Recognition of Shape and Metric Changes in 3-D Computer Models

by
Eric N. Wiebe
Sharolyn A. Converse

NC State University
Raleigh, NC

Presented at the
Human Factors and Ergonomics Society 40th Annual Meeting
Philadelphia, PA
September 2-6, 1996

Introduction

Graphic images of objects have long been a primary means of communication for design and manufacturing information of proposed products in industrial settings. Though originally the medium of exchange for these communications was paper, increasingly this product information is being displayed on computer monitors as three-dimensional (3-D) representations. Very little research has been done on visual performance issues directly related to the use of such computer modeling tools. The primary focus of this research is to use existing models of human object recognition and extend them to address issues of specific relevance to the use of 3-D computer modeling tools in design and engineering environments.

Though there are a number of contemporary theories of object recognition being debated in the research literature, certainly one of the most comprehensive is Recognition-by-Components theory. This theory is of particular interest here for a number of reasons: first, it is a comprehensive theory which substantively addresses all of the primary issues of 3-D object recognition previously discussed; secondly, because its focus is 3-D objects; and, lastly, because its theoretical foundation is based on geometric feature analysis and feature description which is highly parsimonious with how 3-D CAD computer systems are designed and operated.

Irving Biederman's (1987) Recognition-by-Components (RBC) theory is a view-independent, bottom-up, method of decomposing a 2-D pictorial projection into a collection of generalized 3-D solid primitives, called *geons*. This is exactly the same level of geometry at which 3-D CAD operators typically operate (Majchrzak, et al., 1987). There is the potential that the RBC theory could be modified to specifically take into account differences between topological and metric information used to determine the internal representation used in recognition. RBC theory focuses on the topological qualities of 3-D geometric primitives as a means of explaining object recognition. Though global metric changes (e.g., changing the size of the complete object) has been applied to this theory (Biederman & Cooper, 1992; Cooper, Biederman & Hummel, 1992), the effects of metric changes on individual geon relations have not been elaborated. Both the topological (shape) and metric comparison between parts and the ability to recognize differences is a critical analytic task of the computer-aided design (CAD) computer system operator. The need to develop performance models for both types of tasks demands further exploration of RBC theory to evaluate its appropriateness as a model in this setting.

The following hypotheses concern participant's reaction time (RT) when performing a same/different recognition task. First, comparison of objects with topologically unique geons will produce a faster response time than those with differences defined by metric changes. Second, neither metric nor topological change will show sensitivity to rotational disparity of viewpoint. Finally, in addition to the hypotheses concerning RT for both experiments, it is also hypothesized that correctness of response (error rate) will reflect RT trends and not indicate a speed-accuracy trade-off. These hypotheses were explored in the present experiment.

Method

Subjects

Participants consisted of 24 undergraduate students. None had previously taken any technical or engineering graphics courses. In addition, the subjects were screened for at least 20/30 visual acuity.

Experimental Stimuli

A 3-D model of an object was generated in 3-D modeling software (Adobe Dimensions™) on an Apple Macintosh™ computer. The object being presented consisted of a total of four primitive geometric solids, each representing a geon form as specified by Biederman's RBC theory. Figure 1 represents the actual object and orientation of the target stimulus presented to the subjects.

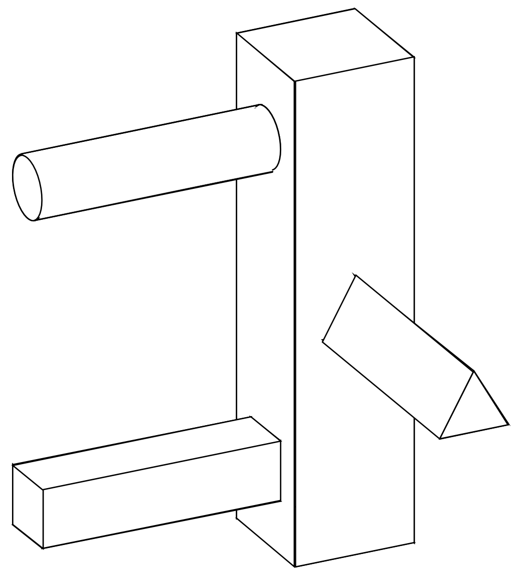


Figure 1. The 3-D computer-generated object used as the target stimulus.

Design

A 2x3x4 mixed model design was used with geometric change of the stimulus object (2), which geon changed (3), and rotational disparity (4) being the three independent variables. Geometric change was the between subjects variable while the other two variables were completely crossed within subjects.

Geometric change represents whether the change between the target object and the test object used in the comparison task was a topological or metric change (i.e. change in shape versus change in size). The change was only be applied to one of the three smaller, horizontal geons attached to the larger, vertical geon (see Figure 1). With topological change, one of the smaller geons in the target object is changed to be the same as one of the other two smaller geons. Alternatively, with metric change, one of the geons in the target was modified to represent a different aspect ratio in the test object.

