CHAPTER FIVE

Visual Salience and Finding Information

Suppose there is a crisis at a large bank because an employee, George, has lost billions on risky stock trades. We need to identify the vice president who is responsible for George's activities, and we have at our disposal an organization chart showing the management hierarchy of the company. This problem can be solved through a straightforward visual thinking process. First, conduct a visual search for the box representing George, then visually trace upward, following the chain of lines and boxes up to the level of vice president.

Another example: Suppose we are looking at the floor plan of a museum building and we wish to find a coffee shop. We locate the symbol for coffee shop on the key at the side of the floor plan, and then we carry out a visual search to find that symbol on the plan. A second more complex visual thinking process will be needed to find a route from where we are currently to the location of our coffee.

In both of these examples, a core activity can be described in terms of a two-step process:

> **Step 1.** A visual query is formulated in the mind of the person, relating to the problem to be solved.
>
> **Step 2**. A visual search of the display is carried out to find patterns that resolve the query.

The visual query can have many different forms, but it always involves reformulating part of the problem so that the solution can be found through a visual pattern search.

Information Visualization. DOI: 10.1016/B978-0-12-381464-7.00005-3

The visual pattern to be found can range from a symbol of a particular shape or color to an arbitrary complex or subtle visual pattern. In all cases, understanding what makes a pattern easy to find is critical in determining how efficiently the query will be executed, and what makes for efficient search is the central theme of this and the next chapters. In explaining this we will be putting flesh on the bare bones of the first two guidelines of this book set out in Chapter 1 and restated here to save the need to look back. [G1.1] *Design graphic representations of data by taking into account human sensory capabilities in such a way that important data elements and data patterns can be quickly perceived.* [G1.2] *Important data should be represented by graphical elements that are more visually distinct than those representing less important information.*

In understanding how visual queries are resolved we gain a deeper understanding of how best to design two of the most common kinds of things used in data visualization—namely, graphical *symbols* and *glyphs*. A graphical *symbol* is a *graphical object that represents an entity*. An example is the coffee shop symbol on the map. If this were to look like a coffee cup it would be an *iconic* symbol. Other examples are the noniconic triangles and squares used to represent data points in statistical graphs. A well-designed symbol set is one where each of the symbols can be readily found, each symbol is distinct from the others, and each symbol is compact.

Whereas symbols have a purely nominal function, *glyphs* also represent quantitative values. A *glyph* is a *graphical object designed to represent some entity and convey one or numerical attributes of that entity*. For information about stocks on the stock exchange, the color of a glyph can be used to show the price-to-earnings ratio, the size of the glyph can display the growth trend, and the shape of the glyph can represent the type of company—square for technology stocks, round for resources, and so on. A well-designed glyph is one that, in addition to being easily found, supports rapid and accurate resolution of visual queries regarding the ordinal, interval, or ratio quantities that are expressed.

Visual search is one of the basic things the visual system is designed for, and it involves the entire visual system. A large part of search is the way the eyes are moved around the scene to pick up information, but as we shall see it also involves the retuning of every visual part of the brain to meet the needs of the query task. There is a kind of mental inner scan, within a fixation, where a few visual patterns are tested for query-resolving properties. We will start with some basic facts about eye movements and then go on to discuss the factors that make something a target of an eye movement, before returning to the overall process.

Eye Movements

Moving our eyes causes different parts of the visual environment to be imaged on the high-resolution fovea where we can see detail. Eye movements are frequent. For example, as you read this page, your eye is making between two and five jerky movements, called *saccades*, per second, and each of these movements can be thought of as a basic act of visual search.

There are three important types of eye movements:

1. **Saccadic movements.** In a visual search task, the eye moves rapidly from fixation to fixation. The dwell period is generally between 200 and 400 msec; the saccade takes between 20 and 180 msec and depends on the angle moved. For eye movements of more than 20 degrees, head movements follow, and this can take half a second or more (Hallett, 1986; Barfield et al., 1995; Rayner, 1998). A typical length of a saccade for someone scanning a scene is about 5 degrees of visual angle. A typical length of a saccade when reading is 2 degrees (Land & Tatler, 2009). The typical length of the saccade that people make when using visualizations depends on the design and the size of the display, but we can expect it to be in the range of 2 to 5 degrees for a well-designed display. As a general principle, visual search will be considerably more efficient for more compact displays because eye movements will be shorter and faster.

[G5.1] To minimize the cost of visual searches, make visualization displays as compact as possible, compatible with visual clarity. For efficiency, information nodes should be arranged so that the average saccade is 5 degrees or less.

2. **Smooth-pursuit movements.** When an object is moving smoothly in the visual field, the eye has the ability to lock onto it and track it. This is called a *smooth-pursuit* eye movement. This ability also enables us to make head and body movements while maintaining fixation on an object of interest.

3. **Convergent movements (also called vergence movements).** When an object moves toward us, our eyes converge. When it moves away, they diverge. Convergent movements can be either saccadic or smooth.

Saccadic eye movements are said to be *ballistic*. This means that once the brain decides to switch attention and make an eye movement, the muscle signals for accelerating and decelerating the eye are preprogrammed; the movement cannot be adjusted in mid-saccade. During a saccadic eye movement, we are less sensitive to visual input. This is called *saccadic suppression* (Riggs et al., 1974). The implication is that certain kinds of events can easily be missed if they occur while we happen to be moving our eyes. This is important when we consider the problem of alerting a computer operator to an event.

Another implication of saccadic suppression is that the brain is usually processing a rapid sequence of discrete images. Our capacity to do this is increasingly being exploited in television advertising, in which more than one cut per second of video has become commonplace. More generally, the staccato nature of seeing means that what we can see *at a single glance* is tremendously important.

Accommodation

For completeness, a different kind of adjustment also requires mention before we move on. When the eye moves to a new target at a different distance from the observer, it must refocus, or accommodate, so that the target is clearly imaged on the retina. An accommodation response typically takes about 200 msec. As we age, however, the ability to accommodate declines and refocusing the eyes must be accomplished by changing eyeglasses or, for users of bifocals or progressive lenses, by moving the head so that a different lens is between the pupil and the object being fixated. Another solution is to use laser surgery to make one eye have a near focus and the other a far focus. In this case, change of focus is accomplished by switching attention from one eye's input to the other. This is a skill that must be learned.

The Eye Movement Control Loop

Seeing can be thought of as a never-ending series of cognitive acts, each of which has the same structure: make an eye movement, pick up some information, interpret that information, and plan the next eye movement. Sometimes the planning occurs in parallel with the interpretation. Figure 5.1 summarizes the major components of this process. As a first step, search queries are constructed to help with whatever task is at hand, and these typically consist of the cognitive construction of a simple pattern to be found. The next step is a visual search for that pattern.

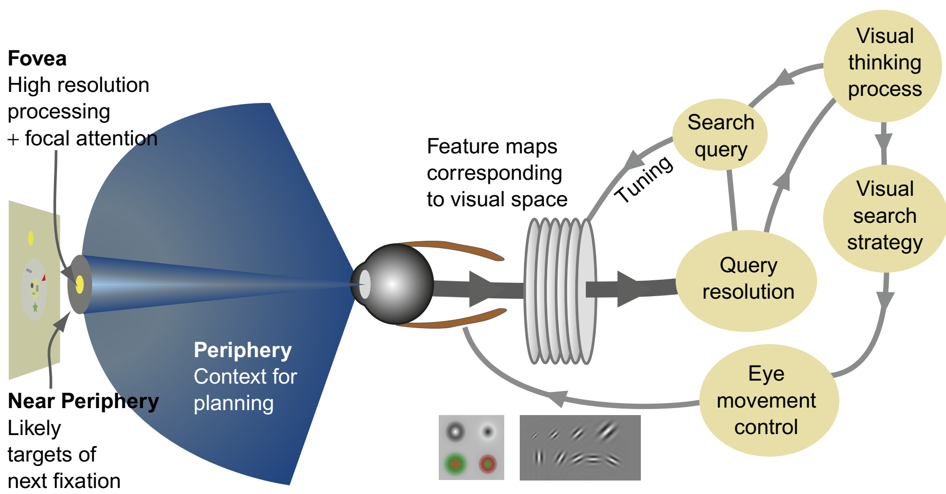


Figure 5.1 The visual search process.

But, how can the brain prepare for an eye movement without already knowing what is at the target location? How do we know where to look next? According to the theory of Wolfe and Gancarz (1996), a heuristic strategy is employed. First, a set of feature maps of the entire visual field is produced in a parallel processing operation mostly done in the V1 area of the primary visual cortex. Each feature map is devoted to a particular kind of feature, for example, vertical contours, blobs of a particular size, or a particular color. Each map is weighted according to the current task. If we are scanning a crowd to look for someone we know who has a yellow raincoat, the feature maps will emphasize yellow blobs. Next, eye movements are executed in sequence, visiting the strongest possible target area (as defined by the feature maps) first and proceeding to the next strongest. Eye movements are also weighted according to the distance from the current focus along with other strategic factors.

Three things determine what is easily findable:

1. ***A priori* salience.** Some patterns excite more neural activity in the feature maps than others.

2. **Top-down salience modification.** Depending on what we are looking for, top-down mechanisms retune the feature maps to increase their sensitivity to certain features; for example, we may wish to find a mostly vertical elongated symbol. The vertical orientation feature map will gain enhanced sensitivity.

3. **Scene gist.** This has less to do with feature maps and more to do with experience. It is something that is discussed in Chapter 11 in the context of eye movement control strategies. The important point for now is that the brain very rapidly recognizes the type of scene that is being viewed (store interior, open landscape, city street), allowing it to activate visual search strategies appropriate to a visual scene (Oliva et al., 2003). If a type of visualization is well known, then the eye movement strategies will be automatically primed for activation. This is part of the skill we develop in repeatedly using a particular style of visualization.

The perceptual mechanisms relating to V1 and V2 are the subject of this chapter, together with the lessons we can learn from them. The lessons of scene gist and the overall strategy can be found in later chapters when we discuss the skills of visual thinking.

V1, Channels, and Tuned Receptors

After preliminary processing in the retina of the eye, visual information passes up the optic nerve through a neural junction at the lateral geniculate nucleus (LGN) and through several stages of processing in the cortex. The first areas in the cortex to receive visual inputs are called, simply, *visual area 1* (V1) and *visual area 2* (V2). Most of the output from area 1 goes on to area 2, and together these two regions make up more than 40% of vision processing (Lennie, 1998). There is plenty of neural processing power, as several billion neurons in V1 and V2 are devoted to analyzing the signals from only 2 million nerve fibers coming from the optic nerves of two eyes. This makes possible the massively parallel simultaneous processing of the entire visual field for incoming signals for color, motion, texture, and the elements of form. It is here that the elementary vocabularies of both vision and data display are defined.

