Effects of parametric manipulation of inter-stimulus similarity on 3D object categorization

Shimon Edelman*

School of Cognitive and Computing Sciences

University of Sussex

Falmer, BN1 9QH, UK

shimone@cogs.susx.ac.uk

Heinrich H. Bülthoff and Isabelle Bülthoff

Max-Planck-Institut für biologische Kybernetik

Spemannstrasse 38

D-72076 Tübingen, Germany

Revised 23 October 1998

Abstract

To explore the nature of the representation space of 3D objects, we studied human performance in forced-choice categorization of objects composed of four geon-like parts emanating from a common center. Two categories were defined by prototypical objects, distinguished by qualitative properties of their parts (bulging vs. waist-like limbs). Subjects were trained to discriminate between the two prototypes (shown briefly, from a number of viewpoints, in stereo) in a 1-interval forced-choice task, until they reached a 90% correct-response performance level. After training, in the first experiment, 11 subjects were tested on shapes obtained by varying the prototypical parameters both orthogonally (Ortho) and in parallel (Para) to the line connecting the prototypes in the parameter space. For the eight subjects who performed above chance, the error rate increased with the Ortho parameter-space displacement between the stimulus and the corresponding prototype; the effect of the Para displacement was weaker. Thus, the parameter-space location of the stimuli mattered more than the qualitative contrasts, which were always present. To find out whether both prototypes or just the nearest one to the test shape influenced the decision, in the second experiment we varied the similarity between the categories. Specifically, in the test stage trials the distance between the two prototypes could assume one of three values (Far, Intermediate, and Near). For the 13 subjects who performed above chance, the error rate (on physically identical stimuli) in the Near condition was higher than in the other two conditions. The results of the two experiments contradict the prediction of theories that postulate exclusive reliance on qualitative contrasts, and support the notion of a representation space in which distances to more than one reference point or prototype are encoded (Edelman, 1998).

1 Introduction

To make sense of the world of shapes it encounters, the visual system must overcome two major computational difficulties. The first of these is the variability in the appearance of a 3D object (and, hence, in the stimulus it presents to the visual system), caused by the varying viewing conditions, such as illumination and pose with respect to the observer. Thus, the same 3D object may look quite different when seen from different viewpoints; to realize that two views belong to the same object, the visual system must reveal their common origin, while ignoring (or making explicit) the conditions that gave rise to their differences.

The second source of problems is the variability in the shape of individual objects belonging to the same category. Just as a series of views of the same object must be perceived as such, a collection of different shapes should be attributed to the same category, if they are sufficiently similar. There is, however, an important distinction between the two cases: whereas the changes in the object appearance precipitated by changing viewpoint can be fully characterized by a handful of parameters (as few as six, in the case of a rigid object), the variation in the shape of objects belonging to the same class is *a priori* unconstrained.

A convenient common approach to the description of the two kinds of computational problems mentioned above is to coach both in terms of certain subspaces of the *measurement space* — the space of all possible outputs of the filters (or receptive fields) at the initial stage of the visual system (Edelman, 1998). Recognizing an image as a view of some object then becomes the problem of deciding the membership of that image in the space of all views of that object, which we call its *view space.* Analogously, the categorization of an image as produced by some member of a class of shapes amounts to pinpointing the location of the image in a *shape space* spanned by all members of that class within the measurement space.

A considerable amount of attention has been given recently to the issues involved in the perception of different views as belonging to the same object, or, using the terminology we just introduced, in the processing of the view spaces of individual objects. In contrast, much less work has been done on the processing of shape spaces generated by object categories. In the present paper, we report two experiments intended to fill this gap.

1.1 View space effects

Psychophysical studies conducted in the past few years led to the characterization of certain basic limitations of the visual system in generalizing shape-based recognition to novel conditions; see (Jolicoeur and Humphrey, 1998) for an extensive review and a discussion. Specifically, it was found that the recognition of novel views of objects tends to be slower and more prone to errors than the recognition of highly familiar views (Rock and DiVita, 1987; Tarr and Pinker, 1989; Bülthoff and Edelman, 1992; Edelman and Bülthoff, 1992; Humphrey and Khan, 1992; Cutzu and Edelman, 1994; Bülthoff et al., 1995; Lawson et al., 1994; Lawson and Humphreys, 1996), This effect persists even when full 3D shape information is available to the subject through, e.g., binocular stereo cues (Edelman and Bülthoff, 1992).

The relevance of the above findings to the understanding of the processes of object recognition has been disputed on the basis of the difference between viewpoint-dependent performance exhibited by the subjects in these experiments, and the viewpoint-invariant performance found in other studies. In particular, Biederman and Gerhardstein (1993) reported essentially viewpoint-invariant performance on some of the objects used previously by Edelman and Bülthoff (1992), to which distinctive single parts have been added. A subsequent detailed investigation, in which the number of distinctive parts was manipulated in addition to object orientation, showed, however, that recognition always becomes poorer with increasing change in viewpoint, although this dependence is at its weakest for objects with one unique part (Tarr et al., 1997).

1.2 Shape space effects and the role of similarity

The assumption that the processes and the representations involved in identifying specific individuals are different from those used for categorization (Jolicoeur, 1990) has been recently put to an explicit test in a series of experiments, in which objective similarity between stimuli (and, consequently, the categorical level of their distinction) varied in a controlled fashion (Edelman, 1995a).

Subjects in those experiments were trained to discriminate between two classes of computer generated 3D objects, one resembling monkeys, and the other dogs. Both classes were defined by the same set of 56 parameters, which encoded sizes, shapes, and placement of the limbs, the ears, the snout, etc. Interpolation between parameter vectors of the class prototypes yielded shapes that changed smoothly between monkey and dog. Within-class variation was induced in each trial by randomly perturbing all the parameters. After the subjects reached 90% correct performance on a fixed canonical view of each object, discrimination performance was tested for novel views that differed by up to 60° from the training view. In all the experiments reported in (Edelman, 1995a), higher inter-stimulus similarity was associated with an increase in the mean error rate and, for misorientation of up to 45°, with an increase in the degree of viewpoint dependence. On the one hand, when the inter-stimulus similarity was low, performance was essentially independent of viewpoint, despite the lack of qualitative (in Biederman's terms) contrasts between the classes in that particular experiment. On the other hand, in an experiment where qualitative contrasts were present, viewpoint dependence was strong, especially in the high-similarity condition. These results suggest that a geon-level (Biederman, 1987; Biederman and Gerhardstein, 1993) difference between stimuli is neither strictly necessary nor always sufficient for viewpoint-invariant performance.