Combined, the two within-subjects variables allowed one of the three smaller geons to be altered and seen at one of four possible relative rotations. The geon-altered object in the test image, when combined with the target image, represented one of twelve possible *Different* pairings seen in an experimental trial. There were also an equal number of *Same* pairings in which the test image was rotated but none of the geons were altered. Examples of different and same pairings are shown in Figure 2.

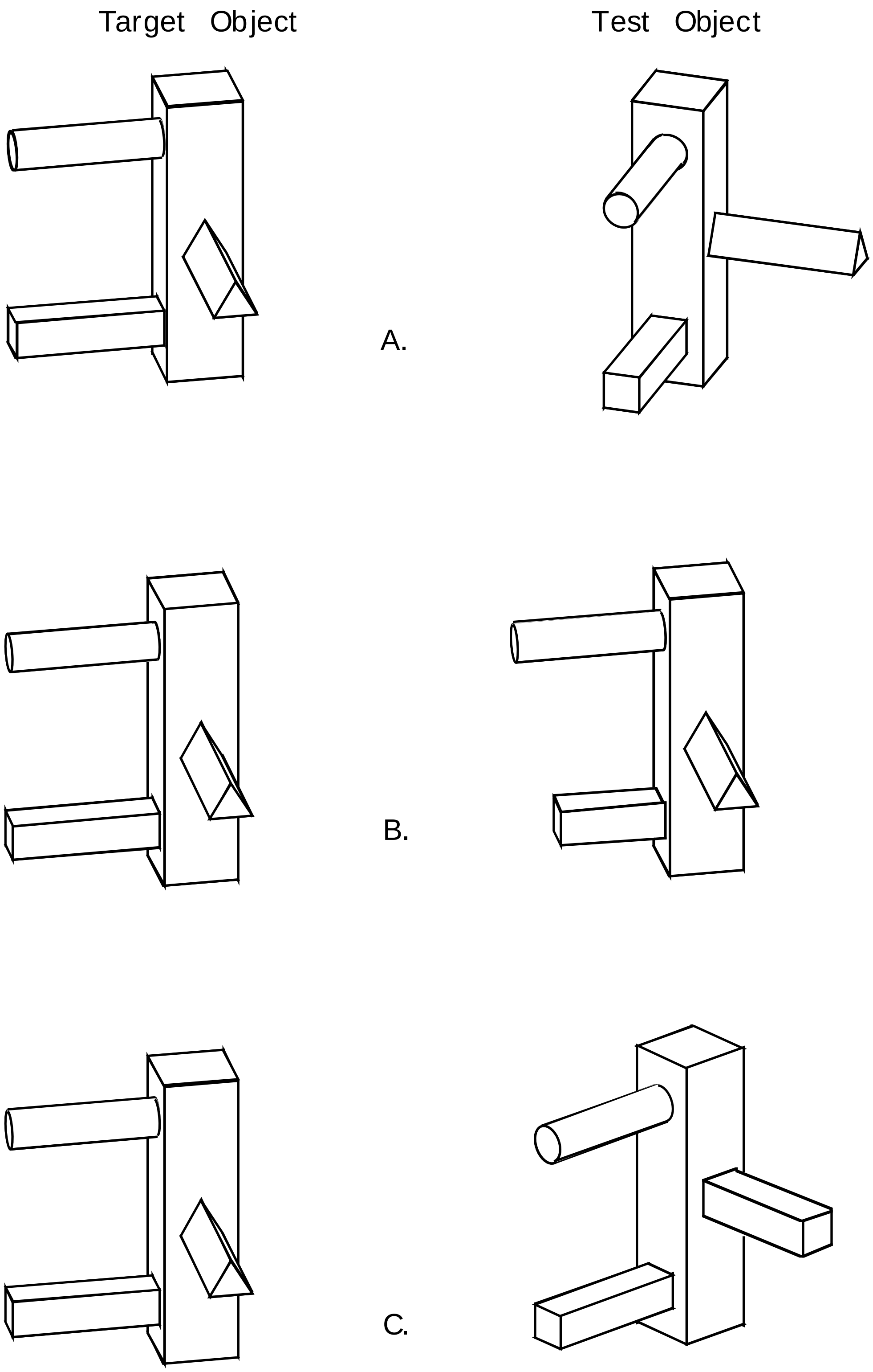


Figure 2. Examples of same and different pairings of stimuli. A) A same pairing with a rotational disparity of 60°. B) A different pairing with a metric change of the square prism. C) A different pairing with a topological change of the triangular prism to a square prism and a rotational disparity of 30°.

Experimental Task

The participants were asked to perform a standard same/different comparison task between two stimuli. The first object seen, the *target stimulus*, is depicted in Figure 1 and is unchanging. The target stimulus was displayed for 1000ms. After an interstimulus interval of 2000ms the second object, the *test stimulus*, was displayed until the subjects responded. The changes represented by the three independent variables were applied only to the test stimulus. Both the target and test stimulus subtended a visual angle of approximately 8 degrees.

