Perception, 2011, volume 40, pages 1290 – 1308

doi:10.1068/p6970

The role of constant curvature in 2-D contour shape representations

Patrick Garrigan, Philip J Kellman§
Department of Psychology, Saint Joseph's University, 5600 City Avenue, Philadelphia, PA 19131, USA; e-mail: pgarriga@sju.edu; § Department of Psychology, University of California, Los Angeles, CA, USA Received 25 March 2011, in revised form 5 September 2011

Abstract. In early cortex, visual information is encoded by retinotopic orientation-selective units. Higher-level representations of abstract properties, such as shape, require encodings that are invariant to changes in size, position, and orientation. Within the domain of open, 2-D contours, we consider how an economical representation that supports viewpoint-invariant shape comparisons can be derived from early encodings. We explore the idea that 2-D contour shapes are encoded as joined segments of constant curvature. We report three experiments in which participants compared sequentially presented 2-D contour shapes comprised of constant curvature (CC) or non-constant curvature (NCC) segments. We show that, when shapes are compared across viewpoint or for a retention interval of 1000 ms, performance is better for CC shapes. Similar recognition performance is observed for both shape types, however, if they are compared at the same viewpoint and the retention interval is reduced to 500 ms. These findings are consistent with a symbolic encoding of 2-D contour shapes into CC parts when the retention intervals over which shapes must be stored exceed the duration of initial, transient, visual representations.

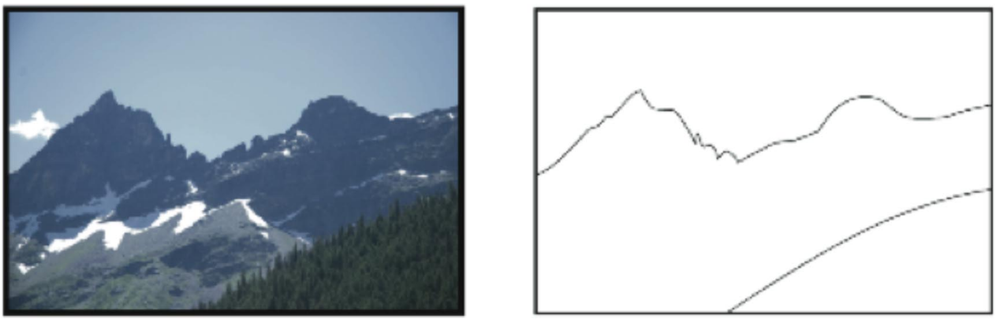
1 Introduction

Among the most important functions of human vision are perception and representation of shapes of objects. At a coarse scale, we use shape to recognize and classify objects, as in determining that something is a suitcase. At finer scale, shape is used to distinguish objects of the same category, as in finding your suitcase among others. Many studies have contributed to our understanding of shape perception, but deep questions remain. Objects can be identified in isolation (with no context) from their shapes without information about color, texture, and reflectance (Wagemans et al 2008). We know, therefore, that the visual system can represent, at some level, the contiguous space that individual objects occupy, independent of other features. Exactly how the visual system does this is unclear. Much discussion has focused on whether shape representations are built from a set of primitives in 'structural models' of object recognition (Marr and Nishihara 1978; Biederman 1987), or from stored views of the objects and a model of the spatial transformations among them (Poggio and Edelman 1990; Riesenhuber and Poggio 1999).

Structural and view-based approaches address important aspects of object recognition, but other more basic issues of shape representation underlie both. In a view-based approach, what are the interpreted parts of images? How are bounding contours represented and distinguished from surface features? In structural description models, efforts have focused on object recognition, usually at a categorical level (eg Biederman 1987). Two cups with different contour curvatures may both be encoded as cups, but what is the code that allows one to choose among a set of cups with subtle differences in shape? Subtle distinctions of this sort require a flexible representation of the shapes of contours and surfaces. How can the shape of a contour presented in isolation be compared to part of the shape of the bounding contour of a complete object? If the parsing into parts takes into account the whole object, then even if part of a full contour and a contour segment matched exactly, their parts-based representations may not.

The point is that, even when very simple shapes (or parts of shapes) are compared, the same questions about representation crop up. Even simple shapes need to be represented in a symbolic code that allows matching across variations in the inputs from which shape is extracted. This code would be realized as distinct patterns of neural activity corresponding to features of the visible stimulus, and must allow for the flexible invariant shape comparisons that humans perform with little conscious effort.

Consider the mountain scene in figure 1a. After viewing for a few moments, we likely have some representation of the shape of the contour of the mountains and also of the forested hilltop in the foreground. It is unlikely, however, that we have represented the details of every point along these contours. The contours in figure 1b may capture much of the gist and would likely be recognizable as a close match in sequential viewing. They are far from a detailed match, however. Note, for example, the lower right parts of the two panels. The contour in the original scene is composed of jagged treetops; yet we somehow extract an overall simplified representation suitable for remembering and recognizing—something more like the simple contour drawn in figure 1b. How we get from the massively detailed information available while we view a scene to a more schematic, symbolic shape representation is an example of the deep and unsolved questions motivating the present work.



(a) (b)

Figure 1. [In color online, see http://dx.doi.org/10.1068/p6970] A mountain scene in Glacier National Park, Montana (a) and corresponding schematic contours depicting these mountains (b).

In this paper we consider shape comparison between pairs of very simple shapes—2-D contour segments. 2-D contours are useful for highlighting a crucial issue in shape perception generally: how do we obtain symbolic descriptions from sub-symbolic encoding in the early stages of visual processing (Kellman and Garrigan 2006)? The problem is quite general and often neglected. The experiments and models in this paper may serve to focus the problem and show what a solution might look like in at least one context.

1.1 *Sub-symbolic and symbolic descriptions*

One function of shape encoding is to reduce complex shape information in the world into simpler compact representations more suitable for recognition and comparison. Forming this code involves a transition from sub-symbolic to symbolic representations, an issue that for vision science might be posed in a neurophysiological context. The earliest stage of visual processing of shape in cortex involves cells sensitive to oriented contrast in local regions of the visual field (eg Hubel and Wiesel 1959). These early representations are much like an uninterpreted pixel map, in that incidental changes to the input that may greatly affect the activity of individual cells often have little perceptual consequence. Two identical shapes viewed from different positions in space or formed from different elements will generate very different neural activity at early stages of visual processing. However, they must converge into a symbolic neural representation that allows us to say that they are the same. This transformation, from a sub-symbolic to a symbolic representation, is not well understood.

The Gestalt psychologists were among the first to emphasize the gap between form descriptions and local sensory elements, noting among other things that shape is a relational property and similar shapes can be given by different elements (eg Koffka 1935). To say a cloud resembles a fish requires a shape code that makes explicit geometric information that is not explicit in the local image information. Consider a local pixel map representation of a fish. Attempting to find any meaningful correspondence between a representation of this sort and the shape of a cloud would be futile. Many features of the pixel map would not only be irrelevant to making a shape match, but they would actually impede detecting the relevant similarity. A cloud that resembles a fish will do so only in a global and abstract sense. Locally, at the level of the features in the image, matches and mismatches will have little relevance.

1.2 *Information, efficiency, and simplicity*

For the purposes of shape comparison, a symbolic code transforms shape information into a more useful format that is invariant to incidental viewing conditions. However, this is only one of the issues that a representation of shape must address in order to plausibly account for human shape recognition performance. Shape recognition requires that shape representations be stored for both short and long periods of time in visual memory. It is widely accepted that both short and long-term visual stores have significant capacity limitations, and so these symbolic representations must also be efficient.