The studies we mentioned so far concentrated on the quantification of the effects of viewpoint on recognition, and on the interaction between these effects and those of similarity among the objects that were to be recognized. While these studies explored the effects of the relative location of the stimuli both in the view space and in the shape space, the former exploration has been more thorough. For example, the experiments of (Bülthoff and Edelman, 1992) involved parametric control over viewpoint along two mutually orthogonal dimensions, whereas the study of (Edelman, 1995a) only manipulated the similarity between the two classes of stimuli, which is a one-dimensional quantity. Thus, in the experiments reported below, we chose to concentrate on a parametric exploration of the effects of shape-space proximity (similarity) between the stimuli, the issues of viewpoint having been deemed of a secondary importance, in view of the previous findings in this field.

2 The ORTHO experiment

The first experiment involved two classes of objects (p_1 class, p_2 class), defined by prototypes p_1 and p_2 (see Figure 1). The objects were jointly parameterized by a number of variables that controlled their appearance. Each object thus corresponded to a point in the parameter space (Figure 2). The shape of the objects was manipulated by combining two orthogonal directions of displacement in the shape (parameter) space — in parallel and in perpendicular to the line connecting p_1 and p_2 (Figure 3). Altogether, 15 exemplar objects for p_2 class that were to serve as the stimuli in the test phase of the experiment were formed by this procedure (Figure 8).

2.1 Motivation

One may note the parallel between the layout of the stimulus space in the present experiment and in the INTER/EXTRA/ORTHO experiments of (Bülthoff and Edelman, 1992). In that study, two groups of views of the same target object, separated by a 75° rotation in depth, were shown to the subjects in the training stage. The subjects then had to discriminate between new views of the target and views of some distractors, with the former situated on the same great circle of the viewing sphere (in between the training views, or outside the arc they defined), or on an orthogonal great circle. The significantly less than perfect performance of the subjects on all three kinds of test views (including those in between the training views) provided evidence against a particular family of theories of generalization of recognition to novel *views*.[1]

In the present study, our aim was to examine various theories of categorization, that is, generalization of recognition to novel *shapes*. Consequently, the shape-space arrangement of the stimuli in our experiments combined key elements of the view-space arrangement used in (Bülthoff and Edelman, 1992) with the introduction of controlled shape contrasts between two classes of stimuli, as it was done in (Edelman, 1995a). Specifically, our two

---

[1] Specifically, this constituted evidence against the exclusive reliance of recognition on linear view interpolation (Ullman and Basri, 1991), which predicts perfect performance on novel views that can be represented as linear combinations of familiar ones.

class prototypes differed by a so-called qualitative contrast (Biederman, 1987): the sign of the bulge of the generalized-cylinder parts. Theories that postulate reliance on such contrasts (e.g., Biederman's Recognition By Components, or RBC) predict viewpoint-invariant near-perfect discrimination performance for the two class prototypes, *and* for stimuli derived from the prototypes by a parameter-space displacement which is orthogonal to the line connecting p_1 and p_2. The same prediction can be derived from theories that postulate involvement of metric features, but do not allow for the possibility of an interaction between the different orthogonal dimensions of the feature space (for an example of such a theory, see Ashby and Perrin, 1988). The reason for this prediction is that a variation which is orthogonal to the difference between p_1 and p_2 should not affect discrimination.

At least one class of theories of representation and recognition does predict an effect for the manipulation described above. Namely, lower performance for stimuli that are farther from the line connecting the two prototypes (Figure 4) is predicted by exemplar-based theories (Nosofsky, 1988; Edelman, 1998)[2] but not by theories that postulate the construction of a decision surface (Ashby and Perrin, 1988).

2.2 Method

Eleven subjects were trained to discriminate between the two prototypes in a 1-interval forced-choice task. In each trial, an image of one of the two prototypes was briefly (300 *msec*) presented on the screen of a Silicon Graphics workstation, in binocular stereo (using LCD shutter glasses synchronized with the display). The subject was required to press the right or the left key on the computer mouse, depending on the class to which the stimulus belonged;

---

[2] These theories may be considered the shape-space analog of the interpolation approach to the recognition of novel views of objects (Poggio and Edelman, 1990).

an incorrect response triggered a beep (only during the training phase of the experiment). The object could be seen from any of four viewpoints, spaced evenly around the viewing sphere.

The subjects were trained for a minimum of 30 trials, until they reached a 90% correct-response performance level (computed on the trailing 30 trials of the session). They were then tested (for 360 trials) on shapes obtained by varying the prototypical parameters both orthogonally (ORTHO) and in parallel (PARA) to the line connecting the prototypes in the parameter space, as described above.

2.3 Results

Eight of the 11 subjects who participated in the experiment performed above chance in the test phase (the mean error rate of these was 23%). For these subjects, the error rate (computed over the four test views and the three repetitions per condition) increased with the ORTHO parameter-space displacement between the stimulus and the corresponding prototype. A General Linear Models analysis (using procedure GLM; SAS, 1989) showed this effect to be significant: $F(2,63) = 2.9$, $p < 0.08$. The effect of the PARA displacement was close to nil: $F(2,63) = 1.7$, $p < 0.19$.[3]

A stronger effect was masked by the large individual differences (the error rates of the eight subjects ranged between 4% and 34%). When these were taken into account (by incorporating the subject variable into the analysis), the effect of ORTHO displacement became stronger: $F(2,28) = 5.0$, $p < 0.01$, and a significant effect of PARA displacement emerged:

---

[3] In the computation of these effects, we collapsed the data over the two directions of ORTHO shift away from the prototype p_2, due to considerations of symmetry. Thus, the number of degrees of freedom in the F statistics was 2 in both cases.

