# [27] In Search of 'Characteristic View' 3D Object Representations in Human Vision Using Ratings of Perceived Differences Between Views

Patrick M Langdon, John E W Mayhew and John P Frisby

AI Vision Research Unit,
University of Sheffield, Sheffield S10 2TN

*Nineteen subjects were required to rate the perceived difference between a reference view of a 3D object and another view of the object. The latter views were created using four different axes of rotation from the reference view. Ratings of perceived difference increased with increasing angle of view point rotation. The main results, however, were that the amount of perceived difference created by a given angle of view point rotation depended greatly on the axis of rotation and that discontinuities were present in the functions relating perceived difference ratings to view point rotation angle. The latter discontinuities were matched by steps in a simple measure of feature visibility. The implications of these results for 'characteristic view' theories of object recognition are discussed. It is noted that the results are inconsistent with models incorporating shortest-path mental rotation.*

## 1 INTRODUCTION

One possible object representation for view-based recognition is a small set of 'characteristic views' each of which contains information about critical features of the object visible from the same general direction (e.g. Minsky, 1975; Koenderink, 1987; Koenderink & van Doorn, 1979; Chakravarty and Freeman, 1982). Mayhew's YASA representation (described in the introduction to this section of this book) also exploits the characteristic view idea. YASA incorporates a hierarchical organisation of clusters of 3D features such as edges, vertices, and surface regions, stable over a range of view points, to map an object's view potential for recognition.

If human object recognition uses features in this general way then it might be expected that views of an object which possess a high degree of feature commonality would be perceived as more similar than those which do not. Moreover, the function relating perceived difference to view point rotation might be expected to reveal 'steps' reflecting any qualitative changes in visible feature commonality as view point angle varies. That is, perceived difference should increase slowly (if at all) with quantitative feature distortions resulting from small changes in view point but decline sharply as the boundary between characteristic views is traversed and substantial qualitative feature changes occur[1].

Hence, the present study investigated whether visibility of face, edge and vertex features predicts perceived differences between views. The stimuli were line drawings of views of the industrial widget used by various sites within the consortium as a test object (figure 1). The views were created by rotating the view point used for a reference view around four different axes. The study had two parts: (a) measurement of differences in feature visibilities between the reference view and other views; and (b) a psychophysical experiment in which subjects were required to rate their perceived differences between the reference view and the other views.

## 2 METHOD

### 2.1 Stimuli

The stimuli were high contrast black and white slides of line drawings of pictures of an object (figure 1) generated by the IBM solid body modeller WINSOM, programmed with the appropriate 3 x 3 rotation matrices. Each slide presented two perspective views of the object, each subtending 3° of the subject's visual angle. There were 43 slides, 36 of which portrayed a reference view paired with a rotated view of the object and 7 which showed a pair of identical reference views. The reference view portrayed the object rotated 45° about the y-axis and -55° about the x-axis (i.e. rotated towards the view point). This view was paired with one of 36 rotated views, the latter taking a random left or right position over conditions.

There were four rotation axes (figure 2). For each axis the object was rotated from the starting orientation of the reference view in 20° steps from 20° to 180° and a view created for each step. For the *CYLINDER* axis condition the object was rotated around an axis of the object which corresponded to the axis of its cylindrical component. This created 9 views, each of which was paired with the reference view to form a slide. In the *HPLOS* (Horizontal Perpendicular to the Line Of Sight) condition the object was rotated around an axis perpendicular to the line of sight and lying in the horizontal plane (parallel to the picture x-axis). In the *LOS* (Line Of Sight) axis condition the object was rotated around the observer's line of sight through the object. Finally, in the *SKEW* axis condition, intended to be an axis which bore no relation to the object or principal axes of the observer's reference frame, the object was rotated around the axis in the direction of the unit axis $x = 0.512$, $y = 0.384$, $z = -0.768$ (figures 1 and 2).

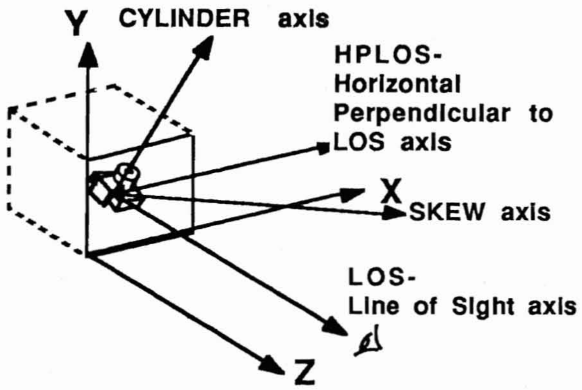

*Figure 2. The four rotation axes used for the generation of the rotated views.*

---

[1] The term qualitative is used here to refer to presence/absence of a feature in a view. The term quantitative is used for changes in the appearance of a feature with changes in view point.

