Discriminating the volume
of motion-defined solids

H. A. H. C. ᴠᴀɴ VEEN, A. M. L. KAPPERS, J. J. KOENDERINK, and P. WERKHOVEN
University of Utrecht, Utrecht, The Netherlands

We investigated the ability of human observers to discriminate an important global 3-D structural property, namely volume, of motion-defined objects. We used convex transparent wire-frame objects consisting of about 12 planar triangular facets. Two objects, vertically separated by 7°, were shown simultaneously on a computer display. Both revolved at 67°/sec around a common vertical axis through their centers of mass. Observers watched the objects monocularly for an average of three full rotations before they responded. We measured volume discrimination as a function of absolute volume (3–48 cm³; 1 m viewing distance) and shape (cubes, rods, and slabs of different regularity). We found that (1) volume discrimination performance can be described by Weber's law, (2) Weber fractions depend strongly on the particular combination of shapes used (regular shapes, especially cubes, are easiest to compare, and similar shapes are easier to compare than different shapes), and (3) humans use a representation of volume that is more veridical and stable in the sense of repeatability than a strategy based on the average visible (2-D) area would yield.

The volume of a solid object is one of its most fundamental structural features. It determines the space the object occupies and can give some indication of its mass. We use our visual sense continuously to make judgments of volume (and mass): Which piece of meat is larger? Do I have as much lemonade as she has? Can I lift this object? Will this pile of clothes fit into that suitcase? And so on. How do we go about making these judgments? It is highly probable that we use all sorts of experience and prior knowledge. This will not be very helpful, however, when we are dealing with unfamiliar objects. Observing the objects from different viewpoints will help us to gain a better 3-D impression. This paper reports on the human ability to discriminate the volumes of various objects when these objects are rotating in front of the observer. The rotation lets the observer see all sides of the objects and make use of structure-from-motion *computations* to determine the 3-D structure of the objects. We varied the size and shape of the objects and measured the veridicality and reproducibility of volume comparisons.

There has been some previous research in this area, but it has mainly involved the use of stationary, real objects. Lauer (1929) asked subjects to compare spheres, cubes, prisms, spherical sectors, and spherical segments, and found that shape does affect estimates of volume (e.g., the volumes of cubes were overestimated relative to spheres). He also concluded that disparity of dimensions within an object was critical, because the volumes of elongated prisms were overestimated relative to cube-like prisms, and the volumes of flat spherical segments were underestimated relative to higher segments. Brunswik (1934, 1956) mentioned an unpublished study by Stevenson on a similar theme. In this study prisms of various elongations were compared with cubes of various sizes. Subjects were given a prism and then asked to select from a set of cubes the one with the same volume as the prism. The volumes of long prisms (sticks) were slightly overestimated relative to cubes, whereas the volumes of thin prisms (plates) were underestimated relative to the volumes of cubes. Brunswik also noted that cube–cube comparisons showed much less variability (low threshold) than comparisons of different shapes.[1] In 1941 Piaget published his famous experiments on the conservation of continuous quantities, describing how young children seemed to base their judgments of the amount of orangeade in a glass on the height of the liquid column, ignoring the width of the column and thus the actual volume of liquid. In the 1960s, the perception of volume was investigated with paradigms involving ratio and magnitude estimation tasks. The data from these experiments are usually analyzed in terms of power functions: Judged volume is assumed to be proportional to a certain power of physical volume. Baird (1970) summarized these investigations, concluding that in most cases the exponent of the power function for volume judgments appears to be smaller than 1, typically 0.7, which is smaller than the values found for area and length judgments. He did not give any information about the response variance. Stanek (1968, 1969) published two short papers on surface and volume judgments. He asked subjects to match the volume of prisms and cubes and found both underestimations and overestimations of prism volume, depending on the specific elongation and size of the prism. Some connections

between estimates of volume and mass have also been investigated. The existence of the so-called size–weight illusion has been known for more than a century now (Charpentier, 1891): The perceived weight of an object depends not only on its physical weight, but also on its size. Recently, Ellis and Lederman (1993) published an article in this journal about the role of haptic versus visual volume cues in the size–weight illusion.

There are two main differences between our work and previously published studies. First, we investigate volume judgments of *synthetic* (computer graphics) objects that are *rotating* relative to the observer. This approach features structure-from-motion as the main depth cue for estimating volume, while eliminating other (uncontrolled) depth cues that are usually available when real objects are used. Previous research has typically employed fairly natural viewing conditions in which the relative motion of object and observer has played an inferior role. Second, we measured the *variance of responses*, or discrimination thresholds. We have been unable to find any systematic study on volume discrimination thresholds in the existing literature. Our research resembles previous work in that it tries to establish the influence of shape on the judgments of volume, albeit that we vary different aspects of shape.

The paper starts with a description of two psychophysical experiments we did with human subjects. We measured veridicality and variance of relative volume judgments as a function of overall scale in the first experiment, and as a function of shape in the second experiment. Then we consider the possible use of projected area in estimating volume. We present the results of an algorithm that bases its volume estimates on the correlation between projected area and volume, and discuss the implications for the interpretation of the psychophysical data. Finally, in the general discussion, we compare our results with those reported in the literature, and discuss what we have learned about the way in which our visual system represents objects like the ones we use.

METHOD

Experimental Setup

The stimuli were generated on an Apple Macintosh IIfx computer and displayed on a 71 Hz Radius TPD/19 high spatial resolution (82 dpi) gray-scale monitor ($1,152 \times 882$ pixels). The subjects were seated in front of the screen at a distance of 1 m. Head motions were not explicitly restricted, but the subjects were required to keep their right eye just in front of the center of the screen. The left eye was covered with a black patch. The room was dimly lit so that the subjects were still able to see the monitor and the table on which it stood.

Stimulus

Temporal and spatial parameters. Two objects, vertically separated by 7º (center–center), were shown simultaneously on a computer display. Object dimensions were controlled parameters of the experiment and resulted in projected sizes of between roughly 0.2º and 8º of visual angle; the projected sizes depended on shape, volume, and instantaneous orientation of the object with respect to the observer. Both objects revolved at a fixed rate of 67º/sec about a common vertical axis through their centers of mass. Every 75 msec a new view (parallel projection) of the objects was shown on the screen.

The objects rotated back and forth in depth with an amplitude of 180º. Thus, information about the period of rotation was available to the subjects. Subjects had to watch each display for at least one full rotation (360º in 5.4 sec). There was no upper limit to the presentation time. We did not prescribe fixation.

It should be noted here that, because we used orthographic projections of wire-frame objects, the direction of rotation was mathematically ambiguous at every instant (see Ullman, 1979). This effect was confirmed by the subjects; they experienced spontaneous reversals as well as opposite rotation directions for the two objects. Furthermore, the orthographic projection itself is not sufficient to define the position in depth of the objects. However, the subjects reported that all objects appeared to be located at the computer display, that is, at 1 m distance (note that the subjects were able to see the monitor boundary), which is of course in accordance with accommodation information. Apparently, since all objects were seen at the same position in depth, apparent size differences were not interpreted as being due to different positions in depth but were probably interpreted as real size differences.