$F(2,28) = 3.3$, $p < 0.05$. Importantly, there was no interaction between these effects and that of subject. A direct comparison of the sums of squares computed by GLM indicated that the effect of ORTHO ($SS = 998.0$) was somewhat stronger than that of PARA ($SS = 657.8$). Figure 5 shows the mean error rate and response time, plotted against the ORTHO and the PARA displacement.

2.4 Discussion

The results of the first experiment clearly indicate that the parameter-space location of the stimuli mattered more than the qualitative contrasts (which were always present) between the two classes that had to be distinguished. Moreover, the stronger effect of the ORTHO relative to the PARA displacement suggests that both prototypes participated in determining the response to the test stimuli. This pattern exactly mimics the distinction between the ORTHO and the INTER/EXTRA effects in the experiments described in (Bülthoff and Edelman, 1992), where the subjects' generalization performance was worse for views extending in the ORTHO direction, compared to the other two. We stress again, however, that there the manipulation of the stimuli was carried out in the view space (that is, the exemplars were rotated versions of the "prototype"), whereas in the present experiment the manipulation was in the shape space (the exemplars differed from the prototype by their shape).

3 The NEAREST-NEIGHBOR experiment

The next experiment we describe was designed to gain further support for the idea that proximities to *both prototypes* contribute to the categorization process. If the visual system indeed relies on the computation of distances between the stimulus and prototypical memory traces in some feature space, it may use those distances in two different manners. The first possibility is that the identity of the prototype nearest to the stimulus is the sole determinant of the response. We term this the NEAREST-NEIGHBOR hypothesis; nearest neighbor decision rules are ubiquitous in computer vision approaches to recognition, and constitute a central, albeit seldom acknowledged, assumption behind computational theories of human object recognition. The second possibility is that a number of close neighbors of the stimulus jointly determine the nature of the response.

To distinguish between these two possibilities, we examined the responses of the subjects to a fixed set of stimuli, while manipulating the location of one of the two class prototypes (see Figure 6). Importantly, the manipulation left the test stimuli themselves always within the half-space dominated (by proximity) by the other, fixed, prototype. According to the NEAREST-NEIGHBOR hypothesis, the performance on the test stimuli should not change under the proposed manipulation. In contrast, theories that postulate the involvement of all sufficiently close prototypes (Poggio and Edelman, 1990) predict that performance will vary as the distance between the two prototypes changes (Edelman, 1995a). Assuming that the system's response is based on matching the pattern of prototype activations elicited by the stimulus to patterns memorized during training, the expected changes in performance are to the worse, as illustrated in Figure 6.

3.1 Method

The stimulus set and the course of each trial were the same as in the first experiment. Of the 15 stimuli associated with prototype p_2, only the five belonging to the middle column were used (see Figure 6). These were crossed with three possible locations of prototype p_1, which we termed FAR, INTERMEDIATE, and NEAR, yielding 15 test conditions (as in the first experiment).[4] Note, however, that the subject's performance was always assessed on the same five physical objects that belonged to p_2 class. The results reported below only pertain to the responses given to those objects.

3.2 Results

Thirteen of the 18 subjects who participated in this experiment performed above chance (mean error rate of these was 18.0%). When averaged over these subjects, the effect of moving prototype p_1 was marginal: $F(2, 180) = 1.6$, $p = 0.2$.[5]

As in the previous experiment, a stronger effect was masked by the large individual differences (the error rates of the 13 subjects ranged between 5% and 34%). When these were taken into account (by incorporating the subject variable into the analysis), the effect of moving p_1 became more significant: $F(2, 96) = 2.9$, $p < 0.06$; importantly, there was no interaction between this effect and that of subject. Pairwise contrasts among the levels of this effect indicated n.s. differences between levels 0 and 1 ($F = 1.67$, $p = 0.2$) and levels 0 and 2 ($F = 1.19$, $p = 0.28$), and a significant difference between levels 1 and 2 ($F = 5.68$, $p < 0.02$; these two levels correspond to the intermediate and the farthest locations of prototype p_1, relative to p_2). Figure 7 shows the mean performance in this experiment, plotted against the PARA displacement of prototype p_1.

---

[4] The location of p_1 in the test phase was manipulated by dividing the entire sequence of test trials into three blocks, each of which corresponded to FAR, INTERMEDIATE, or NEAR condition. The order of the blocks in the test session was randomized across subjects.

[5] Because this experiment concentrated on the effect of the displacement of prototype p_1 in the PARA direction, the data were collapsed over the ORTHO displacement.

3.3 Discussion

The results of this experiment demonstrate the sensitivity of the visual system to the general setting of the categorization task with which it is confronted (cf. the recent review of the role of the task in categorization in Schyns, 1998). If the classification decision were carried out by comparing a representation of the stimuli (which remained fixed throughout the experiment) with that of the closest class prototype (which remained fixed as well), a constant performance would have ensued. We found, however, that the performance has been affected by the relocation of the second prototype relative to the first (closest) one, in clear violation of the prediction of the NEAREST NEIGHBOR hypothesis.

4 General discussion

The objective of the two experiments we described was to gather quantitative data regarding the process whereby the shape of an object is labeled as belonging to one of two classes. To minimize the effects of prior familiarity with the stimuli, we used novel shapes, generated to tight specifications. Although this choice allowed us a high degree of control over the stimulus set, the resulting task turned out to be too difficult for eight out of twenty-seven subjects, who failed to perform above chance during the test phase. This difficulty should be kept in mind while discussing the implications of our findings.

