The Role of Intrinslc Axes in Shape Recognition

Proceedings of the Annual Meeting of the Cognitive Science Society, 3(0)

Wiser, Marianne

1981

The Role of Intrinsic Axes in Shape Recognition

Marianne Wiser
Massachusetts Institute of Technology

This paper presents a model for the mental representation of visual shapes which accounts for their recognition in different orientations, including novel ones. It is motivated by the following considerations.

Some shapes are easier than others to recognize after rotation in space. For example, it is easy to see that Fig. 1a and Fig. 1b represent the same shape in two different positions whereas Fig. 2a and Fig. 2b seem to represent two different shapes. This suggests that a rotation-invariant representation is more readily available for shape 1 than for shape 2. Shapes that are easily recognized in different positions are also perceived as in a specific orientation with respect to an external frame of reference (here, the page); for example, shape 1 looks vertical in Fig. 1a and oblique in Fig. 1b. Shape 2, on the other hand, is not perceived as in any specific orientation.

45''


a b

Fig. 1

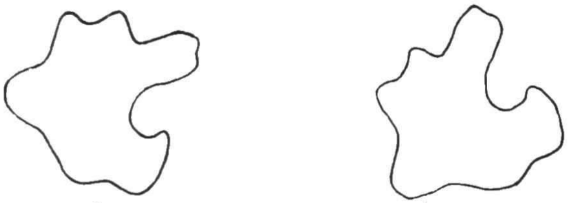

a b

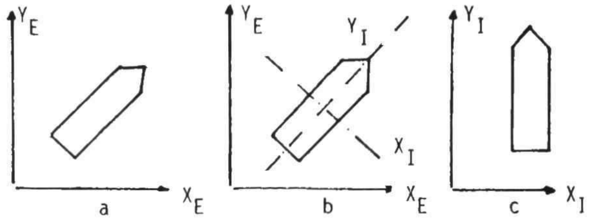
Fig. 2

Shapes like shape 1 are perceived as oriented because they have an axis of their own, determined by their geometrical properties (symmetry, elongation, parallel sides). The position of their axis relative to the side of the page determines their perceived orientation. I will call this axis "intrinsic" because it is fixed within the shape and exists independently of any other direction in space. In contrast, axes like the retinal or gravitational vertical, or the side of the page are "external" because their direction is independent of the shape being perceived. Shape 2 is not perceived as oriented because it lacks axis-determining properties.

Rock (1973) proposes a theory that explains why shapes like shape 2 are difficult to recognize after rotation. When a shape is perceived, it is described in a specific spatial frame of reference (e.g. the page); shapes are compared and recognized on the basis of this description. If the shape is rotated with respect to the frame of reference being used, its description changes and consequently, it is hard to recognize. Rock uses the notion of description within a frame of reference to explain why the change in retinal orientation that occurs when an observer tilts his head does not affect the percept of the shape, whereas the same change, when it results from tilting the shape itself (with the observer's head upright), generally makes the shape look different (as in Fig. 2). He suggests that the description tends to be performed in a gravitational, rather than retinal, frame of reference. As long as the shape has a fixed position in space, its gravitational description does not change and it will, therefore, look the same although its position on the observer's retina changes.

This theory explains why the world does not tilt when we tilt our heads, but it also predicts that we should not be able to recognize a shape rotated in space since, in that case, the description within both the gravitational and the retinal frames of reference change, A fortiori, it cannot explain the differences between shape 1 and shape 2. I propose that, when a shape has an intrinsic axis, it is used as a frame of reference to compute a description of the shape, the intrinsic description, which is independent of its position in the external frame of reference.

More precisely, the intrinsic axis and a perpendicular to it form a system of coordinates or frame of reference: the intrinsic frame of reference (Fig. 3). During the encoding of the shape, the shape is described in terms of intrinsic coordinates, i.e., its elements are localized within the intrinsic frame of reference. Such a description is rotation-invariant because the intrinsic frame of reference is fixed within the shape.



(a) Description in terms of external coordinates (X_E, Y_E)
(b) Position of the intrinsic axis Y_I in the external frame
(c) Description in terms of intrinsic coordinates (X_I, Y_I)

Fig. 3

The intrinsic axis therefore plays a dual role in the perception of the shape. On the one hand, it indicates the position of the shape in the external frame of reference; on the other, it is the frame of reference in which an invariant description of the shape can be built. The intrinsic axis makes it possible to keep separate the invariant part of the information contained in an external description (the identity of the shape) from the source of variation in the external description (the orientation of the shape in the external frame).