General object description. We used convex transparent wire-frame objects consisting of about 12 planar triangular facets. They were presented in high contrast as light line segments on a dark background; the lines had a diameter of 1 pixel, and antialiasing was not used. The objects were rendered as wire frames without hidden line removal. This method of rendering can facilitate the perception of the spatial structure because it allows subjects to see all parts of the objects at the same time. During pilot experiments, subjects sometimes reported they had difficulty in memorizing the structure of the hidden part of opaque objects. This effect was strongly coupled with the angular velocity, apparently because slower rotations require comparisons of parts more separated in time. Although the definition of volume seems to be more intuitive in case of opaque objects than in the case of transparent ones, we did not want to lose the advantage of seeing the complete structure at any instant in time. (Todd, Akerstrom, Reichel, and Hayes, 1988, investigated the difference in the perceived rigidity of opaque and of transparent rotating cylinders and found minor differences for the specific frame rate that we used. Braunstein and Andersen, 1984, also reported only small differences in the veridicality of perception of shape of either opaque or transparent rotating spheres, veridicality being slightly better in the case of transparent objects.) Similar considerations regarding the specific choice of object motion led us to the conclusion that a rotation around a frontoparallel rotation axis was the most suitable one. Any other rotation, or a translation combined with perspective instead of orthographic projections, would always hide certain parts of the object from direct view. We used a vertical rotation axis, but any other frontoparallel axis would do.

Object generation, classification, and manipulation. For every trial two new objects were created as follows: A set of vertices was generated by randomly picking 10 positions within a unit cube. Next, the convex hull of this set of points was computed.[2] Subsequently, points inside the hull were removed: A set of 10 points usually reduced to about eight vertices on the hull. This enclosing surface defined our objects, which at this stage had many uncontrolled properties.

In our experiments we controlled a few structural parameters of these objects: their volume, center of mass, and elongation and flatness. Elongation and flatness were defined using the global length–width–height ratio (LWH ratio) of an object. Our definition of the LWH ratio for a general object can be found in the Appendix. Using this ratio we defined three classes of objects: *cube*-like, *rod*-like, and *slab*-like objects (Figure 1). A cube has an LWH ratio of about 1:1:1. A rod (elongated or prolate object) has an LWH ratio of about 4:1:1. Finally, a slab (flattened or oblate object) has an LWH ratio of about 4:4:1. The factor 4 was used in all experiments. The preposition *about* is used to indicate that for an object to be in a certain class the LWH ratio need not be exactly equal to a certain ratio, but may differ

by as much as a factor of 2 (which equals the square root of the aforementioned factor 4). Thus, an object with LWH ratio 2:1:1 is considered to be halfway between a cube and a rod, whereas 2:2:1 is at the boundary of cube and slab, and 4:2:1 is at the boundary of rod and slab. Clearly, there are many LWH ratios that do not fit into any of the three classes that we defined. For our purposes that does not present a problem. We need only a few well-defined and reasonably different shape categories. (A similar classification scheme, based on the notion of "extendedness," can be found in Willats, 1985; see also Willats, 1992). Roughly speaking, cubes, rods, and slabs correspond to his 3_{111}, 3_{100}, and 3_{110} regions, respectively. We prefer the mathematically well-defined moment-based description, as given in the Appendix.)

To prevent subjects from making use of knowledge about fixed LWH ratios, we scattered all three dimensions by multiplying them by factors uniformly distributed between 1 and 2. In this way all ratios become distorted, but the classification cube-, rod-, or slab-like is still valid. (A 1:1:1 object [cube-like] can, in an extreme case, become 2:1:1 [cube–rod boundary] or 2:2:1 [cube–slab boundary], but it never becomes truly rod- or slab-like.)

There is one object parameter that we have not yet discussed; we call this the regularity of the objects. In addition to the randomly shaped objects discussed above, in the second experiment we also used a set of rectangular (regular) boxes. These boxes are again described by their volume, center of mass, and cube–rod–slab classification. They consist of exactly eight corners (vertices) and are triangulated just like the random objects; that is, each of the six sides is divided into two triangles. This limits the differences between random and regular objects to the regularity alone.

After the object parameters have been defined, it is easy to control them. Elongation and flatness can be manipulated by anisotropic scaling along the principal axes of the object; see Appendix. The position of the center of mass with respect to the observer is easily changed by translating the object in space. The volume can be adjusted by applying an overall scaling of the object. Under these transformations the property of convexity is invariant. Special care was taken to ensure that the object orientation was totally random; in other words, each possible orientation had an equal probability of occurrence (the orientation of the objects can be defined using their principal axes). This randomization of orientation is very important for two reasons. First, we wanted to avoid conditions in which both objects had the same orientation, especially when they were both rods or both slabs. Having roughly the same shape and orientation, the volume difference translates into a difference in *projected* size, which makes the task much easier. Second, in general the randomization ensures that the faces of the regular objects (the boxes) are not aligned with the screen or with the (vertical) rotation axis. A situation in which they are aligned is clearly not preferable.

Procedure

The subjects were instructed to discriminate the volume of two simultaneously rotating objects. They watched the objects for an average of about three full rotations before indicating which of the two had the larger volume by moving a trackball either up or down. No feedback was given.

Each stimulus consisted of a test object and a reference object. Their order on the screen (upper or lower position) was randomly chosen for each trial. Each experimental condition was defined by the volume of the reference object, the regularity of both objects (random or regular), and the cube–rod–slab classification of the reference object and test object separately.

Adaptive stimulus control and the psychometric curve. For each condition we measured, as a function of the ratio of volumes $ln(V_{test}/V_{ref})$, the percentage of trials in which the subject indicated that the test object had the larger volume. We reasoned that if we assumed that subjects would behave similarly for different absolute volumes (scale invariance), then we would have to use a logarithmic scale. (We used this log on the basis of a result from decision theory known as "Jeffrey's rule"; see Jaynes, 1983. In fact, the assumption was confirmed by the experimental data, as we shall point out later.) An adaptive psychometric procedure (Werkhoven & Snippe, in press) was used to estimate the two parameters that characterize our psychometric curve (cumulative normal distribution): the point of subjective equality (PSE) μ, which is the physical log of the volume ratio at which the two objects appear equal in size (the PSE represents a certain bias in the subjective volume of two different objects); and the threshold σ, which is the difference in the log of the physical volume ratio relative to the PSE needed to successfully discriminate the two volumes (84% correct). After each trial, new maximum likelihood estimates of the PSE and threshold were calculated. The next trial was placed either at $\mu + \sigma$ or at $\mu - \sigma$, with equal probability. We used 100 trials per condition per subject. The adaptive procedure usually converged within 50 to 70 trials. Only 1 adaptive procedure out of 69 did not converge sufficiently and was discarded. For each psychometric curve, we computed statistical minimum estimates of the measurement errors in both μ and σ. To allow for an easier interpretation of the results, we do not present the logarithmic μ and σ but we translate their values into *subjectively equal volume ratios* and *volume increment thresholds*, respectively.

