CHAPTER 5

Visual Attention and Information that Pops Out

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. From an image buffer, the massively parallel machinery of early visual processing finds objects based on salient features of images. Once identified, complex objects are scanned in series, one after another, at about the rate of 25 items per second. This means that we can parse somewhere between four and twelve items before the eye jumps to another fixation. Understanding the steps in this process can help us with many visualization tasks. Here are some examples.

A tactical map display used by a military strategist must simultaneously show many different kinds of information about resources, such as equipment, personnel, and environmental conditions that exist in the field. Ideally, with such a display it should be possible either to attend to a single aspect of the data, such as the deployment of tanks, or, by an act of visual attention, to perceive the whole, complex, interwoven pattern. Understanding early vision is critical in understanding how to make information visually distinct or how to make the integrated patterns stand out.

In a scatter plot, each plotted data point can be made to represent many different kinds of information by using a glyph instead of an undifferentiated circle. A *glyph* is a graphical object designed to convey multiple data values. For information about stocks on the stock exchange, the color of an information glyph might be used to show the price-to-earnings ratio, the size of the glyph to display the growth trend, and the shape of the glyph to represent the type of company—square for technology stocks, round for resources, and so on. But what makes a glyph stand out if it is displayed in this way? This type of graphical tool will be most useful if the interesting stocks can be made to stand out and catch the analyst's eye.

Visual search provides one of the great benefits of visualization. It is possible, in less than a second, to detect a single dark pixel in a 500 × 500 array of white pixels. This screen can be replaced every second by another, enabling a search of more than 15 million pixels in a minute.

This is an artificial example, because most search tasks are more complex, but it does highlight the incredible parallel search capacity of the visual system.

Attention is both a low-level and a high-level property of vision. This chapter is concerned with the low-level mechanisms that help us understand what is more readily available to attention. A large body of vision research is related to this problem, and in many cases this information can be translated, in a fairly direct way, into design guidelines for data visualization.

Chapter 11 is concerned with the high-level direction of attention for problem solving.

Searching the Visual Field

A problem with most research into attention, according to a recent book by Arien Mack and Irvin Rock, is that almost all perception experiments (except their own) demand attention in the very design (Mack and Rock, 1998). They 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 and they were attending to 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, so hundreds of subjects were used. 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 at least some expectation and need to see it. 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.

Useful Field of View

The attention process is concentrated around the fovea, where vision is most detailed. However, we can redirect attention to objects within a single fixation, and 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 fovea-center attentional field is the searchlight of attention. When we are reading fine print, we can read the words only at the exact point of fixation. But we can take in the overall shape of a larger pattern at a single glance. In the former case, the searchlight beam is as narrow as the fovea, whereas in the latter it is much wider.

A concept called the *useful field of view* (UFOV) has been developed to define the size of the region from which we can rapidly 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). But Drury and Clement (1978) have shown that for low character densities (less than one per degree of visual angle), the useful visual field 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 and Stress

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 problem handling 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.

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 5.1.) This implies a useful field of at least 40 degrees across for the moving-targets task. However, this was merely the largest field that was measured. There is every reason to suppose that the useful visual field for moving targets is even larger; it may well encompass the entire visual field. Thus, motion of icons in user interfaces can be useful for attracting attention to the periphery of the screen (Bartram et al., 2003).

Reading from the Iconic Buffer

Figure 5.2 shows a collection of miscellaneous symbols. If we briefly flash such a collection of symbols on a screen—say, for one-tenth of a second—and then ask people to name as many of

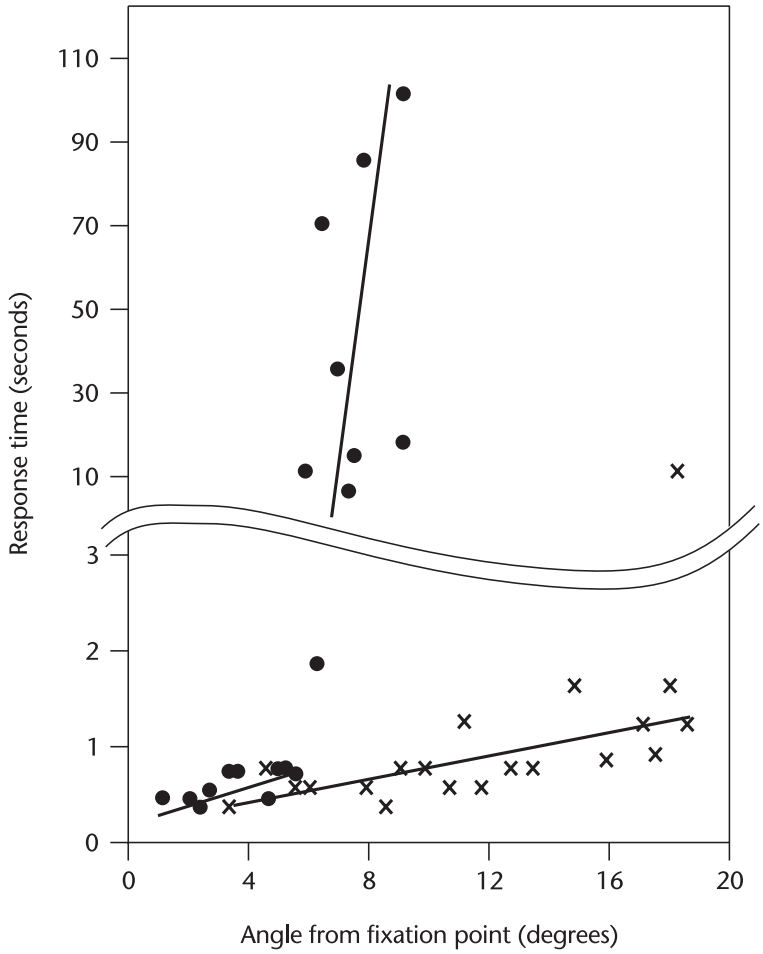


Figure 5.1 Results of a study by Peterson and Dugas (1972). The task was to detect small symbols representing aircraft in a simulation display. The circles show the response times from the appearances of static targets. The crosses show response times from the appearances of moving targets. Note the two different scales.

the symbols as they can, they typically produce a list of three to seven items. Several factors limit the number of items listed. The first is the short-lived visual buffer that allows us to hold the image for about one to two tenths of a second while we read the symbols into our short-term memory. This visual buffer is called *iconic memory*. Its properties were first described in a classic paper by Sperling (1960). See Humphreys and Bruce (1989) for a review. Any information that is retained longer than three-tenths of a second has been read into visual or verbal working memory (discussed in Chapter 11). This is an artificial example, but it has to do with a process that is very general. In each fixation between saccadic eye movements, an image of the world is captured in iconic memory; from this transient store higher-level processes must identify objects,

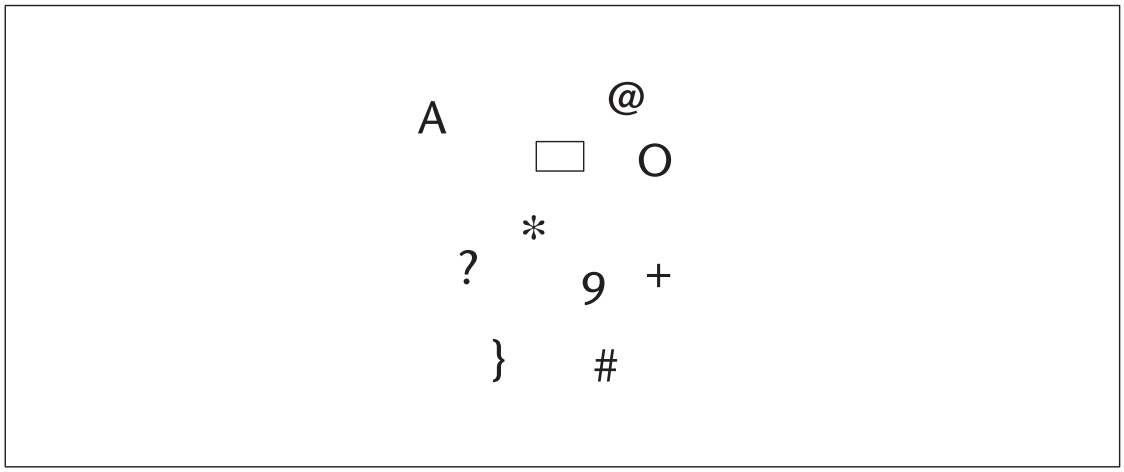


Figure 5.2 How many of these symbols can you remember after a glimpse one-tenth of a second long?

match them with objects previously perceived, and take information into working memory for symbolic analysis.

Preattentive Processing

We can do certain things to symbols to make it much more likely that they will be visually identified even after very brief exposure. Certain simple shapes or colors "pop out" from their surroundings. The theoretical mechanism underlying pop-out is called *preattentive processing* because logically it must occur prior to conscious attention. In essence, preattentive processing determines what visual objects are offered up to our attention. An understanding of what is processed preattentively is probably the most important contribution that vision science can make to data visualization.

Preattentive processing is best introduced with an example. To count the 3s in a table of digits in Figure 5.3(a), it is necessary to scan all the numbers sequentially. To count the 3s in Figure 5.3(b), it is necessary only to scan the red digits. This is because color is preattentively processed.

The typical experiment that is conducted to find out whether something is preattentively processed involves measuring the response time to find a target in a set of 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 independent of the number of distractors. Thus, if time to find the target is plotted against number of distractors, the result should be a horizontal line.

Figure 5.4 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 dark digit in the array of digits shown above is independent of the number of gray digits. The *X*s in Figure 5.4 show the results from processing a feature that is not preattentive. The time to respond depends on the number of distractors. The results of this kind of experiment are not

a 8568972698468976268976435892265986598655489768926989 8
0246299687402655762798678904567923276928546098677209 8
9083457980279075904709827908579084772908759082790875 4
9870985674906897578625984569024379047219079070981145 0
8568972698468976268976445892265986598655489768926989 8

b 8568972698468976268976435892265986598655489768926989 8
0246299687402655762798678904567923276928546098677209 8
9083457980279075904709827908579084772908759082790875 4
9870985674906897578625984569024379047219079070981145 0
8568972698468976268976445892265986598655489768926989 8

Figure 5.3 Preattentive processing. (a) To count the 3s in a table of digits, it is necessary to scan all the numbers sequentially. (b) To count the 3s in the next table, it is necessary only to scan the red digits. This is because color is preattentively processed.

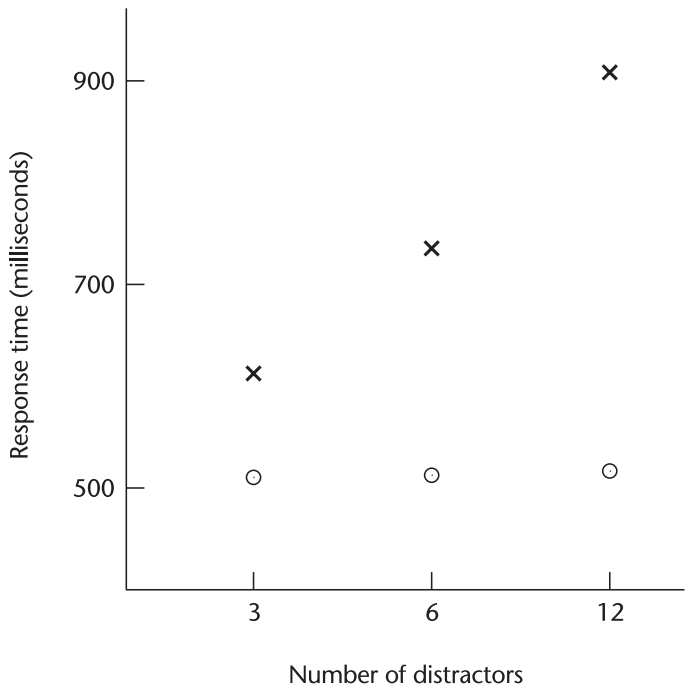


Figure 5.4 Typical results from a study of preattentive processing. The circles show time to perceive an object that is preattentively distinct from its surroundings. Time to process is independent of the number of irrelevant objects (distractors). The *X*s show how time to process non-preattentively distinct targets depends on the number of distractors.

always as perfectly clear-cut as Figure 5.4 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 non-preattentive targets are 40 msec per item and more (Triesman and 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 identify instantaneously some 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 5.5 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 harder to find the targets in the last two boxes in Figure 5.5.