Subjects were told that the goal of the task was to decide whether the target and test objects are the same, ignoring any rotational disparity. They were also be instructed to respond as quickly as possible but to keep their errors under 5% (no more than one wrong response per 20 trials).

Apparatus

The images were displayed on a Macintosh computer with a 20 inch, 1024x860, Sony Trinitron™ monitor. Responses was made by pressing one of two keys on the computer's keyboard. A software package, PsyScope, both controlled the display of the images and recorded RT and correctness of response (errors) (Cohen, MacWhinney, Flatt & Provost, 1993). RT was measured in milliseconds (msec) by the PsyScope software.

Results

Analysis

Both RT and errors were recorded. Data analysis was performed using a mixed model analysis of variance (ANOVA) with geometric change, geon to change, and rotational disparity serving as the independent variables. The dependent variables were RT for correct responses and error rate.

Response Time

The ANOVA performed on RT is summarized in Table 1. A *post hoc* analysis using the Newman-Keuls means test revealed a number of findings. Looking at Change versus Rotation (Figure 3), subjects were significantly faster viewing the Metric change at 0° than Topological change while the reverse was true at 60°. There was no significant difference at 30° and 90°. Looking at just Topological change, subjects were significantly faster at 30° and 60° than at 0° or 90°. Metric change showed different trends with 0° and 30° significantly faster than 60° or 90°.

Controlling for Geon, Rotation, and Change, Geon B tended to follow different trends than Geons A and C (Figure 4). For example, for Metric change, Geon B was significantly faster at 90° than at 0° or 60° while Geon C was significantly faster at 0° than at 90° or 60°. Topological change showed different trends; for example, Geon B was significantly faster at 30° and 0° than at 90°. Unlike for Metric change, Geon A showed similar trends to Geon B with significantly faster response at 0°, 30°, and 60° than at 90°. Geon C showed a slightly different trend, with only 60° showing a significantly faster response than 90°

Table 1
Summary of ANOVA for RT - Experiment 1

Source	**F(df)**	**Significance**
Change	F(1,22)=0.51	$p < .4831$
Rotation	F(3,66)=14.17	$p < .0001$
Geon	F(2,44)=3.11	$p < .0545$
Change x Rotation	F(3,66)=13.06	$p < .0001$
Change x Geon	F(2,44)=3.70	$p < .0326$
Rotation x Geon	F(6,132)=4.45	$p < .0004$
Change x Rotation x Geon	F(6,132)=6.68	$p < .0001$

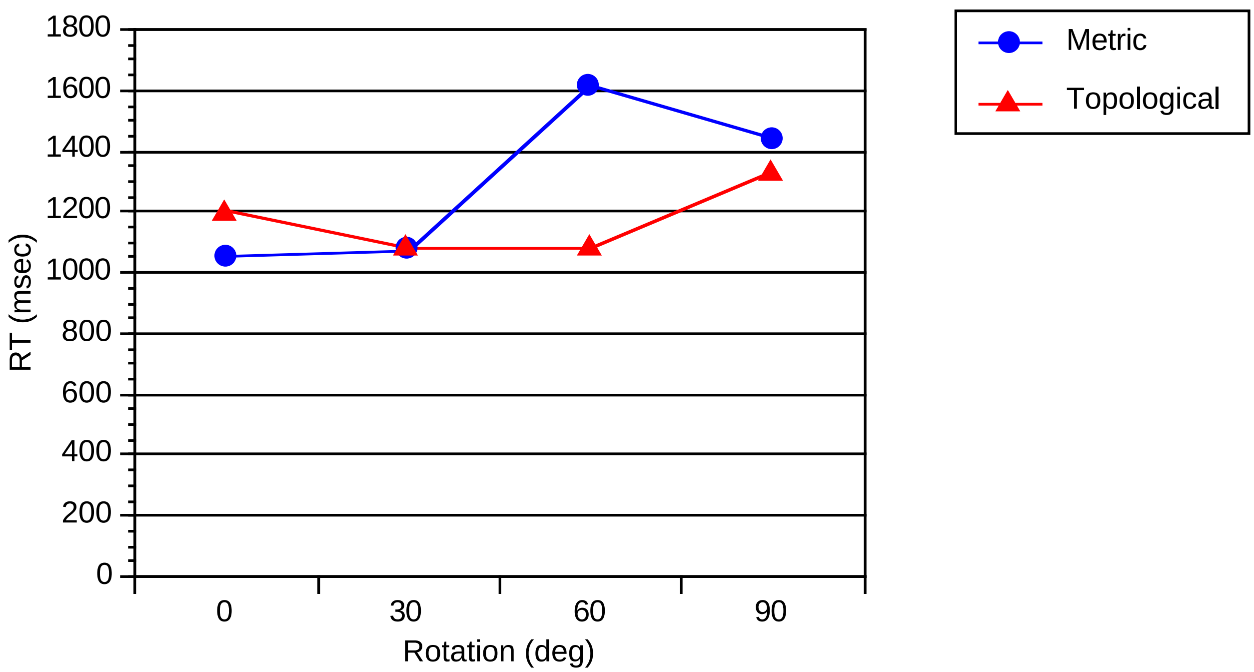


Figure 3. RT for type of Geometric Change by Rotation, collapsed across Geon.