Subjects

Four subjects participated in our experiments: H.V., S.P., M.H., and I.H. The subjects S.P., M.H., and I.H. were emmetropes. The author, H.V., was myopic, but spectacle corrections were used throughout. All subjects had had previous experience with structure-from-motion tasks, but except for H.V., they were only slightly acquainted with the topic of research.

cube   rod   slab
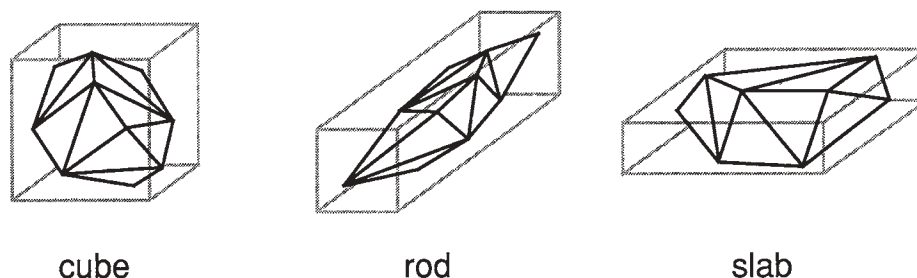
Figure 1. The three drawings depict objects with length-width-height ratios of 1:1:1 (cube), 4:1:1 (rod), and 4:4:1 (slab), from left to right, respectively. Both the regular variety (outside box) as well as the random variety (inside shape) are shown. The objects have different volumes. (To make the drawing clearer, some line elements are removed. In particular, note that the random shapes are presented as opaque instead of transparent, and that the diagonals of the regular shapes are not drawn.)

EXPERIMENT 1

Volume Discrimination as a Function of Overall Scale

Experiment 1 was designed to get a general impression of volume discrimination performance. We varied the absolute size of the objects and used a number of different shapes.

We measured the volume discrimination performance of 3 subjects (S.P., M.H., and I.H.) as a function of reference volume. We defined 15 experimental conditions. Five different values for V_{ref} were used: 3, 6, 12, 24, and 48 cm³. (To prevent subjects from using simple cues based on absolute size, in each trial we chose the actual value of V_{ref} from a uniform distribution of ±30% around one of these values.) Three different shape conditions were used: the reference object was cube-like, rod-like, or slab-like, whereas the test object was always cube-like. We used random shapes only. All 15 conditions were mixed together, resulting in a total set of 1,500 trials per subject. Subjects completed the experiment in about three 2-hour sessions.

Results

The upper panels of Figure 2 show the increment threshold values for volume discrimination for 3 subjects. First,

the thresholds are independent of the value of the reference volume, except possibly for Subject I.H. in the cube versus slab condition. Thus Weber's law for volume discrimination holds within this range of volumes. The average Weber fractions vary from 15% (S.P., cube vs. cube) to 80% (I.H., cube vs. slab), depending both on subject and shape condition. A finding common to all subjects was that the cube versus cube condition resulted in significantly lower Weber fractions (21% ± 2%, averaged over subjects and reference volumes) than the other two shape conditions (cube vs. rod, 45% ± 3%; cube vs. slab, 55% ± 6%).

The subjectively equal volume ratios are presented in the lower panels. It is clear that there are many intersubject variations. For S.P., the ratios are almost one, whereas for I.H. and M.H., the ratio of the cube versus slab condition was about two thirds, indicating a relative overestimation of 50% of cube volume relative to slab volume! For the other shape conditions there was no consistent deviation from veridicality, although M.H. did estimate rods to be slightly larger than cubes (ratio 1.13 ± 0.03, averaged over reference volumes). We can summarize the data on subjectively equal volume ratios by stating that a significant deviation from unity (the veridical value) was found for 2 out of 3 subjects, but only in the slab versus cube condi-
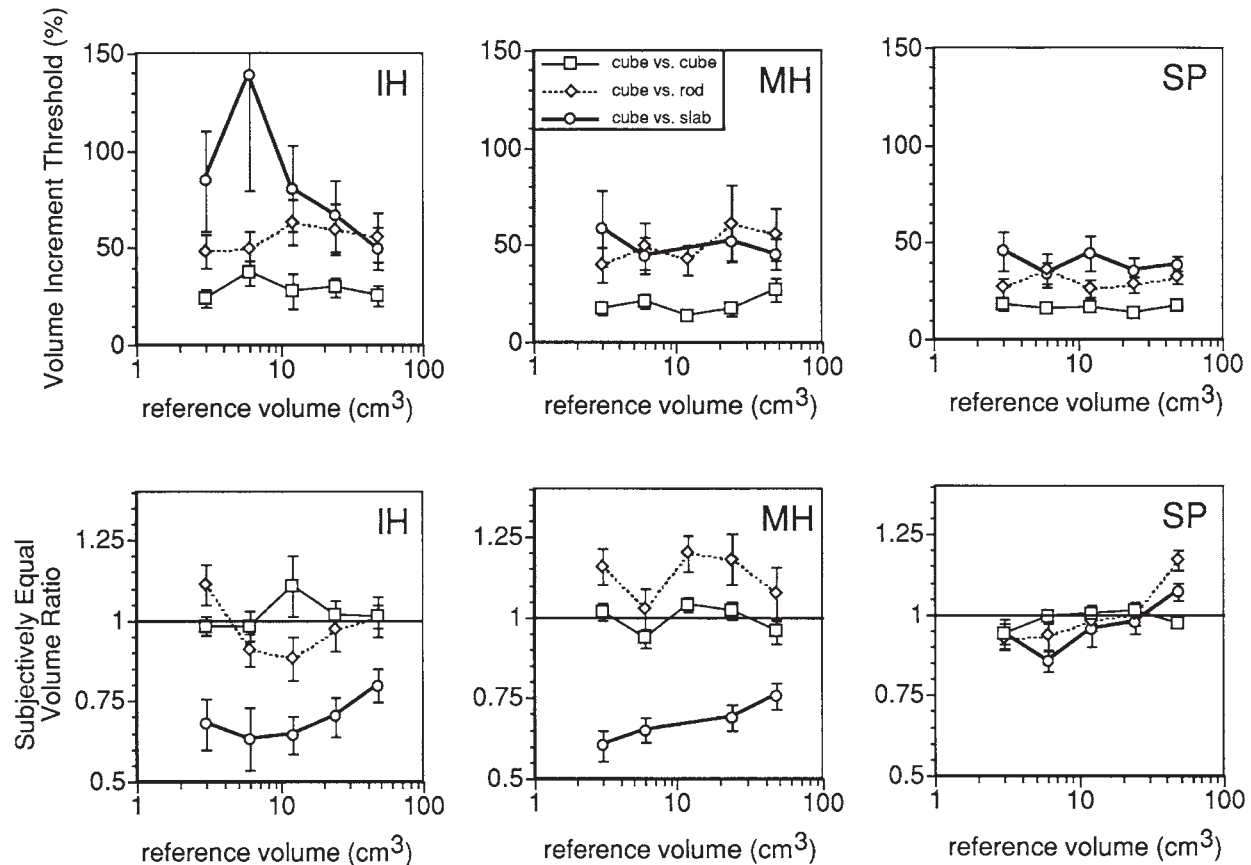