The reason that preattentive processing has attracted so much attention among researchers is that it is thought to be a way of measuring the primitive features that are extracted in early visual processing (Triesman and Gormican, 1988). However, there is a risk of misinterpreting the findings of such studies. 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 pop-out effect. The scientific principle of finding the most parsimonious explanation, known as *Occam's razor*, applies here.

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

Form

- Line orientation
- Line length
- Line width
- Line collinearity
- Size
- Curvature
- Spatial grouping
- Blur
- Added marks
- Numerosity

Color

- Hue
- Intensity

Motion

- Flicker
- Direction of motion

Spatial Position

- 2D position
- Stereoscopic depth
- Convex/concave shape from shading

The results of preattentive processing experiments can be applied directly to the design of symbols for information display. In some cases, it may be desirable that each of many symbols be preattentively distinct from all the others. For example, in the case of a map of the ocean environment, we might wish to be able to scan visually only for scallop beds, only for fish farms, only for cod schools, or only for the fishing boats, assuming that we had all of this data. To make this possible, each type of symbol should be preattentively distinct from the others.

Figure 5.6 shows a set of nine symbols designed so that each is preattentively different from the others. The set could easily be extended—for example, by using blink coding. One thing that is clear from a cursory look at 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, 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 and Humphreys 1987; Duncan and 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. For another example, Chau and Yeh (1995) showed that preattentive segregation by stereoscopic depth decreased as the number of depth layers increased.

It is natural to ask which visual dimensions are preattentively stronger and therefore more salient. Unfortunately, this question cannot be answered, 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 depended on the saturation (vividness) and size of the color patch, as well as the degree of difference from surrounding colors. 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.

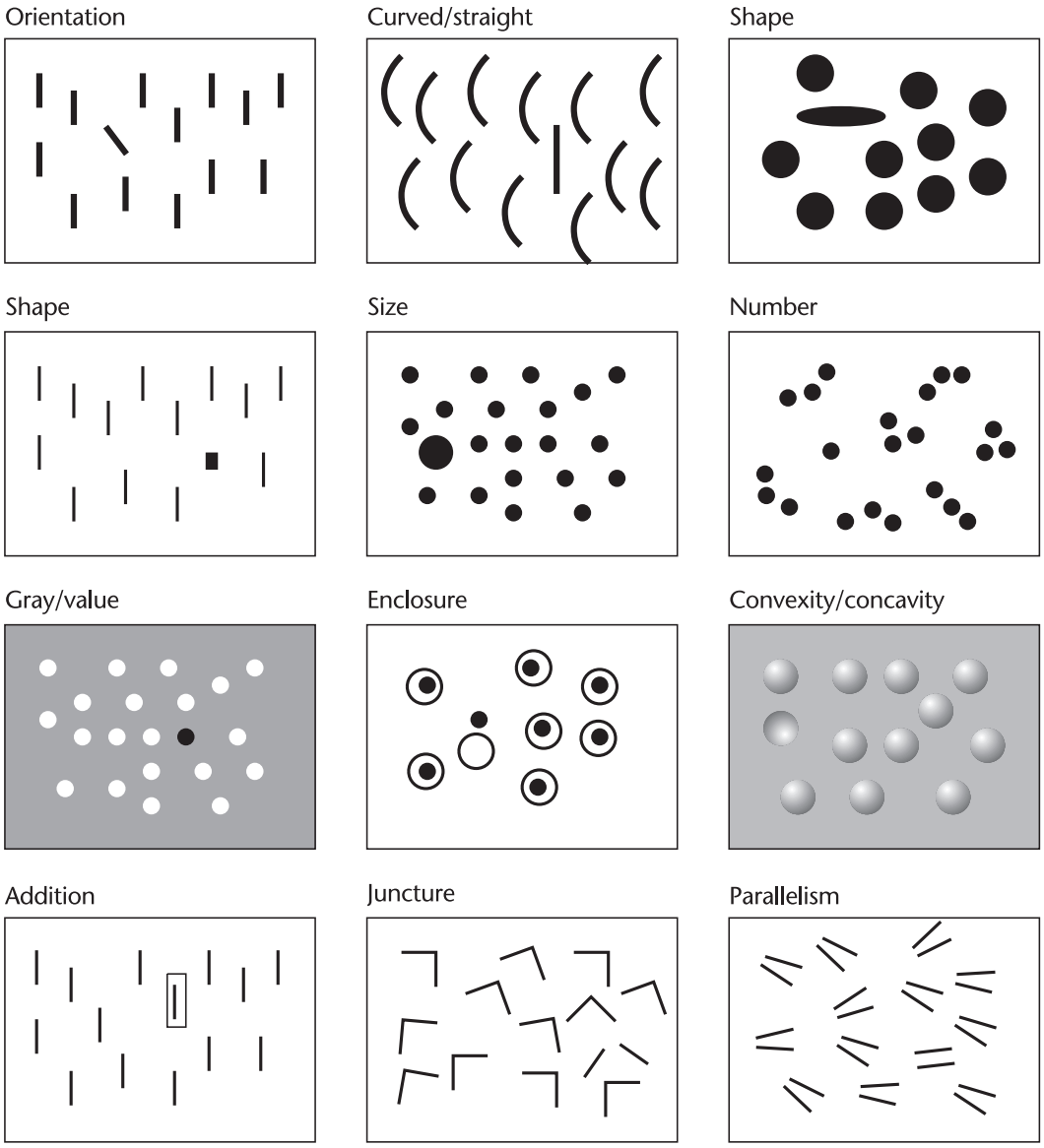


Figure 5.5 Most of the differences shown are preattentively distinguished. Only juncture and parallelism are not.

Numerous studies have addressed the preattentive properties of various combinations of features. It would be impossible to describe all the interactions without writing a complete book on the subject. However, some generalizations are in order. Adding marks to highlight something is generally better than taking them away (Triesman and Gormican, 1988). Thus, it is better to highlight a word by underlining it than to underline all the words in a paragraph except for the target word. It is also the case that simple numerosity is preattentively processed. We can see at

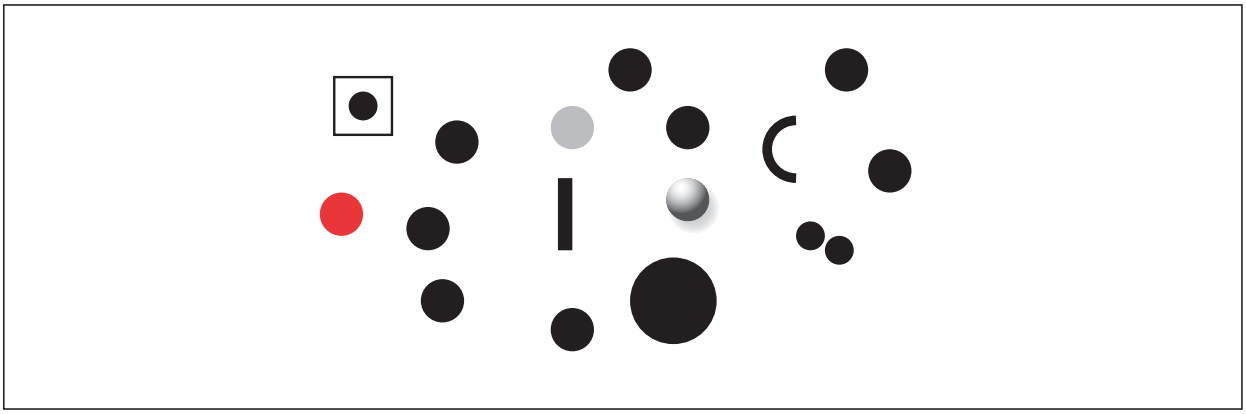


Figure 5.6 A set of symbols in which each of the nine symbol types is preattentively distinct from all the others.

a glance that there are one, two, three, or four objects in a group; this ability appears very early in human development (Dehaene, 1997). Once the number of objects increases beyond four, explicit counting is necessary.

That color is also preattentive has been well established, and the problem of defining a color that will be preattentively distinct from surrounding colors has already been discussed in Chapter 4. To restate a key finding, Bauer et al. (1996) showed that to be preattentively distinct, a color should lie outside the boundary of the region defined by all the other colors in the local part of the display (see Figure 4.19).

Rapid Area Judgments

Most work on preattentive processing has involved the detection of isolated targets. But other tasks can also benefit from rapid processing. In interpreting map data, a common task is to rapidly estimate the area of some region. Healey et al. (1998) showed that fast area estimation can be done on the basis of either the color or the orientation of the graphical elements filling a spatial region. It is a reasonable assumption that all the preattentive cues that have been identified for target identification are also valid for area estimation judgments.

Coding with Combinations of Features

A critical issue for information display is whether more complex patterns can be preattentively processed. For example, what happens if we wish to search for a gray square, not just something that is gray or something that is square? It turns out that this kind of search is slow if the surrounding objects are squares (but not gray ones) and other gray shapes. We are forced to do a serial search of either the gray shapes or the square objects. This is called a *conjunction search*, because it involves searching for the specific conjunction of gray-level and shape attrib-

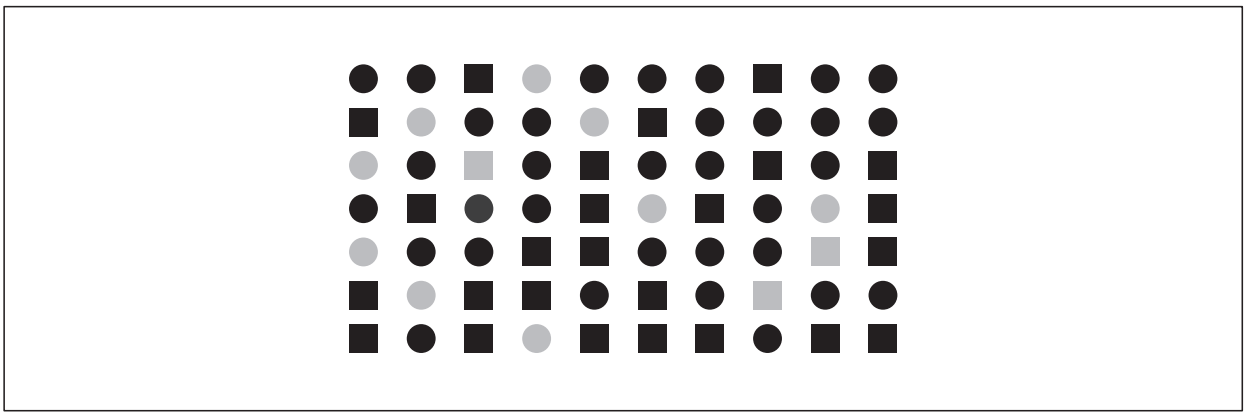


Figure 5.7 Searching for the gray squares is slow because they are identified by conjunction coding.

utes. Figure 5.7 illustrates a conjunction search task in which the targets are represented by three gray squares. Conjunction searches are generally not preattentive, although there are a few very interesting exceptions.

Conjunctions with Spatial Dimensions

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. 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 position on the *XY* plane, stereoscopic depth, shape from shading, or motion.

Spatial grouping on the *XY* plane: Triesman and Gormican (1988) argue that preattentive search can be restricted by the identification of visual clusters. This is a form of conjunction search: the conjunction of space and color. In Figure 5.8(a), we cannot conjunctively search for green ellipses, but in Figure 5.8(b), we can rapidly search the conjunction of lower grouping and gray target. The fact that the target is also elliptical is irrelevant.

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. This may be very useful in producing highlighting techniques allowing for a preattentive search within the set of highlighted items (Bartram and Ware, 2002).

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.

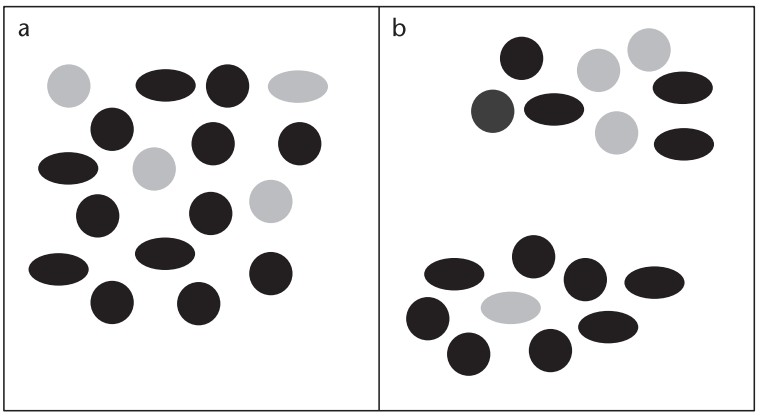


Figure 5.8 Spatial conjunction. The pattern on the left is a classic example of a preattentive conjunction search. To find the gray ellipses, either the gray things or the elliptical things must be searched. However, the example on the right shows that the search can be speeded up by spatial grouping. If attention is directed to the lower cluster, perceiving the gray ellipse is preattentive. This is a preattentive conjunction of spatial location and gray value.

