1

Basic Concepts in Display Systems

1.1 Introduction

Regardless of the type of display used, its size, or the application in question, there are some concepts that are common to just about any display system. In the most general usage of the phrase, "display system" can be taken to mean any system through which information is conveyed to people through visual means. A book, a painting, or a sign could all be considered "display systems" in this sense, although admittedly they are not the sort of things with which this book will primarily be concerned.

But even these simpler examples of "displays" share some basic properties and concepts with the most advanced electronic display now being developed. If we use the above broad definition, and use the convenient term *image* to refer to any information which is being conveyed visually, then several of these basic concepts should become readily apparent.

1.1.1 Basic components of a display system

Remaining within the broad definition of a *display system,* again "a system through which information is conveyed to people through visual means," we could further divide this system into several basic elements. Any practical display system would include, in some form or another, all of the following (Figure 1-1):

1. An image source. This might be a real object, as viewed by an image transducer such as a camera. It might be a computer program (along with the hardware on which that program is running), as in the case of the completely synthetic images seen in many movies. It might even be information which is originally not in viewable form; an example might

be spoken words transcribed to text (in which case other factors, such as the choice of font, etc., also could be considered a part of the "image source").

2. Image processing and rendering. Almost always, the basic information from the source must be processed in some manner before being delivered to the rest of the system, and then must be put into a form which is suitable for the intended display device. An example of basic image processing is the "gamma correction" performed as a basic part of the broadcast television system. The computer graphics (CG) field has given us the term "rendering" to refer to the process of finally changing the image information into the form required by the display; in the case of CG, "drawing" the image in nearly display-ready form into a *frame buffer*.
3. Image storage, compositing, and transmission. Within many, if not most, systems, there are provisions for the storage of the image information prior to display. This can range from the combination of digital memory and video tape storage common in the television industry, to the purely digital frame buffer of the computer graphics system. The image storage portion of a display system (in this broad sense of the term) may be used simply to delay the delivery of the image information to the viewer, but it may also be an important factor in the image processing and rendering step, and shared by that portion of the system. Between the two, the image storage subsystem is also quite often that point at which information from multiple sources is composited into the final single image which will be seen by the viewer. Lastly, the composited image must be delivered to the viewer via a transmission channel suited to the application in question.
4. The display itself. In the context of this particular discussion, also included within the display subsystem are those portions of the complete system which are used to translate the image information between the format in which it is stored (or in which it has been transmitted so far) and the format used by the display device proper, should such a step be required. For example, computer systems will most commonly store image information in digital form, which often must then be converted into an analog video signal for use by the display.
5. The viewer. It may seem odd to be including the viewer in a discussion of a display system, which might normally be assumed to be simply a collection of electronic equipment or some other inanimate construction. The point here is that the viewer is in reality the single most important factor in determining the performance requirements and other factors that define the rest of the system. It is pointless, for instance, to construct a system which could provide detail or differences in color far beyond the capability of human vision to resolve. Some display systems are even defined completely for the specific needs of a particular class of viewer, as in the case of those which aid the visually impaired. In any case, it is of utmost importance to keep in mind that any display system ultimately relies on, and is limited by, the needs and limitations of human vision. Display systems are a *human* interface.

The basic notion brought up in that last section is worth repeating for the overall system. It makes little sense to provide capabilities in any one part of the system that cannot be used by the whole. Conversely, a limit in any of these subsystems is a limit on the whole. Most often, unless it is anticipated that a given subsystem will be used with other "blocks" of varying performance, it is best to attempt to achieve a balance between all parts of the complete system. And again, in the case of display systems specifically, it is ultimately the capabilities of the viewer that will define the requirements for the system.

