6

The Two-dimensional Field: Area

THE SCREEN PROVIDES YOU WITH A NEW, CONCENTRATED LIVING SPACE—a new field for presenting media events. It helps you tame space. You are no longer dealing with the real, amorphous space we walk through and live in every day but with screen space. The viewfinder in a video or digital cinema camera and the various screens on which the media event is projected have definite, fixed borders that define the new aesthetic playing field. You must now clarify and intensify an event within the confines of this new space. Boris Uspensky's insightful comment about the frame in painting applies equally to the video screen: "The frame is the borderline between the internal world of the representation and the world external to the representation."[1]

This chapter examines the structural factors of screen space: aspect ratio and the aesthetics of object size and image size.

Aspect Ratio

Aspect ratio is the relationship of screen width to screen height. As a painter or still photographer, you have free choice in the basic orientation of your picture frame: you may want a frame that is taller than it is wide for the picture of a skyscraper, or a horizontally oriented format for the wide expanse of a desert landscape. You may even choose a round or an irregularly shaped boundary. In video, film, or computer displays, you do not have this format flexibility. The various video and motion picture screens may differ in size and aspect ratio, but the only flexibility is within the screen, not of the screen itself.

HORIZONTAL ORIENTATION

The standard video, film, and computer screens are horizontally oriented. Why?

One reason could be that our peripheral vision is greater horizontally than vertically.[2] But this does not mean that we actually see much more at the sides of our head than what is above or below it, especially because we can look up and down just as easily as sideways. Even if the screen is small enough that it is reflected in full at our retina, we can't see all of its content by simply staring at it. Our eyes must constantly cut from one spot to another in a picture before we can make sense of what we see.[3] A wide screen will not eliminate such a scanning process. It will, however, allow us to use our peripheral vision to become aware of what is going on at the sides of the screen before deciding to actually look at it.

A more compelling reason for the horizontal screen is that we normally live and operate on a horizontal plane. In the context of our everyday pursuits, the world is flat. Gravity makes it relatively easy for us to stay and move on a horizontal plane but rather difficult to conquer vertical space. To travel 200 miles over land is nothing special; to cover the same distance vertically is a supreme achievement. A 600-foot-high tower is a structural adventure; a 600-foot-long building is simply large.

STANDARD ASPECT RATIOS

The two aspect ratios most prevalent in video production are the traditional 4 × 3 and the HDTV 16 × 9. Standard television and many computer screens took over the traditional 4 × 3 aspect ratio of early motion pictures. Regardless of their size, they are all 4 units wide by 3 units high. This aspect ratio is also expressed as 1.33:1—for every unit in height, there are 1.33 units in width. SEE 6.1

The digital video screens, usually called HDTV television screens regardless of whether they are high- or low-definition, have a wide-screen aspect ratio of 16 × 9, or 1.78:1. SEE 6.2 The aspect ratio of the standard wide screen of motion pictures is 5.55 × 3, or 1.85:1. SEE 6.3

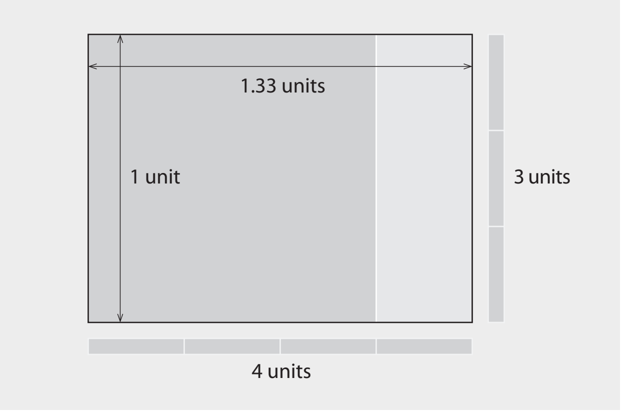
The Panavision 35 or Cinemascope format has a very wide, wraparound aspect ratio of 7 × 3, or 2.39:1. To distinguish it from the older and slightly narrower Cinemascope aspect ratio of 2.35:1, the movie pros call the wider version the 240 ("two-four-O") format. SEE 6.4

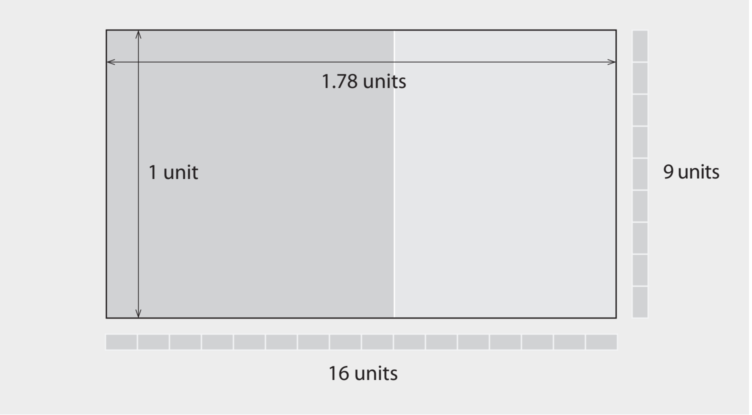
Most U.S. films are shot and projected in the 1.85:1 ***wide-screen format***. The newer, wider Panavision aspect ratio requires an anamorphic lens to stretch the squeezed film frame for projection on the ultrawide screen.

The aspect ratios of mobile video devices (cell phones and mini camcorders) differ widely to serve a variety of multimedia functions, such as texting, framing a shot during video capture, or displaying graphics, tables, and text. Some have a single fixed window in the upper part of the device that can display a variety of differently proportioned images; others have a larger screen whose aspect ratio is a skinny vertical one when held upright and a wide one when turned sideways. As a stretched vertical aspect ratio (from top to bottom), the screen can display vertical vistas or be divided into a number of individual picture spaces; as a horizontal one, it can display wide-screen scenes and, especially, wide texting. SEE 6.5–6.7

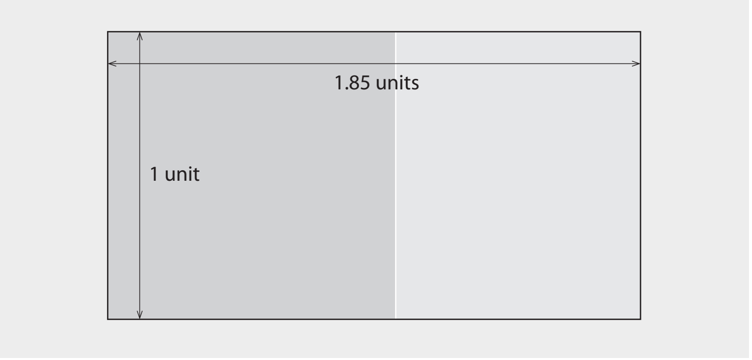
6.1 Standard Television, Computer, and Classic Movie Screen Aspect Ratio

All standard television screens, computer screens, and the classic motion picture screen have an aspect ratio of 4 × 3, which means that they are 4 units wide by 3 units high. The 4 × 3 aspect ratio is also expressed as 1.33:1—for every unit in height, there are 1.33 units in width.

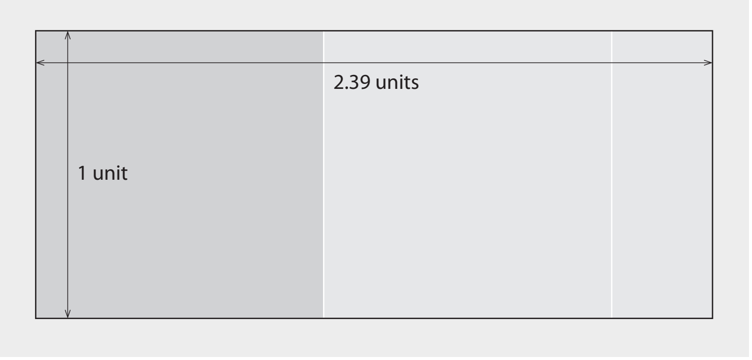




6.2 HDTV Aspect Ratio
The HDTV screen aspect ratio is 16 × 9, or 1.78:1, making this screen more horizontally stretched than the traditional television screen. The HDTV aspect ratio can accommodate wide-screen movie formats without much picture loss at the sides. Many standard digital cameras can switch between the 4 × 3 and the 16 × 9 aspect ratios.



6.3 Wide-screen Motion Picture Aspect Ratio
The aspect ratio of wide-screen motion pictures is 5.55 × 3, or 1.85:1. This ratio provides a horizontally stretched vista and is the standard film aspect ratio in the United States.



6.4 Panavision 35 Motion Picture Aspect Ratio
The aspect ratio of the Panavision film format is 7 × 3, or 2.39:1, which gives a wider panoramic view than the standard wide-screen format.

6.5 Mini Camcorder
There are several mobile video devices whose screens have the classic 4 × 3 aspect ratio.

6.6 Vertical Aspect Ratio
When held upright, this mobile video device can display images in a relatively large screen with a vertical aspect ratio. It can display tall objects or stacked information.

6.7 Horizontal Aspect Ratio
When turned sideways, the same screen becomes a display area with a wide-screen aspect ratio.

6.8 Framing Height and Width in a Single Shot
In the 4 × 3 aspect ratio, the difference between width and height is not so noticeable as to unduly emphasize one screen dimension over the other. It accommodates both horizontal vistas and vertical objects.

FRAMING IN THE 4 × 3 ASPECT RATIO

The 4 × 3 standard for film was established as early as 1889. There is still some controversy, if not confusion, about just why this aspect ratio became the standard.[4] Be that as it may, when you look at this standard not from a technical but from an aesthetic point of view, the reason for this screen format is not so baffling.

The advantage of the 4 × 3 aspect ratio is that the difference between screen width and screen height is not pronounced enough to emphasize one dimension over the other. SEE 6.8 You can frame a horizontally oriented scene without too much wasted vertical screen space, and a vertical scene without worrying too much about how to fill the sides of the screen. SEE 6.9 AND 6.10

The 4 × 3 aspect ratio is especially appropriate for the close-up (CU) and extreme close-up (ECU) framing of a person. In the CU or medium shot (MS), the person seems to be comfortably placed in the 4 × 3 frame; but in the 16 × 9 frame, the subject in the medium shot seems somewhat lost in the wide-screen video space. SEE 6.11 AND 6.12

6.9 Framing a Horizontal View
The 4 × 3 aspect ratio can easily accommodate a horizontally oriented vista.

6.10 Framing a Vertical View
Even a vertically oriented scene can be easily framed within the 4 × 3 aspect ratio.

6.11 Framing a CU in the 4 × 3 Aspect Ratio
The 4 × 3 aspect ratio allows us to frame a close-up of a person without crowding the frame or leaving unused screen space.

6.12 Framing an MS in the 16 × 9 Aspect Ratio
When framing a medium shot, there is much unused space on both sides of the screen (called dead zones), so the person seems oddly lost in video space.

6.13 Framing an ECU in the 4 × 3 Aspect Ratio
The 4 × 3 aspect ratio can accommodate an extreme close-up without leaving undesirable screen space at the sides.

6.14 Framing an ECU in the 16 × 9 Aspect Ratio
When framing a close-up or an extreme close-up in the HDTV aspect ratio, we are again left with a considerable amount of unused screen space on either side of the subject. Also, the person looks somewhat squeezed by the low ceiling of the frame.

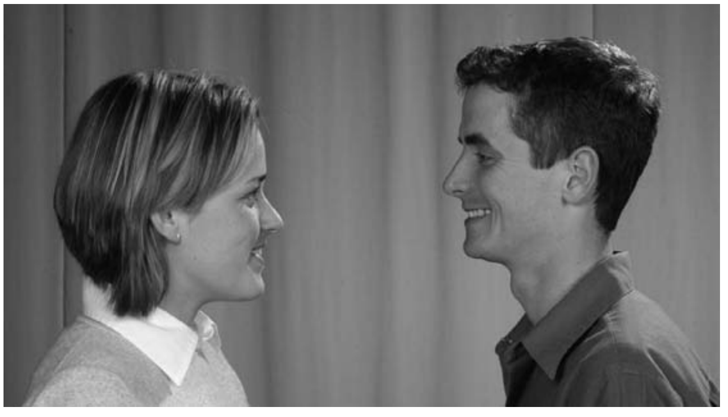
The extreme close-up of the person looks properly intensified when placed in the 4 × 3 aspect ratio but loses intensity in the 16 × 9 frame because of the "dead zones," of leftover screen space on either side. SEE 6.13 AND 6.14 These dead zones are especially noticeable on the larger HDTV screens.

FRAMING IN THE 16 × 9 ASPECT RATIO

Why, then, did the motion picture and HDTV standards committees opt for the wider aspect ratio? One of the initial reasons was certainly film's desire to compete with television and to engulf us in spectacle—something difficult for standard television to achieve.[5] In the meantime directors and camera operators have learned how to use the wide screen effectively for intimate scenes as well. Still, it is the landscape shots in which the wide-screen format shines. As discussed in chapter 4, the term *landscape* includes environmental shots, such as actual landscapes as well as city streets, sports arenas, and interiors.

6.15 Framing a Wide Vista in the 16 × 9 Aspect Ratio

The wide 16 × 9 aspect ratio is ideally suited to show the long row of people waiting for the bus. In a narrower frame, the picture would lose its charm.



6.16 Framing of Two Dialogue Partners in the 16 × 9 Aspect Ratio

On the wide HDTV screen, you can easily frame a head-on medium shot of two people facing each other. The two people can stand at a comfortable distance without being too close to each other.

6.17 Framing of Two Dialogue Partners in the 4 × 3 Aspect Ratio

On the narrower standard television screen, the same setup is difficult to frame. Note how the two people are partially cut by the screen edges. The empty screen center becomes especially prominent.

Although the 4 × 3 aspect ratio can frame horizontal scenes as well as vertical ones, it is obviously not as adept at showing wide vistas. No matter how hard you try, you will find it difficult to show a long row of people queued horizontally or to overwhelm your audience with the great expanse of the ocean, the majestic power of the Grand Canyon, or the intensity of a sold-out football stadium. SEE 6.15

On a much smaller scale, it easier to show people standing or sitting across from each other in a conversation in the wide aspect ratio than on the smaller 4 × 3 screen. There is enough lateral space to have the people at a comfortable distance from each other. When duplicating the shot on the small screen, you usually end up with both conversation partners seemingly glued to the edges of the screen with a definite aesthetic hole in screen-center. On the wide screen, the people have more personal space, more breathing room. SEE 6.16 AND 6.17

The wide aspect ratio also allows you to establish a specific setting, such as a street scene, a restaurant, or a living room, in a single long shot. SEE 6.18 It also gives you more room for lateral (screen-left or screen-right) motion without having to pan with the camera. Such flexibility is greatly appreciated by television

6.18 Establishing a Scene

With the wide screen, you can establish a scene without having to go to an extreme long shot.

6.19 Framing Vertical Vistas

A vertical vista, such as a tall building, is difficult to fit into the wide HDTV screen. One solution is to frame it diagonally.

directors when blocking multicamera productions, such as soap operas and other serial dramas.[6] Lateral motion is relatively ineffective on the small 4 × 3 screen, but it can have powerful effects on the wider 16 × 9 screen.

Framing tall objects such as towers, high-rise buildings, and trees in the wide-screen aspect ratio always presents a problem. On long shots you end up with an inordinate amount of space to fill on both sides. When using closer shots, you need to decide to show either the top or the bottom of the tall object. Such problems are usually solved by tilting the object and placing it on the screen diagonal, by shooting from below and looking up with the camera, or by slowly tilting up or down with the camera. **SEE 6.19** Another way is to create a vertical aspect ratio within the wide screen, which is the subject of the following discussion.

Flexible Aspect Ratio

We cannot alter the physical aspect ratio of a video, movie, or computer screen, but we can rearrange the screen space itself to create a variety of horizontal landscape and vertical portrait formats. The most immediate task is to accommodate wide-screen HDTV and film images into the 4 × 3 screen of the standard television set and to fit the 4 × 3 images of standard video into the wide-screen format. A similar problem exists when placing the standard video and wide-screen HDTV and film aspect ratios into the various small mobile video displays.

Regardless of the aspect ratio of the video screen, we can create secondary frames within the principal screen.

MATCHING ASPECT RATIOS

So far the problem has been to make the 16 × 9 movie images fit the 4 × 3 television screen. But with the increasing number of wide-screen TV sets in use, we face the reverse dilemma when trying to show material that has been produced in the traditional 4 × 3 format.