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.

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.

Highlighting

The purpose of highlighting is to make some information stand out from other information. This is the most straightforward application of preattentive processing results. The problem of highlighting is easy to solve in homogeneous graphical displays. Yellow background highlighting of text works well for text that is black on white because yellow is a high luminance color that maintains text contrast. When the environment is visually complex, already employing color, texture, and shape, the problem becomes complex. As a rule of thumb, use whatever graphical dimension is least used otherwise in the design. For example, if texture is not used elsewhere, use it. 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 and Ware, 2002). However, making things move may be too strong a cue for many applications.

A 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 *seman-*

Figure 5.9 Blur can be used to highlight important information by blurring irrelevant information. *Kosara et al. (2002) call this technique semantic depth of field.*

tic 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 5.9 illustrates, blur works well, although again there are obvious drawbacks. If the designer is not completely sure what the user should attend to, he or she runs the risk of making important information illegible.

Designing a Symbol Set

One way to think about preattentive processing is to understand that we can easily and rapidly perceive the "odd man out" in visual feature space. If a set of symbols is to be designed to represent different classes of objects on a map display, then these symbols should be as distinct as possible. Military operational maps are an obvious example in which symbols can be used to represent many different classes of targets. (Targets are entities of operational importance that may be friendly or hostile.) A simplified example provides an interesting design exercise.

A tactical map might require the following symbols:

- Aircraft targets
- Tank targets
- Building targets
- Infantry position targets

In addition:

- Each of the target types can be classified as friendly or hostile.
- Targets exist whose presence is suspected but not confirmed.

Finally, there is a need to display features of the terrain itself. Roads, rivers, vegetation types, and topography are all important.

In this example, we encounter many of the characteristic problems of symbol set design. Even though this is a great simplification of the requirements of actual command and control displays, there are still many different types of things to be represented. There is a need for various orthogonal classifications (friendly vs. hostile, static vs. mobile). In some circumstances, conjunction search might be desirable (friendly tanks); in others, it would be useful if whole classes of objects could be rapidly estimated.

A solution to this simplified problem is illustrated in Figure 5.10. The actual symbols for the different target types have all been made preattentively distinct using shape. Color has been used to classify the targets preattentively into friendly and hostile ones. Possible targets are indicated by adding a thin rectangular box. Spatial grouping also helps to distinguish between friendly and hostile targets, but this would not always be the case. In a real application of this type, dozens more different symbols may be required on many different backgrounds, making the design tradeoffs much harder.

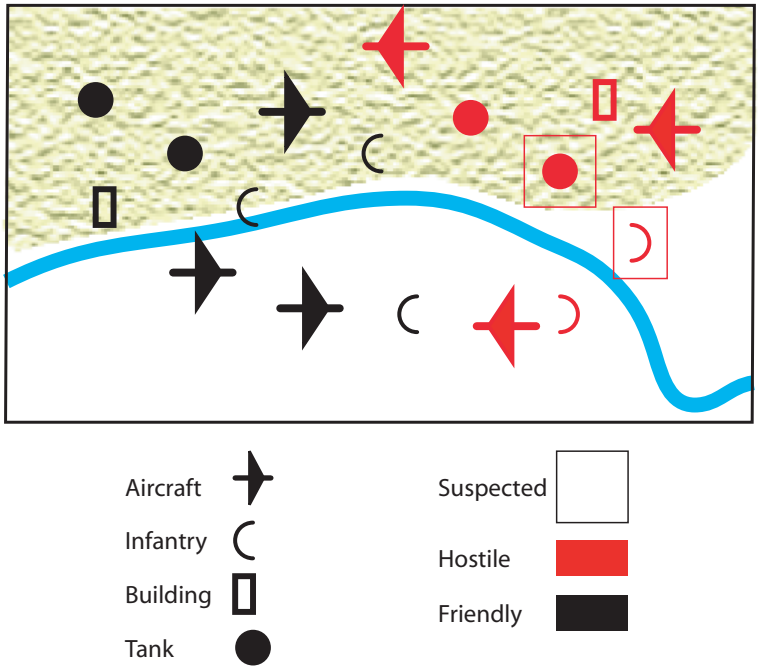


Figure 5.10 A set of symbols for a military command and control display.

Neural Processing, Graphemes, and Tuned Receptors

We now consider the same problem from a neurological perspective. Triesman and others claim that preattentive processing is due to *early* visual processing. What is the neurological evidence for this?

Visual information leaves the retina, passing up the optic nerve, through the neural junction at the lateral geniculate nucleus (LGN), and on to the much richer world 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 areas V1 and V2 are devoted to analyzing the signals from only two million nerve fibers coming from the optic nerves of two eyes. This makes possible the massively parallel simultaneous processing of the 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.

Figure 5.11 is derived from Livingston and Hubel's diagram (1988) that summarizes both the neural architecture and the features processed in this area of the brain. 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 early 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.

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 signal) via the opponent processing channel mechanisms discussed in Chapter 4
- Elements of local stereoscopic depth
- Elements of motion

Moreover, all these properties are extracted for each point in the visual field. In V1 and V2 and many other regions of the brain, neurons are arranged in the form of a spatial map of the retina. It is a highly distorted map, because the fovea is given more space than the periphery of vision. The receptive fields are smaller for cells that process information coming from the fovea than for cells that process information from peripheral regions of the visual field. Nevertheless, for every

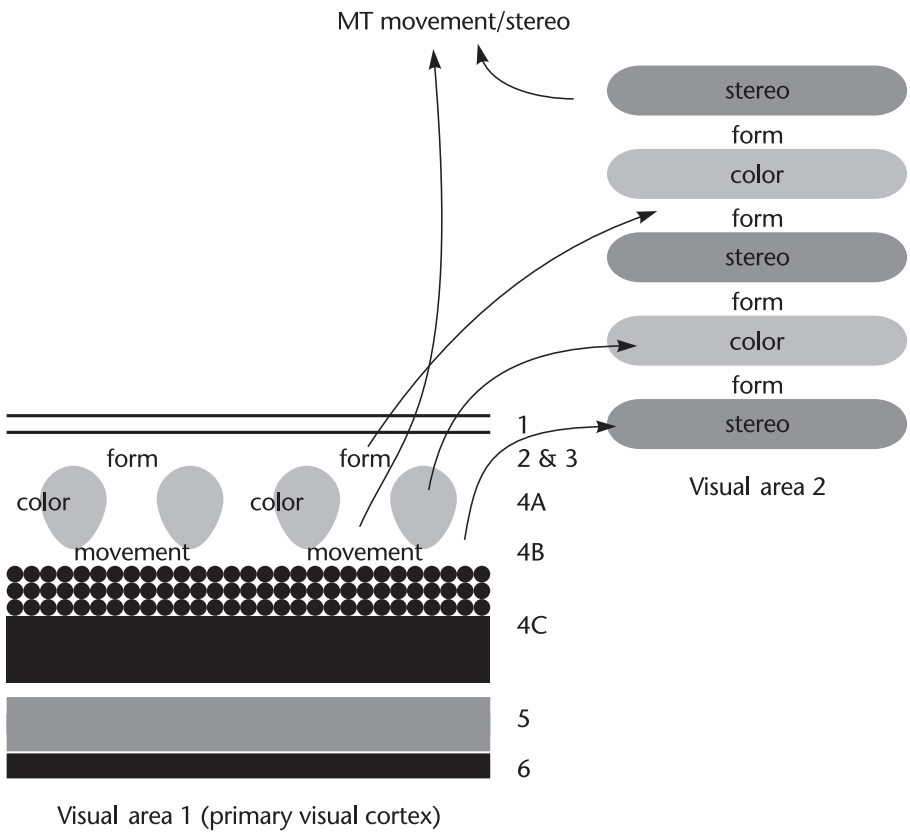


Figure 5.11 Architecture of primary visual areas. *Adapted from Livingston and Hubel (1988).*

point in V1, there is a corresponding area of the visual field in a topographic relationship (adjacency is preserved between areas). It is a massively parallel system in which, for each point in visual space, there are tuned filters for many different orientations, many different kinds of color information, many different directions and velocities of motion, and many different stereoscopic depths.

The Grapheme

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

We use the term *grapheme* to describe a graphical element that is primitive in *visual* terms, the visual equivalent of a phoneme. The basis of the grapheme concept is that the pattern that most efficiently excites a neuron in the visual system is exactly the pattern that the neuron is

tuned to detect (Ware and Knight, 1995). Thus, the most efficient grapheme is one that matches the receptive field properties of some class of neurons. An orientation detector will be excited most efficiently by a pattern whose light distribution is exactly the same as the sensitivity distribution of the cell. This is simply another way of saying that the detector is tuned to that particular pattern. 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 based on the receptive field properties of neurons should be rapidly detected and easily distinguished.

A number of assumptions are implicit in this account. They are worth examining critically. 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. 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 may be 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. Thus, there is a convergence of lines of evidence.

We also assume that *early-stage* neurons are particularly important in determining how distinct things seem. We know that at higher levels of processing in the visual cortex, receptive fields are found that are much more complex; they respond to patterns that appear to be composites of the simple receptive field patterns found at earlier stages. The evidence suggests that composite patterns analyzed further up the visual processing chain, are not, in general, processed as rapidly. It seems natural, then, to think of early-stage processing as forming the graphemes, and of later-stage processing as forming the "words," or objects, of perception.

Much of the preattentive processing work already discussed in this chapter can be regarded as providing experimental evidence of the nature of graphemes. The following sections apply the concept to the perception of visual texture and show how knowledge of early mechanisms enables us to create rules for textures that are visually distinct.

The Gabor Model and Texture in Visualization

A number of electrophysiological and psychophysical experiments show that visual areas 1 and 2 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*). These particular neurons are not color-coded; they respond to luminance changes only.

A simple mathematical model used widely to describe the receptive field properties of these neurons is the *Gabor function*. This function is illustrated in Figure 5.12. It consists of the product of a cosine wave grating and a gaussian. 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

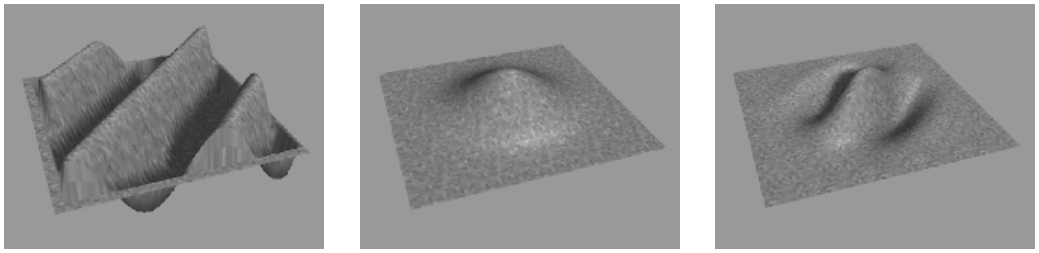


Figure 5.12 Gabor receptive field, composed of cosine and gaussian components. Multiply the cosine wave grating on the left by the gaussian envelope in the center to get the two-dimensional Gabor function shown on the right. This example has an excitatory center flanked by two inhibitory bars.