Figure 2. The results of Experiment 1. The upper row of panels shows, from left to right, the volume increment threshold percentages for Subjects I.H., M.H., and S.P., respectively. The lower row shows the corresponding subjectively equal volume ratios. The horizontal axis indicates the volume of the reference object on a log scale. The test object was always a cube-like random object, whereas the reference object was either a cube-like (square symbol), rod-like (diamond), or slab-like (circle) random object. Error bars represent minimum statistical estimates of the measurement error.

tion: In that case, the volume of the slab seemed to have been underestimated relative to the volume of the cube.

Discussion

First, volume increment thresholds of up to 80% were found, and even for the cube versus cube condition, the average increment threshold was about 20%. In view of our everyday experience with estimating weights (of ingredients for our meals, for instance), one may find these values surprisingly high. On the other hand, it is a well-known psychophysical result that discrimination of 3-D lengths also reveals high thresholds; typical values are about 25%. (Norman, Todd, Perotti, and Title, in press, used stimuli similar to ours to estimate 3-D length discrimination under varying experimental conditions. Werkhoven and van Veen, 1995, found similar thresholds for relief discrimination, which is equivalent to an affine 3-D length task.) Since we can combine three length estimates to get a volume estimate, high thresholds for 3-D length discrimination could have some predictive value for volume discrimination, although the numerous possibilities of combining lengths into volume impede actual quantitative predictions.[3] Nevertheless, the high thresholds might also have resulted from subjects having conceptual difficulties in dealing with the random shapes. It might be a good idea to use objects that are easier to recognize. We attempted this in Experiment 2.

Second, by inspecting the values of the subjectively equal volume ratios, we see that very large deviations from veridicality do occur. These deviations are subject dependent. If we extended the set of shape combinations, we might be able to show consistency of the ratios within each subject. This issue was also addressed in Experiment 2.

Our assumption that performance would be independent of reference volume was clearly proven by the data: The Weber fractions were independent of the absolute volume. Supplementary evidence for this was supplied by a pilot experiment in which we used only symmetrical conditions; that is, test and reference object were both cube-like, both rod-like, or both slab-like. With such a setup, asymmetries of the underlying psychometric curve show up as differences between objectively and subjectively equal volume ratios (bias). Such an asymmetry indicates that Weber's law has been violated. When we analyzed the data (2,065 trials) in terms of volume ratios instead of the logarithm of those ratios, a small but significant (3 sigma) deviation appeared; this was absent when we used the logarithmic scale. Apparently the psychometric curve is best described (it is symmetric) by a cumulative normal distribution of the logarithm of the volume ratio, not of the ratio itself.

EXPERIMENT 2

Shape Dependence

Experiment 2 was designed so that we could investigate more closely the differences in subjects' performance with differently shaped objects.

The reference volume was fixed at 24 cm³ (and scattered as before). We measured the volume discrimination performance of 2 subjects (H.V. and S.P.) as a function of two shape parameters: the amount of regularity of the objects and the cube–rod–slab classification of each object. We defined 12 experimental conditions. Two different levels of regularity were used: either random shapes (like those in Experiment 1) or regular shapes were shown. Furthermore, each object was cube-like, rod-like, or slab-like, resulting in cube–cube, rod–rod, slab–slab, cube–rod, cube–slab, and rod–slab combinations for test and reference shape. All 12 conditions were mixed together, resulting in a total set of 1,200 trials per subject. It took subjects about three sessions to complete the experiment.

Results

As we can see in Figure 3, the most important result of this experiment was that the performance of the subjects improved dramatically when they were tested with regular shapes instead of random shapes. In almost all conditions and for both subjects, the Weber fractions were roughly halved: The average ratio of increment thresholds (regular over random) was 0.49 ± 0.07 for Subject S.P., and 0.56 ± 0.08 for H.V. With the regular shapes, the lowest thresholds were about 10%! Because we measured each condition only once for both subjects, we do not have enough data to test the significance of the small differences between the thresholds for different cube–rod–slab combinations.

The subjectively equal volume ratios show much variation. Within a subject, the ratios for random and regular shapes were not the same. These within-subject differences seemed to be almost as large as those between subjects. No systematic difference between the ratios of regular and random shapes could be found. Note that the subjectively equal volume ratios of the symmetric conditions ought to have been unity on average; therefore, inspection of the measured ratios for these conditions gives us an impression of the accuracy of the data and the validity of the estimated measurement errors: Of the 12 ratios (2 subjects ∗ 3 symmetric conditions ∗ 2 levels of regularity), 6 deviated from unity by less than one measurement error and only 1 differed from unity by more than two measurement errors; in all this indicates a very slight underestimation of the measurement error.

In the discussion of Experiment 1, we mentioned the possibility of performing a consistency check for the ratios within a subject. To measure consistency, we defined a closure relation for the subjectively equal volume ratios R:

$$C = R_{\text{cube}-\text{rod}} \cdot R_{\text{rod}-\text{slab}} \cdot R_{\text{slab}-\text{cube}}.$$