4.1 On the interdependence of the dimensions of the representation space

In the study of visual object processing, the foremost issue is the nature of the representation space wherein tasks such as categorization are carried out. To clarify this issue, we asked, specifically, whether the dimensions of the relevant space are independent. The results of experiment 1 suggest that they are not: our subjects performed worse on shapes that were progressively more different from the class prototype acquired during the training phase, even though this difference was orthogonal to the (qualitative) distinction between the two classes. This finding is consistent with the results of a recent series of studies of categorization in pigeons and people (Mackintosh, 1995; see also McLaren et al., 1994). In those studies, introducing a modification into the feature-space (in our terms, the shape-space) location of the stimuli was shown to affect performance, even when the modification was orthogonal to the learned categorization. Thus, our results indicate that the visual system does not always rely solely on the single most distinctive contrast between the categories, even if this contrast is "qualitative" or "nonaccidental" (using the terminology adopted by Biederman, 1987).

In an unpublished manuscript, Bar and Biederman (1995) claim that a comparison of the effects of nonaccidental and quantitative shape changes on orientation invariance must be accompanied by a proper scaling of the two.[6]. Our main goal is to investigate the effects of shape change along a dimension that is orthogonal to a predefined shape contrast between two class prototypes (rather than to compare numerically the effects of qualitative and

---

[6] Bar and Biederman attempted such scaling by varying the degree of shape change in each case so as to obtain a fixed change in performance at the reference orientation

quantitative manipulations). Scaling, therefore, would only be a concern here if in our stimulus set the nonaccidental contrast were infinitesimal and the other one not. To dispel this concern, we have computed the relevant image-space distances between corresponding views of the stimuli (shown in Figure 8). The distance between images of the two class prototypes, p_1 and p_2, was $d_{\text{diff class}} = 14.95$.[7] In comparison, the distance between prototype p_2 and an outlying exemplar of its class (top row, middle column in Figure 8, left) was $d_{\text{same class}} = 7.21$. Neither of these distances is infinitesimal compared to the other. Thus, we can recapitulate our conclusion: the shape change underlying $d_{\text{same class}}$ (which is only half as large as $d_{\text{diff class}}$) resulted in a significant deterioration in the classification performance, despite being orthogonal to the learned classification task.

4.2 On the shortcomings of "nearest-neighbor" models

Even if the location of the stimulus in a shape representation space — and not merely its location along the line connecting the two class prototypes in that space — determines the subject's performance, the question of the nature of the shape space (that is, the nature of the relevant features) still remains open. Rather than attempting to characterize the features explicitly (an undertaking that is notoriously resistant to a purely psychological approach), we chose to find out whether or not the features that the stimulus shares with the closest prototype alone determine the performance. The outcome of the second experiment reported above suggests that both prototypes contribute to the perceptual categorization decision.

---

[7] The units here are immaterial. Note that this distance provides an objective measure of the smallest magnitude of the dissimilarity between the two objects, attained when the objects are aligned with each other. When averaged over different orientations, this distance is essentially the same for all pairs of shapes that are as similar to each other as our stimuli.

An intriguing computational hypothesis consistent with our findings holds that the internal shape space is spanned by a vector of proximities of the stimulus to a number of "reference" or prototypical objects, whose role can be played by the class prototypes (Edelman and Duvdevani-Bar, 1997; Cutzu and Edelman, 1998). The implications of this hypothesis, according to which the features by which an object is judged are its similarities to other objects,[8] as well as a discussion of its compatibility with recent psychophysical and neurobiological findings on object representation, can be found in (Edelman, 1998).

In summary, the results of the two experiments we reported above contradict the prediction of theories of recognition that postulate exclusive reliance on qualitative contrasts (Biederman, 1987) or on proximity to a decision surface (Ashby and Perrin, 1988; Maddox and Ashby, 1993), and support the notion of a metric representation space, with the subjects' performance determined by proximities to more than one reference point or prototype (Nosofsky, 1988; Nosofsky, 1991; Kruschke, 1992; Edelman, 1995b; McKinley and Nosofsky, 1996; Edelman, 1998).

Acknowledgments

We thank Michael J. Tarr for his comments on the ideas that led to these experiments. This work has been supported in part by the German-Israeli Foundation, through a grant to Heinrich H. Bülthoff, Nathan Intrator, and Shimon Edelman.

---

[8] Note that more than one prototype is needed to pinpoint the location of the stimulus in the shape space; thus, the computational scheme based on this notion is called the Chorus of Prototypes.

References

Ashby, F. G. and Perrin, N. A. (1988). Toward a unified theory of similarity and recognition. *Psychological Review*, 95(1):124–150.

Bar, M. and Biederman, I. (1995). One-shot viewpoint invariance in matching novel objects. unpublished manuscript.

Biederman, I. (1987). Recognition by components: a theory of human image understanding. *Psychol. Review*, 94:115–147.

Biederman, I. and Gerhardstein, P. C. (1993). Recognizing depth-rotated objects: evidence and conditions for 3D viewpoint invariance. *Journal of Experimental Psychology: Human Perception and Performance*, 19:1162–1182.

Bülthoff, H. H. and Edelman, S. (1992). Psychophysical support for a 2-D view interpolation theory of object recognition. *Proceedings of the National Academy of Science*, 89:60–64.

Bülthoff, H. H., Edelman, S., and Tarr, M. J. (1995). How are three-dimensional objects represented in the brain? *Cerebral Cortex*, 5:247–260.

Cutzu, F. and Edelman, S. (1994). Canonical views in object representation and recognition. *Vision Research*, 34:3037–3056.

Cutzu, F. and Edelman, S. (1998). Representation of object similarity in human vision: psychophysics and a computational model. *Vision Research*, 38:2227–2257.

Edelman, S. (1995a). Class similarity and viewpoint invariance in the recognition of 3D objects. *Biological Cybernetics*, 72:207–220.

Edelman, S. (1995b). Representation, Similarity, and the Chorus of Prototypes. *Minds and Machines*, 5:45–68.

Edelman, S. (1998). Representation is representation of similarity. *Behavioral and Brain Sciences*, 21:449–498.

