

Attention and Visual Memory in Visualization and Computer Graphics

Christopher G. Healey, *Senior Member, IEEE*, and James. T. Enns

Abstract—A fundamental goal of visualization is to produce images of data that support visual analysis, exploration, and discovery of novel insights. An important consideration during visualization design is the role of human visual perception. How we “see” details in an image can directly impact a viewer’s efficiency and effectiveness. This article surveys research on attention and visual perception, with a specific focus on results that have direct relevance to visualization and visual analytics. We discuss theories of low-level visual perception, then show how these findings form a foundation for more recent work on visual memory and visual attention. We conclude with a brief overview of how knowledge of visual attention and visual memory is being applied in visualization and graphics. We also discuss how challenges in visualization are motivating research in psychophysics.

Index Terms—Attention, color, motion, nonphotorealism, texture, visual memory, visual perception, visualization.

1 INTRODUCTION

HUMAN perception plays an important role in the area of visualization. An understanding of perception can significantly improve both the quality and the quantity of information being displayed [1]. The importance of perception was cited by the NSF panel on graphics and image processing that proposed the term “scientific visualization” [2]. The need for perception was again emphasized during recent DOE–NSF and DHS panels on directions for future research in visualization [3].

This document summarizes some of the recent developments in research and theory regarding human psychophysics, and discusses their relevance to scientific and information visualization. We begin with an overview of the way human vision rapidly and automatically categorizes visual images into regions and properties based on simple computations that can be made in parallel across an image. This is often referred to as preattentive processing. We describe five theories of preattentive processing, and briefly discuss related work on ensemble coding and feature hierarchies. We next examine several recent areas of research that focus on the critical role that the viewer’s current state of mind plays in determining what is seen, specifically, postattentive amnesia, memory-guided attention, change blindness, inattention blindness, and the attentional blink. These phenomena offer a perspective on early vision that is quite different from the older view that early visual processes are reflexive and inflexible. Instead, they highlight the fact that what we see depends critically on *where*

attention is focused and *what* is already in our minds prior to viewing an image. Finally, we describe several studies in which human perception has influenced the development of new methods in visualization and graphics.

2 PREATTENTIVE PROCESSING

For many years vision researchers have been investigating how the human visual system analyzes images. One feature of human vision that has an impact on almost all perceptual analyses is that, at any given moment detailed vision for shape and color is only possible within a small portion of the visual field (i.e., an area about the size of your thumbnail when viewed at arm’s length). In order to see detailed information from more than one region, the eyes move rapidly, alternating between brief stationary periods when detailed information is acquired—a fixation—and then flicking rapidly to a new location during a brief period of blindness—a saccade. This fixation–saccade cycle repeats 3–4 times each second of our waking lives, largely without any awareness on our part [4], [5], [6], [7]. The cycle makes seeing highly dynamic. While bottom-up information from each fixation is influencing our mental experience, our current mental states—including tasks and goals—are guiding saccades in a top-down fashion to new image locations for more information. *Visual attention* is the umbrella term used to denote the various mechanisms that help determine which regions of an image are selected for more detailed analysis.

For many years, the study of visual attention in humans focused on the *consequences* of selecting an object or location for more detailed processing. This emphasis led to theories of attention based on a variety of metaphors to account for the selective nature of perception, including filtering and bottlenecks [8], [9], [10], limited resources and limited capacity [11], [12], [13], and mental

- C. G. Healey is with the Department of Computer Science, North Carolina State University, Raleigh, NC, 27695.
E-mail: healey@csc.ncsu.edu
- J. T. Enns is with the Department of Psychology, University of British Columbia, Vancouver, BC, Canada, V6T 1Z4.
E-mail: jenns@psych.ubc.ca

effort or cognitive skill [14], [15]. An important discovery in these early studies was the identification of a limited set of visual features that are detected very rapidly by low-level, fast-acting visual processes. These properties were initially called preattentive, since their detection seemed to precede focused attention, occurring within the brief period of a single fixation. We now know that attention plays a critical role in what we see, even at this early stage of vision. The term preattentive continues to be used, however, for its intuitive notion of the speed and ease with which these properties are identified.

Typically, tasks that can be performed on large multi-element displays in less than 200–250 milliseconds (msec) are considered preattentive. Since a saccade takes at least 200 msec to initiate, viewers complete the task in a single glance. An example of a preattentive task is the detection of a red circle in a group of blue circles (Figs. 1a–b). The target object has a visual property “red” that the blue distractor objects do not. A viewer can easily determine whether the target is present or absent.

Hue is not the only visual feature that is preattentive. In Figs. 1c–d the target is again a red circle, while the distractors are red squares. Here, the visual system identifies the target through a difference in curvature.

A target defined by a unique visual property—a red hue in Figs. 1a–b, or a curved form in Figs. 1c–d—allows it to “pop out” of a display. This implies that it can be easily detected, regardless of the number of distractors. In contrast to these effortless searches, when a target is defined by the joint presence of two or more visual properties it often cannot be found preattentively. Figs. 1e–f show an example of these more difficult conjunction searches. The red circle target is made up of two features: red and circular. One of these features is present in each of the distractor objects—red squares and blue circles. A search for red items always returns true because there are red squares in each display. Similarly, a search for circular items always sees blue circles. Numerous studies have shown that most conjunction targets cannot be detected preattentively. Viewers must perform a time-consuming serial search through the display to confirm its presence or absence.

If low-level visual processes can be harnessed during visualization, they can draw attention to areas of potential interest in a display. This cannot be accomplished in an ad-hoc fashion, however. The visual features assigned to different data attributes must take advantage of the strengths of our visual system, must be well-suited to the viewer’s analysis needs, and must not produce visual interference effects (e.g., conjunction search) that mask information.

Fig. 2 lists some of the visual features that have been identified as preattentive. Experiments in psychology have used these features to perform the following tasks:

- *target detection*: viewers detect a target element with a unique visual feature within a field of distractor elements (Fig. 1),
- *boundary detection*: viewers detect a texture bound-

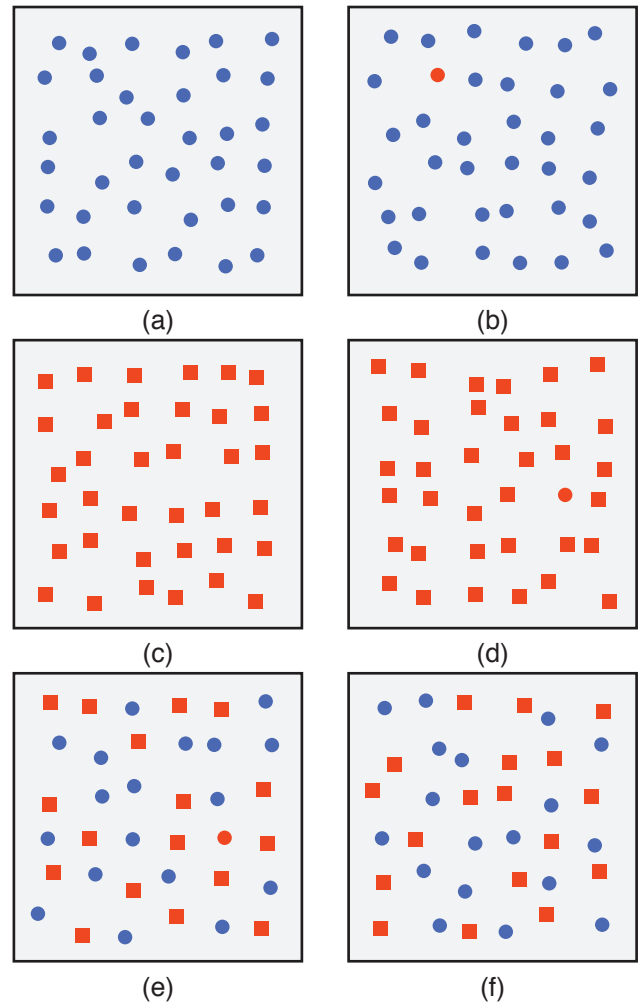


Fig. 1. Target detection: (a) hue target red circle absent; (b) target present; (c) shape target red circle absent; (d) target present; (e) conjunction target red circle present; (f) target absent

ary between two groups of elements, where all of the elements in each group have a common visual property (see Fig. 10),

- *region tracking*: viewers track one or more elements with a unique visual feature as they move in time and space, and
- *counting and estimation*: viewers count the number of elements with a unique visual feature.

3 THEORIES OF PREATTENTIVE VISION

A number of theories attempt to explain how preattentive processing occurs within the visual system. We describe five well-known models: feature integration, textons, similarity, guided search, and boolean maps. We next discuss ensemble coding, which shows that viewers can generate summaries of the distribution of visual features in a scene, even when they are unable to locate individual elements based on those same features. We conclude with feature hierarchies, which describe situations where the visual system favors certain visual

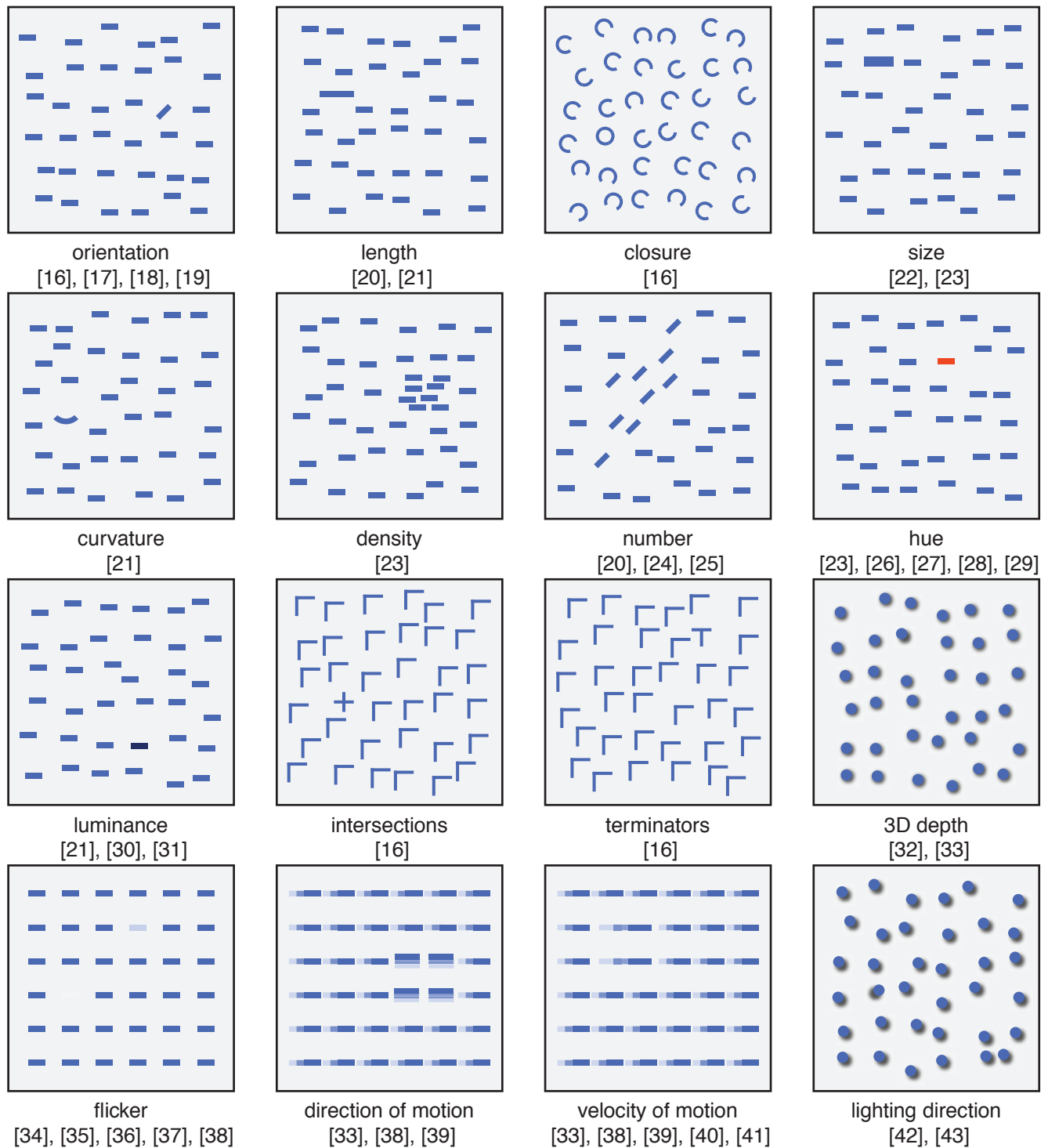


Fig. 2. Examples of preattentive visual features, with references to papers that investigated each feature's capabilities

features over others. Because we are interested equally in *where* viewers attend in an image, as well as to *what* they are attending, we will not review theories focusing exclusively on only one of these functions (e.g., the attention orienting theory of Posner and Petersen [44]).

