

Space of preattentive shape features

Liqiang Huang

Department of Psychology, Chinese University of
Hong Kong, Hong Kong



Four decades of studies in visual attention and visual working memory used visual features such as colors, orientations, and shapes. The layout of their featural space is clearly established for most features (e.g., CIE-Lab for colors) but not shapes. Here, I attempted to reveal the basic dimensions of preattentive shape features by studying how shapes can be positioned relative to one another in a way that matches their perceived similarities. Specifically, 14 shapes were optimized as n -dimensional vectors to achieve the highest linear correlation (r) between the log-distances between $C(14, 2) = 91$ pairs of shapes and the discriminabilities (d') of these 91 pairs in a texture segregation task. These d' values were measured on a large sample ($N = 200$) and achieved high reliability (Cronbach's $\alpha = 0.982$). A vast majority of variances in the results ($r = 0.974$) can be explained by a three-dimensional SCI shape space: *segmentability*, *compactness*, and *spikiness*.

Human observers extract a set of preattentive features, such as colors, shapes, and orientations, in parallel from the stimuli that are presented to their visual system (Treisman & Gelade, 1980; Wolfe, 1994). In the past four decades, these features have been commonly used to study visual attention, visual working memory, and other related topics in cognitive psychology. In some sense, these features serve as the building blocks of findings and theories in these areas.

This study examined shape features in a task that requires preattentive processing of large numbers of shapes at once. Previous studies have shown that several shape features appear to be processed preattentively, including intersection, closure, line terminations, and concavities (e.g., Chen, 1982; Donnelly, Humphreys, & Riddoch, 1991; Elder & Zucker, 1993, 1994; Hulleman, Winkel, & Boselie, 2000; Julesz, 1984, 1986; Julesz & Bergen, 1983; Pomerantz & Pristach, 1989).¹ The present study aims to provide a unified space for understanding the perceptual space of these shape features.

Public significance

Along with colors and orientations, shapes are very commonly used as visual features in studies of visual attention. However, there has been a lack of understanding of the layout of the shape space. This stands in contrast to other important feature dimensions such as color—the layout of a color space (e.g., CIE-Lab space) is very well known. For example, in a systematic review of visual features, Wolfe (1998) pointed out that “the most problematical basic feature is shape or form. . . . The heart of the problem is a lack of widely agreed understanding of the layout of shape space” (p. 33). This problem still remains unsolved up to now. In the present study, I attempted to reveal the basic dimensions of shape features by studying how shapes can be positioned relative to one another in a way that matches their perceived similarities. The present results suggested that, just like colors, shape features are governed by a simple and lawful space. On the practical side, these results provide guidance on how to choose distinguishable shape sets for visual symbols or glyphs in data visualization tasks (e.g., scatterplots).

Citation: Huang, L. (2020). Space of preattentive shape features. *Journal of Vision*, 20(4):10, 1–20, <https://doi.org/10.1167/jov.20.4.10>.

Preattentive shapes features versus postattentive shape processing

Shape is an umbrella term. Conceptually, there is a clear distinction (e.g., Wolfe & Bennett, 1997) between preattentive shape features that can be extracted in parallel and efficiently used in guiding attention (e.g., Julesz, 1984, 1986) versus the postattentive processing of an object's shape (e.g., Biederman, 1987) that can only occur when it is in the current focus of attention.²

Experimentally, these two can be dissociated from each other in the sense that the preattentive shape features' advantage (over other aspects of shapes that are not extracted as preattentive features) is substantial in attention-demanding tasks (e.g., visual search or texture segregation) but negligible in one-to-one comparison tasks. For example, Huang (2015a; see also Huang, 2015b, 2015c, 2020b, for related discussions) systematically examined how 16 stimulus types are processed in eight different tasks. These 16 stimulus types include both “preattentive shapes features” (e.g., four stimulus items involving the presence and absence of the intersection, the hole, and the line

shape features. *Journal of Vision*, 20(4):10, 1–20,



termination) and also “Ts in different orientations” that are not processed as preattentive shape features. Huang (2015a) found that in a perceptual discrimination (i.e., one-to-one comparison) task, the preattentive shape features were processed no better than Ts (thresholds 48 vs. 47 ms). However, in a visual search task that required processing eight items, the former was processed dramatically better than the latter (thresholds 81 vs. 335 ms).

These results suggest that the preattentive shape feature is not a proportional reduction of all aspects of shapes but is a special subset of the latter set of the information. By making an analogy to the search engine of an academic database, preattentive shape features correspond to the keywords that can be used to make a search, whereas the “postattentive shape processing” corresponds to the complete content of an individual article that can only be displayed one at a time. Only the keywords but not all details of the articles can be used to search for a target article. Moreover, the keywords of an article are not the proportional reduction of all the contents of that article but a special subset of it.

To summarize, among all the aspects of shapes, only some special subset of information (preattentive shape features) can be extracted in parallel and used to guide attention. It is the purpose of the present study to elucidate the exact content of this subset. Connections to postattentive shapes processing and other studies in the general domain of shape processing will be discussed in the Discussion.

Shape space

An important concept in studies of visual features is the *feature space*. Generally speaking, in a feature space, each feature corresponds to an n -dimensional vector, and so we can clearly see the relations between the features in this n -dimensional space. For example, in a CIE-Lab color space, each color corresponds to a vector in a three-dimensional (3D) space.

Although we have a fairly good understanding of the spaces for most features (for reviews, see Wolfe, 1998; Wolfe & Horowitz, 2004, 2017), shape space remains a mystery. In 1970, it was already discussed as an old and established problem in Leonard Zusne’s book *Visual Perception of Form*. In a systematic review of the research on visual features, Wolfe (1998) pointed out that “the most problematical basic feature is shape or form. . . . The heart of the problem is a lack of widely agreed understanding of the layout of shape space” (p. 33). This problem still remains unsolved up to now.

Why is shape space so elusive? The root of the difficulty seems to be the lack of a starting point. The visual processing of other features (e.g., colors) may be complex and nonobvious, but the features themselves can usually be organized in a fairly straightforward way. For example, the majority of saturated colors can be

found in the one-dimensional spectrum of wavelengths of light. In contrast, the shape of an object can include many details and cannot boil down to one or two obvious dimensions. To formulate a hypothesis on the structure of color space and to explain the functional role of colors, one can consider the different ways the spectrum of light may be nonlinearly mapped. However, with shape space, we do not know where to start. Therefore, with a lack of previous knowledge on its layout, understanding shape space has to be achieved by an open-ended exploration.

Discriminability versus featural distance

In this study, I attempted to reveal the basic dimensions of preattentive shape features by studying how the set of shapes I tried here can be positioned relative to one another in a way that matches their perceived similarities.

As shown in Figure 1a, if shapes are represented by vectors in an n -dimensional shape space that is *uniform*, then there should be a linear *discriminability* versus *featural distance* relationship: The more discriminable two shapes are from each other in a visual task (i.e., greater discriminability), the further away these two shape vectors are from each other in the shape space (i.e., longer featural distance). For example, if the cross-circle difference is more discriminable than the crescent-triangle difference, then it is reasonable to assume that in the shape space, the cross-circle distance is greater than the crescent-triangle distance.

Without prior knowledge on the scale of the discriminability for shapes or the scale of featural distance in shape space, I approached them in the most typical ways in psychophysics.³ Specifically, I assume d' is a linear scale of discriminability (e.g., Macmillan & Creelman, 2004), and following the classic Weber-Fechner law, I assume log-distance (i.e., the logarithmic value of the distance) is a linear scale of the featural distance.

Dimensionality reduction

In this modeling, the number of dimensions of the shape space can be determined as the *minimum number of dimensions required to explain a major portion of the variance in discriminabilities between all the possible pairs of shapes among this set*. For example, if we study three shapes, A, B, and C, and measure the three discriminabilities between the three pairs of shapes, we may find the values of d'_{AC} , d'_{AB} , and d'_{BC} can be adequately explained by placing these three shapes in a one-dimensional space (Figure 1b); alternatively, we may find that all three discriminabilities are roughly equal to each other, and therefore they have to be explained by placing the three shapes in a

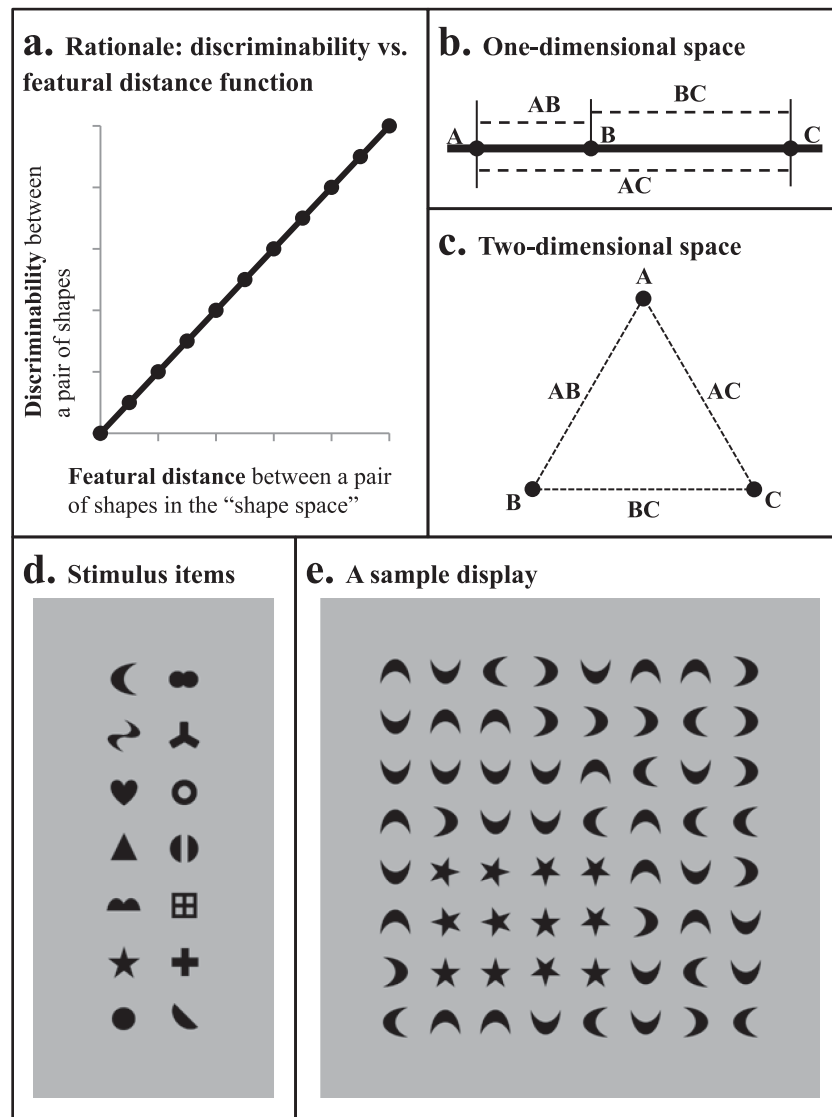


Figure 1. Rationale and method. The rationale of this study is illustrated (a). If shapes are represented by vectors in n -dimensional uniform shape space, then there should be a *linear discriminability-featural distance relationship*: The more discriminable two shapes are from each other, the further away they are from each other in the shape space. Next, the number of dimensions of the shape space can be determined as the minimum number of dimensions required to explain a major portion of the variance in discriminabilities between all the possible pairs of shapes among this set. For example, if we study three shapes, A, B, and C, and measure the three discriminabilities between the three pairs of shapes, we may find the values of d'_{AC} , d'_{AB} , and d'_{BC} can be adequately explained by placing these three shapes in a one-dimensional space (b); alternatively, we may find that all three discriminabilities are roughly equal to each other, and so they therefore have to be explained by placing these three shapes in a two-dimensional space (c). The present study included 14 typical shapes (d). In each trial, a target shape and a background shape were chosen from these 14 shapes and respectively used to generate a rectangular target-shape array inside a background array. Observers were asked to judge, in a very brief (100 ms) and masked display, whether the rectangle was vertical or horizontal (e).

two-dimensional space (Figure 1c). To generalize, if we study a large set of shapes and n is the minimum number of dimensions needed to explain a major portion of the variance in discriminabilities between all the possible pairs, then we can reasonably conclude that shape space is an n -dimensional space.