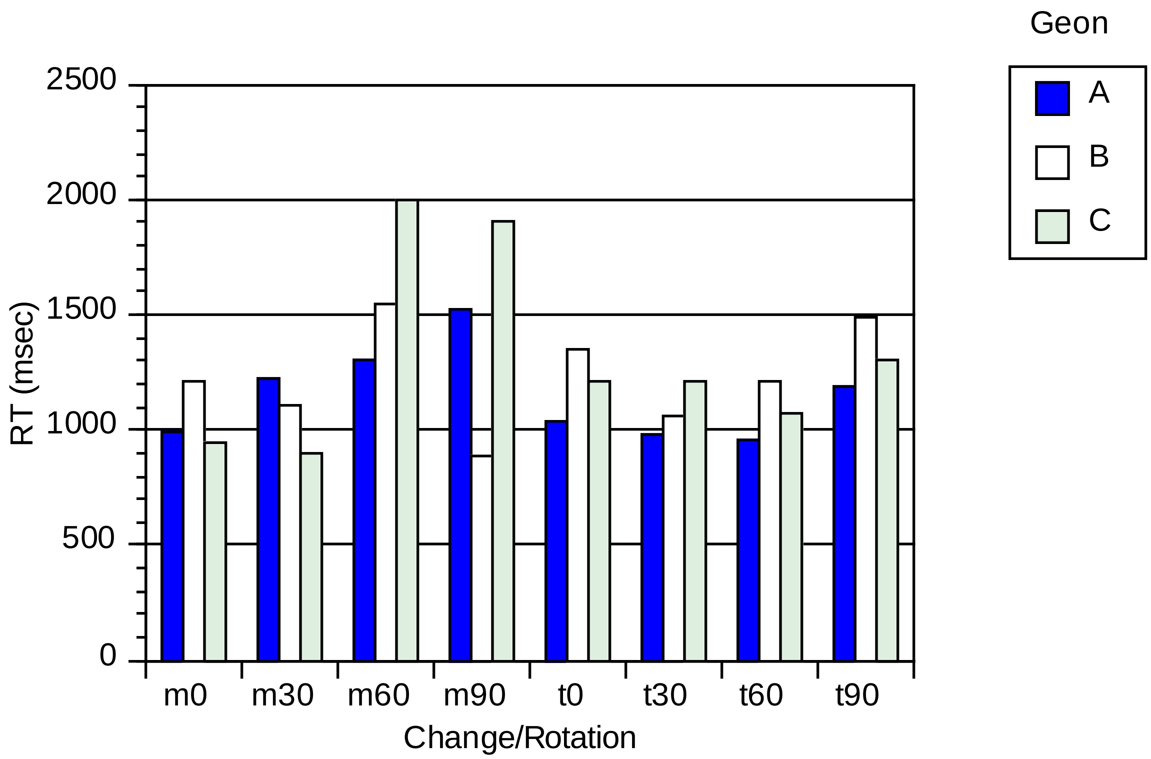


Figure 4. RT for Geon by type of Geometric Change and Rotation.

Error Rate

The ANOVA performed on error rate is summarized in Table 2. A *post hoc* analysis using the Newman-Keuls means test revealed a number of findings. Looking at Change versus Rotation (Figure 5), subjects viewing Topological change did not have a significantly lower error rate at 0° than Metric change, but did at 30°, 60°, and 90°, with the most obvious difference at 60°. Looking at just Topological change, there was no significant differences between any of the Rotational disparities while for Metric change, 60° had a significantly higher error rate than 0°, 30°, or 90°.

A comparison of Change and Geon (Figure 6) showed that Geon A had no significant difference between Topological and Metric change, but both Geon B and C showed significantly lower error rates for Topological Change. Looking at just Topological change, there were no significant differences between Geons A, B, and C while with Metric change there was a significantly higher error rate with Geon C than with Geons A and B.

Controlling for Geon, Rotation, and Change revealed that Geon C at 60° was a significant contributor to the increased error rates for the Metric Change group (Figure 7). For the Metric group, Geon C at 60° had a significantly higher error rate than at 90°, 30° or 0°. In addition, Geon C had a significantly higher error rate than Geons A or B at 60°. Also of note was the comparison of the error rates for the Metric group at 0° and at 90°. At 0°, Geon B had a significantly higher error rate than either Geon A or C. The trend was somewhat reversed at 90°, where Geon C had a significantly higher error rate than Geon B. Comparisons between the Metric and Topological groups also revealed some significant differences. Geon B at 0° had a significantly higher error rate for the Metric group than for the Topological group; at 60°, all three Geons had significantly higher error rates for the Metric group; and at 90°, Geon C had a significantly higher error rate for the Metric group. In no instances did a Geon at a particular Rotation have a higher error rate for the Topological group.

Table 2
Summary of ANOVA for error rate - Experiment 1.