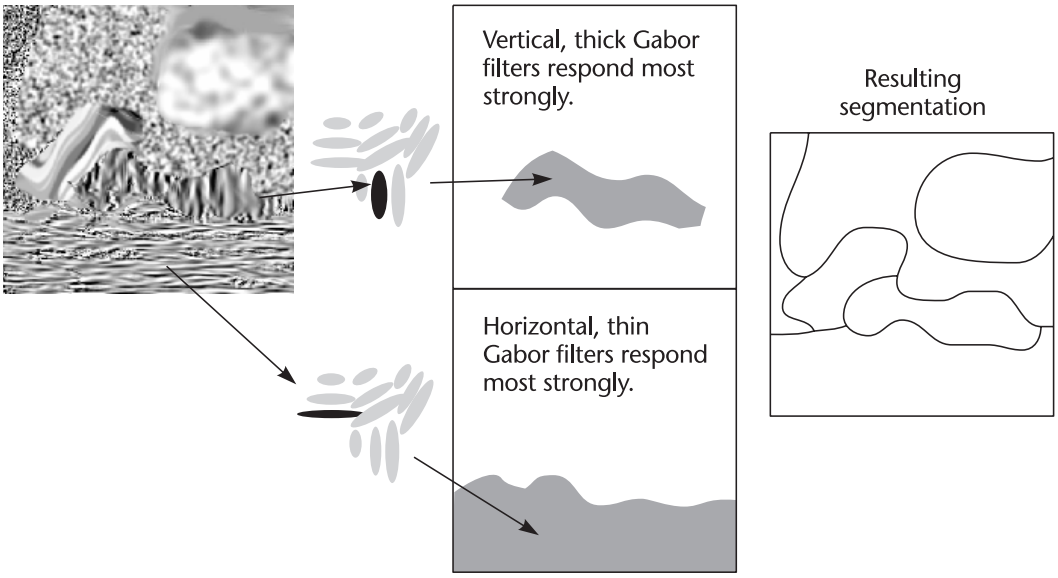


Figure 5.13 The texture segmentation model. Two-dimensional arrays of Gabor detectors filter every part of the image for all possible orientations and sizes. Areas exciting particular classes of detectors form the basis of visually distinct segments of the image.

inhibitory bars. The opposite kind of neuron also exists, with an inhibitory center and an excitatory surround.

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 function has two components, as illustrated in Figure 5.12: a cosine wave and a gaussian envelope. Multiply them together, and the result is a Gabor function. Mathematically, a Gabor function has the following form (simplified for ease of explanation):

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

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. 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.

Texture Segmentation

One way to apply the Gabor model is in understanding how the visual system *segments* the visual world into regions of distinct visual texture. Suppose we wish to understand how people perceptually differentiate types of vegetation based on the visual textures in a black-and-white satellite image. A model based on Gabor filters provides a good description of the way people perform this kind of texture segmentation task (Bovik et al., 1990; Malik and Perona, 1990).

The segmentation model is illustrated in Figure 5.13. It has three main stages. In the first stage, banks of Gabor filters respond strongly to regions of texture where particular spatial frequencies and orientations predominate. In a later stage, the output from this early stage is low-pass-filtered. (This is a kind of averaging process that creates regions, each having the same general characteristic. At the final stage, the boundaries are identified between regions with strongly dissimilar characteristics.) This model predicts that we will divide visual space into regions according to the predominant spatial frequency and orientation information. A region with large orientation and size differences will be the most differentiated. Also, regions can be differentiated based on the texture contrast. A low-contrast texture will be differentiated from a high-contrast texture with the same orientation and size components.

Tradeoffs in Information Density: An Uncertainty Principle

A famous vision researcher, Horace Barlow, developed a set of principles that have become influential in guiding our understanding of human perception. The second of these, called "the second dogma" (Barlow, 1972), 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. John Daugman (1984) showed mathematically that Gabor detectors satisfy the requirements of the Barlow dogma. They 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.

Daugman (1985) has also shown that a fundamental uncertainty principle is related to the perception of position, orientation, and size. Given a fixed number of detectors, resolution of size can be traded for resolution of orientation or position. We have shown that same principle applies to the synthesis of texture for data display (Ware and Knight, 1995). A gain in the ability to display orientation information precisely inevitably comes at the expense of precision in displaying size information. Given a constant density of data, orientation or size can be specified precisely, but not both.

Figure 5.14 illustrates this tradeoff, expressed by changing the shape and size of the gaussian multiplier function with the same sinusoidal grating. When the gaussian is large, the spatial frequency is specified quite precisely, as shown by the small image in the Fourier transform. When the gaussian is small, position is well specified but spatial frequency is not, as shown by the large image in the Fourier transform. The lower two rows of Figure 5.14 show how the gaussian envelope can be stretched to specify either the spatial frequency or the orientation more precisely. Although a full mathematical treatment of these effects is beyond the scope of this book, the main point is that there are fundamental limits and tradeoffs related to the ways texture can be used for information display. To restate them simply, large display glyphs can only show position imprecisely; precise orientation can be shown at the expense of precise size information, and both trade off against precision in position.

Texture Coding Information

If texture perception can be modeled and understood using the Gabor function as a model of a detector, the same model should be useful in *producing* easily distinguished textures for information display. The ideal grapheme for generating visual textures will be the Gabor function expressed as a luminance profile, as shown in Figure 5.15. A neuron with a Gabor receptive field will respond most strongly to a Gabor pattern with the same size and orientation. Therefore, textures based on Gabor primitives should be easy to distinguish.

Primary Perceptual Dimensions of Texture

A completely general Gabor model has parameters related to orientation, spatial frequency, contrast, and the size and shape of the gaussian envelope. However, in human neural receptive fields, the gaussian and cosine components tend to be coupled so that low-frequency cosine components have large gaussians and high-frequency cosine components have small gaussians (Caelli and Moraglia, 1985). This allows us to propose a simple three-parameter model for the perception and generation of texture.

Orientation *O*: The orientation of the cosine component

Scale *S*: The size = 1/(spatial frequency component)

Contrast *C*: An amplitude or contrast component

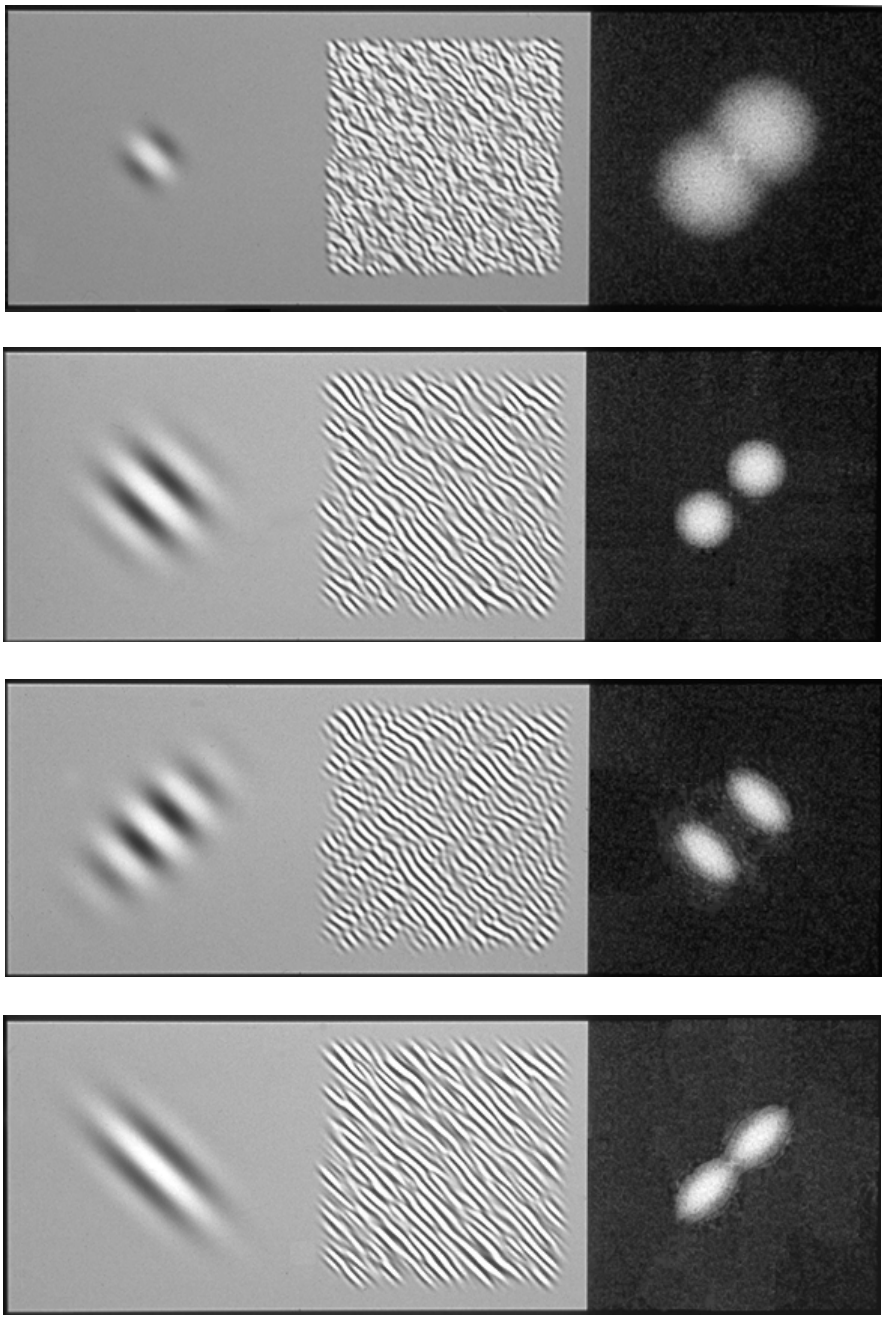


Figure 5.14 In the left-hand column, the same cosine pattern is paired with different gaussian multipliers. In the center column are textures created using each Gabor function by reducing the size by a factor of 5 and spattering it in the field. In the right-hand column are 2D Fourier transforms of the textures.

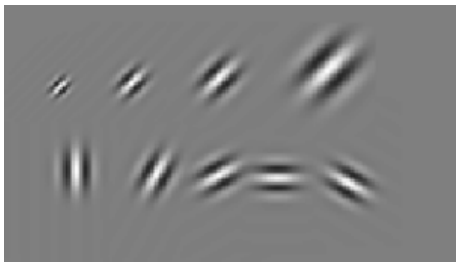


Figure 5.15 Gabor receptive fields shown as gray-scale images. Different sizes and orientations are represented for each part of the visual field.

Generation of Distinct Textures

With this simple model, it is straightforward to generate textures using Gabor functions as primitives. These textures can be varied in orientation, size (1/frequency), or contrast.

One method is to randomly splatter down Gabor functions whose orientation, size, and contrast have been determined by data values for the region in space where each splatter lands (Ware and Knight, 1995). When enough splatters have been accumulated in this way, we will have a continuous map that can represent up to three variables (a trivariate map). We can also map an additional variable to hue, producing a four-variable map.

Data value 1 → Orientation

Data value 2 → Size

Data value 3 → Contrast

Data value 4 → Hue

Figure 5.16 provides an example showing a magnetic field displayed using orientation and size manipulations. Color coding is added to the Gabor textures to illustrate field strength. A word of caution—Figure 5.16 illustrates a direct application of low-level visual theory, but it should not be taken as an optimal display. It is based on a feature-level model; to understand how to better show flow patterns, we need to move up the visual system and consider how patterns are formed from features. A more effective approach to vector field visualization, through pattern perception, is discussed in Chapter 6.

Note that textures need not be made of Gabor patterns for the method or the theory to work. It is only necessary that texture elements have a dominant orientation and spatial frequency. It is also important to note that the fundamental tradeoffs in our ability to represent spatial information using texture are independent of whether or not the Gabor model of texture perception is correct. To take a simple example, if we consider that texture elements, or *textons*, can be made from small graphical shapes representing data, the number of such shapes that can be drawn per unit area is inversely proportional to their size. The location of the packet of

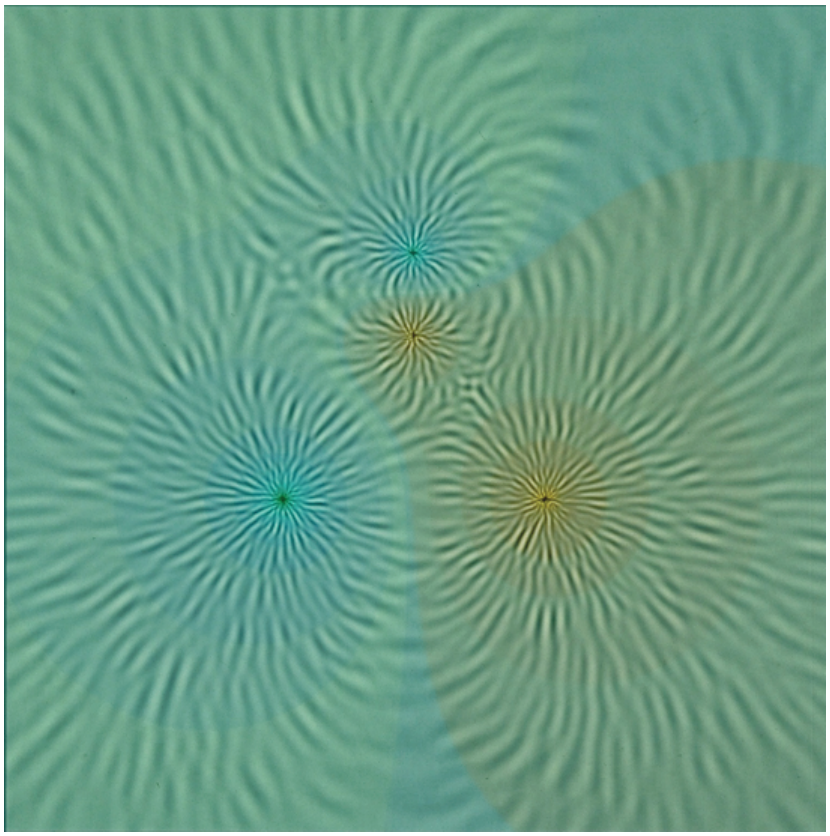


Figure 5.16 Magnetic field shown using Gabor textures.

information can be specified only to a precision determined by the size of the object representing that information.