Method

An overview of the method

The present study included 14 shape items (Figure 1d). In each trial, 2 shapes were chosen from

these 14 shapes. A target shape was used to fill in a rectangular array, whereas a background shape was used to fill in the rest of a background array (Figure 1e). Observers were asked to judge, in a very brief (100 ms) and masked display, whether the rectangle was vertical or horizontal. This attention-demanding “texture segregation” task (Beck, 1966, 1980) was chosen so that the performance critically measured preattentive shape features rather than postattentive processing of shapes. The items were randomly presented in one of four possible orientations to make sure that the task relied on the shapes of the items in an orientation-invariant sense.

Participants

A total of 200 university students (average age = 20 years, 147 females), all of whom had a normal or corrected-to-normal vision, participated in this study’s experiments. I planned for this large sample size because there were only four trials per condition per observer (see below), and I aimed for a target reliability (Cronbach’s α) of at least 0.97. Fourteen observers were excluded because their performance was below a predetermined criterion (overall accuracy < 0.6), and the data of the remaining 186 observers were analyzed.

All procedures were in accordance with the national ethical standards on human experimentation and with the Declaration of Helsinki of 1975, as revised in 2008, and were approved by the research ethics committee of the Chinese University of Hong Kong. Informed consent was obtained from each observer.

Apparatus

The stimuli were presented on a computer monitor, and the observers viewed the display from a distance of about 60 cm.

Choice of shapes

Reasons for using two-dimensional filled geometrical shapes

As shown in Figure 1d, a total of 14 shapes were used in the present study. The scope of these shapes was limited to two-dimensional (2D) geometrical filled shapes. Alternatively, other important categories of shape stimuli could have been used, such as real-world objects (e.g., animals, letters, and characters), 3D shapes, and 2D unfilled line-drawing shapes. The present testing was restricted to one specific category because the exhaustive testing of all possible pairs of shapes requires a large number of observers. If

the scope is increased from 14 to 50 shapes, then approximately 2,700 observers will be needed to reach the same target reliability.

Given this restriction, the 2D geometrical shapes are chosen for two reasons. First, this is the set that has been linked directly with the preattentive shape features (e.g., Chen, 1982; Donnelly, Humphreys, & Riddoch, 1991; Elder & Zucker, 1993, 1994; Hulleman, Winkel, & Boselie, 2000; Julesz, 1984, 1986; Julesz & Bergen, 1983; Pomerantz & Pristach, 1989). On the other hand, no preattentive shape feature has been established for any other categories such as animals, letters, or 3D shapes (see Enns & Rensink, 1990, for an exception; but see Zhang, Huang, Serap, & Rosenholtz, 2015), making them less promising for the initial exploration.

Second, the category of 2D geometrical shapes is a conceptually more fundamental category of shapes, and the understanding of the shape features in other categories requires a prior understanding of this fundamental category. For example, if we have identified a shape feature in 3D shape, we will need to know whether that can be attributed to their 2D profiles. Similarly, if we have identified a shape feature in letters, we will need to know whether that can be attributed to their geometrical shapes.

Within 2D geometrical shapes, there are still variations. For example, shapes of unfilled line drawings could have been used. In the present study, the unfilled and filled shapes are not considered together because they will differ obviously from each other on colors and spatial frequency distributions. Between these two options, the filled shapes are chosen because the presence/absence of the holes of the shapes can be interpreted more unambiguously.

To summarize, it seems fruitful to start with 2D filled geometrical shapes as the initial exploration. Of course, other categories are potentially important and need to be explored in future studies.

Reasons for choosing these 14 specific shape items

This set of shapes is not a class of shapes generated by varying in one or two predetermined dimensions, as usually used in previous studies, but a collection of intentionally diverse items. For the present purpose, I wanted to maximize the diversity between items as much as possible in the scope of 2D filled geometrical shapes, with the hope that they can capture all dimensions of preattentive shape features in this open-ended exploration.

These shape items were chosen by considering two criteria.

First, shape items were included to represent known shape features as well as other conceptual dimensions that are potentially important. For example, the **O** was included to represent the feature of closure. The **+** was included to represent the feature of intersection. The **⊞**

was included to represent the conjunction of the closure and the intersection. The ◀ and ▶ were included to represent the feature of concavity. The ◐ was included to represent the dimension of “discontinued parts” (as against the other shapes that are one continued region), and the ◑ and ◒ were included to be the borderline cases on this dimension of continuity. The ◓ and ◔ were included to represent “elongated shapes.” Several shapes were chosen to vary on the dimensions of curvature (i.e., whether the edges are curved or straight lines), including the shapes consisting of curved edges (e.g., ●, ○), the shapes consisting of straight lines (e.g., ◻, ▲), and those in between (e.g., ◕, ◖).

Second, some shapes were included because they are among the most typical examples of the layperson’s concept of “shapes” (e.g., ▲, ★, ●, ◐, ◑).

Although it is impossible to perfectly match these items on their sizes and spatial frequencies, informal adjustments have been made to try to minimize these differences so that their contributions to the task performances can be kept at a minimal level.⁴

Stimuli and task

As shown in Figure 1e, in each stimulus display, a total of 64 items were presented as one 8×8 array on the center of the display. This array measured 8.5×8.5 cm. In each trial, 2 (out of all 14) shapes were chosen, one as the “target shape” and the other as the “background shape.” Twelve target-shape items were presented in a rectangular array (3×4 or 4×3), whereas the rest of the 8×8 array was filled by background-shape items. The position of the target-shape array was randomized with the constraint that it never resided on the leftmost column, rightmost column, topmost row, or bottommost row. Observers attempted to report whether the target-shape array was vertical or horizontal.

This texture segregation task (Beck, 1966, 1980) was chosen as an attention-demanding task⁵ to characterize the roles of shapes as basic visual features. For this purpose, one-to-one comparison tasks (e.g., pairwise comparison or same-different discrimination) were naturally excluded from the options. Another classic option for an attention-demanding task is the visual search task (e.g., Treisman & Gelade, 1980; Wolfe, 1994). In the present study, texture segregation was chosen as the better option than visual search because I wanted the task to require the observers to simultaneously select multiple items that were the same shape but presented in different orientations. This ensured that the task measured the shape in an orientation-invariant sense rather than focusing on a shape template at a specific orientation; for example, a crescent should be seen as the same shape regardless of whether it is facing upward or leftward. For this

purpose, each item was randomly presented in one of four orientations (0° , 90° , 180° , 270°).

The stimulus items were presented briefly, and accuracy was measured as the index of performance.

Procedure

Each trial started with a white fixation cross. The fixation cross was presented in the center of the display for 800 ms and was followed by the stimulus display, which was presented for 100 ms and then masked. The mask remained until a response was made. The observers were asked to report whether the rectangular target-shape array was vertical (pressing “j”) or horizontal (pressing “k”). They were asked to respond as accurately as possible but were under no time pressure (i.e., “unsped” responses).

Each observer completed five blocks (182 trials per block). In total, there were $14 \times 13 = 182$ conditions of choosing a target shape and a background shape out of the 14 shapes, and it was arranged so that each condition appeared once in each block.

The first block was regarded as practice and excluded from the analysis. In this practice block, the stimulus duration gradually decreased from 800 ms to 100 ms, so the observers could have an opportunity to learn about the task gradually.

Results

Data

I measured accuracies in a total of 182 conditions. The reliability of these accuracies (Cronbach’s α) was 0.982. This provided a solid foundation for subsequent modeling. The Cronbach’s α is an index of the reliability across individual items. It should be clarified that, for the present purpose, the “items” are not individual shapes but individual observers.

These accuracies were converted to 182 d' values (Figure 2). This conversion was done because the d' , but not accuracy, is usually a linear measure of discriminability in psychophysics (e.g., Macmillan & Creelman, 2004).

The discriminability between each pair of shapes (A, B) is reflected in two conditions (Target-A-Background-B vs. Target-B-Background-A). For example, the “star-target-crescent-background” and “crescent-target-star-background” displays both reflected the discriminability between star and crescent. Therefore, the 182 d' values were two-to-one averaged into 91 d' values (see Figure 2), which were then used in optimization as indexes of the discriminabilities between the 91 pairs of shapes.

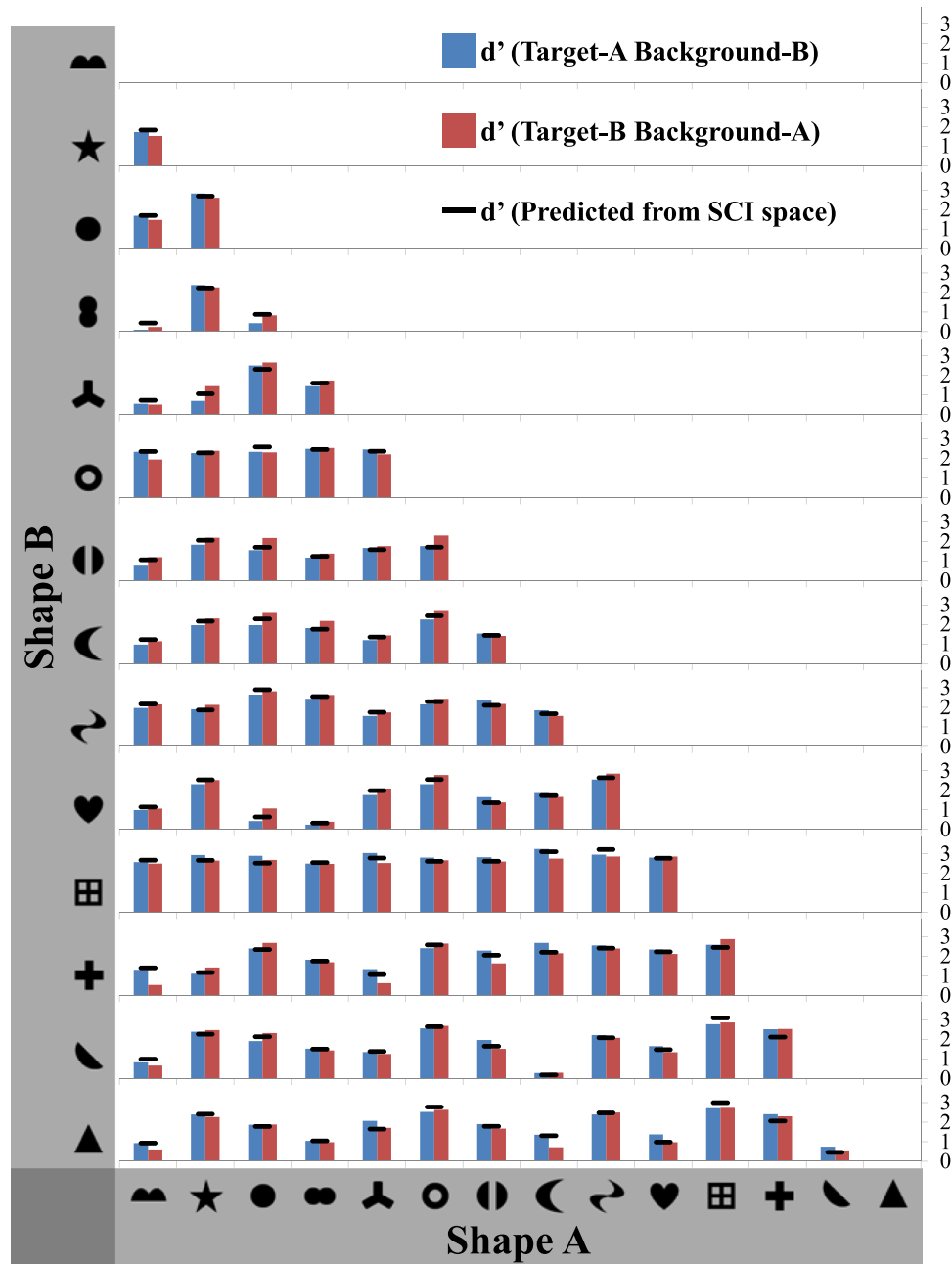


Figure 2. Data. For a pair of shape (A and B), the blue bar represents the d' of performing a texture segregation task for the Target-A-Background-B condition, whereas the red bar represents the Target-B-Background-A condition. The black line represents the predicted d' from the SCI shape space. See text for details.