Source	**F(df)**	**Significance**
Change	F(1,22)=20.72	$p < .0002$
Rotation	F(3,66)=4.26	$p < .0083$
Geon	F(2,44)=9.14	$p < .0005$
Change x Rotation	F(3,66)=7.47	$p < .0002$
Change x Geon	F(2,44)=8.87	$p < .0006$
Rotation x Geon	F(6,132)=6.19	$p < .0001$
Change x Rotation x Geon	F(6,132)=3.19	$p < .0058$

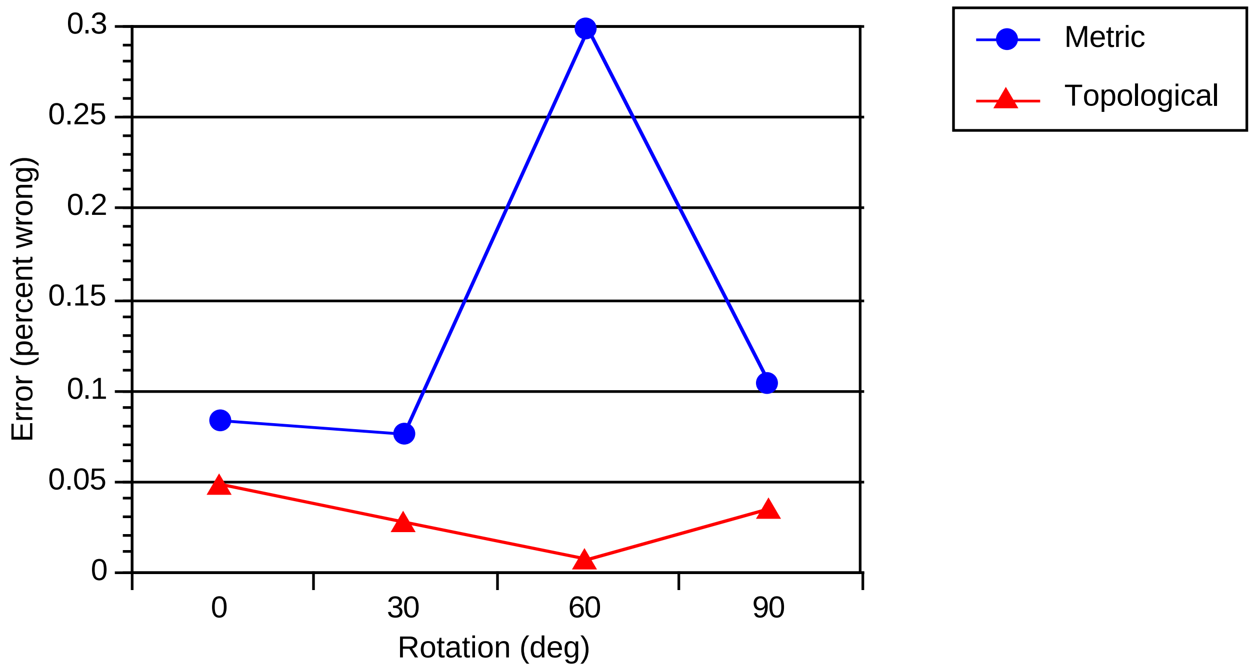


Figure 5. Error rate for type of Geometric Change by Rotation, collapsed across Geon.