Edelman, S. and Bülthoff, H. H. (1992). Orientation dependence in the recognition of familiar and novel views of 3D objects. *Vision Research*, 32:2385–2400.

Edelman, S. and Duvdevani-Bar, S. (1997). Similarity, connectionism, and the problem of representation in vision. *Neural Computation*, 9:701–720.

Humphrey, G. K. and Khan, S. C. (1992). Recognizing novel views of three-dimensional objects. *Can. J. Psychol.*, 46:170–190.

Jolicoeur, P. (1990). Identification of disoriented objects: a dual-systems theory. *Mind and Language*, 5:387–410.

Jolicoeur, P. and Humphrey, G. K. (1998). Perception of rotated two-dimensional and three-dimensional objects and visual shapes. In Walsh, V. and Kulikowski, J., editors, *Perceptual constancies*, chapter 10. Cambridge University Press, Cambridge, UK. in press.

Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, 99(1):22–44.

Lawson, R., Humphreys, G., and Watson, D. G. (1994). Object recognition under sequential viewing conditions: evidence for viewpoint-specific recognition procedures. *Perception*, 23:595–614.

Lawson, R. and Humphreys, G. W. (1996). View specificity in object processing: evidence from picture matching. *Journal of Experimental Psychology: Human Perception and Performance*, 22:395–416.

Mackintosh, N. J. (1995). Categorization by people and pigeons: The twenty-second Bartlett memorial lecture. *Quarterly Journal of Experimental Psychology*, 48B:193–210.

Maddox, W. T. and Ashby, F. G. (1993). Comparing decision bound and exemplar models of categorization. *Perception and Psychophysics*, 53:49–70.

McKinley, S. C. and Nosofsky, R. M. (1996). Selective attention and the formation of linear decision boundaries. *Journal of Experimental Psychology: Human Perception and Performance*, 22:294–317.

McLaren, I. P. L., Leevers, H. J., and Mackintosh, N. J. (1994). Recognition, categorization, and perceptual learning (or, how learning to classify things together helps one to tell them apart). In Umiltá, C. and Moscovitch, M., editors, *Attention and Performance*, volume XV, chapter 35, pages 889–909. MIT Press.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 14:700–708.

Nosofsky, R. M. (1991). Tests of an exemplar model for relating perceptual classification and recognition memory. *Journal of Experimental Psychology: Human Perception and Performance*, 17:3–27.

Poggio, T. and Edelman, S. (1990). A network that learns to recognize three-dimensional objects. *Nature*, 343:263–266.

Rock, I. and DiVita, J. (1987). A case of viewer-centered object perception. *Cognitive Psychology*, 19:280–293.

SAS (1989). *User's Guide, Version 6.* SAS Institute Inc., Cary, NC.

Schyns, P. G. (1998). Diagnostic recognition: task constraints, object information, and their interactions. *Cognition*, 67:147–179.

Tarr, M. J., Bülthoff, H. H., Zabinski, M., and Blanz, V. (1997). To what extent do unique parts influence recognition across changes in viewpoint? *Psychological Science*, 8:282–289.

Tarr, M. J. and Pinker, S. (1989). Mental rotation and orientation-dependence in shape recognition. *Cognitive Psychology*, 21:233–282.

Ullman, S. and Basri, R. (1991). Recognition by linear combinations of models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13:992–1005.

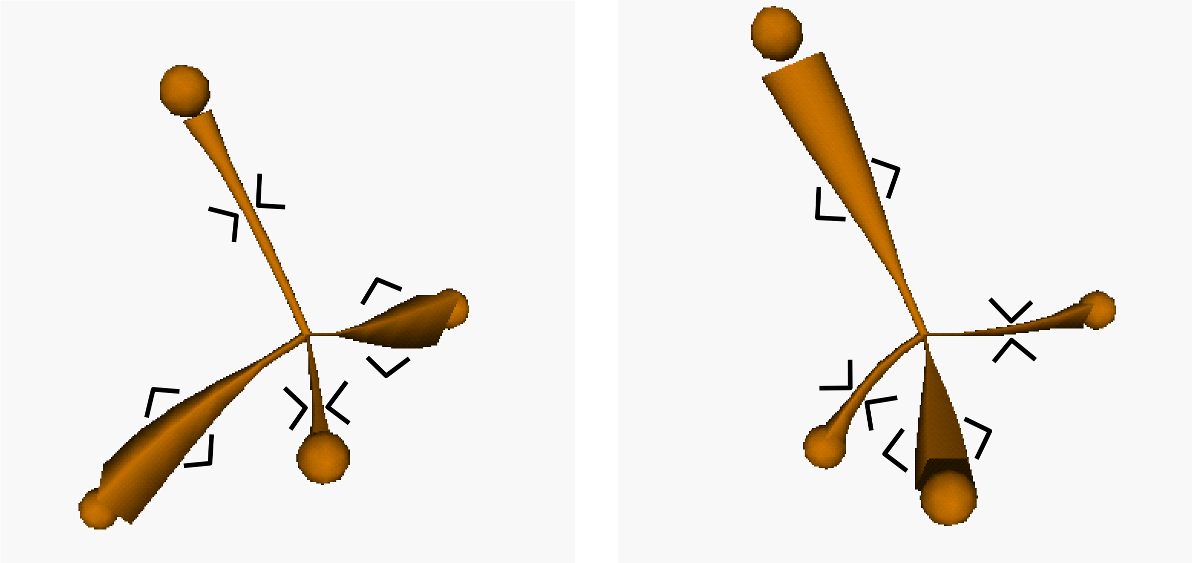

Figure 1: Two prototypical **hedgehog**-like objects, similar to those we used in our experiments (both "hedgehogs" are shown at the same orientation). Left: p_1 class, right: p_2 class. Each object is composed of a number of limbs protruding from a common center; the limbs are generalized cylinders, similar to Biederman's (1987) geons. The two prototypes are distinguished by qualitative (nonaccidental) contrasts (sign of bulge $<>$/waist $><$ of the limbs; see Figure 3). In addition, a number of quantitative parameters such as the degree of bulge/waist, the amount of taper, length, etc., control the exact shape of each instance object. Note that the qualitative contrasts emerge from the accumulation of quantitative changes, as illustrated in Figure 2.