By the time it gets to the LGN, the signal has already been decomposed by the concentric receptive fields discussed in the previous chapter that convert the signal into red–green, yellow–blue, and dark–light differences. These signals are then passed on to V1 where slightly more complex patterns are processed.

Figure 5.2 is derived from Livingston and Hubel's diagram (1988) that summarizes both the neural architecture and the features processed in V1 and V2. A key concept in understanding this diagram is the tuned receptive field. In Chapter 3, we saw how single cell recordings of cells in the retina and the LGN reveal cells with distinctive concentric receptive fields. Such cells are said to be tuned to a particular pattern of a white spot surrounded by black or a black spot surrounded by white. In general, a tuned filter is a device that responds strongly to a certain kind of pattern and responds much less, or not at all, to other patterns. In the primary visual cortex, some cells respond only to elongated blobs with a particular position and orientation, others respond most strongly to blobs of a particular position moving in a particular direction at a particular velocity, and still others respond selectively to color.

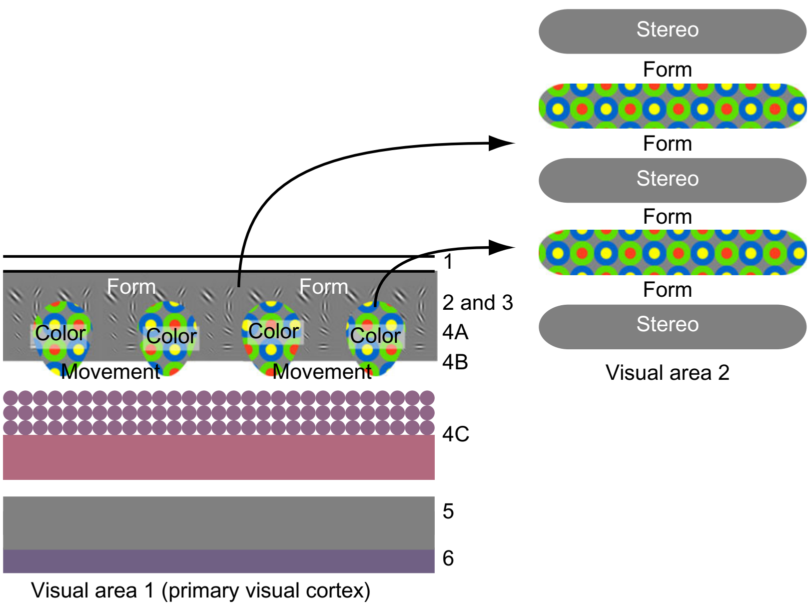


Figure 5.2 Architecture of the primary visual cortex. (*Redrawn from Livingston & Hubel (1988)*.)

There are cells in V1 and V2 that are differentially tuned to each of the following properties:

- Orientation and size (with luminance) via the Gabor processor described later in this chapter
- Color (two types of signals) via the opponent processing channel mechanisms discussed in Chapter 4
- Elements of local stereoscopic depth
- Elements of local motion

In V1 and V2 and many other regions of the brain, neurons are arranged in the form of a spatial map of the retina, meaning that adjacency relationships are preserved. A feature that is close to another feature in the image on the retina is processed by nearby neurons in V1. These maps are highly distorted, however, because the fovea is given far more space in the cortex than regions in the periphery of vision. In cortical regions devoted to the fovea, receptive fields are much smaller. It is a system in which, for each point in visual space, neurons are tuned for many different orientations, many different kinds of color information, many different directions and velocities of motion, and many different stereoscopic depths.

Notice that here we have been talking about V1 as containing a single map of the visual field, but in fact it contains a set of semi-independent feature maps, all spatially co-registered.

The Elements of Form

It is useful to think of the things that are extracted by early stage visual processing as the elements of form and pattern perception. Phonemes are the smallest elements in speech recognition, the components from which meaningful words are made. In a similar way, we can think of orientation detectors, color detectors, and so on as the elements from which meaningful perceptual objects are constructed.

An important point that can be derived from Figure 5.3 is that color and the elements of form (orientation and size) are processed separately and therefore are easy to visually separate. It is also the case that moving patterns are visually separate from static patterns. These different properties are said to have different *channels*, meaning that information expressed in one channel, the color of a symbol, does not interfere with information expressed in another, the orientation of a symbol. There are three basic high-level channels that match the areas shown in Figure 5.2—namely, color, form, and motion. We can use this fact to establish a basic principle of display design.

> **[G5.2]** Use different visual channels to display aspects of data so that they are visually distinct.

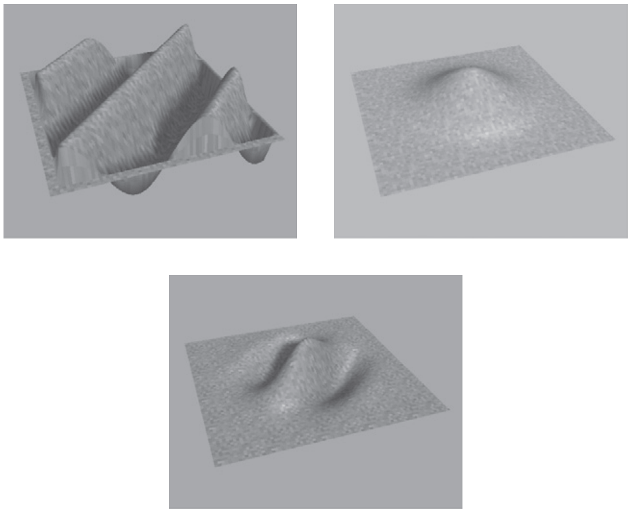


Figure 5.3 Gabor model of a V1 receptive field. Multiply the cosine wave grating on the upper left figure by the Gaussian envelope in the upper right figure to get the two-dimensional Gabor function shown on the bottom figure. The result is an excitatory center flanked by two inhibitory bars.

Once we understand the kinds of patterns the tuned cells of the visual cortex respond to best, we can apply this information to create efficient visual patterns. Patterns that can be separated based on the receptive field properties of V1 neurons should be rapidly found and easily distinguished. A number of assumptions are implicit in this account. They are worth critical examination.

One basic assumption is that the rate at which single neurons fire is the key coding variable in terms of human perception. This assumption can certainly be questioned (Montemurro et al., 2008). It may be that what is important is the way in which groups of neurons fire, or perhaps the temporal spacing or synchronization of cell firings. In fact, there is evidence that these alternative information codings are important, perhaps critical. Nevertheless, few doubt that neurons that are highly sensitive to color differences (in terms of their firing rates) are directly involved in the processing of color and that the same thing is true for motion and shape. Moreover, as we shall see, the behavior of neurons fits well with studies of how people perceive certain kinds of patterns.

Early-stage neurons are particularly important in determining how distinct things appear. We know that at higher levels of processing in the visual cortex, there are receptive fields that are much more complex; they respond to patterns that are composites of the simple receptive field patterns found at earlier stages. These more complex composite patterns, analyzed further up the visual processing chain, are not, in general, processed as rapidly.

The Gabor Model and Visual Distinctness

A number of electrophysiological and psychophysical experiments show that V1 and V2 contain large arrays of neurons that filter for orientation and size information at each point in the visual field. These neurons have both a preferred orientation and a preferred size (they are said to have spatial and orientation tuning). They are either weakly color coded or not color coded, responding to luminance patterns only.

A simple mathematical model used widely to describe the receptive field properties of these neurons is the Gabor function (Barlow, 1972; Daugman, 1984). The Gabor function has two components as illustrated in Figure 5.3: a cosine wave and a Gaussian envelope. Multiply them together, and the result is a function that responds strongly to bars and edges of a particular orientation and not at all to edges of a bar or edges at right angles to that orientation. Roughly, this can be thought of as a kind of fuzzy bar detector. It has a clear orientation, and it has an excitatory center, flanked by inhibitory bars. The opposite kind of neuron also exists, with an inhibitory center and an excitatory surround, as well as other variants.

Mathematically, a Gabor function has the following form (simplified for ease of explanation):

$$R = C\cos\left(\frac{Ox}{S}\right)\exp\left(-\frac{x^2+y^2}{S}\right) \tag{5.1}$$

The C parameter gives the amplitude or contrast value, S gives the overall size of the Gabor function by adjusting both the wavelength of the cosine grating and the rate of decay of the Gaussian envelope, and O is a rotation matrix that orients the cosine wave. Other parameters can be added to position the function at a particular location in space and adjust the ratio of the Gaussian size to the sine wavelength; however, orientation, size, and contrast are most significant in modeling human visual processing.

In an influential paper, Barlow (1972) developed a set of principles that have become influential in guiding our understanding of human perception. The second of these, called the "second dogma," provides an interesting theoretical background to the Gabor model. In the second dogma, Barlow asserted that the visual system is simultaneously optimized in both the spatial–location and spatial–frequency domains. Gabor detectors optimally preserve a combination of spatial information (the location of the information in visual space) and oriented-frequency information. A single Gabor detector can be thought of as being tuned to a little packet of orientation and size information that can be positioned anywhere in space. John Daugman (1984) showed mathematically that Gabor detectors satisfy the requirements of the Barlow dogma.

Many things about low-level perception can be explained by this model. Gabor-type detectors are used in theories of the detection of contours at the boundaries of objects (form perception), the detection of regions that have different visual textures, stereoscopic vision, and motion perception. The Gabor-type detector yields a basic set of properties out of which all more complex patterns are built. This stage of visual processing also determines some of the basic rules that make patterns distinctive at all subsequent levels of processing.

One thing that Gabor functions do is process parts of the image in terms of different spatial frequencies (see Chapter 2), and this has led to the concept of *spatial frequency channels*. These are subchannels of the form channel that encodes texture and the elements of shape. The halfwidth of the spatial tuning curve is approximately a period change (in the sinusoid) of a factor of 3, and the total number of spatial frequency channels is about 4. Wilson and Bergen (1979) determined these values using a masking technique, which essentially determines the extent to which one type of information interferes with another. The resulting estimation of spatial frequency channels is illustrated in Figure 5.4. The idea of spatial frequency channels, however, is different from the concept of separate channels made up of color, shape and texture or motion. It is more meaningful to think of spatial frequency channels as subchannels of the broader shape channel.

A single Gabor-type neuron is also broadly tuned with respect to orientation. Orientation tuning-in appears to be about 30 degrees (Blake & Holopigan, 1985); therefore, objects that differ from one another by more than 30 degrees in orientation will be more easily distinguished. Orientation can also be considered as a subchannel of form.

Probably none of the perceptual channels we shall discuss is fully independent; nevertheless, it is certainly the case that some kinds of information are processed in ways that are more independent than others. A channel that is independent from another is said to be orthogonal to it. Here, the concept is applied to the spatial information carried by Gabor detectors.

Because all information passes through spatial frequency channels, it is important to keep different kinds of information as separate as possible in terms of their frequency components and orientations.