Optimization

The coordinates of the 14 shapes as n -dimensional vectors were optimized to achieve the highest correlation (Pearson’s r) between the featural distances (i.e., the logarithmic value of the Euclidean distance between the vectors, following the Weber-Fechner law) between the 91 pairs and the 91 discriminabilities between them. These coordinates were found through an optimization algorithm. The algorithm

attempts to find the maximum (or minimum) of a function by varying input values and calculating the value of the function. Optimization is widely used in science and engineering (Nocedal & Wright, 2006). The present study adopted a straightforward general-purpose algorithm⁶: dimension-by-dimension fixed-step-size movements, implemented as a four-level loop (see Figure 3). The MATLAB script of this optimization procedure is available on the Open Science Framework project page at <https://osf.io/khgdt/>.

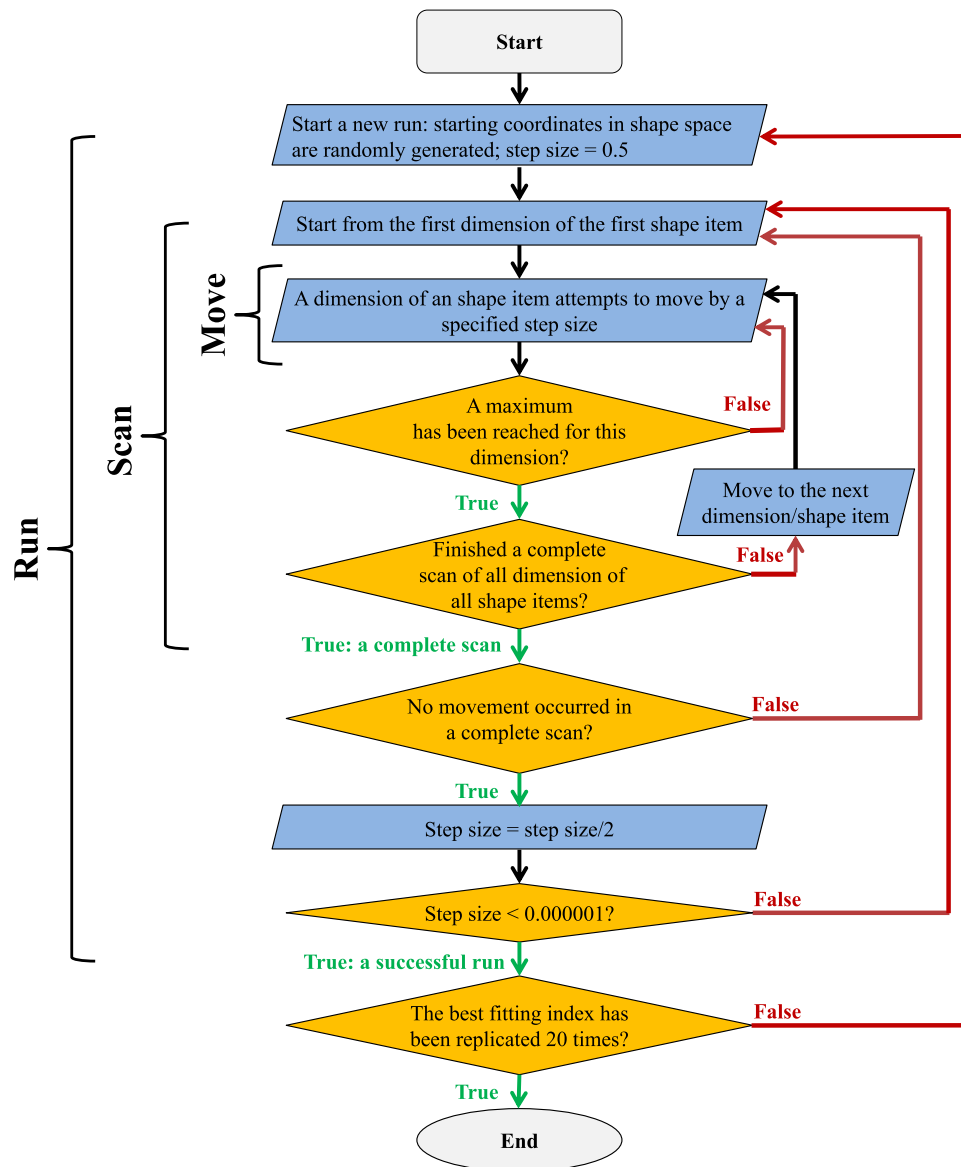


Figure 3. The program flowchart of the optimization algorithm. The present study adopted a general-purpose straightforward algorithm: dimension-by-dimension fixed-step-size movements, implemented as a four-level loop. Here, 14 vectors were used to represent the coordinates of the 14 shape items in the shape space. On the level of whole optimization procedure, multiple runs were started from different sets of random starting vectors and the runs repeated until a best-fitting r had been replicated 20 times. On the level of a run, at the beginning of each run, the step size was set as 0.5. A run consisted of multiple scans. These scans were repeated until meeting the criterion of successfully reaching a maximum of the fitting index r (step size < 0.000001) or until meeting some criteria of terminating a run. On the level of a scan, in each scan, the program sequentially scanned through all dimensions of all vectors. If, in a complete scan, there was no change for any of the dimensions of any of the 14 vectors because the fitting index (r) was already at a maximum, then the step size was halved at the end of this scan. On the level of a move, at this most elementary level, a dimension of a vector attempted to move by the specified step size until the fitting index (r) had reached a maximum. After each movement, the scale of vectors was standardized.

Level of the whole optimization procedure

In this algorithm, 14 vectors were used to represent the coordinates of the 14 shape items in the shape space. The whole optimization procedure consisted of multiple runs that were started from different sets

of random starting vectors ($M = \text{zero vector}$, $SD = 1$). The runs repeated until a best-fitting r had been replicated 20 times; this was set to make sure that the global, rather than a local, maximum was achieved.

Level of a run

At the beginning of each run, the step size was set as 0.5. A run consisted of multiple scans. These scans were repeated until meeting the criterion of successfully reaching a maximum of the fitting index r (step size < 0.000001) or until meeting some criteria of terminating a run.⁷

Level of a scan

In each scan, the program sequentially scanned through all dimensions of all vectors. If, in a complete scan, there was no change for any of the dimensions of any of the 14 vectors because the fitting index (r) was already at a maximum, then the step size was halved at the end of this scan.

Level of a move

At this most elementary level, a dimension of a vector attempted to move by the specified step size until the fitting index (r) had reached a maximum. After each movement, the scale of vectors was standardized.

Results of optimization

Next, for $n = 1$ to 6, I used this optimization algorithm to find the coordinates of the 14 shapes in n -dimensional space to achieve the highest correlation (r) between the log-distances between the 91 pairs of shapes and the 91 d' values of these pairs. As shown in Figure 4a, the shape space is most reasonably characterized by three dimensions. The fitting is already very good ($r = 0.974$) and improves relatively little for the subsequent dimensions. Therefore, I will focus on the 3D model in further discussion. The possibilities of a greater (or a smaller) number of dimensions will be considered in the Overfitting and Underfitting subsections in the Discussion.

Rotation of the shape space

The optimization in the present study is based on the discriminability-featural distance relation. Therefore, this optimization index remains the same if the layout of these vectors is affected by rigid motion (translation, rotation, reflection) or uniform scaling. Most importantly, to improve the conceptual interpretability of the dimensions, the layout of the vectors needs to be rotated so that the three axes match three interpretable shape features. Given the lack of previous knowledge on the layout of the shape space, there is no objectively optimal way of doing this. Therefore, I rotated it manually with the goal of

matching the previously established dimensions: the intersection and closure. Clearly, this can be achieved fairly well, and the resulting third dimension also has a fairly clear conceptual meaning (spikiness). This will probably need to be finely adjusted in the future when a more objective method is developed to determine the exact direction of the three axes.

Discussion

(S)egmentability-(C)ompactness-Sp(i)kiness (SCI) shape space

The optimization for 3D shape space generated 14 3D vectors, each representing a shape in the shape space. The 91 predicted discriminabilities calculated from these vectors are also given as black lines in Figure 2, and they are generally quite close to the measured discriminabilities (red and blue bars in Figure 2). Following the rationale illustrated in Figure 1a, I also plotted the relationship between the featural distances (in log scale) between the 91 pairs of shapes in the 3D shape space and the 91 discriminabilities (d') between them (Figure 4b). There is clearly a strong relationship between the two.

The best-fitting coordinates of the 14 shapes as 3D vectors are plotted in Figure 4c (and also in Figure 4d, which plotted the same data in two panels). The layout was orthogonally rotated so that the three axes match three interpretable shape features. This 3D shape space had also been translated and uniformly scaled so that all scores from all three dimensions could fit into the range [0, 10]. These transformations were conducted only for the purpose of making the scores more intuitive and had no effect on either the fitting index or the interpretations of dimensions.

