European Conference on Visual Perception, Trieste, Italy, 1999. Z. Liu is supported in part by a National Science Foundation grant IBN-9817979 and a National Eye Institute grant R03 EY14113. D. Kersten is supported by National Institutes of Health grant EY02857. We thank Sara L. van Driest, Chris Wainman, and Ingrid Wu for their assistance in conducting the experiments.

Z. Liu can be reached by fax, 310-206-5895, or e-mail, zili@psych.ucla.edu. D. Kersten can be reached by fax, 612-626-2079, or e-mail, kersten@umn.edu.

REFERENCES

1. J. J. Koenderink, A. J. van Doorn, and A. M. L. Kappers, "Depth relief," Perception **24**, 115–126 (1995).
2. J. J. Koenderink, A. J. van Doorn, and A. M. L. Kappers, "Pictorial surface attitude and local depth comparisons," Percept. Psychophys. **58**, 163–173 (1996).
3. M. Leyton, *Symmetry, Causality, Mind* (MIT, Cambridge, Mass., 1992).
4. H. I. Bailit, P. L. Workman, J. D. Niswander, and C. J. MacLean, "Dental asymmetry as an indicator of genetic and environmental conditions in human populations," Hum. Biol. **42**, 626–638 (1970).
5. P. A. Parsons, "Fluctuating asymmetry: an epigenetic measure of stress," Biol. Rev. **65**, 131–145 (1990).
6. A. P. Møller, "Female swallow preference for symmetrical male sexual ornaments," Nature **357**, 238–240 (1992).
7. R. Thornhill, "Fluctuating asymmetry and the mating system of the Japanese scorpionfly *Panorpa japonica*," Anim. Behav. **44**, 867–879 (1992).
8. A. L. R. Thomas, "On avian asymmetry: the evidence of natural selection for symmetrical tails and wings in birds," Proc. R. Soc. London Ser. B **252**, 245–251 (1993).
9. A. L. R. Thomas, "The aerodynamic costs of asymmetry in the wings and tails of birds: asymmetric birds can't fly round tight corners," Proc. R. Soc. London Ser. B **254**, 181–189 (1993).
10. K. Grammer and R. Thornhill, "Human (*Homo sapiens*) facial attractiveness and sexual selection: the role of symmetry and averageness," J. Comp. Psychol. **108**, 233–242 (1994).
11. P. J. Watson and R. Thornhill, "Fluctuating asymmetry and sexual selection," Trends Ecol. Evol. **9**, 21–25 (1994).
12. M. Enquist and A. Arak, "Symmetry, beauty and evolution," Nature **372**, 169–172 (1994).
13. R. Thornhill and S. W. Gangestad, "Human fluctuating asymmetry and sexual behavior," Psych. Sci. **5**, 297–302 (1994).
14. G. Rhodes, F. Profitt, J. M. Grady, and A. Sumich, "Facial symmetry and the perception of beauty," Psychon. Bull. Rev. **5**, 150–163 (1998).
15. A. P. Møller and R. Thornhill, "Bilateral symmetry and sexual selection: a metaanalysis," Am. Nat. **151**, 174–192 (1998).
16. C. W. Tyler, ed., *Human Symmetry Perception and Its Computational Analysis* [VSP (VNU Science Press) BV, Utrecht, The Netherlands, 1996].
17. D. P. Carmody, C. F. Nodine, and P. J. Locher, "Global detection of symmetry," Percept. Mot. Skills **45**, 239–249 (1977).
18. B. Jenkins, "Component processes in the perception of bilaterally symmetric dot textures," J. Exp. Psychol. **9**, 258–269 (1983).
19. B. S. Tjan and Z. Liu, "Symmetry discrimination of faces," in *ARVO Abstracts. Investigative Ophthalmology and Visual Science*, Vol. 39, Suppl. 4 (Association for Research in Vision and Ophthalmology, Fort Lauderdale, Fla., 1998), p. s170.
20. Z. Liu and B. S. Tjan, "Near-bilateral symmetry impedes symmetry discrimination," in *Proceedings of the European Conference on Visual Perception*, Vol. 27 (Suppl.) (Pion Ltd., London, 1998), p. 6.
21. D. C. Knill and D. Kersten, "Ideal perceptual observers for computation, psychophysics and neural networks," in *Pattern Recognition by Man and Machine*, R. Watt, ed., Vol. 14 of *Vision and Visual Dysfunction*, J. Cronly-Dillon, gen. ed. (MacMillan, London, 1991), Chap. 7.
22. R. A. Eagle and A. Blake, "Two-dimensional constraints on three-dimensional structure from motion tasks," Vision Res. **35**, 2927–2941 (1995).
23. Z. Liu, D. C. Knill, and D. Kersten, "Object classification for human and ideal observers," Vision Res. **35**, 549–568 (1995).
24. B. S. Tjan, W. L. Braje, G. E. Legge, and D. Kersten, "Human efficiency for recognizing 3-D objects in luminance noise," Vision Res. **35**, 3053–3070 (1995).

25. Z. Liu, D. Kersten, and D. C. Knill, “Dissociating stimulus information from internal representation—a case study in object recognition,” Vision Res. **39**, 603–612 (1999).
26. W. Knecht, Z. Liu, and D. Kersten, “Symmetry detection for 2D projections of 3D wire objects,” in *ARVO Abstracts: Investigative Ophthalmology*, Vol. 34, Suppl. 4 (Association for Research in Vision and Ophthalmology, Rockville, Md., 1993), p. s1867.
27. D. M. Green and J. A. Swets, *Signal Detection Theory and Psychophysics* (Krieger, Huntington, N.Y., 1974).
28. A. Burgess and H. B. Barlow, “The precision of numerosity discrimination in arrays of random dots,” Vision Res. **23**, 811–820 (1983).
29. H. B. Barlow, “The efficiency of detecting changes of density in random dot patterns,” Vision Res. **18**, 637–650 (1978).
30. A. B. Watson and D. G. Pelli, “QUEST: a Bayesian adaptive psychometric method,” Percept. Psychophys. **33**, 113–120 (1983).
31. H. B. Barlow and B. C. Reeves, “The versatility and absolute efficiency of detecting mirror symmetry in random dot displays,” Vision Res. **19**, 783–793 (1979).
32. T. Vetter, T. Poggio, and H. H. Bülthoff, “The importance of symmetry and virtual views in three-dimensional object recognition,” Curr. Biol. **4**, 18–23 (1994).
33. P. Wenderoth, “The effects on bilateral-symmetry detection of multiple symmetry, near symmetry, and axis orientation,” Perception **26**, 891–904 (1997).
34. R. A. Johnstone, “Female preference for symmetrical males as a by-product of selection for mate recognition,” Nature **372**, 172–175 (1994).
35. T. Poggio and T. Vetter, “Recognition and structure from one 2D model view: observations on prototypes, objects classes and symmetries,”Tech. Rep., Artificial Intelligence Memo 1347 and Center for Biological Information Processing Paper 69 (MIT, Cambridge, Mass., 1992).
36. P. Verghese and S. L. Stone, “Perceived visual speed constrained by image segmentation,” Nature **381**, 161–163 (1996).
37. I. Bülthoff, H. H. Bülthoff, and P. Sinha, “Top-down influences on stereoscopic depth-perception,” Nat. Neurosci. **1**, 254–257 (1998).
38. Z. Liu and D. Kersten, “2D observers for human 3D object recognition?” Special issue on Models of Recognition, Vision Res. **38**, 2507–2519 (1998).