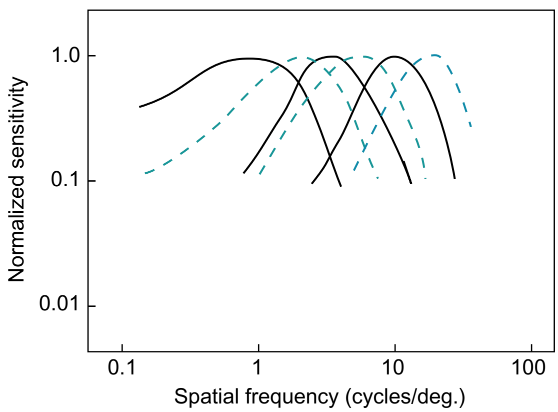


Figure 5.4 Wilson and Bergen (1979) spatial channels.

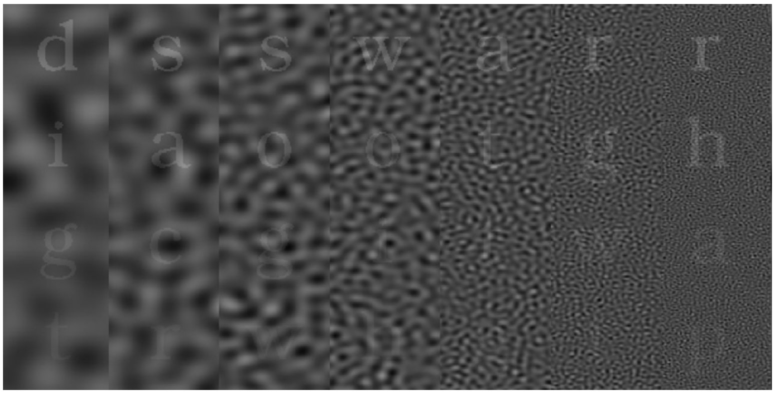


Figure 5.5 The letters are harder to see where they lie on top of visual noise that has spatial frequency components similar to the letters. (*From Solomon & Pelli (1994). Reproduced with permission.*)

Figure 5.5 shows the letters of the alphabet on top of a random visual noise pattern that has a range of spatial frequencies from low to high (Solomon & Pelli, 1994). As can be seen, the letters are difficult to perceive where the background has spatial frequency components similar to the letters. This is an example of visual interference between spatial frequency subchannels.

[G5.3] To make symbols easy to find, make them distinct from their background and from other symbols; for example, the primary spatial frequency of a symbol should be different from the spatial frequency of the background texture and from other symbols.

A Differencing Mechanism for Fine Discrimination

The very broad spatial and orientation tuning of Gabor-type detectors implies that we should not be able to discriminate small-sized orientation differences, yet this is clearly not the case. When people get enough time, they can resolve far smaller differences than they can with brief exposures. Given time, the resolvable size difference for a Gabor pattern is a size change of about 9% (Caelli & Bevan, 1983). The resolvable orientation difference is about 5 degrees (Caelli & Bevan, 1983). These resolutions are much smaller than the channel-tuning functions would predict. Neural differencing mechanisms can account for the higher resolution. The explanation for finer discriminations is *differencing mechanisms*, higher-level processes that sharpen up the output from individual receptors. The mechanism is based on inhibition. If a neuron has an excitatory input from one neuron and an inhibitory input from another with a slightly different tuning, the resulting difference signal is much more sensitive to spatial tuning than either of the original signals. This kind of sharpening is common

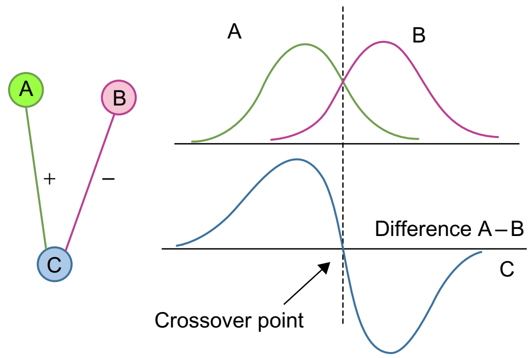


Figure 5.6 Differences between signals from neurons A and B are created by an excitatory and an inhibitory connection to neuron C.

in neural systems; it appears in color systems, edge detection, and size comparisons. Figure 5.6 illustrates the concept. Neurons A and B both have rather broadly tuned and somewhat overlapping response functions to some input pattern. Neuron C has an excitatory input from A and an inhibitory input from B. The result is that C is highly sensitive to differences between A and B at the crossover point.

The differencing mechanism explains why the visual system is exquisitely sensitive to differences, but not to absolute values. It also explains contrast effects because if one of the signals is rendered less sensitive, through lateral inhibition, the crossover point moves, but such fine discriminations are processed more slowly than the basic low-level responses. So, for rapid target finding, it is important that targets be distinct in orientation by 30 degrees or more and in size by a factor of two.

Feature Maps, Channels, and Lessons for Visual Search

To summarize to this point, because different kinds of visual properties are processed separately they can be thought of as forming separate feature maps, roughly at the V1 level. These maps cover the entire visual field, and there are many of them, each based on a different kind of feature. There is a map for redness, a map for greenness, a map for vertical orientation, a map for horizontal orientation, a map for motion, and so on.

When we are looking for something, a target set of feature properties is defined made up of the kinds of features that are found in feature maps (Eckstein et al., 2007). Eye movements are directed to feature map regions that best match the target properties. Figure 5.7 illustrates the idea. On the left is a set of symbols. On the right is how this image appears in a few of the feature maps. A search for red objects yields three candidate targets, and a search for black objects yields three different targets. A search for a left-slanted shape yields two strong and two weak targets. The oblique edges of the triangular symbols produce the weak signals, and these will somewhat distract in a search for the left-oriented bars.

Based on what we have learned so far, we can derive a number of lessons that can be applied to symbol set design. Low-level feature properties are critical.

[G5.4] Make symbols as distinct from each other as possible, in terms of both their spatial frequency components and their orientation components.

[G5.5] Make symbols as distinct as possible from background patterns in terms of both their spatial frequency components and their orientation components.

Figure 5.8 illustrates guideline G5.5 with a number of examples of scatterplots. The ones on the left use symbol shapes that are typical in many plotting packages.

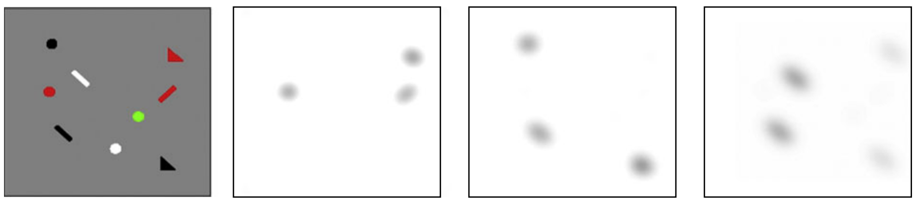


Figure 5.7 The symbols shown on the left are processed via a set of feature maps and the result directs eye movements.

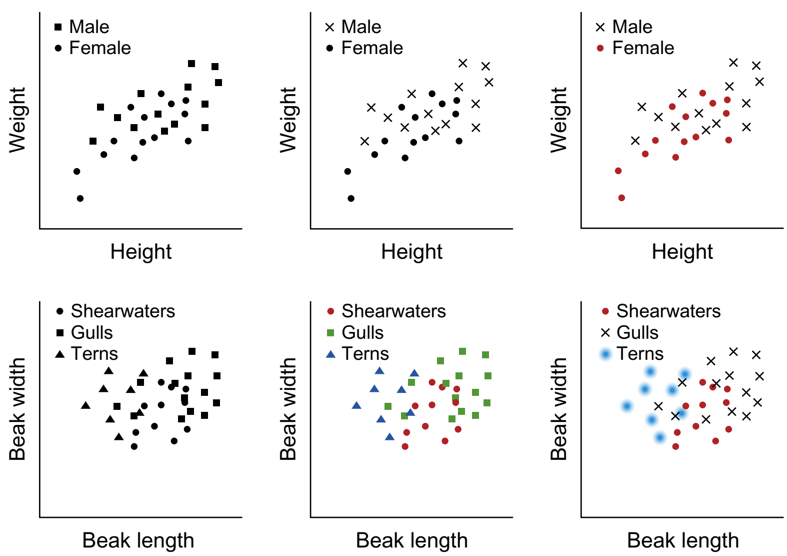


Figure 5.8 Feature channels can be used to make symbols more distinct from one another. The graphs on the right use redundant color coding in addition to more distinctive shapes.

The squares and circles are not very distinct because the differences are encoded in high spatial frequencies (see Figure 2.25 in Chapter 2, which shows how spatial sensitivity declines with high spatial frequencies). If the symbols were made larger they would be more distinct. The other examples in the center and the right have much more distinctive spatial subchannel components. Some use both color and form to increase separation in feature space.

Preattentive Processing and Ease of Search

Neuropsychology can only tell us so much about what makes shapes distinctive, because although the field is advancing rapidly the level of effort required for each discovery is huge. Inevitably, neuropsychological theory lags behind results from direct experiments using psychophysical methods with human observers. Psychophysics is the study of human responses to physically defined stimuli. There have been many experiments in which human observers are asked if a particular shape appears in a pattern of other shapes that are flashed briefly in front of their eyes. These studies have led to the concept of *preattentive processing* that is central to how we understand visual distinctiveness (Treisman, 1985).

Preattentive processing is best introduced with an example. To count the 3s in the table of digits shown in Figure 5.9(a), it is necessary to scan all the numbers sequentially. To count the 3s in Figure 5.9(b), it is necessary only to scan the red digits. This is because color is preattentively processed. Certain simple shapes or colors seem to pop out from their surroundings. The theoretical mechanism underlying popout was called *preattentive processing* because early researchers thought that it must occur prior to conscious attention, although a more modern view is that attention is integral, and we shall return to this point. In essence, preattentive processing determines what visual objects are offered up to our attention and easy to find in the next fixation

45929078059772098775972655665110049836645
27107462144654207079014738109743897010971
43907097349266847858715819048630901889074
25747072354745666142018774072849875310665

(a)

4592907805977209877597265566511004983**3**6645
2710746214465420707901473**3**8109743897010971
4**3**907097**3**49266847858715819048630901889074
2574707**3**547456661420187740728498753**3**10665

(b)

Figure 5.9 Preattentive processing. (a) To count the 3s in this table of digits, it is necessary to scan the numbers sequentially. (b) To count the 3s in this table, it is only necessary to scan the red 3s because they pop out from their surroundings.

(Findlay & Gilchrist, 2005), so prior attention is part of the phenomenon. Still, although the term is misleading, we shall continue to use it because of its widespread adoption. In any case, the phenomena described by the term are very real and of critical importance.