This consistency definition is based on the assumption that the ratio of volumes and not of volume itself is the basic entity describing human volume discrimination. This view is supported by the results of Experiment 1: Weber fractions and subjectively equal volume ratios were almost independent of absolute volume over at least one decade. When the ratios of the three mixed shape conditions are consistent within a subject, the value of C should be unity. We found 0.93 ± 0.07 (random), 0.89 ± 0.04 (regular) for H.V., and 1.20 ± 0.09 (random), 0.97 ± 0.04 (regular) for
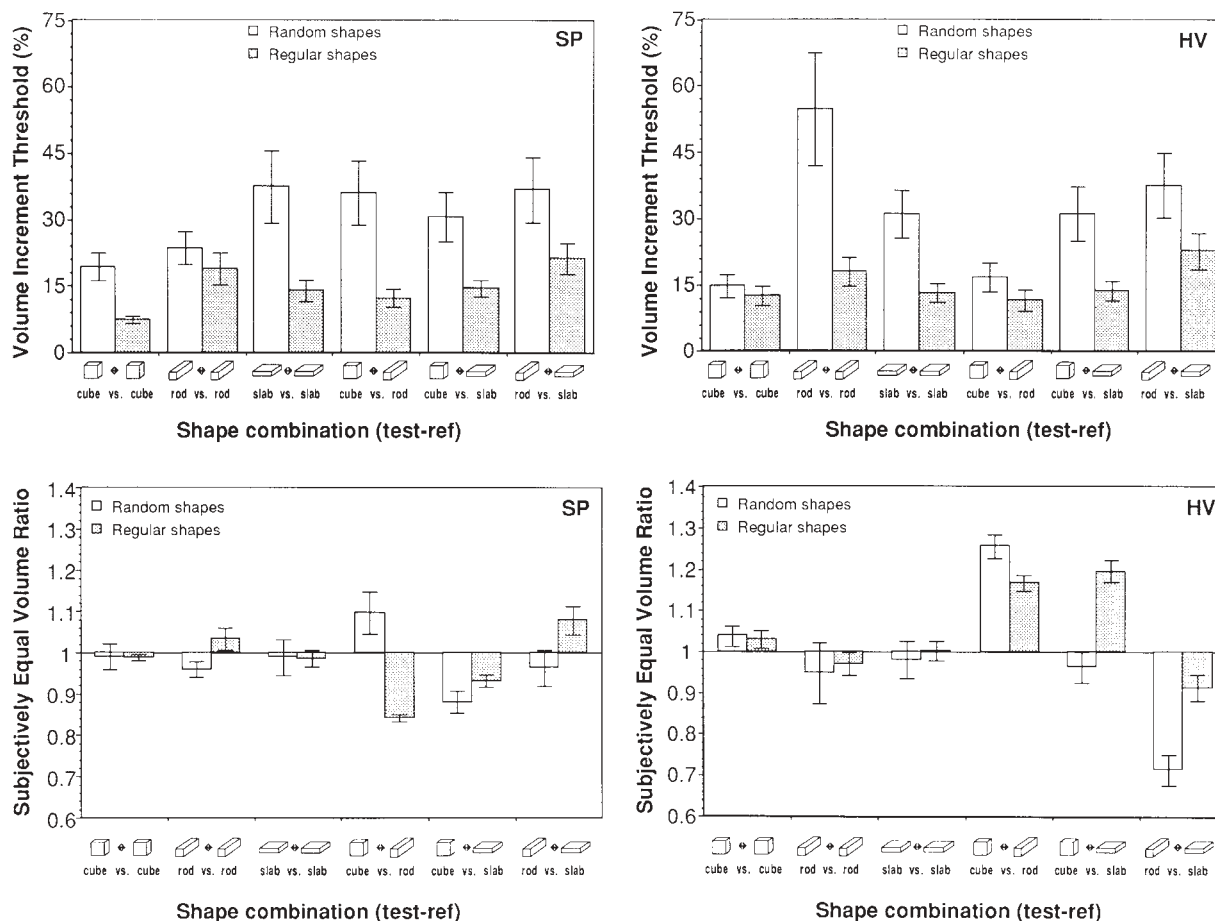
Figure 3. The results of Experiment 2. The upper row of panels shows the volume increment threshold percentages for Subjects S.P. (left) and H.V. (right). The lower row shows the corresponding subjectively equal volume ratios. The data are presented as a function of the shape condition. The various cube–rod–slab combinations are specified along the horizontal axis. The two bars for each shape combination correspond to random and regular shapes (see legend). Error bars represent minimum statistical estimates of the measurement error.

S.P. These values are not close enough to unity to demonstrate consistency of shape-bounded overestimations and underestimations of volume within each subject.

Discussion

What could be the reason for the differences in subjects' performance in the regular and the random shape conditions? The complexity of the projected scenes is equal, because we used equal numbers of line elements for both types of objects. The main differences are in 3-D: The regular objects were more symmetrical than the random ones. It cannot be deduced from the current set of experiments whether the improved performance was due to the subjects making good use of the parallelism of facets, the equal length of parallel edges, or the right angles while "computing" volume, or to the easier recognition and remembrance of more symmetric objects in general. One would need to test objects with characteristics in the range between those of the regular and the random ones. An alternative explanation based on area estimates (2-D cue) and the

variation of shapes within a shape category will be explored in the next section.

SIMULATIONS
Area-Based Volume Judgments

One aspect of our experiments that we have not yet discussed is whether area cues may perhaps be used in estimating volume. The argument is that projected area is related to surface area, and the latter is related to volume. An observer might find it easier to estimate projected area than to estimate volume directly. In this section, we report our attempt to establish the usefulness of this cue by testing an algorithm that uses the average projected area during a single rotation about a fixed axis as the direct correlate of the volume of the corresponding object.

The Relation Between Area and Volume

The objects used in the previous experiments were all convex and therefore their 3-D surface area was finite. For

these types of objects, a theorem (see van de Hulst, 1964, pp. 110–111) states that the average projected area (2-D) of such an object (averaged over all possible orientations of the object in 3-D space; *not* an average during one rotation!) is one quarter of the 3-D surface area. If, in addition, we appreciate the fact that surface area is related to volume, a link between projected area and volume emerges. The relation between surface area and volume depends on the shape of the object. To summarize so far, we can write,

$$V = \gamma(\text{shape}) \cdot S^{3/2}$$

$$S = 4 <A>,$$

in which V is the volume, S is the surface area, $<A>$ is the average projected area, and γ is a *shape*-dependent factor relating volume to surface area. This factor γ is 0.094 for spheres, 0.068 for cubes, 0.052 for rods, and 0.048 for slabs.[4] (The exponent $3/2$ is easily understood: When we scale each dimension of an object with a factor λ, its surface area scales with λ^2 and its volume with λ^3. So, in order to keep γ scale independent, we have to take the surface area to the power of $3/2$.)

Of course, a few difficulties arise when a subject tries to use the projected area cue. First, the subject needs to know the factors γ for the shapes that he/she is judging. This would mean judging shape before judging volume. Apparently, this is quite likely, because subjects can recognize the shape of the objects to at least some extent. Second, an accurate judgment of volume requires area estimation using projections from as many different directions as possible. This is clearly a big problem, because we presented the subjects with a very limited set of projections, namely those corresponding to a rotation around a fixed axis. Consider the rods, for example: Depending on the orientation of the rod relative to the rotation axis, its projected area can be constant and at its maximum during the rotation (axis of rotation aligned with the length axis of the rod), or it can vary from maximum to minimum (axis of rotation orthogonal to the length axis of the rod), or anything in between (random axis of rotation). Clearly this will result in very large differences in the average projected area of one and the same object! From this we must conclude that the estimation of volume based solely on the average projected area during one single rotation is a very unreliable method. However, it is much easier to find out how the method works for a large number of trials. The average projected area in such a large group of trials, and therefore also in a large group of relative orientations, eventually becomes equal to one quarter of the 3-D surface area. Therefore a reasonable average performance might develop. To check this "average usefulness" of the area cue, we ran several simulations of the previous psychophysical experiment, replacing the human subject by an algorithm that used the area cue. Although it is hard to anticipate the outcome of the simulations with regard to the thresholds, the values of the subjectively equal volume ratios for the different shape combinations can be predicted from the definition of the algorithm. For the regular shapes (with the standard LWH ratios) we can compute these ratios: 1.19

(cube–rod), 1.26 (cube–slab), and 1.06 (rod–slab). Even before actually running the simulations we can see that these values are quite different from the values found with human subjects. Predictions for the random shapes are much harder to make because their shapes are only defined statistically.