3.1 Feature Integration Theory

Treisman was one of the first attention researchers to systematically study the nature of preattentive processing,

focusing in particular on the image features that led to selective perception. She was inspired by a physiological finding that single neurons in the brains of monkeys responded selectively to edges of a specific orientation and wavelength. Her goal was to find a behavioral consequence of these kinds of cells in humans. To do this, she focused on two interrelated problems. First, she tried to determine which visual properties are detected preattentively [21], [45], [46]. She called these properties

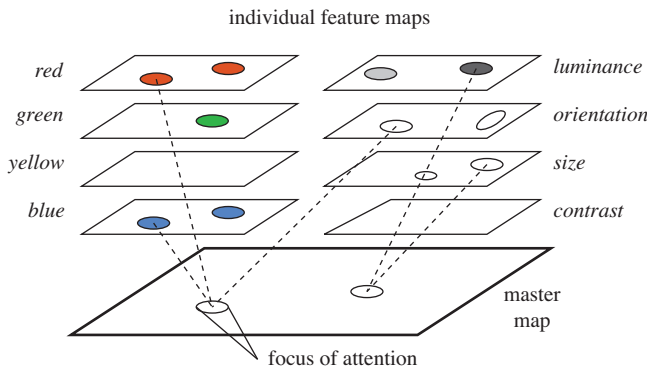


Fig. 3. Treisman's feature integration model of early vision—individual maps can be accessed in parallel to detect feature activity, but focused attention is required to combine features at a common spatial location [22]

“preattentive features” [47]. Second, she formulated a hypothesis about how the visual system performs preattentive processing [22].

Treisman ran experiments using target and boundary detection to classify preattentive features (Figs. 1 and 10), measuring performance in two different ways: by response time, and by accuracy. In the response time model viewers are asked to complete the task as quickly as possible while still maintaining a high level of accuracy. The number of distractors in a scene is varied from few to many. If task completion time is relatively constant and below some chosen threshold, independent of the number of distractors, the task is said to be preattentive (i.e., viewers are not searching through the display to locate the target).

In the accuracy version of the same task, the display is shown for a small, fixed exposure duration, then removed. Again, the number of distractors in the scene varies across trials. If viewers can complete the task accurately, regardless of the number of distractors, the feature used to define the target is assumed to be preattentive.

Treisman and others have used their experiments to compile a list of visual features that are detected preattentively (Fig. 2). It is important to note that some of these features are asymmetric. For example, a sloped line in a sea of vertical lines can be detected preattentively, but a vertical line in a sea of sloped lines cannot.

In order to explain preattentive processing, Treisman proposed a model of low-level human vision made up of a set of feature maps and a master map of locations (Fig. 3). Each feature map registers activity for a specific visual feature. When the visual system first sees an image, all the features are encoded in parallel into their respective maps. A viewer can access a particular map to check for activity, and perhaps to determine the amount of activity. The individual feature maps give no information about location, spatial arrangement, or relationships to activity in other maps, however.

This framework provides a general hypothesis that explains how preattentive processing occurs. If the target

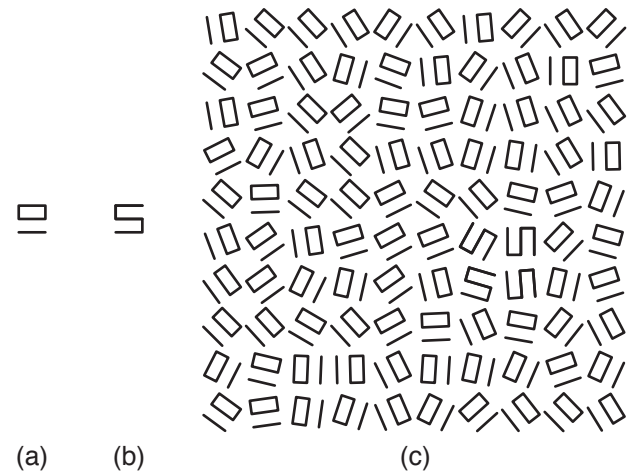


Fig. 4. Textons: (a,b) two textons *A* and *B* that appear different in isolation, but have the same size, number of terminators, and join points; (c) a target group of *B*-textons is difficult to detect in a background of *A*-textons when random rotation is applied [49]

has a unique feature, one can simply access the given feature map to see if any activity is occurring. Feature maps are encoded in parallel, so feature detection is almost instantaneous. A conjunction target can only be detected by accessing two or more feature maps. In order to locate these targets, one must search serially through the master map of locations, looking for an object that satisfies the conditions of having the correct combination of features. Within the model, this use of focused attention requires a relatively large amount of time and effort.

In later work, Treisman has expanded her strict dichotomy of features being detected either in parallel or in serial [21], [45]. She now believes that parallel and serial represent two ends of a spectrum that include “more” and “less,” not just “present” and “absent.” The amount of difference between the target and the distractors will affect search time. For example, a long vertical line can be detected immediately among a group of short vertical lines, but a medium-length line may take longer to see.

Treisman has also extended feature integration to explain situations where conjunction search involving motion, depth, color, and orientation have been shown to be preattentive [33], [39], [48]. Treisman hypothesizes that a significant target–nontarget difference would allow individual feature maps to ignore nontarget information. Consider a conjunction search for a green horizontal bar within a set of red horizontal bars and green vertical bars. If the red color map could inhibit information about red horizontal bars, the search reduces to finding a green horizontal bar in a sea of green vertical bars, which occurs preattentively.

3.2 Texton Theory

Julész was also instrumental in expanding our understanding of what we “see” in a single fixation. His start-

ing point came from a difficult computational problem in machine vision, namely, how to define a basis set for the perception of surface properties. Julesz's initial investigations focused on determining whether variations in order statistics were detected by the low-level visual system [37], [49], [50]. Examples included contrast—a first-order statistic—orientation and regularity—second-order statistics—and curvature—a third-order statistic. Julesz's results were inconclusive. First-order variations were detected preattentively. Some, but not all, second-order variations were also preattentive. as were an even smaller set of third-order variations.

Based on these findings, Julesz modified his theory to suggest that the early visual system detects three categories of features called textons [49], [51], [52]:

- 1) Elongated blobs—Lines, rectangles, or ellipses—with specific hues, orientations, widths, and so on.
- 2) Terminators—ends of line segments.
- 3) Crossings of line segments.

Julesz believed that only a difference in textons or in their density could be detected preattentively. No positional information about neighboring textons is available without focused attention. Like Treisman, Julesz suggested that preattentive processing occurs in parallel and focused attention occurs in serial.

Julesz used texture segregation to demonstrate his theory. Fig. 4 shows an example of an image that supports the texton hypothesis. Although the two objects look very different in isolation, they are actually the same texton. Both are blobs with the same height and width, made up of the same set of line segments with two terminators. When oriented randomly in an image, one cannot preattentively detect the texture boundary between the target group and the background distractors.

3.3 Similarity Theory

Some researchers did not support the dichotomy of serial and parallel search modes. They noted that groups of neurons in the brain seemed to be competing over time to represent the same object. Work in this area by Quinlan and Humphreys therefore began by investigating two separate factors in conjunction search [53]. First, search time may depend on the number of items of information required to identify the target. Second, search time may depend on how easily a target can be distinguished from its distractors, regardless of the presence of unique preattentive features. Follow-on work by Duncan and Humphreys hypothesized that search ability varies continuously, and depends on both the type of task and the display conditions [54], [55], [56]. Search time is based on two criteria: T-N similarity and N-N similarity. T-N similarity is the amount of similarity between targets and nontargets. N-N similarity is the amount of similarity within the nontargets themselves. These two factors affect search time as follows:

- as T-N similarity increases, search efficiency decreases and search time increases,

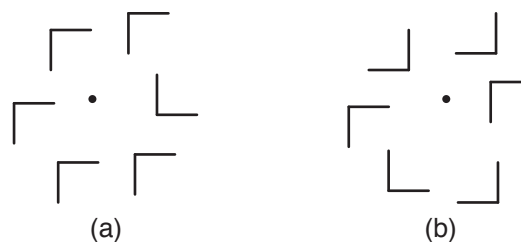


Fig. 5. N-N similarity affecting search efficiency for an L-shaped target: (a) high N-N (nontarget-nontarget) similarity allows easy detection of the target L; (b) low N-N similarity increases the difficulty of detecting the target L [55]

- as N-N similarity decreases, search efficiency decreases and search time increases, and
- T-N and N-N similarity are related; decreasing N-N similarity has little effect if T-N similarity is low; increasing T-N similarity has little effect if N-N similarity is high.

Treisman's feature integration theory has difficulty explaining Fig. 5. In both cases, the distractors seem to use exactly the same features as the target: oriented, connected lines of a fixed length. Yet experimental results show displays similar to Fig. 5a produce an average search time increase of 4.5 msec per distractor, versus 54.5 msec per distractor for displays similar to Fig. 5b. To explain this, Duncan and Humphreys proposed a three-step theory of visual selection.

- 1) The visual field is segmented in parallel into structural units that share some common property, for example, spatial proximity or hue. Structural units may again be segmented, producing a hierarchical representation of the visual field.
- 2) Access to visual short-term memory is a limited resource. During target search a template of the target's properties is available. The closer a structural unit matches the template, the more resources it receives relative to other units with a poorer match.
- 3) A poor match between a structural unit and the search template allows efficient rejection of other units that are strongly grouped to the rejected unit.

Structural units that most closely match the target template have the highest probability of access to visual short-term memory. Search speed is therefore a function of the speed of resource allocation and the amount of competition for access to visual short-term memory. Given this, we can see how T-N and N-N similarity affect search efficiency. Increased T-N similarity means more structural units match the template, so competition for visual short-term memory access increases. Decreased N-N similarity means we cannot efficiently reject large numbers of strongly grouped structural units, so resource allocation time and search time increases.

Interestingly, similarity theory is not the only attempt to distinguish between preattentive and attentive results based on a single parallel process. Nakayama and his

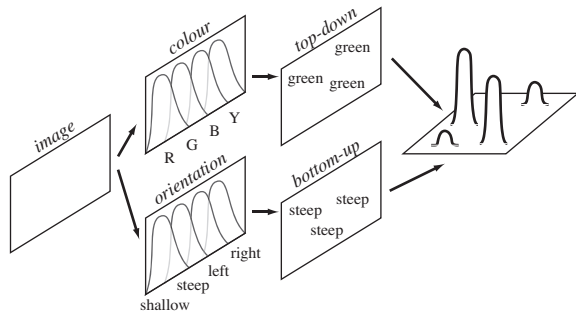


Fig. 6. Guided search for steep green targets, an image is filtered into categories for each feature map, bottom-up and top-down activation “mark” target regions, and an activation map combines the information to draw attention to the highest “hills” in the map [61]

colleagues proposed the use of stereo vision and occlusion to segment a three-dimensional scene, where preattentive search could be performed independently within a segment [33], [57]. Others have presented similar theories that segment by object [58], or by signal strength and noise [59]. The problem of distinguishing serial from parallel processes in human cognition is one of the longest-standing puzzles in the field, and one that researchers often return to [60].

3.4 Guided Search Theory

More recently, Wolfe et al. has proposed the theory of “guided search” [48], [61], [62]. This was the first attempt to actively incorporate the goals of the viewer into a model of human search. He hypothesized that an activation map based on both bottom-up and top-down information is constructed during visual search. Attention is drawn to peaks in the activation map that represent areas in the image with the largest combination of bottom-up and top-down influence.

As with Treisman, Wolfe believes early vision divides an image into individual feature maps (Fig. 6). Within each map a feature is filtered into multiple categories, for example, colors might be divided into red, green, blue, and yellow. Bottom-up activation follows feature categorization. It measures how different an element is from its neighbors. Top-down activation is a user-driven attempt to verify hypotheses or answer questions by “glancing” about an image, searching for the necessary visual information. For example, visual search for a “blue” element would generate a top-down request that is drawn to blue locations. Wolfe argued that viewers must specify requests in terms of the categories provided by each feature map [18], [31]. Thus, a viewer could search for “steep” or “shallow” elements, but not for elements rotated by a specific angle.