Spatial-Frequency Channels, Orthogonality, and Maps

Sometimes we may wish to display many different kinds of information in a single map. For example, we might wish to show sea-surface temperature and sea-surface salinity at the same time. Naturally, we would prefer that the different sources of information do not interfere with one another. It would be unfortunate if regions of high salinity appeared to have a greater apparent temperature than they really have, due to visual crosstalk between the way we display temperature and the way we display salinity. Thus, our goal is to create display methods that are *perceptually independent*.

The concept of the *visual processing channel* can be taken directly from vision research and applied to the independence problem. We have already discussed the concept of color channels in Chapter 4. Here, the same idea is applied to spatial information. The idea is that information carried on one channel should not interfere with information displayed on another. It is

probably not the case that any of the perceptual channels we shall discuss are 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.

A given Gabor-type neuron is broadly tuned with respect to orientation and size. The half-width 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 four. 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 5.17.

Orientation tuning-in appears to be about ±30 degrees (Blake and Holopigan, 1985). Therefore, textures that differ from one another by more than 30 degrees in orientation will be easily distinguished.

These experimental results can be applied to problems in information display. For textured regions to be visually distinct, the dominant spatial frequencies should differ by at least a factor of 3 or 4, and the dominant orientations should differ by more than 30 degrees, all other factors (such as color) being equal. In general, the more displayed information differs in spatial frequency and in orientation, the more distinct that information will be. In practical applications, this means

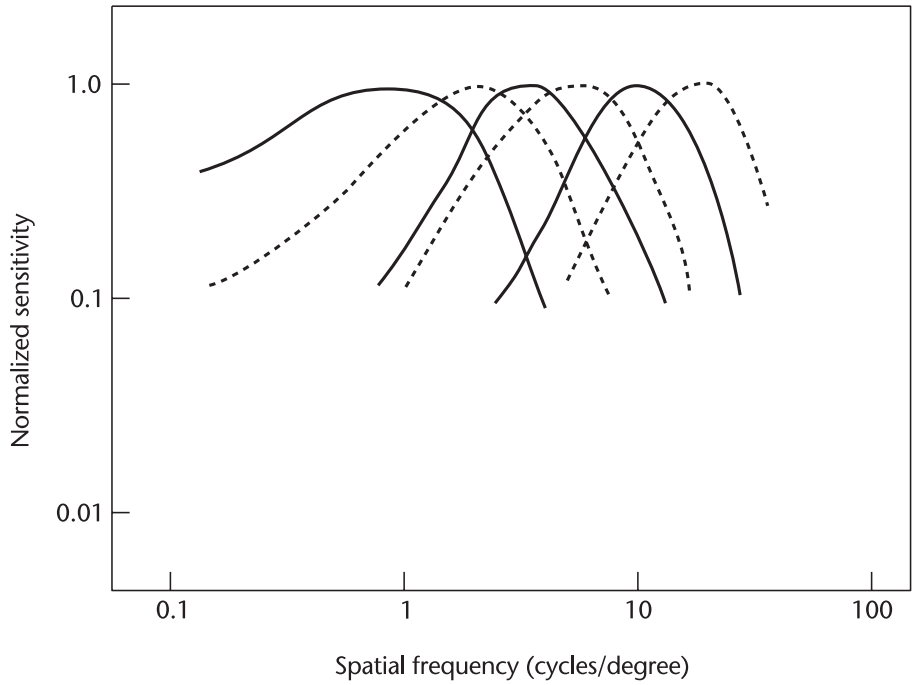


Figure 5.17 Wilson and Bergen spatial channels.

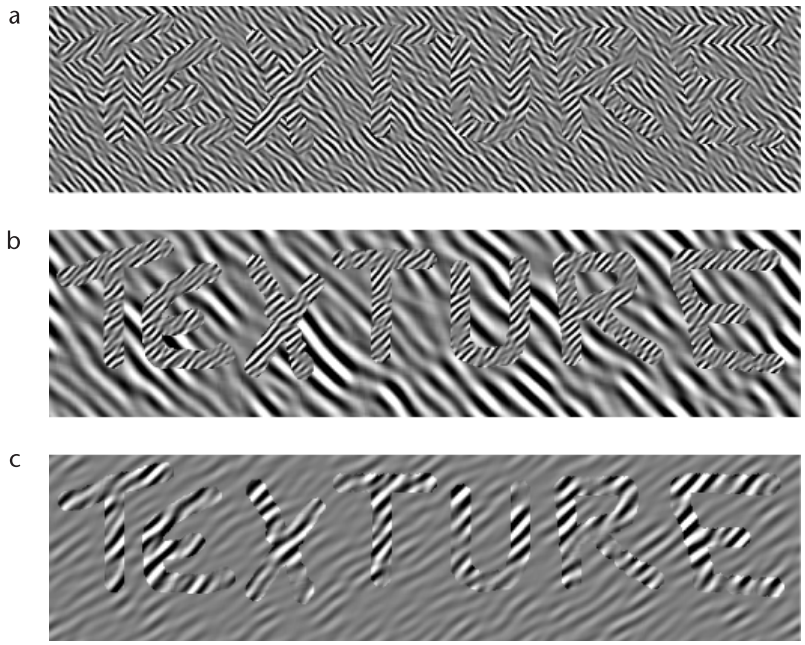


Figure 5.18 The word *TEXTURE* is visible only because of texture differences between the letters and the background; overall luminance is held constant. (a) Only texture orientation is altered. (b) Texture orientation and size are altered. (c) Texture contrast is altered.

that if we want different regions to be distinct because of their texture, the dominant orientations of the patterns should be made as different as possible. In Figure 5.18(a), only orientation is changed between different regions of the display, and although the word *TEXTURE* appears distinct from its background, it is weak. The difference appears much stronger when both the spatial frequency and the orientation differ between the figure and the background, as in Figure 5.18(b). The third way that textures can be made easy to distinguish is by changing the contrast, as illustrated in Figure 5.18(c).

Texture Resolution

The model of texture segmentation described previously predicts performance when people are asked to rapidly classify regions of a display. However, if we ask how small a difference people can *resolve,* we need a different model. When people are allowed to stare at two regions of a display for as long as they like, they can resolve far smaller differences than those perceived in brief presentations.

The *resolvable* size difference for a Gabor pattern is a size change of about 9% (Caelli et al., 1983). The resolvable orientation difference is about 5 degrees (Caelli and Bevan, 1983). These resolutions are much smaller than the channel-tuning functions would predict. This implies that higher-level mechanisms are present to sharpen up the output from individual receptors. The

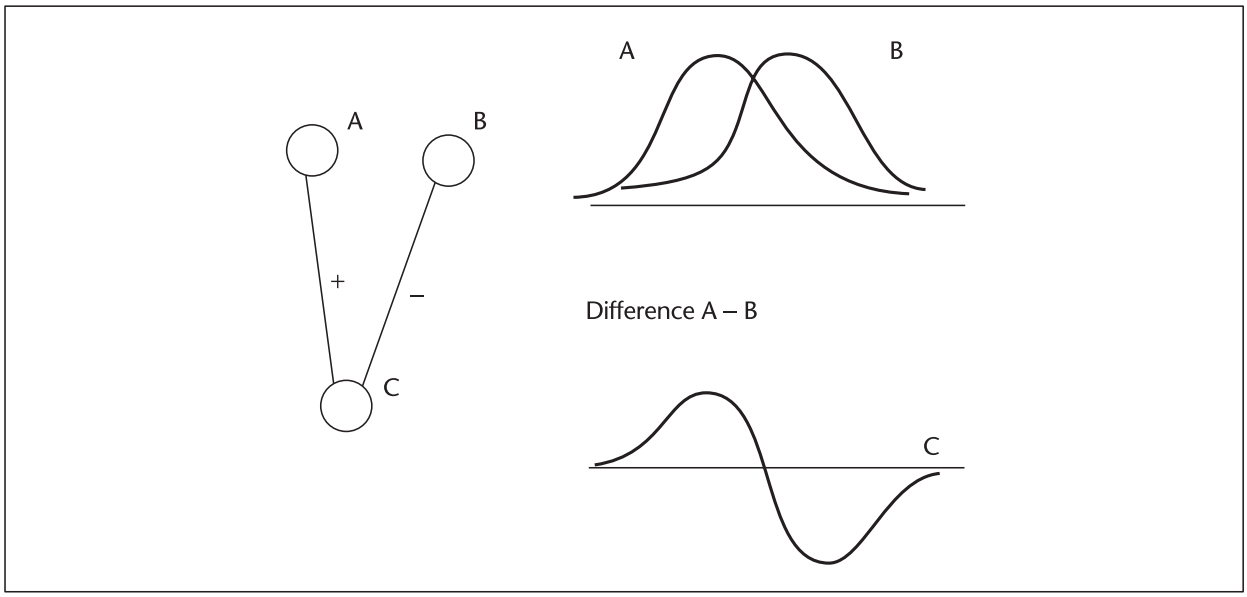


Figure 5.19 Differences between two signals are created by an excitatory and an inhibitory connection.

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 in neural systems; it appears in color systems, edge detection, and heading detection (for navigation). Figure 5.19 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.

Texture Contrast Effects

Textures can appear distorted because of contrast effects, just like the luminance contrast illusions that were described in Chapter 3. Thus, a given texture on a coarsely textured background will appear finer than the same texture on a finely textured background. This phenomenon is illustrated in Figure 5.20. The effect is predicted by higher-order inhibitory connections. It will cause errors in reading data that is mapped to texture element size. Texture orientation can cause contrast illusions in orientation, and this, too, may cause misperception of data. See Figure 5.21.

Other Dimensions of Visual Texture

Although there is considerable evidence to suggest that orientation, size, and contrast are the three dominant dimensions of visual texture, it is clear that the world of texture is much richer

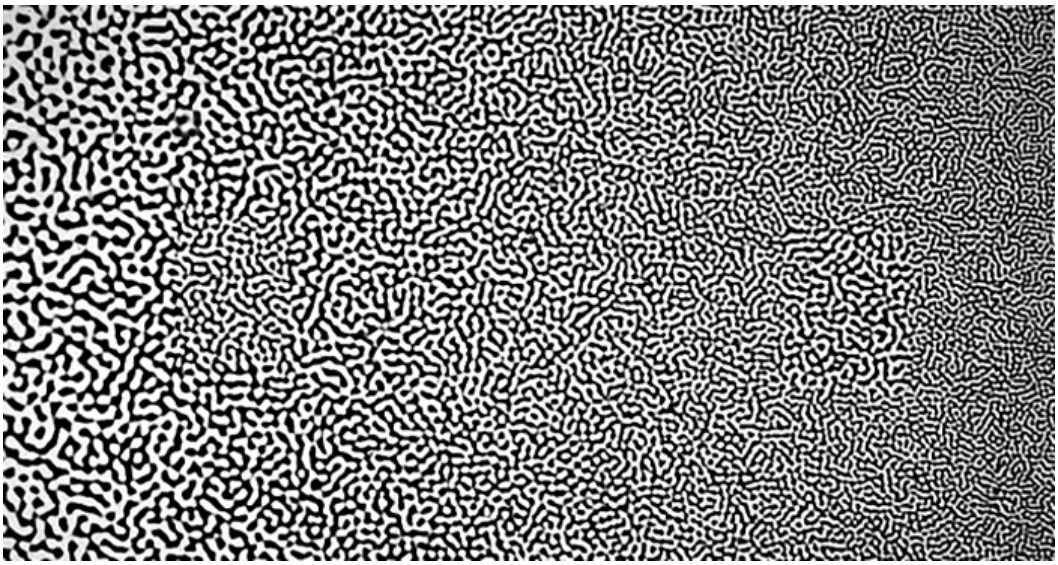


Figure 5.20 Texture contrast effect. The two patches left of center and right of center have the same texture granularity, but texture contrast makes them appear different.