Segmentability

The first dimension is shown as the x -axis of Figure 4c (and also x -axis of the left graph of Figure 4d). It matches quite well with the conceptual dimension of “intersection” that has been well established in previous studies (Julesz, 1984, 1986; Julesz & Bergen, 1983). Shapes including clear intersections (\boxplus and \oplus) scored very high on this dimension, and most of the other shapes scored low. Interestingly, the shape \star scored as high as \oplus , and the shape \blacktriangleright also scored moderately. It seems that this dimension generally responds to “joints” on which multiple segments are connected, rather than strictly to the crossing of two lines. This observation is consistent with Wolfe and DiMase’s (2003) finding that, preattentively, an intersection is indistinguishable from a four-segment shape that does not include an intersection. Taken together, although

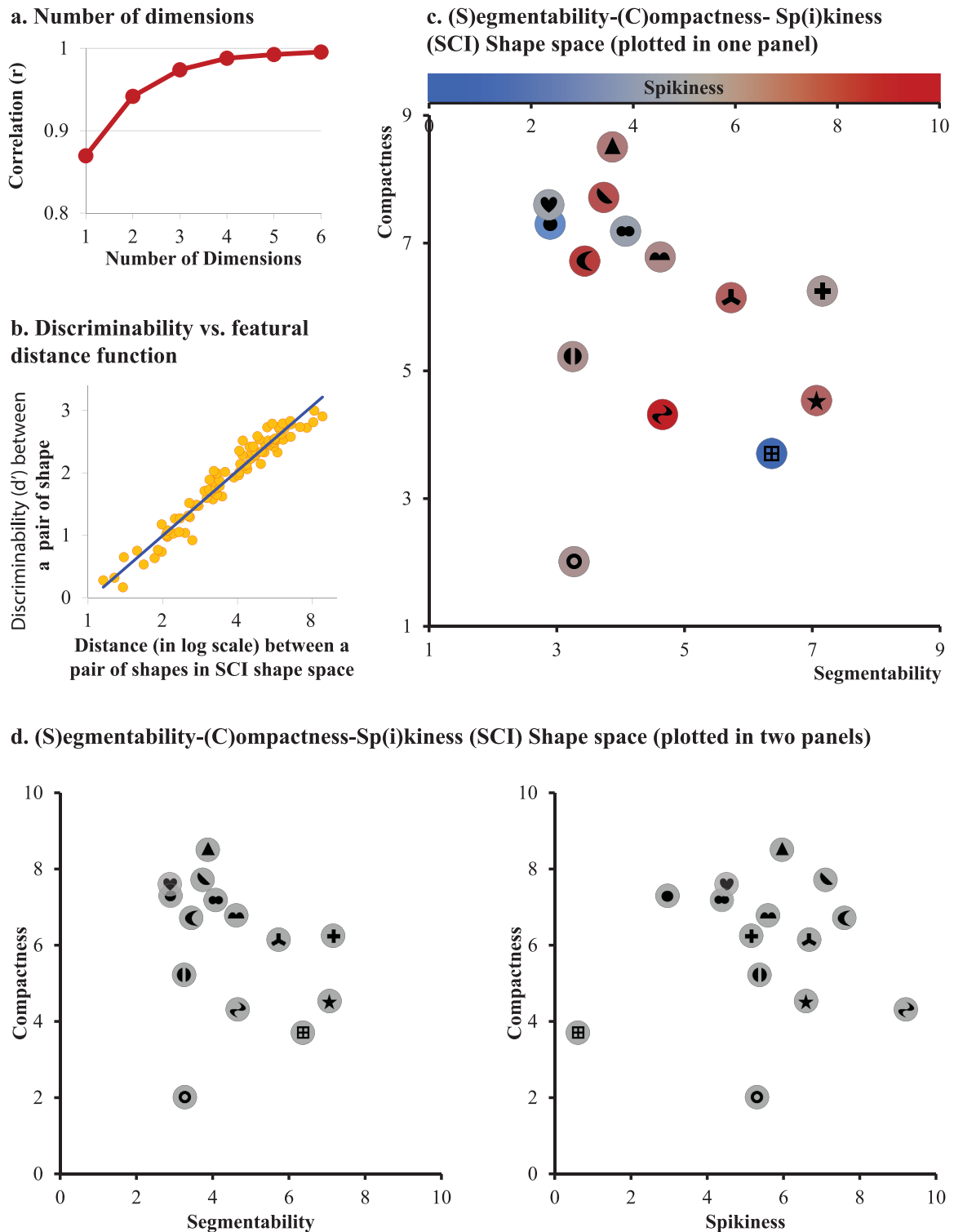


Figure 4. Results and analysis. For number of dimensions = 1 to 6, I used an optimization algorithm to determine the best model (i.e., highest correlation). It seems that the shape space is most reasonably characterized by three dimensions because the fitting is already very good ($r = 0.974$) and improves relatively little for the subsequent dimensions (a). Following the rationale illustrated in Figure 1a, there is a strong correlation between the log-distances between the 91 pairs of shapes in this 3D shape space and the 91 discriminabilities (d') between them (b). The best-fitting coordinates of the 14 shapes as 3D vectors were plotted in (c) (and also in (d) that plotted the same data in two graphs). The first dimension (x -axis of (c) and also x -axis of the left graph of (d)) matches well with the conceptual dimension of “segmentability” (e.g., \oplus vs. \bullet). The second dimension (y -axis of (c) and also y -axis of both graphs of (d)) matches well with the conceptual dimension of “compactness” (e.g., \circ vs. \blacktriangle). The third dimension (hue of bubbles of (c) and also x -axis of the right graph of (d)) matches well with the conceptual dimension of “spikiness” (e.g., \curvearrowright vs. \bullet). Together, they constitute a (s)egmentability-(c)ompactness-sp(i)kiness (SCI) space.

this dimension was initially identified by looking for the presence/absence of intersections, the nature of this dimension is reflected more precisely by the concept of “segmentability.”

Compactness

The second dimension is shown as the y -axis of Figure 4c (and also y -axis of both graphs of Figure 4d). It matches fairly well with the conceptual dimension of “closure” (i.e., holes) that has been well established in previous studies (Chen, 1982; Donnelly, Humphreys, & Riddoch, 1991; Elder & Zucker, 1993, 1994; Pomerantz & Pristach, 1989). Shapes including clear holes (⊕ and ○) scored very low on this dimension, and most of the other shapes scored high. Interestingly, three shapes (★, ●, and ↶) scored moderately. It seems that this dimension also reflects the “semicovered spaces” that can be included by connecting two points in the shape (e.g., gaps, or concavities; Hulleman, Winkel, & Boselie, 2000). In other words, although this dimension was initially identified by looking for the presence/absence of closures (holes), the notions of closure and concavity perhaps can be unified and should be more precisely described as a type of “compactness” index: the ratio of “the area of a shape” and “the area of the smallest convex shape covering it.” Therefore, this dimension is formally termed *compactness*. It is not entirely clear why ▲ scored higher than ● on this compactness dimension even if they should be equal (i.e., both 1) on the compactness index defined above. This may reflect noise or some conceptual aspect of this dimension that we do not yet understand. Future work will be needed to find out.

Spikiness

The third dimension is represented by the hues of bubbles in Figure 4c (and also x -axis of the right graph of Figure 4d), and it seems to reflect the *spikiness* of a shape: Those shown by the reddish bubbles (★, ↶, ◀, ↷, and ♣) have sharp spikes, whereas those shown by the bluish bubbles (●, ♥, ⊕, and ●) have smooth contours without any sharp spikes. This notion of spikiness is very consistent with the recent finding that both human observers and their lateral occipital complex (LOC) area are very sensitive to the spiky parts of the objects (de Beeck, Torfs, & Wagemans, 2008). Ecologically, the spikiness of an object perhaps indicates a potential danger of grasping an object,⁸ and this is perhaps why spikiness is processed as a basic visual feature. This spikiness dimension seems to be driven nonlinearly by the presence of at least one or two very sharp spikes, rather than linearly by the number of spikes. For example, the ↶ and ◀ score higher than ★, even if the latter has many more spikes.

This dimension of spikiness seems to be related to the concept of “line termination,” which has been proposed as a basic shape feature (Julesz, 1984; Julesz & Bergen, 1983). Clearly, a line termination is usually also a spiky part (e.g., ↶), so the previous findings concerning line termination can generally be explained in terms of spikiness. However, this third dimension of shape seems to be more generally consistent with the concept of spikiness than with the specific definition of line termination. For example, ● should score higher than ↶ in terms of the number of line terminations, but the former is arguably less spiky because the spiky parts do not point outward and are therefore not dangerous for grasping. It is not entirely clear why ⊕, which has four corners, scored even lower than a disk with a completely smooth surface (●). Perhaps, this is because it appears as a standard box that can be placed more stably and held more readily than a ball, so it is in some sense even less dangerous than a ball. Future work will be needed to explore this hypothesis.

(S)egmentability-(C)ompactness-Sp(i)kiness (SCI) shape space

To summarize, the three dimensions of this shape space respectively match the concept of segmentation, compactness, and spikiness. Therefore, taking the first letters of segmentability and compactness and the “I” from spikiness, this 3D shape space is named SCI shape space.

Overfitting

Overfitting is an important concern for complex models. For the present study, the correlation (r) that can be reached by the optimization algorithm will always improve with a larger number of dimensions because more dimensions always give more room to explain the variations in the data. However, some of these additional dimensions may have provided artificial explanations for “random noises” rather than substantive underlying mechanisms. It is therefore important to assess the issue of overfitting for the SCI shape space. Specifically, are three dimensions too many? Is the third dimension an artificial one or not?

The central problem of overfitting is the failure of making good predictions for new data. I considered two types of new data: the generalization to a new group of observers and the generalization to new shapes.

First, for generalization to a new group of observers, I used an out-of-sample fitting procedure to assess this problem. In this procedure, I randomly split the data into two parts (i.e., 93 observers for each), conducted an optimization on Part 1, and calculated the usual in-sample fitting index on which the optimization was based,

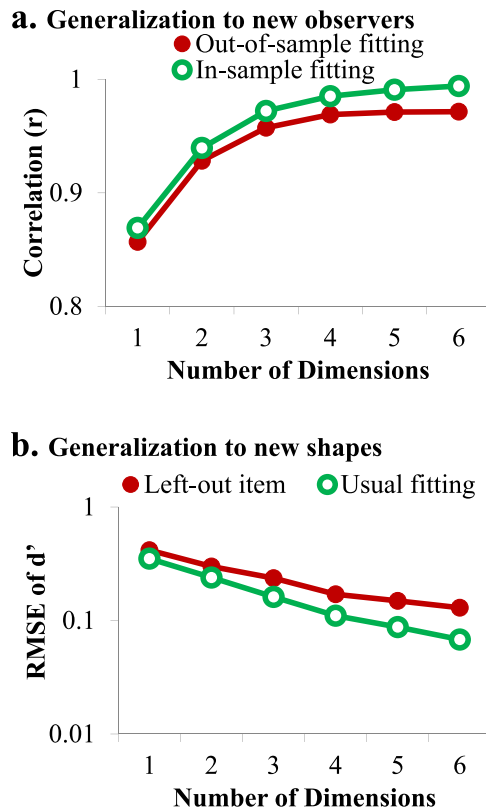


Figure 5. Assessments of overfitting. The central problem of overfitting is the failure of making good predictions for new data. For generalization to a new group of observers (a), I used an out-of-sample fitting procedure to assess this problem. Unsurprisingly, the in-sample fitting was better than the out-of-sample fitting, and the difference grew with more dimensions. However, it is also clear that the difference was relatively trivial ($r = 0.972$ vs. 0.957 for the 3D model), and this suggests that the SCI shape space can be generalized to a new group of observers. For generalization to new shapes (b), I used a leave-one-out fitting procedure to assess this problem. Unsurprisingly, when a shape was left out, the fitting of the shape was worse than the usual fitting, and the difference grew with more dimensions. However, it is also clear that the difference was modest (a 49% increase of RMSE), and this suggests that the SCI shape space can be generalized to new shapes.

$r_{(\log\text{-distances Part 1, discriminabilities Part 1)}$, as well as the out-of-sample fitting index, $r_{(\log\text{-distances Part 1, discriminabilities Part 2)}$. This was repeated 15 times, and the average results are plotted in Figure 5a. Unsurprisingly, the in-sample fitting was better than the out-of-sample fitting, and the difference grew with more dimensions. However, it is also clear that the difference was relatively trivial ($r = 0.972$ vs. 0.957 for the 3D model), and this suggests that the SCI shape space can be generalized fairly well to a new group of observers.

Second, for generalization to new shapes, the set of shapes is too small to be split into two subsets.