The activation map is a combination of bottom-up and top-down activity. The weights assigned to these two values are task dependent. Hills in the map mark regions that generate relatively large amounts of bottom-up or

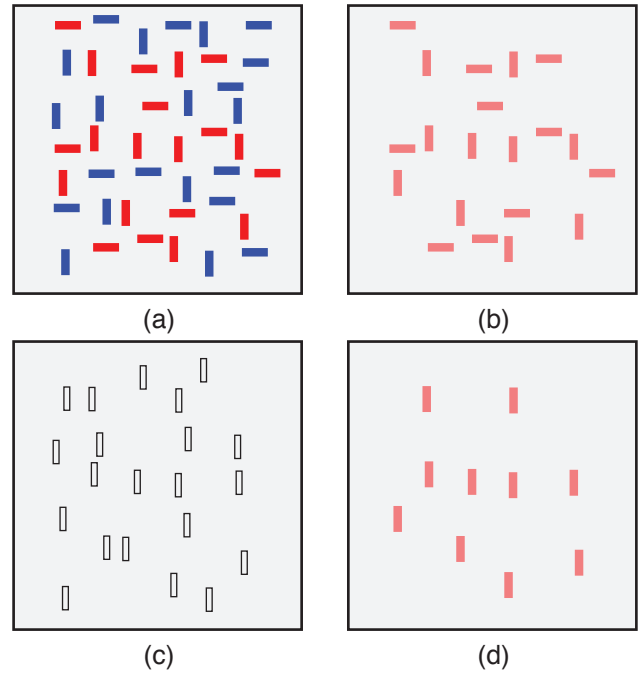


Fig. 7. Boolean maps: (a) red and blue vertical and horizontal elements; (b) map for “red”, color label is red, orientation label is undefined; (c) map for “vertical”, orientation label is vertical, color label is undefined; (d) map for set intersection on “red” and “vertical” maps [64]

top-down influence. A viewer’s attention is drawn from hill to hill in order of decreasing activation.

In addition to traditional “parallel” and “serial” target detection, guided search explains similarity theory’s results. Low N-N similarity causes more distractors to report bottom-up activation, while high T-N similarity reduces the target element’s bottom-up activation. Guided search also offers a possible explanation for cases where conjunction search can be performed preattentively [33], [48], [63]: viewer-driven top-down activation may permit efficient search for conjunction targets.

3.5 Boolean Map Theory

A new model of low-level vision has been presented by Huang et al. to study why we often fail to notice features of a display that are not relevant to the immediate task [64], [65]. This theory carefully divides visual search into two stages: selection and access. Selection involves choosing a set of objects from a scene. Access determines which properties of the selected objects a viewer can apprehend.

Huang suggests that the visual system can divide a scene into two parts: selected elements and excluded elements. This is the “boolean map” that underlies his theory. The visual system can then access certain properties of the selected elements for more detailed analysis.

Boolean maps are created in two ways. First, a viewer can specify a single value of an individual feature to select all objects that contain the feature value. For

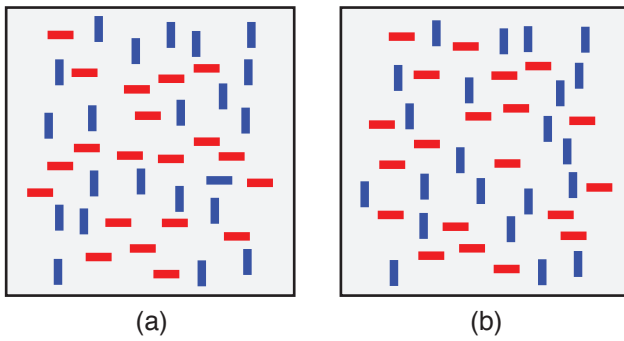


Fig. 8. Conjunction search for a blue horizontal target with boolean maps, select “blue” objects, then search within for a horizontal target: (a) target present; (b) target absent

example, a viewer could look for red objects, or vertical objects. If a viewer selected red objects (Fig. 7b), the color feature label for the resulting boolean map would be “red”. Labels for other features (e.g., orientation, size) would be undefined, since they have not (yet) participated in the creation of the map. A second method of selection is for a viewer to choose a set of elements at specific spatial locations. Here the boolean map’s feature labels are left undefined, since no specific feature value was used to identify the selected elements. Figs. 7a–c show an example of a simple scene, and the resulting boolean maps for selecting red objects or vertical objects.

An important distinction between feature integration and boolean maps is that, in feature integration, presence or absence of a feature is available preattentively, but no information on location is provided. A boolean map encodes the specific spatial locations of the elements that are selected, as well as feature labels to define properties of the selected objects.

A boolean map can also be created by applying the set operators union or intersection on two existing maps (Fig. 7d). For example, a viewer could create an initial map by selecting red objects (Fig. 7b), then select vertical objects (Fig. 7c) and intersect the vertical map with the red map currently held in memory. The result is a boolean map identifying the locations of red, vertical objects (Fig. 7d). A viewer can only retain a single boolean map. The result of the set operation immediately replaces the viewer’s current map.

Boolean maps lead to some surprising and counter-intuitive claims. For example, consider searching for a blue horizontal target in a sea of red horizontal and blue vertical objects. Unlike feature integration or guided search, boolean map theory says this type of combined feature search is more difficult because it requires two boolean map operations in series: creating a blue map, then creating a horizontal map and intersecting it against the blue map to hunt for the target. Importantly, however, the time required for such a search is constant and independent of the number of distractors. It is simply the sum of the time required to complete the two boolean map operations.

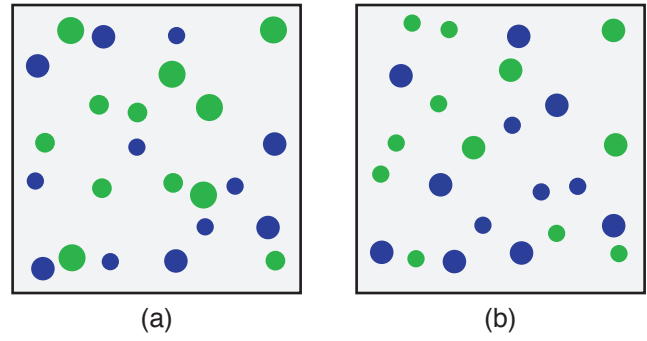


Fig. 9. Estimating average size: (a) average size of green elements is larger; (b) average size of blue elements is larger [66]

Fig. 8 shows two examples of searching for a blue horizontal target. Viewers can apply the following strategy to search for the target. First, search for blue objects, and once these are “held” in your memory, look for a horizontal object within that group. For most observers it is not difficult to determine the target is present in Fig. 8a and absent in Fig. 8b.

3.6 Ensemble Coding

All the preceding characterizations of preattentive vision have focused on how low level-visual processes can be used to guide attention in a larger scene and how a viewer’s goals interact with these processes. An equally important characteristic of low-level vision is its ability to generate a quick summary of how simple visual features are distributed across the field of view. The ability of humans to register a rapid and in-parallel summary of a scene in terms of its simple features was first reported by Ariely [66]. He demonstrated that observers could extract the average size of a large number of dots from a single glimpse of a display. Yet, when observers were tested on the same displays and asked to indicate whether a specific dot of a given size was present, they were unable to do so. This suggests that there is a preattentive mechanism that records summary statistics of visual features without retaining information about the constituent elements that generated the summary.

Other research has followed up on this remarkable ability, showing that rapid averages are also computed for the orientation of simple edges seen only in peripheral vision [67], for color [68] and for some higher-level qualities such as the emotions expressed—happy versus sad—in a group of faces [69]. Exploration of the robustness of the ability indicates the precision of the extracted mean is not compromised by large changes in the shape of the distribution within the set [68], [70].

Fig. 9 shows examples of two average size estimation trials. Viewers are asked to report which group has a larger average size: blue or green. In Fig. 9a each group contains six large and six small elements, but the green elements are all larger than their blue counterparts, resulting in a larger average size for the green group. In

Fig. 9b the large and small elements in each group are the same size, but there are more large blue elements than large green elements, producing a larger average size for the blue group. In both cases viewers responded with 75% accuracy or greater for diameter differences of only 8–12%. Ensemble encoding of visual properties may help to explain our experience of gist, the rich contextual information we are able to obtain from the briefest of glimpses at a scene.

This ability may offer important advantages in certain visualization environments. For example, given a stream of real-time data, ensemble coding would allow viewers to observe the stream at a high frame rate, yet still identify individual frames with interesting relative distributions of visual features (i.e., attribute values). Ensemble coding would also be critical for any situation where viewers want to estimate the amount of a particular data attribute in a display. These capabilities were hinted at in a paper by Healey et al. [71], but without the benefit of ensemble coding as a possible explanation.

3.7 Feature Hierarchy

One of the most important considerations for a visualization designer is deciding how to present information in a display without producing visual confusion. Consider, for example, the conjunction search shown in Figs. 1e–f. Another important type of interference results from a feature hierarchy that appears to exist in the visual system. For certain tasks one visual feature may be “more salient” than another. For example, during boundary detection Callaghan showed that the visual system favors color over shape [72]. Background variations in color slowed—but did not completely inhibit—a viewer’s ability to preattentively identify the presence of spatial patterns formed by different shapes (Fig. 10a). If color is held constant across the display, these same shape patterns are immediately visible. The interference is asymmetric: random variations in shape have no effect on a viewer’s ability to see color patterns (Fig. 10b). Luminance-on-hue and hue-on-texture preferences have also been found [23], [47], [73], [74], [75].

Feature hierarchies suggest the most important data attributes should be displayed with the most salient visual features, to avoid situations where secondary data values mask the information the viewer wants to see.

Various researchers have proposed theories for how visual features compete for attention [76], [77], [78]. They point to a rough order of processing: (1) determine the 3D layout of a scene; (2) determine surface structures and volumes; (3) establish object movement; (4) interpret luminance gradients across surfaces; and (5) use color to fine-tune these interpretations. If a conflict arises between levels, it is usually resolved in favor of giving priority to an earlier process.

4 VISUAL EXPECTATION AND MEMORY

Preattentive processing asks in part, “What visual properties draw our eyes, and therefore our focus of attention

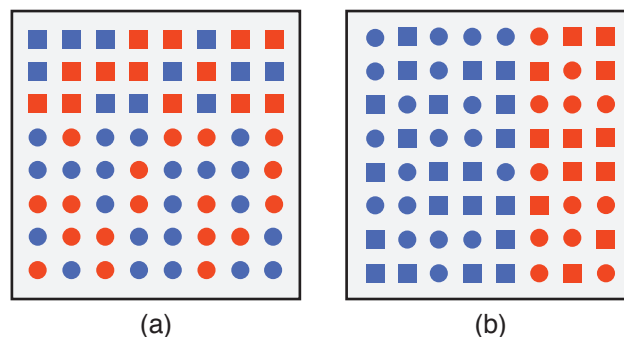


Fig. 10. Hue-on-form hierarchy: (a) horizontal form boundary is masked when hue varies randomly; (b) vertical hue boundary preattentively identified even when form varies randomly [72]

to a particular object in a scene?” An equally interesting question is, “What do we remember about an object or a scene when we stop attending to it and look at something else?” Many viewers assume that as we look around us we are constructing a high-resolution, fully detailed description of what we see. Researchers in psychophysics have known for some time that this is not true [21], [47], [61], [79], [80]. In fact, in many cases our memory for detail between glances at a scene is very limited. Evidence suggests that a viewer’s current state of mind can play a critical role in determining what is being seen at any given moment, what is not being seen, and what will be seen next.

4.1 Eye Tracking

Although the dynamic interplay between bottom-up and top-down processing was already evident in the early eye tracking research of Yarbus [4], some modern theorists have tried to predict human eye movements during scene viewing with a purely bottom-up approach. Most notably, Itti and Koch [6] developed the saliency theory of eye movements based on Treisman’s feature integration theory. Their guiding assumption was that during each fixation of a scene, several basic feature contrasts—luminance, color, orientation—are processed rapidly and in parallel across the visual field, over a range of spatial scales varying from fine to coarse. These analyses are combined into a single feature-independent “conspicuity map” that guides the deployment of attention and therefore the next saccade to a new location, similar to Wolfe’s activation map (Fig. 6). The model also includes an inhibitory mechanism—inhibition of return—to prevent repeated attention and fixation to previously viewed salient locations.

The surprising outcome of applying this model to visual inspection tasks, however, has not been to successfully predict eye movements of viewers. Rather, its benefit has come from making explicit the failure of a purely bottom-up approach to determine the movement of attention and the eyes. It has now become almost routine in the eye tracking literature to use the Itti and Koch

model as a benchmark for bottom-up saliency, against which the top-down cognitive influences on visual selection and eye tracking can be assessed (e.g., [7], [81], [82], [83], [84]). For example, in an analysis of gaze during everyday activities, fixations are made in the service of locating objects and performing manual actions on them, rather than on the basis of object distinctiveness [85]. A very readable history of the technology involved in eye tracking is given in Wade and Tatler [86].

Other theorists have tried to use the pattern of eye movements during scene viewing as a direct index of the cognitive influences on scene perception. For example, in the scanpath theory of Stark [5], [87], [88], [89], the saccades and fixations made during initial viewing become part of the lasting memory trace of a scene. Thus, according to this theory, the fixation sequences during initial viewing and then later recognition of the same scene should be similar. Much research has confirmed that there are correlations between the scanpaths of initial and subsequent viewings. Yet, at the same time, there seem to be no negative effects on scene memory when scanpaths differ between views [90].