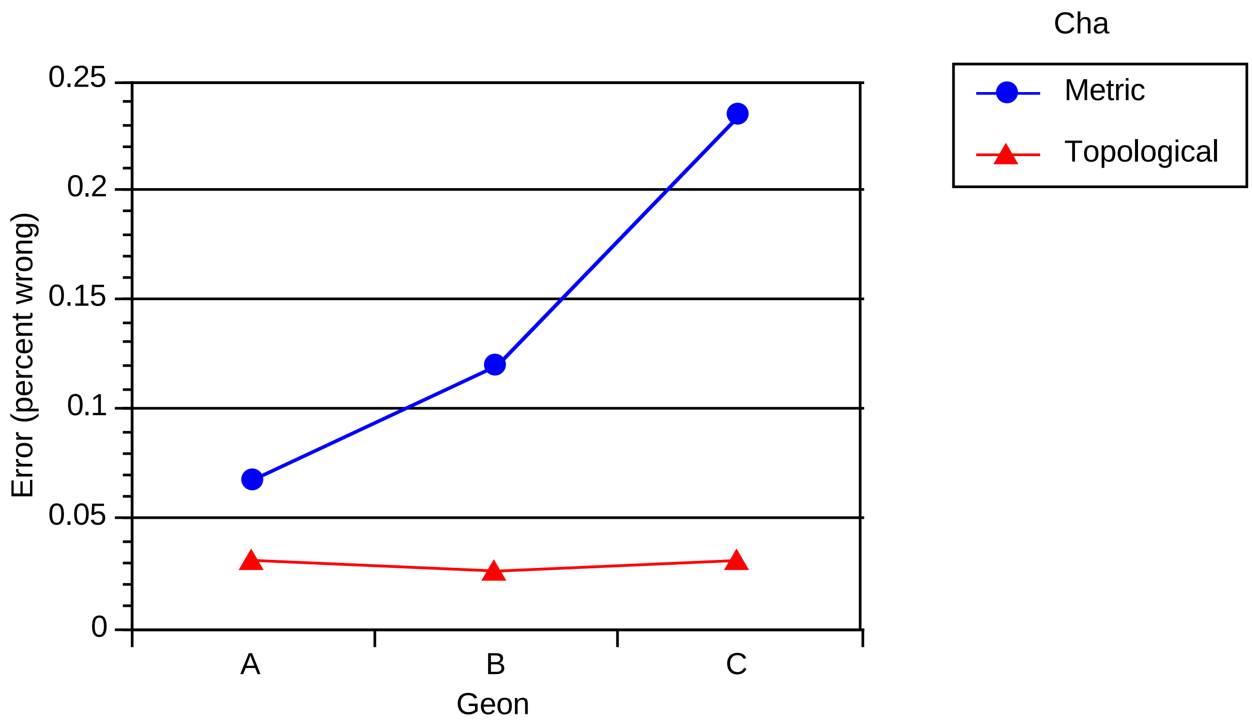


Figure 6. Error rate for type of Geometric Change by Geon, collapsed across Rotation.

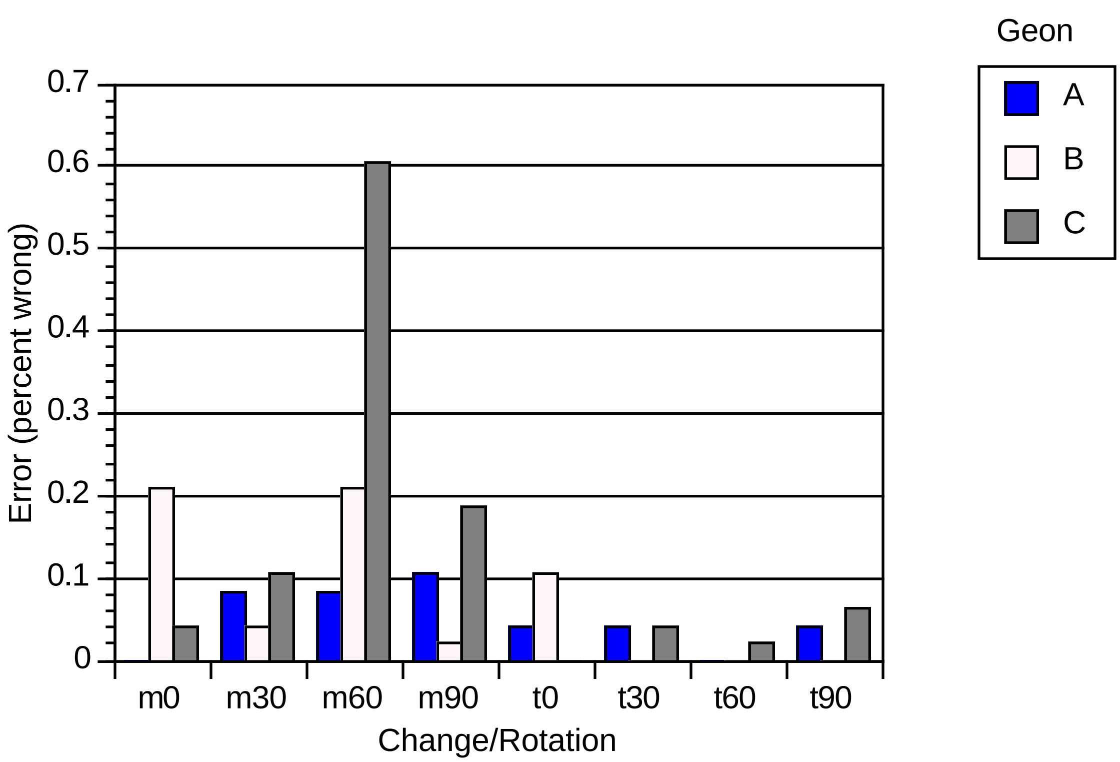


Figure 7. Error rate for Geon by type of Geometric Change and Rotation.

Discussion

The results of this experiment clearly indicate that topological and metric changes have differing effects on 3-D object recognition. At the same time, the results do seem to indicate that both topological and metric change RTs are sensitive to rotational disparity. This finding of RT sensitivity to rotation for topological change could be seen as being in conflict with the results of previous work by Biederman and his colleagues (Biederman, 1987; Biederman & Gerhardstein, 1993). The significant difference in RT between levels of rotational disparity within the Topological group can largely be explained by the interaction of the Geon variable with Rotation. The poorer performance in recognizing changes at 90° can largely be attributed to Geon B (see Figure 8). In this orientation, the end face of the prism — which is critical to the recognition of the geon — is largely obscured. Similarly, the somewhat poorer performance at 0° can be largely attributed to Geons A and C which have their end faces highly foreshortened. At 30° and 60°, the end faces of all three geons are clearly visible. This follows Biederman's caveat that obscuration of critical geon features will lead to RT sensitivity to viewpoint (Biederman & Cooper, 1991). With this explanation of the interaction of Rotation and Geon, RT data seem to indicate that the Topological group used a viewpoint-invariant, object-based model of recognition. Error rates of the Topological group revealed nothing that would contradict the RT data.