Windowboxing The simplest way to accommodate a 4 × 3 picture in an HDTV aspect ratio is to reduce the size of the picture and place it in the center of the wide-screen frame, with the leftover space of the 16 × 9 frame surrounding it. This technique is called ***windowboxing***, although the 4 × 3 image looks more like a picture with a matte around it than a window. The problem with such a framing is, of course, the reduced size of the 4 × 3 image. **SEE 6.20**

Pillarboxing Fitting a full-sized 4 × 3 image in the middle of a 16 × 9 screen is called ***pillarboxing***. When using this method, the 4 × 3 picture is simply inserted into the center of the wider screen so that the top and the bottom of its frame coincide with the 16 × 9 picture space. The leftover empty sides of the screen are blanked with vertical bars, often called side bars or dead zones. These side bars are vertical and look like pillars, hence the term pillarboxing. **SEE 6.21**

Letterboxing ***Letterboxing*** is used to make a wide-screen presentation, such as a 1.85:1 movie or 16 × 9 (1.78:1) digital video, fit a 4 × 3 (1.33:1) screen or some kind of vertical mobile video display.

A wide-screen letterbox is created by showing the whole width and height of the original format and masking the top and the bottom of the screen with black, white, or colored bars. **SEE 6.22**

You may also see motion pictures windowboxed on the HDTV screen, which makes little sense, especially because the wide-screen HDTV format was developed in large part to accommodate motion pictures without cropping. Such

6.20 Windowboxing

Maintaining the aspect ratio of the original image but reducing it in size and then placing it in the center of the primary (usually 16 × 9) screen is called windowboxing.

6.21 Pillarboxing

Fitting a 4 × 3 format into a 16 × 9 aspect ratio leaves empty bands on both sides of the screen. This change of aspect ratio is called pillarboxing.

6.22 Letterboxing
When showing a full frame of a wide-screen image on the standard television screen, we must leave borders on the top and the bottom of the screen. This change of aspect ratio is called letterboxing.

6.23 Letterboxing on Mobile Video Display
The vertical format of this mobile video display is changed to a 4 × 3 video format through letterboxing. Note that the so-called dead zones are quite alive with written information.

maneuvers are performed not for aesthetic reasons, however, but for technical ones. A smaller image saves bandwidth, which allows more information to be transmitted or squeezed onto recording media.

Because the small display of mobile multimedia phones may be square or even vertical, letterboxing is the typical way of adjusting the display to accommodate the 4 × 3 format of standard television or the 16 × 9 format of HDTV or wide-screen movies. Often the dead zones are quite alive with lettering, which sometimes relates to what is displayed on the screen or shows unrelated text messages. **SEE 6.23**

You will find that on a fairly large screen, such aspect ratio accommodations do not do too much damage to the overall aesthetic intent and usually preserve the integrity of the original production. On a small screen, however, the letterboxed or pillarboxed images shrink so much that they inevitably lose most or all of their landscape impact. Realize, however, that such a loss of aesthetic energy is more a function of the small screen size than the manipulation of aspect ratios.

Cutting, stretching, and squeezing More drastic ways of making wide-screen images fit the aspect ratio of standard video is to simply crop the wide-screen image on either side. If the film or video was shot with standard television in mind, not much is lost. A good director will have kept most of the important action screen-center and more of the decorative material, which Peter Ward calls "visual fluff," on the extreme sides.[7] If the movie or video was shot explicitly for the wide screen, however, such a technique is quite intrusive. You become especially aware of this crude method of amputating picture information when letters or words from titles, credits, or subtitles are missing. When you enlarge the standard 4 × 3 television image to fill the sides of a wide-screen television receiver, you inevitably lose headroom and cut off some feet at the bottom of the screen.

Fit, Fill, Copy, Scale

A relatively simple way to address this problem is to digitally stretch or squeeze the image to fit a specific aspect ratio without the telltale dead zones on the sides or on the top and the bottom of the screen. Unfortunately, when stretching the 4 × 3 image to fit the wide screen, the people on-screen instantly gain at least 25 pounds and vertical objects look wider than in the original. All

6.24 Digital Stretching
Sometimes the 4 × 3 image is digitally stretched to make it fit the 16 × 9 HDTV aspect ratio. The unfortunate consequence of this simple solution is that objects lose their normal proportions: people look chubby, horizontal lines get longer, and vertical lines get wider.

6.25 Digital Squeezing
To fit the wide-screen 16 × 9 image into the standard 4 × 3 aspect ratio, the image is sometimes digitally squeezed. People look thinner and taller, and objects look vertically stretched.

Big NO-NO; Unless content is designed for this and creatively applied, DON'T. Too often looks like you didn't understand or care enough to fix it.

But I agree with this -->

horizontal lines get wider, as well. SEE 6.24 When squeezing a wide-format image into a 4 × 3 aspect ratio without the dead zones, the people and vertical objects get thinner, and the horizontal lines get shorter. SEE 6.25 These stretching and squeezing methods become especially noticeable when there is written material on the screen. We as viewers may eventually accept such aesthetic anomalies if we are told often enough that this is the new and fashionable look of high-tech picture manipulation.

One of the more sophisticated and costly techniques for adapting a wide-screen presentation to normal television is the ***pan-and-scan*** process, whereby the more important portions of the wide-screen frame are scanned and made to fit the 4 × 3 aspect ratio of the standard video screen.

To satisfy the aesthetic requirements of the tiny displays of mobile media, various techniques have been developed to "repurpose" visual content.[8] This means that a series of shots is analyzed for various aesthetic criteria and then reedited so that they will show up effectively on the small screen. This is much more frequently a matter of resizing content or highlighting the area of interest, however, than of changing the aspect ratio.

Whatever technique is employed, such a recomposition of each frame is as much an aesthetic intrusion as re-editing the entire movie. It is therefore much easier for you to frame each shot with the "4 × 3 protect" in mind, which means to keep, at least for the time being, the central action more or less confined to the 4 × 3 aspect ratio.

ASPECT RATIOS OF MOBILE VIDEO MEDIA

Although the viewing screens of mobile video devices have different aspect ratios, all must be able to accommodate the 4 × 3 and 16 × 9 formats. Much like the larger screens of television sets, the small screen of the mobile video media use letter- and pillarboxing for displaying content framed in one or the other aspect ratio.

Because of the small size of the mobile video displays, the difference between the two aspect ratios is hardly noticeable. In any case, the aesthetic consequences of such space manipulation are minimal. For example, the extended width of a large 16 × 9 screen is definitely superior to the 4 × 3 format when showing a typical landscape shot, such as a city skyline, a mountain range, or a river. As

6.26 **4 × 3 Mobile Video Display**
The small 4 × 3 screen of a mobile video device can show a landscape but not carry its emotional impact.

6.27 **16 × 9 Mobile Video Display**
Although the shot now captures the full 16 × 9 aspect ratio, the difference in size from the 4 × 3 ratio is so minute that it will not make us feel the landscape any more intensely.

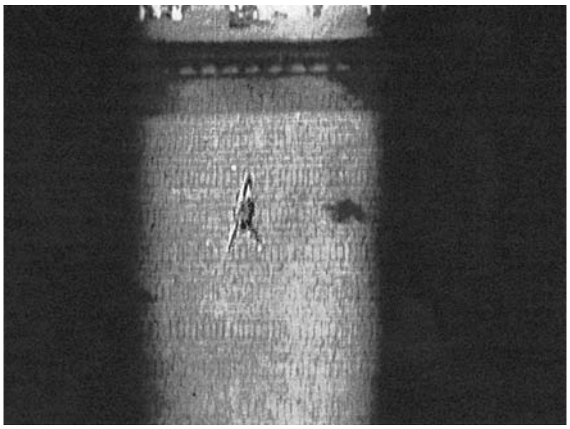
mentioned earlier, the extended screen width intensifies the landscape character of a horizontally oriented scene. On the small screen, however, the difference in screen size is so minute that it fails to cause such an affective response. We may *see* a slight difference in how much of the scene is captured by the shot, but we won't *feel* it. **SEE 6.26 AND 6.27**

SECONDARY FRAMES

Since the early days of the motion picture, directors have attempted to change the fixed aspect ratio of the screen whenever the pictorial content demanded it. Especially in the era of silent film, when story and concepts had to be communicated by visual means only, attempts were made to break out of the static 4 × 3 aspect ratio.

> American film producer and director **D. W.** (David Wark) **Griffith** (1875–1948) was the innovator of several filmic techniques, such as using various camera positions to change the field of view (long shots and close-ups) and the angle from which the event was filmed. His *Birth of a Nation* (1915) established the feature-length film.

Artificial masking D. W. Griffith used masking during a spectacular battle scene in his 1916 film *Intolerance.* By blacking out both sides of the screen, he created a secondary frame and changed the fundamentally horizontal aspect ratio to a vertical one. This vertical orientation greatly intensified a soldier's fall from high atop the walls of Babylon. Yes, he used the pillarboxing technique long before the digital age. **SEE 6.28**



6.28 **Changing Aspect Ratio Through Artificial Masking**
One of the more successful masking effects occurred in 1916 in the spectacular battle scene in D. W. Griffith's *Intolerance.* To intensify a soldier's fall from high atop the walls of Babylon, Griffith changed the horizontal aspect ratio to a vertical one by masking both sides of the screen. Obviously, pillarboxing is nothing new.

6.29 Changing Aspect Ratio Through Organic Masking

Rather than artificially block out the sides of the wide movie screen to change its aspect ratio, we can use scenic elements such as trees or buildings as masking devices. In this case, the foreground buildings change the horizontal aspect ratio to a vertical one that dramatizes the high-rise building.

Organic masking A less obvious method of masking is filling the sides of the screen with natural scenic elements, such as buildings, trees, or furniture. This technique is especially effective when framing close-ups or vertical scenes within the wide aspect ratio. **SEE 6.29**

SCREENS WITHIN THE SCREEN

You can create a secondary frame of any aspect ratio within the principal frame and simply leave the space around the secondary frame empty (or color it in some way to suggest a common background). This is one of the favorite ways in advertising to draw attention to the central message. **SEE 6.30**

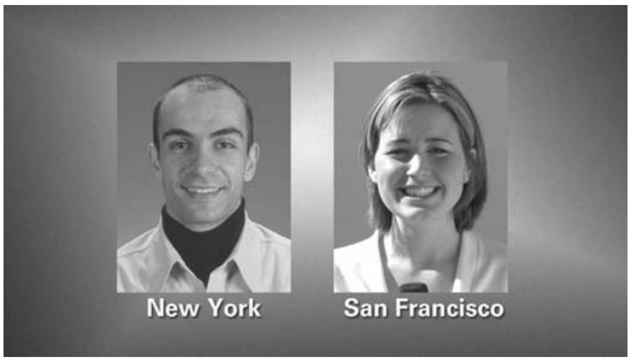
Today the use of different aspect ratios within the main frame is common in video presentations, especially in news and advertisements. Through the use of digital video effects equipment, we can easily create separate and discrete picture areas within the television frame, each carrying a different static or moving image. **SEE 6.31**

By displaying multiple images with their own aspect ratios on the principal video screen, you are able to show simultaneous yet spatially separated events or relate various subjects or points of view in a single shot. Note, however, that such displays do not change or expand the original aspect ratio of the screen. **SEE 6.32**

However you look at it, a divided screen cannot replace the effect of an expanded field through multiple screens. We discuss this distinction in chapter 8.

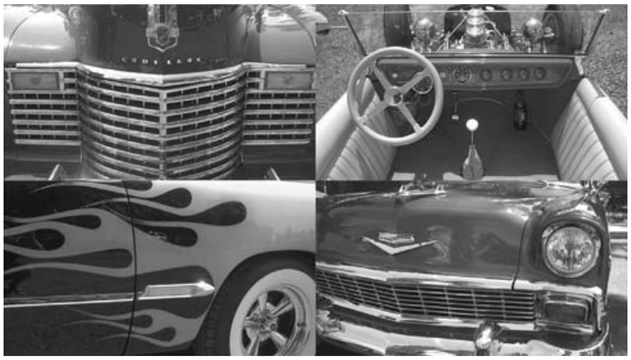
6.30 Secondary Frame Within the Principal Frame

A secondary frame within the main video display can have a variety of aspect ratios. In this case, the secondary frame has a vertical orientation. This is similar to windowboxing, but the secondary frame does not have to be centered.



6.31 Changing Aspect Ratio with Secondary Frames

Through digital video effects, we can create many secondary picture areas within the video display. Both secondary frames have a different aspect ratio from the principal 16 × 9 screen.



6.32 Split Screen
So long as we are aware of the actual screen borders, the four separate screen areas divide the major display but do not expand it as would separate smaller screens.

Moving camera A moving camera, of course, can easily overcome the basic restrictions of the fixed aspect ratio. To show the height of a tower, for example, you can start with a close-up of the bottom of the tower and then gradually tilt up to reveal its top. Panning sideways, you can reveal an equally wide horizontal vista. Such gradual revelations are often more dramatic than a long shot of the same scene.

The Aesthetics of Size

When you watch a movie first on a large screen in the theater and then on a small video screen, do you see giants on the large screen and Lilliputians on the small screen? Of course not. As with color constancy, which makes us see colors as uniform despite variations, our perception is guided by ***size constancy***, which means that we perceive people and their environments as normal sized regardless of whether they appear in a long shot or a close-up on a large movie screen or a small video screen, or whether we are relatively close to or far away from the screen. So long as we know by experience how large or small an object should be, we perceive it as its normal size regardless of screen size, relative image size, or perceived object distance. As you will see, however, the physical size of the screen does have a great influence on how we perceive its projected images.

To learn more about the aesthetics of size, we take a closer look at two important factors: object size and image size.

Size constancy is truly a contextual phenomenon. It is the context within which we see an object that supplies us with the necessary clues for perceiving the object's actual shape and size. In the absence of other clues, it is the perceived distance that keeps the size of an object constant.

Keep the same scale between the project assets!

Object Size

How do you know how big an object that appears on a video or movie screen really is? You can't tell. **SEE 6.33** Because you know how big a playing card actually is, you

6.33 Interpreting Object Size: Knowledge of Object
How big are these playing cards? We don't really know. All we can do is assume that they are the same size.

6.34 Comparison to Established Size Standard
Now we know that they are different-sized playing cards. But we still don't know which of the three cards is the normal-sized one.

6.35 Size Reference to Human Being
As soon as you see the cards in relation to the universal reference—a human being (or parts thereof, such as these hands)—you can immediately judge the actual size of the objects. The middle card is the normal-sized one.

probably assume that you are seeing close-ups of standard-sized playing cards. Not so. As you can see in the figure, the shots are of cards that differ considerably in size. **SEE 6.34** Which of the three cards is normal size? Although you now know that there are three different sizes of cards, do you know just how big they actually are? You still can't tell for sure. The only sure way of indicating actual size is to put the cards in the context of a universal size reference: a human being. For example, a hand holding the three cards is an immediate and fairly accurate index of the cards' actual sizes. **SEE 6.35**

But we also have other, perhaps more subtle perceptual clues to judging size when objects appear as screen images: knowledge of the object, relation to screen area, and scale.

KNOWLEDGE OF OBJECT

If you are familiar with the object displayed on the screen, you automatically translate the screen image into the known size of the object regardless of whether the object appears in a long shot or a close-up (see figure 6.33).

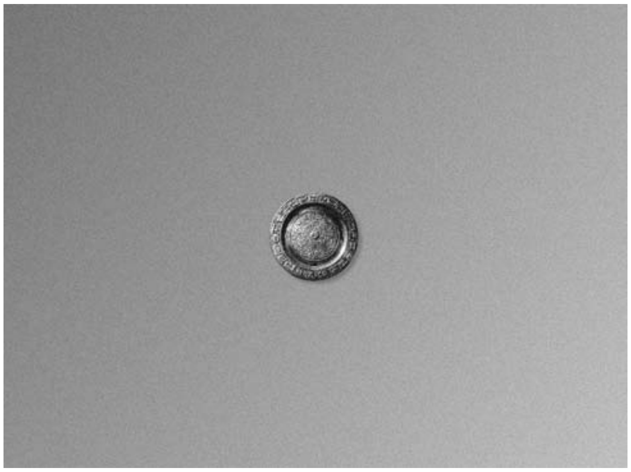
RELATION TO SCREEN AREA

If we don't know the object on the screen, we tend to judge its size by how much screen area it occupies. If it takes up a large portion of the screen, we perceive it as relatively large. **SEE 6.36** If it covers only a small area, we perceive it as relatively small. **SEE 6.37** But even if we know the object, we tend to feel more graphic weight and, with it, more aesthetic energy when the object is shown in a close-up than in a long shot.

Let's assume for a moment that you are asked by an advertising agency to get a tight close-up of a tiny, powerful new watch battery so that viewers can read the manufacturer's name stamped on it. Would you agree to such a shot? In the light of the theories put forth here, such a close-up would certainly prove counterproductive. A screen-filling close-up of the battery would take up so much of the picture area that we would perceive the battery as both large and heavy.

A relatively simple solution might be to pull back for a long shot so that the battery is quite small relative to the surrounding screen space and then key the name of the manufacturer above or below it. Even more effectively, you could have someone hold it on a fingertip, using the hand as a universal size reference. You could then move in for a tight close-up, because the size relation between

6.36 Screen Area as Size Reference: Close-up
If we don't know the object, we tend to judge its size by how much screen area it occupies. The more it occupies, the larger it seems to be. In this close-up, the disc seems large.