One of the most profound demonstrations that eye gaze and perception were not one and the same was first reported by Grimes [91]. He tracked the eyes of viewers examining natural photographs in preparation for a later memory test. On some occasions he would make large changes to the photos during the brief period—20–40 msec—in which a saccade was being made from one location to another in the photo. He was shocked to find that when two people in a photo changed clothing, or even heads, during a saccade, viewers were often blind to these changes, even when they had recently fixated the location of the changed features directly.

Clearly, the eyes are not a direct window to the soul. Research on eye tracking has shown repeatedly that merely tracking the eyes of a viewer during scene perception provides no privileged access to the cognitive processes undertaken by the viewer. Researchers studying the top-down contributions to perception have therefore established methodologies in which the role of memory and expectation can be studied through more indirect methods. In the sections that follow, we present five laboratory procedures that have been developed specifically for this purpose: postattentive amnesia, memory-guided search, change blindness, inattention blindness, and attentional blink. Understanding what we are *thinking*, *remembering*, and *expecting* as we look at different parts of a visualization is critical to designing visualizations that encourage locating and retaining the information that is most important to the viewer.

4.2 Postattentive Amnesia

Wolfe conducted a study to determine whether showing viewers a scene prior to searching it would improve their ability to locate targets [80]. Intuitively, one might assume that seeing the scene in advance would help with target detection. Wolfe's results suggest this is not true.

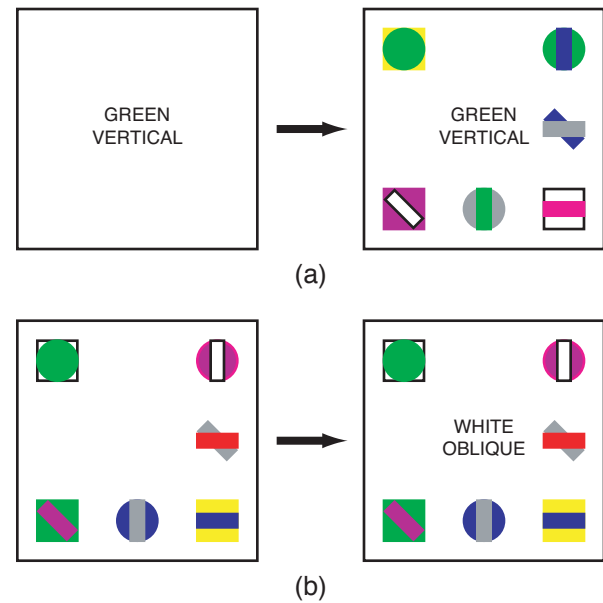


Fig. 11. Search for color-and-shape conjunction targets: (a) text identifying the target is shown, followed by the scene, green vertical target is present; (b) a preview is shown, followed by text identifying the target, white oblique target is absent [80]

Wolfe believed that if multiple objects are recognized simultaneously in the low-level visual system, it must involve a search for links between the objects and their representation in long-term memory (LTM). LTM can be queried nearly instantaneously, compared to the 40–50 msec per item needed to search a scene or to access short-term memory. Preattentive processing can rapidly draw the focus of attention to a target object, so little or no searching is required. To remove this assistance, Wolfe designed targets with two properties (Fig. 11):

- 1) Targets are formed from a conjunction of features—they cannot be detected preattentively.
- 2) Targets are arbitrary combinations of colors and shapes—they cannot be semantically recognized and remembered.

Wolfe initially tested two search types:

- 1) *Traditional search*. Text on a blank screen described the target. This was followed by a display containing 4–8 potential target formed by combinations of colors and shapes in a 3×3 array (Fig. 11a).
- 2) *Postattentive search*. The display was shown to the viewer for up to 300 msec. Text describing the target was then inserted into the scene (Fig. 11b).

Results showed that the preview provided no advantage. Postattentive search was as slow (or slower) than the traditional search, with approximately 25–40 msec per object required for target present trials. This has a significant potential impact for visualization design. In most cases visualization displays are novel, and their contents cannot be committed to LTM. If studying a display offers no assistance in searching for specific data values, then preattentive methods that draw attention to

areas of potential interest are critical for efficient data exploration.

4.3 Attention Guided by Memory and Prediction

Although research on postattentive amnesia suggests that there are few, if any, advantages from repeated viewing of a display, several more recent findings suggest there are important benefits of memory during search. Interestingly, all of these benefits seem to occur outside of the conscious awareness of the viewer.

In the area of contextual cuing [92], [93], viewers find a target more rapidly for a subset of the displays that are presented repeatedly—but in a random order—versus other displays that are presented for the first time. Moreover, when tested after the search task was completed, viewers showed no conscious recollection or awareness that some of the displays were repeated or that their search speed benefited from these repetitions. Contextual cuing appears to involve guiding attention to a target by subtle regularities in the past experience of a viewer. This means that attention can be affected by incidental knowledge about global context, in particular, the spatial relations between the target and non-target items in a given display. Visualization might be able to harness such incidental spatial knowledge of a scene by tracking both the number of views and the time spent viewing images that are later re-examined by the viewer.

A second line of research documents the unconscious tendency of viewers to look for targets in novel locations in the display, as opposed to looking at locations that have already been examined. This phenomenon is referred to as *inhibition of return* [94] and has been shown to be distinct from strategic influences on search, such as choosing consciously to search from left-to-right or moving out from the center in a clockwise direction [95].

A final area of research concerns the benefits of resuming a visual search that has been interrupted by momentarily occluding the display [96], [97]. Results show that viewers can resume an interrupted search much faster than they can start a new search. This suggests that viewers benefit from implicit (i.e., unconscious) perceptual predictions they make about the target based on the partial information acquired during the initial glimpse of a display.

Rapid resumption was first observed when viewers were asked to search for a T among L-shapes [97]. Viewers were given brief looks at the display separated by longer waits where the screen was blank. They easily found the target within a few glimpses of the display. A surprising result was the presence of many extremely fast responses after display re-presentation. Analysis revealed two different types of responses. The first, which occurred only during re-presentation, required 100–250 msec. This was followed by a second, slower set of responses that peaked at approximately 600 msec.

To test whether search was being fully interrupted, a second experiment showed two interleaved displays,

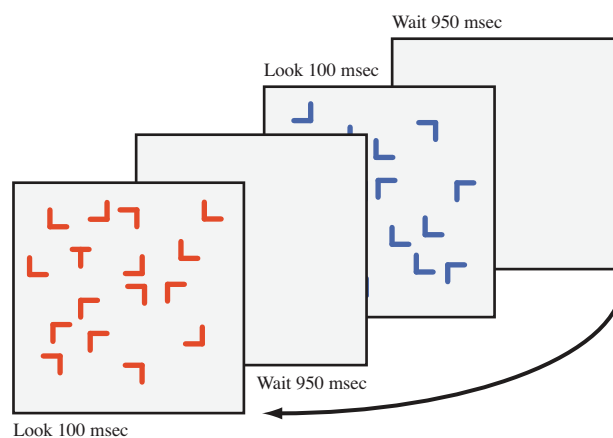


Fig. 12. A rapid responses re-display trial, viewers are asked to report the color of the T target, two separate displays must be searched [97]

one with red elements, the other with blue elements (Fig. 12). Viewers were asked to identify the color of the target T—that is, to determine whether either of the two displays contained a T. Here, viewers are forced to stop one search and initiate another as the display changes. As before, extremely fast responses were observed for displays that were re-presented.

The interpretation that the rapid responses reflected perceptual predictions—as opposed to easy access to memory of the scene—was based on two crucial findings [98], [99]. The first was the sheer speed at which a search resumed after an interruption. Previous studies on the benefits of visual priming and short term memory show responses that begin at least 500 msec after the onset of a display. Correct responses in the 100–250 msec range call for an explanation that goes beyond mere memory. The second finding was that rapid responses depended critically on a participant’s ability to form implicit perceptual predictions about what they expected to see at a particular location in the display after it returned to view.

For visualization, rapid response suggests that a viewer’s domain knowledge may produce expectations based on the current display about where certain data might appear in future displays. This in turn could improve a viewer’s ability to locate important data.

4.4 Change Blindness

Both postattentive amnesia and memory-guided search agree that our visual system does not resemble the relatively faithful and largely passive process of modern photography. A much better metaphor for vision is that of a dynamic and ongoing construction project, where the products being built are short-lived models of the external world that are specifically designed for the current visually guided tasks of the viewer [100], [101], [102], [103]. There does not appear to be any general purpose vision. What we “see” when confronted with a new

scene depends as much on our goals and expectations as it does on the light that enters our eyes.

These new findings differ from the initial ideas of preattentive processing: that only certain features are recognized without the need for focused attention, and that other features cannot be detected, even when viewers actively search for them. More recent work in preattentive vision has shown that the visual differences between a target and its neighbors, what a viewer is searching for, and how the image is presented can all have an effect on search performance. For example, Wolfe's guided search theory assumes both bottom-up (i.e., preattentive) and top-down (i.e., attention-based) activation of features in an image [48], [61], [62]. Other researchers like Treisman have also studied the dual effects of preattentive and attention-driven demands on what the visual system sees [45], [46]. Wolfe's discussion of postattentive amnesia points out that details of an image cannot be remembered across separate scenes except in areas where viewers have focused their attention [80].

New research in psychophysics has shown that an interruption in what is being seen—a blink, an eye saccade, or a blank screen—renders us “blind” to significant changes that occur in the scene during the interruption. This change blindness phenomena can be illustrated using a task similar to one shown in comic strips for many years [101], [102], [103], [104]. Fig. 13 shows three pairs of images. A significant difference exists between each image pair. Many viewers have a difficult time seeing any difference and often have to be coached to look carefully to find it. Once they discover it, they realize that the difference was not a subtle one. Change blindness is not a failure to see because of limited visual acuity; rather, it is a failure based on inappropriate attentional guidance. Some parts of the eye and the brain are clearly responding differently to the two pictures. Yet, this does not become part of our visual experience until attention is focused directly on the objects that vary.

The presence of change blindness has important implications for visualization. The images we produce are normally novel for our viewers, so existing knowledge cannot be used to guide their analyses. Instead, we strive to direct the eye, and therefore the mind, to areas of interest or importance within a visualization. This ability forms the first step towards enabling a viewer to abstract details that will persist over subsequent images.

Simons offers a wonderful overview of change blindness, together with some possible explanations [103].

- 1) **Overwriting.** The current image is overwritten by the next, so information that is not abstracted from the current image is lost. Detailed changes are only detected at the focus of attention.
- 2) **First Impression.** Only the initial view of a scene is abstracted, and if the scene is not perceived to have changed, it is not re-encoded. One example of first impression is an experiment by Levins and Simon where subjects viewed a short movie [105], [106]. During a cut scene, the central character was

switched to a completely different actor. Nearly two-thirds of the subjects failed to report that the main actor was replaced, instead describing him using details from the initial actor.

- 3) **Nothing is Stored.** No details are represented internally after a scene is abstracted. When we need specific details, we simply re-examine the scene. We are blind to change unless it affects our abstracted knowledge of the scene, or unless it occurs where we are looking.
- 4) **Everything is Stored, Nothing is Compared.** Details about a scene are stored, but cannot be accessed without an external stimulus. In one study, an experimenter asks a pedestrian for directions [103]. During this interaction, a group of students walks between the experimenter and the pedestrian, surreptitiously taking a basketball the experimenter is holding. Only a very few pedestrians reported that the basketball had gone missing, but when asked specifically about something the experimenter was holding, more than half of the remaining subjects remembered the basketball, often providing a detailed description.
- 5) **Feature Combination.** Details from an initial view and the current view are merged to form a combined representation of the scene. Viewers are not aware of which parts of their mental image come from which scene.

Interestingly, none of the explanations account for all of the change blindness effects that have been identified. This suggests that some combination of these ideas—or some completely different hypothesis—is needed to properly model the phenomena.

Simons and Rensink recently revisited the area of change blindness [107]. They summarize much of the work-to-date, and describe important research issues that are being studied using change blindness experiments. For example, evidence shows that attention is required to detect changes, although attention alone is not necessarily sufficient [108]. Changes to attended objects can also be missed, particularly when the changes are unexpected. Changes to semantically important objects are detected faster than changes elsewhere [104]. Low-level object properties of the same kind (e.g., color or size) appear to compete for recognition in visual short-term memory, but different properties seem to be encoded separately and in parallel [109]—similar in some ways to Treisman's original feature integration theory [21]. Finally, experiments suggest the locus of attention is distributed symmetrically around a viewer's fixation point [110].

Simons and Rensink also described hypotheses that they felt are not supported by existing research. For example, many people have used change blindness to suggest that our visual representation of a scene is sparse, or altogether absent. Four hypothetical models of vision were presented that include detailed representations of a scene, while still allowing for change blindness.