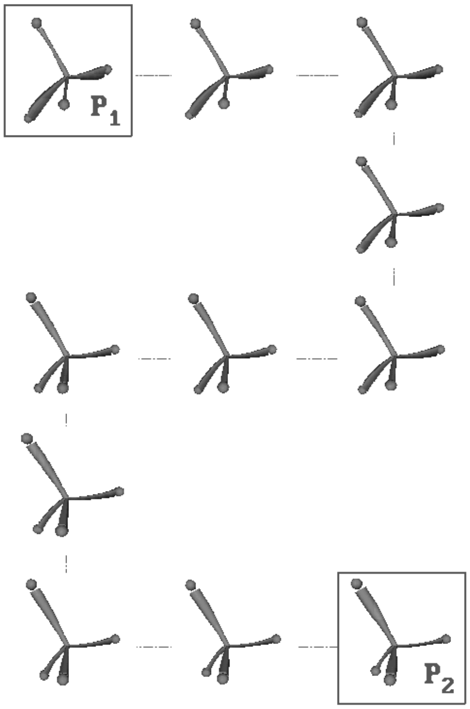

Figure 2: Each object can be represented by a single point in a parameter (shape) space. Changing the quantitative parameters ("morphing") corresponds to a movement of the shape-space representation of the object. This figure illustrates the morphing sequence that connects the two prototype objects. Although the changes between the successive images are minute, they accumulate to make up easily perceptible (and, eventually, "qualitative") differences between the endpoints of the sequence.

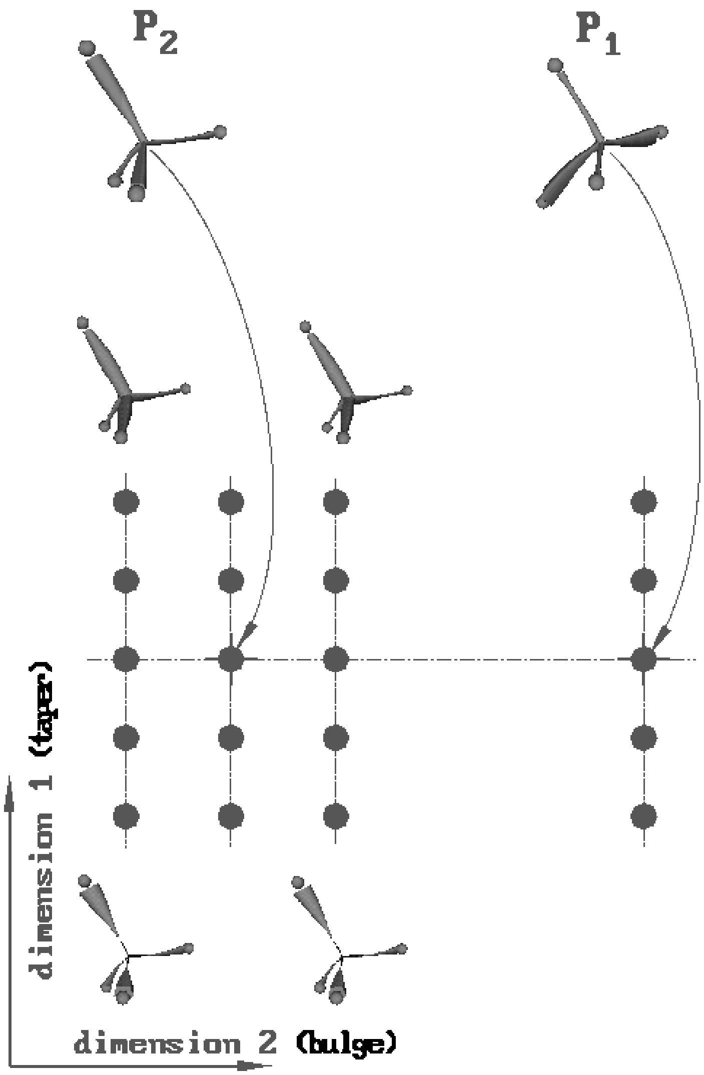

Figure 3: The parameter-space arrangement of stimuli. The parameter-space locations of the two prototypical objects are marked by p_1 and p_2. The two orthogonal directions of shape variation are bulge (increase/decrease) and taper (proximal to distal or vice versa). Specifically, the shift from p_1 to p_2 corresponds to a gradual change from a waist-like to a bulging profile of the hedgehog's limbs or vice versa; the orthogonal direction corresponds to an equally gradual change of limb shape that tapers from the proximal towards the distal end to a shape that tapers in the opposite direction. See Figure 8 for an illustration of the entire array of stimuli corresponding to this parameter-space pattern.