Both Biederman & Gerhardstein (1993) and Marr & Nishihara (1978) discuss the importance of the primary axis (defined typically as the longest dimensional axis and the one supporting the most number of degrees of symmetry) in identifying the topological characteristics of a 3-D form. This primary axis may also play a role in recognition of metric changes, especially when it is the dimensional axis being altered. Though the 'true length' of the primary axis of a geon is unaffected by viewpoint, the projected length of the geon subtends a smaller and smaller visual angle until it reaches zero when the primary axis is parallel to the viewpoint axis. It is interesting to note that the viewpoint of maximum foreshortening (distortion) of the primary axis is the viewpoint which shows the end face with the least amount of distortion (see Figure 9). This inverse relationship of distortion of the primary axis and the end face is seen in a comparison of the Metric and Topological groups at different rotations. That recognition performance for the Metric group was dependent on projected length of the geons seems to indicate that recognition of metric changes depended, at least in part, on an internal model that was viewpoint dependent. Superior performance for the 0° rotational disparity in both RT and error rate also indicated that overall alignment of the pictorial representations between the target and test views may also play a role.

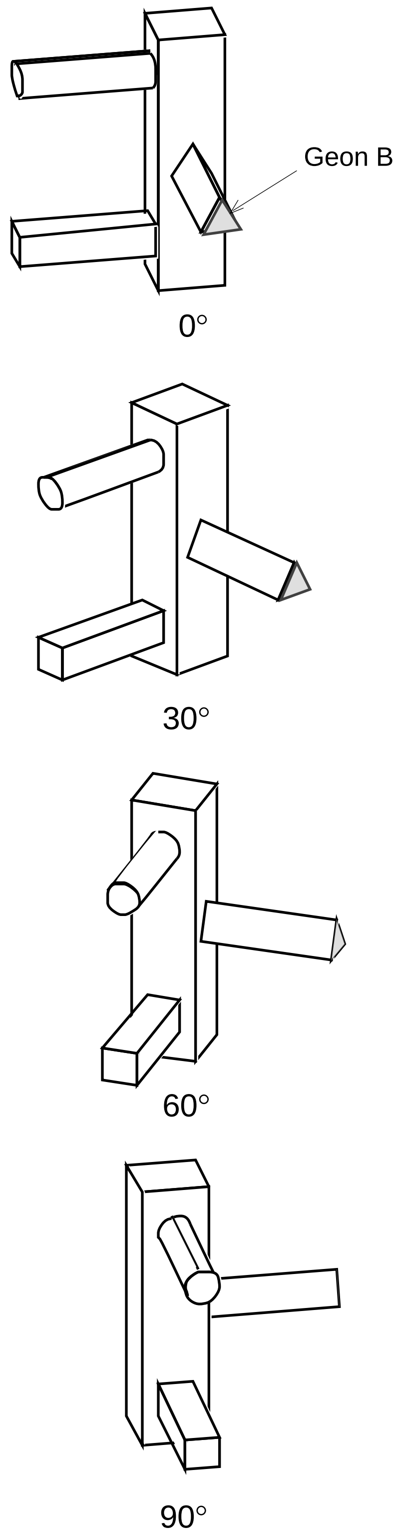
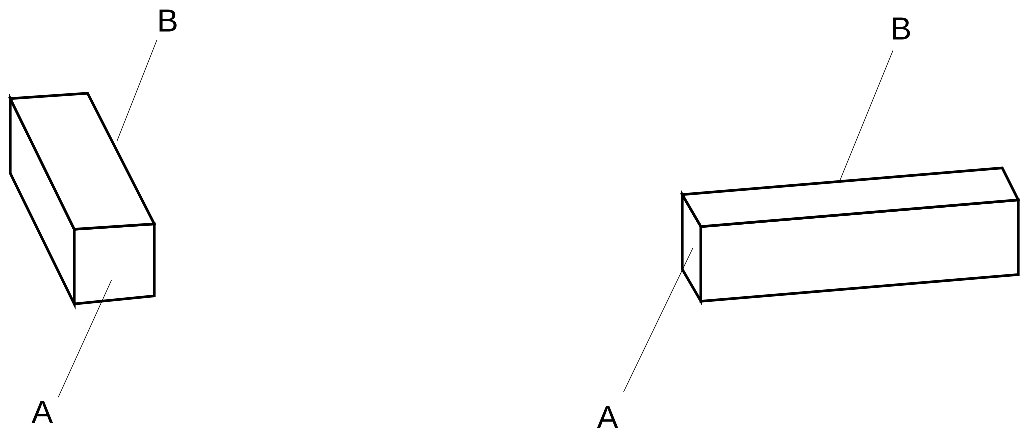


Figure 8. Differing rotations have varying levels of distortion/obscuration of the end polygon of Geon B.