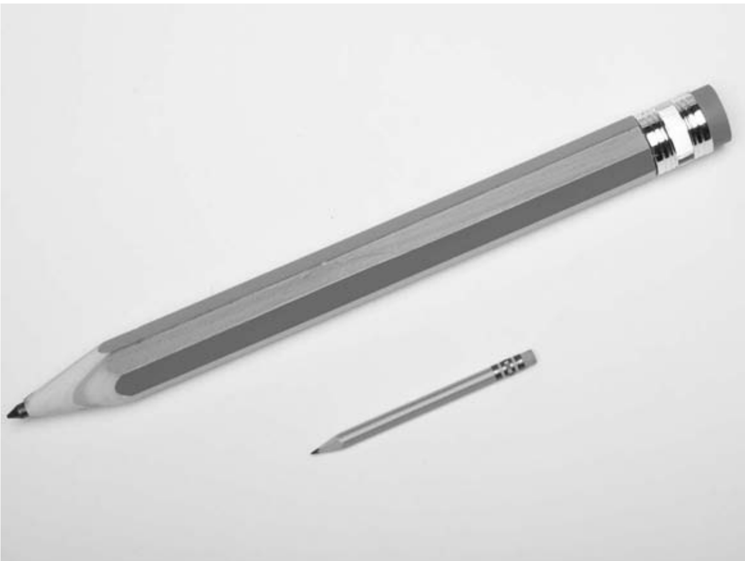


6.37 Screen Area as Size Reference: Long Shot
In this long shot, the disc occupies relatively little screen space, so it appears to be small.

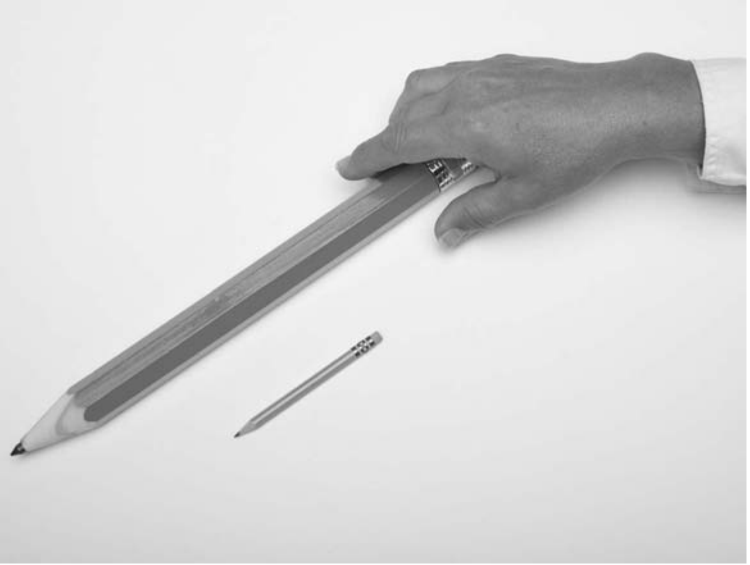
hand and battery is established and will remain largely intact. You can then key into the frame the manufacturer's name.

SCALE

As in our real environment, we make continual judgments about object size by seeing the object on-screen in relation to other (usually known) objects that appear in the same shot. We seem to decide on a size standard against which the other sizes are judged and hierarchically ordered. The convenient feature about scale is that it remains constant, regardless of whether you perceive it on the large movie screen or the tiny mobile video display. To establish a scale, however, we must have some norm, some clue that tells us just what is large and what is small. In the absence of such contextual information, even a comparison of two known objects leaves you in the dark about their actual size. SEE 6.38



6.38 Screen Frame as Scale Reference
In the absence of more accurate contextual clues, we use the proximity of the object to the frame edge as a clue. The larger pencil is probably shown in a close-up.



6.39 Hand as Scale Reference

As with the playing cards, the hand as a contextual reference helps establish an accurate scale. The large pencil was not a normal-sized one on a close-up but an oversized one on a medium shot.

The two pencils obviously differ in size. But how? In the absence of more obvious clues, we try to estimate the apparent distance of an object and then calculate its size.[9] If we assume that the larger pencil is a normal-sized one, it must be on a close-up, and the smaller one must be a little pencil. This stands to reason because we come up against more small pencils than supersized ones. But when we introduce our universal size reference—a hand—we see that we were, indeed, fooled by this shot and that our established scale was based on wrong assumptions. **SEE 6.39**

We usually get a great variety of contextual clues that help us establish a reasonably accurate scale for screen-displayed object size. But if it is contextual clues that prompt us to establish a scale for object size, couldn't we provide false clues that trick the viewers into a fallacious scale so that they perceive small things as large and large things as small? Yes, and we do this every time we use small models for simulating large-scale events or put an event into a computer-generated environment. The real problem, however, is not establishing a scale but maintaining it throughout the screen event. A model of a skyscraper must look big in relation to all contextual event factors, such as people, cars, and trees. The important, and not always easy, additional task is to make the movement of all people and objects fit the established scale. A small model of a large battleship usually bounces much too fast in relation to the real movement of the waves. A model of a spaceship must move extremely slowly if we are to believe that it is huge. Fortunately, today we can manipulate such movement quite easily with digital wizardry.

Cartoons always operate on the contextual principle of scale: the close-up of a foot trampling houses, trees, bridges, and everything else in its way becomes that of a giant if we take the landscape as the norm. But if we take the foot as the standard-sized reference, the landscape becomes small like the one of the unfortunate Lilliputians in *Gulliver's Travels.*

As with the perception of other aesthetic phenomena, our translation of screen images into actual object size is greatly influenced by other contextual clues—in this case, by the framework of the story.

Image Size

Have you ever felt ill at ease when watching a motion picture spectacle on the small video screen or, worse, on the mobile video display? Some motion pictures designed for the large, wide movie screen do not seem to come across as well on a standard-sized television screen.

Despite our hardwired ability to perceive scale as constant regardless of how big the pictures are, physical size still has considerable influence on how we perceive and feel about screen images.

The overwhelmingly large images on the horizontally stretched movie screen present a spectacle, no matter what is shown. People as well as things attain dramatic proportions, not only physically but also psychologically. The landscape aspect of a scene—that is, its physical environment—carries at least as much energy as the people who play in it. On the large movie screen, the simple event of a man walking down a country road becomes a grand act; on the small video screen, it remains a simple gesture.

In a movie about space exploration, the spaceship floating majestically across the giant screen most likely has as much or more impact on the viewer as the fate of the crew. Large HDTV screens favor, similar to the large movie screen, landscape over people. But on the small video screen, the people and their actions command attention and supply the primary energy; the environment remains relatively peripheral if not incidental. Cinema and large-screen video derive their energy from landscape as well as people; in small-screen video, it is primarily people with all their complexity who power a scene, while the landscape aspects take a back seat.

Does this mean that we cannot show landscape-type movies on a mobile video device? Not necessarily. First, the computer interface, which often divides its page into many diminutive pictures and text blocks, has already conditioned us for small-screen static and moving images. Second, it depends on how you are looking at the movie. Trying to watch a complex, full-length film while sitting on a bus would probably not be the most stirring experience, even if the bus ride were fairly long. Although you would be able to follow the plot enough to fear for the hero and loath the villain, the small images and inadequate sound would probably fail to communicate the story's full dramatic impact or give you the gut-level experience of large, high-energy images and sound.

But if your friend sends you the same movie via your mobile video device with a request to analyze the continuity of its editing, the small screen and its tiny images do not present an insurmountable handicap, even while riding on a bus. You no longer are trying to immerse yourself emotionally in the program but are scrutinizing it more intellectually. When you do this, the experience of watching the small mobile video display is not much different from looking at landscapes or portraits in miniature paintings or through the viewfinder of a film or video camera when framing a shot. In either case, the aesthetics of size and the emotional impact on how we perceive screen images of people and their surroundings are no longer an issue.

But how can we optimize the images on the mobile video display, assuming that mobile video is used as a diversion rather than an instrument of analysis or communicating text messages?

MOBILE VIDEO MEDIA AS COMPANION

!!!

Much like watching a program on your small television set in the kitchen while doing the dishes, or listening to your iPod while jogging, your mobile video media can be a pleasant companion or an occasional diversion from your primary activity. In this case, you would do much better by choosing a video that has been specifically designed not only for the small screen but also for such viewing conditions. But what are these design factors? Let's have a look at the major ones.

Close-ups Like the standard television screen, the mobile video display requires a predominance of close-ups. On the most basic level, if the shot is too wide and you can't see what it contains, you need to get closer. Long shots are always hard to see on a small screen, and they lack aesthetic energy. In fact, the quickest way to deflate a shot of its energy even on a larger screen is to zoom back from a close-up to a wider shot. On the small screen, you need to use close-ups as much as possible, regardless of whether you show people or environment. Yes, the infamous "talking heads" gain new prominence. In fact, this is one of the prerequisites for presenting interviews. There is literally no room on a mobile video display to have your guests sit comfortably apart from each other on a loose medium shot during the interview or walk side-by-side through a picturesque countryside on a long shot.

Inductive sequencing As you will learn in chapter 11, the inductive visual approach means that you present a story in a series of close-up details rather than move from an establishing long shot to a few eventual close-ups. The shot sequence must be very dense, which necessitates cramming a lot of shots into a relatively short time period, much like commercial spots do. Because inductive sequencing is mostly done with close-ups, it is essential for mobile video.

Brief running times The stories must be told in a very short time span or at least in relatively short segments, from 30 seconds to a maximum of three minutes. If this seems too short to you, just look at the average news story, which rarely exceeds 20 seconds, video clips included. Again, realize that your stories are mostly watched by somebody walking to the subway, riding on a bus or in a car, or sitting in a classroom (before the professor arrives). A series of brief stories that, like soap operas, are self-contained but have a continuing plot are most effective. You can also provide the necessary continuity by using the same character in different stories. A researcher from the London School of Economics calls this "content snacking."[10]

Dense audio track Because of the extremely limited visual space, you need to pay particular attention to the audio track. The audio must not only reproduce what is being said but also provide ambient sounds. These environmental sounds will help expand the screen and convey what is occurring in on- and off-screen space as well as provide much needed aesthetic energy (see chapter 15).

You probably noticed that all these design factors for mobile video are already applied to movie trailers and commercial spots. Despite your best efforts, however, it is virtually impossible to generate the aesthetic energy of large images on the tiny mobile video display. If you try to do it by boosting the audio level—a technique that works well in a movie theater—the sound has a tendency to separate from the picture and take on its own life.

IMAGE SIZE AND RELATIVE ENERGY

Despite our facility for size constancy, image size still influences how we feel about certain screen events. The large images on the panoramic movie screen *feel* more overpowering than a small video image. They are visually "louder" than the pictures on the video screen. When an image is large, it simply has more aesthetic energy than the same image does when small.

Just think of how you felt in the presence of an exceptionally large thing: a huge skyscraper, a giant ocean liner, a crowded football stadium, a towering redwood tree, a majestic mountain, or the Grand Canyon. You may not have felt smaller physically than in any other environment, but the grandiosity of the object probably overwhelmed you somehow. The mere energy of enormousness can inspire awe. Large things do not seem to be as manageable as small things; you have less control over them, which makes them appear more powerful.[11]

When you watch a documentary on a small standard (4 × 3) video screen and then see it on an equally small wide-screen (16 × 9) receiver, you probably notice relatively little energy change. When you switch from the small screen to a large HDTV screen, however, the energy change is readily apparent. This is why some movies that emphasize landscape (from actual landscapes to spaceships or battle scenes) must be seen on the large screen to feel the total impact. Even if you use proper conversion methods for aspect ratios, squeezing such large images into the small video screen reduces not only image size but also, if not especially, event energy. Close-ups, inductive sequencing, and a dense audio track help generate some aesthetic energy on the small video screen, but they cannot compete with the large movie images and high-volume surround sound.

SUMMARY

This chapter explored the two-dimensional field—the area of the video and film screens and some of the basic structural elements and characteristics we confront in this field. The video or film screen provides us with a new, concentrated living space—an aesthetic field in which to clarify and intensify an event. We examined the structural factors of screen space: aspect ratio and the aesthetics of object size and image size.

Contrary to the picture area in painting and still photography, which can have any shape and orientation, the standard video and film screens are rectangular and horizontal. The classic film screen and the standard video screen have a 4 × 3 (1.33:1) aspect ratio: the screen is 4 units wide by 3 units high, regardless of size. This aspect ratio permits easy framing of horizontally as well as vertically oriented objects or scenes. Most HDTV presentations and motion pictures are now shot and projected in a horizontally stretched format. The aspect ratio for HDTV is 16 × 9 (1.78:1); for standard motion picture projection, it is 5.55 × 3 (1.85:1).

Standard video and movie screens are horizontally oriented because our peripheral vision is greater horizontally than it is vertically, and in keeping with our everyday experience of living on a horizontal plane. The aspect ratios of small mobile video displays can be horizontal or even vertical in a great variety of aspect ratios.

There are several, albeit not always satisfactory, solutions to making a standard 4 × 3 presentation fit the 16 × 9 screen and vice versa. Windowboxing is to center the smaller screen within the actual video screen. The window can have any format but is usually 4 × 3 or 16 × 9. Letterboxing is making a 16 × 9 image fit a 4 × 3 screen by masking the unused areas on top and bottom of the screen with horizontal stripes, sometimes called dead zones. Pillarboxing is used to make a 4 × 3 picture fit a 16 × 9 screen, leaving the vertical stripes on either side of the screen.

With digital special-effects equipment in video, we can easily generate secondary frames that have various aspect ratios and configurations. In film we can either black out the sides of the screen to reduce the horizontal stretch or fill in the sides with scenic objects to change the aspect ratio more unobtrusively.

The aesthetics of size includes the perception of object size when shown on the screen. Our perception is guided by size constancy, which means that we perceive people and their environments as normal sized regardless of whether they appear in a long shot or a close-up on a large movie screen or a small video screen, or whether we are relatively close to or far away from the screen. The major clues to the actual size of an object when shown on-screen are knowledge of the object, relation to the screen area, and contextual scale.

When we know an object, we perceive its screen image as actual size. When we don't know the object, we tend to translate the image size into actual object size as perceived on-screen: the more screen area the object takes up, the larger it seems and vice versa. When several objects are shown on-screen, we establish a size standard and judge all other objects accordingly. We also use the perceived distance of an object for establishing scale. The most common size reference for objects is a human being.

When shown on the movie screen, the image size of an object is many times larger than when shown on even a large video screen. Because of size constancy, we perceive the objects as equal in size. Nevertheless, the large movie image usually carries much more aesthetic energy than does the small video image. People as well as things attain spectacular dimensions on the wide screen. On the standard video screen, human actions gain prominence whereas the mere spectacle of things is de-emphasized.

The small display area of mobile video media can show landscape movies but fails to communicate the emotional impact and the aesthetic energy. When a mobile video device is used as a companion, these major aesthetic factors need to be considered in creating content: close-ups, inductive sequencing, brief running times, and a dense audio track. Despite these production factors, however, it is virtually impossible to generate the aesthetic energy of large images on the tiny mobile video display.

NOTES

1. Boris Uspensky, *A Poetics of Composition,* trans. by Valentina Zavarin and Susan Wittig (Berkeley: University of California Press, 1973), p. 143.
2. When you're standing still, your normal peripheral vision is considerably wider on a horizontal plane (about 180 degrees) than on a vertical one (about 70 degrees). Expressed in a rectangular aspect ratio, this comes to approximately 2.53:1. When moving, however, our peripheral vision diminishes rapidly. See Arthur Seiderman and Steven Marcus, *20/20 Is Not Enough* (New York: Alfred A. Knopf, 1990), pp. 112–13, 142–43.
3. Such eyeball jumps are called saccadic movement. James J. Gibson says that it is the combination of saccadic movement and temporary fixation that constitutes the act of scanning. See James J. Gibson, *The Ecological Approach to Visual Perception* (Hillsdale, N.J.: Lawrence Erlbaum, 1986), pp. 211–13.
4. John Belton, *Widescreen Cinema* (Cambridge, Mass.: Harvard University Press, 1992), pp. 15–18. See also Mark Schubin, "Searching for the Perfect Aspect Ratio," *SMPTE Journal* (August 1996): 460–78; and Mark Schubin, "No Answer," *Videography* (March 1996): 18–33.
5. Peter Ward, *Picture Composition for Film and Television,* 2nd ed. (Boston and London: Focal Press, 2003), pp. 90–105. Ward argues eloquently against the "spectacle" idea and considers it a throwback to the 1950s. He suggests that the wider-screen aspect ratio was developed simply "for its ability to engage the audience."
6. Steven D. Katz, *Film Directing: Shot by Shot* (Studio City, Calif.: Michael Wiese Productions, 1991), p. 215.
7. Ward, *Picture Composition,* p. 83.
8. There are several projects under way to develop a technique that can adjust, or "repurpose," video productions and films to fit various screen aspect ratios and sizes. For how aesthetic decisions are made in such a process, see Chitra Dorai and Svetha Venkatesh (eds.), *Media Computing: Computational Media Aesthetics* (Boston: Kluwer Academic, 2002).
9. E. H. Gombrich, *The Image and the Eye* (Ithaca, N.Y.: Cornell University Press, 1982), p. 198.
10. "Content snacking" was coined by Dr. Shani Orgad, a media scholar at the London School of Economics and Political Science. See Shani Orgad, *This box was made for walking: how will mobile TV transform viewers' experience and change advertising?* Nokia Mobile TV report 2006, no. 2519. Available at *www.nokia.com/NOKIA_COM_1/Press/Press_Events/mobile_tv_report,_november_10,_2006/Mobil_TV_Report.pdf.*
11. See Alva Noë, *Action in Perception* (Cambridge, Mass.: MIT Press, 2004), p. 105.