Therefore, to assess this issue, I used a leave-one-out fitting procedure to assess this problem. In this procedure, I took one “test shape” out and used optimization to fully determine a shape space on the basis of the other 13 shapes. Then I put the test shape back to see how well this test shape could fit with this previously optimized shape space (i.e., fitting of the 13 d' values relevant to this test shape) and compared it against the usual situation in which all 14 shapes were optimized together. Here, the linear correlation coefficient (r) could not be used as an unbiased index to compare the fitting of a subset to that of the whole set,⁹ so I used the root-mean-square error (RMSE). This procedure was repeated for all 14 shapes, and the average results are plotted in Figure 5b. Unsurprisingly, when a shape was left out, the fitting of the shape was worse than the usual fitting, and the difference grew with more dimensions. However, it is also clear that the difference was modest (a 49% increase of RMSE for the 3D model) and this suggests that the SCI shape space can be generalized reasonably well to new shapes. Of course, this leave-one-out fitting procedure only provides evidence for the generalization within this category of shapes, namely, other 2D filled geometrical shapes. With this restriction, it seems unlikely that the current diverse set has missed the variations in a whole dimension. In the “additional shape sets” section below, we will consider the issue of generalization to other types of shapes, such as 3D shapes and shapes of real-world objects (e.g., animal silhouettes, letters, and characters).

In addition to the quantified assessments, conceptual clarity is essential to the prevention of overfitting. As mentioned above, the three dimensions of SCI shape space have fairly clear conceptual interpretations. For comparison, the best-fitting two-dimensional model is plotted in Figure 6, and it is clearly inadequate and overly compressed. Some of the fundamentally different variations (\oplus vs. \bullet difference; \circ vs. \bullet difference) have been forced to reside on almost the same dimension, which is conceptually inappropriate. Therefore, 3D (i.e., SCI) shape space is conceptually better than the two-dimensional shape space.

It is worth mentioning that another important type of generalization is the generalization to new tasks. The present study has only tried one task. Therefore, we cannot really come up with quantified assessments on how well the SCI will generalize to other tasks. Conceptually, I predict that the SCI should be able to generalize very well to other attention-demanding tasks (e.g., visual search), but probably only partly to one-to-one comparison tasks (e.g., pairwise comparison or same-different discrimination) because the latter tasks measures postattentive processing of shapes. It will be interesting to see how much (and what aspect) of SCI can generalize to these one-to-one comparison tasks.

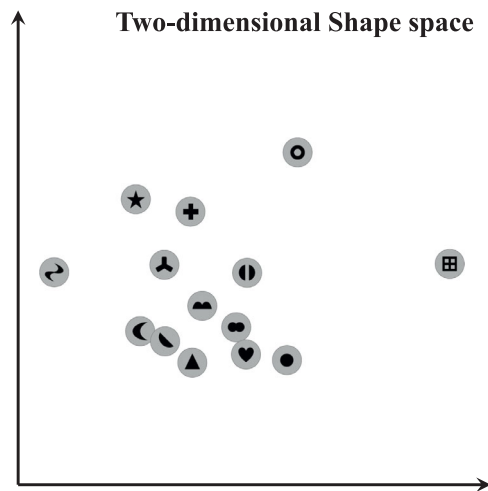


Figure 6. The best-fitting two-dimensional shape space. It is clearly inadequate and overly compressed. Some of the fundamentally different variations (+ vs. • difference; ○ vs. ● difference) have been forced to reside on almost the same dimension, which is conceptually inappropriate. Therefore, 3D (i.e., SCI) shape space is conceptually better than the two-dimensional shape space.

Underfitting

Although overfitting is a more prevalent problem in modeling, underfitting is also an important potential issue that needs to be considered. Underfitting occurs when a model cannot adequately capture an underlying trend of the data (e.g., a linear model is used on data with an apparent nonlinear trend). For the present study, there is indeed a slightly nonlinear trend in the fitting of SCI shape space (see Figure 4b), and the fitting (r) can be improved from 0.974 to 0.976 if a quadratic function is used. Although this nonlinearity has reached significance ($p < 0.01$), the magnitude of improvement is too trivial, so I decided that, without a good conceptual interpretation, it should not be implemented in SCI shape space.

It should also be considered whether a fourth dimension should be added to the SCI shape space. The best-fitting four-dimensional model is plotted in Figure 7, and the layout is rotated so that the first three dimensions approximately match the three dimensions of the SCI shape space.¹⁰ As can be seen in the y -axis of Figure 7b, the fourth dimension is not associated with any clear conceptual interpretation. Therefore, 3D (i.e., SCI) shape space is also conceptually better than the four-dimensional shape space.

Theoretical implications

The most important theoretical contribution of the SCI space is determining the layout of the space of

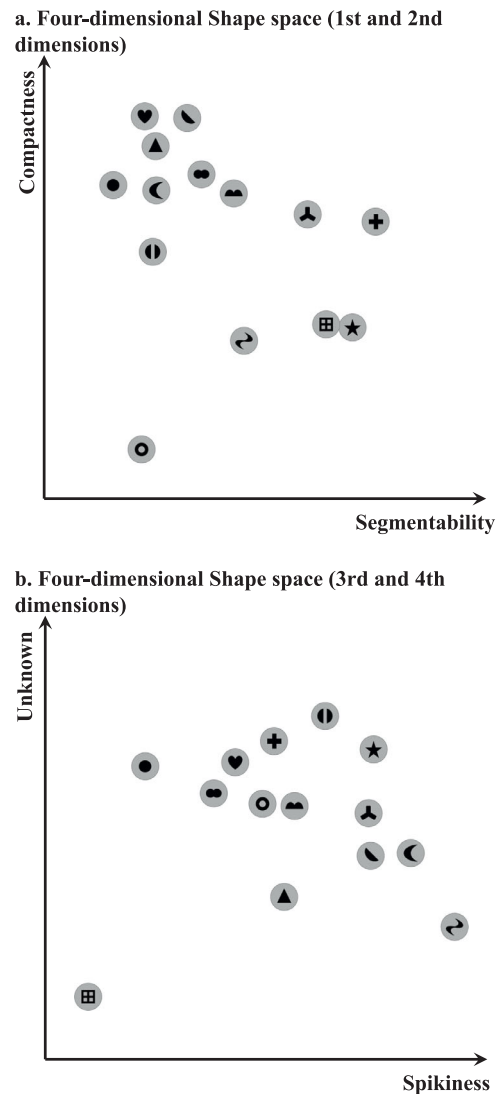


Figure 7. The best-fitting four-dimensional shape space. The layout is rotated so that the first three dimensions approximately match the three dimensions of the SCI shape space. As can be seen in the y -axis of (b), the fourth dimension is not associated with any clear conceptual interpretation. Therefore, 3D (i.e., SCI) shape space is conceptually better than the four-dimensional shape space.

preattentive shape features. An important merit of the SCI space is its very good fit to empirical data ($r = 0.974$). This suggests that, just like colors, shapes are governed by a simple and lawful space.

This SCI space advances our knowledge in two ways. First, previous studies on preattentive shape features have mainly focused on identifying several isolated conceptual distinctions (e.g., the presence/absence of a closure, an intersection, or a line termination). The SCI space has not only confirmed these dimensions but, more important, has also placed them in a unified and quantified framework. Second, the present finding points to important modifications to these concepts,

from “intersection” to “segmentability,” from “closure” and “concavity” to “compactness,” and from “line termination” to “spikiness.”

Practical implications

This SCI shape space also has potential applications. A practical guide is provided in Figure 8a, which lists six shapes that are most well separated in SCI shape space. These shapes are listed in descending orders. Therefore, for a set of three clearly distinguishable shapes, the first three should be used and so forth. These can be used by researchers who want to find a set of distinguishable shape items.

This set of distinguishable shape items can be used to improve the discriminabilities between shape symbols. One obvious potential case is the symbols used in scatterplots. Figures 8b and 8c respectively show sample displays in which default symbols of Microsoft Excel and those of Tableau are used, whereas Figure 8d shows one in which these six distinguishable shapes are used. It seems that the series of data can be distinguished more easily from each other in Figure 8d than in Figures 8b and 8c.

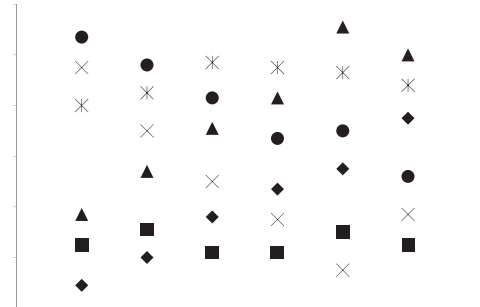
In recent years, there have been growing interests in an in-depth integration between the basic studies of human perception and the applied studies on data visualization (e.g., Nothelfer, Gleicher, & Franconeri, 2017; Szafir et al., 2016). The SCI shape space also has the immediate potential for making such connections. For example, the default set of shape items in professional data visualization software (e.g., Tableau) has included more unique shapes than those in Microsoft Excel. However, as shown above, they can be further enriched by using the distinguishable shapes in Figure 8a. Specifically, the first six Tableau shapes can be divided into two sets (Set 1: \diamond , \circ , \square ; Set 2: $+$, \times , \times), and the three items in either set are not very distinguishable from each other because they do not differ greatly on segmentability, compactness, or spikiness. Therefore, this set can be improved by replacing some of the existing shapes with the discriminable shapes mentioned above (i.e., \curvearrowright , \blacktriangleleft , \bullet , \boxplus).

Two recent studies on data visualization are particularly relevant. First, Burlinson, Subramanian, and Goolkasian (2018) have successfully applied the perceptual distinction between open and closed shapes (i.e., the feature “closure” or “hole” that has been included in the dimension compactness in the present study) to the domain of data visualization. Specifically, their observers were presented two intermixed set of shapes and were asked to perform three tasks that were similar to typical visualization tasks: (a) average value (determine which of the two sets of shapes had a higher position on the y -axis), (b) numerosity (determine which of the two sets of shapes contained

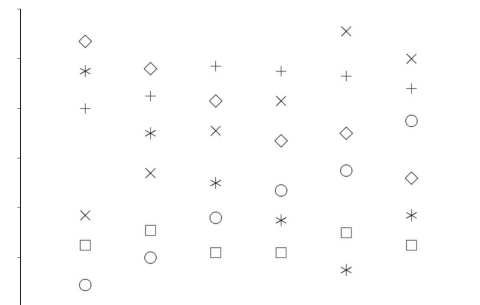
a. Shapes that are well-separated in SCI space



b. Sample display (Microsoft Excel shapes)



c. Sample display (Tableau shapes)



d. Sample display (six shapes above)

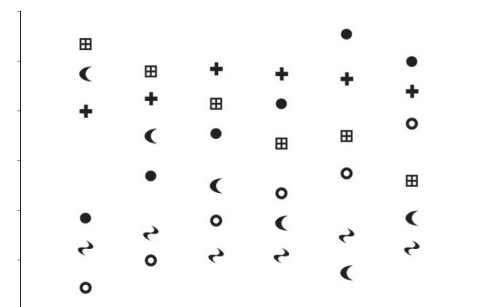


Figure 8. Clearly distinguishable shapes. Shapes are often used as signs and symbols. For these purposes, it is usually desirable to use shapes that are clearly distinguishable from each other. Panel (a) provides a practical guide of 6 shapes that are most well separated in SCI shape space. These shapes are listed in descending order of distinguishability. Therefore, for a set of three clearly distinguishable shapes, the first three should be used, and so forth. Panels (b) and (c) respectively show sample displays in which default symbols of Microsoft Excel and those of Tableau are used, whereas (d) shows one in which these six distinguishable shapes are used. It seems that the series of data can be distinguished more easily from each other in (d) than in (b) and (c).

more elements), and (c) trend judgments (determine which set of shapes showed a linear relationship). Their results confirmed that differentiating the two sets of shapes with an open-closed distinction consistently improved performance in these tasks. Future work should test the other two shape dimensions in the SCI space (segmentability and spikiness) on the tasks used by [Burlinson et al. \(2018\)](#) to verify that those distinctions can also improve the efficiency of these data visualization tasks. While the tasks of [Burlinson et al. \(2018\)](#) involved more complex processing (further processing/judgments on the selected subset) compared to the present tasks, differentiating sets by differences in segmentability and spikiness should still improve performance, because the core processing bottleneck should still be the shape-based selection of a subset of items in the display.