A typical experiment conducted to find out whether some pattern is preattentively distinct involves measuring the response time to find a target among a set of other symbols called *distractors*—for example, finding the 3s in a set of other numbers. If processing is preattentive, the time taken to find the target should be equally fast no matter how many distracting nontargets there are. So, if time to find the target is plotted against number of distractors, the result should be a horizontal line. Figure 5.10 illustrates a typical pattern of results. The circles illustrate data from a visual target that is preattentively different from the distractors. The time taken to detect whether there is a red digit in the array of digits shown in Figure 5.9 is independent of the number of black digits. The Xs in Figure 5.10 show the results from processing a feature that is *not* preattentively distinct. In this case, time to respond *increases* with number of distractors suggesting sequential processing. The results of this kind of experiment are not always as perfectly clear cut as Figure 5.10 would suggest. Sometimes there is a small, but still measurable, slope in the case of a feature that is thought to be preattentive. As a rule of thumb, anything that is processed at a rate faster than 10 msec per item is considered to be preattentive. Typical processing rates for nonpreattentive targets are 40 msec per item and more (Treisman & Gormican, 1988).

Why is this important? In displaying information, it is often useful to be able to show things "at a glance." If you want people to be able to instantaneously identify some

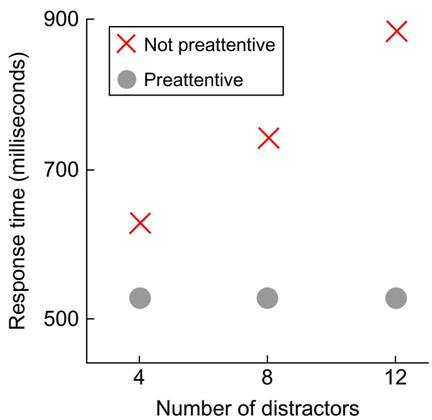


Figure 5.10 Typical results from a pattern of preattentive processing. The circles show time to perceive an object that is preattentively distinct from its surroundings. In this case, time to process is independent of the number of irrelevant objects (distractors). The Xs show how time to process nonpreattentively distinct targets increases with the number of distractors.

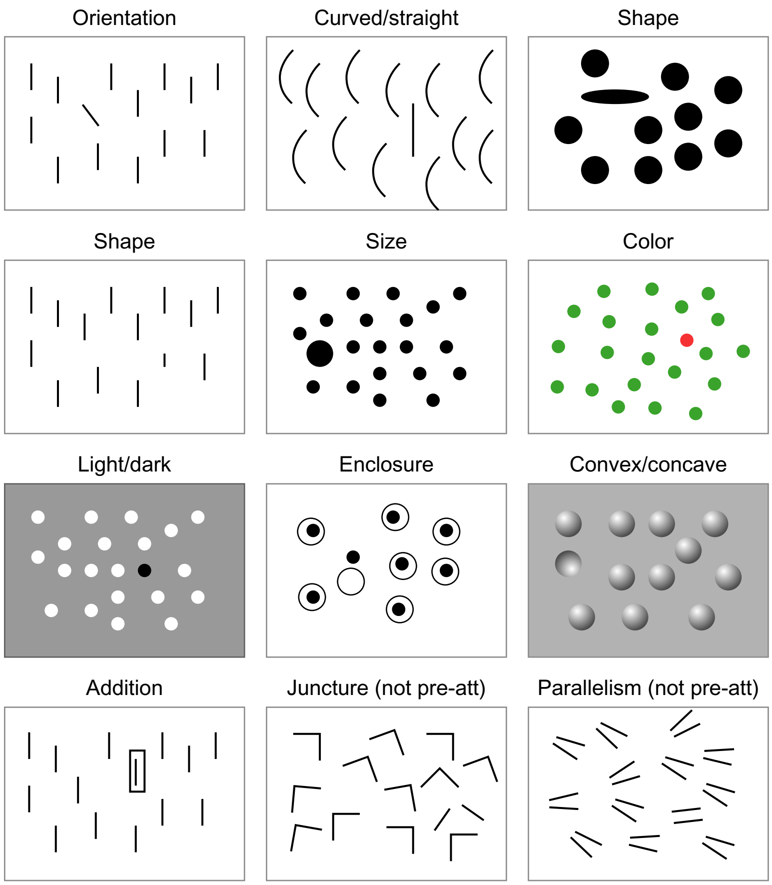


Figure 5.11 Most of the preattentive examples given here can be accounted for by the processing characteristics of neurons in the primary visual cortex.

mark on a map as being of type A, it should be differentiated from all other marks in a preattentive way. There have been literally hundreds of experiments to test whether various kinds of features are processed preattentively. Figure 5.11 illustrates a few of the results. Orientation, size, basic shape, convexity, concavity, and an added box around an object are all preattentively processed. However, the junction of two lines is not preattentively processed; neither is the parallelism of pairs of lines, so it is more difficult to find the targets in the last two boxes in Figure 5.11.

The features that are preattentively processed can be organized into a number of categories based on form, color, motion, and spatial position.

- Line orientation
- Line length
- Line width

- Size
- Curvature
- Spatial grouping
- Blur
- Added marks
- Numerosity (one, two, or three objects)
- Color
- Hue
  - Intensity
- Motion
  - Flicker
  - Direction of motion
- Spatial position
  - Two-dimensional position
  - Stereoscopic depth
- Convex/concave shape from shading

Originally, studying preattentive features was thought to be a way of measuring the primitives of early visual processing (Treisman & Gormican, 1988). The list given above, however, is considerably longer than the one that resulted from neuropsychological studies. Still, in most of these instances, the target is different from the surrounding nontargets in terms of the basic channels introduced earlier (color, size, contrast, orientation).

There is a risk of misinterpreting the findings of psychophysical studies and proposing a new kind of detector for every distinct shape. To take a single example, curved lines can be preattentively distinguished from straight lines. Despite this, it may be a mistake to think that there are curved line detectors in early vision. It may simply be the case that cells responsive to long, straight-line segments will not be strongly excited by the curved lines. Of course, it may actually be that early vision curvature detectors do exist; it is just that the evidence must be carefully weighed. It is not a good idea to propose a new class of detector for everything that exhibits the popout effect. The scientific principle of finding the most parsimonious explanation, known as Occam's razor, applies here.

It is also important to note that not all preattentive effects are equally strong. There are degrees of popout. In general the strongest effects are based on color, orientation, size, contrast, and motion or blinking, corresponding to the findings of neuropsychology. Effects such as line curvature tend to be weaker. Also, there are degrees of difference. Large color differences have more popout than small ones. Some popout effects occur with no instruction and are difficult to miss, such as the red 3s in Figure 5.9 and blinking points, but other patterns labeled preattentive require considerable attention for them to be seen. So the term *preattentive* should not be taken too literally because prior attention must be given to prime the relevant properties using the tuning mechanisms we have already discussed.

[G5.6] Use strong preattentive cues before weak ones where ease of search is critical.

Attention and Expectations

A problem with most research into attention, according to a book by Arien Mack and Irvin Rock (1998), is that almost all perception experiments (except their own) demand attention in the very design. The authors have a point. Typically, a subject is paid to sit down and pay close attention to a display screen and to respond by pressing a key when some specified event occurs. This is not everyday life. Usually we pay very little attention to what goes on around us. To understand better how we see when we are not primed for an experiment, Mack and Rock conducted a laborious set of experiments that only required one observation from each experiment. They asked subjects to look at a cross for a fraction of a second and report when one of the arms changed length. So far, this is like most other perception studies, but the real test came when they flashed up something near the cross that the subjects had *not* been told to expect. Subjects rarely saw this unexpected pattern, even though it was very close to the cross they were attending to in the display. Mack and Rock could only do this experiment once per subject, because as soon as subjects were asked if they had seen the new pattern they would have started looking for "unexpected" patterns. Hundreds of subjects had to be used, but the results were worth it; they tell us how much we are likely to see when we are looking for something else. The answer is, not much.

The fact that most subjects did not see a wide range of unexpected targets tells us that humans do not perceive much unless we have a need to find something and a vague idea of what that something looks like. In most systems, brief, unexpected events will be missed. Mack and Rock initially claimed from their results that there is no perception without attention; however, because they found that subjects generally noticed larger objects, they were forced to abandon this extreme position.

The question of which visual dimensions are preattentively stronger and therefore more salient cannot be answered in a simple way, because it always depends on the strength of the particular feature and the context. For example, Callaghan (1989) compared color to orientation as a preattentive cue. The results showed that the preattentiveness of the color

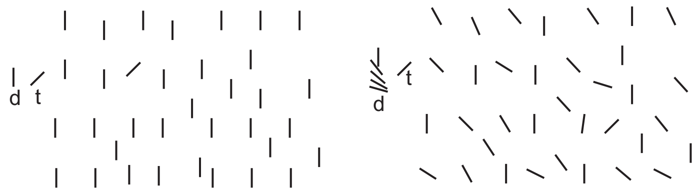


Figure 5.12 On the left, the right-slanted bar pops out; on the right, it does not. Yet, most of the distractors on the right have an orientation that is more different from the target orientation than the distracters on the left.

depended on the saturation (vividness) and size of the color patch, as well as the degree of difference from surrounding colors. So it is not just a question of color versus orientation, but exactly how the color differs from other colors in the set. Similarly, the preattentiveness of line orientation depends on the length of the line, the degree to which it differs from surrounding lines, and the contrast of the line pattern with the background. Figure 5.12 shows how an oblique line stands out from a set of vertical lines. When the same oblique line is in a set of lines of various orientations it is much more difficult to see, even though the difference in orientation between the target and the distractor set is just as large or larger. One thing that is clear from this example is that preattentive symbols become less distinct as the variety of distractors increases. It is easy to spot a single hawk in a sky full of pigeons, mostly because it has a different motion pattern, but if the sky contains a greater variety of birds, the hawk will be more difficult to see.

Studies have shown that two factors are important in determining whether something stands out preattentively: the degree of difference of the target from the nontargets and the degree of difference of the nontargets from each other (Quinlan & Humphreys, 1987; Duncan & Humphreys, 1989). For example, yellow highlighting of text works well if yellow is the only color in the display besides black and white, but if there are many colors the highlighting will be less effective.

[G5.7] For maximum popout, a symbol should be the only object in a display that is distinctive on a particular feature channel; for example, it might be the only item that is colored in a display where everything else is black and white.