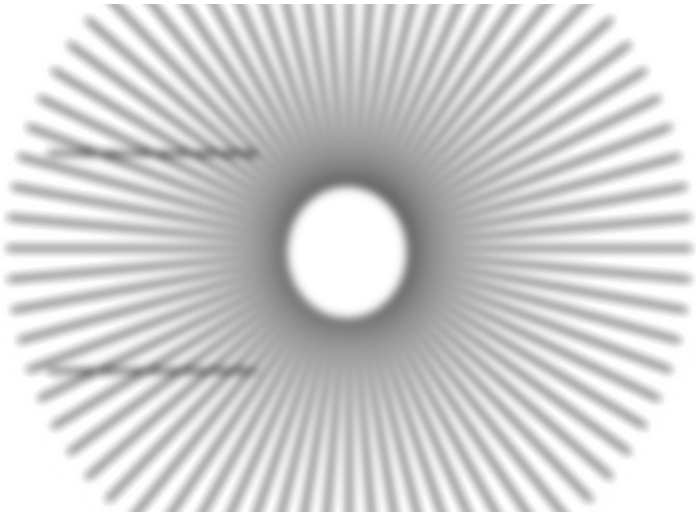


Figure 5.21 The two horizontal lines are parallel, but they appear tilted due to orientation contrast.

than this. The dimensionality of visual texture is very high, as a visual examination of the world around us attests. Think of the textures of wood, brick, stone, fur, leather, and other natural materials. One of the important additional texture dimensions is certainly randomness (Liu and Picard, 1994). Textures that are regular have a very different quality from random ones.

Texture Field Displays

We would do well to learn to use texture more effectively in information displays. The world of visual texture is arguably as rich and expressive as the world of color. Users of GISs commonly require the display of many overlapping variables on the same map, such as geological information, vegetation type, surface topography, and magnetic anomalies. In light of the theory of parallel feature processing, we are now in a position to say something about various solutions that apply visual texture to such problems.

The Exvis tool (Pickett and Grinstein, 1988) mapped data values to various attributes of stick-figure icons such as those shown in Figure 5.22. This package had many display options, including changing the angles of line segments relative to each other, or relative to a reference orientation, and changing the line segment widths. These glyphs could then be displayed in a dense field over a plane producing a visual texture. Although the Exvis developers implemented the capability to map data to icon colors, they worked mostly with angles (Pickett et al., 1995).

What does early visual processing tell us about the Exvis glyph? The theory of visual texture segmentation based on low-level Gabor detectors suggests a problem. With the Exvis glyph, multiple segments of a single glyph can have the same or similar orientations, although each represents a different data dimension. These line segments will be visually confounded when the glyphs are densely displayed, ensuring that unrelated aspects of the underlying data will be visually confounded. Because the orientation tuning of V1 neurons indicates that glyph element orientations should be separated by at least 30 degrees, and because a line-oriented segment will be confused with an identical segment rotated through 180 degrees, fewer than six orientations can be rapidly distinguished.

Weigle et al. (2000) developed a technique called *oriented sliver textures* specifically designed to take advantage of the parallel processing of orientation information. Each variable in a multivariate map was mapped to a 2D array of slivers where all the slivers had the same orientation. Differently oriented 2D sliver arrays were produced for each variable. The values of each scalar map were shown by controlling the amount of contrast between the sliver and the background. Combining all of the sliver fields produced the visualization illustrated in Figure 5.23.

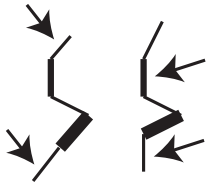


Figure 5.22 The Exvis data glyph used to form visual textures. Different variables are mapped to the angle between line segments and their thickness.

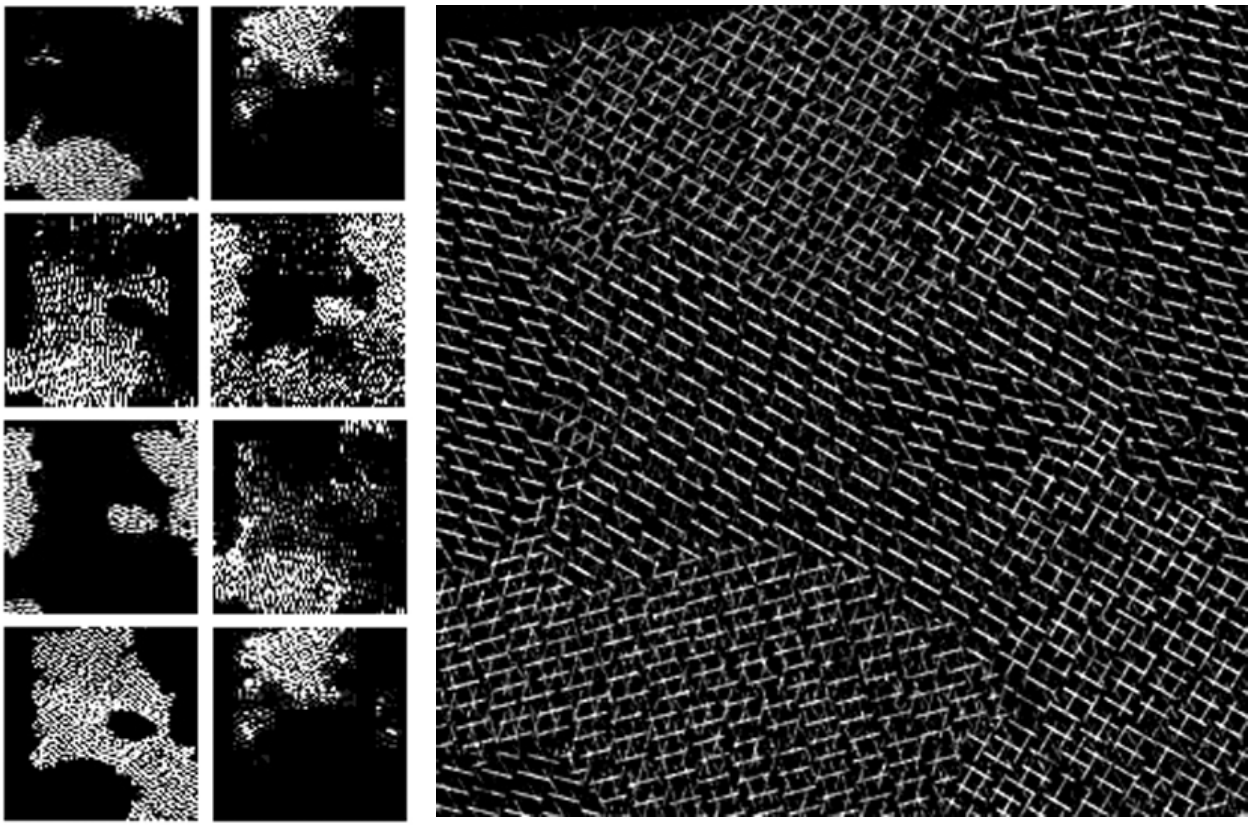


Figure 5.23 The sliver plot of Weigle et al. (2000). Each of the variables shown in the thumbnail patterns in the left part of the above figure is mapped to a differently oriented sliver field. *Courtesy of Chris Weigle.*

The right-hand part of this figure shows the combination of the eight variables illustrated in the thumbnail patterns shown on the left. Weigle et al. conducted a study showing that if slivers were oriented at least 15 degrees from surrounding regions, they stood out clearly. However, the experiment was only carried out with a single sliver at each location (unlike Figure 5.23). To judge the effectiveness of the sliver plot for yourself, try looking for each of the thumbnail patterns in the larger combined plot. The fact that many of the patterns cannot easily be seen suggests that the technique is not effective for so many variables. The tuning of orientation-sensitive cells suggests that slivers should be at least 30 degrees apart to be clearly readable (Blake and Holopigan, 1985), perhaps more, but in Figure 5.23 some differ by only 15 degrees.

Figure 5.24 shows another sliver plot with only three orientations. This adds a colored background and also uses slivers having both positive and negative contrast with the background. It is easier to see the different patterns in this example.

Two other examples of high-dimensional data display from Laidlaw and his collaborators (Laidlaw et al. 1998) (Figures 5.25 and 5.26) were created using a very different design strategy.

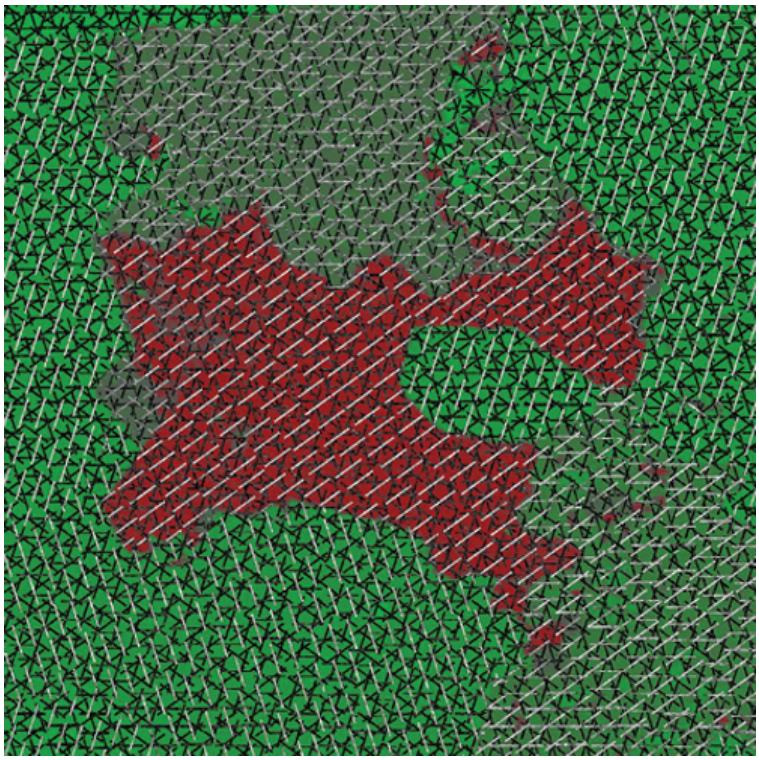


Figure 5.24 Another example of a sliver plot. Three variables are mapped to three differently oriented slivers. A fourth variable is mapped to the background color. *Courtesy of Chris Weigle.*