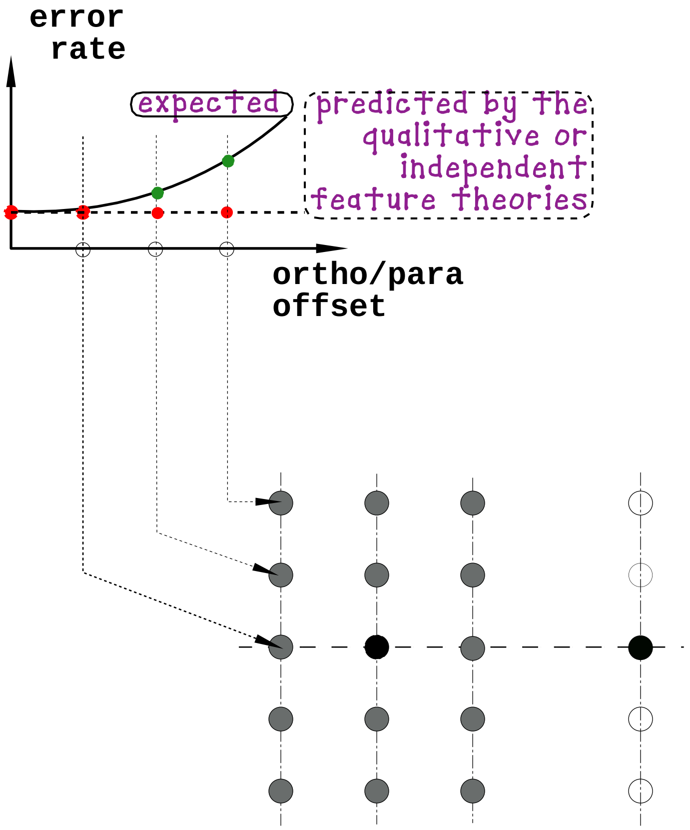

Figure 4: The parameter-space arrangement of the stimuli and the expected performance in the Ortho experiment. The subjects were trained to discriminate between the two prototypes, p_1 and p_2. They were then tested on the discrimination of stimuli produced by a shape-space variation orthogonal to the contrast between the two prototypes. Black dots: prototypes, grey dots: p_2 class objects used in the experiments, white dots p_1 class objects (not used). See section 2 for a discussion of the predicted results and the actual findings.

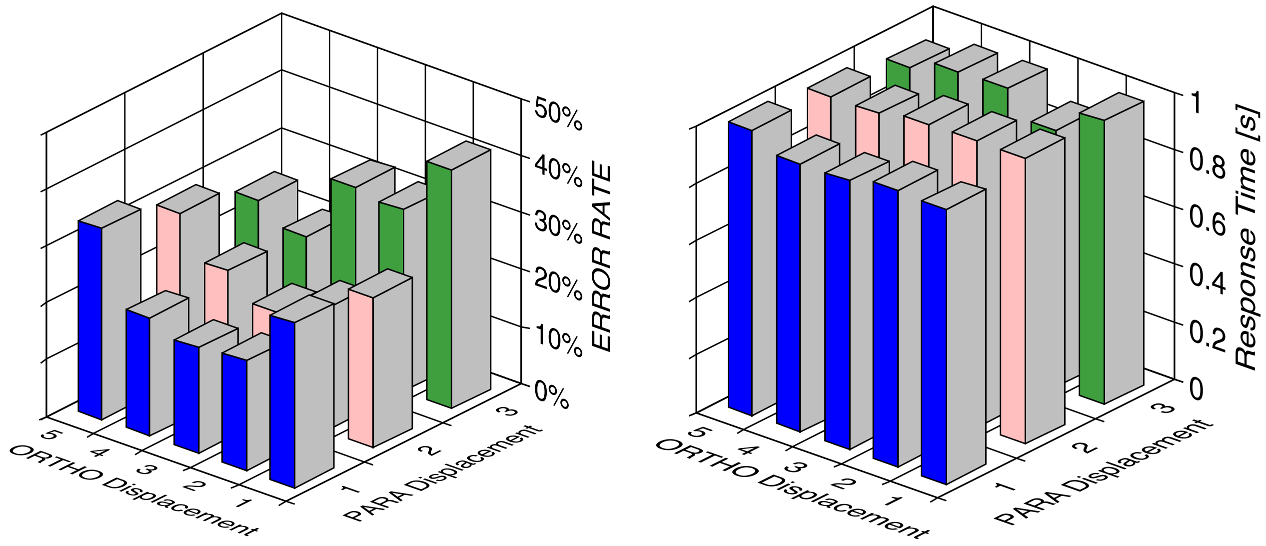

Figure 5: The mean performance of the eight subjects who responded above chance in the first experiment, plotted against the Ortho and the Para displacement (see section 2.3). The three Para displacement values, denoted symbolically by the numerals 1,2,3, appear along the abscissa in the plots; the five Ortho values correspond to the ordinate. The location of prototype p_2 corresponds to the point whose coordinates are $(2, 3)$. Altogether, the 15 data points are arranged in a 3×5 grid around prototype p_2 (see Figure 3); the direction towards the other prototype in these plots is along the increasing abscissa values. *Left:* error rate. Note the general increase in the error rate for test stimuli that are closer to the other prototype. *Right:* response time.

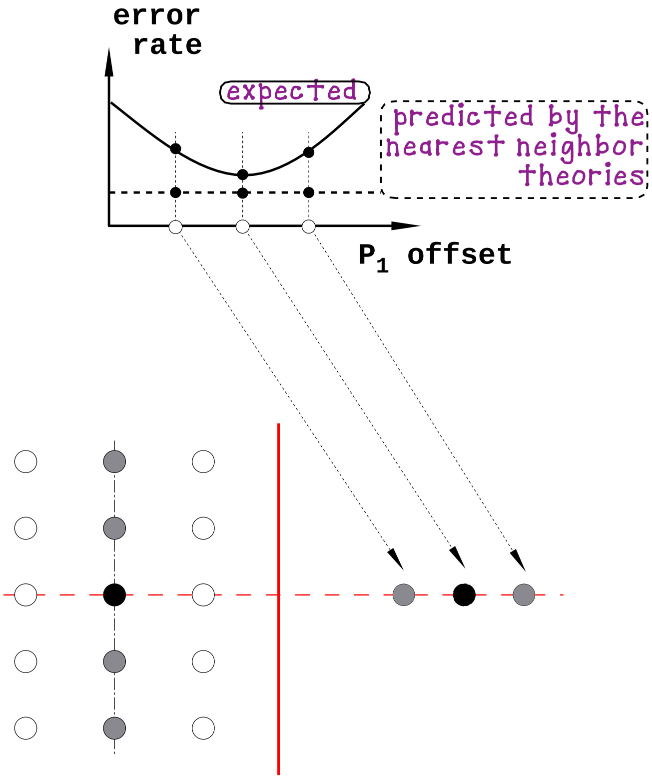

Figure 6: The parameter-space arrangement of the stimuli and the expected performance in the NEAREST-NEIGHBOR experiment. The two prototypes, p_1 and p_2, are as before. In this experiment, the location of p_1 relative to p_2 varied along the line connecting the two prototypes. Performance (discrimination between the two classes) was tested for the same physical stimuli, whose location in the parameter space corresponds to the middle column in the 3×5 grid of points surrounding p_2. Black dots: prototypes, grey dots: p_2 and p_1 class objects used in this experiment, white dots: p_2 class objects not used here. For a discussion of the expected performance, see section 3.