The efficiency of encoded information was formalized in Shannon's (1948) theory of communication, and has since been developed in the theory that the brain evolved to most efficiently encode behaviorally important sensory information (eg Barlow 1961). Psychologists studying perceptual processes were also quick to adopt information theoretic techniques. Within the domain of shape, it was Attneave (1954) who first illustrated both that different parts of shapes appeared to convey different amounts of information, and therefore that shapes had inherent redundancy. This was demonstrated most famously by the drawing of Attneave's cat, in which straight lines connecting the points of maximal curvature replace the full drawing of a cat. The vastly simpler drawing is still obviously a cat, and therefore the information that was lost was clearly not necessary for object identification at the basic categorical level (although it would likely be necessary for finer discriminations).

For our purposes, the physical world contains shape information that is communicated to, and represented by, the brain. Representation in the brain is subject to capacity limitations and noise that limit how much the brain can learn about the physical world from the sensory input. One important factor that influences how resources can be most efficiently allocated is redundancy among signals, which in our case are shapes. If shapes tend to have similar characteristics (eg convexity), or if part of a shape can be predicted by a nearby part of that shape (eg smooth contours), then this redundancy can be leveraged to form more efficient representations.

Unfortunately, sampling the distribution of the features of shapes has proven to be a difficult problem. Feldman and Singh (2005) applied information theory to contour-shape representation within a limited context. They proposed coding 2-D contour shapes in terms of the local curvature, and then defined the 2-D shape surprisal as the information gained by encoding some segment of the contour. Their measure uses a probability distribution of change in orientation along a contour, with no change having highest probability. Regions of high curvature are most informative (because they are least probable) and straight lines are the least informative (because they are most probable). In the case of closed contours, the most probable curvature is shifted in the direction closing the contour, and consequently concave minima have greater information than equivalent convex maxima. They relate this measure to the contour segmentation algorithm proposed by Hoffman and Richards (1984), and subsequent psychophysical evidence suggesting that concave minima are perceptually important for shape segmentation (Hoffman and Singh 1997; Singh et al 1999).

This measure does not incorporate non-local characteristics of shapes, which would be required for fully measuring the information content of a contour. The Gestalt notion of simplicity (Koffka 1935; Hochberg and McAlister 1953), however, tells us that some non-local relations among features are perceptually important. 'Simplicity' is realized by explicitly representing a relation that is present in the stimulus. The apparent behavioral significance of non-local spatial relationships, like symmetry, good continuation and closure, suggest that, however difficult it may be to measure the probability of these relationships from natural images, the visual system incorporates them into representations of shape.

Simplicity in perception, such as the Gestalt rules, express a different kind of efficiency. They do not require sampling statistics in the physical world; rather, they suggest that perceptual organizations are compact. Early characterizations of what made a figure simple were primarily intuitive (eg Wertheimer 1938). Kohler (1938) proposed a more principled explanation of perceptual simplicity based on isomorphism between physical forces in the world and 'brain fields', but these ideas were quite general and have received little scientific specification or support. Later approaches gave more precise formal accounts of simplicity (eg Hochberg and McAlister 1953). Coding theory (Leeuwenberg 1971) and structural information theory (SIT; van der Helm and Leeuwenberg 1991) formalized simplicity by defining operations to describe visual patterns, so that for a given pattern the simplest of multiple possible descriptions is the preferred perceptual encoding. Van der Helm and Leeuwenberg (1991) suggested that iteration, symmetry, and alternation are three regularities the visual system uses to encode visual arrays efficiently. These relations form the basis of SIT because they are especially accessible—an argument similar to the one used to justify the set of geons (Biederman 1987).

The notion of efficiency embodied in simplicity approaches such as SIT avoids the ill-defined problem of sampling the joint probability of all the features of all shapes. It has been often applied to simplified patterns (eg a series of dots—van der Helm et al 1992; joined line segments—Hanssen et al 1993) that are useful for demonstrating that the visual system is sensitive to particular relations. It has also been tested using a variety of tasks, such as judging how three-dimensional an object looks (Butler 1982), choosing among a set of possible segmentations of patterns (van der Helm et al 1992), and explicitly judging pattern complexity (Hanssen et al 1993). SIT has been shown to predict the hidden parts of objects in 2-D occlusion displays (van Lier 1999) and the symmetries employed in various classes of artwork (van der Helm 2011).

1.3 *Task dependence*

The foregoing suggests that more than one kind of efficiency is relevant to shape coding: utilizing probable characteristics of shapes in the world and representing shapes using simple or compact codes. One complicating issue that remains, however, is that all of these factors may depend, in part, on the task to which the representations will be applied.

The best strategy for how to form efficient representations of object shape will depend in part on task conditions, like the particular set of shapes being represented and the type of discrimination being made. Shape representations must, therefore, be adaptable to constraints imposed by the task. If at the time of encoding the observer knows the task for which visual information will be used, the visual system should encode information that is likely to be useful and discard information that is likely to be irrelevant. In some cases, this will involve making particular information explicit.

For example, an uninterpreted pixel representation of an image may be used for tasks such as reconstituting an image or transmitting it to a printer. From a more ecological perspective, detailed shape information may be required for making subtle distinctions between similar shapes, but for other tasks relatively crude shape information will suffice.

Researchers have suggested that human shape representations are hierarchically structured so that a single type of shape representation can be used for a variety of shape-related tasks that vary in their particular demands (eg Leeuwenberg and van der Helm 1991). This idea is inherent in the 'generalized cylinder' representation of 3-D objects (Marr and Nishihara 1978). As Marr (1982) discusses, at one level in the hierarchy, simpler representations can distinguish between quadrupeds, bipeds, and birds. At higher levels in the hierarchy, representations become more complex, but can be used to make subtler distinctions between eg types of quadrupeds, like cows, horses, and giraffes.

1.4 *The current research*

Below, we explore one possible representation for a limited class of shapes—2-D contours, and a single task—shape recognition. We believe this representation has the efficiency and flexibility described above. Further, it may afford a plausible account of the step from early sub-symbolic visual encodings to symbolic shape tokens in human vision. We present three experiments and modeling efforts that converge on the idea that an important kind of shape representation in human vision involves representing contour shape as segments of constant curvature. In previous work (Kellman and Garrigan 2006), we described how constant curvature segments—which can be tokens in a symbolic representation—may be derived from (sub-symbolic) detectors of local oriented-contrast known to populate early visual cortical areas, especially V1 and V2. The results presented here may thus shed light on the sub-symbolic to symbolic transition in human shape coding and on recent investigations that have found psychophysical and neurophysiological evidence for a special role of constant curvature.

1.5 *Rationale for the experiments*

To preview, all of the experiments presented in this paper involve a comparison between a previously viewed shape's representation and a currently visible shape. We provide evidence that the human visual system uses a particular code to represent and store 2-D contour shapes for the purpose of shape recognition. We further investigate the conditions under which representing shapes using this code is necessary. However, before presenting the experiments, we must first make some assumptions about how the use of a particular shape code will affect behavior in our experiments.

When comparing a visible shape to the representation of a previously presented shape, performance will suffer if important differences between the two shapes are not encoded. Consequently, the visual system is designed to maximize performance by detecting and leveraging redundancy in physical shapes to form efficient representations. However, it is not reasonable to expect that the visual system can detect and utilize all of the redundancy available in any particular shape. When a task requires forming a concise representation of shape, redundancy in shape information will only lead to increased performance if the visual system can detect and use that particular form of redundancy.