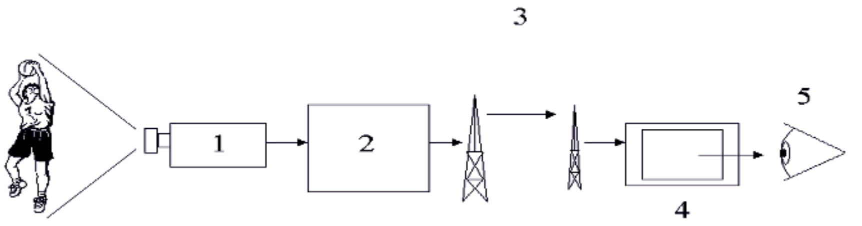

Figure 1-1 The elements of any imaging system include: (1) the image source, in this case a real object as seen by an image capture device (a camera); (2) image processing; (3) image storage and transmission; (4) image display; and (5) last but certainly not least, the viewer. As the entire purpose of an imaging system is to deliver visual information to the viewer, the viewer becomes the most important factor in the entire system, and dictates the required performance of the rest of the chain.

1.2 Imaging Concepts

The previous section defined a display system as one that is intended to convey image information to a viewer, but did not say anything about that image information itself. In what form should we expect it to be? How much information will there be, and at what rate must it be conveyed? What do we mean by "image" in the first place?

The most common-sense answer to the last question is that "image" simply means "what people see". We will look more deeply into just how human vision operates in a later chapter, but at this point this simple definition will serve as the basis for examining some concepts from imaging science. Humans see the world in color, and can discern motion, apparently at a fairly high level of resolution, and in three dimensions – how does this ability define that which is captured or produced through electronic or other means for human viewing?

First, while humans *do* possess visual systems that provide "three-dimensional" information, we should recognize that this comes from having two "image receptors" (eyes), each of which actually only captures a two-dimensional view of the world. We gain information about the third dimension through comparing these two views. So our basic working definition of *image* might be "a two-dimensional visual representation", using the word "visual" to hide a multitude of issues regarding the limitations of vision. This agrees with the common-sense definition as well: a picture on a wall is an image, that which appears on a television screen is an image, and the contents of a page of a book is an image.

More importantly, we draw a distinction between *image* and *reality*; the light reflected from an object, and captured by a lens, may form an image of that object on a sheet of paper – but the image is not the real object, and does not contain all of the information about that object. It contains only the information relating to the appearance of the object – "that which we see" – and often only a part of that.

1.2.1 Vector-scan and raster-scan systems; pixels and frames

But how can we represent these images in a form which can be handled by mechanical or electronic equipment? We can answer that by looking at two means through which people have created images – or pictures – for years. The most obvious means of creating a picture is as a child does, through drawing lines which come together to form recognizable representations of shapes, objects, letters, and so forth. We can "teach" a machine to draw in this manner, simply by defining the image space (the "surface" onto which the image will be drawn) properly. We might, for example, create a two-dimensional coordinate system, and then "tell" the machine (through its programming, controls, or whatever) to create the image as follows:

1. Draw a red line from point (2,4) to point (4,4).
2. Draw a red line from point (2,4) to point (2,0).
3. Draw a red line from point (2,2) to point (3,2)…

and so forth. Rather than defining the lines' starting and ending points, we could also simply define the start point, the direction (relative to some agreed-upon reference), and the length of the line; the results would clearly be the same. This sort of drawing is called a *vector representation*, and display devices which create images in this manner are said to be *vector-scan* displays (Figure 1-2). The classic mechanical plotter or analog oscilloscope are excellent examples. For a more current example, we should note that computer graphics systems often use vector representations of lines and objects within an image, even if the final display device does not use this form.

But there's another way that people have used for centuries to make images. You can make a recognizable picture by placing small dots or tiles of pure color next to each other to

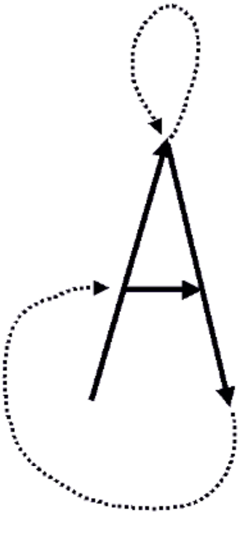

Figure 1-2 *Vector scanning*. The dotted lines represent portions of the scan in which no line is being drawn or illuminated, but rather the drawing device is being repositioned to create the next line.

form the desired shape or pattern; this is called a *mosaic*. To get a machine to draw a mosaic, it is again simplest if we define the image surface in terms of a two-dimensional coordinate system. If that is done, then we can instruct the machine as follows:

1. Place a blue tile at (4,3).
2. Place a blue tile at (3,2).
3. Place a green tile at (1,5)…

and so forth until the desired image has been created (Figure 1-3). (We assume in this exercise that all the tiles are the same size, and properly sized so that each one fills the "grid square" to which it is assigned.) But what if we are trying to create this mosaic from a "real" scene? One way to do this would be to project the image of the scene to be created, through a lens as was described above, onto a sheet of paper on which was also drawn a grid of lines using the chosen coordinate system. A person could then simply look at the projected image, determine the color, brightness, etc., at each point on the grid, and use that information to create the mosaic. (We assume we have tiles of any possible color, brightness, and so forth; this implies a very understanding tile supplier, not to mention a rather capable tile-making machine.)

This system of creating images in the manner of a mosaic is actually the basis behind practically all electronic imaging and display systems in use today. Information about the image is taken at a number of regularly spaced sampling points; this information may then be processed, translated, etc., until finally sent to a display which may itself create the desired

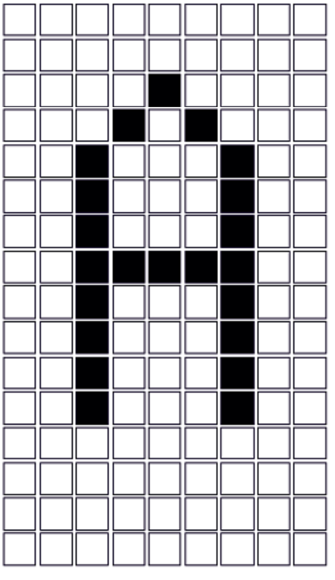

Figure 1-3 Creating an image as a mosaic. This forms the basis for the raster-scan display types; however, it is important to note that the picture elements ("pixels") we will be discussing are actually to be considered as dimensionless point samples, rather than something analogous to square tiles.

output through controlling each of a regular array of points or cells making up the display's "screen". The term *picture element,* most often condensed into the convenient word *pixel*, has been used to refer both to the individual samples of the original image at each point, and to those individual elements in the display device which make up the final displayed image. It is important not to confuse these two meanings of the word – *pixel* as a point sample of an image, and *pixel* as a physical part of a display device – as they have significant differences. The most important of these is that, in the former definition, *pixel* truly refers to a *point sample* taken from the original image; it has neither size nor shape, but is only the information regarding the characteristics of the image at a theoretically dimensionless point. This is very different from the *pixel* of many display devices, which most definitely has a fixed size, shape, and other restrictions. As will be seen later, this distinction is at the heart of many differences in image appearance among different display technologies and imaging techniques.

At this point, we should also realize that there is a more efficient means of conveying the data in such a "mosaic" system than specifying the coordinates for each tile separately. Assuming that each location within the 2-D image space must receive a "tile", or pixel, it is much easier simply to specify a starting point – the first "tile to be placed", or the first pixel location to be sampled – and then to proceed through the array in a predefined, regular manner. If the space is defined by a rectangular coordinate system, an obvious method is to proceed by sampling each location in a row, and through each row in turn. This is the basic description of the *raster scan* technique (Figure 1-4). Almost every display system in use today employs raster scanning; the vast majority of these begin the scan at the upper left corner of the image, and proceed through all the pixels in a given horizontal row and thus through each row in the supposed array of pixels. One pass through the full array is generally referred to as a *frame*.