Second, [Demiralp, Bernstein, and Heer \(2014\)](#) have measured “perceptual kernels” (perceptual distance matrices between a set of visual items) from various data visualization tasks. In [Demiralp et al.’s \(2014\)](#) terms, the present study essentially has measured a perceptual kernel consisting of 14 shapes. However, there are two important differences between [Demiralp et al.’s \(2014\)](#) perceptual kernels and the SCI shape space. First, their tasks are a one-to-one comparison (i.e., pairwise rating) rather than attention-demanding tasks, so their perceptual kernels may be more closely related to postattentive shape processing as discussed in the introduction. Second, the present study was intentionally designed to capture the diverse variations in the whole shape space, but [Demiralp et al. \(2014\)](#) have focused on several specific points of the shape space (e.g., triangles, crosses). Accordingly, they found that these shape items form a few clear-cut clusters (e.g., the triangles or the crosses) in the shape space (i.e., palette). Given the two differences, it is difficult to align the shape spaces between these two studies. Nevertheless, it seems the shape space or palette in [Demiralp et al.’s \(2014\)](#) figure 1 can be approximately mapped to the SCI space of the present study. Specifically, the present set of shapes (and perhaps some variations on these items in ways that are similar to [Demiralp et al., 2014](#)) should be tried in these one-to-one comparison data visualization tasks to see whether their findings can be generalized to shape items in other parts of shape space.

Other shape sets?

As discussed above, in addition to the 2D geometrical filled shapes, other important shape categories should be considered in the future. The list includes real-world objects (e.g., animals), letters and characters, 3D shapes, and 2D unfilled line-drawing shapes.

First, for real-world objects such as animals, we can perhaps create their silhouettes as experimental

stimuli, so that the processing is mainly based on shape rather than other cues (e.g., colors, spatial frequency distributions). In addition, letters and characters are technically shapes and they are very important to us functionally. As discussed in [Wolfe and Horowitz \(2017\)](#), there is weak or no evidence for these types of real-world stimuli to play any special role in guiding attention. If this is true, then these shape sets will basically reveal the same SCI shape space as the present experiment does. Future studies will be needed to either confirm this hypothesis or reveal new dimensions or new findings in some other way (e.g., asymmetry of dimensions).

Second, many researchers study 3D shapes (e.g., [Pizlo, 2010](#)), including both real-world objects (e.g., a rabbit viewed from different perspectives) and artificial objects. Given the limited support for the 3D shapes as preattentive visual features ([Enns & Rensink, 1990](#); but see [Zhang et al., 2015](#)), it seems doubtful that they will make any special contribution to the space of preattentive shape features other than what can be extracted from their 2D profiles. Future studies will be needed to either confirm this hypothesis or reveal new dimensions or other new findings.

Third, the shape space revealed from 2D unfilled line-drawing shapes may differ significantly from what is found in the present study. Specifically, the dimension of compactness may be weakened because all shapes now contain a large portion of “empty spaces.” Future studies will be needed to retest the line-drawing versions of the present 14 shapes. The results can shed important light on the question of how visual mechanisms recover surfaces from line-drawing shapes.

Shape dimensions in postattentive processing of shapes

In terms of the above-stated distinction between preattentive shape features and postattentive processing of shapes, the studies of visual shapes in the area of object recognition generally fall into the category of postattentive shape processing. In these studies, observers usually perform a one-to-one comparison of the detailed shapes.

Some of these studies have studied the dimensions of shape processing. But they usually dealt with specific predefined dimensions rather than attempted to reveal all the shape dimensions. For example, [Ons, De Baene, and Wagemans \(2011\)](#) studied whether two predefined dimensions (aspect ratio and medial axis curvature) are integrally encoded or not. Clearly, these two dimensions were not intended to cover all the dimensions of shape processing.

More generally, it seems doubtful that there could be a general-purpose several-dimensional space for

postattentive shape processing after all. By definition, the detailed shape is an irreducible high-dimensional vector. A shape as simple as an irregular hexagon has 10 free dimensions (x and y for six vertexes minus the x and y of the overall object), and there is no way they can be reduced because two irregular hexagons will only be identical if they are matched on all these dimensions. So there is no hope of building a 3D shape space. In addition, these dimensions depend on the subclass of shapes, and the shape space for each of these subclasses would necessarily be different. Therefore, it needs to be clearly stated that SCI is only a space of preattentive shape features and is not intended to be a general-purpose shape space for postattentive processing of shapes in tasks requiring one-to-one comparison of visual details.

Having made this disclaimer, it does seem likely that a lot of useful insights can be learned from studies of postattentive processing of shapes, as will be discussed in the next section.

Connections to previous studies on postattentive shape processing

The previous studies on postattentive shape processing have a few important implications on the SCI shape space.

First, an important theme in studies of shape processing is to try to develop formal ways of defining the elements of a shape (or elementary processes in processing a shape). [Pentland \(1986\)](#) and [Biederman \(1987\)](#) respectively proposed superquadrics and geons as building elements of shapes. They can certainly account for a lot of variations in the concept of shape. However, as [Hoffman and Singh \(1997\)](#) pointed out, they tend to be “limited in scope.” [Leyton \(1988\)](#) tried to come up with a “grammar” of describing the shapes in terms of protrusion, indentation, squashing, and internal resistance. It offers some potential insights on the ways of understanding the origin of shape dimensions from a functional point of view. Subsequently, [Hoffman and Singh \(1997\)](#) developed a way of using negative minima to define the parts of an object, and [Feldman and Singh \(2005\)](#) formally defined the Shannon information on the border of objects. These quantitative models have the advantage of making precise predictions. For example, [Feldman and Singh \(2005\)](#) showed that a minimum carries more information than a maximum, which perhaps explains why the concavity is a basic shape feature ([Hulleman, Winkel, & Boselie, 2000](#)). Future development of SCI space should be developed in a similar direction. Specifically, image-based quantitative models should be developed to try to define formally and extract the SCI values from the stimulus items and fit those to

the behavior data. The segmentability can perhaps be defined in terms of the “number of parts” and the “strength of connection,” which can be quantified by the values of minimum points on the border. Compactness can perhaps be defined in terms of the ratio of the area of the shape to the area of the smallest convex shape covering it. Spikiness can perhaps be defined in terms of the values of one or two global maxima on the border. Obviously, the present set of 14 shapes is too limited for this purpose. A much larger set of shape items will be required.

Second, in another set of studies, researchers ([Cortese & Dyre, 1996](#); [Shepard & Cermak, 1973](#); [Wilkinson, Wilson, & Habak 1998](#); [Zahn & Roskies, 1972](#)) defined a radial frequency analysis (i.e., Fourier analysis in polar-coordinate) and showed that perceived similarities among a set of shapes match well with the parameter space defined by this analysis. On the one hand, this notion of radial frequency seems relevant to some or even all the three dimensions proposed in the present study. Higher-frequency shapes perhaps tend to form more segments, to have more semicovered spaces, and to be spikier. On the other hand, the difference between these two lines of studies is also obvious. The 81 shapes used in [Shepard and Cermak \(1973\)](#) seem to differ from each other only in terms of detailed shapes but not in terms of any preattentive shape feature. The shapes in [Wilkinson, Wilson, and Habak \(1998\)](#) do seem to differ from each other preattentively. But it seems those differences boil down to differences in spatial frequency, which is something we try to exclude as a confounding factor in this study. As discussed above, the present study has intentionally avoided using a coherent class of shapes, so it will be difficult to test the applicability of radial frequency analysis directly. But future work should be conducted to explore the potential connection between these two lines of studies.

Third, a large set of studies has shown that the shapes are encoded in V2, V4, IT, and LOC areas in monkey and human observers ([Anzai, Peng, & Van Essen, 2007](#); [de Beeck, Torfs, & Wagemans, 2008](#); [de Beeck, Wagemans, & Vogels, 2001](#); [Hegd e & Van Essen, 2000](#); [2005](#); [Pasupathy & Connor, 2002](#); [Yamane, Carlson, & Connor, 2008](#)). Generally, these findings make it plausible that there are one or more shape maps in the brain that will encode shapes as preattentive features. These studies have not systematically explored the dimensions in the encoding of these neurons (or brain areas), so it is difficult to directly compare these two lines of studies. However, there seems to be evidence in favor of coding of segmentability ([Anzai et al., 2007](#); [Hegd e & Van Essen, 2000](#); [Pasupathy & Connor, 2002](#)), compactness ([Hegd e & Van Essen, 2000](#)), and spikiness ([de Beeck, Torfs, & Wagemans, 2008](#)).

Residual and potential mechanisms

Although the SCI shape space fits the data very well, the fitting ($r = 0.974$) is still less than perfect. A large portion of the fitting errors is likely to be simply caused by noises in measurement. Nevertheless, it is still worth considering potential residual mechanisms and, when possible, assessing their contributions. To start, the average residual errors are plotted (in Figure 9a) for each shape as the target item and also for each shape as the background item. With one exception (▲), there appears to be a negative relationship between the average residual errors of a shape as the target item and of that shape as the background item.

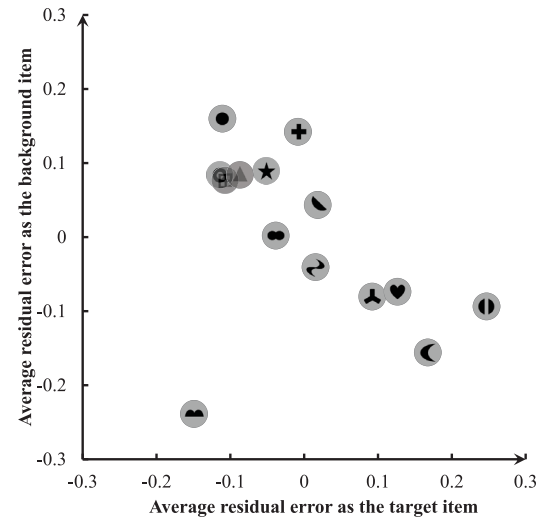
To see why these items differ from each other in this way, the difference between being the target and background items is plotted in Figure 9b. The residual errors are clearly affected by the rotational symmetry of a shape, namely, how much a shape looks the same after rotations. For the six shapes with the largest negative differences, four of them are identical for all four orientations (⊞, +, ●, and ○), whereas for each of the other two shapes (▲ and ★), the four appearances of the four orientations do not appear to be visually very different from each other (see Figure 9b). On the other hand, for each of the two shapes with the largest positive differences (◐ and ◑), the four (or two) appearances of the different orientations do appear to be very different from each other.