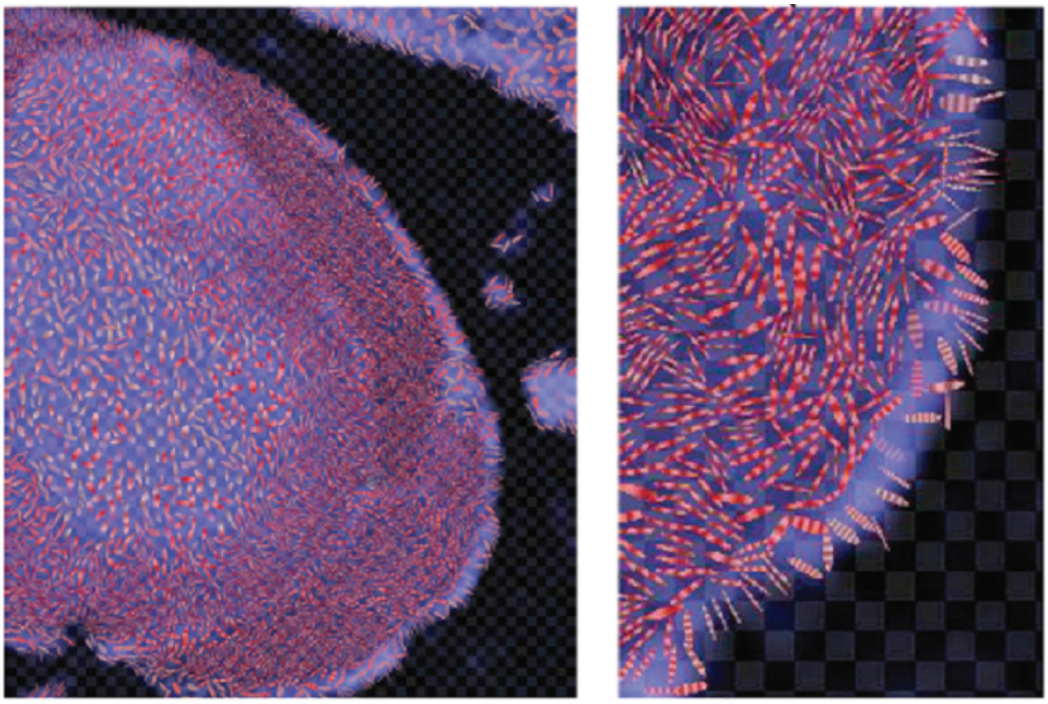


Figure 5.25 A cross section of a mouse spinal column. Seven variables are shown at each location. Part of the image is enlarged on the right. See text for description. *Courtesy of David Laidlaw.*

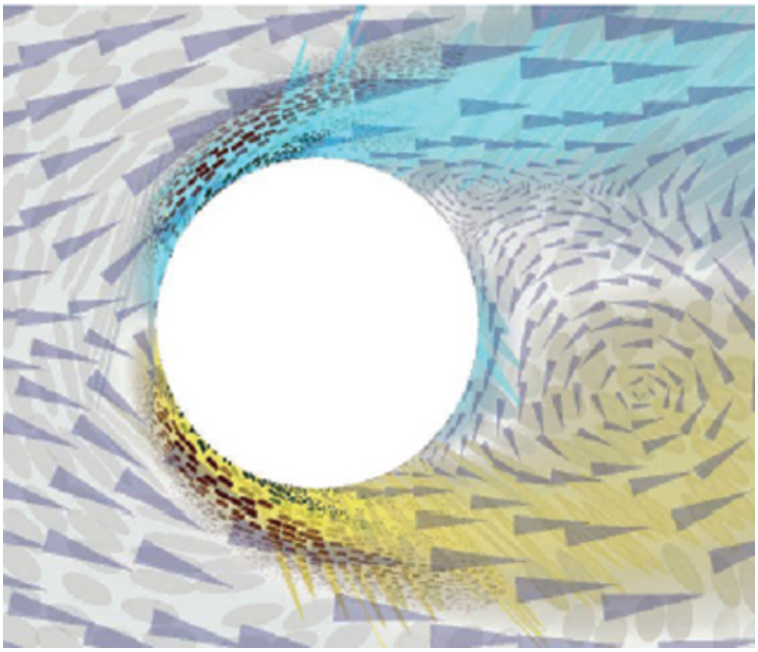


Figure 5.26 A flow visualization showing six variables relating to fluid flow around a cylinder. *Courtesy of David Laidlaw.*

Rather than attempting to create a simple general technique (like slivers), both figures were hand-crafted in a collaboration between the scientist and the designer. Figure 5.25 shows a cross section of a mouse spinal column. The data has seven values at each location in the image. The image is built up in layers: image intensity, sampling rate determines the grid, elliptical shapes show the in-plane component of principal diffusion and anisotropy, texture on the ellipses shows absolute diffusion rate.

The image in Figure 5.26 is a flow visualization. It displays six variables relating to the flow pattern of a fluid around a cylinder. These values are 2D velocity (two values are mapped to arrow direction and area), vorticity (one value is mapped to color and texture on ellipses), and deformation rate tensor (three values are mapped to shape and orientation of ellipses).

Without specific knowledge of mouse physiology or fluid dynamics, it is impossible to judge the success of these examples. Nevertheless, they provide a vivid commentary on the tradeoffs involved in trying to display high-dimensional multivariate maps. The first point to be made is that none of the preceding three examples (Figures 5.24, 5.25, and 5.26) shows much detail, and there is a good reason for this. We only have one luminance channel, and luminance variation is the only way of displaying detailed information. If we choose to use texture (or any kind of glyph field), we inevitably sacrifice the ability to show detail, because to be clear each glyph element must be displayed using luminance contrast. Larger glyphs mean that less detail can be shown.

There are also tradeoffs when displaying orientation. It may be only possible to display a single orientation clearly at each point in space for the purpose of showing flow patterns. Figure 5.26 suggests that if we need to show differently oriented glyphs in the same region, the glyphs must be widely spaced. This reduces the data density further. Also, Figure 5.26 suggests that the colors of different glyph layers must be very carefully chosen to be dissimilar. This, in turn, severely restricts how color-coding can be used on individual glyphs. In Figure 5.25, each of the elliptical glyphs is textured to display an additional variable. However, the texture striations are at right angles to the ellipse major axes. This camouflages the glyphs, making their orientation more difficult to see. The use of texture will inevitably tend to camouflage glyph shape; if the textures are oriented, the problem will be worse. In general, the more similar the spatial frequencies of the different pattern components, the more likely they are to disrupt one another visually.

The complexity of the design tradeoffs suggests that the problem of creating complex visualizations will be more of a craft than a science for quite some time. The problem is too difficult for automatic assignments of data maps to graphical attributes to be successful. Still, the designer needs to be aware of the perceptual tradeoffs in order to make informed decisions about the best choice of glyph size, shape, and color distribution.

It is also worth pointing out that there are some perceptual dimensions that may be used in addition to color, shape, and texture. In some cases, it is helpful to use stereoscopic depth and motion in displaying multidimensional data. Stereoscopic depth, especially if used with a high-resolution display, can undoubtedly help us perceptually segment data layers. So can motion. Making all of the points in a data layer move coherently, even by a small amount, may make it possible to visually attend to either the static layer or the moving layer (as shown by the possibility of preattentive conjunction search with motion).

Glyphs and Multivariate Discrete Data

In the previous section, we saw how texture could be used to represent continuous map data. In Chapter 4, it was shown that color could be used in a similar way. However, sometimes multivariate *discrete* data is the subject of interest. For example, a marketing specialist may have data for every person in a particular geographical area, including estimates of income, educational level, employment category, and location of residence. The marketer would like to see each person on a map in such a way that the concentrations of individuals with particular sets of attributes can easily be seen. In this way, neighborhoods to be blanketed with flyers might be selected most effectively.

To create a glyph, multiple data attributes are mapped in a systematic way to show the different aspects of the appearance of the graphical object. In the aforementioned marketing example, income might be mapped to the glyph's size, education level to its color, employment category to its shape, and geographic location to the x,y location where the glyph is plotted. All the previously discussed results relating to preattentive detection of size, orientation, and color-coding of data apply to the design of glyphs.

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. Will the color-coding scheme interfere with our perception of glyph size and therefore distort perceived income level? 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 variables. Very little work has been done on interactions among three or more display variables.

Restricted Classification Tasks

In restricted classification tasks, observers are shown sets of three glyphs that are constructed according to the diagram shown in Figure 5.27. Two of the glyphs (A and B) are made the same on one variable. 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, A 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 creates an integral perception of shape. Thus, in Figure 5.28(a), the ellipses B and C appear to be more similar to each other than to the circle A, even though the width of B matches the width of A. If 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 gray value are separable. Thus, in Figure 5.28(b), either the two gray shapes or the two elliptical shapes will be categorized together. With separable dimensions, it is easy to attend to one dimension or the other.

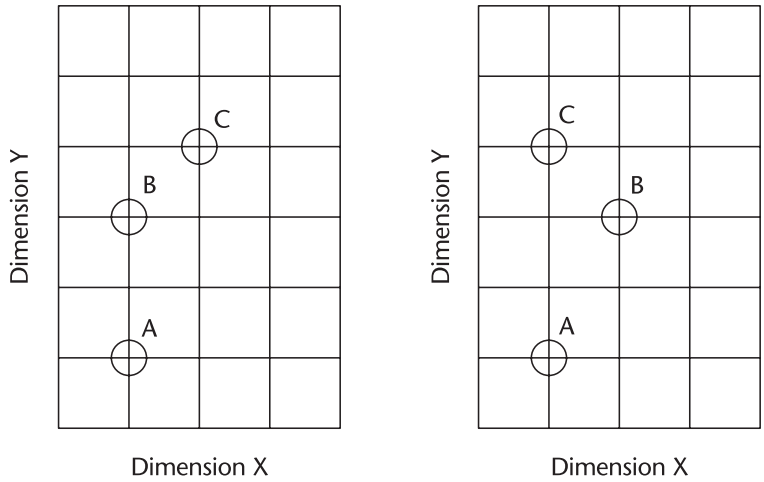


Figure 5.27 When we are considering integral and separable visual dimensions, it is useful to consider a space defined by two display dimensions. One might be size; the other might be color. Or one might be hue and the other might be saturation, both defined in color space.

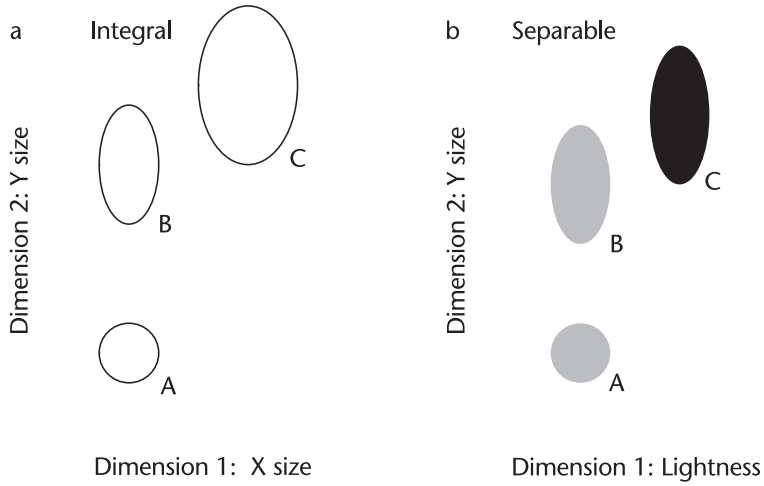


Figure 5.28 (a) The width and height of an ellipse are perceived integrally; therefore, B and C are perceived as more similar. (b) The gray value and the height of an ellipse are perceived as separable; therefore, A and B, which have identical lightness, are perceived as more similar.

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 classify visual patterns rapidly 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

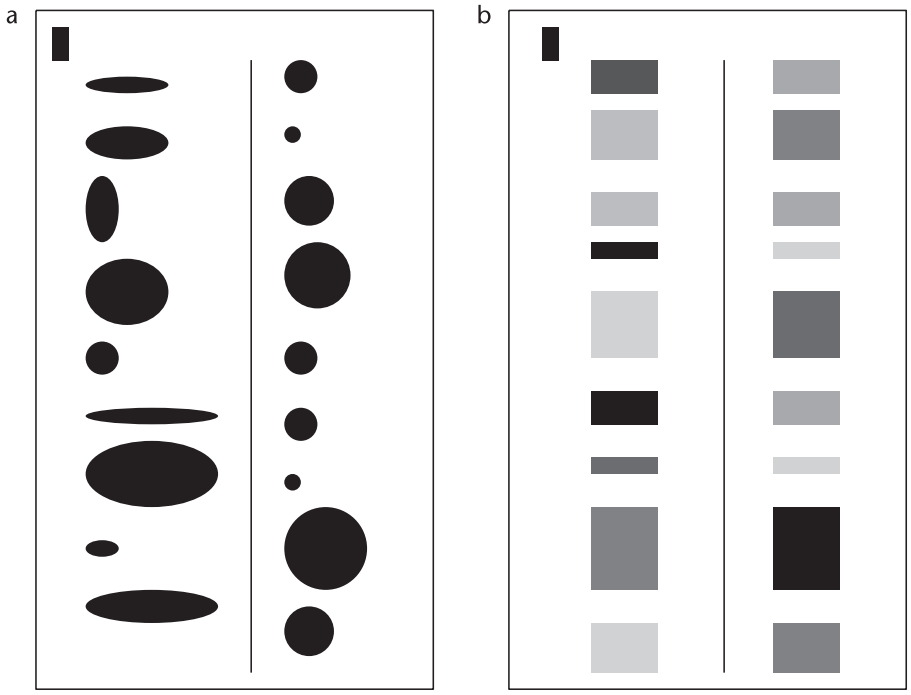


Figure 5.29 Patterns for a speeded classification task. Subjects are required to respond positively only to those glyphs that have the same height as the black bar in the upper-left corner. (a) Integral dimensions. In the first column, a second integral dimension is randomly coded by horizontal size (interference condition). In the second column, width information is redundantly coded with height information. (b) Separable dimension. In the first column, gray information is not correlated with height. In the second column, gray level is a redundant code.