Fig. 13. Change blindness, a major difference exists between each pair of images; (a–b) object added/removed; (c–d) color change; (e–f) luminance change

A detailed representation could rapidly decay, making it unavailable for future comparisons; a representation could exist in a pathway that is not accessible to the comparison operation; a representation could exist and be accessible, but not be in a format that supports the comparison operation; or an appropriate representation could exist, but the comparison operation is not applied even though it could be.

4.5 Inattentional Blindness

A related phenomena called inattentional blindness suggests that viewers can completely fail to perceive visually salient objects or activities. Some of the first experiments

on this subject were conducted by Mack and Rock [101]. Viewers were shown a cross at the center of fixation and asked to report which arm was longer. After a very small number of trials (two or three) a small “critical” object was randomly presented in one of the quadrants formed by the cross. After answering which arm was longer, viewers were then asked, “Did you see anything else on the screen besides the cross?” Approximately 25% of the viewers failed to report the presence of the critical object. This was surprising, since in target detection experiments (e.g., Figs. 1a–d) the same critical objects are identified with close to 100% accuracy.

These unexpected results led Mack and Rock to mod-



Fig. 14. Images from Simons and Chabris's inattention blindness experiments, showing both superimposed and single-stream video frames containing a woman with an umbrella, and a woman in a gorilla suit [111]

ify their experiment. Following the critical trial another two or three noncritical trials were shown—again asking viewers to identify the longer arm of the cross—followed by a second critical trial and the same question, “Did you see anything else on the screen besides the cross?” Mack and Rock called these divided attention trials. The expectation is that after the initial query viewers will anticipate being asked this question again. In addition to completing the primary task, they will also search for a critical object. In the final set of displays viewers were told to ignore the cross and focus entirely on identifying whether a critical object appears in the scene. Mack and Rock called these full attention trials, since a viewer’s entire attention is directed at finding critical objects.

Results showed that viewers were significantly better at identifying critical objects in the divided attention trials, and were nearly 100% accurate during full attention trials. This confirmed that the critical objects were salient and detectable under the proper conditions.

Mack and Rock also tried placing the cross in the periphery and the critical object at the fixation point. They assumed this would improve identifying critical trials, but in fact it produced the opposite effect. Identification rates dropped to as low as 15%. This emphasizes that subjects can fail to see something, even when it is directly in their field of vision.

Mack and Rock hypothesized that “there is no perception without attention.” If you do not attend to an object in some way, you may not perceive it at all. This suggestion contradicts the belief that objects are organized into elementary units automatically and prior to attention being activated (e.g., Gestalt theory). If attention is intentional, without objects first being perceived there is nothing to focus attention on. Mack and Rock’s experiments suggest this may not be true.

More recent work by Simons and Chabris recreated a classic study by Neisser to determine whether inatten-

tional blindness can be sustained over longer durations [111]. Neisser’s experiment superimposed video streams of two basketball games [112]. Players wore white shirts in one stream and black shirts in the other. Subjects attended to one team—either white or black—and ignored the other. Whenever the subject’s team made a pass, they were told to press a key. After about 30 seconds of video, a third stream was superimposed showing a woman walking through the scene with an open umbrella. The stream was visible for about 4 seconds, after which another 25 seconds of basketball video was shown. Following the trial, only a small number of observers reported seeing the woman. When subjects only watched the screen and did not count passes, 100% noticed the woman.

Simons and Chabris controlled three conditions during their experiment. Two video styles were shown: three superimposed video streams where the actors are semi-transparent, and a single stream where the actors are filmed together. This tests to see if increased realism affects awareness. Two unexpected actors were also used: a woman with an umbrella, and a woman in a gorilla suit. This studies how actor similarity changes awareness (Fig. 14). Finally, two types of tasks were assigned to subjects: maintain one count of the bounce passes your team makes, or maintain two separate counts of the bounce passes and the aerial passes your team makes. This varies task difficulty to measure its impact on awareness.

After the video, subjects wrote down their counts, and were then asked a series of increasingly specific questions about the unexpected actor, starting with “Did you notice anything unusual?” to “Did you see a gorilla/woman carrying an umbrella?” About half of the subjects tested failed to notice the unexpected actor, demonstrating sustained inattention blindness in a dynamic scene. A single stream video, a single count task, and a woman actor all made the task easier, but in every case at least one-third of the observers were blind to the unexpected event.

4.6 Attentional Blink

In each of the previous methods for studying visual attention, the primary emphasis is on how human attention is limited in its ability to represent the details of a scene, and in its ability to represent multiple objects at the same time. But attention is also severely limited in its ability to process information that arrives in quick succession, even when that information is presented at a single location in space.

Attentional blink is currently the most widely used method to study the availability of attention across time. Its name—“blink”—derives from the finding that when two targets are presented in rapid succession, the second of the two targets cannot be detected or identified when it appears within approximately 100–500 msec following the first target [113], [114].

In a typical experiment, visual items such as words or pictures are shown in a rapid serial presentation at a single location. Raymond et al. [114] asked participants to identify the only white letter (first target) in a 10-item per second stream of black letters (distractors), then to report whether the letter “X” (second target) occurred in the subsequent letter stream. The second target was present in 50% of trials and, when shown, appeared at random intervals after the first target ranging from 100–800 msec. Reports of both targets were required after the stimulus stream ended. The attentional blink is defined as having occurred when the first target is reported correctly, but the second target is not. This usually happens for temporal lags between targets of 100–500 msec. Accuracy recovers to a normal baseline level at longer intervals.

Curiously, when the second target is presented immediately following the first target (i.e., with no delay between the two targets), reports of the second target are quite accurate [115]. This suggests that attention operates over time like a window or gate, opening in response to finding a visual item that matches its current criterion and then closing shortly thereafter to consolidate that item as a distinct object. The attentional blink is therefore an index of the “dwell time” needed to consolidate a rapidly presented visual item into visual short term memory, making it available for conscious report [116].

Change blindness, inattention blindness, and attentional blink have important consequences for visualization. Significant changes in the data may be missed if attention is fully deployed or focused on a specific location in a visualization. Attending to data elements in one frame of an animation may render us temporarily blind to what follows at that location. These issues must be considered during visualization design.

5 VISUALIZATION AND GRAPHICS

How should researchers in visualization and graphics choose between the different vision models? In psychophysics, the models do not compete with one another. Rather, they build on top of one another to address common problems and new insights over time. The models differ in terms of *why* they were developed, and in *how* they explain our eye’s response to visual stimuli. Yet, despite this diversity, the models usually agree on *which* visual features we can attend to. Given this, we recommend considering the most recent models, since these are the most comprehensive.

A related question asks how well a model fits our needs. For example, the models identify numerous visual features as preattentive, but they may not define the difference needed to produce distinguishable instances of a feature. Follow-on experiments are necessary to extend the findings for visualization design.

Finally, although vision models have proven to be surprisingly robust, their predictions can fail. Identifying these situations often leads to new research, both in

visualization and in psychophysics. For example, experiments conducted by the authors on perceiving orientation led to a visualization technique for multivalued scalar fields [19], and to a new theory on how targets are detected and localized in cognitive vision [117].

5.1 Visual Attention

Understanding visual attention is important, both in visualization and in graphics. The proper choice of visual features will draw the focus of attention to areas in a visualization that contain important data, and correctly weight the perceptual strength of a data element based on the attribute values it encodes. Tracking attention can be used to predict where a viewer will look, allowing different parts of an image to be managed based on the amount of attention they are expected to receive.

Perceptual Salience. Building a visualization often begins with a series of basic questions, “How should I represent the data? How can I highlight important data values when they appear? How can I ensure that viewers perceive differences in the data accurately?” Results from research on visual attention can be used to assign visual features to data values in ways that satisfy these needs.

A well known example of this approach is the design of colormaps to visualize continuous scalar values. The vision models agree that properties of color are preattentive. They do not, however, identify the *amount* of color difference needed to produce distinguishable colors. Follow-on studies have been conducted by visualization researchers to measure this difference. For example, Ware ran experiments that asked a viewer to distinguish individual colors and shapes formed by colors. He used his results to build a colormap that spirals up the luminance axis, providing perceptual balance and controlling simultaneous contrast error [118]. Healey conducted a visual search experiment to determine the number of colors a viewer can distinguish simultaneously. His results showed that viewers can rapidly choose between up to seven isoluminant colors [119]. Kuhn et al. used results from color perception experiments to recolor images in ways that allow colorblind viewers to properly perceive color differences [120]. Other visual features have been studied in a similar fashion, producing guidelines on the use of texture—size, orientation, and regularity [121], [122], [123]—and motion—flicker, direction, and velocity [38], [124]—for visualizing data.

An alternative method for measuring image salience is Daly’s visible differences predictor, a more physically-based approach that uses light level, spatial frequency, and signal content to define a viewer’s sensitivity at each image pixel [125]. Although Daly used his metric to compare images, it could also be applied to define perceptual salience within an visualization.

Another important issue, particularly for multivariate visualization, is feature interference. One common approach visualizes each data attribute with a separate visual feature. This raises the question, “Will the visual

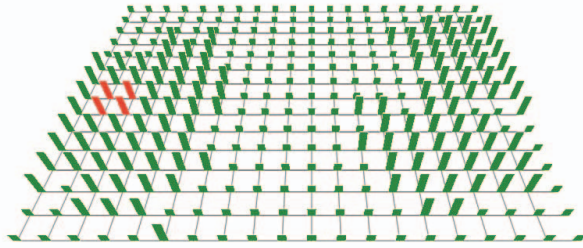


Fig. 15. A red hue target in a green background, non-target feature height varies randomly [23]

features perform as expected if they are displayed together?” Research by Callaghan showed that a hierarchy exists: perceptually strong visual features like luminance and hue can mask weaker features like curvature [73], [74]. Understanding this feature ordering is critical to ensuring that less important attributes will never “hide” data patterns the viewer is most interested in seeing.

Healey and Enns studied the combined use color and texture properties in a multidimensional visualization environment [23], [119]. A 20×15 array of paper-strip glyphs was used to test a viewer’s ability to detect different values of hue, size, density, and regularity, both in isolation, and when a secondary feature varied randomly in the background (e.g., Fig. 15, a viewer searches for a red hue target with the secondary feature height varying randomly). Differences in hue, size, and density were easy to recognize in isolation, but differences in regularity were not. Random variations in texture had no affect on detecting hue targets, but random variations in hue degraded performance for detecting size and density targets. These results suggest that feature hierarchies extend to the visualization domain.

New results on visual attention offer intriguing clues about how we might further improve a visualization. For example, recent work by Wolfe showed that we are significantly faster when we search a familiar scene [126]. One experiment involved locating a loaf of bread. If a small group of objects are shown in isolation, a viewer needs time to search for the bread object. If the bread is part of a real scene of a kitchen, however, viewers can find it immediately. In both cases the bread has the same appearance, so differences in visual features cannot be causing the difference in performance. Wolfe suggests that semantic information about a scene—our gist—guides attention in the familiar scene to locations that are most likely to contain the target. If we could define or control what “familiar” means in the context of a visualization, we might be able to use these semantics to rapidly focus a viewer’s attention on locations that are likely to contain important data.

Predicting Attention. Models of attention can be used to predict where viewers will focus their attention. In photorealistic rendering, one might ask, “How should I render a scene given a fixed rendering budget?” or “At what point does additional rendering become im-

perceptible to a viewer?” Attention models can suggest where a viewer will look and what they will perceive, allowing an algorithm to treat different parts of an image differently, for example, by rendering regions where a viewer is likely to look in higher detail, or by terminating rendering when additional effort would not be seen.

One method by Yee et al. uses a vision model to choose the amount of time to spend rendering different parts of a scene [127]. Yee constructed an error tolerance map built on the concepts of visual attention and spatiotemporal sensitivity—the reduced sensitivity of the visual system in areas of high frequency motion—measured using the bottom-up attention model of Itti and Koch [128]. The error map controls the amount of irradiance error allowed in radiosity-rendered images, producing speedups of $6\times$ versus a full global illumination solution, with little or no visible loss of detail.

Directing Attention. Rather than predicting where a viewer will look, a separate set of techniques attempt to direct a viewer’s attention. Santella and DeCarlo used nonphotorealistic rendering (NPR) to abstract photographs in ways that guide attention to target regions in an image [129]. They compared detailed and abstracted NPR images to images that preserved detail only at specific target locations. Eye tracking showed that viewers spent more time focused close to the target locations in the NPRs with limited detail, compared to the fully detailed and fully abstracted NPRs. This suggests that style changes alone do not affect how an image is viewed, but a meaningful abstraction of detail *can* direct a viewer’s attention.

Bailey et al. pursued a similar goal, but rather than varying image detail, they introduced the notion of brief, subtle modulations presented in the periphery of a viewer’s gaze to draw the focus of attention [130]. An experiment compared a control group that was shown a randomized sequence of images with no modulation to an experiment group that was shown the same sequence with modulations in luminance or warm-cool colors. When modulations were present, attention moved within one or two perceptual spans of a high-lighted region, usually in a second or less.