Following from this idea, the nature of the representation of shape can be probed by asking subjects to compare shapes that either have or do not have a particular type of redundancy, but are otherwise equivalent. If the visual system forms representations of shape that utilize a particular redundancy, its presence should lead to better performance on tasks that require the formation of such a representation. This theoretical approach is not novel to our experiments. Victor and Conte (2004) similarly introduced various forms of statistical structure into textures and measured how visual working memory performance changed for textures with and without particular redundancy. Using logic similar to ours, they inferred from performance differences the extent to which visual working memory made use of these statistical regularities for forming and storing representations of the textures.

1.5.1 *Stimuli.* Following Feldman and Singh, we test our hypothesis using open 2-D contours drawn in the picture plane coded in terms of local curvature. In the following three experiments, we present subjects with contours that either have or do not have a specific type of redundancy. The redundant set has contours composed of a series of circular arc segments, so that the curvature at adjacent points within each segment is the same. The non-redundant set has contours composed of a series of curved segments that are constructed so that the curvatures at any two adjacent points within the same segment are never the same. We are interested in whether or not the human visual system uses redundancy in the form of constant-curvature extents (iteration of curvature in SIT terminology) for efficient encoding and comparison of 2-D contour shapes.

Unlike studies using arrays of elements, the relational structure of these stimuli is between elements that are not spatially separated or marked by a feature (eg corners), but instead must be extracted. We employ a within-subjects design in all experiments, and the subjective report of subjects indicated they were unaware of the structure of the stimulus set (that half of the shapes were of one type, and half of another). These facts are important because they suggest that our subjects did not employ novel strategies that differed between the conditions of our experiment.

1.5.2 *Task structure.* We use sequentially presented shape comparison to investigate whether perceptual representations of single contours leverage redundancy in the spatial relationships among adjacent segments of the contour to form efficient representations of shape. Unlike detection paradigms and subjective judgments of stimulus characteristics, our task requires comparing a currently visible stimulus to the stored representation of a previously presented stimulus. This choice is motivated by the idea that the processing stage at which an efficient, compact representation is most critical is in the capacity-limited storage of visual working memory (eg Alvarez and Cavanagh 2004). Consequently, we also test the conditions under which the need to form concise representations of shape is mitigated by the use of earlier, high-capacity visual representations. Specifically, across three experiments we vary how the shapes are presented (from different viewpoints or from the same viewpoint) and the duration over which the shape information must be stored.

1.5.3 *The hypothesis.* Shapes in each set are matched in terms of the magnitude of the difference in the 'same/different' pairs. If the visual system codes shapes in a manner that is sensitive to the constant curvature relation, one stimulus set has a smaller information load relative to the other stimulus set. Our prediction is that subjects will be better able to store the shapes of contours in working memory when they are composed of regions of constant curvature than equivalent contours without this redundancy. This will lead to increased performance for the redundant set, especially when the shapes are presented from different viewpoints and separated by longer inter-stimulus durations. However, if the comparison involves shapes presented at the same viewpoint and separated by shorter inter-stimulus durations, the capacity limitations of visual working memory will not limit performance. Consequently, the ability to form a concise representation of the first shape (to be compared to the second shape) will have a reduced or eliminated effect on recognition performance. Therefore, the performance advantage for the redundant set will also be reduced or eliminated.

1.5.4 *Scope.* We limit our study to open, 2-D contours primarily because 2-D contours presented in the picture plane are less susceptible to individual biases in perceived shape due to differences between subjects in the weighting of depth cues (Todd 2004; Liu and Todd 2004), some of which can be controlled well in the lab (perspective and stereo) and others less so (focus and accommodation). Smooth, open, 2-D contours also do not, in general, have semantic associations that can influence visual memory performance (eg Verhaeghen et al 2006).

2-D contours are an important class of shapes, however, and lead in fundamental processes that precede object recognition like figure/ground assignment (Peterson and Gibson 1991) and boundary completion (Yin et al 1997). 2-D surface contours and bounding contours also carry important information about the 3-D shape of objects (Barrow and Tenenbaum 1981; Koenderink 1984). Furthermore, it is reasonable to expect that what can be learned about efficiently coding simple shapes will be related to the coding of more complex shapes.

2 Experiment 1

2.1 *Method*

2.1.1 *Subjects.* Ten undergraduates with normal or corrected-to-normal vision participated in this study. There were no additional exclusion criteria. All subjects received course credit for participation; no subjects had any other incentive or received any other compensation.

2.1.2 *Apparatus.* The apparatus for this experiment was a Macintosh computer presenting stimuli using MATLAB (The MathsWorks, Inc., Natick, MA) and the Psychophysics Toolbox (Brainard 1997; Pelli 1997). Stimuli were displayed on a standard monitor at 1152×870 pixel resolution. Subjects sat in a comfortable chair and had their head stabilized in a chin-rest positioned 78 cm from the monitor.

2.1.3 *Stimuli.* Stimuli used in this experiment were pairs of thin, white, open contours presented in sequence on a black screen. All contours were smooth (had no corners). The members of each contour pair were both either composed of joined segments of constant curvature (CC) or they both had varying curvature at all adjacent locations on each contour (NCC).

CC shapes were constructed by choosing five radii of curvature and five angular spans. All segments appeared from inspection to look like parts of a circle (ie the curvature was not high or low enough to introduce artifacts due to monitor resolution or to make the segments appear as straight lines). The magnitude of curvature was allowed to vary randomly within this range, but the polarity of the curvature was always different between adjacent segments. The angular span of each segment varied between 30° and 100°. All CC shapes were then uniformly scaled to fit within a 9.1 deg $\times$ 9.1 deg square oriented upright on the monitor.

For each shape a 'different' case was generated, to be used for same/different judgments. The 'different' case was formed from its matched pair by increasing or decreasing one segment's curvature by between 25% and 75%. The angular span of all segments remained the same because changing this parameter tended to encourage the exclusive use of global stimulus characteristics (eg the overall orientation of the figure) for making the same/different judgments.

NCC shapes were generated by non-uniform scaling of the CC shapes. To do this, the CC shapes were scaled by a factor of 1/2 along the vertical or horizontal axis, and by a factor of 3/2 along the other, orthogonal axis. By generating a single NCC shape for every CC shape, the difference between NCC shapes and CC shapes on corresponding 'different' trials was, to a reasonable extent, matched. However, the non-uniform scaling of the NCC curvature shapes caused them to end up being tall and thin or short and wide versions of the CC shapes, depending on whether the largest scaling factor was along the vertical or horizontal axis. To compensate for this added variability in aspect ratio for the NCC shapes, half of the CC shapes were uniformly scaled by a factor of $1/\sqrt{2}$. This scaling factor was used because it ensures that the average size of the CC shapes was equal to the average size of the NCC shapes. Examples of the different types of shape are shown in figure 2.

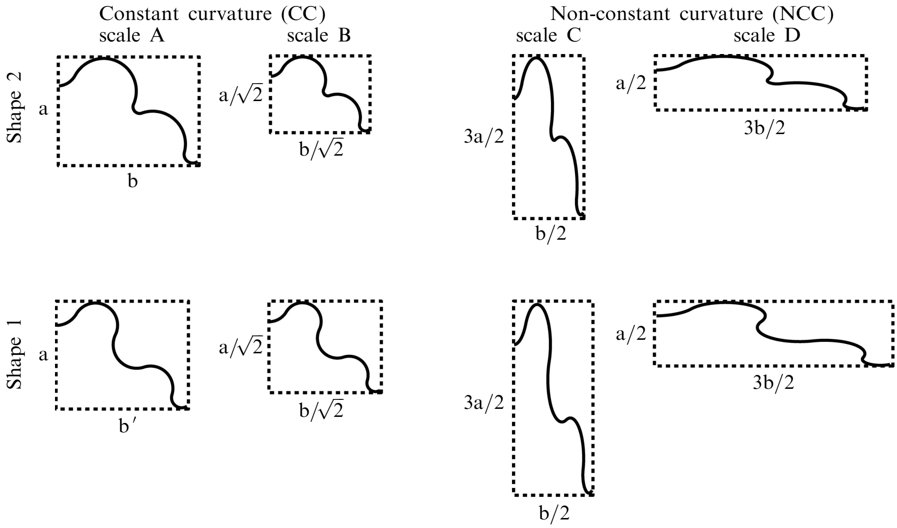


Figure 2. Stimuli: for a single same/different pair (shapes 1 and 2), all scaling factors are shown for both a constant curvature (CC) version and non-constant curvature (NCC) version of the shape. CC shapes (left) can be presented with one of two different scale factors, as can NCC shapes (right). The dotted line bounding box was not presented with the shapes; rather it is drawn here to illustrate the axes along which shapes were scaled.