As view point is changed there occasionally occur views which are 'singular' or 'degenerate', in the sense that they offer a radically unrepresentative appearance available only from a single or very restricted range of view points (Koenderink 1987; Chakravarty, 1982). An extreme example is a face-on view of a block where only one face is visible. Views of this type were avoided as far as possible but the simple form of view sampling adopted did not preclude them altogether: see, for example, the 140° views for the *CYLINDER* and *HPLOS* axes.

### 2.2 Measures of Feature Differences Between Views

A count of the number of common and distinctive features between stimulus pairs was made by hand for three feature classes: faces, edges and vertices. Each feature was numbered and a feature was counted as common when the same feature was visible in two views. A 3D vertex was counted as visible even when its appearance was different (e.g. a 3D vertex projecting into a 2D arrow junction which changes to a 2D Y-junction with view point rotation).

It was possible for views to possess the same set of common features but differ in that one view contained additional distinctive features. Hence, the following definition of feature difference was used:

(Features in reference view) - (Common - Distinctive)

For present purposes, a 'combined' feature difference measure was obtained based on all three feature types measured: faces, edges, and vertices. Separate analyses for each feature type, however, have provided a similar overall picture (Langdon, 1989).

### 2.3 Experimental Design and Procedure

Nineteen volunteer subjects with normal or corrected to normal vision took part. They saw the stimuli at eye-level back-projected on to a ground glass screen which they viewed from a distance of about 150 cm while seated.

The experimental session began with subjects being invited to examine a small hand-held model of the test object for 2 minutes. They were then shown an identical pair of reference views and told these corresponded to zero difference on the perceived difference scale which they would be taught to use. Then followed a set of ten example views, taken from view points different from those used in the experimental stimulus sets. Subjects examined these for one minute to gain familiarity with the range of the scale. This scale was defined to them in formal instructions in the following way: 0 was described as as *"a pair of identical views"* and 100 as *"the most different views imaginable"*, and judgements were to be made on the basis of *"the visual appearance of the object's shape"*. For the experiment proper subjects were instructed to call out their ratings as quickly as possible. Response latencies were measured from the onset of stimulus to verbal response, the latter recorded with a throat microphone. The latency data are reported in Langdon (1989).

Each subject performed 86 trials, the same 42 slides being presented in two blocks. The first and every sixth slide presented an identical pair of reference views to remind subjects of the meaning of zero difference on the perceived difference scale (see below). The order of presentation of all remaining slides was randomised for each subject.

## 3 RESULTS

### 3.1 Feature Difference Measures

Figure 3a shows how the combined feature difference measure varied with rotation angle around each axis. Data from the LOS axis are not shown because for that axis feature visibility remained constant over all view points. Inevitably, feature difference increased with rotation angle for all axes but two aspects of the graphs are worthy of special note, as follows:

(a) The feature difference scores from the different rotation axes tend to form separate sub-populations. That is, the data points are not scattered homogeneously around a single regression line. Those from the *HPLOS* and *CYLINDER* axes are fairly well intermingled but the *SKEW* axis points fall distinctly below them.

(b) There are some indications of step changes in feature visibility. For example, in the *HPLOS* case there is a clear step between 40° and 60°, and another between 120° and 140°. For the *SKEW* axis, there is step between 60° and 80° and a suggestion of another between 100° and 140°. The *CYLINDER* axis presents much more of a straight line, except for the suggestion of a step between 20° and 40°.

### 3.2 Perceived Difference Ratings

The means of each subject's two ratings for each stimulus were analysed using a two-factor ANOVA, with group means and standard errors plotted in figure 3b. The ±1 standard error bars around these means reveal remarkably good concordance between subject's ratings, suggesting that they interpreted the task similarly.

As was to be expected, perceived difference increased with increased view point rotation for all axes ($F_{5,92} = 135.5$, $p<0.001$). The important points of detail to note from figure 3b are as follows:

(a) There were considerable disparities in overall mean perceived difference ratings for the different axes ($F_{2,30} = 81.6$, $p<0.001$). Paired comparisons between means revealed that the *CYLINDER* axis ratings were lower than all the others, and that the *HPLOS* axis ratings were higher than those for the *LOS* and *SKEW* axes (all at $p<0.01$, Newman-Keuls).

(b) The interaction between axis and rotation angle is highly significant ($F_{9,164} = 5.9$, $p<0.001$). In the *HPLOS* case, a sharp decrease in slope is noticeable at 60°; the curve flattens off at that point although there is some suggestion of another discontinuity between 120° and 140°. For the *SKEW* axis, there is a shallow step between 60° and 80°, and some suggestion of an even shallower one between 120° and 140°. The *CYLINDER* axis data fall more or less on a straight line except for a 'bump' at the 140° point which may be a reflection of the rather degenerate character of the 140° view.