Directing attention is also useful in visualization. For example, a pen-and-ink sketch of a dataset could include detail in spatial areas that contain rapidly changing data values, and only a few exemplar strokes in spatial areas with nearly constant data (e.g., some combination of stippled rendering [131] and dataset simplification [132]). This would direct a viewer’s attention to high spatial frequency regions in the dataset, while abstracting in ways that still allow a viewer to recreate data values at any location in the visualization.

5.2 Visual Memory

The effects of change blindness and inattention blindness have also generated interest in the visualization and graphics communities. One approach tries to manage

these phenomena, for example, by trying to ensure viewers do not “miss” important changes in a visualization. Other approaches take advantage of the phenomena, for example, by making rendering changes during a visual interrupt or when viewers are engaged in an attention-demanding task.

Avoiding Change Blindness. Being unaware of changes due to change blindness has important consequences, particular in visualization. One important factor is the size of a display. Small screens on laptops and PDAs are less likely to mask obvious changes, since the entire display is normally within a viewer’s field of view. Rapid changes will produce motion transients that can alert viewers to the change. Larger-format displays like powerwalls or arrays of monitors make change blindness a potentially much greater problem, since viewers are encouraged to “look around.” In both displays, the changes most likely to be missed are those that do not alter the overall gist of the scene. Conversely, changes are usually easy to detect when they occur within a viewer’s focus of attention.

Predicting change blindness is inextricably linked to knowing where a viewer is likely to attend in any given display. Avoiding change blindness therefore hinges on the difficult problem of knowing what is in a viewer’s mind at any moment in time. The scope of this problem can be reduced using both top-down and bottom-up methods. A top-down approach would involve constructing a model of the viewer’s cognitive tasks. A bottom-up approach would use external influences on a viewer’s attention to guide it to regions of large or important change. Models of attention provide numerous ways to accomplish this. Another possibility is to design a visualization that combines old and new data in ways that allow viewers to separate the two. Similar to the “nothing is stored” hypothesis, viewers would not need to remember detail, since they could re-examine a visualization to reacquire it.

Nowell et al. proposed an approach for adding data to a document clustering system that attempts to avoid change blindness [133]. Topic clusters are visualized as mountains, with height defining the number of documents in a cluster, and distance defining the similarity between documents. Old topic clusters fade out over time, while new clusters appear as a white wireframe outline that fades into view and gradually fills with the cluster’s final opaque colors. Changes persist over time, and are designed to attract attention in a bottom-up manner by presenting old and new information—a surface fading out as a wireframe fades in—using unique colors and large color differences.

Harnessing Perceptual Blindness. Rather than treating perceptual blindness as a problem to avoid, some researchers have instead asked, “Can I change an image in ways that are hidden from a viewer due to perceptual blindness?” It may be possible to make significant changes that will not be noticed if viewers are looking elsewhere, if the changes do not alter the overall gist of

the scene, or if a viewer’s attention is engaged on some non-changing feature or object.

Cater et al. were interested in harnessing perceptual blindness to reduce rendering cost. One approach identified central and marginal interest objects, then ran an experiment to determine how well viewers detected detail changes across a visual interrupt [134]. As hypothesized, changes were difficult to see. Central interest changes were detected more rapidly than marginal interest changes, which required up to 40 seconds to locate.

Similar experiments studied inattention blindness by asking viewers to count the number of teapots in a rendered office scene [135]. Two iterations of the task were completed: one with a high-quality rendering of the office, and one with a low-quality rendering. Viewers who counted teapots were, in almost all cases, unaware of the difference in scene detail. These findings were used to design a progressive rendering system that combines the viewer’s task with spatiotemporal contrast sensitivity to choose where to apply rendering refinements. Computational improvements of up to 7× with little perceived loss of detail were demonstrated for a simple scene.

The underlying causes of perceptual blindness are still being studied [107]. One interesting finding is that change blindness is not universal. For example, adding and removing an object in one image can cause change blindness (Fig. 13a), but a similar add–remove difference in another image is immediately detected. It is unclear what causes one example to be hard and the other to be easy. If these mechanisms are uncovered, they may offer important guidelines on how to produce or avoid perceptual blindness.

5.3 Current Challenges

New research in psychophysics continues to provide important clues about how to visualize information. We briefly discuss some areas of current research in visualization, and results from psychophysics that may help to address them.

Visual Acuity. An important question in visualization asks, “What is the information-processing capacity of the visual system?” Various answers have been proposed, based mostly on the physical properties of the eye (e.g., see [136]). Perceptual and cognitive factors suggest that what we perceive is not the same as the amount of information the eye can register, however. For example, work by He et al. showed that the smallest region perceived by our attention is much coarser than the smallest detail the eye can resolve [137], and that only a subset of the information captured by the early sensory system is available to conscious perception. This suggests that even if we can perceive individual properties in isolation, we may not be able to attend to large sets of items presented in combination.

Related research in visualization has studied how physical resolution and visual acuity affect a viewer’s

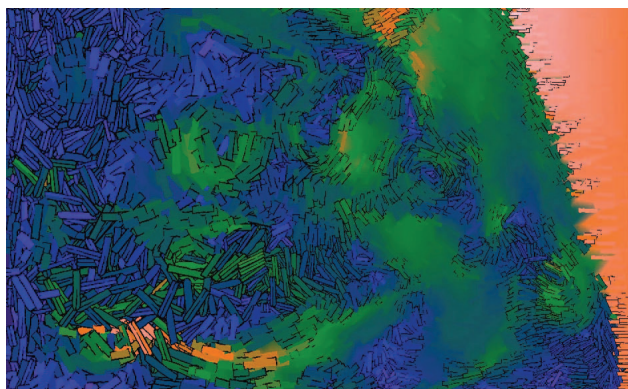


Fig. 16. A nonphotorealistic visualization of a supernova collapse rendered using indication and detail complexity [139]

ability to see different luminance, hue, size, and orientation values. These boundaries confirm He et al.'s basic findings. They also define how much information a feature can represent for a given on-screen element size—the element's physical resolution—and viewing distance—the element's visual acuity [138].

Aesthetics. Recent designs of “aesthetic” images have produced compelling results, for example, the nonphotorealistic abstractions of Santella and DeCarlo [129], or experimental results that show that viewers can extract the same information from a painterly visualization that they can from a glyph-based visualization (Fig. 16) [140].

A natural extension of these techniques is the ability to measure or control aesthetic improvement. It might be possible, for example, to vary perceived aesthetics based on a data attribute's values, to draw attention to important regions in a visualization.

Understanding the perception of aesthetics is a long-standing area of research in psychophysics. Initial work focused on the relationship between image complexity and aesthetics [141], [142], [143]. Currently, three basic approaches are used to study aesthetics. One measures the gradient of endogenous opiates during low level and high level visual processing [144]. The more the higher centers of the brain are engaged, the more pleasurable the experience. A second method equates fluent processing to pleasure, where fluency in vision derives from both external image properties and internal past experiences [145]. A final technique suggests that humans understand images by “empathizing”—embodying through inward imitation—their creation [146]. Each approach offers interesting alternatives for studying the aesthetic properties of a visualization.

Engagement. Visualizations are often designed to try to engage the viewer. Low-level visual attention occurs in two stages: orientation, which directs the focus of attention to specific locations in an image, and engagement, which encourages the visual system to linger at the location and observe visual detail. The desire to engage is based on the hypothesis that, if we orient viewers to

an important set of data in a scene, engaging them at that position may allow them to extract and remember more detail about the data.

The exact mechanisms behind engagement are currently not well understood. For example, evidence exists that certain images are spontaneously found to be more appealing, leading to longer viewing (e.g., Hayden et al. suggest that a viewer's gaze pattern follows a general set of “economic” decision-making principles [147]). Ongoing research in visualization has shown that injecting aesthetics may engage viewers (Fig. 16), although it is still not known whether this leads to a better memory for detail [139].

6 CONCLUSIONS

This paper surveys past and current theories of visual attention and visual memory. Initial work in preattentive processing identified basic visual features that capture a viewer's focus of attention. Researchers are now studying our limited ability to remember image details and to deploy attention. These phenomena have significant consequences for visualization. We strive to produce images that are salient and memorable, and that guide attention to important locations within the data. Understanding what the visual system sees, and what it does not, is critical to designing effective visual displays.

ACKNOWLEDGMENTS

The authors would like to thank Ron Rensink and Dan Simons for the use of images from their research.

REFERENCES

- [1] C. Ware, *Information Visualization: Perception for Design*, 2nd Edition. San Francisco, California: Morgan Kaufmann Publishers, Inc., 2004.
- [2] B. H. McCormick, T. A. DeFanti, and M. D. Brown, “Visualization in scientific computing,” *Computer Graphics*, vol. 21, no. 6, pp. 1–14, 1987.
- [3] J. J. Thomas and K. A. Cook, *Illuminating the Path: Research and Development Agenda for Visual Analytics*. Piscataway, New Jersey: IEEE Press, 2005.
- [4] A. Yarbus, *Eye Movements and Vision*. New York, New York: Plenum Press, 1967.
- [5] D. Norton and L. Stark, “Scan paths in saccadic eye movements while viewing and recognizing patterns,” *Vision Research*, vol. 11, pp. 929–942, 1971.
- [6] L. Itti and C. Koch, “Computational modeling of visual attention,” *Nature Reviews: Neuroscience*, vol. 2, no. 3, pp. 194–203, 2001.
- [7] E. Birmingham, W. F. Bischof, and A. Kingstone, “Saliency does not account for fixations to eyes within social scenes,” *Vision Research*, vol. 49, no. 24, pp. 2992–3000, 2009.
- [8] D. E. Broadbent, *Perception and Communication*. Oxford, United Kingdom: Oxford University Press, 1958.
- [9] H. von Helmholtz, *Handbook of Physiological Optics, Third Edition*. Dover, New York: Dover Publications, 1962.
- [10] A. Treisman, “Monitoring and storage of irrelevant messages in selective attention,” *Journal of Verbal Learning and Verbal Behavior*, vol. 3, no. 6, pp. 449–459, 1964.
- [11] J. Duncan, “The locus of interference in the perception of simultaneous stimuli,” *Psychological Review*, vol. 87, no. 3, pp. 272–300, 1980.
- [12] J. E. Hoffman, “A two-stage model of visual search,” *Perception & Psychophysics*, vol. 25, no. 4, pp. 319–327, 1979.