It should be noted that the above description, based on a uniform rectangular array of sample points or "pixels", is a close match to the reality of practical fixed-format display

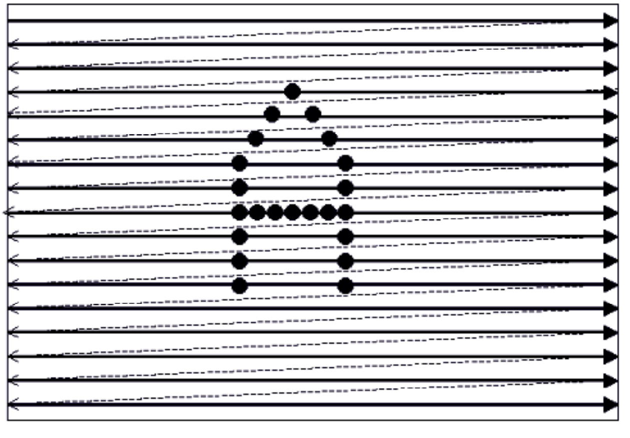

Figure 1-4 A *raster-scanned* image. In this form of display, the "drawing" mechanism follows a regular path across the imaging area, and the picture elements are placed or illuminated as required to build the desired image. Again, the dotted lines represent times during which no "drawing" is performed, but rather the drawing mechanism is being repositioned.

devices such as LCD panels, etc.. However, the notion of constructing an image through a regular, repeated scanning structure is not necessarily tied to the idea of fixed sampling points. In so-called "purely analog" systems such as the original television implementations, the image is scanned on a continuous basis, without distinct, discrete sampling points as have been assumed above. (There is much more to be said on the distinctions, both real and imagined, between "analog" and "digital" systems; this will be covered in a later chapter.) Even in systems which do employ discrete pixels, there is no requirement that these be in a simple rectangular array. Other sampling structures have been proposed, and in many cases may provide certain advantages over the common rectangular structure.

1.2.2 Spatial formats vs. resolution; fields

Regardless of whether or not the array of sampling points (pixels) is rectangular (i.e., the pixels are arrayed orthogonally, in linear rows and columns), a concept often encountered in imaging discussions is the notion of "square" pixels. This does not mean that the pixels are literally square in shape – again, a pixel in the strictest sense of the word is a dimensionless, and therefore shapeless, point sample of the original image. Instead, a sampling grid or *spatial format* is said to be "square" if the distance between samples is the same along both axes.

Images do not have to be sampled in a "square" manner, and in fact many sampling standards (especially those used in digital television) do not use square sampling, as we will see later. However, it is often easier to manipulate the image data if a given number of pixels can be assumed to represent the same physical distance in both directions, and so square sampling or square-pixel image format is almost always used in computer-generated graphics.

The term *format*, or *spatial format*, is used to refer to the overall "size" of the image, in terms of the number of pixels horizontally and vertically covered by the defined image space. For example, a common format used in the computer graphics industry is 1024 × 768, which means that the image data contains 1024 samples per horizontal row, or line, and 768 of these lines (or, in other words, there are 768 pixels in each vertical column of the array of pixels). Unfortunately, the convention used in the computer industry is exactly opposite to that used in the television industry; TV engineers more often refer to image formats by quoting the number of lines (the number of samples along the vertical axis) first, and then the number of pixels in each line. What the computer graphics industry calls a "1024 × 768" image, television would often call a "768 × 1024" image.

Readers who are familiar with the computer industry will also notice that what we are here referring to as a "spatial format" is often called a "resolution" by computer users. This is actually a misuse of the term *resolution*, which already has a very well-established meaning in the context of imaging. There is a very important distinction to be made between the *format* of a sampled image, and the *resolution* of that image or system, and so we will try to avoid giving in to the common, incorrect usage of the latter term.

Resolution properly refers simply to the amount of details which can be resolved in an image or by an imaging system; it is usually expressed in terms of the number of basic image elements (lines, pairs of lines, or pixels - also "dots" or "samples") per a definite physical distance. "Dots per inch", often seen in specifications for computer printers, is a legitimate measure of resolution – dots or pixels per line is not, since a "line" in a image does not have a definite, fixed physical length. Changing from a spatial format of, for instance, 800 × 600 pixels to one of 1600 × 1200 pixels does *not* necessarily double the resolution. It would do so

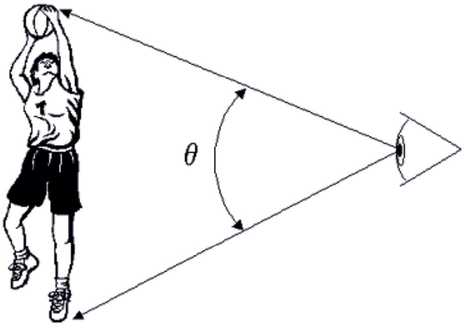

Figure 1-5 The portion of the visual field occupied by a given image may be expressed in terms of *visual angle,* or in *visual degrees*, which is simply a measure of the angle subtended by the image from the viewer's perspective.

only if the physical dimensions of the image and/or the objects within it remain constant, and so finer details within the image are now resolved.