Highlighting and Asymmetries

Another issue relating to making targets distinctive comes from research that has revealed *asymmetries* in some preattentive factors; for example, *adding* marks to highlight a symbol is generally better than taking them away (Treisman & Gormican, 1988). If all of the symbols in a set except for a target object have an added mark, the target will be less distinctive. It is better to highlight a word by underlining it than to underline all the words in a paragraph except for the target word. Another asymmetry

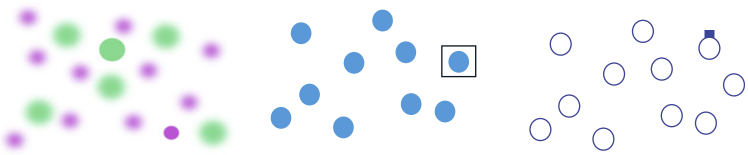


Figure 5.13 A number of highlighting methods that use positive asymmetric preattentive cues: sharpness, added surrounding feature, added shape.

is the finding that a big target is easier to see surrounded by small targets than a small target surrounded by big targets. Several examples are given in Figure 5.13.

[G5.8] Use positively asymmetric preattentive cues for highlighting.

When a visual design is complex, employing color, texture, and shape, the highlighting problem becomes more difficult. If all of the fonts in a display have the same size, for example, an increase in size can be used for highlighting.

[G5.9] For highlighting, use whatever feature dimension is used least in other parts of the design.

Modern computer graphics permit the use of motion for highlighting. This can be very effective when there is little other motion in the display (Bartram & Ware, 2002; Ware & Bobrow, 2004); however, making things move may be too strong a cue for many applications, although quite subtle motion can be effective.

[G5.10] When color and shape channels are already fully utilized, consider using motion or blink highlighting. Make the motion or blinking as subtle as possible, consistent with rapid visual search.

A relatively new idea for highlighting is the use of blur. Kosara et al. (2002) suggested blurring everything else in the display to make certain information stand out. They call the technique *semantic depth of field*, because it applies the depth-of-focus effects that can be found in photography to the display of data according to semantic content. As Figure 5.13 illustrates, blur works well, although again there is an obvious potential drawback to the technique. By blurring, the designer runs the risk of making important information illegible, as it is usually not possible to reliably predict the interests of the viewer.

Coding with Combinations of Features

So far we have been concentrating on using a single visual channel to make symbols distinct, or to highlight; often, though, we may wish to make objects distinctive using two or more channels. There are two issues here. The first is using redundant coding for extra distinctiveness. The second is, what can we expect if we use more complex patterns in symbol design?

Coding with Redundant Properties

We can choose to make something distinct on a single feature dimension, such as color, or we can choose to make it distinct on several dimensions, such as color, size, and orientation. This is called *redundant coding*. It means that someone can search based on any or all of the properties. The degree to which search is improved by redundant coding is a complex issue; sometimes the benefit is a simple addition and sometimes it is less than additive. It depends on what visual properties are being employed and the background. Nevertheless, there is almost always a benefit to redundant coding (Eriksen & Hake, 1955; Egeth & Pachella, 1969). Figure 5.8 gives examples of redundant coding of symbols in scatterplots.

[G5.11] To make symbols in a set maximally distinctive, use redundant coding wherever possible; for example, make symbols differ in both shape and color.

What Is Not Easily Findable: Conjunctions of Features

So far we have been discussing what can easily be found, but what kinds of things are difficult to spot? The answer is that, even if visual patterns get just a little bit more complex, a search can change from being almost instantaneous to something requiring much longer serial processing. What happens, for example, if we wish to search for a red square, not just something that is red or something that is square? Figure 5.14 illustrates a conjunction search task in which the targets are three red squares. It turns out that this kind of search is slow if the surrounding objects are squares (but not red ones) and other red shapes. We are forced to do a serial search of *either* the red shapes or the square objects. This is called a *conjunction* search, because it involves searching

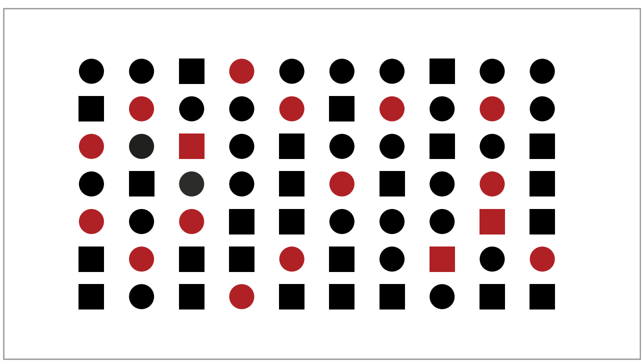


Figure 5.14 Searching for the red squares is slow because they are identified by a conjunction of shape and color.

for the specific conjunction of redness *and* shape attributes (Treisman & Gelade, 1980). This is very different from redundant coding, where parallel search can be carried out on one *or* the other. Conjunction searches are generally not preattentive, although there are a few very interesting exceptions that we will get to shortly.

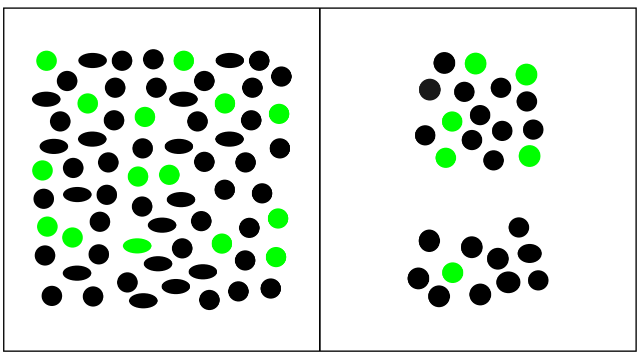
The fact that conjunction searches are slow has broad implications. It means, among other things, that we cannot learn to rapidly find more complex patterns. Even though we may have hundreds or thousands of hours of experience with a particular symbol set, searching for conjunctions of properties is still slow, although a modest speedup is possible (Treisman et al., 1992).

> **[G5.12]** If symbols are to be preattentively distinct, avoid coding that uses conjunctions of basic graphical properties.

Highlighting Two Data Dimensions: Conjunctions That Can Be Seen

Although early research suggested that conjunction searches were never preattentive, it has emerged that there are a number of preattentive dimension pairs that do allow for conjunctive search. Interestingly, these exceptions are all related to space perception. Searches can be preattentive when there is a conjunction of spatially coded information and a second attribute, such as color or shape. The spatial information can be a position on the *XY* plane, stereoscopic depth, shape from shading, or motion.

Spatial grouping on the *XY* plane. Treisman and Gormican (1988) argued that preattentive search can be guided by the identification of spatial clusters. This led to the discovery that the conjunction of space and color can be searched preattentively. In Figure 5.15(a), we cannot conjunctively search for green ellipses, but in Figure 5.15(b), we can rapidly search the conjunction of lower cluster and green target.



(a) (b)

Figure 5.15 (a) With the conjunction of shape and color, search is slow. (b) If we search the lower group for the green object, the search is fast. This is also a conjunction.

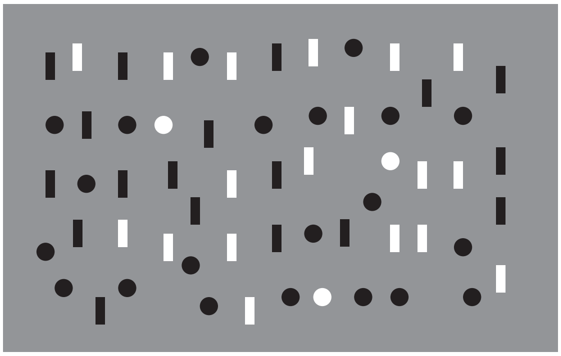


Figure 5.16 The white circles are a conjunction of shape and luminance polarity; nevertheless, they can be found preattentively.

Stereoscopic depth. Nakayama and Silverman (1986) showed that the conjunction of stereoscopic depth and color, or of stereoscopic depth and movement, can be preattentively processed.

Luminance polarity and shape. Theeuwes and Kooi (1994) showed that luminance polarity with targets lighter and darker than a gray background can support a preattentive conjunction search. In Figure 5.16, the white circles can be searched in parallel (a conjunction of whiteness and shape).

Convexity, concavity, and color. D'Zmura et al. (1997) showed that the conjunction of perceived convexity and color can be preattentively processed. In this case, the convexity is perceived through shape-from-shading information.

Motion. Driver et al. (1992) determined that motion and target shape can be preattentively scanned conjunctively. Thus, if the whole set of targets is moving, we do not need to look for nonmoving targets. We can preattentively find, for example, the red moving target. This may be very useful in producing highlighting techniques that allow for a preattentive search within the set of highlighted items (Bartram & Ware, 2002; Ware & Bobrow, 2004).

An application in which preattentive spatial conjunction may be useful is found in geographic information systems (GISs). In these systems, data is often characterized as a set of layers—for example, a layer representing the surface topography, a layer representing minerals, and a layer representing ownership patterns. Such layers may be differentiated by means of motion or stereoscopic depth cues.

[G5.13] When it is important to highlight two distinct attributes of a set of entities, consider coding one using motion or spacial grouping and the other using a property such as color or shape.

Ware and Bobrow (2005) used a conjunction coding method in an interactive network visualization application. To make it possible to trace paths in a visually impenetrable mass of hundreds of nodes and links, we added a feature whereby when someone touched a node a subnetwork of closely linked nodes and edges jiggled by a small amount (motion coding). This made the subnetwork stand out strongly from the background information. Previously found subnetworks were highlighted in a more conventional way using color coding. We found that it was easy for people to focus either on the recently selected subnetwork or on the previously selected subnetwork or on a sub-subnetwork that was both recently and previously selected (a conjunction).

Integral and Separable Dimensions: Glyph Design

Another body of theory that is relevant to glyph design is the theory of *integral and separable dimensions*, developed by Garner (1974). The kind of multidimensional coding that occurs in the use of glyphs raises questions about the perceptual independence of the display dimensions. In many ways, the lessons are the same as from channel theory (visual dimensions are much the same as channels), but Garner's theory provides a useful alternative description. We will use it to discuss approaches to glyph design.

Sometimes we need a symbol to do more than simply stand for something. Sometimes it is useful if symbols can convey how large, hot, or wet something is. Figure 5.17 shows an example in which the size of each circle represents the population of a country, and the color represents the geographic region to which that country belongs. In

Figure 5.17 In this visualization, the color of each circle represents a geographic region and the size of the circle represents population. (*From* www.gapminder.org.)

this case, color functions as a nominal coding, whereas size represents a quantity, a ratio coding. Symbols that represent quantity are called *glyphs*. To create a glyph, one or more quantitative data attributes are mapped in a systematic way to the different graphical properties of an object.

Garner's theory helps us answer questions such as, "Will the color-coding scheme interfere with our perception of glyph size and therefore distort perceived population level?" or "What if we use both color and size to represent a single variable—will this make the information clearer?" The concept of integral vs. separable visual dimensions tells us when one display attribute (e.g., color) will be perceived independently from another (e.g., size).