Therefore, a plausible account for these target/background differences is that they reflect the degrees of homogeneity when a shape fills in the target array (or the background). This is clearly related to the classic finding on similarity effect in visual search (Duncan & Humphreys, 1989). Duncan and Humphreys (1989) found that visual search performance is enhanced by the similarity between the nontarget items, and this is because it is easier to ignore a homogeneous array as the background. In the present study, a more homogeneous background region also leads to better performance because it is easier to ignore. On the other hand, a more homogeneous target region leads to worse performance because it is easier to (mistakenly) ignore the homogeneous target region.

The target/background difference (as shown in Figure 9b) is correlated fairly strongly ($r = 0.58$) with the residual errors in the 182 conditions and is clearly an important residual mechanism.

There are other residual or potential mechanisms. First, unlike colors, shapes cannot be defined in a way that is perfectly orthogonal from other features. For example, the items differ in their sizes, and there are various aspects of it (e.g., area of the black region, the diameter of minimum covering circle). The items also differ in terms of the spatial frequency components. It is impossible to control all of them simultaneously.

a. Average residual error as a target/background item.



b. Difference for being target vs. background items

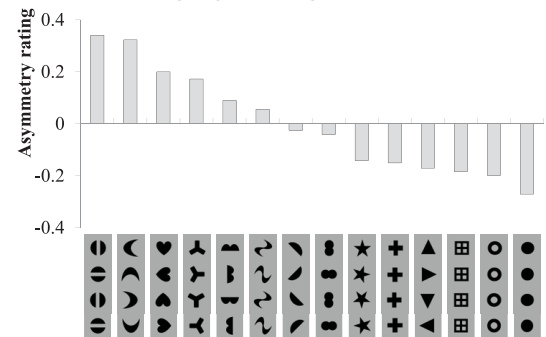










Figure 9. Residual errors. Panel (a) plotted the average residual errors for each shape as the target item and also for each shape as the background item. Panel (b) plotted the difference between being the target and background items. A positive value indicates that the performance tended to be better when a shape was the target shape than was the background shape. The residual errors are clearly affected by the rotational symmetry of a shape, namely, how much a shape looks the same after rotations. For the six shapes with the largest negative differences, four of them are identical for all four orientations (⊞, +, ●, and ○), whereas for each of the other two shapes (▲ and ★), the four appearances of the four orientations do not appear to be visually very different from each other. On the other hand, for each of the two shapes with the largest positive differences (◐ and ◑), the four (or two) appearances of the different orientations do appear to be very different from each other.

Nevertheless, these factors are unlikely to have played important roles. For one thing, if one or more items do enjoy such an extra difference (e.g., noticeably larger than other items), then use of these items will lead to better performance both as the target item and as the background item. Clearly, there is no such item in Figure 9a.

Second, previous studies showed that the asymmetry of dimensions of visual features can have substantial impacts on observers' performances (Treisman & Souther, 1985). If such feature asymmetry does contribute significantly to the performances, we would have seen clear indications in Figure 9b. For example, if there is a noticeable asymmetry for spikiness, we would have seen a significant correlation between the shape items' asymmetry ratings (as shown in Figure 9b) and their coordinates on the dimension of spikiness. However, the correlation is not significant between the asymmetry rating and any of the SCI dimensions. Therefore, feature asymmetry is unlikely to have played an important role.

Third, some of the items may be more familiar to the observers. For example, the shape  resembles the Chinese character 田 (meaning "field") and is familiar to the observers. Again, this factor is unlikely to have played an important role: The shape  is not special in Figure 9a.

Fourth, the present set of shape items has some degree of symmetry, either axial or rotational or both. So perhaps this special set is not representative of the general shapes that are usually not symmetric. Although this issue should be explored in the future, it is unlikely to have made a significant difference in the layout of the SCI shape space. Generally speaking, although symmetry has been shown to play a very important role in the processing of visual shapes (Wagemans, 1997, 2003), these studies usually deal with the postattentive processing of shapes. In other words, these studies have demonstrated that symmetry judgment of a single visual pattern is parallel and effortless without the scrutiny of the elements in this pattern¹¹ but have *not* demonstrated that symmetry of a visual pattern is a preattentive feature in the sense that multiple visual patterns can be processed in parallel. The only study that set out to test this question (Olivers & van der Helm, 1998) asked observers to search for an symmetric pattern among several asymmetric patterns (or do the opposite) and found that the times required to accomplish this task are highly dependent on the number of distractor patterns (often greater than 100 ms/item). From these results, symmetry does not appear to be a preattentive visual feature. In addition, the potential role of symmetry in SCI shape space can also be tentatively assessed in the present results. If symmetry is a dimension of preattentive shape feature, there will be an additional advantage to distinguish between items that are more symmetric (i.e., identical appearance across all four orientations: , , , and ) and items that are less so (i.e., very different appearance in these four orientations:  and ) and this would have emerged as a separate dimension on the shape space. However, there is no indication of this happening, confirming the opinion that symmetry is not a preattentive visual feature (Olivers & van der Helm, 1998).

Fifth, in the current modeling, I assumed that d' is a perfectly linear measure for the discriminability, and the data have largely confirmed this. However, this assumption may be violated for various reasons (e.g., averaging across good performers and bad performers, making occasional errors for nonperceptual reasons). Unfortunately, it is impossible to estimate and correct the impacts of these potential factors because there were only four trials per condition per observer. Nevertheless, this factor is unlikely to have played an important role because the fitting of the model would otherwise be significantly worse.

Sixth, one difficulty for perfect control lies in the irreducible nature of the dimensions in the detailed shapes. The SCI shape space is meant to capture the roles of shapes as preattentive features, namely, the aspects of shapes that can be extracted in parallel and used to guide attention. However, it is not implausible that the one-to-one comparison of visual details will always play a trivial, but nonzero, role in the performance of even an attention-demanding task. Nevertheless, this factor is unlikely to have played an important role because the fitting of the model would otherwise be significantly worse.

Conclusion

In the present study, I attempted to reveal the basic dimensions of preattentive shape features by studying how shapes can be positioned relative to one another in a way that matches their perceived similarities. A great majority of mechanisms ($r = 0.974$) can be explained by a 3D shape space: segmentability, compactness, and spikiness. On the theoretical side, this study has placed previously identified dimensions (e.g., intersection, closure, line termination, concavity) in a unified and quantified framework and has pointed to important modifications to these concepts (intersection → segmentability, closure and concavity → compactness, line termination → spikiness). On the practical side, the study has provided a guide on finding a set of distinguishable shape items.

Keywords: shape, visual feature, visual attention, texture segregation

Acknowledgments

The work described in this article was fully supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (Project No. CUHK 14617615). The writing of this article was substantially helped by the Humanities and Social Sciences Prestigious Fellowship (CUHK

34000817), also from the Research Grants Council. Data and script for data analysis are available on the Open Science Framework project page at <https://osf.io/khgd/>. The author is grateful to Steve Franconeri, Jeremy Wolfe, James Pomerantz, and one anonymous reviewer for very helpful comments.

Commercial relationships: none.

Corresponding author: Liqiang Huang.

Email: lqhuang@psy.cuhk.edu.hk.

Address: Department of Psychology, Chinese University of Hong Kong, Hong Kong.

Footnotes

¹More recent studies have rarely focused on these shape features themselves but have nevertheless kept on using shapes as exemplars of features to study other important questions in visual attention and visual working memory (e.g., Bae & Flombaum, 2013; Cowan, Blume, & Saults, 2013; Gajewski & Brockmole, 2006; Huang, 2020a; Kong & Fougny, 2019; Ort, Fahrenfort, & Olivers, 2017; Yu, Xiao, Bemis, & Franconeri, 2019).

²Consistent with this conceptual distinction, there is also a critical methodological difference between these two ways of studying shape processing. Those that study shape as preattentive features usually use attention-demanding tasks (e.g., visual search, Wolfe, 1998; or texture segregation, Beck, 1966, 1980), which require the observers to process the shapes of a large array of objects, whereas those who study postattentive processing of shapes usually use tasks that require one-to-one comparison of shapes (e.g., object recognition task).

³Here it is assumed that the shape space is uniform and linear. It is certainly possible that the shape space is nonlinear in some way, and future explorations and adjustments may be needed. However, without prior evidence in favor of a specific type of nonlinear model, my initial model assumes linearity.

⁴The adjustment of sizes was done by asking 10 naive observers to adjust the sizes of these items until they appear to be equally large, and the logarithmic mean of the 10 results was used for each shape item. The adjustment of spatial frequencies was done by asking another 10 naive observers to pick out those items that appear to be in noticeably more fine-grained or more coarse-grained texture. Only one item (♣) was consistently chosen as being fine-grained (6 out of the 10 observers), so the lines were widened (to become the current version) to remove this apparent difference.

⁵One may notice that early researchers (e.g., Julesz) often referred to “effortless texture segregation.” So there is an apparent conflict between those statements and the present claim that the texture segregation is an attention-demanding task. Actually, although the wording may be the opposite, they actually intend to convey the same message. When referring to “effortless texture segregation,” what these researchers tried to express was not that the texture segregation task is always effortless for all types of visual stimuli but that it was an impressive finding that the texture segregation task, which is usually attention demanding, is effortless for some special types of visual stimuli.

⁶This is in some sense a “brute-force” algorithm that is possibly suboptimal in terms of computational efficiency. However, the present function is high-dimensional and nonlinear, making the application of the typical algorithms (e.g., gradient descent) potentially problematic. So it is overall a better solution to adopt this general-purpose “brute-force” algorithm. It should be mentioned that this optimization algorithm is just a tool to find the shape spaces (i.e., coordinates). It is the latter, not the former, that is the “model of the present study.” Therefore, this general-purpose straightforward algorithm is sufficient for the present purpose.

⁷Specifically, a run is terminated if one of two criteria was met. First, if the number of scans exceeded 1,000 in a run, then this run was terminated.

Second, whenever the step was halved, it was assessed whether the fitting index (r) was as good as expected (from previous best-fitting runs, allowing some zooms); if not, then this run was terminated. These criteria were implemented to throw out unpromising runs. Testing on simulated data showed that they are both effective (i.e., reasonably good hit) and safe (i.e., very low false alarm rate).

⁸This is very relevant to the recent studies that showed that attention is automatically affected by the convenience of grasping an object (e.g., Gozli, West, & Pratt, 2012).

⁹This is because the correlation (r) is an assessment of the global relationship. Therefore, if data are split into two subsets, the r of the entire set is not a linear combination of the two r of the two subsets.

¹⁰This matching is only approximate. This is because the present model is not a linear model. Therefore, a lower-dimension model is not exactly an orthogonal projection of a higher-dimension model.

¹¹Even this parallel and effortless perception of the symmetry of a single visual pattern is subject to certain limits. Huang and Pashler (2002; see also Huang & Pashler, 2007, for a generalized account) showed that multiple features in a single pattern have to be assessed sequentially. Huang, Pashler, and Junge (2004) showed that the symmetry judgment is coarse in the sense that the processing of elements is significantly reduced in denser patterns.