7

The Two-dimensional Field: Forces Within the Screen

In our effort to tame space aesthetically, that is, to create a picture space, you should now consider the video, computer, and motion picture screens a new field of operations. Within this fixed and clearly defined space operate specific field forces that are quite different from those of an undefined space, such as our actual three-dimensional environment. To clarify and intensify events within this new operational field, you must understand these major field forces: main directions, magnetism of the frame, asymmetry of the frame, figure and ground, psychological closure, and vectors.

Main Directions: Horizontal and Vertical

One of the most basic field forces is how we perceive and structure events that occupy primarily horizontal and vertical screen space. Because we ordinarily move about in a horizontal world, and also spend a great portion of our lives in a horizontal position while sleeping, a horizontal placement within the screen, and horizontal lines, seems to suggest normalcy, calm, tranquility, and rest. **SEE 7.1** Vertical space, on the other hand, is harder to manage. Thus vertical lines seem more dynamic, powerful, and exciting than horizontal ones. **SEE 7.2**

These aesthetic principles have been applied throughout civilization. The Gothic cathedral, for example, was built with a vertical orientation to direct the human experience upward, toward heaven and God. **SEE 7.3** Renaissance buildings, in contrast, emphasized the importance of human endeavors and are therefore principally horizontal in orientation. **SEE 7.4**

Today's high-rise buildings operate much in the spirit of the Gothic cathedrals. Though their vertical orientation may have been primarily motivated by the high cost of real estate, they nevertheless reflect the contemporary human spirit—adventurous with an admirable earth-defying dynamism. **SEE 7.5**

Because of gravity and the fact that we are used to standing upright on level ground, we like to see our environment mirror our experience and be portrayed as a stable series of horizontals and verticals. In fact, most of our natural or constructed physical environment is organized into verticals that are perpendicular to the level ground. **SEE 7.6**

7.1 Main Direction: Horizontal
Horizontal lines suggest calm, tranquility, and rest. We feel normal when operating on this familiar horizontal plane.

7.2 Main Direction: Vertical
Vertical lines are more powerful and exciting than horizontal ones. Their defiance of gravity charges them with extra energy.

7.3 Vertical Orientation: Gothic Cathedral
The extreme vertical orientation of Gothic cathedrals and their imposing size were designed to remind people of their insignificance relative to God and to direct their spirit upward toward heaven.

7.4 Horizontal Orientation: Renaissance Building
The horizontal orientation of Renaissance buildings appropriately reflected people's renewed interest in human affairs.
Courtesy of the DeBellis Collection, San Francisco State University.

In the Gothic period (roughly from 1150 to 1400), life on earth was simply a preparation for the "real life" in heaven. People lived and worked strictly *in dei gloriam*—for the glory of God. The vertical orientation of the Gothic cathedrals reflects this mind-set.

The Renaissance (roughly from 1400 to 1600), with its rebirth of classical Greek architecture, emphasized the importance of humanity. *In dei gloriam* was supplanted by *in hominis gloriam*. Horizontal buildings are in keeping with this new attitude of glorifying the human spirit.

TILTING THE HORIZONTAL PLANE

Our sense of vertical and horizontal accuracy is so keen that we can, for example, judge whether a picture hangs straight or crooked with uncanny precision even without the aid of a level. It's no wonder that when we see a tilt to the horizontal plane within the screen, we become somewhat disturbed if not disoriented. Our normal and hence secure upright position on a level horizontal plane is threatened by what we perceive. As the horizon starts tilting, we lose our usually reliable and stable reference: the earth. When the horizon tilts, we immediately seek a new and more stable reference regardless of whether it makes sense. For example, when sitting in an airplane that banks sharply, we tend to assign stability to our immediate environment—the airplane—and not to the earth. Consequently, the horizon, rather than the airplane, seems to be doing the tilting.

Lacking a new stable reference, such a tilting effect may cause in us considerable psychophysical discomfort. For example, when we

7.5 Vertical Orientation: High-rise Building
Modern high-rises, like their Gothic cousins the cathedrals, are bold and earth defying. They mirror the dynamic spirit of contemporary people.

7.6 Horizontal/Vertical Environment
A combination of horizontals and verticals reflects our everyday world. Most of our normal physical environment—the buildings we live in, our furniture, doors, and windows—are verticals perpendicular to level ground.

7.7 Level Horizon: Stability
If you want to emphasize stability, conservatism, and reliability, you must show the buildings standing on level ground.

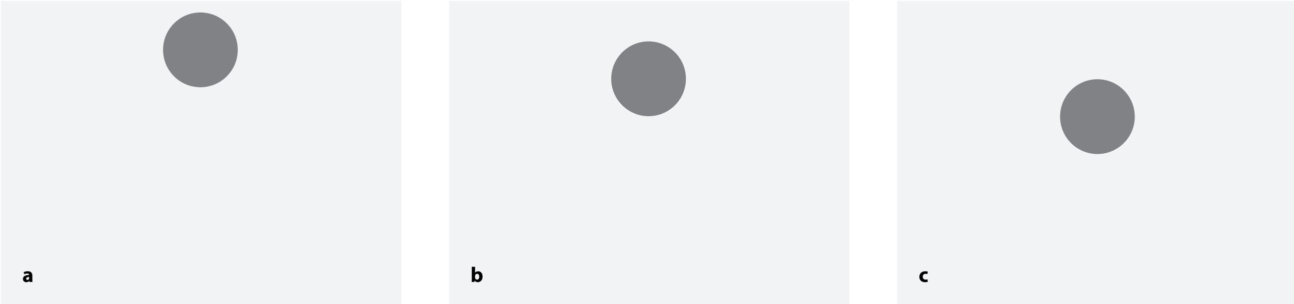
7.8 Tilted Horizon: Dynamism
If you want to suggest energy, activity, and progress, you can render the scene dynamic by tilting the horizontal plane.

sit close to a large movie screen, the tilted horizon effect can, in extreme cases, cause nausea.

Simply by canting the camera and thus tilting the horizontal plane, you can easily destabilize a scene or make an otherwise uninteresting building or object look dynamic. SEE 7.7 AND 7.8 You can also suggest people's extreme physical or mental stress by having them operate on a tilted horizontal plane. SEE 7.9

There is no video of a rock concert that does not have the horizon tilt at least a few dozen times. Because this adds intensity to the already high-energy scene, such an aesthetic device is usually justified. If it is indiscriminately applied to news or documentary productions, however, tilting the horizon often proves counterproductive. A tilted horizon will not liven up a dull interview or help a wooden political candidate get elected.

7.9 Tilted Horizon: Stress
By tilting the horizontal plane, you can intensify the feeling of physical or mental stress of someone running from danger.



7.10 Magnetism of the Frame: Top

The edges of the screen exert a strong pull on objects near them. The disc in (a) seems to be glued to the top edge despite the normal downward gravitational pull. The pull of the frame is strong enough to hold the disc up even when there is more space between the disc and the edge (b). Only when the distance between the top edge and the disc reaches a certain point does it become so great that gravity takes over (c).

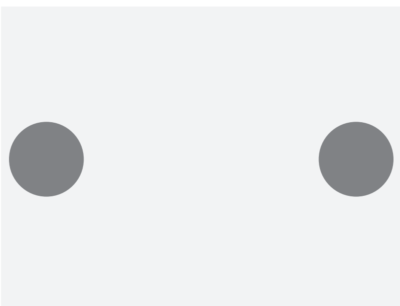
Magnetism of the Frame

The borders of a picture field act like magnets: they have a tendency to attract objects near them. This ***magnetism of the frame*** can be so strong that it counteracts our natural reaction to gravitational pull. **SEE 7.10** Note how the disc seems to be glued to the upper border in figure 7.10a. The gravitational pull comes into play only after the disc is a considerable distance away from the upper border and its magnetic attraction (figure 7.10c).

The sides of the screen also exert a strong pull. **SEE 7.11** As you can see in figure 7.11, the discs do not seem to be pulled down by gravity as you might expect; instead they are attracted by the magnetism of the frame's sides. Because the screen corners combine the magnetism of two sides, they exert an especially strong pull. Obviously, you should avoid compositions whose dominant lines lead directly to the corner of the screen. **SEE 7.12**

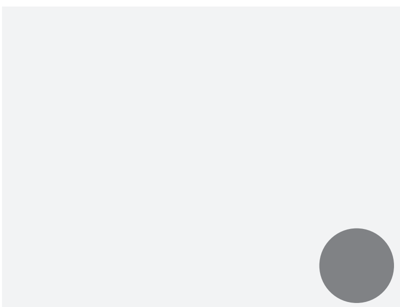
The most stable position for the disc is clearly screen-center. Here it is farthest away from the magnetic pull of the edges; and the force of the pull, however weak, is equally distributed. **SEE 7.13**

If the disc is large and wedged into the screen, it is subject to the magnetism of all four edges. We have the feeling that it wants to burst out of the frame's confinement and expand. **SEE 7.14** Of course, the magnetism of the screen edges is



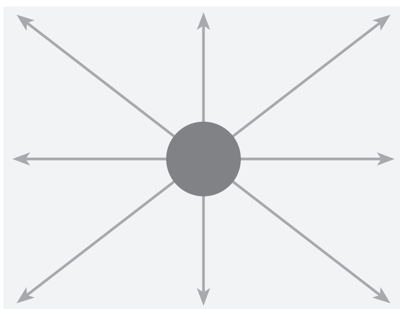
7.11 Magnetism of the Frame: Sides

The pull of the side edges is so strong that it easily overrides gravitational pull.



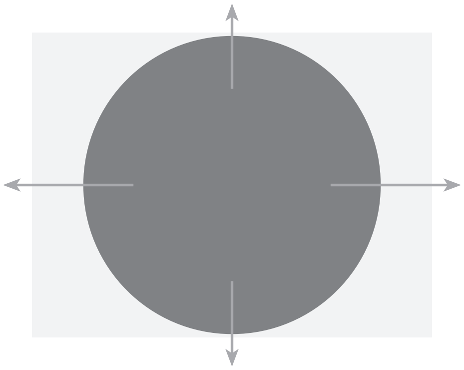
7.12 Magnetism of the Frame: Corners

The screen corners exert an especially strong magnetic pull.

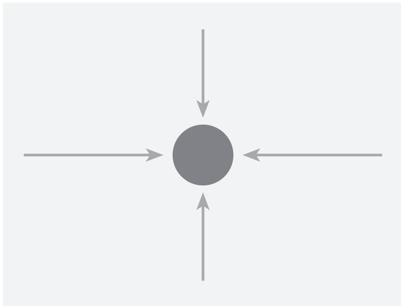


7.13 Neutralized Magnetism: Screen-center

When the disc is centered within the frame, all magnetic pulls are equalized. The disc is at rest.



7.14 Large Object: Expansion
When an object is wedged into the screen, it is subject to the magnetism of the entire frame and therefore tends to expand and look unusually large.



7.15 Small Object: Compression
When the object is small enough and relatively far from the screen edges, the magnetism of the frame no longer operates. On the contrary, the space surrounding the object compresses it to make it appear quite small.

not the only reason why we perceive the disc as unusually large. It also occupies a relatively large screen area, which, as discussed in chapter 6, indicates a large object.

Conversely, when the small disc is centered within the picture field and far from the screen edges, it is no longer subject to their magnetic pull. The concentrated "heavy" negative space surrounding the object also seems to compress it, making it appear smaller than it really is. **SEE 7.15** This principle was also applied in figures 6.36 and 6.37.

Let's see how to work with the magnetism of the frame when composing a shot.

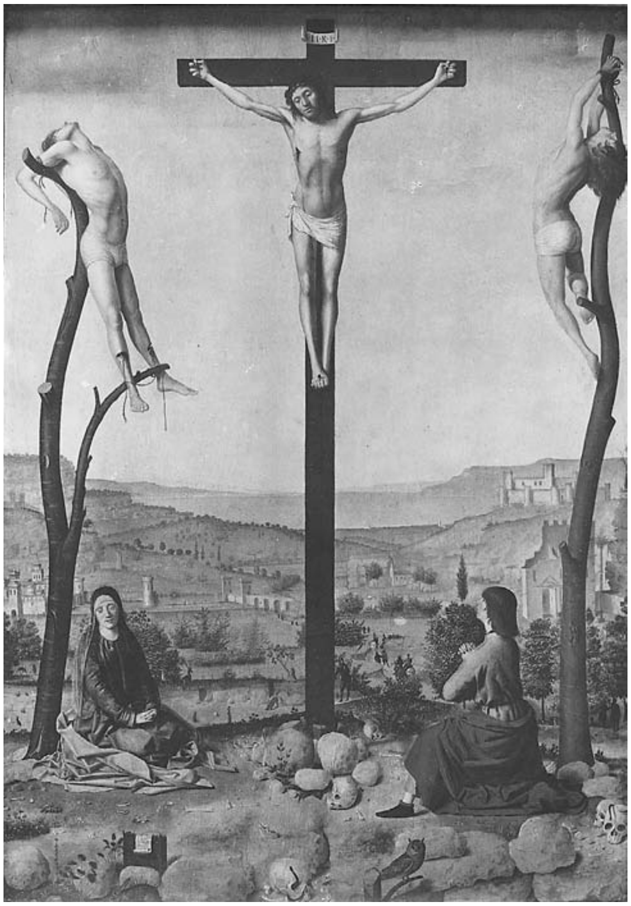
HEADROOM

A close-up with too little space between the top of the head and the upper edge of the frame looks awkward. Why? Because the head seems to be glued to the upper screen edge. We are as uncomfortable with such a composition as we are standing up in a room with a very low ceiling. **SEE 7.16** You need to leave some "breathing space," called ***headroom***, to counteract the magnetic pull of the upper screen edge. **SEE 7.17** But if you leave too much headroom, the head is pushed down into the lower half of the frame through the combined forces of gravitational pull and the magnetism of the lower screen edge.[1] **SEE 7.18**

7.16 No Headroom
Without headroom the attraction of the upper screen edge pulls the head firmly against it.

7.17 Proper Headroom
To counteract the pull of the upper screen edge, you must leave enough headroom.

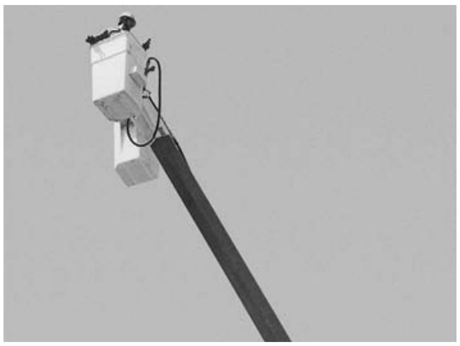
7.18 Too Much Headroom
Here you have too much headroom. The magnetism of the frame has now capitulated to gravitational pull, so the shot looks bottom-heavy.



7.19 Using the Pull of the Top Edge

The old masters used the magnetism of the frame to intensify their messages. By reducing headroom to a minimum, the figures seem to be more suspended than if they were lower in the painting.

Antonello da Messina (c. 1430–1479) *Crucifixion.*/Koninklijk Museum voor Schone Kunsten, Antwerp, Belgium/Scala/Art Resource, NY.



7.20 Lack of Headroom for Intensification

To emphasize the precariousness of somebody working up high, you can use the magnetism of the upper edge to advantage. In this shot the minimal headroom maximizes the pull of the upper screen edge.

How much headroom should you leave? Just enough to neutralize, but not eliminate, the magnetic pull of the upper screen edge. But just where is that? Technically, your headroom framing will be approximately correct if you place the eyes of the person in the upper third of the frame's height. Unfortunately, you will have little opportunity during a production to pull out a tape measure to pinpoint the upper third of the frame. You are better off learning how to judge it intuitively.