- [13] U. Neisser, *Cognitive Psychology*. New York, New York: Appleton-Century-Crofts, 1967.
- [14] E. J. Gibson, *Principles of Perceptual Learning and Development*. Upper Saddle River, New Jersey: Prentice-Hall, Inc., 1980.
- [15] D. Kahneman, *Attention and Effort*. Upper Saddle River, New Jersey: Prentice-Hall, Inc., 1973.
- [16] B. Julesz and J. R. Bergen, "Textons, the fundamental elements in preattentive vision and perception of textures," *The Bell System Technical Journal*, vol. 62, no. 6, pp. 1619–1645, 1983.
- [17] D. Sagi and B. Julesz, "Detection versus discrimination of visual orientation," *Perception*, vol. 14, pp. 619–628, 1985.
- [18] J. M. Wolfe, S. R. Friedman-Hill, M. I. Stewart, and K. M. O'Connell, "The role of categorization in visual search for orientation," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 18, no. 1, pp. 34–49, 1992.
- [19] C. Weigle, W. G. Emigh, G. Liu, R. M. Taylor, J. T. Enns, and C. G. Healey, "Oriented texture slivers: A technique for local value estimation of multiple scalar fields," in *Proceedings Graphics Interface 2000*, Montréal, Canada, 2000, pp. 163–170.
- [20] D. Sagi and B. Julesz, "The 'where' and 'what' in vision," *Science*, vol. 228, pp. 1217–1219, 1985.
- [21] A. Treisman and S. Gormican, "Feature analysis in early vision: Evidence from search asymmetries," *Psychological Review*, vol. 95, no. 1, pp. 15–48, 1988.
- [22] A. Treisman and G. Gelade, "A feature-integration theory of attention," *Cognitive Psychology*, vol. 12, pp. 97–136, 1980.
- [23] C. G. Healey and J. T. Enns, "Large datasets at a glance: Combining textures and colors in scientific visualization," *IEEE Transactions on Visualization and Computer Graphics*, vol. 5, no. 2, pp. 145–167, 1999.
- [24] C. G. Healey, K. S. Booth, and J. T. Enns, "Harnessing preattentive processes for multivariate data visualization," in *Proceedings Graphics Interface '93*, Toronto, Canada, 1993, pp. 107–117.
- [25] L. Trick and Z. Pylyshyn, "Why are small and large numbers enumerated differently? A limited capacity preattentive stage in vision," *Psychology Review*, vol. 101, pp. 80–102, 1994.
- [26] A. L. Nagy and R. R. Sanchez, "Critical color differences determined with a visual search task," *Journal of the Optical Society of America A*, vol. 7, no. 7, pp. 1209–1217, 10 1990.
- [27] M. D'Zmura, "Color in visual search," *Vision Research*, vol. 31, no. 6, pp. 951–966, 1991.
- [28] M. Kawai, K. Uchikawa, and H. Ujike, "Influence of color category on visual search," in *Annual Meeting of the Association for Research in Vision and Ophthalmology*, Fort Lauderdale, Florida, 1995, p. #2991.
- [29] B. Bauer, P. Jolicoeur, and W. B. Cowan, "Visual search for colour targets that are or are not linearly-separable from distractors," *Vision Research*, vol. 36, pp. 1439–1446, 1996.
- [30] J. Beck, K. Prazdny, and A. Rosenfeld, "A theory of textural segmentation," in *Human and Machine Vision*, J. Beck, K. Prazdny, and A. Rosenfeld, Eds. New York, New York: Academic Press, 1983, pp. 1–39.
- [31] J. M. Wolfe and S. L. Franzel, "Binocularity and visual search," *Perception & Psychophysics*, vol. 44, pp. 81–93, 1988.
- [32] J. T. Enns and R. A. Rensink, "Influence of scene-based properties on visual search," *Science*, vol. 247, pp. 721–723, 1990.
- [33] K. Nakayama and G. H. Silverman, "Serial and parallel processing of visual feature conjunctions," *Nature*, vol. 320, pp. 264–265, 1986.
- [34] J. W. Gebhard, G. H. Mowbray, and C. L. Byham, "Difference lumens for photic intermittence," *Quarterly Journal of Experimental Psychology*, vol. 7, pp. 49–55, 1955.
- [35] G. H. Mowbray and J. W. Gebhard, "Differential sensitivity of the eye to intermittent white light," *Science*, vol. 121, pp. 137–175, 1955.
- [36] J. L. Brown, "Flicker and intermittent stimulation," in *Vision and Visual Perception*, C. H. Graham, Ed. New York, New York: John Wiley & Sons, Inc., 1965, pp. 251–320.
- [37] B. Julesz, *Foundations of Cyclopean Perception*. Chicago, Illinois: University of Chicago Press, 1971.
- [38] D. E. Huber and C. G. Healey, "Visualizing data with motion," in *Proceedings of the 16th IEEE Visualization Conference (Vis 2005)*, Minneapolis, Minnesota, 2005, pp. 527–534.
- [39] J. Driver, P. McLeod, and Z. Dienes, "Motion coherence and conjunction search: Implications for guided search theory," *Perception & Psychophysics*, vol. 51, no. 1, pp. 79–85, 1992.
- [40] P. D. Tynan and R. Sekuler, "Motion processing in peripheral vision: Reaction time and perceived velocity," *Vision Research*, vol. 22, no. 1, pp. 61–68, 1982.
- [41] J. Hohnsbein and S. Mateeff, "The time it takes to detect changes in speed and direction of visual motion," *Vision Research*, vol. 38, no. 17, pp. 2569–2573, 1998.
- [42] J. T. Enns and R. A. Rensink, "Sensitivity to three-dimensional orientation in visual search," *Psychology Science*, vol. 1, no. 5, pp. 323–326, 1990.
- [43] Y. Ostrovsky, P. Cavanagh, and P. Sinha, "Perceiving illumination inconsistencies in scenes," *Perception*, vol. 34, no. 11, pp. 1301–1314, 2005.
- [44] M. I. Posner and S. E. Petersen, "The attention system of the human brain," *Annual Review of Neuroscience*, vol. 13, pp. 25–42, 1990.
- [45] A. Treisman, "Search, similarity, and integration of features between and within dimensions," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 17, no. 3, pp. 652–676, 1991.
- [46] A. Treisman and J. Souther, "Illusory words: The roles of attention and top-down constraints in conjoining letters to form words," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 14, pp. 107–141, 1986.
- [47] A. Treisman, "Preattentive processing in vision," *Computer Vision, Graphics and Image Processing*, vol. 31, pp. 156–177, 1985.
- [48] J. M. Wolfe, K. R. Cave, and S. L. Franzel, "Guided Search: An alternative to the feature integration model for visual search," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 15, no. 3, pp. 419–433, 1989.
- [49] B. Julesz, "A theory of preattentive texture discrimination based on first-order statistics of textons," *Biological Cybernetics*, vol. 41, pp. 131–138, 1981.
- [50] B. Julesz, "Experiments in the visual perception of texture," *Scientific American*, vol. 232, pp. 34–43, 1975.
- [51] B. Julesz, "Textons, the elements of texture perception, and their interactions," *Nature*, vol. 290, pp. 91–97, 1981.
- [52] B. Julesz, "A brief outline of the texton theory of human vision," *Trends in Neuroscience*, vol. 7, no. 2, pp. 41–45, 1984.
- [53] P. T. Quinlan and G. W. Humphreys, "Visual search for targets defined by combinations of color, shape, and size: An examination of task constraints on feature and conjunction searches," *Perception & Psychophysics*, vol. 41, no. 5, pp. 455–472, 1987.
- [54] J. Duncan, "Boundary conditions on parallel search in human vision," *Perception*, vol. 18, pp. 457–469, 1989.
- [55] J. Duncan and G. W. Humphreys, "Visual search and stimulus similarity," *Psychological Review*, vol. 96, no. 3, pp. 433–458, 1989.
- [56] H. J. Müller, G. W. Humphreys, P. T. Quinlan, and M. J. Riddoch, "Combined-feature coding in the form domain," in *Visual Search*, D. Brogan, Ed. New York, New York: Taylor & Francis, 1990, pp. 47–55.
- [57] Z. J. He and K. Nakayama, "Visual attention to surfaces in three-dimensional space," *Proceedings of the National Academy of Sciences*, vol. 92, no. 24, pp. 11 155–11 159, 1995.
- [58] K. O'Craven, P. Downing, and N. Kanwisher, "fMRI evidence for objects as the units of attentional selection," *Nature*, vol. 401, no. 6753, pp. 548–587, 1999.
- [59] M. P. Eckstein, J. P. Thomas, J. Palmer, and S. S. Shimozaki, "A signal detection model predicts the effects of set size on visual search accuracy for feature, conjunction, triple conjunction, and disjunction displays," *Perception & Psychophysics*, vol. 62, no. 3, pp. 425–451, 2000.
- [60] J. T. Townsend, "Serial versus parallel processing: Sometimes they look like Tweedledum and Tweedledee but they can (and should) be distinguished," *Psychological Science*, vol. 1, no. 1, pp. 46–54, 1990.
- [61] J. M. Wolfe, "Guided Search 2.0: A revised model of visual search," *Psychonomic Bulletin & Review*, vol. 1, no. 2, pp. 202–238, 1994.
- [62] J. M. Wolfe and K. R. Cave, "Deploying visual attention: The Guided Search model," in *AI and the Eye*, T. Troscianko and A. Blake, Eds. Chichester, United Kingdom: John Wiley & Sons, Inc., 1989, pp. 79–103.
- [63] J. M. Wolfe, K. P. Yu, M. I. Stewart, A. D. Shorter, S. R. Friedman-Hill, and K. R. Cave, "Limitations on the parallel guidance of visual search: Color \times color and orientation \times orientation conjunctions," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 16, no. 4, pp. 879–892, 1990.

- [64] L. Huang and H. Pashler, "A boolean map theory of visual attention," *Psychological Review*, vol. 114, no. 3, pp. 599–631, 2007.
- [65] L. Huang, A. Treisman, and H. Pashler, "Characterizing the limits of human visual awareness," *Science*, vol. 317, pp. 823–825, 2007.
- [66] D. Ariely, "Seeing sets: Representation by statistical properties," *Psychological Science*, vol. 12, no. 2, pp. 157–162, 2001.
- [67] L. Parkes, J. Lund, A. Angelucci, J. A. Solomon, and M. Morgan, "Compulsory averaging of crowded orientation signals in human vision," *Nature Neuroscience*, vol. 4, no. 7, pp. 739–744, 2001.
- [68] S. C. Chong and A. Treisman, "Statistical processing: Computing the average size in perceptual groups," *Vision Research*, vol. 45, no. 7, pp. 891–900, 2005.
- [69] J. Haberman and D. Whitney, "Seeing the mean: Ensemble coding for sets of faces," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 35, no. 3, pp. 718–734, 2009.
- [70] S. C. Chong and A. Treisman, "Representation of statistical properties," *Vision Research*, vol. 43, no. 4, pp. 393–404, 2003.
- [71] C. G. Healey, K. S. Booth, and J. T. Enns, "Real-time multivariate data visualization using preattentive processing," *ACM Transactions on Modeling and Computer Simulation*, vol. 5, no. 3, pp. 190–221, 1995.
- [72] T. C. Callaghan, "Interference and dominance in texture segregation," in *Visual Search*, D. Brogan, Ed. New York, New York: Taylor & Francis, 1990, pp. 81–87.
- [73] T. C. Callaghan, "Dimensional interaction of hue and brightness in preattentive field segregation," *Perception & Psychophysics*, vol. 36, no. 1, pp. 25–34, 1984.
- [74] T. C. Callaghan, "Interference and domination in texture segregation: Hue, geometric form, and line orientation," *Perception & Psychophysics*, vol. 46, no. 4, pp. 299–311, 1989.
- [75] R. J. Snowden, "Texture segregation and visual search: A comparison of the effects of random variations along irrelevant dimensions," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 24, no. 5, pp. 1354–1367, 1998.
- [76] F. A. A. Kingdom, "Lightness, brightness and transparency: A quarter century of new ideas, captivating demonstrations and unrelenting controversy," *Vision Research*, vol. 51, no. 7, pp. 652–673, 2011.
- [77] T. V. Papathomas, A. Gorea, and B. Jülész, "Two carriers for motion perception: Color and luminance," *Vision Research*, vol. 31, no. 1, pp. 1883–1891, 1991.
- [78] T. V. Papathomas, I. Kovacs, A. Gorea, and B. Jülész, "A unified approach to the perception of motion, stereo, and static flow patterns," *Behavior Research Methods, Instruments, & Computers*, vol. 27, no. 4, pp. 419–432, 1995.
- [79] J. Pomerantz and E. A. Pristach, "Emergent features, attention, and perceptual glue in visual form perception," *Journal of Experimental Psychology: Human Perception & Performance*, vol. 15, no. 4, pp. 635–649, 1989.
- [80] J. M. Wolfe, N. Klempen, and K. Dahlen, "Post attentive vision," *The Journal of Experimental Psychology: Human Perception and Performance*, vol. 26, no. 2, pp. 693–716, 2000.
- [81] E. Birmingham, W. F. Bischof, and A. Kingstone, "Get real! Resolving the debate about equivalent social stimuli," *Visual Cognition*, vol. 17, no. 6–7, pp. 904–924, 2009.
- [82] K. Rayner and M. S. Castelano, "Eye movements," *Scholarpedia*, vol. 2, no. 10, p. 3649, 2007.
- [83] J. M. Henderson and M. S. Castelano, "Eye movements and visual memory for scenes," in *Cognitive Processes in Eye Guidance*, G. Underwood, Ed. Oxford, United Kingdom: Oxford University Press, 2005, pp. 213–235.
- [84] J. M. Henderson and F. Ferreira, "Scene perception for psycholinguistics," in *The Interface of Language, Vision and Action: Eye Movements and the Visual World*, J. M. Henderson and F. Ferreira, Eds. New York, New York: Psychology Press, 2004, pp. 1–58.
- [85] M. F. Land, N. Mennie, and J. Rusted, "The roles of vision and eye movements in the control of activities of daily living," *Perception*, vol. 28, no. 11, pp. 1311–1328, 1999.
- [86] N. J. Wade and B. W. Tatler, *The Moving Tablet of the Eye: The Origins of Modern Eye Movement Research*. Oxford, United Kingdom: Oxford University Press, 2005.
- [87] D. Norton and L. Stark, "Scan paths in eye movements during pattern perception," *Science*, vol. 171, pp. 308–311, 1971.
- [88] L. Stark and S. R. Ellis, "Scanpaths revisited: Cognitive models direct active looking," in *Eye Movements: Cognition and Visual Perception*, D. E. Fisher, R. A. Monty, and J. W. Senders, Eds. Hillsdale, New Jersey: Lawrence Erlbaum and Associates, 1981, pp. 193–226.
- [89] W. H. Zangemeister, K. Sherman, and L. Stark, "Evidence for a global scanpath strategy in viewing abstract compared with realistic images," *Neuropsychologia*, vol. 33, no. 8, pp. 1009–1024, 1995.
- [90] K. Rayner, "Eye movements in reading and information processing: 20 years of research," *Psychological Bulletin*, vol. 85, pp. 618–660, 1998.
- [91] J. Grimes, "On the failure to detect changes in scenes across saccades," in *Vancouver Studies in Cognitive Science: Vol 5. Perception*, K. Akins, Ed. Oxford, United Kingdom: Oxford University Press, 1996, pp. 89–110.
- [92] M. M. Chun and Y. Jiang, "Contextual cueing: Implicit learning and memory of visual context guides spatial attention," *Cognitive Psychology*, vol. 36, no. 1, pp. 28–71, 1998.
- [93] M. M. Chun, "Contextual cueing of visual attention," *Trends in Cognitive Science*, vol. 4, no. 5, pp. 170–178, 2000.
- [94] M. I. Posner and Y. Cohen, "Components of visual orienting," in *Attention & Performance X*, H. Bouma and D. Bouwhuis, Eds. Hillsdale, New Jersey: Lawrence Erlbaum and Associates, 1984, pp. 531–556.
- [95] R. M. Klein, "Inhibition of return," *Trends in Cognitive Science*, vol. 4, no. 4, pp. 138–147, 2000.
- [96] J. T. Enns and A. Lleras, "What's next? New evidence for prediction in human vision," *Trends in Cognitive Science*, vol. 12, no. 9, pp. 327–333, 2008.
- [97] A. Lleras, R. A. Rensink, and J. T. Enns, "Rapid resumption of an interrupted search: New insights on interactions of vision and memory," *Psychological Science*, vol. 16, no. 9, pp. 684–688, 2005.
- [98] A. Lleras, R. A. Rensink, and J. T. Enns, "Consequences of display changes during interrupted visual search: Rapid resumption is target-specific," *Perception & Psychophysics*, vol. 69, pp. 980–993, 2007.
- [99] J. A. Jungé, T. F. Brady, and M. M. Chun, "The contents of perceptual hypotheses: Evidence from rapid resumption of interrupted visual search," *Attention, Perception & Psychophysics*, vol. 71, no. 4, pp. 681–689, 2009.
- [100] H. E. Egeth and S. Yantis, "Visual attention: Control, representation, and time course," *Annual Review of Psychology*, vol. 48, pp. 269–297, 1997.
- [101] A. Mack and I. Rock, *Inattentional Blindness*. Menlo Park, California: MIT Press, 2000.
- [102] R. A. Rensink, "Seeing, sensing, and scrutinizing," *Vision Research*, vol. 40, no. 10–12, pp. 1469–1487, 2000.
- [103] D. J. Simons, "Current approaches to change blindness," *Visual Cognition*, vol. 7, no. 1/2/3, pp. 1–15, 2000.
- [104] R. A. Rensink, J. K. O'Regan, and J. J. Clark, "To see or not to see: The need for attention to perceive changes in scenes," *Psychological Science*, vol. 8, pp. 368–373, 1997.
- [105] D. T. Levin and D. J. Simons, "Failure to detect changes to attended objects in motion pictures," *Psychonomic Bulletin and Review*, vol. 4, no. 4, pp. 501–506, 1997.
- [106] D. J. Simons, "In sight, out of mind: When object representations fail," *Psychological Science*, vol. 7, no. 5, pp. 301–305, 1996.
- [107] D. J. Simons and R. A. Rensink, "Change blindness: Past, present, and future," *Trends in Cognitive Science*, vol. 9, no. 1, pp. 16–20, 2005.
- [108] J. Triesch, D. H. Ballard, M. M. Hayhoe, and B. T. Sullivan, "What you see is what you need," *Journal of Vision*, vol. 3, no. 1, pp. 86–94, 2003.
- [109] M. E. Wheeler and A. E. Treisman, "Binding in short-term visual memory," *Journal of Experimental Psychology: General*, vol. 131, no. 1, pp. 48–64, 2002.
- [110] P. U. Tse, D. L. Shienberg, and N. K. Logothetis, "Attentional enhancement opposite a peripheral flash revealed using change blindness," *Psychological Science*, vol. 14, no. 2, pp. 91–99, 2003.
- [111] D. J. Simons and C. F. Chabris, "Gorillas in our midst: Sustained inattention blindness for dynamic events," *Perception*, vol. 28, no. 9, pp. 1059–1074, 1999.
- [112] U. Neisser, "The control of information pickup in selective looking," in *Perception and its Development: A Tribute to Eleanor J. Gibson*, A. D. Pick, Ed. Hillsdale, New Jersey: Lawrence Erlbaum and Associates, 1979, pp. 201–219.