2.1.4 *Procedure.* The experiment was conducted in a dark room and consisted of 60 practice trials and 320 experimental trials during which data were recorded. Before the practice session, subjects were told that they would be judging the shapes of contours and that they should indicate with a button press if two sequentially presented contours had the same or different shape. Subjects were told that two contours had the same shape only if their forms matched exactly, allowing for translation, rotation, and uniform scaling (ie scaling by the same amount along the vertical and horizontal axes). Several examples of 'same' and 'different' shapes drawn on paper were shown to the subjects.

The beginning of each trial was signaled by a short beep heard through earphones. 250 ms after the beep, a contour appeared in the upper left quadrant of the screen. 1000 ms later, the contour would disappear and another contour was presented in the lower right quadrant of the screen. The second contour remained visible until the subject made a response indicating whether they perceived the two stimuli as having the same shape or different shapes. After a response was made, the screen was cleared to all black, audio feedback indicating if the response was correct or incorrect was given, and a new trial began.

2.2 *Results and discussion*

Proportion of correct responses for each condition, averaged across subjects, is plotted in figure 3. Error bars indicate ± 1 SEM. Across subjects, proportion correct on trials with CC stimuli was reliably better than proportion correct on trials with NCC stimuli (0.67 versus 0.59; $t_9 = 6.769$, $p < 0.01$). The proportion of correct responses within

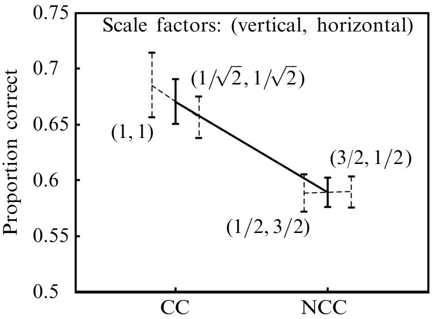


Figure 3. Results, experiment 1: proportion of correct same/different judgments are shown for both the CC stimuli (left) and the NCC stimuli (right). For each condition, mean performance is shown, flanked by the data shown separately for each of the two scale factors in that condition. For all data, error bars are ±1 SEM.

each condition is also shown for the various scale factors. We found no evidence of differences in performance between the two scale factors within each condition ($t_9 = 1.180$, $p = 0.268$; $t_9 = 0.085$, $p = 0.934$).

These results indicate that the CC shapes were significantly easier to classify as same/different than NCC shapes. The manner in which the stimuli were generated, and the lack of any difference between the various scale factors within each condition, suggests that this result is not due to an unintentional stimulus artifact. Specifically, some NCC shapes scaled to a greater degree vertically than horizontally tend to be perceived as contours slanted in depth. First, there was no difference in performance between the two scale factors in the NCC condition. Second, this depth effect may actually arise because subjects are compelled to encode the NCC as CC shapes sloped in depth, if the 2-D frontal-parallel projections of the shapes are consistent with this interpretation (which in our case they are). We are currently investigating this possibility. Regardless of the potential depth cue in half of the NCC shapes, the results of experiment 1 suggest that the human visual system exploits redundancy in the form of adjacent regions of constant curvature to form more compact or more precise representations of contour shape.

One reason why a compact representation of shape may be necessary is to reduce the amount of data that must be stored in order to allow for shape comparisons across viewpoints. In experiment 1, we simulated changes in viewpoint by translating, rotating, and scaling the shapes between presentations. In experiment 2 we directly test the importance of changes in viewpoint for demonstrating the performance advantage for classifying pairs of CC shapes compared to pairs of NCC shapes.

3 Experiment 2

3.1 *Method*

Experiment 2 was the same as experiment 1, except as noted.

3.1.1 *Procedure.* The differences between experiments 1 and 2 are in the spatial transformations and timing between the presentations of the two shapes on each trial. Specifically, unlike experiment 1, there was no translation, rotation, or scaling of the first contour relative to the second. Also, since both shapes were presented without applying any of the spatial transformations used in experiment 1, an interval of 1000 ms was added between presentation of the first and second shape to eliminate the perception of apparent motion or deformation of a single contour, rather than the perception of two distinct contour shapes.

3.2 *Results and discussion*

Proportion of correct responses for each condition, averaged across subjects (figure 4, left) remained reliably higher in the CC condition than in the NCC condition (0.74 versus 0.68; $t_9 = 2.801$, $p < 0.05$). A mixed factorial ANOVA with shape type as a within-subjects, and experiment (1 versus 2) as a between-subjects factor revealed that average

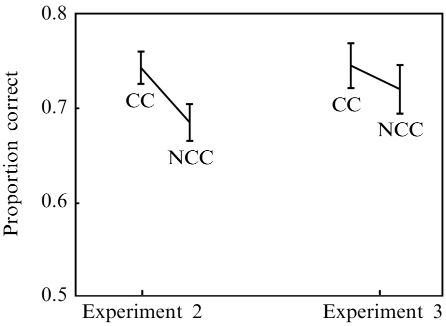


Figure 4. Results, experiments 2 and 3. Proportion of correct same/different judgments are shown for experiment 2 (left) and experiment 3 (right). Data for the CC stimuli and the NCC stimuli are plotted separately. For all data, error bars are ±1 SEM.

proportion of correct responses in experiment 2 was also reliably better than in experiment 1 ($F_{1,18} = 11.51$, $p < 0.05$).

The results of experiment 2 are consistent with the idea that shape information must be compressed in order to fit within the capacity limits of visual working memory. These results fit with the observation that some comparisons that are possible for simultaneously present visual stimuli are not possible if one or both of the stimuli must be encoded, stored, and retrieved from visual memory (Farah et al 1994). For example, very subtle shape discriminations between two shapes may be possible if they are simultaneously visible and subjects are allowed to inspect them for an extended period of time.

In experiment 2, we eliminated the spatial transformations and still found that CC shapes were easier to compare than NCC shapes. This suggests that representing the first shape during the 1000 ms interstimulus interval required forming a representation that leveraged the redundancy of the CC parts. Consequently, this result suggests that the redundant shapes in the CC condition were more easily compared even when the comparison could be made based on an image-like representation of the shapes, rather than a representation that allows for normalizing across translation, rotation, and scaling.

It would, of course, be trivial to detect differences between two sequentially presented shapes if there was no temporal gap separating them (similar to how change blindness requires a blank interval between the two displays—Rensink et al 1997), provided the shapes did not undergo any spatial transformations. One question that remains is whether the benefit of CC shapes over NCC shapes is dependent on the interstimulus interval, beyond the case of zero (or very short) interstimulus intervals. In experiment 3, we investigate this issue.

4 Experiment 3

4.1 *Method*

Experiment 3 was the same as experiment 2, except as noted.