The experiment reported here provides evidence that the intrinsic axis and the intrinsic description are part of the mental representation of shapes. In this experiment, I will not attempt to distinguish between retinal and gravitational frames. Both are external, as opposed to the intrinsic frame. The subjects in the experiment were tested with head upright so that the retinal and gravitational frames coincided. I will refer to this frame as retinal, to simplify.

Experiment

A learning and recognition paradigm was used to test the hypothesis that the recognition of shapes with intrinsic axes is based on their intrinsic descriptions. Two models for the memory representation of shapes are contrasted: the retinal encoding model and the intrinsic encoding model. In the retinal encoding model, the intrinsic axis plays no role in the processing of the shapes, which are treated like shapes without intrinsic axes. They are stored in memory as retinal descriptions (i.e.

spatial descriptions in terms of retinal coordinates). In the intrinsic encoding model, it is the intrinsic description which is stored in memory. Shapes with intrinsic axes were presented for learning in either a vertical, oblique, or horizontal position. They were simple two dimensional non-sense shapes (Fig. 4).

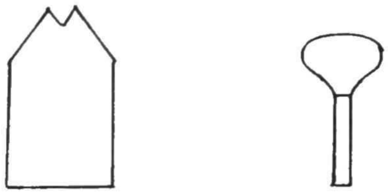


Fig. 4

45/4 The task was to recognize them among distractors when they were presented again in either the same or one of the other two positions. The subjects had to say "yes" as fast as possible if they recognized a shape they had learned and "no" if the did not. They were warned before the recognition session (but not before the learning session) that some shapes would be rotated. A "yes" shape was seen in one of 9 combinations of orientations: VV (vertical during learning/vertical during testing), VO (vertical during learning/oblique during testing), VH (vertical during learning/horizontal during testing), OV (oblique during learning/vertical during testing), etc.

For shapes seen in an oblique or horizontal position, the intrinsic frame of reference does not coincide with the retinal frame, and, thus, the intrinsic description differs from the retinal one. Depending on whether the retinal or the intrinsic description is used to represent the shape in memory, a different pattern of reaction times was predicted in each of the 9 conditions.

The retinal encoding model predicts that RT's should be faster when the learning and testing orientations are the same (i.e. when the retinal description of the target matches the description in memory) than when they differ. This is justified by numerous experiments in which shapes without intrinsic axes were used (Dearborn, 1899; Shinar and Owen, 1973; Rock, 1973). Therefore conditions VV, OO and HH should be equally fast and faster than conditions VO, VH, OV, OH, HV, and HO.

The intrinsic model on the other hand predict that shapes tested vertically should be faster to recognize than shapes tested in an oblique or horizontal position, irrespective of the orientation in which they were learned. This prediction is based on the assumption that the encoding of the intrinsic description requires more processing when the shape is oblique or horizontal than when it is vertical. The geometrical properties that determin the intrinsic axis, such as symmetry, are detected faster about a vertical than an oblique axis (Julez 1971) and the building of the intrinsic description itself involves a shift in perceptual frame of reference when the shape is not vertical. Therefore conditions VV, OV and HV should be equally fast and faster than conditions VO, VH, OO, OH and HH.

Since the models tested in this experiment cocern the memory representation of the shapes, the analysis of the "yes" responses only will be considered. It was run on RT's in msec. for correct responses. Fig 5 shows that the pattern of RT's supports the intrinsic encoding model for horizontal as well as oblique shapes.

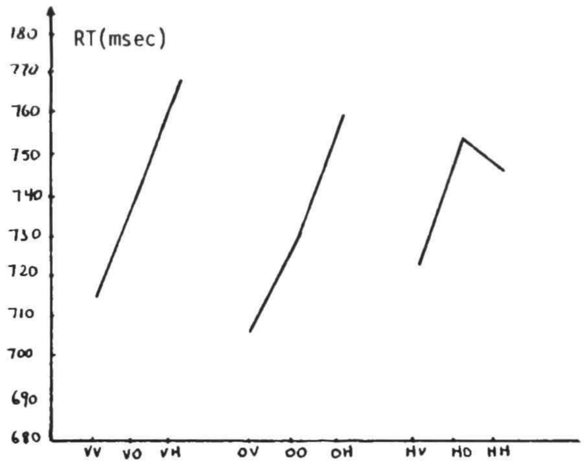


Fig. 5

An ANOVA shows that recognition time depends upon testing orientation only: vertical shapes are faster than oblique shapes and oblique shapes faster than horizontal shapes (F 21, P 10^{-6}, df 140). The effect of learning orientation and the interaction between learning and testing orientations are non-significant. As predicted by the intrinsic encoding model, vertical shapes are easy to recognize, irrespective of the orientation in which they were learned whereas oblique and horizontal shapes take longer.