- [113] D. E. Broadbent and M. H. P. Broadbent, "From detection to identification: Response to multiple targets in rapid serial visual presentation," *Perception and Psychophysics*, vol. 42, no. 4, pp. 105–113, 1987.
- [114] J. E. Raymond, K. L. Shapiro, and K. M. Arnell, "Temporary suppression of visual processing in an RSVP task: An attentional blink?" *Journal of Experimental Psychology: Human Perception & Performance*, vol. 18, no. 3, pp. 849–860, 1992.
- [115] W. F. Visser, T. A. W. Bischof and V. DiLollo, "Attentional switching in spatial and nonspatial domains: Evidence from the attentional blink," *Psychological Bulletin*, vol. 125, pp. 458–469, 1999.
- [116] J. Duncan, R. Ward, and K. Shapiro, "Direct measurement of attentional dwell time in human vision," *Nature*, vol. 369, pp. 313–315, 1994.
- [117] G. Liu, C. G. Healey, and J. T. Enns, "Target detection and localization in visual search: A dual systems perspective," *Perception & Psychophysics*, vol. 65, no. 5, pp. 678–694, 2003.
- [118] C. Ware, "Color sequences for univariate maps: Theory, experiments, and principles," *IEEE Computer Graphics & Applications*, vol. 8, no. 5, pp. 41–49, 1988.
- [119] C. G. Healey, "Choosing effective colours for data visualization," in *Proceedings of the 7th IEEE Visualization Conference (Vis '96)*, San Francisco, California, 1996, pp. 263–270.
- [120] G. R. Kuhn, M. M. Oliveira, and L. A. F. Fernandes, "An efficient naturalness-preserving image-recoloring method for dichromats," *IEEE Transactions on Visualization and Computer Graphics*, vol. 14, no. 6, pp. 1747–1754, 2008.
- [121] C. Ware and W. Knight, "Using visual texture for information display," *ACM Transactions on Graphics*, vol. 14, no. 1, pp. 3–20, 1995.
- [122] C. G. Healey and J. T. Enns, "Building perceptual textures to visualize multidimensional datasets," in *Proceedings of the 9th IEEE Visualization Conference (Vis '98)*, Research Triangle Park, North Carolina, 1998, pp. 111–118.
- [123] V. Interrante, "Harnessing natural textures for multivariate visualization," *IEEE Computer Graphics & Applications*, vol. 20, no. 6, pp. 6–11, 2000.
- [124] L. Bartram, C. Ware, and T. Calvert, "Filtering and integrating visual information with motion," *Information Visualization*, vol. 1, no. 1, pp. 66–79, 2002.
- [125] S. J. Daly, "Visible differences predictor: An algorithm for the assessment of image fidelity," in *Human Vision, Visual Processing, and Digital Display III*. San Jose, California: SPIE Vol. 1666, 1992, pp. 2–15.
- [126] J. M. Wolfe, M. L.-H. Võ, K. K. Evans, and M. R. Greene, "Visual search in scenes involves selective and nonselective pathways," *Trends in Cognitive Science*, in press, 2011.
- [127] H. Yee, S. Pattanaik, and D. P. Greenberg, "Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments," *ACM Transactions on Graphics*, vol. 20, no. 1, pp. 39–65, 2001.
- [128] L. Itti and C. Koch, "A saliency-based search mechanism for overt and covert shifts of visual attention," *Vision Research*, vol. 40, no. 10–12, pp. 1489–1506, 2000.
- [129] A. Santella and D. DeCarlo, "Visual interest in NPR: An evaluation and manifesto," in *Proceedings of the 3rd International Symposium on Non-Photorealistic Animation and Rendering (NPAR 2004)*, Annecy, France, 2004, pp. 71–78.
- [130] R. Bailey, A. McNamara, S. Nisha, and C. Grimm, "Subtle gaze direction," *ACM Transactions on Graphics*, vol. 28, no. 4, pp. 100:1–100:14, 2009.
- [131] A. Lu, C. J. Morris, D. S. Ebert, P. Rheingans, and C. Hansen, "Non-photorealistic volume rendering using stippling techniques," in *Proceedings Visualization 2002*, Boston, Massachusetts, 2002, pp. 211–218.
- [132] J. D. Walter and C. G. Healey, "Attribute preserving dataset simplification," in *Proceedings of the 12th IEEE Visualization Conference (Vis 2001)*, San Diego, California, 2001, pp. 113–120.
- [133] L. Nowell, E. Hetzler, and T. Tanasse, "Change blindness in information visualization: A case study," in *Proceedings of the 7th IEEE Symposium on Information Visualization (InfoVis 2001)*, San Diego, California, 2001, pp. 15–22.
- [134] K. Cater, A. Chalmers, and C. Dalton, "Varying rendering fidelity by exploiting human change blindness," in *Proceedings of the 1st International Conference on Computer Graphics and Interactive Techniques*, Melbourne, Australia, 2003, pp. 39–46.
- [135] K. Cater, A. Chalmers, and G. Ward, "Detail to attention: Exploiting visual tasks for selective rendering," in *Proceedings of the 14th Eurographics Workshop on Rendering*, Leuven, Belgium, 2003, pp. 270–280.
- [136] K. Koch, J. McLean, S. Ronen, M. A. Freed, M. J. Berry, V. Balasubramanian, and P. Sterling, "How much the eye tells the brain," *Current Biology*, vol. 16, pp. 1428–1434, 2006.
- [137] S. He, P. Cavanagh, and J. Intriligator, "Attentional resolution," *Trends in Cognitive Science*, vol. 1, no. 3, pp. 115–121, 1997.
- [138] A. P. Sawant and C. G. Healey, "Visualizing multidimensional query results using animation," in *Proceedings of the 2008 Conference on Visualization and Data Analysis (VDA 2008)*. San Jose, California: SPIE Vol. 6809, paper 04, 2008, pp. 1–12.
- [139] L. G. Tateosian, C. G. Healey, and J. T. Enns, "Engaging viewers through nonphotorealistic visualizations," in *Proceedings 5th International Symposium on Non-Photorealistic Animation and Rendering (NPAR 2007)*, San Diego, California, 2007, pp. 93–102.
- [140] C. G. Healey, J. T. Enns, L. G. Tateosian, and M. Remple, "Perceptually-based brush strokes for nonphotorealistic visualization," *ACM Transactions on Graphics*, vol. 23, no. 1, pp. 64–96, 2004.
- [141] G. D. Birkhoff, *Aesthetic Measure*. Cambridge, Massachusetts: Harvard University Press, 1933.
- [142] D. E. Berlyne, *Aesthetics and Psychobiology*. New York, New York: Appleton-Century-Crofts, 1971.
- [143] L. F. Barrett and J. A. Russell, "The structure of current affect: Controversies and emerging consensus," *Current Directions in Psychological Science*, vol. 8, no. 1, pp. 10–14, 1999.
- [144] I. Biederman and E. A. Vessel, "Perceptual pleasure and the brain," *American Scientist*, vol. 94, no. 3, pp. 247–253, 2006.
- [145] P. Winkielman, J. Halberstadt, T. Fazendeiro, and S. Catty, "Prototypes are attractive because they are easy on the brain," *Psychological Science*, vol. 17, no. 9, pp. 799–806, 2006.
- [146] D. Freedberg and V. Gallese, "Motion, emotion and empathy in esthetic experience," *Trends in Cognitive Science*, vol. 11, no. 5, pp. 197–203, 2007.
- [147] B. Y. Hayden, P. C. Parikh, R. O. Deaner, and M. L. Platt, "Economic principles motivating social attention in humans," *Proceedings of the Royal Society B: Biological Sciences*, vol. 274, no. 1619, pp. 1751–1756, 2007.



aesthetics related to visual analysis and data management.



Princeton University and he is a Fellow of the Royal Society of Canada.

Christopher G. Healey is an Associate Professor in the Department of Computer Science at North Carolina State University. He received a B.Math from the University of Waterloo in Waterloo, Canada, and an M.Sc. and Ph.D. from the University of British Columbia in Vancouver, Canada. He is an Associate Editor for *ACM Transactions on Applied Perception*. His research interests include visualization, graphics, visual perception, and areas of applied mathematics, databases, artificial intelligence, and

James T. Enns is a Distinguished Professor at the University of British Columbia in the Department of Psychology. A central theme of his research is the role of attention in human vision. He has served as Associate Editor for *Psychological Science* and *Consciousness & Cognition*. His research has been supported by grants from NSERC, CFI, BC Health & Nissan. He has authored textbooks on perception, edited volumes on cognitive development, and published numerous scientific articles. His Ph.D. is from