4.1.1 *Procedure.* Stimuli were presented in the same manner as in experiment 2, except that the interval between presentation of the first and second contour was reduced from 1000 ms to 500 ms. Inspection of the presentation sequence indicated that this interval was long enough so that the image sequence was not perceived as the motion or deformation of a single contour, ie the sequentially presented contours appeared as two distinct shapes.

4.2 *Results and discussion*

Proportion of correct responses for each condition, averaged across subjects (figure 4, right) was not reliably different between the CC and NCC conditions (0.75 versus 0.72; $t_9 = 1.457$, $p = 0.179$).

We contend that in experiments 1 and 2, CC shapes were more easily remembered than NCC shapes because the visual system leverages the redundancy in the CC shapes when forming symbolic representations that are stored in visual working memory. The results of experiment 3 suggest that an earlier, high-capacity, transient representation, like that described in classic work by Sperling (1960), persists for 500 ms after a shape disappears, but not for 1000 ms. This interval may approximate the processing time needed to form a more concise, durable representation.

Besides this main result, experiment 3 also serves as an indirect verification of the equivalence of the same/different pairs of stimuli in the CC and NCC conditions. Under the conditions of experiment 3, subjects were no better at discriminating either shape type, indicating that the difference in performance between conditions in experiments 1 and 2 cannot be accounted for by an unintentional difference between the CC and NCC shapes. Furthermore, the similarity in performance across conditions in experiment 3 cannot be attributed to ceiling effects since the highest measured performance for any subject in any condition was 85.6% correct (the highest performance for any of the subjects in any condition of experiments 1 and 2 was 81.2% correct).

5 General discussion

The results of experiment 1 suggest that the representation of 2-D contour shapes in visual working memory leverages the redundancy present in shapes composed of constant curvature parts. The CC shapes are redundant in that, within each part, the curvature at one point can be predicted exactly from the curvature at any other point. The fact that this redundancy leads to better shape-matching performance suggests that the visual system uses CC parts for representing shapes when they are scaled, rotated, and translated between presentations.

Experiments 2 and 3 showed that performance is better for shapes composed of constant-curvature segments even when transformations do not occur between presentations, but not when there are no transformations and the inter-stimulus presentation time is small (500 ms).

Together, these results are consistent with the proposed purpose of the visual system's sensitivity to the 'same curvature' relation, specifically to reduce the amount of information that needs to be stored in visual working memory and more permanent representations. We propose that no difference in shape-comparison performance was observed in experiment 3 because subjects were using a high-capacity transient representation of visual information like those described in Sperling (1960) and Ferber and Emrich (2007). This type of representation does not have the capacity limits that are believed to constrain visual working memory (Luck and Vogel 1997; Alvarez and Cavanagh 2004), but may not support the spatial reasoning that is necessary for shape comparisons across transformations.

5.1 *Conclusions*

5.1.1 *Quantitative interpretation of the results.* The connection between the nature of the representation and performance on a task that involves using that representation is not always clear. In our case, the proposed CC parts-based representation is consistent with at least two processing approaches. The two approaches illustrate the tradeoff between data compression and fidelity (in our case, how faithfully the shape representation encodes the shape information in the stimulus), an issue commonly encountered in computer science. Specifically, the results of experiment 1 could have occurred because NCC shapes need to be represented with more parts (and therefore require more memory capacity) or that NCC shape representations are more distorted relative to the actual shape stimulus.

Here we consider these two possibilities, by presenting two models that illustrate the tradeoff between representation complexity and fidelity. The first model, model A, forms representations that limit the amount of distortion between the stimulus and its representation, but have high complexity in terms of number of allowable parts in the representation. The second model, model B, forms representations with significantly limited number of parts, but must therefore tolerate distortion between stimuli and their corresponding representations when an input shape cannot be perfectly represented with the maximum allowable number of parts. Both models benefit from the use of a CC representation, rather than more complicated curve fitting algorithms. First, both models form their parts only by considering adjacent curvature values along the contour. Second, the goodness of fit of the representation to the stimulus is done considering only the curvature values within each part. Both of these factors greatly simplify the implementation of these curve fitting procedures relative to more sophisticated methods.

Model A and model B both predict lower recognition performance for NCC shapes relative to CC shapes, but for different reasons. Model A can adequately represent CC shapes with high fidelity using a small number of parts because of the match between the shapes and the parts used to represent them. In contrast, NCC shapes must be represented using a larger number of parts in order to achieve a similar level of fidelity because of the mismatch between the shapes and the parts used to represent them. Consequently, NCC shapes have more complex representations. Model B can represent CC shapes with higher fidelity using a fixed number of parts than NCC shapes, again because of the match between the CC shapes and the parts used to represent them. In contrast, NCC shapes cannot be represented with a fixed number of CC parts without introducing distortion between the stimulus and its representation. Consequently, even though the complexities of the representations of CC and NCC shapes are the same, the representations of CC shapes more faithfully encode their geometry.

These models illustrate the tradeoff between fidelity and complexity. Exactly how the visual system handles this tradeoff is beyond the scope of this paper. The best overall strategy will include factors like the range of complexity of contour shapes, the capacity of visual working memory for stimuli of varying type (eg as described in Alvarez and Cavanagh 2004), and the magnitudes of typical shape differences that are to be detected. A detailed description of how these models can be implemented, including sample inputs and outputs is presented in the Appendix.

5.1.2 *Segments in natural scenes.* As the computational models described above show, the visual system's use of CC segments for coding contour shape will be more or less effective depending on the contours to which this representation is applied. Analysis of the spatial dependence of oriented contrast in natural scenes shows that co-circular segments are indeed more probable than other configurations (Sigman et al 2001), but not necessarily that circles are commonly found in natural images. Chow et al (2002) measured the relative orientations of line segments in artificial scenes composed of ellipses sampled from a distribution of eccentricities, which included some probability of ellipses with eccentricity $= 0$, ie circles. They found that even a small number of circles will tend to skew the distribution toward a maximum orientation consistent with co-circularity, since the probability density of relative orientations on a single circle is concentrated entirely at the value in the distribution corresponding to co-circularity.

Natural images are, however, not composed of circles or ellipses. Even if a circle were present in a scene, the projection of that circle onto an image (the relevant shape from which the statistics measured by Sigman et al 2001 were sampled) would be most likely elliptical. Similarly, the bounding contours of most objects are clearly not ellipses.

Chow et al (2002) interpret their findings to suggest that the fact that co-circularity arises in natural images as the most probable arrangement of two separated contour segments results from the abundance of smooth closed contours in natural scenes. This interpretation is consistent with the idea that a co-circularity-based representation may have good correspondence with relatively few parts (and therefore low memory load) when applied to the contour shapes found in natural images, even though these contours are not particularly likely to be circular. This notion is consistent with the prominent role that co-circularity plays as a constraint in Parent and Zucker's (1989) model of contour extraction from natural images. In addition, Parent and Zucker (1989) note that, while higher-order derivatives of curvature might be useful for more concisely representing a larger class of contours, it is difficult to characterize these properties of contours presented in natural scenes, making such a representation impractical.

5.1.3 *Psychophysics.* Shape, of course, is more complicated than the relative orientation of two contour segments in an image. The difficulty of sampling the higher-order statistics necessary for developing a principled theory of form perception from analysis of natural scenes have naturally led researchers to test various plausible theories of form perception behaviorally. For example, all objects have a closed bounding contour (Elder and Zucker 1993). Perhaps the visual system uses this fact to extract and group contour segments in natural images. In fact, Kovacs and Julesz (1993) found that detecting a contour formed from a group of oriented, spatially separated elements hidden in a field of randomly oriented and positioned elements was easier if the contour was closed. However, more recent analysis shows that the benefits of closure, a non-local characteristic, may actually be accounted for using local statistics, like density and eccentricity of elements (Tversky et al 2004).