## 4 DISCUSSSION

The fact that both perceived difference ratings and feature differences scores rise with view point rotation is not at all surprising and does not in itself provide interesting evidence that our simple feature visibility measure predicts perceived difference/similarity. For an examination of this issue, various details of the data need to be considered.

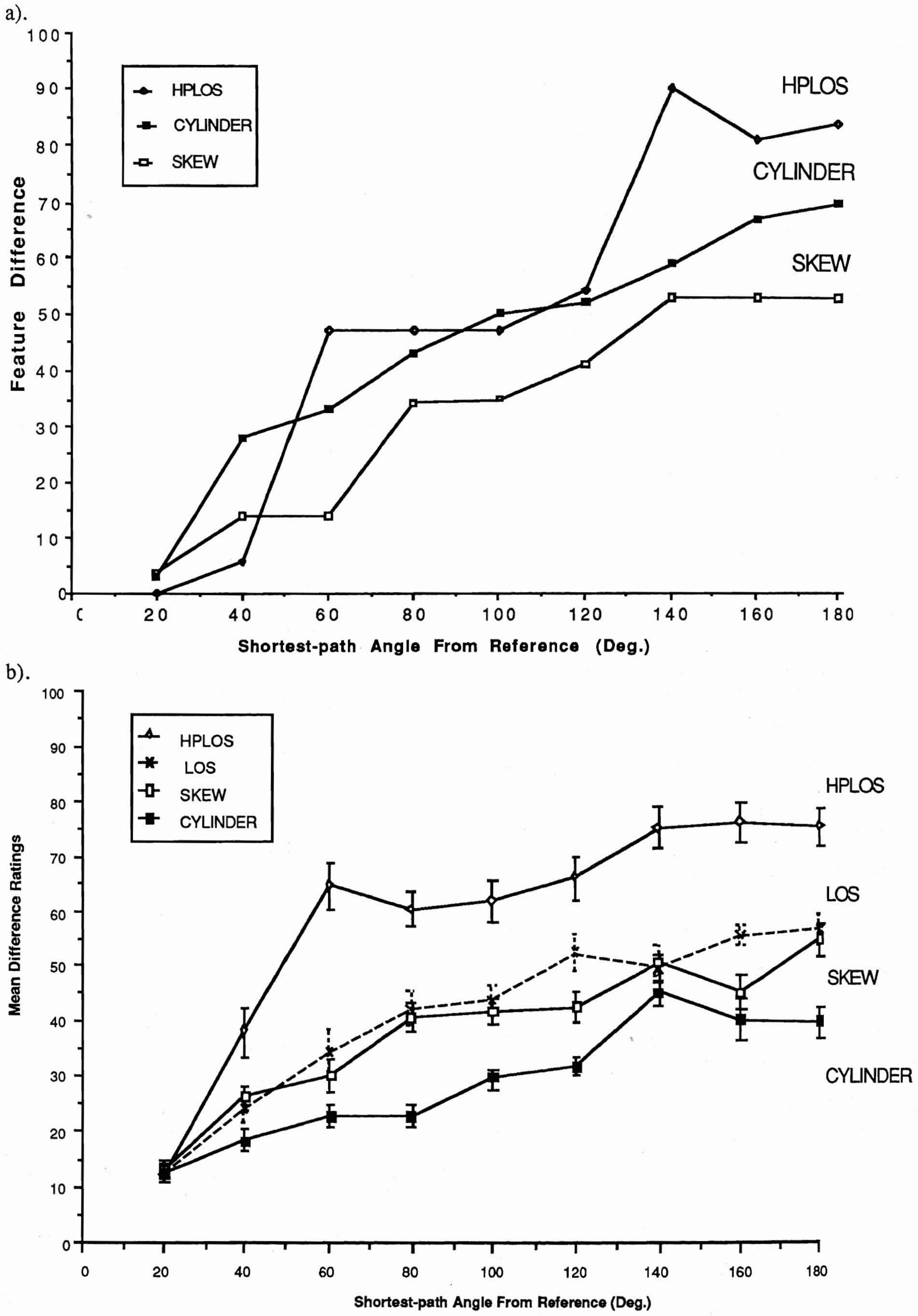

*Figure 3 Relationships of rotation angle from reference with (a) Combined face, edge and vertex feature difference (top), and (b) Perceived difference (bottom).*

(a) Both the ratings and the feature difference data for the various axes tended to form separate sub-populations, with the *HPLOS* axis having the greatest effect in each case. However, whereas *SKEW* axis rotations had least effect on feature difference scores, that position was clearly taken by the *CYLINDER* axis for the perceived difference ratings. This qualitative dissimilarity suggests that our simple feature difference score is a rather poor predictor of perceived difference ratings, but it can be argued that mean overall ratings could be influenced in part by other factors. For example, the *CYLINDER* axis conditions stayed closest throughout the rotation angle range to the 'three-quarters canonical perspective' view customarily preferred for 3D objects (Palmer, Rosch and Chase, 1981). This might have led subjects to judge them overall as 'less different'. Hence, perhaps more weight should be given to the changes within each axis rotation condition than to overall mean differences between them.