Method

Stimulus, procedure, and design were the same as those in Experiment 2. The subject was replaced by an algorithm that computed the projected area of each object, averaged over one complete rotation. Thus, the judgment of volume was replaced by a judgment of average projected area. Our decision to average over exactly one full rotation was somewhat arbitrary; subjects always watched at least one full rotation, but typically somewhere between two and three rotations. The use of either more or less than one full rotation would obviously influence the results of the simulations. However, in comparing the results of the algorithm with those obtained with human subjects, we decided that the most important point was to keep the number of *different* projections equal.

Results

The results are presented in Figure 4. The different stimulus conditions have two marked effects on the threshold levels. First, the thresholds for the cube–cube combination are 2 to 4 times lower than the thresholds for the other shape combinations. Second, thresholds are significantly lower for regular objects than for random ones; the average ratio of regular threshold over random threshold was 0.80 ± 0.10.

We computed the factor γ for each individual trial during the simulations and found a range of almost 1.5 decades. This means that for each single trial, volume judgments based on the average projected area can be wrong by more than an order of magnitude! However, after averaging over 100 trials per condition, we concluded that the algorithm is reasonably accurate.

The subjectively equal volume ratios of the symmetric conditions were equal to unity, as we anticipated. We can compute the closure of the ratios of the nonsymmetric conditions to check for internal consistency. Of course we do not expect the algorithm to yield any significant deviation. We found 1.00 ± 0.09 (random) and 0.93 ± 0.07 (regular). These values are indeed close to unity, closer than the results with human subjects in Experiment 2. The subjectively equal volume ratios that we measured for the regular mixed shape conditions were 1.21 ± 0.05 (cube–rod), 1.33 ± 0.04 (cube–slab), and 1.02 ± 0.05 (rod–slab), which is in agreement with the values that we predicted above using the γ factors.

Discussion

The main goal of running the simulations was to find out something about the amount of volume information carried by projected area, which then indicates the usefulness of this area cue for human observers. We already concluded that it is mainly average information that is conveyed; the area cue is not very effective for a single volume judgment. When average performance over 100 trials was compared, the algorithm showed somewhat higher thresh-
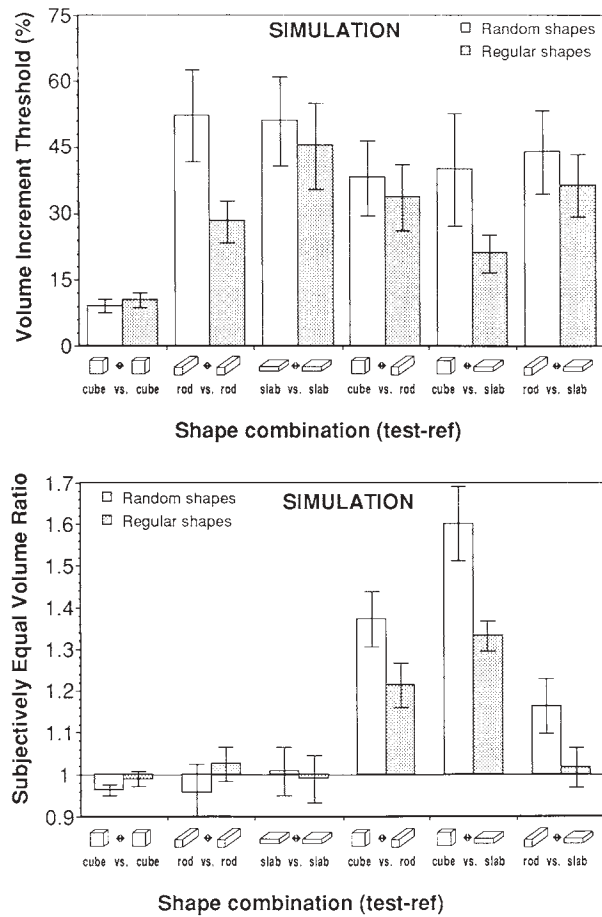
Figure 4. The results of the simulations. The upper panel shows the volume increment threshold percentages and the lower panel shows the corresponding subjectively equal volume ratios. The data are presented as a function of the shape condition. The various cube–rod–slab combinations are specified along the horizontal axis. The two bars for each shape combination correspond to random and regular shapes (see legend). Error bars represent minimum statistical estimates of the measurement error.

olds than did the human subjects, except in the cube versus cube conditions. The improvement in performance from random to regular shapes was less drastic in the case of the algorithm than in the case of human subjects: The average Weber fraction dropped to 80% of the random object fractions whereas it dropped to 53% for the human subjects. It is still surprising, however, that the absolute thresholds produced by the simulations, which are based on a rather primitive and imprecise strategy, were only moderately larger than the human thresholds.

A direct comparison of the subjectively equal volume ratios that we measured with human subjects with those produced by the algorithm will not be very fruitful, because the differences between subjects are quite large. Furthermore, since the subjects clearly recognized the objects in some sense (cube-like, etc.), their judgments might be affected by feedback and training. A deviation from veridicality then has more to do with subjects' past experiences and less to do with a specific strategy in judging volume.

We did not verify whether feedback can indeed influence a subject's behavior. Nevertheless, the fact that the subjectively equal volume ratios differed between the human and algorithmic discrimination suggests that humans do use a representation of volume that is not based on projected area alone. Moreover, this representation is on average more veridical than if it were based on projected area alone (the ratios were closer to unity than those found with the algorithm; cf. Figures 3–4, lower panels).

We must conclude that in machine vision, projected area can indeed be used as an estimator for volume, but that it is not sufficient to describe human performance in volume discrimination tasks.

GENERAL DISCUSSION

Summary

We presented results of psychophysical experiments testing the abilities of human observers to discriminate the volumes of rotating objects, as affected by the shape of the objects. We found that the volume discrimination thresholds could be described by Weber's law. The Weber fractions depended on the particular combination of shapes so that the task was easiest for regular objects, especially cubes (Weber fraction ~10%), and most difficult for irregular objects with different LWH ratios (Weber fractions sometimes up to 80%). Furthermore, we designed an algorithm using the average projected area of a rotating object to compute its volume and tried to address the relevance of this relationship to human observers. We found that the repeatability of human volume judgments varies with shape in a way that is rather well predicted by the algorithm. We also found that human observers were generally more veridical than the algorithm predicts.