Psychophysical evidence, however, suggests that circles are perceptually 'special'. Pizlo et al (1996) compared detection performance for a curve formed from dots arranged in straight lines, dots arranged in circular arcs, and dots arranged in various types of irregular paths. They found that straight lines were easiest to detect, but that circular arcs were easier to detect than irregular paths (provided the change in curvature along the irregular path was not too small). More relevant to these studies, however, Pizlo et al also found that circular arcs were significantly easier to detect than all the irregular paths they tested when the subject was given prior information about the shape of the target. This result is consistent with our result showing that CC shapes are only easier to compare when the shape must be stored in visual working memory.

In more recent experiments, Achtman et al (2003) showed, using a Gabor-path detection paradigm, that circular paths were more easily detected than radial or spiral paths. Similar detection-threshold advantages for circles have been found using Glass patterns (Wilson et al 1997; Seu and Ferrera 2001; Kurki and Saarinen 2004). Other psychophysical evidence suggests that the primitives of global form perception include both circular and spiral pooling mechanisms (Webb et al 2008).

5.1.4 *Neurophysiology.* Recently, evidence from single-cell recordings has shown that intermediate visual areas such as V4 may be representing object-oriented contour curvature (Pasupathy and Connor 1999, 2001, 2002). Full-fledged representations of shapes are presumably derived from the collective output of these cells. It is difficult, though, given the small number of cells that can be sampled and the large set of contour shapes in the world, to parameterize 'shape cell' receptive fields and then form an ecological theory of why these cells code shape in the way that they do.

There is, however, behavioral and neuroimaging evidence suggesting that circles have special status in higher-level visual areas. Gallant et al (2000) report a case study of a patient with no impairment in processing orientation, but significant impairment in spiral discrimination, and in discrimination between circles and radially distorted circles. This result may indicate a circle deficit but could also indicate a more fundamental curvature deficit. More recently, Dumoulin and Hess (2007) used fMRI to show that fields of Gabor patches oriented along circular paths elicited significantly larger BOLD response in human areas V3/VP and V4 than similar fields of Gabor patches with matched local curvatures, but non-circular paths.

6 Conclusion

We have shown that the visual system can recognize CC shape better than NCC shapes when the shapes were compared across different viewpoints or across intervals of 1000 ms but from the same viewpoint. Performance was not significantly different between CC and NCC shapes when the shapes were compared across intervals of 500 ms and from the same viewpoint. Together, these results suggest that visual working memory represents contour shape in terms of regions of constant curvature, but that earlier more-transient representations do not.

The proposed representation has a number of desirable properties. First, it fits with the idea of efficient coding in the classical sense, or in terms of the simplicity principle (though it does not distinguish between the two). Within the framework of classical information theory, just as Feldman and Singh (2005) proposed that the distribution of curvature is peaked at zero, we could similarly propose that the distribution of change in curvature is also peaked at zero. Recent evidence from the analysis of natural images suggests that co-circularity may emerge as part of an efficient coding of contour shape (Zucker et al 2011). Within the framework of SIT, CC contours are simpler because they can be expressed with a concise code by utilizing the iteration relation.

Regardless of the best account of efficiency, CC coding reduces the amount of information that needs to be encoded for shape recognition in a manner that is flexible enough to handle less-precise coarse coding (for relatively easy discriminations) and more-precise fine coding (for more subtle discriminations). Also, as the models (detailed in the Appendix) show, this representation can be derived from local encodings along the contour. These properties fit with other work indicating that constant curvature representations could be readily extracted from relations of local orientation-sensitive units found in early cortical areas, and that such a representation leads naturally to scale-invariant contour shape representations (Kellman and Garrigan 2006). Consequently, a constant curvature representation is one way in which local encodings, like those found in early visual cortex, can be used to form the more abstract, symbolic representations of visual perception.

As applied to the closed bounding contours of real objects, this representation could complement other models of object representation. Specifically, models that describe part segmentation and structure [eg medial axis type representation—(eg Blum 1973; Feldman and Singh 2006), 'parts of recognition' (Hoffman and Richards 1984)] could be augmented with a simple coding scheme for the shape of the contour at specific important locations. Together, these models could be used to generate from the bounding contour a functional set of parts, make explicit the important perceptual relations among them and, finally, form a concise viewpoint-invariant description of whole shapes.

References

Achtman R L, Hess R F, Wang Y-Z, 2003 "Sensitivity for global shape detection" *Journal of Vision* **3**(10) 616–624

Alvarez G, Cavanagh P, 2004 "The capacity of visual short term memory is set both by visual information load and by number of objects" *Psychological Science* **15** 106–111

Attneave F, 1954 "Some informational aspects of visual perception" *Psychological Review* **61** 183

Barlow H B, 1961 "Possible principles underlying the transformation of sensory messages", in *Sensory Communication* Ed. W A Rosenblith (Cambridge, MA: MIT Press) pp 217–234

Barrow H G, Tenenbaum J M, 1981 "Interpreting line drawings as three-dimensional surfaces" *Artificial Intelligence* **17** 75 – 116
Biederman I, 1987 "Recognition-by-components: a theory of human image understanding" *Psychological Review* **94** 115 – 147
Blum H, 1973 "Biological shape and visual science (part I)" *Journal of Theoretical Biology* **38** 205 – 287
Brainard D H, 1997 "The Psychophysics Toolbox" *Spatial Vision* **10** 433 – 436
Butler D I, 1982 "Predicting the perception of three-dimensional objects from the geometrical information in drawings" *Journal of Experimental Psychology: Human Perception and Performance* **8** 674 – 692
Chow C C, Jin D Z, Treves A, 2002 "Is the world full of circles?" *Journal of Vision* **2**(8) 571 – 576
Dumoulin S O, Hess R F, 2007 "Cortical specialization for concentric shape processing" *Vision Research* **47** 1608 – 1613
Elder J H, Zucker S W, 1993 "The effect of contour closure on the rapid discrimination of two-dimensional shapes" *Vision Research* **33** 981 – 991
Farah M J, Rochlin R, Klein K L, 1994 "Orientation invariance and geometric primitives in shape recognition" *Cognitive Science* **18** 325 – 344
Feldman J, Singh M, 2005 "Information along contours and object boundaries" *Psychological Review* **112** 243 – 252
Feldman J, Singh M, 2006 "Bayesian estimation of the shape skeleton" *Proceedings of the National Academy of Sciences of the USA* **103** 18014 – 18019
Ferber S, Emrich S M, 2007 "Maintaining the ties that bind: The role of an intermediate visual memory store in the persistence of awareness" *Cognitive Neuropsychology* **24** 187 – 210
Gallant J L, Shoup R E, Mazer J A, 2000 "A human extrastriate cortical area that is functionally homologous to Macaque area V4" *Neuron* **27** 227 – 235
Hanssen A, Leeuwenberg E, Helm P van der, 1993 "Metrical information load of lines and angles in line patterns" *Psychological Research* **55** 191 – 199
Helm P A van der, 2011 "The influence of perception on the distribution of multiple symmetries in nature and art" *Symmetry* **3** 54 – 71
Helm P A van der, Leeuwenberg E L J, 1991 "Accessibility, a criterion for regularity and hierarchy in visual pattern codes" *Journal of Mathematical Psychology* **35** 151 – 213
Helm P A van der, Lier R J van, Leeuwenberg E L J, 1992 "Serial pattern complexity: irregularity and hierarchy" *Perception* **21** 517 – 544
Hochberg J, McAlister E, 1953 "A quantitative approach to figural 'goodness'" *Journal of Experimental Psychology* **46** 361 – 364
Hoffman D D, Richards W A, 1984 "Parts of recognition" *Cognition* **18** 65 – 96
Hoffman D D, Singh M, 1997 "Salience of visual parts" *Cognition* **63** 29 – 78
Hubel D H, Wiesel T N, 1959 "Receptive fields of single neurons in the cat's striate cortex" *Journal of Physiology* **148** 574 – 591
Kellman P J, Garrigan P, 2006 "Segmentation, grouping, and shape: Some Hochbergian questions", in *Perception: Essays in Honor of Julian Hochberg* Ed. M A Peterson (New York: Oxford University Press) pp 542 – 554
Koenderink J J, 1984 "What does the occluding contour tell us about solid shape?" *Perception* **13** 321 – 330
Kohler W, 1938 "Physical Gestalten", in *A Source Book of Gestalt Psychology* Ed. W D Ellis (New York: Humanities Press) pp 17 – 54
Koffka K, 1935 *Principles of Gestalt Psychology* (New York: Harcourt, Brace, & World)
Kovacs I, Julesz B, 1993 "A closed curve is much more than an incomplete one: effect of closure in figure – ground segmentation" *Proceedings of the National Academy of Sciences of the USA* **90** 7495 – 7497
Kurki I, Saarinen J, 2004 "Shape perception in human vision: specialized detectors for concentric spatial structures" *Neuroscience Letters* **360** 100 – 102
Leeuwenberg E L J, 1971 "A perceptual coding language for visual and auditory patterns" *American Journal of Psychology* **84** 307 – 319
Leeuwenberg E L J, Helm P A van der, 1991 "Unity and variety in visual form" *Perception* **20** 595 – 622
Lier R van, 1999 "Investigating global effects in visual occlusion: From a partly occluded square to a tree-trunk's rear" *Acta Psychologica* **102** 203 – 220
Liu B, Todd J T, 2004 "Perceptual biases in the interpretation of 3D shape from shading" *Vision Research* **44** 2135 – 2145
Luck S J, Vogel E K, 1997 "The capacity of visual working memory for features and conjunctions" *Nature* **390** 279 – 281
Marr D, 1982 *Vision* (San Francisco, CA: W H Freeman)