With *integral display dimensions*, two or more attributes of a visual object are perceived holistically and not independently. An example is a rectangular shape, perceived as a holistic combination of the rectangle's width and height. Another is the combination of green light and red light; this is seen holistically as yellow light, and it is difficult to respond independently to the red and green components.

With *separable dimensions*, people tend to make separate judgments about each graphical dimension. This is sometimes called *analytic processing*. Thus, if the display dimensions are the diameter of a ball and the color of a ball, they will be processed relatively independently. It is easy to respond independently to ball size and ball color. Integral and separable dimensions have been determined experimentally in a number of ways.

Three experimental paradigms are discussed here. All are related to interactions between pairs of graphical qualities, such as size and color. Very little work has been done on interactions among three or more graphical qualities.

Restricted Classification Tasks

In restricted classification tasks, observers are shown sets of three glyphs that are constructed according to the diagram shown in Figure 5.18. Two of the glyphs (A and B) are made the same on one graphical feature dimension. A third glyph (C) is constructed so that it is closer to glyph B in feature space, but this glyph differs from the other two in both of the graphical dimensions. Subjects are asked to group the two glyphs that they think go together best. If the dimensions are integral, B and C are grouped together because they are closest in the feature space. If they are separable, A and B are grouped together because they are identical in one of the dimensions (analytic mode). The clearest example of integral dimensions is color space dimensions. If dimension X is the red–green dimension and dimension Y is the yellow–blue dimension of color space, subjects tend to classify objects (roughly) according to the Euclidean distance between the colors (defined according to one of the uniform color spaces discussed in Chapter 4). Note that even this is not always the case, as the evidence of color categories (also discussed in Chapter 4) shows.

The width and height of an ellipse create an integral perception of shape. Thus, in Figure 5.19(a, top) the ellipses appear to be more similar to each other than to the circle, even though the width of the circle matches the width of the first ellipse. If

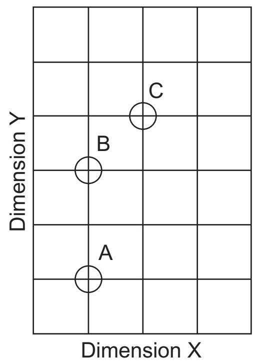


Figure 5.18 It is useful to think in terms of two display dimensions when considering the integral–separable concept. One dimension might be color, while another might be some aspect of shape.

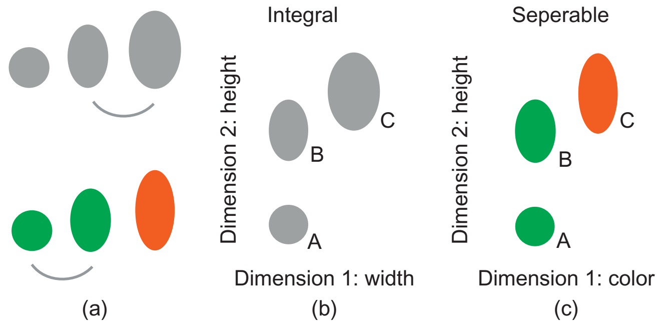


Figure 5.19 (a) The width and height of an ellipse are perceived integrally, so the ellipses are seen as more similar to each other (because they have the same shape) than the pair having the same width. The color and height of a shape are perceived separably, so the two green shapes are seen as most similar. (b, c) Space plots of the two examples.

the two dimensions are separable, subjects act in a more analytic manner and react to the fact that two of the objects are actually identical on one of the dimensions. Shape and color are separable. Thus, in Figure 5.19(a, below) either the green shapes or the two elliptical shapes will be categorized together. With separable dimensions, it is easy to attend to one dimension or the other.

Speeded Classification Tasks

Speeded classification tasks tell us how glyphs can visually interfere with each other. In a speeded classification task, subjects are asked to quickly classify visual patterns according to only one of the visual attributes of a glyph. The other visual attribute can be set up in two different ways; it can be given random values (interference condition), or it can be coded in the same way as the first dimension (redundant coding). If the data dimensions are integral, substantial interference occurs in the first case. With redundant coding, classification is generally speeded for integral dimensions. With separable codes, the results are different. There is little interference from the

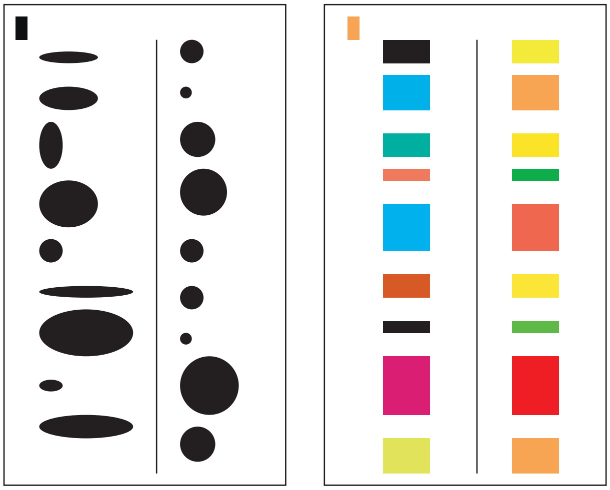


Figure 5.20 Sets of patterns for a speeded classification task. In both cases, (a) and (b), participants are required to respond positively to only those glyphs that have the same *height* as the bar in the upper corner. The interference condition is on the left in both (a) and (b). (a, left) The variable widths interfere with classification based on height. (b, left) The variable color does not interfere with classification based on height. (a, right) Redundant size coding speeds classification. (b, right) Redundant color and size coding does not speed classification.

irrelevant graphical dimension, but there is also little advantage in terms of speeded classification when redundant coding is used. Of course, in some cases, using redundant separable codes may still be desirable; for example, if both color and shape are used for information coding, color-blind individuals will still have access to the information. Figure 5.20 gives examples of the kinds of patterns that are used in integral–separable dimension experiments, illustrating these points.

The lessons to be learned from integral–separable dimension experiments are easy to apply in cases in which each data entity has only two attributes.

[G5.14] If it is important for people to respond holistically to a combination of two variables in a set of glyphs, map the variables to integral glyph properties.

[G5.15] If it is important for people to respond analytically to a combination of variables, making separate judgments on the basis of one variable or the other, map the variables to integral glyph properties.

Figure 5.21 shows how integral dimensions can help us perceive the combination of two variables. The body mass index is a common measure of obesity. This index is a ratio of height squared to weight. If we use two integral values, ellipse height and ellipse width, to show height squared and weight respectively, then we can arrange the plot in such a way that the ideal height-to-weight relationship is a perfect circle. Someone who is overweight will be represented as a squashed ellipse, while someone who is very thin will be represented by a tall ellipse. On the left side of Figure 5.21, we can see at a glance who is overweight and who is underweight.

The right-hand side of Figure 5.21 shows the same data represented using two separable variables: red–green variation for weight and vertical size for height. This is a poor choice, as it is very difficult to see who is overweight and who is underweight.

We can also apply the lessons of integral and separable dimensions to data glyphs designed to represent many variables. Figure 5.22 shows a field of data glyphs from Kindlmann and Westin (2006) in which three variables are mapped to color and many more are mapped to the shape of the glyphs. Detailed knowledge of the application would be required to decide if this is a good representation, but this is not our concern

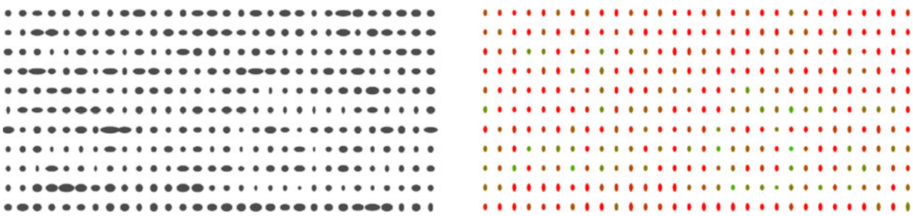


Figure 5.21 Height and weight data from 400 elderly Dutch people is displayed. On the left, height squared is mapped to the height of each ellipse and the weight is mapped to the width. On the right, weight is mapped to color and the width is held constant (red is more, green is less).

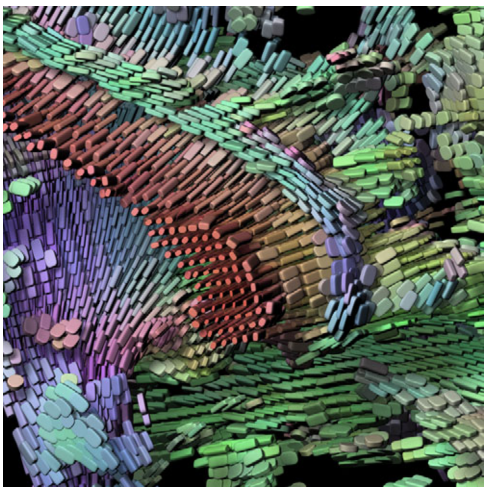


Figure 5.22 This map of a tensor field from Kindlmann and Westin (2006) has some variables mapped to the color of the lozenge-like glyphs and some variables mapped to their shape and orientation.

here. The point of showing it is to illustrate how the color-mapped variables tend to be seen integrally and independently (separably) from the shape variables, which also tend to be viewed holistically, making up the lozenge shapes.

Integral–Separable Dimension Pairs

The preceding analysis presented integral and separable dimensions as if they were qualitatively distinct. This overstates the case; a continuum of integrality–separability more accurately represents the facts. Even between the most separable dimension pairs, there is always some interference between different data values presented using the different channels. Likewise, the most integral dimension pairs can be regarded analytically to some extent. We can, for example, perceive the degree of redness and the degree of yellowness of a color—for example, orange or pink. Indeed, the original experimental evidence for opponent color channels was based on analytic judgments of exactly this type (Hurvich, 1981).