Discussion

It is almost impossible to compare our results with those from previous literature on volume judgments, because we use a completely different experimental setup and analysis. However, as has been reported in the literature, we do indeed find that shape influences volume judgments. As we have argued before, the effect of shape on relative underestimations and overestimations of volume is too variable over subjects and conditions for us to draw detailed conclusions. Past research has almost always used large groups of subjects and has reported (stable) average biases instead of individual ones. For instance, Stanek (1969) found that stationary real objects, much like our regular rods in Experiment 2, were judged to have the same volume as cubes that were 10% smaller. This is an average over 60 subjects. We found the bias to vary across subjects and conditions (random vs. regular) as shown in the lower panels of Figure 3. However, the result found by Lauer (1929) and Stevenson (cited in Brunswik, 1934, 1956) that the volume of flat objects tends to be underestimated with respect to that of more cube-like objects might be reflected in the data of Experiment 1, in which 2 out of 3 subjects were found to significantly underestimate the volume of slabs relative to cubes. We found one study that actually

reported on the response variability; this response variability is closely related to our discrimination thresholds. Stevenson (cited in Brunswik, 1934, 1956) noted that in his experiments, cube–cube comparisons had much less variability (lower threshold) than comparisons of different shapes, including rod- and slab-like objects. This is indeed in agreement with our results.

Area-based volume judgments. Using projected area measures as a replacement of volume can be interpreted as an ignorance of the third dimension. We could equally well say, however, that we did not ignore the third dimension but instead assumed the unknown depth of an object to be directly related to its width, or to the square root of its projected area. Whether we compare projected area or projected area to the power of $\frac{3}{2}$ is clearly irrelevant. Such relations between depth and width have been proposed in the literature before (Caudek & Proffitt, 1993; Rubin, Solomon, & Hochstein, 1995). Whatever the interpretation, the experimental results prove that human subjects do better than a simple area-based analysis of volume. To obtain more reliable judgments with area-based algorithms, one could for instance incorporate some extra information about the γ factor (which requires a shape judgment), preferably depending on the specific orientation of the object (which requires an orientation judgment as well). This requirement of shape and orientation judgments in addition to an area estimate means making quite a complex extension of the basic method, and we will not discuss this any further. There might be some practically relevant applications of area-based volume computations in general. A machine vision system could combine these with knowledge about the chemical constituents of an object in order to compute its mass. The advantage of using vision to judge mass instead of measuring weight would be that vision can be used at any distance from the object, whereas the measurement of weight requires some force-exerting device near the object. It is obvious that the area method fails with nonconvex objects, but in general at least an upper bound of volume (mass) can be computed.

REFERENCES

Baird, J. C. (1970). *Psychophysical analysis of visual space*. Oxford: Pergamon Press.

Banach, S. (1951). *Mechanics*. (E. J. Scott, Trans.) (Monografie matematyczne 24). Warsaw: Nakl. Polskiego Tow. Matematycznego.

Braunstein, M. L., & Andersen, G. J. (1984). Shape and depth perception from parallel projections of three-dimensional motion. *Journal of Experimental Psychology: Human Perception & Performance*, **10**, 749-760.

Brunswik, E. (1934). *Wahrnehmung und Gegenstandswelt; Grundlegung einer Psychologie vom Gegenstand her* [Perception and the world of objects; foundation of a psychology based on objects). Vienna: Franz Deuticke.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: University of California Press.

Caudek, C., & Proffitt, D. R. (1993). Depth perception in motion parallax and stereokinesis. *Journal of Experimental Psychology: Human Perception & Performance*, **19**, 32-47.

Charpentier, A. (1891). Analyse expérimentale de quelques éléments de la sensation de poids [Experimental analysis of several aspects of the perception of weight]. *Archives de Physiologie Normales et Pathologiques*, **23**, 122-135.

Dudani, S. A., Breeding, K. J., & McGhee, R. B. (1977). Aircraft identification by moment invariants. *IEEE Transactions on Computers*, **1**, 39-45.

Ellis, R. R., & Lederman, S. J. (1993). The role of haptic versus visual volume cues in the size-weight illusion. *Perception & Psychophysics*, **53**, 315-324.

Franken, R. B., & Larrabee, C. B. (1928). *Packages that sell*. New York: Harper.

Gurevich, G. B. (1964). *Foundations of the theory of algebraic invariants*. Groningen, The Netherlands: P. Noordhoff.

Hu, M. K. (1962). Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, **8**, 179-187.

Jaynes, E. T. (1983). Prior probabilities. In R. D. Rosenkrantz (Ed.), *E. T. Jaynes: Papers on probability, statistics and statistical physics* (pp. 114-130). Dordrecht, The Netherlands: D. Reichel. (Reprinted from *IEEE Transactions on Systems Science and Cybernetics*, **SSC-4**, 1968, 227-241)

Koenderink, J. J., & van Doorn, A. J. (1991). Affine structure from motion. *Journal of the Optical Society of America A*, **8**, 377-385.

Lauer, L. (1929). Untersuchungen über die scheinbare Größe von Körpern [Investigations concerning the apparent size of objects]. *Archiv für die Gesamte Psychologie*, **68**, 295-324.

Norman, J. F., Todd, J. T., Perotti, V. J., & Tittle, J. S. (in press). The visual perception of 3D length. *Journal of Experimental Psychology: Human Perception & Performance*.

Piaget, J. (1941). *The child's conception of number*. London: Routledge.

Prokop, R. J., & Reeves, A. P. (1992). A survey of moment-based techniques for unoccluded object representation and recognition. *Computer Vision, Graphics, and Image Processing: Graphical Models & Image Processing*, **54**, 438-460.

Routh, E. J. (1960). *Dynamics of a system of rigid bodies*. New York: Dover.

Rubin, N., Solomon, S., & Hochstein, S. (1995). Restricted ability to recover three-dimensional global motion from one-dimensional local signals: Theoretical observations. *Vision Research*, **35**, 569-578.

Sadjadi, F. A., & Hall, E. L. (1980). Three-dimensional moment invariants. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **2**, 127-136.

Stanek, R. J. (1968). Surface and volume judgements of three-dimensional shapes. *Psychonomic Science*, **11**, 121-122.

Stanek, R. J. (1969). A parametric study of volume and surface judgments. *Perception & Psychophysics*, **6**, 16-18.

Todd, J. T., Akerstrom, R. A., Reichel, F. D., & Hayes, W. (1988). Apparent rotation in three-dimensional space: Effects of temporal, spatial, and structural factors. *Perception & Psychophysics*, **43**, 179-188.

Todd, J. T., & Norman, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics*, **50**, 509-523.