Marr D, Nishihara H K, 1978 "Representation and recognition of the spatial organization of three-dimensional structure" *Proceedings of the Royal Society of London, Series B* **200** 269 – 294
Parent P, Zucker S W, 1989 "Trade inference, curvature consistency and curve detection" *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11** 823 – 839
Pasupathy A, Connor C E, 1999 "Responses to contour features in macaque area V4" *Journal of Neurophysiology* **82** 2490 – 2502
Pasupathy A, Connor C E, 2001 "Shape representation in area V4: Position-specific tuning for boundary conformation" *Journal of Neurophysiology* **86** 2505 – 2519
Pasupathy A, Connor C E, 2002 "Population coding of shape in area V4" *Nature Neuroscience* **5** 1332 – 1338
Pelli D G, 1997 "The VideoToolbox software for visual psychophysics: Transforming numbers into movies" *Spatial Vision* **10** 437 – 442
Peterson M A, Gibson B S, 1991 "The initial identification of figure – ground relationship: Contributions from shape recognition processes" *Bulletin of the Psychonomic Society* **29** 199 – 202
Pizlo Z, Salach-Golyska M, Rosenfeld A, 1996 "Curve detection in a noisy image" *Vision Research* **37** 1217 – 1241
Poggio T, Edelman S, 1990 "A network that learns to recognize 3D objects" *Nature* **343** 263 – 266
Rensink R A, O'Regan J K, Clark J J, 1997 "To see or not to see: the need for attention to perceive changes in scenes" *Psychological Science* **8** 368 – 373
Riesenhuber M, Poggio T, 1999 "Hierarchical models of object recognition in cortex" *Nature Neuroscience* **2** 1019 – 1025
Seu L, Ferrera V P, 2001 "Detection thresholds for spiral glass patterns" *Vision Research* **41** 3785 – 3790
Shannon C E, 1948 "A mathematical theory of communication" *Bell System Technical Journal* **27** 379 – 423, 623 – 656
Sigman M, Cecchi G, Gilbert C, Magnasco M, 2001 "On a common circle: natural scenes and Gestalt rules" *Proceedings of the National Academy of Sciences of the USA* **98** 1935 – 1940
Singh M, Seyranian G D, Hoffman D D, 1999 "Parsing silhouettes: The short-cut rule" *Perception & Psychophysics* **61** 636 – 660
Sperling G, 1960 "The information available in brief visual presentations" *Psychological Monographs* **74** 1 – 29
Todd J T, 2004 "The visual perception of 3D shape" *Trends in Cognitive Science* **8** 115 – 121
Tversky T, Geisler W S, Perry J S, 2004 "Contour grouping; closure effects are explained by good continuation and proximity" *Vision Research* **44** 2769 – 2777
Verhaeghen P, Palfai T, Johnson M P, 2006 "Verbal labeling as an assimilation mnemonic for abstract visual stimuli: The sample case of recognition memory for Chinese characters" *Memory and Cognition* **34** 795 – 803
Victor J D, Conte M M, 2004 "Visual working memory for image statistics" *Vision Research* **44** 541 – 556
Wagemans J, De Winter J, Op de Beeck H, Ploeger A, Beckers T, Vanroose P, 2008 "Identification of everyday objects on the basis of silhouette and outline versions" *Perception* **37** 207 – 244
Webb B S, Roach N W, Peirce J W, 2008 "Masking exposes multiple global form mechanisms" *Journal of Vision* **8**(9) 1 – 10
Wertheimer M, 1938 "Laws of organization in perceptual forms", in *A Source Book of Gestalt Psychology* Ed. W D Ellis (New York: Harcourt Brace & Co.) pp 71 – 88
Wilson H R, Wilkinson F, Asaad W, 1997 "Concentric orientation summation in human form vision" *Vision Research* **37** 2325 – 2330
Yin C, Kellman P J, Shipley T F, 1997 "Surface completion complements boundary interpolation" *Perception* **26** 1459 – 1479
Zucker S, Lawlor M, Holtmann-Rice D, 2011 "Third order edge statistics reveal curvature dependency", paper presented at the 11th Annual Meeting of the Vision Sciences Society, May, Naples, FL, USA, 6 – 11 May

Appendix
Model A

Model A demonstrates one way to form a representation that enforces a minimum fidelity. Model A segments CC and NCC contour shapes into parts based on the difference in curvature between adjacent elements within each segment. The input to the model is a set of signed curvatures κ sampled uniformly at M locations along the contour. Since all computations will be performed on the curvature values, the initial coding of the contour has M elements. The output from the model is a set of CC parts that represent the shape stimulus in terms of a single curvature and angular extent for each part. This simple model has one parameter, t, the maximum allowable change in curvature between any two adjacent points within a segment.

Model A begins with the contour segmented in to M parts, where each part has a single curvature value and the same arbitrary extent, which we set to 1. The number of initial parts, M, is determined by the sampling rate of curvature along the contour. Next, the magnitude of the relative change in curvature at each point along the contour, Δ_x, is measured:

$$\Delta_x = \left| \frac{\kappa_x - \kappa_{x+1}}{\kappa^{\ x+1}} \right| ,$$

and the location x along the contour where Δ is smallest is located. If this value is less than t, then the segments meeting at points x and $x+1$ are joined. Next, Δ_x is again searched, but this time not considering curvature values that have already been joined into the same segment. The smallest value of Δ_x is again compared to t, and if this value is again less than t, then the two segments meeting at points x and $x+1$ are joined. This process continues until no two adjacent points that do not belong to the same segment meet the criterion for being joined. Finally, the shape of each part is then represented as the mean curvature within the part and an extent equal to the number of sampled curvature values within that part. Ultimately, the contour is represented in terms of a set of N parts, with each part having a single curvature κ_n and extent m_n, with

$$M = \sum_{n=1}^{N} m_n ,$$

assuming again that the extent of each of the initial M parts was set to 1, and so the entire contour has extent M.