PULL OF THE TOP EDGE

Leaving headroom is not mandatory, however. Under certain circumstances you can use the magnetic pull of the upper edge to your advantage. You can learn more about the pull of the top edge of a frame by studying the paintings of old masters. For example, in his painting *Crucifixion,* Antonello da Messina placed the figures, especially that of Christ, so close to the upper edge that its magnetic pull comes into action and heightens the agony of hanging from the cross. **SEE 7.19**

In a similar way, you can emphasize the precariousness of someone operating at uncomfortable heights by deliberately leaving inadequate headroom. This lack of space will pull the person to the upper edge and make him hang in space more than if you had framed him with the customary headroom. **SEE 7.20** You can also stretch aesthetically a fashion model by wedging her between the top and the bottom edges of the frame.

PULL OF THE SIDE EDGES

The major problem that occurs as a result of the magnetic pull of the side edges is that objects that seem normally spaced in a long shot will often look too far apart in a tighter shot. You can see this in an establishing shot of an interview, when the host and the guest seem to be sitting at a comfortable distance from each other. **SEE 7.21** But as soon as you move in with the camera for a tighter shot, the two people now look too far apart, seemingly glued to the left and right edges of the screen. **SEE 7.22** They have become part of the screen edge and created a secondary frame, emphasizing the empty center portion of the picture. This problem is similar to that of framing two people facing each other, as shown in figure 6.17.

When working in the 4 × 3 aspect ratio, this magnetism problem of pulling people apart is ever-present. One way of solving it is to use over-the-shoulder shots or cross-shots for two people talking to each other (see figures 11.22 to 11.25). In the wider 16 × 9 aspect ratio, however, this problem is greatly reduced or even eliminated because you have enough space between the people and the screen edges to neutralize the magnetism of the frame.

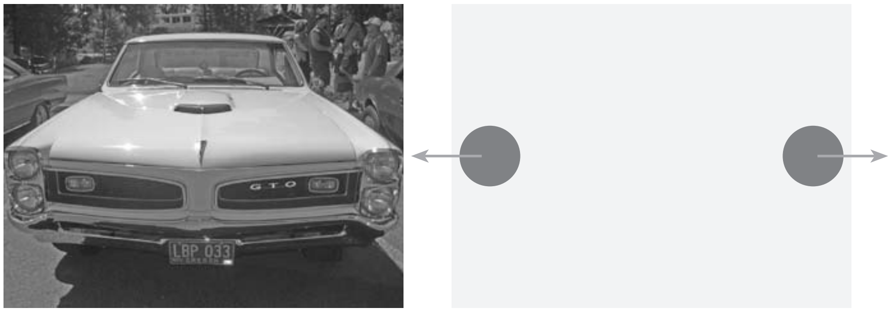
7.21 Normal Distance in Long Shot
In this establishing shot, the people seem to be sitting at a comfortable distance from each other.

7.22 Edge Pull in Medium Shot
When you move in for a tighter shot, however, the people are pulled apart. The magnetism of the side edges intensifies this separation.

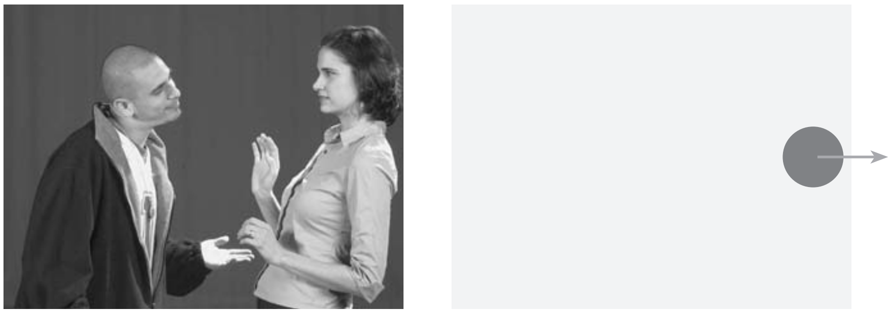
Fortunately, the pull of the sides is not all bad and can work for, rather than against, effective composition. For example, when you want to emphasize the width of an object, the pull of the screen's sides becomes a definite graphic asset. The combined pull of the left and right edges seems to stretch the old road cruiser horizontally, epitomizing the pride of its owner for having such a big car. **SEE 7.23** You can also use the magnetic force of a screen edge to arrest motion and keep someone rooted to the spot. **SEE 7.24**

You encounter a similar problem if you hang pictures at a "normal" distance from one another on a set. The edges of the screen attract the pictures, so they appear pulled apart.

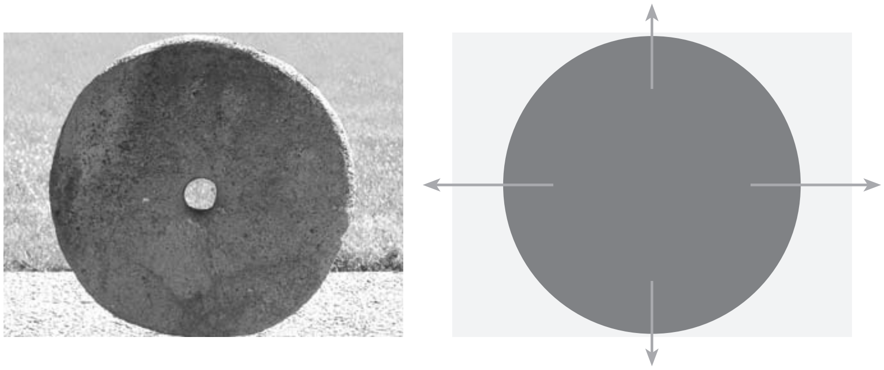
To make pictures look normally spaced in a medium shot, you must crowd them. The attraction of graphic mass among the pictures then counteracts, at least to some extent, the pull of the frame.



7.23 Positive Pull of Side Edges
When you want to emphasize the width of a large vintage car, the pull of the screen edges becomes a definite graphic asset.



7.24 Positive Pull of Side Edge
Here the pull of the frame works to your advantage. To emphasize visually that the woman has little chance of avoiding this man, she is pinned against the right screen edge.



7.25 Positive Pull of All Screen Edges
The magnetism of all four edges helps emphasize the enormity and the weight of this millstone. It begs to burst out of the confinement of the frame.

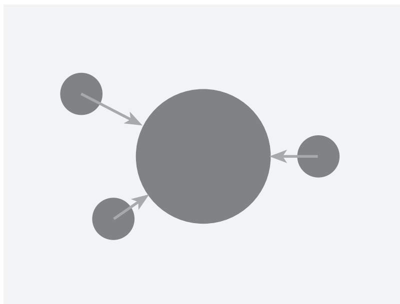
PULL OF ENTIRE FRAME

When an object is wedged into the screen so that all four edges are allowed to pull, the object seems to be as uncomfortable as we are when looking at it. Its confinement makes it look bigger and heavier, which can also be a positive effect. For example, the weight of the millstone in figure 7.25 is properly intensified through its tight framing. **SEE 7.25** A similar thing happens when you frame a close-up of a head in this way. Because of the magnetic pull of the frame, the head seems unusually large, as if it's trying to burst out of the screen. Obviously, you should not frame a head in this fashion.[2] **SEE 7.26**

ATTRACTION OF MASS

All screen images have a certain ***graphic mass***. Usually, larger images with highly saturated colors have a larger graphic mass than smaller ones with less saturated colors. The larger the graphic mass, the greater its ***graphic weight***, that is, the “heavier” (more prominent) the image seems within the frame. Graphic mass attracts graphic mass. The larger the mass, the greater its attractive power. Also, a larger graphic mass attracts smaller ones and not vice versa. Finally, a larger graphic mass is more stable and less likely to move than smaller ones. **SEE 7.27**

7.26 Negative Pull of All Screen Edges
Wedging this women's head into the frame intensifies the perfect oval shape of the head but also makes it appear bulky.



7.27 Attraction of Mass
Graphic mass attracts graphic mass: the larger the mass, the larger its attractive power. The larger graphic mass attracts the smaller one and not vice versa. A large mass is more independent than smaller masses. The larger graphic mass is also more stable (less likely to move) than the smaller ones.

Asymmetry of the Frame

We do not seem to look at the left and right sides of any screen in the same way, nor do we pay equal attention to objects that are located on either half of the screen. We carry this perceptual peculiarity over to the video, film, and computer screens, where we consider it one of the structural screen forces. The two sides of the screen seem structurally unequal. This concept is known as ***asymmetry of the frame***.

UP/DOWN DIAGONALS

Take a look at the following figures. **SEE 7.28 AND 7.29** Which of the diagonals seems to go up and which seems to go down?

7.28 Diagonal 1
Does this diagonal suggest an uphill or a downhill slope?

7.29 Diagonal 2
Does this diagonal suggest an uphill or a downhill slope?

Now look at the next set of figures. **SEE 7.30 AND 7.31** Which diagonal do you now perceive to go up and which down?

7.30 Diagonal 3
Does this diagonal suggest an uphill or a downhill slope?

7.31 Diagonal 4
Does this diagonal suggest an uphill or a downhill slope?

You probably chose diagonals 1 and 4 as going up and diagonals 2 and 3 as going down. Why we perceive a specific diagonal as going uphill and another as downhill is anyone's guess. Regardless of whether we learned to write from left to right, right to left, or top to bottom, or whether we are right- or left-handed and right- or left-brained, we seem to "read" the diagonals from left to right.

7.32 Uphill Motion Aided by Uphill Diagonal
The uphill diagonal helps pull the car to the top of the hill.

7.33 Uphill Motion Impeded by Downhill Diagonal
This truck has to overcome the graphic down-diagonal, which therefore emphasizes the truck's power.

Although any movement along either slant can override this graphic up/down sensation, you can nevertheless use the up/down slants to intensify motion along the diagonals. For example, if you would like to show the ease with which a car moves up a hill, you should have the car go on a diagonal from lower left to upper right. This way the uphill diagonal helps pull the car to the top of the hill. **SEE 7.32** On the other hand, off-road vehicles, trucks, or bulldozers seem to need more power and effort to climb a hill from right to left than from left to right because they now have to overcome the natural flow of the graphic downhill slant. **SEE 7.33**

If you want to intensify the apparent speed of a runaway truck, have it move along the graphic downhill diagonal (left to right) instead of against the uphill diagonal from right to left. But wouldn't we see the truck race downhill no matter what? Of course, but we are dealing with the finer points of media aesthetics to intensify a scene, however subtle such intensification may be.

SCREEN-LEFT AND SCREEN-RIGHT ASYMMETRY

Quickly glance at the advertisement in the next figure. **SEE 7.34** What was the woman holding? When you take another look at the picture, do you have a tendency to focus more on the product than the woman? Now look at the next figure. **SEE 7.35** Yes, it is a flopped image except for the lettering. But did you now pay just a little more attention to the woman holding the box than to the box itself? If you did, this is a common perceptual response.

Even if the screen is not divided symmetrically, we tend to pay more attention to the right side. Despite the considerable academic controversy surrounding this aesthetic phenomenon,[3] we can use the up- and downhill diagonal as a reasonable explanation of why the right picture area seems to be more conspicuous than the left. In practice this means that if you have a choice, you should place the more important event on the right side of the screen. If, in an interview show, you consider the guest more important than the host, place the guest screen-right and the host screen-left. Most prominent hosts, however, do not want to be upstaged, so they occupy the more conspicuous screen-right position.

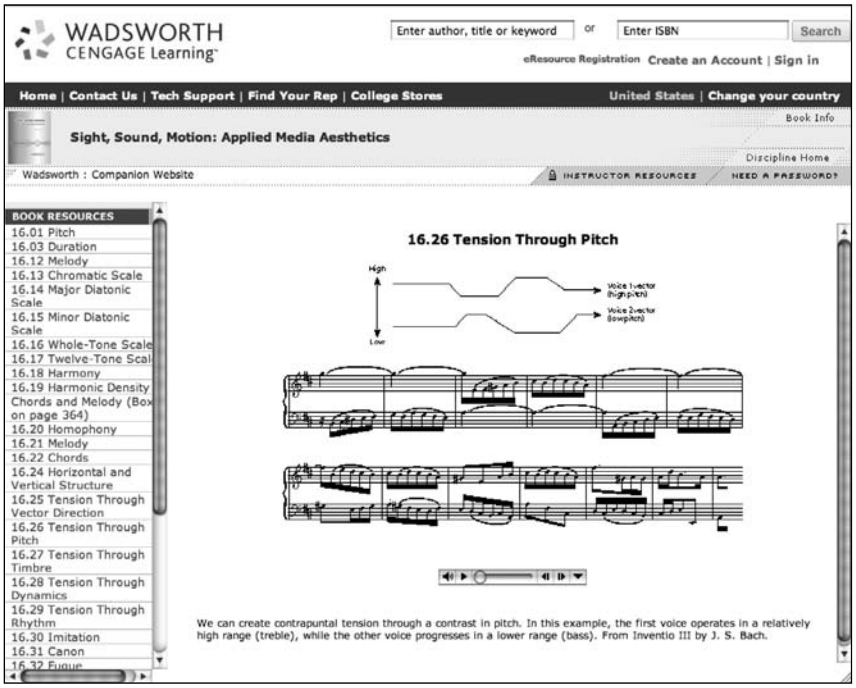
Why, then, do we see so many newscasters on screen-left with the less important graphic information of the box on screen-right? Should you not put the primary information source—the newscaster—on screen-right and leave the screen-left side for the computer-generated box? From an aesthetic point of view, the picture is more balanced when the high-energy information source, in this

7.34 Emphasis Through Screen Asymmetry
At first glance do you pay more attention to the person or the object?

7.35 Shift of Emphasis
When the picture is flopped, do you still see it in the same way, or has your attention shifted from the object to the person?

case the newscaster, appears on the weaker screen-left side and the lower-energy graphic on the stronger screen-right side. From a communication point of view, however, the primary information source—the newscaster—should definitely be on screen-right, with the illustrative material relegated to the weaker left side. Otherwise we run the risk of having, however subconsciously, the often crude and cryptic graphic representation of the news stories accepted as the primary message rather than the more-detailed accounts by the newscaster.

The asymmetry principle is especially important when designing Web pages. Generally, you should keep the navigation instructions on screen-left and the text (including essential illustrations) on screen-right. Because the navigation information is one of the primary interactive tasks to access the desired content, it does not need to be favored by the right side. The content itself, however, needs to be the focus of attention and is therefore best placed on the stronger screen-right side. **SEE 7.36**



7.36 Asymmetry on Web Page
Because the right side of the screen commands more attention, keep all essential content on the right. Put the navigation information on the less prominent left side.

You probably notice that the asymmetry principle is dependent on screen size. On the tiny mobile media display, it does not seem to affect even our subconscious perception whether the newscaster or talk-show host is placed screen-left or screen-right or where the text is placed relative to the navigation instructions.

Figure and Ground

As is our everyday perception, one of the most elemental structural forces operating within the screen is the ***figure/ground principle***. As discussed in chapter 1, our perceptual tendency is to organize our environment into stable reference points (usually the background) against which we can assess the less-stable elements (the figures in front of it). Organizing our environment into a figure/ground relationship is one of our most fundamental perceptual activities and is so automatic that we are usually unaware of it. It helps stabilize our environment and make it manageable.

The letters on this page and the illustrations appear as figures, with the page as the ground. If your book rests on a table, the book becomes the figure and the supporting table is the ground. As you can see, the figure/ground relationship is contextual and hierarchical. Depending on what you determine the figure to be, the ground will change accordingly. Contrary to the static and fixed figure/ground relationship in a painting or still photograph, the on-screen relationship changes, as it does in real life, with the camera's point of view. Thus the figure in a medium shot can become the ground in a close-up. **SEE 7.37 AND 7.38**

FIGURE/GROUND CHARACTERISTICS

What exactly distinguishes a figure from its ground? Look at the well-known figure/ground example on the next page. **SEE 7.39** If you perceive the vase as the figure, the gray area becomes the ground. If you consider the two profiles as the figure, the white area becomes the ground. When you see the vase as the figure, it takes on certain characteristics, as does the gray ground. The characteristics reverse when you perceive the gray profiles as the figure. Here are some of the major characteristics of figure/ground perceptions:[4]

- The figure is thinglike. You perceive it as an object. The ground is not; it is merely part of the "uncovered" screen area.

7.37 Figure/Ground Relationship in Medium Shot
With a moving camera or a changing field of view, the figure/ground relationship alters continuously. In this shot the car is the figure and the bushes are the ground.

7.38 Figure/Ground Relationship in Close-up
When you move in for a close-up, the hood ornament becomes the figure and the car's hood becomes the ground.

7.39 Figure and Ground

We assign figure and ground depending on what we perceive. For example, if you see a white vase here, the gray area becomes the background. But if you see two gray profiles, the white element becomes the ground. Notice how figure and ground switch characteristics, depending on which you assign to be the figure and which the ground.