Resolution may also be expressed in terms of the same sorts of image elements (dots, lines, etc.) per a given portion of the visual field, since this is directly analogous to distance within the image. The amount of the visual field covered by the image is often expressed in terms of *visual degrees,* defined as shown in Figure 1-5. A common means of indicating visual acuity – basically, the resolution capability of the viewer, himself or herself – is in terms of *cycles per visual degree*. Expressing resolution in cycles (or a related method, in terms of *line pairs*) recognizes that detail is perceived only in contrast to the background or to contrasting detail; one means of determining the limits on resolution, for example, is to determine whether or not a pattern of alternating white and black lines, or similarly a sinusoidal variation between white and black, is perceived as a distinct pattern or if it simply blurs into the appearance of a continuous gray shade. (It should be noted at this point that resolution limits are often different between cases of white/black, or *luminance*, variations, and differences between contrasting colors.)

1.2.3 Moving images; frame rates

To this point, we have considered only the case of a static image; a "still picture", which is often only a two-dimensional representation of a "real" scene. But we know that reality is not static – so we must also face the problem of representing motion in our images. This could be done by simply re-drawing the moving objects within the image (and also managing to restore the supposedly static background appropriately), but most often it is simply assumed that motion will be portrayed by replacing the entire image, at regular intervals, with a new image. This is how motion pictures operate, by showing a sequence of what are basically still photographs, and relying on the eye/brain system to interpret this as a convincing representation of smooth motion. Each individual image in the series is referred to as a *frame* (borrowing terminology originally used in the motion picture industry), and the rate at which they are displayed is the *frame rate*. Later, we will see that there may be two separate "frame rates" to consider in a display system – first, the *update rate*, which is how rapidly new images can be provided to or created in the *frame buffer*, which is the final image store before the display itself, and the *refresh rate*, which is how rapidly new images are actually produced on the "screen" of the display device. These are not necessarily identical, and certain artifacts may arise from the exact rates used and their relationship to one another.

Quite often in display systems, the entire frame will not be handled as a single object; it is often necessary to transmit or process component portions of the frame separately. For instance, the color information of the frame may be separated out into its primary components, or (for purposes of reducing the rate at which data must be transmitted), the frame may be broken apart spatially into smaller components. The term most often used to refer to these partial frames is *fields*, which may generically be defined as any well-defined subcomponent of the frame, but into which each of the frames can similarly be separated. (In other words, components which are unique only to a given frame or series of frames are not generally referred to as fields.)

We should note at this point that our discussion of images has now extended into a third dimension – that of time – and that the complete visual experience we typically receive from a display system actually comprises a regular three-dimensional array of samples in both time and space (Figure 1-6). Each frame represents a sample at a specific point in time; each pixel within that frame is a sample within a two-dimensional image space, representing certain visual qualities of the image at that point in space at that time. As in any sampling-based system, various difficulties and artifacts may be produced through the nature of the sampling methodology itself – the sampling rate, the characteristics of the actual implementation of the sampling devices, etc.. These will be covered in considerable detail in later chapters, but it is important to realize throughout that we are almost always in display work dealing with a series of samples rather than continuous data.

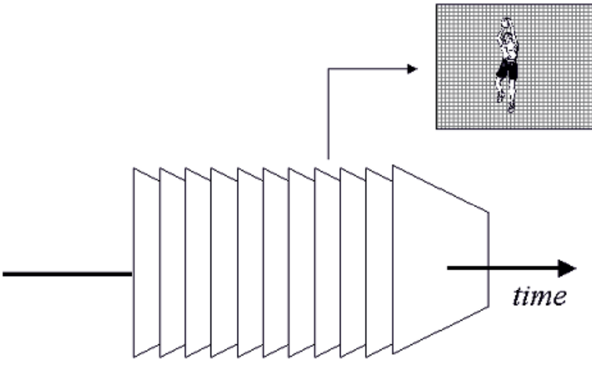

Figure 1-6 Many imaging applications involve the transmission of successive frames of information (as in the case of moving imagery), which introduces a third dimension (time) into the situation. In such cases, not only is each frame to be considered as a 2-D array of samples of the original image, but the frame itself may considered as one sample in time out of the sequence which represents a changing view.