Figure 5.23 provides a list of display dimension pairs arranged on an integral–separable continuum. At the top are the most integral dimensions. At the bottom are

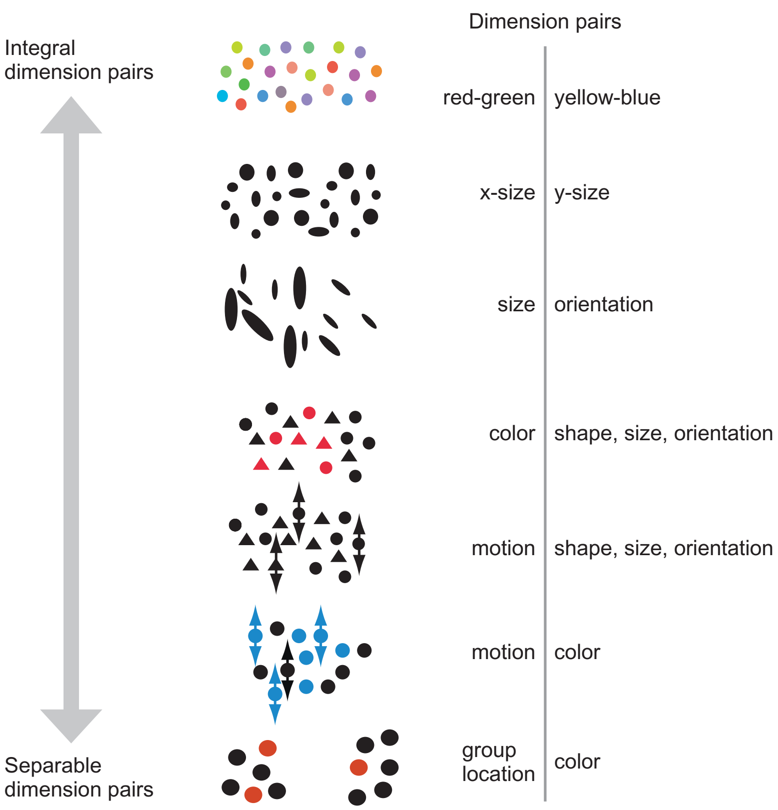


Figure 5.23 Examples of glyphs coded according to two display attributes. At the top are more integral coding pairs. At the bottom are more separable coding pairs.

the most separable dimensions. Some display dimensions are not represented in Figure 5.23 because of too little evidence. For example, one method of separating values is to use stereoscopic depth. It seems likely that stereoscopic depth is quite separable from other dimensions, especially if only two depth layers are involved.

As a theoretical concept, the notion of integral and separable dimensions is undoubtedly simplistic; it lacks mechanism and fails to account for a large number of exceptions and asymmetries that have been discovered experimentally. Also, it says essentially the same thing as channel theory, and channel theory has a firm neuropsychological basis. The beauty of the integral–separable distinction lies in its simplicity as a design guideline.

Representing Quantity

Some visual qualities increase continuously, such as size, brightness, or height above the ground, and are said to be monotonic. Some visual qualities are not monotonic. Orientation is one. It is meaningless to say that one orientation is greater or less than another. The same is true of the phase angle between two oscillating objects. As the phase difference is increased, the objects first appear to move in opposite directions, but as the phase difference continues to increase, they appear to move together again. Phase is cyclic, just as line orientation is cyclic. Hue also lacks a natural order.

Monotonic display variables naturally express relations, such as greater than or less than, if they have a quality that we associate with increasing value. For example, in a three-dimensional data space, the up direction is defined by gravity, and using up to represent a greater quantity of some variable will be readily interpreted, but the left and right directions do not have as clear a value. In the West, we read left to right but this is learned. Other languages, such as Arabic, have right-to-left ordering.

> **[G5.16]** When designing a set of glyphs to represent quantity, mapping to any of the following glyph attributes will be effective: size, lightness (on a dark background), darkness (on a light background), vividness (higher saturation) of color, or vertical position in the display.

If we map a data variable to some visual attribute such as length, area, volume, or lightness, we should be able to judge relative quantities at a glance, although not very accurately. Because of the simultaneous contrast effects discussed previously, grayscales are particularly subject to error. Judged length and area are also subject to contrast effects, although these will be smaller. Length will generally be judged more accurately than area, particularly if people are required to judge the relative area of different shapes (Cleveland & McGill, 1984). Perceived volume is judged very poorly, as shown in a study by Ekman and Junge (1961), who found perceived volume to be

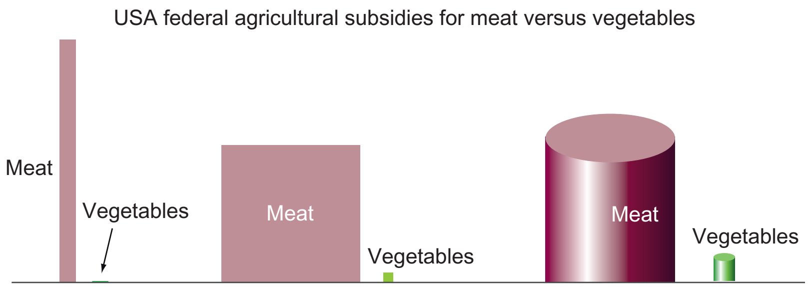


Figure 5.24 The same information is shown using length, area, and volume. Research shows that the quantities shown in the volume display on the right will be mostly judged according to the relative area of the images, not according to volume, resulting in large errors.

proportional to the actual volume raised to a power of 0.75. Because this is close to the relationship between volume and area, they concluded subjects were actually using area rather than volume for their judgments. The advantage to using area over length is that area is capable of conveying larger variations; for example, a 1-mm square can be compared to a 1-cm square, giving a ratio of 100:1. If length were used instead and the larger quantity were represented by 2 cm, then the smaller quantity would have to be represented by a length of only 0.2 mm, something barely visible. Figure 5.24 illustrates this point with data representing U.S. federal subsidies for meat and vegetables. Area representation is useful when relative amounts vary greatly.

> **[G5.17]** Ideally, use glyph length or height, or vertical position, to represent quantity. If the range of values is large, consider using glyph area as an alternative. Never use the volume of a three-dimensional glyph to represent quantity.

Mathematicians and engineers use the logarithmic plot in cases where there is a large variation in magnitude, but this is not a perceptual solution, relying instead on a technical understanding of the effects of the logarithmic transformation.

Representing Absolute Quantities

Visualization is mostly about seeing patterns in data, and this means that seeing if a particular variable is relatively larger or smaller than another is what is critical, rather than reading an absolute quantity. This is a good thing because the kinds of representation we have been discussing do not work well for representing quantities. Generally, only three to five distinct values can be reliably read using simple graphical variables such as color, size, or lightness. This means that glyphs using simple mappings are unsuitable for presenting data where values must be read from a display.

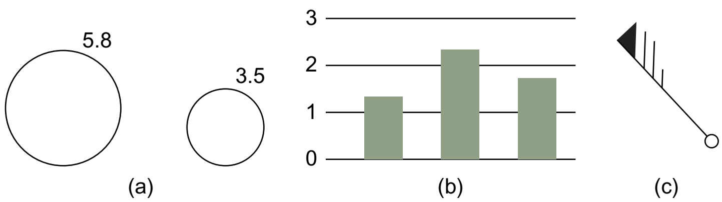


Figure 5.25 Three different ways that more exact numerical values can be read from a diagram.

There are a number of solutions to the problem of representing quantities. One is simply to add numbers to a glyph, or a numerical scale; see Figure 5.25(a, b). But, unless it is done carefully, the numbers will add visual noise, obscuring important patterns in data. A second solution is to create a glyph that by its shape conveys numerical values. The best known example of this is the wind barb, which is shown in Figure 5.25(c). A wind barb is a glyph widely used in meteorology that is a kind of hybrid of perceptual features and symbolic features. The shaft of the barb represents the direction of the wind. The "feathers" of the barb encode wind speed, so that someone familiar with the code can read off the wind speed to an accuracy of 5 knots. Given that surface wind speeds range up to about 150 knots, this means that wind barbs have about 30 steps of resolution, far better than any simple variation in size or color. The wind barb, however, has perceptual problems. The barb feathers greatly interfere with the perception of wind direction and because of this wind barbs are very poor at showing patterns in the winds.

Multidimensional Discrete Data: Uniform Representation versus Multiple Channels

This is a good place to step back and look at the general problem of multivariate discrete data display in light of the concepts that have been presented here and in previous chapters. It is worth restating this problem. We are provided with a set of entities, each of which has values on a number of attribute dimensions. For example, we might have 1000 beetles, each measured on 30 anatomical characteristics, or 500 stocks, each described by 20 financial variables. The reason for displaying such data graphically is often data exploration—to find meaning in the diversity. In the case of the beetles, the meaning might be related to their ecological niche. In the case of the stocks, the meaning is likely to lie in opportunities for profit. In either case, we are likely to be interested in patterns in the data, such as clusters of beetles that share similar attribute values.

If we decide to use a glyph display, each entity becomes a graphical object and data attributes are mapped to graphical attributes of each glyph. The problem is one of mapping data dimensions to the graphical attributes of the glyph. The work on preattentive processing, early visual processing, and integral and separable dimensions

Table 5.1 *Graphical attributes that may be useful in glyph design.*

Visual Variable	Dimensionality	Comment
Spatial position	Three dimensions: X, Y, Z	
Color	Three dimensions: defined by color opponent theory	Luminance contrast is needed to specify all other graphical attributes.
Shape	Size and orientation are basic but there may be more usable dimensions	The dimensions of shape that can be rapidly processed are unknown; however, the number is certainly small.
Surface texture	Three dimensions: orientation, size, and contrast	Surface texture is not independent of shape or orientation; uses one color dimension.
Motion coding	Approximately two to three dimensions; more research is needed, but phase is critical	
Blink coding	One dimension	Motion and blink coding are highly interdependent.

suggests that a rather limited set of visual attributes is available to us if we want to understand the values rapidly. Table 5.1 lists the most useful low-level graphical attributes that can be applied to glyph design, with a few summary comments about the number of dimensions available.

Many of these display dimensions are not independent of one another. To display texture, we must use at least one color dimension (luminance) to make the texture visible. Blink coding will certainly interfere with motion coding. Overall, we will probably be fortunate to display eight types of dimensional data clearly, using color, shape, spatial position, and motion to create the most differentiated set possible.

There is also the issue of how many resolvable steps are available in each dimension. The number here is also small. When we require rapid preattentive processing, only a handful of colors are available. The number of orientation steps that we can easily distinguish is probably about four. The number of size steps that we can easily distinguish is no more than four, and the values for the other data dimensions are also in the single-digit range. It is reasonable, therefore, to propose that we can represent about 2 bits of information for each of the eight graphical dimensions. If the dimensions were truly independent, this would yield 16 displayable bits per glyph (64,000 values). Unfortunately, conjunctions are generally not preattentive. If we allow no conjunction searching, we are left with four alternatives on each of eight dimensions, yielding only 32 rapidly distinguishable alternatives, a far smaller number. Anyone who has tried to design a set of easily distinguishable glyphs will recognize this number to be more plausible.

There is also the issue of the semantics associated with design choices, such as whether to use color or size to represent a particular attribute. Temperature, for example, has a natural mapping to color because of the association of redness with greater heat and blueness to lesser heat. Using the size of a glyph to represent temperature would normally be a poor design choice; however, there is a natural mapping between an increase in the amount of some variable and vertical size, or height above a baseline (Pinker, 2007). Orientation information, such as the direction of flow in a vector field is best represented by the orientation of a glyph—if a glyph representation is chosen, using color to represent orientation would normally be a poor design choice.

> **[G5.18]** In general, the use of heterogeneous display channels is best combined with meaningful mappings between data attributes and graphical features of a set of glyphs.

Stars and Whiskers

Sometimes no natural mappings to channels exist and it is desirable that there be a more symmetric mapping of data dimensions to the visual properties of a glyph. In this case, bar charts and star and whisker plots can be considered.