(b) In general, there is a suprisingly good match between steps in the two cases. The clearest correlation occurs for *HPLOS* at 40-60°. Another reasonably strong linkage occurs at 60-80° for the *SKEW* axis. Matches are rather less pronounced for the steps in the *HPLOS* and *SKEW* feature difference curves at around 140° but there is suggestive evidence of parallel steps in the perceived diffrence curves, at any rate for *HPLOS*. Also, the failure to find any steps in either case for the *CYLINDER* axis over the range 40-180° is in keeping with the general conclusion we draw, namely that the psychophysical data do show some interesting and encouraging signs of transitions which are broadly in keeping with the characteristic view idea. (Note: we have already suggested that the 'bump', not step, located at 140° on the *CYLINDER* perceived difference curve could be caused by subjects' responding to the degeneracy of that view.)

(c) Since no feature differences at all result from increasing rotation angles around the LOS axis, feature visibility completely fails to account both for the monotonic increase in perceived difference obtained for this axis and its mid-way position in the overall perceived difference ordering of conditions (figure 3b). It might be argued, however, that perceived difference ratings for this axis were a special case for which subjects used their ratings to reflect sensitivity to stimulus orientation rather than feature differences.

We conclude that our simple model of feature visibility provides a remarkably good account of some aspects of the perceived difference judgments, namely the locations of step discontinuities. However, it is also clear that not all the details of the judgement data can be predicted by simple feature visibility - but perhaps that would anyway be too much to expect given the intrinsically ill-defined and general character of the 'perceived difference' rating scale which subjects were asked to use.

To sum up, we regard the results as providing some limited support for characteristic view theories founded on simple lists of visible features.

## 5 MENTAL ROTATION

Driven by the observation that subjects seemed sensitive to rotation of the image in the LOS condition even though feature visibility there remained constant, we have also attempted to interpret the data within the context of the literature on mental rotation. However, both analogue and propositional mental rotation theories (e.g. Shepard and Metzler, 1971; Just and Carpenter, 1976) would predict that perceived difference should vary with rotation angle in the same way for all rotation axes. This is because all rotated views were generated using shortest-path rotations around their respective axes, and hence could be brought into congruence using the same-sized shortest-path rotations; and shortest-path rotation is assumed in those mental rotation models. Yet it is clear that marked disparities in perceived difference ratings were obtained for the four axes. Langdon (1989) shows how these ratings can be well-described by a 'spin-precession' mental rotation model (Parsons, 1987). Langdon suggests that the perceived difference ratings could reflect two simultaneous mental rotations: one (precession) brings a major axis of the object, here the CYLINDER axis, into alignment, while the other comprises a rotation around that axis (spin).

## Acknowledgements

Patrick Langdon was supported by a SERC IT-linked studentship. We are grateful to IBM UK Ltd for the use of the WINSOM body modeller.

## REFERENCES

Chakravarty, I. and Freeman, H. (1982) Characteristic views as a basis for three-dimensional object recognition. *SPIE, Robot Vision* **336** 37- 45.

Just, M.A. and Carpenter, P.A. (1976). Eye fixations and cognitive processes.*Cognitive Psychology* **8** 441-480.

Koenderink, J.J., & Van Doorn, A.J. (1979). Internal representation of solid shape with respect to vision. *Biological Cybernetics* **32** 211- 216.

Koenderink, J.J. (1987). An internal representation for solid shape based on the topological properties of the apparent contour. In Whitman R. and Ullman, S. (Eds.) *Image Understanding*. Ablex Publ. Pp. 85-86.

Langdon, P.M. (1989). *Perceived similarity judgments between multiple views of 3D objects and 'Characteristic View' theories of recognition.* PhD thesis, University of Sheffield.

Minsky, M. (1975). A framework for representing knowledge. In Winston, P.H. (Ed.). *The psychology of computer vision.* New York: McGraw-Hill.

Palmer, S. Rosch, E and Chase P. (1981) Canonical perspective and the perception of objects. In Long, J. and Baddeley, A.D. (Eds.) *Attention and Performance IX.* Hillsdale, N.J.: Larence Erlbaum Associates Inc. Pp. 135-151.

Parsons, L.M. (1987) Visual discrimination of abstract mirror-reflected three-dimensional objects at many orientations. *Perception & Psychophysics* **42** 49-59.

Shepard, R.N. and Metzler, J. (1971). Mental rotation of three dimensional objects. *Science* **171** 701- 703.