1.2.4 Three-dimensional imaging

Extending the functioning of display systems in a third spatial dimension, and thus providing images with the appearance of solid objects, has long been a goal of many display researchers, and indeed considerable progress has been made in this field. While true "3-D" displays are not yet commonplace, there has been more than enough work here for some additional terminology to have arisen. At the very least, rendering images in three dimensions (i.e., keeping track of 3-D spatial relationships when creating images via computer) has been used for many years. It is fairly common at this point, for instance, to treat the image space as extending not just in two dimensions but in three (Figure 1-7). In this case, the sample points might no longer be called "pixels", but rather *voxels* – a term derived from "volume pixel". (It should also be obvious at this point that in extending the image space into three dimensions, we have literally increased the amount of information to be handled geometrically.)

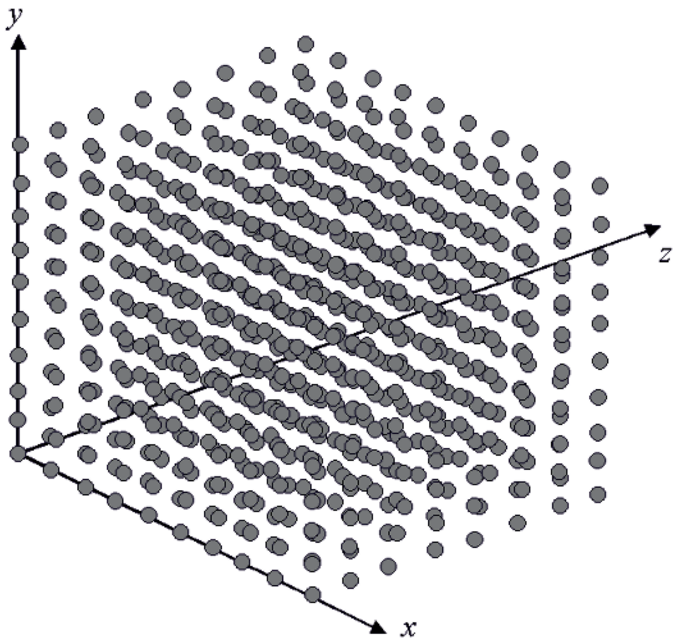

Figure 1-7 In a truly "three-dimensional" imaging system – in the sense of one having three spatial dimensions – the samples are viewed as occupying a regular array in space, and are generally referred to as "voxels" (for *volume pixel*). It is again important to keep in mind that these samples, as in the case of the pixels of a 2-D image, are dimensionless, and each represents the image information as taken from a single point within the space.

1.3 Transmitting the Image Information

Having reviewed some of the basic concepts regarding imaging, and especially how images are sampled for use within electronic display systems, it is time to consider some of the basics in communicating the information between parts of the complete system – and in particular, how this information is best transmitted to the display device itself.

As was noted earlier, almost all modern image sensors and display systems employ the "mosaic" model for sampling and constructing images. In other words, "real-world" images are sampled using a regular, two-dimensional array of sampling points, and the resulting information about each "pixel" within that array may be used to re-create the image on a display device employing a similar array of basic display elements. In practically all modern display systems, the image information is sent in a continuous stream, with all pixels in the array transmitted repeatedly in a fixed standard order. Again, this is generally referred to as a *raster scan* system.