In the whisker plot, each data value is represented by a line segment radiating out from a central point, as shown in Figure 5.26(a). The length of the line segment denotes the value of the corresponding data attribute. A variant of the whisker plot is the star plot (Chambers et al., 1983). This is the same as the whisker plot but with the ends of the lines connected, as in Figure 5.26(b). It is possible to show a large number of variables with whisker or star plots, but this does not mean that the results will be intelligible. If there are a large number of whisker glyphs in a display, there will be visual interference between all contours having a similar orientation, with the star plot being the worst in this regard. In order to minimize interference between similarly oriented contours, a much smaller number of whiskers is recommended—four is probably the maximum. It may also be useful to change the amount of "energy" in glyph segments by

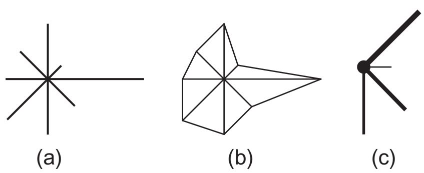


Figure 5.26 (a) Whisker plot. (b) Star plot. (c) Whisker plot with only four variables and varying width.

altering the line width as well as the length of the line; see Figure 5.26(c). A more common alternative to star and whisker plots is small bar charts, such as miniature versions of Figure 5.25(b). These have the advantage that orientation does not have to be taken into account in judging the represented quantities, and the bars can be color coded to make it easier to distinguish the variables.

When the data has many dimensions, a much better tool for its analysis is the parallel coordinates plot discussed in later chapters. This uses interactivity to minimize interference between dimensions.

The Searchlight Metaphor and Cortical Magnification

We now return to the topic of visual search with which we began this chapter. Consider the eyeball as an information-gathering searchlight, sweeping the visual world under the guidance of the cognitive centers that control our attention. Information is acquired in bursts, a snapshot for each fixation. More complex, nonpreattentive objects are scanned in series, one after another, at about the rate of 40 items per second. This means that we can typically parse somewhere between three and six items before the eye jumps to another fixation.

Useful Field of View

The attention process is concentrated around the fovea, where vision is most detailed; however, we can to some extent redirect attention to objects away from the fovea. The region of visual space we attend to expands and contracts based on task, the information in the display, and the level of stress in the observer. A metaphor for the fovea-centered attentional field is the *searchlight of attention*. When we are reading fine print, the searchlight beam is the size of the fovea, perhaps one centimeter from the point of fixation. If we are looking at a larger pattern, the searchlight beam expands. A concept called the *useful field of view* (UFOV) has been developed to define the size of the region from which we can quickly take in information. The UFOV varies greatly, depending on the task and the information being displayed. Experiments using displays densely populated with targets reveal small UFOVs, from 1 to 4 degrees of visual angle (Wickens, 1992). Drury and Clement (1978), however, have shown that for low target densities (less than one per degree of visual angle) the UFOV can be as large as 15 degrees. Roughly, the UFOV varies with target density to maintain a constant number of targets in the attended region. With greater target density, the UFOV becomes smaller and attention is more narrowly focused; with a low target density, a larger area can be attended.

Tunnel Vision, Stress, and Cognitive Load

A phenomenon known as *tunnel vision* has been associated with operators working under extreme stress. In tunnel vision, the UFOV is narrowed so that only the most important information, normally at the center of the field of view, is processed. This phenomenon has been specifically associated with various kinds of nonfunctional behaviors that occur during decision making in disaster situations. The effect can be demonstrated quite simply. Williams (1985) compared performance on a task that required intense concentration (high foveal load) to one that was simpler. The high-load task involved naming a letter drawn from six alternatives; the low-load task involved naming a letter drawn from two alternatives. They found a dramatic drop in detection rate for objects in the periphery of the visual field (down from 75% correct to 36% correct) as the task load increased. The Williams data shows that we should not think of tunnel vision strictly as a response to disaster. It may generally be the case that as cognitive load goes up, the UFOV shrinks.

[G5.19] When designing user interrupts, peripheral alerting cues must be made stronger if the cognitive load is expected to be high.

The Role of Motion in Attracting Attention

A study by Peterson and Dugas (1972) suggests that the UFOV function can be far larger for detection of moving targets than for detection of static targets. They showed that subjects can respond in less than 1 second to targets 20 degrees from the line of sight, if the targets are moving. If static targets are used, performance falls off rapidly beyond about 4 degrees from fixation (see Figure 5.27). This implies a UFOV of at least 40 degrees for the moving-targets task.

Motion as a User Interrupt

As we conduct more of our work in front of computer screens, there is an increasing need for signals that can attract a user's attention. Often someone is busy with a primary task, perhaps filling out forms or composing e-mail, while at the same time events may occur on other parts of the display that require attention. These *user interrupts* can alert us to an incoming message from a valued customer or a signal from a computer agent that has been out searching the Internet for information on the latest flu virus.

There are four basic visual requirements for a user interrupt:

1. A signal should be easily perceived, even if it is outside of the area of immediate focal attention.
2. If the user wishes to ignore the signal and attend to another task, the signal should continue to act as a reminder.
3. The signal should not be so irritating that it makes the computer unpleasant to use.
4. It should be possible to endow the signal with various levels of urgency.

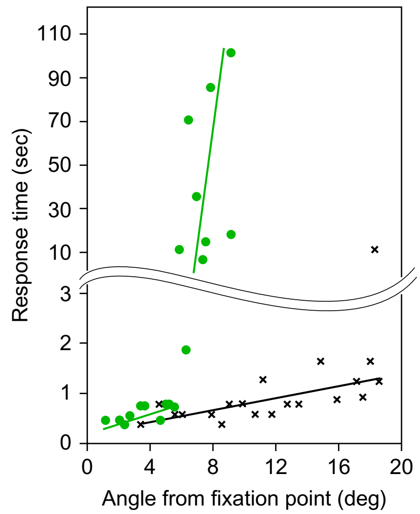


Figure 5.27 Results of a study by Peterson and Dugas (1972). The task was to detect small symbols representing aircraft in a simulation display. The circles show the response times from the appearances of static targets. The crosses show response times from the appearances of moving targets. Note the two different scales.

Essentially, the problem is how to attract the user's attention to information outside the central parafoveal region of vision (approximately the central 6 degrees). For a number of reasons, the options are limited. We have a low ability to detect small targets in the periphery of the visual field. Peripheral vision is color blind, which rules out color signals (Wyszecki & Stiles, 1982). Saccadic suppression (blindness during eye movements) means that some transitory event occurring in the periphery will generally be missed if it occurs during a saccadic movement (Burr & Ross, 1982). Taken together, these facts suggest that a single, abrupt change in the appearance of an icon is unlikely to be an effective signal.

The set of requirements suggests two possible solutions. One is to use auditory cues. In certain cases, these are a good solution, but they are outside the scope of this book. Another solution is to use blinking or moving icons. In a study involving shipboard alarm systems, Goldstein and Lamb (1967) showed that subjects were capable of distinguishing five flash patterns with approximately 98% reliability and that they responded with an average delay of approximately 2.0 seconds. Anecdotal evidence, however, indicates that a possible disadvantage of flashing lights or blinking cursors is that users find them irritating. Unfortunately, many web page designers generate a kind of animated chart junk; small, blinking animations with no functional purpose are often used to jazz up a page. Moving icons may be a better solution. Moving targets are detected more easily in the periphery than static targets (Peterson and Dugas, 1972). In a series of dual-task experiments, Bartram et al. (2003) had subjects carry out a primary task, either text editing or playing Tetris® or solitaire, while simultaneously monitoring for a change in an icon at the side of the display in the periphery of the visual field. The results showed that having an icon move was far more effective in attracting a user's attention than having it change color or shape. The advantage increased as the signal was farther from the focus of attention in the primary task. Another advantage of moving or blinking signals is that they can persistently attract attention, unlike a change in an icon, such as the raising of a mailbox flag, which fades rapidly from attention. Also, although rapid motions are annoying, slower motions need not be and they can still support a low level of awareness (Ware et al., 1992).

Interestingly, more recent work has suggested that it may not be motion *per se* that attracts attention, but the appearance of a new object in the visual field (Hillstrom and Yantis, 1994; Enns et al., 2001). This seems right; after all, we are not constantly distracted in an environment of swaying trees or people moving about on a dance floor. It also makes ecological sense; when early man was outside a cave, intently chipping a lump of flint into a hand axe, or when early woman was gathering roots out on the grasslands, awareness of emerging objects in the periphery of vision would have had clear survival value. Such a movement might have signaled an imminent attack. Of course, the evolutionary advantage goes back much further than this. Monitoring the periphery of vision for moving predators or prey would provide a survival advantage for most animals. Thus, the most effective reminder might be an object that moves into view, disappears, and then reappears every so often. In a study that measured the eye movements made while viewing multimedia presentations, Faraday and Sutcliffe (1997) found that the onset of motion of an object generally produced a shift of attention to that object.

Conclusion

This chapter has provided an introduction to the early stages of vision, in which billions of neurons act in parallel to extract elementary aspects of form, color, texture, motion, and stereoscopic depth. The fact that this processing is done for each point of the visual field means that objects differentiated in terms of these simple low-level features pop out and can be easily found. These low-level filters are not unbiased; they are tuned by the effects of top-down attention. This means that to a great extent what we need to see as well as what we expect to see will have a large influence on what we actually see.

For glyphs to be seen rapidly, they must stand out clearly from all other objects in their near vicinity on at least one coding dimension. In a display of large symbols, a small symbol will stand out. In a display of blue, green, and gray symbols, a red symbol will stand out. Because only simple basic visual properties guide visual search, glyphs and symbols that are distinctive in terms of more complex combinations of features cannot be easily found.

The lessons from this chapter have to do with fundamental tradeoffs in design choices about whether to use color, shape, texture, or motion to display a particular set of variables. These basic properties provide a set of channels that can be used to code information.

There is more visual interference within channels. The basic rule is that, in terms of low-level properties, "like" interferes with "like." If we have a set of small symbols on a textured background, a texture with a grain size similar to that of the symbols will make them difficult to see.

There is more separability between channels. If we wish to be able to read data values from different data dimensions, each of these values should be mapped to a different display channel. Mapping one variable to color and another to glyph orientation will make them independently readable. If we map one variable to height and another to width, they will be read more holistically. If we have a set of symbols that are difficult to see because they are on a textured background, they can be made to stand out by using another coding channel; having the symbols oscillate will also make them distinct. The way to differentiate variables readily is to employ more perceptual channels. Unfortunately, although this solves one problem, it creates another. We have to decide which variable to map to color, to shape, and to texture, and we have to worry about which mappings will be most intuitive for the intended audience. These are difficult design decisions.

This page intentionally left blank