- The figure lies in front of the ground. The white vase obviously lies on top of the gray background.
- The line that separates the figure from the ground belongs to the figure, not the ground.
- The figure is less stable than the ground; the figure is more likely to move.
- The ground seems to continue behind the figure.

Unless you possess a Zenlike state of mind that allows you to reconcile opposing elements, you will not be able to perceive figure and ground simultaneously. You must opt to pay attention to either the figure or the ground. At best you can oscillate between one figure/ground structure and the other.

Because of this urge to organize our environment into figure/ground relationships, you can achieve especially startling effects by rendering the figure and the ground purposely ambiguous or by reversing the figure/ground relationship. Some artists, like M. C. Escher, created a great variety of drawings and paintings that play with such figure/ground reversals.[5] Modern designers often come up with similarly startling creations by using a figure/ground reversal. SEE 7.40

As you can see, we are coerced by this logo into perceiving parts of the ground as the figure. It is almost impossible not to see the *A* and the *O* as figures

7.40 Figure/Ground Reversal

If the design is simple, a designer can achieve a startling effect by having some of the ground appear as the figure. You will find it hard to consider the *A* and the *O* as the ground rather than a figure.

7.41 Ambiguous Figure/Ground Relationship
Through electronic effects, such as this superimposition, you can change the figure/ground relationship in ways that give the image new meaning.

although they are actually parts of the common ground—the white page. The *A* begins to disappear, however, when you cover up the *E*. All of a sudden, a new pyramid-like figure appears that looks more like the top of a sharp pencil than the holes in the letter *A*.

Superimposition too derives its strength from rendering the figure/ground relationship purposely vague. It is often difficult to determine which of the two superimposed images is the figure and which is the ground. It's no wonder that a "super" is such a popular effect when suggesting dream sequences. SEE 7.41 Through a variety of digital effects, you can reverse the figure and the ground at will and thus create powerful and enigmatic images.

We are so used to seeing moving figures relative to a stable ground that we maintain this relationship even if the ground moves against a stationary figure. We can easily trick the viewer into perceiving a car racing through downtown streets simply by shooting the stationary car against a rear projection or a computer-generated background of moving streets. We discuss this figure/ground reversal more thoroughly in chapter 13.

The figure/ground principle also applies to the structuring of sound. We usually try to establish specific foreground sounds against more general background sounds. Figure/ground relationships in sound structures are explored more fully in chapter 15.

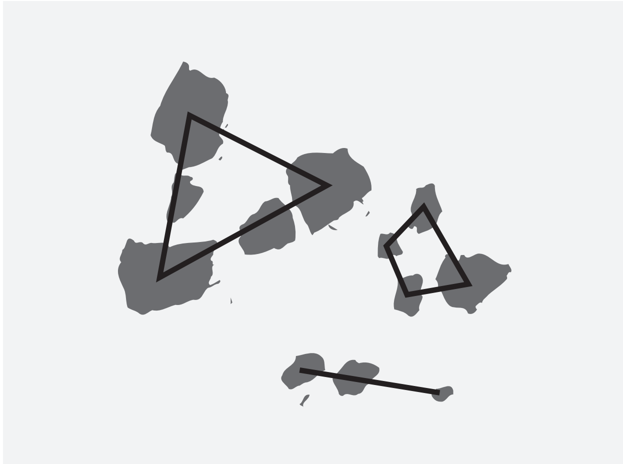
Psychological Closure

In our quest for perceptual sanity, if not survival, we continuously seek to stabilize our infinitely complex and often chaotic environment. One of our built-in survival mechanisms is our tendency to mentally fill in gaps in visual information to arrive at complete and easily manageable patterns and configurations. This perceptual activity is called ***psychological closure***, or closure for short.[6]

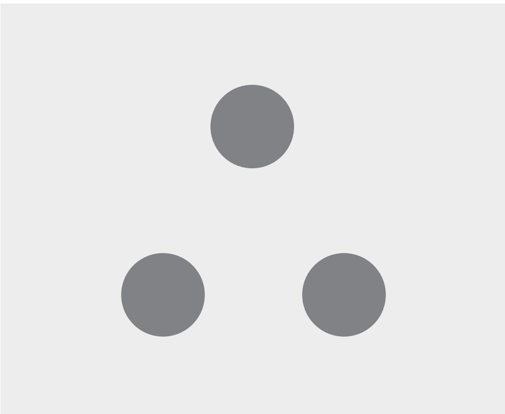
Take a look at the random shapes on the next page. SEE 7.42 As soon as you see them, your perceptual mechanism automatically tries to group them into some sort of simple pattern. What patterns do you see in figure 7.42? Don't force yourself—let the pattern come to you. Now use a pencil to trace the most obvious connections.

You probably arrived at a series of relatively simple and stable figures, such as a triangle, a trapezoid, and a line, by taking a minimum number of visual cues and mentally filling in the missing information—by applying closure. SEE 7.43

7.42 Perceiving Patterns
Look at these random shapes. You will probably find it hard not to group this random array into some kind of order.



7.43 Psychological Closure: Pattern
Most likely, you organized the shapes into a large triangle at the upper left, a trapezoid at the right, and a line at the bottom.

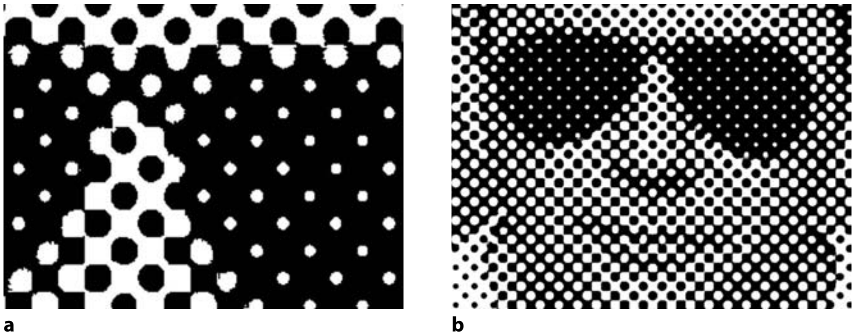


7.44 Psychological Closure: Subjective Completion
Although the page shows only three dots, you must try hard *not* to apply psychological closure and perceive a triangle.

Let's repeat this process in a much simpler visual field. SEE 7.44 Although the figure shows only three dots, you automatically perceive a triangle; in fact, it takes considerable effort *not* to see a triangle. As you can see, through applying closure you constructed a new stable pattern—a triangle—from a minimum number of clues—three dots.

GESTALT

The pattern that results from applying psychological closure is often called a ***gestalt*** (German for "form," "configuration," or "shape"). A gestalt is a perceptual whole that transcends its parts. As soon as you perceive a triangle (the whole) instead of three dots (the parts), the dots have become part of the gestalt; they are no longer independent elements but have become part of a triangle. This is why *gestalt* is often defined as a whole that is larger than, or at least different from, the sum of its parts.[7] The relationship of the parts to the whole becomes especially apparent in music. Three notes played in sequence sound quite different when played simultaneously as a chord. The individual elements (the three notes) have now been molded into a gestalt (the chord) and have thus surrendered their individual functions to that of the new structure.



7.45 Minimum Information

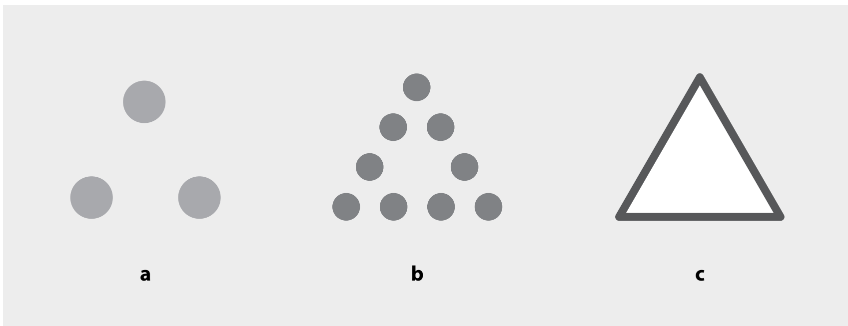
You need a minimum amount of information to perceive a meaningful pattern. Notice how the seemingly random dots (a) suddenly make sense when you add enough information (b).

Note that to apply psychological closure we need a minimum amount of information. If the information falls below the required minimum, the stimulus elements remain random, and we cannot perceive a pattern. For example, three dots are the minimum number of elements needed to perceive a triangle. One dot less would produce a line but never a triangle. As soon as we have enough information, however, the pattern becomes immediately and inevitably apparent, and we experience an almost instant switch from chaos to order. **SEE 7.45**

HIGH- AND LOW-DEFINITION IMAGES

A ***high-definition*** image has more picture information than a low-definition image. **SEE 7.46** Figure 7.46a shows the lowest-definition image of a triangle, figure 7.46b shows a higher-definition image, and figure 7.46c shows the highest. Commercial motion pictures display high-definition images; the large screen lets you see a great amount of event detail, and the film images have high resolution. Similarly, ***high-definition television (HDTV)*** produces images of remarkable clarity.[8]

A ***low-definition*** image requires a great deal more closure than a high-definition image to arrive at a gestalt. In contrast to film and HDTV, the images of STV (standard television, with 480 visible scanning lines) are definitely low-definition. First, their picture resolution is low. Second, even a large television



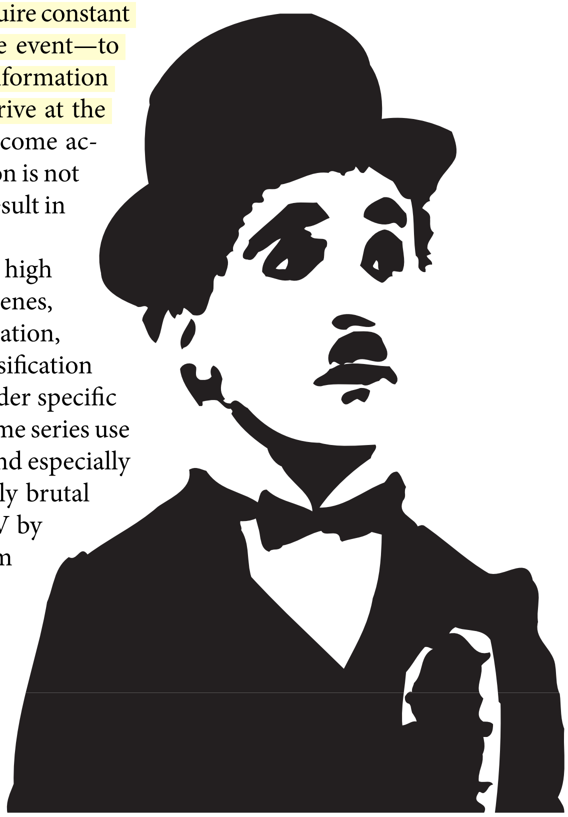
7.46 Low- and High-definition Images

The triangle in (a) has the minimum information: three dots. One dot less and you would perceive a line. The triangle in (b) has more information. You need to apply less psychological closure to arrive at the gestalt of a triangle. The triangle in (c) has the maximum information. You can see its shape without applying any closure.

screen is small compared with a motion picture screen. Third, the contrast ratio—the steps between the lightest and darkest picture areas—is limited. Fourth, to maintain a reasonable level of aesthetic energy, the event details are normally presented as a series of close-ups rather than event overviews, as is frequently the case in motion pictures. Obviously, standard video viewing requires a great deal more psychological closure than watching a film or large-screen HDTV. Because of the miniature screen size, watching video on a mobile media device is unavoidably low-definition, even if the pictures on the small display are touted as high-definition.

Is this innately bad, and should we try to produce images that are maximally high-definition? Not necessarily. Very much like impressionist paintings, standard television derives much of its visual intensity and our perceptual involvement from its low-definition images. Precisely because low-definition images require constant psychological closure, we as viewers are required to *work* with the event—to continuously fill in missing external (form) and internal (story) information to make sense of the picture and story narrative so that we can arrive at the appropriate gestalt. We no longer remain passive spectators but become active, perceptually hardworking participants.[9] Thus the communication is not necessarily inhibited by the low-definition images, which may well result in a desirable intensification of the event.

HDTV practitioners have already discovered that the medium's high resolution is not always a desirable attribute. Especially in intimate scenes, it is often better to soften the image through filters or focus manipulation, very much like the desaturation of color. This participation and intensification aspect of low-definition images has prompted graphic artists to render specific scenes or commercials purposely low-definition. Many television crime series use low-definition effects, such as fast-falloff lighting, color distortion, and especially electronically enhanced high-contrast images, to intensify especially brutal scenes. Some directors soften the sharp "in-the-face" look of HDTV by creating a very shallow depth of field (with long-focal-length zoom positions) or by reducing the saturation of the brilliant HDTV colors, which are otherwise hailed as one of the laudable achievements of high-definition video. Other examples are graphics that carry only minimal information so as to elicit maximum psychological closure. SEE 7.47 Low-definition fashion drawings may encourage you to project how you would ideally like the clothing to look more than high-definition photographs.



7.47 Images Rendered Low-definition
This low-definition image compels you to apply psychological closure. In fact, you will find it hard *not* to see the line of the right side of the face, although it is not actually present.

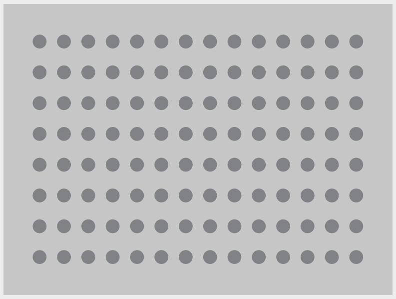
You should realize, however, that the constant mental activity necessary to make sense of a low-definition presentation requires considerable mental effort, however subconscious, on the part of the viewer, which can lead to fatigue. This effort for closure is one of the reasons why you may find it much more tiring to watch a three-hour film on small-screen television than on the large, high-definition screen in a movie theater—even if you have the same amount of popcorn and soft drinks.

FACILITATING CLOSURE

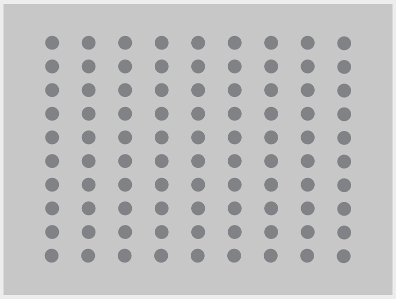
A low-definition image is helpful only if it facilitates, rather than inhibits, closure. When framing a shot, arrange the picture elements in such a way that they can be easily completed in the viewer's mind even in off-screen space; or you can group them into simple figures, that is, into recognizable patterns of basic geometric shapes. This is explored more fully in chapter 8.

We know that people apply psychological closure to formulate specific patterns. But is the process of psychological closure predictable? Can we predetermine the stimuli necessary for someone to perceive a particular pattern? Can we deduce any principles of psychological closure? Yes. Three gestalt psychologists—**Wolfgang Köhler** (1887–1967), **Max Wertheimer** (1880–1943), and **Kurt Koffka** (1886–1941)—helped develop three major principles of psychological closure: proximity, similarity, and continuity.[10]

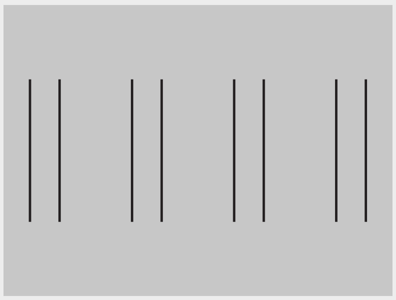
Proximity When similar elements lie in close proximity to one another, we tend to see them together. Because of attraction of mass, we connect more readily those elements that lie closer together than those that lie farther apart.



Here we see horizontal rather than vertical lines because the horizontal dots lie closer together than the vertical dots.

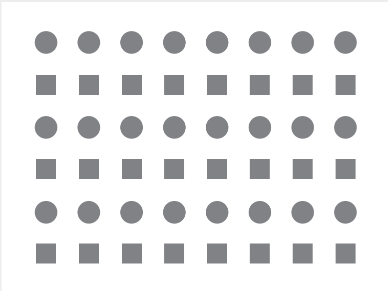


Now we perceive vertical lines because the vertical dots are closer together than the horizontal ones.

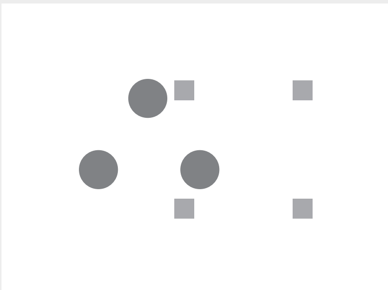


Here we tend to see four narrow columns rather than three fat ones.

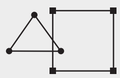
Similarity Similar shapes are seen together.