For any form of interface to the display device, then, the data capacity required of the interface in a raster-scan system may be calculated by considering:

1. The amount of information contained in each sample, or pixel. This is normally stated in *bits* (binary digits), as is common in information theory, but no assumption should be made as to whether the interface in question uses analog or digital signalling based on this alone. Typically, each pixel will contain at least 8, and more often as much as 24–32 bits of information.
2. The number of samples in each image transmitted, i.e., the number of pixels per *frame* or *field*. For the typical two-dimensional rectangular pixel array, this is easily obtained by multiplying the number of pixels per line or row by the number of lines or rows per frame. For example, a single "1280 × 1024" frame contains 1,310,720 pixels.
3. The field or frame rate required. As has been noted, most display systems required repeated and regular updating of the complete image; the required rate is determined by a number of factors, including the characteristics of the display itself (e.g., the likelihood of "flicker" in that display type or application) and/or the need to portray motion realistically. These factors result in typical rates being in the upper tens of frames per second.
4. Any overhead or "dead time" required by the display; as used here, this term refers to any limitations on the amount of time which can be devoted to the transmission of valid image data. A good example is in the "blanking time" required in the case of CRT displays, which can reduce the time available for actual image data transmission by almost a third. (This is discussed in more detail in Chapter 5.)

For example, a system using an image format of 1024 × 768 pixels, and operating at a frame rate (or refresh rate) of 75 frames/s, with 24 bits/pixel, and with 25% of the available time expected to be "lost" to overhead requirements, would require an interface capable of supporting a peak data rate of

$$\frac{1024 \text{ pixels/line} \times 768 \text{ lines/frame} \times 24 \text{ bits/pixel} \times 75 \text{ frames/s}}{0.75} = \sim 1.89 \text{ Gbits/s}$$

(This actually represents a rather medium-performance display system, so one lesson that should be learned at this point is that display systems tend to require fairly capable interfaces!)

It should be noted at this point that these sorts of requirements are very often erroneously referred to as "bandwidth"; this is an incorrect usage of the term, and is avoided here. Properly speaking, "bandwidth" refers only to the "width" of an available transmission channel in the frequency domain, i.e., the range of the available frequency space over which information may be transmitted. The rate at which information is provided over that channel is the *data rate*, which may not exceed the *channel capacity*.

The information capacity of any channel may be calculated from the available bandwidth of that channel *and* the noise level, and is given by a formula developed by Claude Shannon in the 1930s:

$$\text{Channel capacity (bits/s)} = \text{BW} \log_2(1 + \text{SNR})$$

where BW is the bandwidth of the channel in Hz, and SNR is the ratio of the signal power to the noise power in that channel. This distinction between the concepts of "bandwidth" and "channel capacity" will become very important when considering certain display applications, such as the broadcast transmission of television signals.

It is important to note that this formula gives the theoretical maximum capacity of a given channel; this maximum is never actually achieved in practice, and how closely it may be approached depends very strongly on the encoding system used. Note also that this formula says nothing about whether the data are transmitted in "analog" or "digital" form; while these two options for display interfaces are compared in greater detail in later chapters, it is important to realize that these terms fundamentally refer only to two possible means of encoding information. To assign certain attributes, advantages, or disadvantages to a system based solely on whether it is labelled "analog" or "digital" is often a serious mistake, as will also be discussed later.

In discussions of display interfaces and timings, the actual data rate is rarely encountered; it is assumed that the information on a per-pixel basis remains constant (or at least is limited to a fixed maximum number of bits), and so it becomes more convenient to discuss the pixel rate or *pixel clock* instead. (The term "pixel clock" usually refers specifically to an actual discrete clock signal which is used in the generation of the output image information or signal, but "pixel clock rate" is often heard even in applications in which no actual, separate clock signal is present on the interface.) In the example above, the pixel rate required is simply the data rate divided by the number of bits per pixel, or about 78.6 Mpixels/s. (If referred to as a pixel *clock* rate, this would commonly be given in Hz.)

Having covered the basics of imaging – at least as it relates to the display systems we are discussing here – and display interface issues, it is now time to consider two other areas which are of great importance in establishing an understanding of the fundamentals of display systems and their applications. These are the characteristics of human vision itself, and the role these play in establishing the requirements and constraints on our displays, and the surprisingly complex issue of color and how it is represented and reproduced in these systems. These are the topics of the next two chapters.