A. End face near true size and shape

B. Primary axis foreshortenen

A. End face distorted

B. Primary axis near true length

Figure 9. Optimal viewpoint for topological recognition (left) versus optimal viewpoint for metric recognition (right).

One significant observation that is still left unexplained is the very high error rate and RT of geon C at 60° for the Metric group. This single geon/rotation combination explains a considerable amount of the interaction between the Change, Rotation, and Geon variables, especially for error rate. Foreshortening alone is unlikely to explain this finding since geon A at 90° has essentially the same amount of foreshortening. If metric change recognition is viewpoint dependent, geon C has the least amount of pictorial overlap with the larger, vertical geon. This overlap and the relationship between the edges on the vertical geon and the horizontal geons may be important factors in recognition performance (see Figure 10).

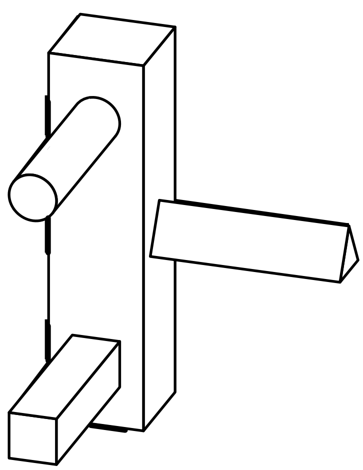


Figure 10. Geon A (top) intersects with different edges of the main geon than Geon C (bottom).

This experiment parallels the findings of Jolicoeur (1992) and Sanford, Barfield & Foley (1987) in that the application of a view-dependent or view-independent model of recognition is interrelated, in part, on the task. Though the dimensions of task definition were different than this experiment, other researchers (e.g., Hollands, 1995; Wickens, LaClair & Sarno, 1995; Wickens, Merwin & Lin, 1994) also came to the conclusion that 3-D displays were not universally applicable and performance was highly dependent on task.

The implication for CAD operators and other workers using 3-D computer models in analytic tasks is that performance — both in terms of RT and error rate — is dependent on whether the task involves metric or topological change judgments. In addition, factors such as the orientation of the 3-D model may also interact with the type of judgment required of the operator. Metric change has been a primary method of conveying change for 2-D graphic displays, but the addition of a third dimension in which this change can take place must be carefully thought out. Both the absolute orientation of dimensions which can change length and the orientation relative to comparison stimuli are factors which must be considered. Use of topological change in 3-D models, though it clearly showed advantages in this experiment, must also be applied with caution. Besides the difficulty in mapping topological change to interval scales, viewpoints which obscure key features of the form will lead to a likely failure in change identification.

References

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. Psychological Review, 94, 115-147.

Biederman, I., & Cooper, E. E. (1991). Priming contour-deleted images: Evidence for intermediate representations in visual object recognition. Cognitive Psychology, 23, 393-419.

Biederman, I., & Cooper, E. E. (1992). Size invariance in visual object priming. Journal of Experimental Psychology: Human Perception and Performance, 18, 121-133.

Biederman, I., & Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: Evidence and conditions for three-dimensional viewpoint invariance. Journal of Experimental Psychology Human Perception and Performance, 19, 1162-1182.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. Behavioral Research Methods, Instruments & Computers, 25, 257-271.

Cooper, E. E., Biederman, I., & Hummel, J. E. (1992). Metric invariance in object recognition: A review and further evidence. Canadian Journal of Psychology, 46, 191-214.

Hollands, J. G. (1995). Displaying quantitative information in two and three dimensions. Paper presented at the Human Factors and Ergonomics Society 39th Annual Meeting, San Diego, CA.

Jolicoeur, P. (1992). Identification of disoriented objects: A dual-systems theory. In G. W. Humphreys (Ed.), Understanding vision: An interdisciplinary perspective, (pp. 180-198). Oxford, UK: Blackwell.

Majchrzak, A., Chang, T., Barfield, W., Eberts, R., & Salvendy, G. (1987). Human aspects of computer-aided design. Philadelphia: Taylor & Francis.

Marr, D., & Nishihara, H. K. (1978). Representation and recognition of three dimensional shapes. Proceedings of the Royal Society of London, Series B, 200, 269-294.

Sanford, J., Barfield, W., & Foley, J. (1987). Empirical studies of interactive computer graphics: Perceptual and cognitive issues. Paper presented at the Human Factors Society 31st Annual Meeting.

Wickens, C. D., LaClair, M., & Sarno, K. (1995). Graph-task dependencies in three-dimensional data: Influence of three-dimensionality and color. Paper presented at the Human Factors and Ergonomics Society 39th Annual Meeting, San Diego, CA.

Wickens, C. D., Merwin, D. F., & Lin, E. (1994). Implications of graphics enhancements for the visualization of scientific data: dimensional intergality, stereopsis, motion, and mesh. Human Factors, 36, 44-61.