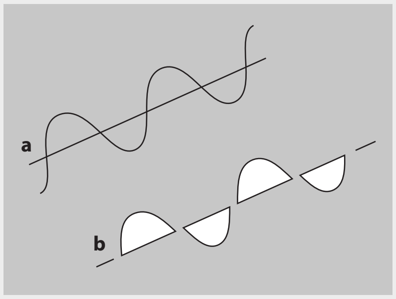
All these dots are equally spaced, yet we see horizontal lines because we tend to see similarly shaped objects together.



Here the similarity overrides the proximity. We see a triangle and a square intersecting.



Continuity Once a dominant line is established, its direction is not easily disturbed by other lines cutting across it.



We see a curved line being intersected by a straight line (a) rather than four oddly shaped forms attached to each other (b).

These three principles of psychological closure—proximity, similarity, and continuity—are all based on our desire to establish ***visual rhythm***. I propose this overriding perceptual principle:

> *We tend to perceive together those elements that are easily recognizable as occurring at a certain frequency (number of similar elements) within a certain interval (distance from one another) or that pursue a dominant line.*

Vectors

Probably the strongest forces operating within the screen are directional forces that lead our eyes from one point to another within, or even outside of, the picture field. These forces, called vectors, can be as coercive as real physical forces.[11] Each vector has a certain magnitude, or strength as to directional certainty and power. A ***vector*** is therefore a force with a direction and a magnitude.[12]

A vector on-screen indicates a main direction that has been established either by implication—such as with arrows, things arranged in a line, or people looking in a specific direction—or by actual screen motion, such as a man running from screen-left to screen-right or toward or away from the camera.

For screen displays, where you must deal with implied as well as real motion, the proper understanding and handling of vectors becomes extremely important.

Once you have grasped what vectors are and how they interrelate and interact with other visual and aural elements, you can use them effectively not only to control screen directions but also to build screen space and event energy within a single frame or over a series of frames. A solid understanding of the vector theory will help you immensely in preproduction placement of cameras and in postproduction editing.

VECTOR FIELD

In structuring ***on-screen space***, we no longer work with isolated vectors but instead with a ***vector field***: a combination of vectors operating within a single picture field (frame), from picture field to picture field (from frame to frame), from picture sequence to picture sequence, from screen to screen (when you use multiple screens), and from on-screen to off-screen events.

You can also find vectors in color, in music, and even in the structure of a story; in fact, a vector is any aesthetic element that leads us in a specific space/time—or even emotional—direction. More complex vector fields include ***external vectors***, which operate within or without the screen, and ***internal vectors***, which operate within ourselves, such as feelings and empathetic responses.

For the present, however, let's concentrate on the visual vectors that operate in the on- and off-screen space of film, video, and computers.

VECTOR TYPES

If you carefully examine the various ways that visual vectors operate, you will discover three principal types: graphic vectors, index vectors, and motion vectors.

Graphic vectors A ***graphic vector*** is created by a stationary element that guides our eyes in a certain direction. It is driven by the principle of continuity. The direction of a graphic vector is ambiguous, however. A simple line is a graphic vector, but we can scan the line from left to right or (with a little more effort) from right to left. When we establish a point of origin for the line, we increase the magnitude of the graphic vector. Its directionality is now determined by a point of origin but is still somewhat ambiguous. **SEE 7.48A AND 7.48B** The pipe structure supporting the roof of a ferry boat pier and the lines of the high-rise building give basic directional orientation, but they do not unequivocally point in a single direction, so their directional magnitude remains relatively low. **SEE 7.49 AND 7.50**

Index vectors **In contrast to a graphic vector, an** ***index vector*** is created by something that points unquestionably in a specific direction. Examples are arrows or people pointing or looking in a particular direction. Index vectors also include

7.48 Line as Graphic Vector

A graphic vector is created by a line or by an arrangement of objects that guides our eyes in a certain direction.

This graphic vector (b) has a more definite direction than those in (a), but it is still not unidirectional.

7.49 Graphic Vectors: Roof Structure
This pipe structure that supports a roof generates a multitude of graphic vectors.

7.50 Graphic Vectors: High-rise Building
The vertical lines of this building create strong graphic vectors. The horizontal lines of the windows are also graphic vectors.

7.51 Index Vector: Pointing
An index vector is created by anything that points unequivocally in a specific direction. Somebody pointing forms a high-magnitude index vector.

7.52 Index Vector: Looking
Somebody looking in a specific direction forms an index vector.

7.53 Index Vector: Arrow
An arrow is a high-magnitude index vector.

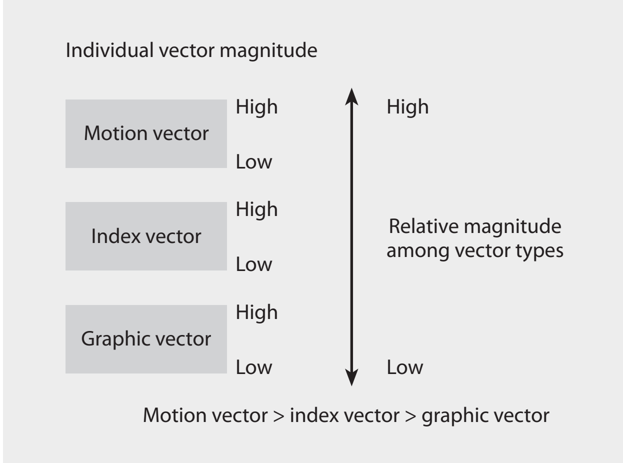
7.54 Index Vector: Motorcycle
Note that a still shot of a person or an object in motion is not a motion vector but rather a strong index vector. Even a blurred photo suggesting movement is still an index vector. A motion vector is created only by something that is actually moving or is perceived as moving on-screen.

still photographs of runners or somebody riding a motorcycle. SEE 7.51–7.54 Note that in figure 7.54 you don't see the motorcycle actually moving (despite the blur indicating motion), so it is an index rather than a motion vector.

Motion vectors A ***motion vector*** is created by an object that is actually moving or seen as moving on-screen. Obviously, a motion vector cannot be illustrated with a still picture in a book. Even if you blur a still photo, as in figure 7.54, or imply motion by keeping the object in focus and blurring the background or foreground, it does not move and is therefore not a motion vector. Imagine that a ladybug has landed on the page you just read and is crawling around figure 7.54. The page with the blurred photo of the motorcycle will inevitably remain the stable ground, and the ladybug is the figure in motion. The importance of this distinction is that motion vectors will command immediate attention, overriding all other vectors.

VECTOR MAGNITUDE

The magnitude of a vector is a product of its relative strength, that is, its directional certainty and perceived directional force. ***Vector magnitude*** is determined primarily by screen direction, graphic mass, and perceived object speed. Although each vector type can be strong or weak (have a high or low magnitude), motion vectors generally have a higher magnitude than index vectors, which in turn have a higher magnitude than graphic vectors. **SEE 7.55**



7.55 Vector Magnitude
Although each vector type can be strong or weak (have a high or low magnitude), motion vectors generally have a higher magnitude than index vectors, which in turn have a higher magnitude than graphic vectors.

Vector magnitude and screen direction As pointed out earlier, a line indicates a general direction (horizontal, vertical, curved) but is not precise as to its direction. Such a graphic vector therefore has a relatively low magnitude. An index vector that is generated by a one-way-street sign, someone pointing in a particular direction, or a directional arrow in a parking garage has a much higher magnitude. Motion vectors have a relatively high magnitude because their screen direction is most conspicuous.[13]

Something moving or pointing directly toward or away from the camera results in a ***z-axis vector***, so called because the vector direction follows the ***z-axis***—the virtual line extending from the camera lens to the horizon. Whereas a ***z-axis index vector***, such as somebody looking or pointing directly at the camera, has a very low magnitude, a ***z-axis motion vector*** can have a high or low magnitude. A racecar hurtling toward the camera definitely has a higher magnitude than the first uncertain steps of a child toddling toward the camera (representing Mother's arms). Depending on the context, a fast zoom-in or zoom-out can also produce a high-magnitude z-axis motion vector, although the screen motion is now implied rather than actual.

In production you simply estimate the relative magnitude of a vector. Because the context in which you judge vector magnitudes is a given, your estimation of which vector has a higher magnitude, or your efforts to lower or raise a specific vector magnitude, can be fairly accurate. But if you analyze a film or video for the specific magnitudes of its vectors, you may need to develop a more accurate and reliable measuring technique.[14]

Vector magnitude and graphic mass The larger the graphic mass that is in motion, the higher its vector magnitude. As stated earlier, a large graphic mass is less likely to be disturbed in its course than a small one. Because its directionality is more certain, its vector magnitude is higher than for small objects.

Vector magnitude and perceived object speed The faster the speed of an object, the higher its vector magnitude. This principle applies to all motion vector directions, including vectors that move along the z-axis. A runner or cyclist racing across the screen, a galloping horse, and a skier hurtling down a steep slope—all are strong motion vectors of high magnitude. The image of a racecar roaring down the track makes for a higher-magnitude motion vector than does a couple's leisurely stroll through the park. But like all elements in applied media aesthetics, vectors are context-dependent. For example, if you see the racecar moving across the screen in an extreme long shot, its magnitude is definitely lower than the relatively slow motion of a police officer who walks along the street in a tighter long shot, unaware of the villain who lurks in the shadows.

7.56 Continuing Index Vectors
Continuing vectors point in the same direction. These two people create continuing index vectors because they are looking in the same direction. The vectors are continuing, even if they are shown in separate successive shots.

VECTOR DIRECTIONS

Index and motion vectors can be continuing, converging, or diverging. Continuing vectors point in the same direction, converging vectors point toward each other, and diverging vectors point away from each other.

Continuing vectors When two or more index and/or motion vectors point in the same direction, they are ***continuing vectors***. The continuity can be established by the same vector type (graphic, index, or motion) or by a combination of vector types. The continuity can be established in a single shot or a shot sequence. If, for example, you see two people looking in the same direction in a single shot, their gazes constitute continuing index vectors. **SEE 7.56**

If you show a child watching her balloon escaping into the sky, you have continuous vectors within a single shot. Both the index vector of the child's gaze and the motion vector of the balloon's ascent point in the same direction. These vectors remain continuous even if you show the same event in separate shots. For example, the first shot could show a close-up of the balloon slipping out of the girl's hand and how she looks upward, with the balloon moving out of the frame. The second shot could then show the balloon climbing higher into the sky.

Although graphic vectors have a relatively weak magnitude (their direction is basically ambiguous), they can be continuous, especially in a shot series or in multiple screens. Let's assume that you are shooting a river scene with cliffs as the background. The horizon is a definite graphic vector that cuts across the screen at a particular spot. To ensure the necessary continuity from shot to shot, the horizon should not jump up or down in subsequent shots. The graphic vector (horizon line) must be continuous. **SEE 7.57**

Converging vectors In contrast to continuing vectors, which support each other in establishing or maintaining a specific single direction, ***converging vectors*** point toward each other either in a single shot or a series of shots. A simple conversation between two people creates converging index vectors whether you show them in a single shot or in individual close-ups. **SEE 7.58–7.60** Two cars racing toward each other represent converging motion vectors. Converging motion vectors are especially useful to increase the aesthetic energy of a high-magnitude event, such as two trains on a collision course.

Diverging vectors If index or motion vectors point away from each other, they are ***diverging vectors***. Obviously, two people pointing or looking in opposite directions create diverging index vectors. Cars traveling in opposite directions

7.57 Continuing Graphic Vectors
Although graphic vectors are low-magnitude, which means that their direction is ambiguous, you can nevertheless see them as continuous. In this three-screen display, the horizon lines must match to form a single graphic vector.

7.58 Converging Index Vectors in Two-shot

Converging vectors point toward each other. These two people looking at each other form converging index vectors.

7.59 Index Vector: Screen-right

In a close-up sequence, you need to maintain the same index vectors as in the two-shot. Here the person on the left effects a screen-right index vector.

7.60 Index Vector: Screen-left

To make the index vectors converge over the shot sequence, the screen-right person must look screen-left, establishing an index vector that converges with that of the screen-left person.

are examples of diverging motion vectors regardless of whether they appear in a single shot or a shot series. **SEE 7.61**

7.61 Diverging Index Vectors in Single Shot

Diverging vectors point in opposite directions. These two people form diverging index vectors by looking away from each other.

Whether z-axis vectors (index or motion) are perceived as converging or diverging depends on the event context. For example, if you first see two people looking at each other in a two-shot and then see them in z-axis close-ups, your ***mental map***—that is, where you perceive things to be in on- and off-screen space—will most likely tell you that the z-axis index vectors of the close-ups are still converging; you will still see the two people as looking at each other even as they look directly into the camera. If the two-shot shows them looking away from each other, you will most likely see the successive close-ups of the two people looking into the lens as diverging. This time your mental map is set on divergence, and you will perceive the people as looking away from each other even in the separate close-ups. For more examples of z-axis vectors and their behavior, see chapters 8 and 17.

So far we have discussed the basic types of external (screen) vectors and their major attributes—magnitude and continuing, converging, and diverging directions. In the following chapters, you will learn more about how vectors operate within the screen, from shot to shot, and from screen to screen.

The proper manipulation of vector fields is one of the most important aspects of your quest to clarify, intensify, and interpret an event for your viewers.

S U M M A R Y

There are six major types of field forces: main directions, magnetism of the frame, asymmetry of the frame, figure and ground, psychological closure, and vectors.

Horizontal lines suggest calm and normalcy. Vertical lines suggest power, formality, and strength. A combination of vertical and horizontal reflects our normal world; it follows the pull of gravity and suggests that we are standing upright on level ground. A tilted horizontal plane implies instability or powerful dynamism in an otherwise stable, or even static, shot.

Magnetism of the frame refers to how the borders of the screen—especially the corners—exert a strong pull on objects within the frame. Moreover, graphic mass attracts graphic mass (all types of screen images). A larger graphic mass (a screen image occupying a relatively large screen area) is usually more stable than a smaller one (occupying a relatively small screen area). The smaller graphic mass is dependent, and the larger is more independent.

A diagonal going from the bottom of screen-left to the top of screen-right is perceived as an uphill slant; a diagonal from top-left to bottom-right implies a downhill slant. This suggests that we tend to start at the left and finish at the right when looking at a picture field. As the destination area, the right side of the screen seems more prominent: we generally pay more attention to an object when it is placed on screen-right.

We always try to organize a picture field into a stable ground against which less stable figures operate. The figure exhibits certain spatial and graphic characteristics, the most important of which is that the figure seems to lie in front of the ground. Certain figure/ground reversals can be distracting but can also contribute to startling and expressive effects.

True to our tendency to stabilize the world we perceive around us, we tend to organize pictorial elements into a pattern of simple geometrical figures, such as triangles, squares, and the like. We can perceive such patterns, even if we have only minimal information, by mentally filling in the missing information—a process known as psychological closure, or closure for short.

The pattern we arrive at through psychological closure is called a gestalt. In a gestalt the parts have assumed a different structural role than when they were independent.

A vector is a force with a direction and a magnitude. All screen events exhibit one or a combination of the three principal vectors: graphic, index, and motion.

A graphic vector is created by simple lines or stationary elements that are arranged so that the viewer sees them as lines. Although graphic vectors suggest a general direction, their actual direction is ambiguous. An index vector is created by an object that unquestionably points in a specific direction. A motion vector is created by an object that actually moves or is perceived as moving on-screen. A still photograph of a moving object, even when blurred, is an index vector, not a motion vector.

Each vector can have a variety of magnitudes (strengths), depending on how conspicuous it is, how emphatically it points in a certain screen direction, its graphic mass, and its perceived object speed. In general, however, graphic vectors have less magnitude than index vectors, which have less magnitude than motion vectors.

Vectors can be continuing, converging, or diverging. Continuing vectors point in the same direction either in the same shot or in a shot sequence. Converging vectors point or move toward each other. Diverging vectors point or move away from each other. How we perceive converging or diverging z-axis vectors depends on the event context.

NOTES

1. Headroom follows the same principles as demonstrated in figure 7.10.
2. See also chapter 8.
3. This asymmetry of screen-right and screen-left—that is, whether one side of the picture draws more attention than the other regardless of picture content—has been a source of confusion and debate for some time. Alexander Dean (1893–1949), who taught play directing at Yale and was quite particular about the asymmetry of the stage, claims that the audience-left side of the stage is "stronger" than the right side because the audience has a tendency to look to the left first and then to the right. Alexander Dean, *Fundamentals of Play Directing* (New York: Farrar and Rinehart, 1946), p. 132.

   Heinrich Wölfflin (1864–1945), a prolific writer on various subjects in art history and theory, claims that the right side of a painting is "heavier" than the left. He says that we have a tendency to read over the things on the left quickly to come to the right side, where "the last word is spoken." He demonstrates convincingly with several illustrations how the character of a painting changes when its sides are reversed. Heinrich Wölfflin, *Gedanken zur Kunstgeschichte* (Ideas for Art History) (Basel: Benno Schwabe, 1940), pp. 82–96.