So, model A generates a CC parts-based representation of the input contour segment with arbitrarily high fidelity, but also a potentially large number of parts. As expected, given a reasonable value of t, this model forms representations with a small number of parts when the input is a 5-segment CC shape. Specifically, the representation typically has 5 parts, since the input consisted of 5 CC segments, although occasionally an additional small part could result from sampling issues at the contour endpoint.

Model A's performance is illustrated in figure A1. Input contours (above the arrows) and outputs drawn from their CC representations (below the arrows) are shown. Three CC input contour examples are shown on top and three NCC input contour examples are shown on bottom. In the redrawn representations, CC segments are marked. The numbers of parts that form the representation of the CC and NCC shapes shown represent typical performance. In a simulation of 500 contours, the median number of parts in the CC shape segmentation was 5 (range: 5 – 6 parts), and in the NCC shape segmentation was 15 (range: 8 – 25 parts).

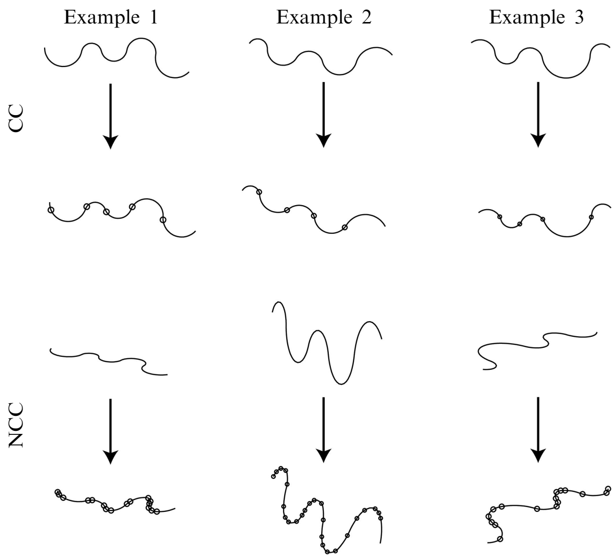


Figure A1. Example performance of model A. Model A forms CC representations of contour shapes with a fixed target fidelity, but no limit on the number of parts. Example input shapes and corresponding shapes drawn from the model's output are shown for CC inputs (top) and NCC inputs (bottom). Note that the number of parts is significantly larger for the NCC shapes (see text).

Model B

Model B demonstrates one way to form a representation that limits its complexity, in terms of number of parts. Like model A, the input to model B is a set of M elements, each corresponding to a single, signed curvature value, with M determined by the curvature sampling rate. We then choose a minimum segment length, m; essentially resampling the contour into n parts, with $n = M/m$. For our purposes, $m \ll M$ and the model's output is not significantly affected by the choice of m, provided m is small relative to M. This resampling allows for an initial calculation of the variance of the curvature within each part, which is required by the model.

Model B proceeds by combining adjacent segments, generating a representation of the contour with increasingly larger, but of course fewer, parts. The current number of segmented parts of the contour at any time is n, with the contour initially segmented into a set of n parts $\{p_1 ... p_n\}$. Each part p_x corresponds to a set of m_x curvature values $s_x = \{\kappa_1 ... \kappa_{m_x}\}$. On each iteration, the model joins two adjacent parts. The two parts that are joined are the two that will result in a new set of parts with lowest total curvature index of dispersion, summed across all n remaining parts. The index of dispersion for each part is calculated as the absolute value of the variance divided by the mean, and so the total curvature index of dispersion, D, for a contour segmented into n parts is

$$D = \sum_{x=1}^{n} \left| \frac{\operatorname{var}(s_x)}{\bar{s}_x} \right| .$$

This process defines a set of parts with low relative variance of curvature within each part. The use of the dispersion (rather than, eg, the variance) causes the model to allow higher variance within a segment if the segment has higher magnitude mean curvature. This prevents the model from forming many small parts at locations along the contour with high curvature.

This process continues until the total number of parts forming the representation is $P_{\max}$. The shape of each part is then represented as the mean curvature within that part and its extent which, as in model A, is equal to the number of curvature values used to define that part. So, model B generates a CC parts-based representation of the input contour segment with the best fidelity (that can be found through this deterministic approach) and a fixed number of parts.

Performance of model B is illustrated in figure A2. Input contours (above the arrows) and outputs drawn from their CC representations (below the arrows) are shown. Three CC input contour examples are shown on top and three NCC input contour examples are shown on bottom. In the shapes redrawn from the CC representations, the segmentation into $P_{\max}$ parts, in this case with $P_{\max} = 10$, is marked with open circles. Notice that for the CC shapes shown on top, fewer parts could have been used without much (or perhaps any) loss in fidelity. On the bottom, some NCC shapes are better approximated with 10 CC segments than others.

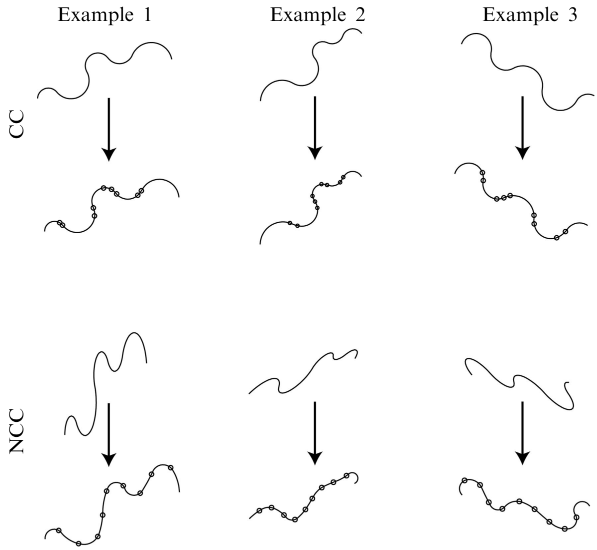


Figure A2. Example performance of model B. Model B forms CC representations of contour shapes with a fixed number of parts. In the examples shown, the number of parts in the representations was set to 10. Note that the distortion between the input shapes and shapes redrawn from the model output is greater for NCC shapes (see text).

As expected, both model A and model B generate better representations when the input contour is formed from CC segments. Model A represents CC contour shapes using fewer parts. Model B represents CC contour shapes in a way that better preserves the input contour shapes, given a fixed number of parts. These models are, of course, not the only way to generate CC representations that have a required fidelity (model A) or maximum number of parts (model B). They are intended to illustrate the two proposed reasons for why CC shapes are easier to recognize than NCC shapes in experiments 1 and 2.

As stated earlier, it is hard to determine how the visual system would, in practice, choose between fidelity and complexity of a stored representation. Aside from fixed parameters, eg the capacity of visual working memory, there are clearly situations in which a high-fidelity representation is required, perhaps at the expense of not encoding the entire shape. In other situations, a lower fidelity representation that fits within the capacity of visual working memory might be preferred. Ultimately, the choice depends on the task for which the shape representation is being used, information that may or may not be available at the time the shape is encoded.

ISSN 0301-0066 (print) ISSN 1468-4233 (electronic)

PERCEPTION

VOLUME 40 2011

www.perceptionweb.com