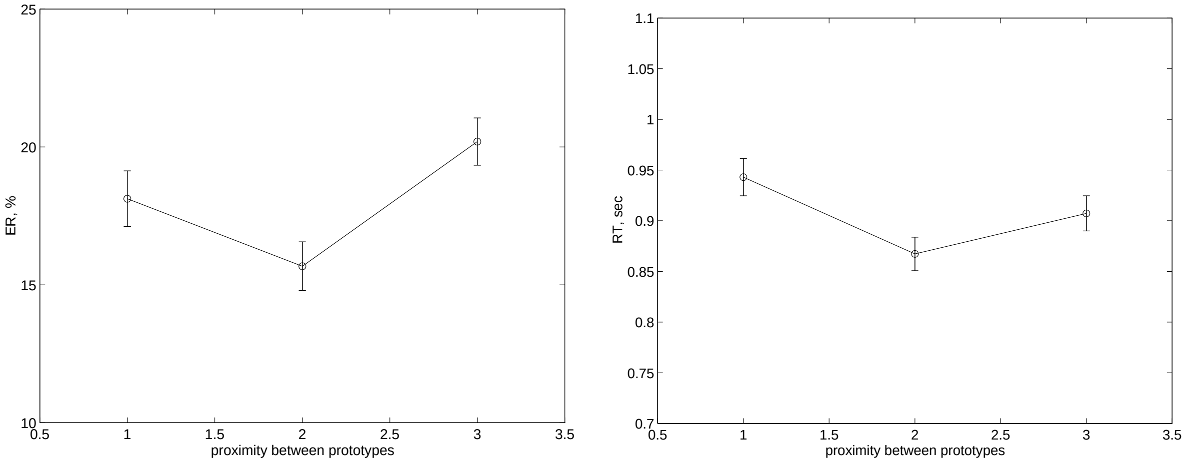

Figure 7: The mean performance of the 13 subjects who responded above chance in the second experiment, plotted against the Para displacement of prototype p_1 (see section 3.2). *Left:* error rate. *Right:* response time. The error bars show the standard deviation of the corresponding means. The three values along the abscissa (prototype proximities 1, 2, 3) correspond, respectively, to the Far, Intermediate, and Near conditions, described in the text.

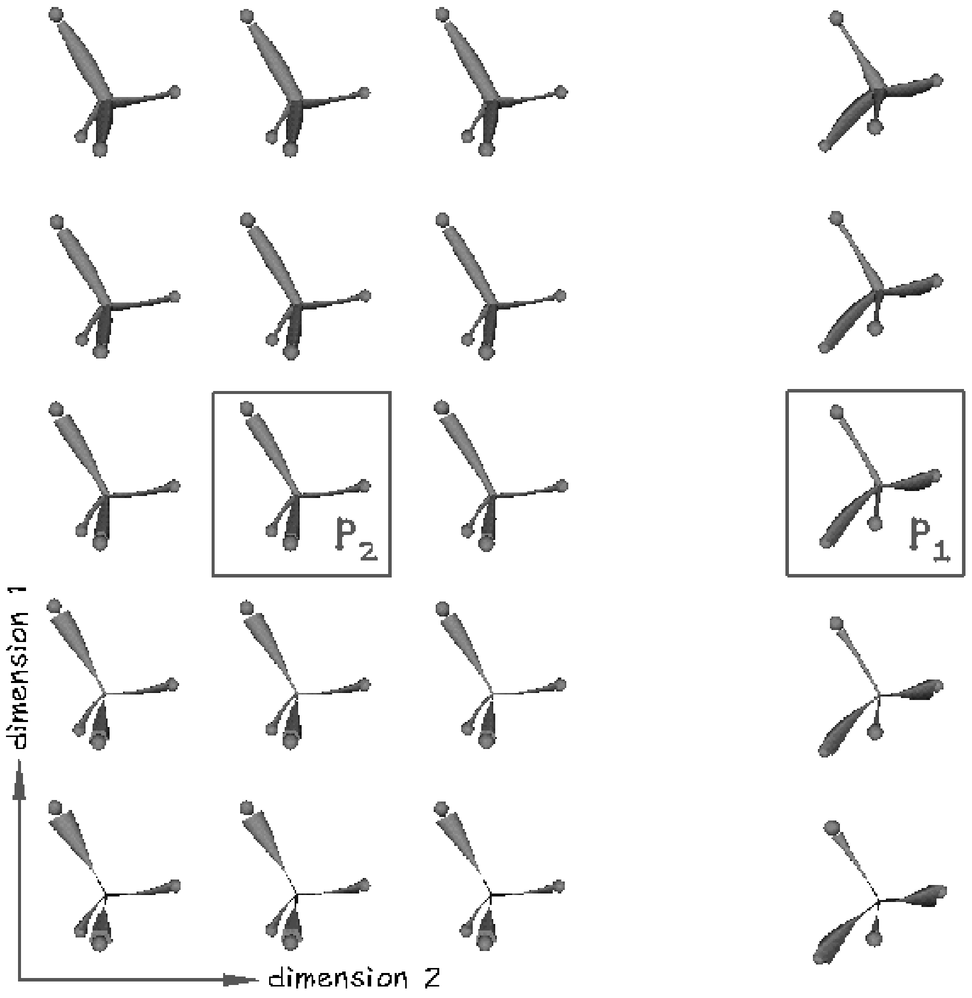

Figure 8: All the stimuli objects. The perception of variants of the learned prototype objects was probed with 15 exemplars made out of prototype p_2. All these exemplars, whose shape-space locations form a 3×5 grid centered on p_2, are illustrated here.