The paired comparison of all combinations of learning and testing orientations (Student's T-correlated scores test) further supports the intrinsic model and infirms the retinal model. The crucial comparisons are between VO and OV and between VH and HV. The retinal encoding model predicted that they should be equally slow because they involve the same angular shift between learning and testing orientations. This is obviously not the case. OV and HV are as fast as VV and significantly faster than VO (t=2.7,df=111,P=.01) and VH (t=3.6, df=111,P<.01) respectively. These results support the hypothesis that the intrinsic description is stored in memory, irrespective of the learning orientation, and is retreived faster when the shape to be recognized is vertical than when it is oblique or horizontal.

Non-vertical shapes go through the same extra-processing whether the test orientation is the same as the learning orientation or not; OO and HH are slower than OV and HV respectively, the latter significantly so (t=2.04,df=111,P=.04). Thus, even when retinal descriptions match, a shape is recognized on the basis of its intrinsic description.

There is a strong linear dependence of RT's upon the angle between the intrinsic axis of the target and the vertical, for shapes learned vertically and obliquely. (The reason why it does not hold for shapes learned horizontally will not be discussed here.) This suggests that mental rotation might be involved in the encoding of the intrinsic description: the intrinsic description of a non-vertical shape might be obtained by mentally rotating the shape until its intrinsic axis is vertical in the perceptual frame of reference. The average linear slope (1.5°/msec.) is consistent with the rates of mental rotation found in other experiments (Cooper and Shepard, 1973; Shinar and Owen, 1973).

Conclusions

The experiment above shows that the intrinsic axes of shapes play an important role in their mental representation. The intrinsic description is stored in memory and retrieved during recognition irrespective of the orientation in which a shape is seen.

To store the intrinsic description in memory is both economical and effective. The memory representation is unique and allows the shapes to be recognized in novel orientations. I have found in other experiments that the intrinsic description is also used to compare shapes presented simultaneously in different orientations. These experiments also support the hypothesis that mental rotation is involved in the encoding of the intrinsic description: in lateralization studies, the perception of non-vertical shapes shows a right-hemisphere effect.

The intrinsic description appears to be inherent to the mental representation of shapes with intrinsic axes. Experiments in which shapes are presented very briefly and have to be identified among similar shapes in the same orientation presented immediately afterwards, show that oblique and horizontal shapes require longer exposure times to be identified correctly. This orientation effect can be attributed to the extra-processing involved in the computation of the intrinsic axes and/or the intrinsic description of non-vertical shapes. If shapes could be identified on the basis of a retinal description, indentification should be independent of orientation. Palmer (1978) has shown that the descriptions on which recognition is based are high-level, articulated descriptions. The experiments just described suggest that the intrinsic description must be encoded before recognition occurs. This in turn suggests that the only articulated descriptions are relative to the intrinsic frame of reference.

The following model can now be outlined for the processing of two-dimensional shapes with intrinsic axes. I will assume that the mental representation of shapes consists of a series of descriptions in spatial frames of reference, from low-level and local to high-level and articulated. Low-level descriptions are encoded in the retinal frame of reference (the earliest one being the distribution of light on the retina). Intrinsic axes are computed at an early stage of the processing. It has been shown that symmetry can be computed from local descriptions (Julez, 1971) and I have found that perceived elongation is based on principal axes of inertia which can also be computed from low-level descriptions. The position of the intrinsic axis relative to the retinal vertical determines the perceived orientation of the shape. The perceptual frame of reference is then shifted to the intrinsic frame and higher-level descriptions are elaborated within the intrinsic frame. The intrinsic description is stored in memory. It is the description on which recognition is based.

References

Cooper, L.A. and Shepard, R.N. "Chronometric studies of the rotation of mental images," in W.G. Chase (ed.), Visual information processing. New York: Academic Press, 1973.

Dearborn, G. "Recognition under objective reversal," Psycho. rev., 1899, 6, 98-107

Julez, B. Foundations of cyclopean perception. Chicago: Univ. of Chicago Press, 1971

Palmer, S.E. "Structural aspect of visual similarity, Mem. cogn., 6, 91-97, 1978

Rock, I. Orientation and form. New York: Academic Press, 1973

Shinar, D. and Owen, D.H. "Effects of form rotation on the speed of classification: The development of shape constancy," Perc. psychophys., 1973, 14, 149-154