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 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 5.29 gives examples of the kinds of patterns that are used in experiments.

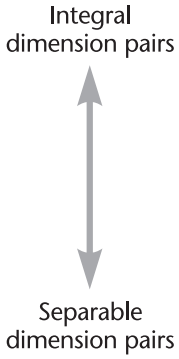
The lessons to be learned from integral—separable dimension experiments are straightforwardly applied to cases in which each data entity has two attributes. If we want people to respond holistically to a combination of two variables, using integral dimensions will be better. If we want people to respond analytically, making judgments on the basis of one variable or the other, using separable dimensions will be better.

Integral–Separable Dimension Pairs

The preceding analysis has 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. There is always some interference between different data values presented using different graphical attributes of a single visual object, even between the most separable dimensions. Likewise, the most integral dimensions 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 instance, orange or pink. Indeed, the original experimental evidence for opponent color channels was based on analytic judgments of exactly this type (Hurvich, 1981).

Figure 5.30 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 the most separable dimensions. Other possible display dimensions are not represented, because of too little evidence for blue and stereoscopic depth. However, it seems likely that stereoscopic depth is quite separable from other dimensions if only two depth layers are involved. The most separable way of coding information, as indicated at the bottom of the list, is to use spatial position to code one of the data dimensions and to use size, shape, or color to code the other. This is exactly what is done in a bar chart in which each bar represents a single value. Figure 5.31 illustrates some of the dimension pairs.

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. Eventually, it is to be expected that a more complete body of theory will emerge to account for the ways in which different kinds of visual information are combined. The beauty of the integral—separable distinction lies in its simplicity as a design guideline.



red–green	yellow–blue
red–green	black–white
shape height	shape width
shape	size
color	size
direction of motion	shape
color	shape
color	direction of motion
x,y position	size, shape, or color

Figure 5.30 This table lists some of the display dimension pairs ranked in order from highly integral to highly separable.

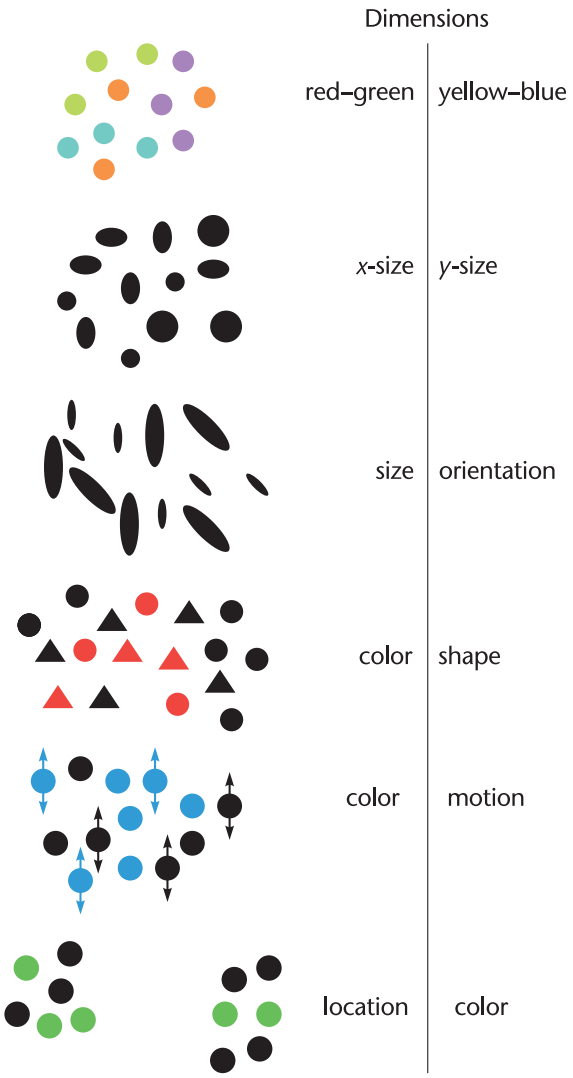


Figure 5.31 Examples of glyphs coded according to two display attributes. At the top are more integral coding pairs. At the bottom are more separable coding pairs.

Monotonicity of Visual Attributes

Some visual qualities increase continuously, like size, brightness, or the up direction, 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 direction that we associate with increasing value. For example, in a 3D data space, the up direction is defined by gravity, and using *up* to represent a greater quantity of some variable will be readily interpreted. The axis representing direction *toward* and *away from* the viewpoint is similarly well defined, 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. For representing simple quantity, a mapping to any of the following attributes will be effective: size, lightness (on a dark background), darkness (on a light background), vividness (higher saturation) of color, or vertical height above the ground plane. For each of these, an inverse mapping will lead to confusion.

Multidimensional Discrete Data

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 the previous chapter. 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 for data exploration. We hope 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.

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 dimension to the graphical attributes of the glyph. The work on preattentive processing, early visual processing, and integral and separable dimensions suggests that a rather limited set of visual attributes is available to us if we want to understand the values rapidly. Figure 5.32 is a list of 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 to make the texture visible. Blink coding will certainly interfere with motion coding. Overall, we will probably be fortunate to display eight 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, no more than eight 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

Visual variable	Dimensionality	Comment
Spatial position of glyph	3 dimensions: *X, Y, Z.*	
Color of glyph	3 dimensions: defined by color opponent theory.	Luminance contrast is needed to specify all other graphical attributes.
Shape	2–3? Dimensions unknown.	The dimensions of shape that can be rapidly processed are unknown. However, evidence suggests that size and degree of elongation are two primary ones.
Orientation	3 dimensions: corresponding to orientation about each of the primary axes.	Orientation is not independent of shape. One object can have rotation symmetry with another.
Surface texture	3 dimensions: orientation, size, and contrast.	Not independent of shape or orientation. Uses up one color dimension.
Motion coding	2–3? Dimensions largely unknown, but phase may be useful.	
Blink coding: The glyph blinks on and off at some rate.	1 dimension.	Motion and blink coding are highly interdependent.

Figure 5.32 Graphical attributes that may be used in glyph design.

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.

Stars, Whiskers, and Other Glyphs

There is a family of glyph designs for multidimensional discrete data displays that is interesting to analyze from a perception perspective. In the whisker plot, each data value is represented by a line segment radiating out from a central point, as shown in Figure 5.33(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 5.33(b). In general, it is better to use a very small number of orientations, perhaps only three, for really rapid classification of glyphs, as shown in Figure 5.34(b). It may be possible to increase the number of rapidly distinguishable orientations by inverting the luminance polarity of half of the bars, as in Figure 5.34(a). Color and position in space can be used to display other data dimensions. If we map

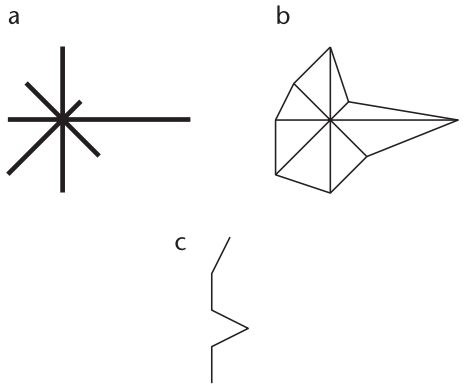


Figure 5.33 Three glyph designs: (a) The whisker or fan plot. (b) A star plot. (c) An Exvis stick icon.

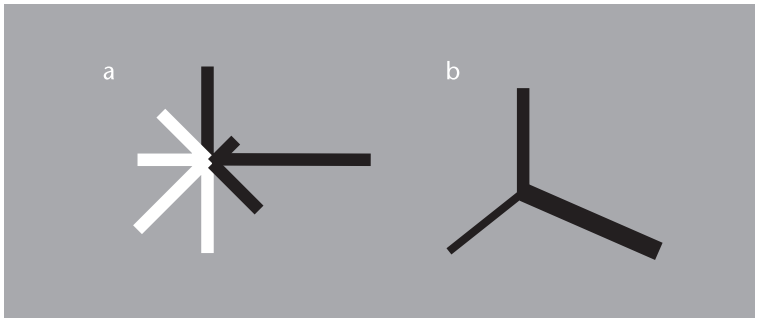


Figure 5.34 (a) It may be possible to increase the number of distinct orientations in a glyph display by changing the luminance polarity of half the line segments. (b) Changing the widths as well as the lengths of segments may also be effective.

three data dimensions to the position of each glyph and two dimensions to the color of the glyph, we can represent eight-dimensional data clearly and effectively. It may also be useful to change the amount of "energy" in glyph segments by altering the line width as well as the length of the line.

When large numbers of glyphs are present in a display, the glyph field becomes a texture field and the theory discussed earlier will apply.

Conclusion

This chapter has provided an introduction to the early stages of vision, in which literally 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 noticed easily. Understanding such preattentive processes is the key to designing elements of displays that must be rapidly attended to. Making an icon or a symbol significantly different from its surroundings on one of the preattentive dimensions ensures that it can be detected by a viewer without effort and at high speed.

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. Here is a short summary of the key lessons we have learned from low-level vision:

- Low-level channels tell us about coding dimensions. We can usefully consider color, elements of form (orientation, size), position, simple motion, and stereoscopic depth as separate channels.
- 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 or a red symbol will stand out.
- 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 hard 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 data dimension. Mapping one variable to color and another to glyph orientation will make them independently readable. If we map one variable to X-direction size and another to Y-direction size, they will be read more holistically. If we have a set of symbols that are hard 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.

Unfortunately, there are no universal rules for mapping multiattribute data to glyphs. The simple techniques, such as star plots, do not allow us to interpret the data rapidly, because we have mapped too much information to line segments having similar orientations that interfere visually with each other. 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.

In this chapter, we have dealt mainly with how attention is directed within a single fixation of the eye. Attention is also central in controlling eye movements and is a fundamental concept in the processes of visual thinking. We revisit the topic of attention in Chapter 11, which explores how we solve problems through visual thinking.

We have arrived at a transition point in this book. To this point, we have discussed mostly the massively parallel processing of low-level features of early vision and the elementary coding of information. We now turn our attention to the way the brain extracts a few complex objects from elemental information and subjects them to sophisticated analysis. We will also discuss how the brain finds elaborate patterns in data, and eventually we will look at the ways in which information should be integrated and displayed for solving complex problems.

CHAPTER 6

Static and Moving Patterns

Data mining is about finding patterns that were previously unknown or that depart from the norm. The stock-market analyst looks for any pattern of variables that may predict a future change in price or earnings. The marketing analyst is interested in perceiving trends and patterns in a customer database. When we look for patterns, we are making visual queries that are key to visual thinking. Sometimes the queries are vague; we are on the lookout for a variety of structures in the data. Sometimes they are precise, as when we look for a positive trend in a graph. In data exploration, seeing a pattern can often lead to a key insight, and this is the most compelling reason for visualization.

What does it take for us to see a group? How can 2D space be divided into perceptually distinct regions? Under what conditions are two patterns recognized as similar? What constitutes a visual connection between objects? These are some of the perceptual questions addressed in this chapter. The answers are central to visualization, because most data displays are two-dimensional and pattern perception deals with the extraction of structure from 2D space.

Consider again our three-stage model of perception (illustrated in Figure 6.1). At the early stages of feature abstraction, the visual image is analyzed in terms of primitive elements of form, motion, color, and stereoscopic depth. At the next 2D pattern perception stage, the contours are discovered and the visual world is segmented into distinct regions, based on texture, color, motion, and contour. Next, the structures of objects and scenes are discovered, using information about the connections between component parts, shape-from-shading information, and so on. Pattern perception can be thought of as a set of mostly 2D processes occurring between feature analysis and full object perception, although aspects of 3D space perception, such as stereoscopic depth and structure-from-motion, can be considered particular kinds of pattern perception. Finally, objects and significant patterns are pulled out by attentional processes to meet the needs of the task at hand.