Ullman, S. (1979). *The interpretation of visual motion*. Cambridge, MA: MIT Press.

van de Hulst, H. C. (1964). *Light scattering by small particles*. New York: Wiley.

van Veen, H. A. H. C., & Werkhoven, P. (in press). Metamerisms in structure-from-motion perception. *Vision Research*.

Werkhoven, P., & Snippe, H. P. (in press). An efficient adaptive procedure for psychophysical discrimination experiments. *Behavior Research Methods, Instruments, & Computers*.

Werkhoven, P., & van Veen, H. A. H. C. (1995). Extraction of relief from visual motion. *Perception & Psychophysics*, **57**, 645-656.

Willats, J. (1985). Drawing systems revisited: The role of denotation systems in children's figure drawings. In N. H. Freeman & M. V. Cox (Eds.), *Visual order: The nature and development of pictorial representation* (pp. 78-100). Cambridge: Cambridge University Press.

Willats, J. (1992). Seeing lumps, sticks and slabs in silhouettes. *Perception*, **21**, 481-496.

NOTES

1. The shape dependency of human volume estimation also has some practical applications. In a book called *Packages That Sell*, Franken and Larrabee (1928) discussed some very interesting packaging issues, including the influence of the shape (and color) of a package on its esti-

mated size. For the specific case that they investigated, they found that customers estimated a flat round can to be larger than a taller one, although the contents of the cans were equal. The manufacturers thereupon decided to sell their codfish cakes in the flat can.

2. We generate and define the convex hull of a set of points in space in two steps. First, we take the planes defined by all possible sets of three different points [$n \times (n-1) \times (n-2) / 3!$ planes for n points]. Second, we select those planes that have all points on one side of the plane (points on the plane are ignored). When there are at least four noncoplanar points, these selected planes enclose a volume (intersection of half spaces), and the enclosing surface is called the *convex hull*. In our case, this surface always consists of a collection of connected triangular facets. If there are only a few points, almost all of them are on the hull, while only a few are inside. The number of facets is always equal to twice the number of vertices minus four.

3. A full metric representation of the objects is not necessary for comparing their volumes. A collective affine representation would suffice, since the ratio of volumes is a relative affine invariant of weight 1 (Gurevich, 1964; see also, Koenderink & van Doorn, 1991). Todd and Norman (1991) tried to find out whether the representation of visual space that humans use in structure-from-motion experiments is closer to an affine representation or to a metric one. Their conclusion was in favor of an affine representation, partly because they found the lowest thresholds with tasks requiring only an affine representation instead of a full metric representation. The high thresholds that we found seem to contradict this conclusion. At the very least, it shows the difficulty of a quantitative comparison of thresholds obtained with affine and metric tasks. Recently, van Veen and Werkhoven (in press) showed that within one experiment, thus comprising one task and one type of stimulus, responses reminiscent of an affine as well as of a metric representation could be found, depending on specific stimulus parameters.

4. This statement is valid only for the *regular* cubes, rods, and slabs with LWH ratios 1:1:1, 4:1:1, and 4:4:1, respectively. The *random* objects as well as objects with other LWH ratios had slightly different γ values.

APPENDIX
How to Define Length–Width–Height Ratios

It is easy to assign a length, width, and height to regular objects like boxes or ellipsoids. But what about irregularly shaped objects? Below we will use a decomposition into 3-D solid moments to find a generalization of the LWH ratio that is useful for a large set of objects.

A well-known class of methods used for object description and recognition is based on moment invariants (Dudani, Breeding, & McGhee, 1977; Hu, 1962; Prokop & Reeves, 1992). These moments describe object (or image) properties that are invariant under certain types of transformations, typically rotations, translations, scaling, and so on. This invariance is indeed very useful in dealing with certain recognition problems. Moments are also commonly used in mechanics to characterize bodies by their spatial distributions of mass (see, e.g., Banach, 1951; Routh, 1960). We will use moments of three different orders to describe the objects that we generated in our experiments. An analogy with the use of moments in mechanics will serve as an aid to understanding the mathematics.

The space occupied by a 3-D geometric solid can be described by a distribution function $\rho(x,y,z)$ that is 0 outside the object and 1 inside. We can define a set of Cartesian[1] solid[2] moments

$\{M_{pqr} | p, q, r = 0,1,2 \ldots \infty\}$ of this distribution function; this set in general forms a complete description of $\rho(x,y,z)$ (see Hu, 1962, and Sadjadi & Hall, 1980):

$$M_{pqr} = \iiint\limits_{space} x^p y^q z^r \rho(x,y,z)dxdydz,$$
$$p,q,r = 0,1,2 \ldots$$

These moments can be grouped on the basis of the sum of their indices, $s = p + q + r$, also called the "order of the moments." In mechanics, usually only orders up to and including 2 are used; this is also sufficient for our purposes.

The zeroth order moment M_{000} ($s = 0$) equals the *volume* of the object (or its mass, depending on the way $\rho(x,y,z)$ is defined). The first order term ($s = 1$) is a vector in space:

$$\begin{pmatrix} M_{100} \\ M_{010} \\ M_{001} \end{pmatrix},$$

which equals M_{000} times the *center-of-mass* vector. Clearly we can use these two terms to describe the size and position of an object. To adjust the values of these parameters to fit our needs, we can apply isotropic scaling and translation to the object.

The second order term ($s = 2$) is somewhat more complex; it is represented by a tensor of rank 2, which can be denoted as a 3×3 real valued symmetric matrix. This matrix has three orthogonal eigenvectors, known in mechanics as the *principal axes*, corresponding to three eigenvalues known as the *principal moments of inertia* (with reference to planes perpendicular to these axes; these values are scaled by a factor M_{000}). The direction of the principal axes can be used to define the orientation of an object (see Prokop & Reeves, 1992). Since the moments of inertia are proportional to the square of the dimensions of the object measured along the corresponding principal axes,[3] these moments can be used to define the LWH ratio of an object. For example, consider an ellipsoid: The principal axes coincide with the three axes of symmetry of the ellipsoid, and the ratio of the moments of inertia is equal to the ratio of the three radii squared. Thus, we will use the (double) ratio of the square roots of the principal moments of inertia as our definition of a LWH ratio. This ratio can be manipulated by anisotropic scaling along the principal axes.

APPENDIX NOTES

1. Some people use rotational, orthogonal, or complex moments instead of Cartesian moments, depending on their specific needs. For instance, the use of orthogonal moments allows for an easier inverse transformation, that is, computing the distribution function from a given (limited) set of orthogonal moments (see Prokop & Reeves, 1992).

2. *Solid* moments as opposed to *surface* moments; we integrate over 3-D space, not over the surface of the objects.

3. This is true only when we compute *central* moments, that is, moments relative to planes passing through the center-of-mass of the object. However, we can easily correct for the difference.