References

- Anzai, A., Peng, X., & Van Essen, D. C. (2007). Neurons in monkey visual area V2 encode combinations of orientations. *Nature Neuroscience*, *10*, 1313, <http://dx.doi.org/10.1038/nn1975>.
- Bae, G. Y., & Flombaum, J. I. (2013). Two items remembered as precisely as one how integral features can improve visual working memory. *Psychological Science*, *24*, 2038–2047.
- Beck, J. (1966). Perceptual grouping produced by changes in orientation and shape. *Science*, *154*, 538–540, <http://dx.doi.org/10.1126/science.154.3748.538>.
- Beck, J. (1982). Texture segregation. In J. Beck (Ed.), *Organization and representation in perception* (pp. 285–318). Hillsdale, NJ: Erlbaum.
- Biederman, I. (1987). Recognition-by-components— a theory of human image understanding. *Psychological Review*, *94*, 115–147, <http://dx.doi.org/10.1037/0033-295X.94.2.115>.
- Burlinson, D., Subramanian, K., & Goolkasian, P. (2018). Open vs. closed shapes: New perceptual categories? *IEEE Transactions on Visualization and Computer Graphics*, *24*, 574–583, <http://dx.doi.org/10.1109/TVCG.2017.2745086>.
- Chen, L. (1982). Topological structure in visual perception. *Science*, *218*, 699–700, <http://dx.doi.org/10.1126/science.7134969>.
- Cortese, J. M., & Dyre, B. P. (1996). Perceptual similarity of shapes generated from Fourier descriptors. *Journal of Experimental Psychology: Human Perception and Performance*, *22*, 133, <http://dx.doi.org/10.1037//0096-1523.22.1.133>.
- Cowan, N., Blume, C. L., & Saults, J. S. (2013). Attention to attributes and objects in working

- memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 39, 731, <http://dx.doi.org/10.1037/a0029687>.
- de Beeck, H. O., Wagemans, J., & Vogels, R. (2001). Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nature Neuroscience*, 4, 1244.
- de Beeck, H. P. O., Torfs, K., & Wagemans, J. (2008). Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. *Journal of Neuroscience*, 28, 10111–10123.
- Demiralp, C., Bernstein, M. S., & Heer, J. (2014). Learning perceptual kernels for visualization design. *IEEE Transactions on Visualization and Computer Graphics*, 20, 1933–1942, <http://dx.doi.org/10.1109/TVCG.2014.2346978>.
- Donnelly, N., Humphreys, G. W., & Riddoch, M. J. (1991). Parallel computation of primitive shape descriptions. *Journal of Experimental Psychology: Human Perception and Performance*, 17, 561, <http://dx.doi.org/10.1037//0096-1523.17.2.561>.
- Duncan, J., & Humphreys, G. W. (1989). Visual-search and stimulus similarity. *Psychological Review*, 96, 433–458.
- Elder, J., & Zucker, S. (1993). The effect of contour closure on the rapid discrimination of 2-dimensional shapes. *Vision Research*, 33, 981–991.
- Elder, J., & Zucker, S. (1994). A measure of closure. *Vision Research*, 34, 3361–3369, [http://dx.doi.org/10.1016/0042-6989\(94\)90070-1](http://dx.doi.org/10.1016/0042-6989(94)90070-1).
- Enns, J. T., & Rensink, R. A. (1990). Sensitivity to three-dimensional orientation in visual search. *Psychological Science*, 1, 323–326.
- Feldman, J., & Singh, M. (2005). Information along contours and object boundaries. *Psychological Review*, 112, 243, <http://dx.doi.org/10.1037/0033-295X.112.1.243>.
- Gajewski, D. A., & Brockmole, J. R. (2006). Feature bindings endure without attention: Evidence from an explicit recall task. *Psychonomic Bulletin & Review*, 13, 581–587, <http://dx.doi.org/10.3758/Bf03193966>.
- Gozli, D. G., West, G. L., & Pratt, J. (2012). Hand position alters vision by biasing processing through different visual pathways. *Cognition*, 124, 244–250.
- Hegde, J., & Van Essen, D. C. (2000). Selectivity for complex shapes in primate visual area V2. *Journal of Neuroscience*, 20, RC61–RC61.
- Hoffman, D. D., & Singh, M. (1997). Saliency of visual parts. *Cognition*, 63, 29–78.
- Huang, L. (2015a). Visual features for perception, attention, and working memory: toward a three-factor framework. *Cognition*, 145, 43–52, <http://dx.doi.org/10.1016/j.cognition.2015.08.007>.
- Huang, L. (2015b). Color is processed less efficiently than orientation in change detection but more efficiently in visual search. *Psychological Science*, 26, 646–652, <http://dx.doi.org/10.1177/0956797615569577>.
- Huang, L. (2015c). Visual features: Featural strength and visual strength are two dissociable dimensions. *Scientific Reports*, 5, 13769, <http://dx.doi.org/10.1038/srep13769>.
- Huang, L. (2020a). Unit of visual working memory: A Boolean map provides a better account than an object does. *Journal of Experimental Psychology: General*, 149, 1–30, <http://dx.doi.org/10.1037/xge0000616>.
- Huang, L. (2020b). *FVS 2.0: A framework for understanding the factors of visual processing in 25 tasks*. Manuscript in preparation.
- Huang, L., & Pashler, H. (2002). Symmetry detection and visual attention: A "binary-map" hypothesis. *Vision Research*, 42, 1421–1430, [http://dx.doi.org/10.1016/s0042-6989\(02\)00059-7](http://dx.doi.org/10.1016/s0042-6989(02)00059-7).
- Huang, L., & Pashler, H. (2007). A Boolean map theory of visual attention. *Psychological Review*, 114, 599–631.
- Huang, L., Pashler, H., & Junge, J. A. (2004). Are there capacity limitations in symmetry perception? *Psychonomic Bulletin & Review*, 11, 862–869.
- Hulleman, J., Te Winkel, W., & Boselie, F. (2000). Concavities as basic features in visual search: Evidence from search asymmetries. *Perception & Psychophysics*, 62, 162–174.
- Julesz, B. (1984). A brief outline of the texton theory of human vision. *Trends in Neurosciences*, 7, 41–45.
- Julesz, B. (1986). Texton gradients—the texton theory revisited. *Biological Cybernetics*, 54, 245–251, <http://dx.doi.org/10.1007/bf00318420>.
- Julesz, B., & Bergen, J. R. (1983). Textons, the fundamental elements in preattentive vision and perception of textures. *Bell System Technical Journal*, 62, 1619–1645.
- Kayaert, G., Biederman, I., Op de Beeck, H. P., & Vogels, R. (2005). Tuning for shape dimensions in macaque inferior temporal cortex. *European Journal of Neuroscience*, 22, 212–224, <http://dx.doi.org/10.1111/j.1460-9568.2005.04202.x>.
- Kong, G., & Fougny, D. (2019). Visual search within working memory. *Journal of Experimental Psychology: General*, 148, 1688–1700.
- Leyton, M. (1988). A process-grammar for shape. *Artificial Intelligence*, 34, 213–247.

- Macmillan, N. A., & Creelman, C. D. (2004). *Detection theory: A user's guide*. New York: Psychology Press.
- Nocedal, J., & Wright, S. (2006). *Numerical optimization*. New York: Springer Science & Business Media.
- Nothelfer, C., Gleicher, M., & Franconeri, S. (2017). Redundant encoding strengthens segmentation and grouping in visual displays of data. *Journal of Experimental Psychology: Human Perception and Performance*, *43*, 1667, <http://dx.doi.org/10.1037/xhp0000314>.
- Olivers, C. N. L., & van der Helm, P. A. (1998). Symmetry and selective attention: A dissociation between effortless perception and serial search. *Perception & Psychophysics*, *60*, 1101–1116.
- Ons, B., De Baene, W., & Wagemans, J. (2011). Subjectively interpreted shape dimensions as privileged and orthogonal axes in mental shape space. *Journal of Experimental Psychology Human Perception and Performance*, *37*, 422.
- Ort, E., Fahrenfort, J. J., & Olivers, C. N. (2017). Lack of free choice reveals the cost of having to search for more than one object. *Psychological Science*, *28*, 1137–1147.
- Pasupathy, A., & Connor, C. E. (2002). Population coding of shape in area V4. *Nature Neuroscience*, *5*, 1332, <http://dx.doi.org/10.1038/nn972>.
- Pentland, A. P. (1986). Perceptual organization and the representation of natural form. *Artificial Intelligence*, *28* 293–331.
- Pizlo, Z. (2010). *3D shape: Its unique place in visual perception*. Cambridge, MA: MIT Press.
- Pomerantz, J. R., & Pristach, E. A. (1989). Emergent features, attention, and perceptual glue in visual form perception. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 635, <http://dx.doi.org/10.1037//0096-1523.15.4.635>.
- Shepard, R. N., & Cermak, G. W. (1973). Perceptual-cognitive explorations of a toroidal set of free-form stimuli. *Cognitive Psychology*, *4*, 351–377.
- Szafir, D. A., Haroz, S., Gleicher, M., & Franconeri, S. (2016). Four types of ensemble coding in data visualizations. *Journal of Vision*, *16*, 11–11, <http://dx.doi.org/10.1167/16.5.11>.
- Treisman, A., & Souther, J. (1985). Search asymmetry: A diagnostic for preattentive processing of separable features. *Journal of Experimental Psychology: General*, *114*, 285.
- Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97–136.
- Wagemans, J. (1997). Characteristics and models of human symmetry detection. *Trends in Cognitive Sciences*, *1*, 346–352, [http://dx.doi.org/10.1016/S1364-6613\(97\)01105-4](http://dx.doi.org/10.1016/S1364-6613(97)01105-4).
- Wagemans, J. (1995). Detection of visual symmetries. *Spatial vision*, *9*(1), 9–32.
- Wilkinson, F., Wilson, H. R., & Habak, C. (1998). Detection and recognition of radial frequency patterns. *Vision Research*, *38*, 3555–3568, [http://dx.doi.org/10.1016/s0042-6989\(98\)00039-x](http://dx.doi.org/10.1016/s0042-6989(98)00039-x).
- Wolfe, J. M. (1994). Guided Search 2.0—a revised model of visual-search. *Psychonomic Bulletin & Review*, *1*, 202–238, <http://dx.doi.org/10.3758/BF03200774>.
- Wolfe, J. M. (1998). Visual search. In H. Pashler (Ed.), *Attention* (Vol. 1, pp. 13–73). London, UK: University College London Press.
- Wolfe, J. M., & Bennett, S. C. (1997). Preattentive object files: Shapeless bundles of basic features. *Vision Research*, *37*, 25–43, [http://dx.doi.org/10.1016/s0042-6989\(96\)00111-3](http://dx.doi.org/10.1016/s0042-6989(96)00111-3).
- Wolfe, J. M., & DiMase, J. S. (2003). Do intersections serve as basic features in visual search? *Perception*, *32*, 645–656.
- Wolfe, J. M., & Horowitz, T. S. (2004). What attributes guide the deployment of visual attention and how do they do it? *Nature Reviews Neuroscience*, *5*, 495.
- Wolfe, J. M., & Horowitz, T. S. (2017). Five factors that guide attention in visual search. *Nature Human Behaviour*, *1*, 0058.
- Yamane, Y., Carlson, E. T., Bowman, K. C., Wang, Z., & Connor, C. E. (2008). A neural code for three-dimensional object shape in macaque inferotemporal cortex. *Nature Neuroscience*, *11*, 1352–1360, <http://dx.doi.org/10.1038/nn.2202>.
- Yu, D., Xiao, X., Bemis, D. K., & Franconeri, S. L. (2019). Similarity grouping as feature-based selection. *Psychological Science*, *30*, 376–385.
- Zahn, C. T., & Roskies, R. Z. (1972). Fourier descriptors for plane closed curves. *IEEE Transactions on Computers*, *100*, 269–281.
- Zhang, X., Huang, J., Yigit-Elliott, S., & Rosenholtz, R. (2015). Cube search, revisited. *Journal of Vision*, *15*, 9, <http://dx.doi.org/10.1167/15.3.9>.
- Zusne, L. (1970). *Visual perception of form*. New York: Academic Press.