Rudolf Arnheim speaks of the well-known tendency to perceive the area in the left corner of a visual field as the point of departure and thus the entire picture as organized from left to right. (Rudolf Arnheim, *The Power of the Center* [Berkeley: University of California Press, 1982], p. 37.) But then he goes on to say that "the left side of the field, corresponding to the projection areas in the right half of the brain, is endowed with special weight" and that "objects placed on the left assume special importance." (Arnheim, *Power of the Center,* p. 38.)

In an experimental study, Nikos Metallinos and Robert Tiemens conclude that there is "some evidence that the retention of visual information in a newscast is enhanced when the visual elements are placed on the left side of the screen" but that there is in general "minimal support to the asymmetry of the screen theories." (Nikos Metallinos and Robert K. Tiemens, "Asymmetry of the Screen: The Effect of Left Versus Right Placement of Television Images," *Journal of Broadcasting* 21, no. 1 [1977]: 30.) Very much aware of the contextual nature of aesthetic communication, however, they warn that "no final conclusions can be made on the basis of mere placement of the visual elements" and that "such factors as size, color, shape, vectors (directional forces), and how individual subjects perceive these qualities must also be considered." (Metallinos and Tiemens, "Asymmetry," p. 32.)

More recent eye-tracking experiments are generally inconclusive and demonstrate more the observer's primary interest in the content of the picture than the asymmetry-of-the-screen phenomenon. See Bruce E. Goldstein, *Sensation and Perception,* 7th ed. (Belmont, Calif.: Wadsworth, 2007), pp. 122–31.

4. Goldstein, *Sensation and Perception,* pp. 99–101.

5. Graphic artist M. C. Escher used the figure/ground reversal in a highly inventive and unconventional way. See M. C. Escher et al., *The World of M. C. Escher* (New York: Harry N. Abrams, 1972). See also Bruno Ernst, *De Toverspiegel van M. C. Escher* (The Magic Mirror of M. C. Escher) (Munich: Heinz Moos Verlag, 1978). Even if you don't read Dutch or German, the book contains many of Escher's lesser-known drawings.

6. Wolfgang Köhler, *The Selected Papers of Wolfgang Köhler,* ed. by Mary Henle (New York: Liveright, 1971), pp. 149-65.

7. Max Wertheimer, "Untersuchungen zur Lehre von der Gestalt" (Studies for the Teachings of Gestalt), *Psychologische Forschung* 4 (1923), pp. 301–50. As an eminent member of the Gestalt school, Wertheimer introduced the Law of Prägnanz, which postulates that psychological organization can only be as good as the prevailing conditions allow. *Good* in this case means that the resulting gestalt is often symmetrical, simple, and relatively stable.

8. Although the terms *definition* and *resolution* are frequently used interchangeably, there is an important technical distinction. Definition refers to the perceived sharpness in an image; resolution is the measurement of this sharpness, such as the number of pixels in a specific picture area.

9. Marshall McLuhan, *Understanding Media: The Extensions of Man* (New York: McGraw-Hill, 1964), pp. 312–14.

10. Richard Zakia, *Perception and Imaging,* 3rd ed. (Boston: Focal Press, 2007), pp. 27–53.

11. The term *vector* in a non-science context was first used by Kurt Lewin to refer to psychological forces; it was later used by Andrew Paul Ushenko to mean an aesthetic force. See Kurt Lewin, *A Dynamic Theory of Personality,* trans. by Donald Adams and Karl Zener (New York: McGraw-Hill, 1935); and Andrew Paul Ushenko, *Dynamics of Art* (Bloomington: Indiana University Press, 1953), pp. 60–119.

12. The definition given is an aesthetic one. If you were dealing in physics (the origin of the vector concept), the definition would be that a vector is a physical quantity with both a magnitude and a direction.

13. All vector types are clearly illustrated in the interactive DVD-ROM *Zettl's VideoLab 3.0* (Belmont, Calif.: Wadsworth, 2004). Click on EDITING→ Continuity→ vectors.

14. See Chitra Dorai and Svetha Venkatesh (eds.), *Media Computing: Computational Media Aesthetics* (Boston: Kluwer Academic, 2002).

8

Structuring the Two-dimensional Field: Interplay of Screen Forces

STRUCTURING THE TWO-DIMENSIONAL FIELD MEANS MAKING THE screen forces work for you rather than against you. This enables you to show events on-screen with clarity and impact. Painters and still photographers strive for a total composition effect in a picture. They try to arrange static pictorial elements so that they look and feel inevitably right. Once such an arrangement is achieved, the composition is finished; it will not change.

Such is not the case in video and film—or even on the computer screen. Most pictorial elements shift constantly within the frame and often change from one picture to the next (from shot to shot, from scene to scene, or during scrolling). You are no longer dealing with structural permanence but largely with structural change.

The old method of composition, that is, the pleasing arrangement of essentially static pictorial elements within a single frame, does not suffice for video and film. You must now think in terms of structuring a dynamic visual field of ***on-screen space*** and consider visual elements that move about the screen and that need to provide structural continuity between previous and subsequent images. In video and film, the guiding compositional principle is not necessarily the screen space of a frame but of a sequence. For example, the composition of a single shot may look wrong when you examine it by itself, but when you see it as part of a shot sequence its composition becomes perfectly acceptable.[1] SEE 8.1 AND 8.2

Structuring the screen space of a Web page presents a unique problem. Often the information does not fit in the relatively small frame of the computer monitor, and the user must scroll up or down and to the sides to reveal pictures and written information. You should therefore consider even a single Web page design not as a static image, similar to a magazine page, but as partial images that maintain their structural integrity when scrolled from one segment to another. In this way their structure is influenced by their sequence much like a brief video scene.

Does this mean that all traditional compositional principles are invalid? Not at all, but they must be adapted so that they fulfill the more complex tasks of structuring static, moving, and sequential images within the two-dimensional field. You must now go beyond the traditional canons of good composition (area proportion, object proportion, and balance) to begin seeing the contextual interaction of the various screen forces, such as the magnetism of the frame, graphic mass, and vectors.

8.1 Composition in Shot Sequence: Context Lacking
The composition in this shot is definitely wrong (the man stands much too close to the right side of the frame).

8.2 Composition in Shot Sequence: Context Provided
The man's position makes sense as soon as you provide the proper context for such a framing, such as the woman's entering the shot from screen-left.

This chapter explores the interplay of screen forces of the two-dimensional field: stabilizing the field through distribution of graphic mass and magnetic force, stabilizing the field through distribution of vectors, stages of balance, object framing, unusual compositions, multiple screens, and dividing the screen.

Stabilizing the Field Through Distribution of Graphic Mass and Magnetic Force

One of the most basic ways of stabilizing the two-dimensional field is to bring into balance the forces of ***graphic mass*** and the ***magnetism of the frame.*** As discussed in chapter 7, every graphic mass operating within a clearly defined two-dimensional field, such as the video, film, or computer screen, carries a graphic weight, which is somewhat akin to the actual weight of an object. ***Graphic weight*** is determined by the dimension of the object (how much area the object takes up relative to the total screen area), its basic shape and orientation, its location within the frame, and its color. SEE 8.3

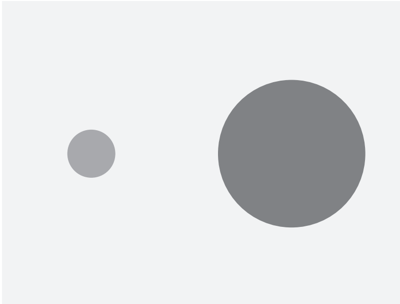
8.3 Factors Influencing Graphic Weight

Factor		Heavy	Light
Dimension		Large	Small
Shape		Simple, geometrically compact	Irregular, diffused
Orientation		Vertical	Horizontal
Location		Corner Upper part of frame Right	Centered Lower part of frame Left
Color	*Hue*	Warm	Cold
	Saturation	Strong	Weak
	Brightness	Dark	Light

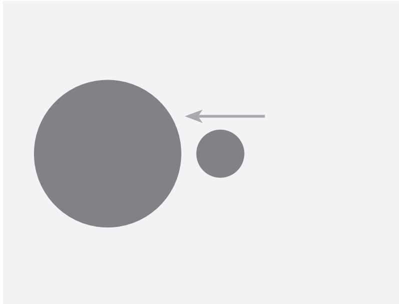
A screen object does not, however, have to display all these elements. Object size alone is enough to give an image graphic weight. **SEE 8.4** For example, the disc on the right side of the frame in figure 8.4 is obviously "heavier" (carries more graphic weight) than the one on the left.

If objects are close enough to each other that the attraction of mass comes into play (as explored in chapter 7), we tend to combine the graphic weights of both objects. **SEE 8.5**

The proximity of the object to the screen edges brings another significant graphic force into play: the magnetism of the frame. Obviously, the closer the object is to the edge of the screen, the more powerful the magnetic force will be regardless of its relative weight. A strong index vector, however, can quite easily override the magnetic force of the frame. Let's look at some simple interactions of graphic weight and magnetism of the frame.



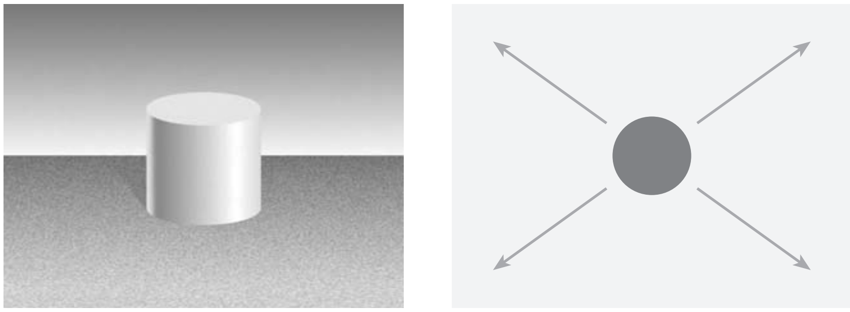
8.4 Graphic Mass and Weight
The larger, darker, and therefore heavier graphic mass on the right side of the screen outweighs the smaller, lighter object on the left.



8.5 Mass Attracts Mass
As in physics, in visual design graphic mass also attracts graphic mass. The attracted objects combine into a single, greater mass, which then has greater graphic weight.

SCREEN-CENTER

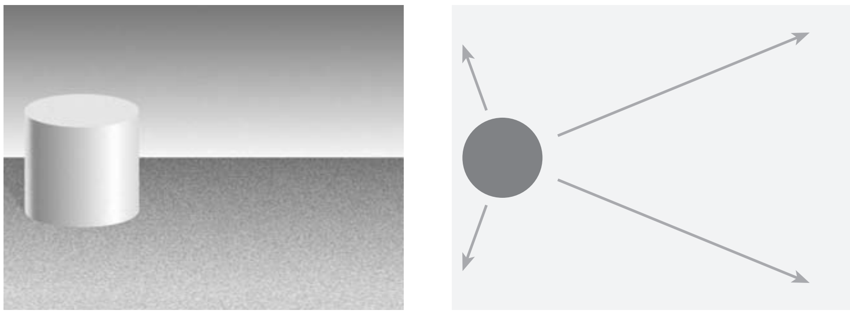
The most stable position of an object is screen-center. When the graphic mass is located in the center of the screen, the surrounding areas and the magnetic forces of the screen edges and corners are symmetrically distributed. **SEE 8.6**



8.6 Screen-center Position
The screen-center position provides maximum stability.

OFF-CENTER

As soon as you move the object to one side, however, the graphic weight increases and the magnetism of the frame comes into play. The more the object moves off-center, the greater its graphic weight—and the attraction of the frame increases. When this happens the picture begins to look unbalanced. But you can certainly use this "wrong" framing to increase the tension in the dynamic picture field. **SEE 8.7**



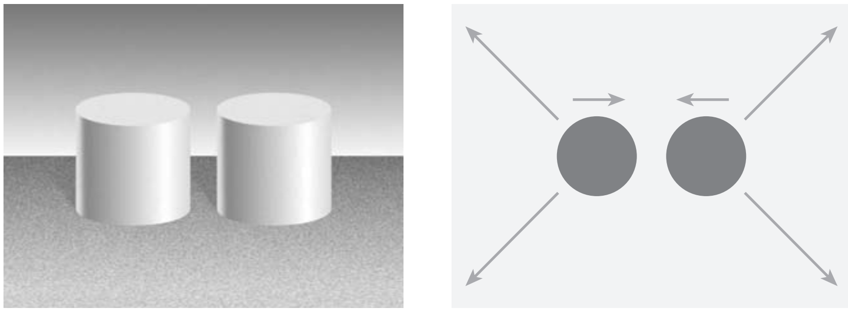
8.7 Off-center Position
When the object moves off-center, its graphic weight increases, and the magnetism of the frame comes into play.

COUNTERWEIGHTING

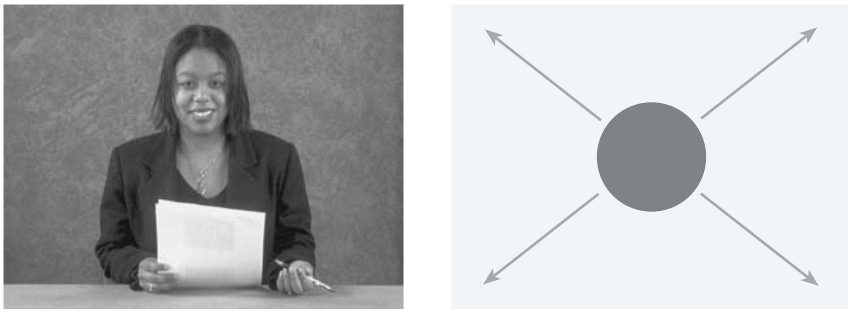
You can correct an imbalance of graphic weight by centering the object through camera or object movement so that the pull of the frame is neutralized (see figure 8.6) or by counterweighting it with another object or other graphic element (such as lettering or even beam of light) of similar graphic weight. **SEE 8.8** The attraction of mass caused by the two objects makes them gravitate toward screen-center. The closer the objects are to each other, the more apt you are to perceive them as belonging together and having a single graphic weight.

Now let's replace the abstract cylinders with a newscaster. The most stable position for the newscaster is obviously screen-center. Picture stability is important here to signify the newscaster's authority and credibility. **SEE 8.9** Putting the newscaster to one side of the screen makes little sense unless you are making room for a co-anchor or for additional pictorial material, such as the box insert that is customarily placed above and to the side of the newscaster's shoulder. **SEE 8.10**

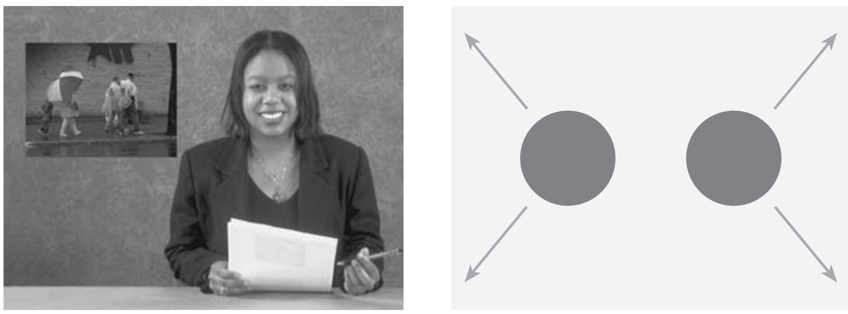
The graphic weight of the secondary frame obviously provides sufficient graphic weight to balance the off-center newscaster. But when you remove the box during the newscast, you need to once more center the anchor to restore balance. A good camera operator can make this shift so smoothly that most viewers are unaware of this rebalancing act.



8.8 Counterweighting
You can achieve balance by counterweighting an object with another object of similar graphic weight. Note that attraction of mass also operates here.



8.9 Newscaster in Screen-center
The most stable and sensible position for a single newscaster is in screen-center because this placement emphasizes the newscaster's credibility.



8.10 Off-center Balance with Box
The traditional box insert necessitates the off-center positioning of the newscaster. If the new element has sufficient graphic weight, the shot will remain balanced.

You may, however, want to place titles and other still graphic images off-center to boost the picture's graphic tension and energy. The asymmetrical distribution of graphic mass and vectors is discussed later in this chapter.

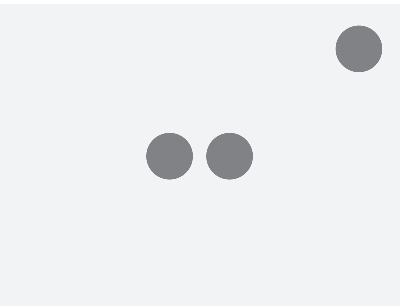
Stabilizing the Field Through Distribution of Vectors

When trying to stabilize the two-dimensional field, you need to focus not only on graphic weight but also, if not especially, on the distribution of vectors. They are such powerful structural elements that they usually override the lesser forces of graphic weight and magnetism of the frame.

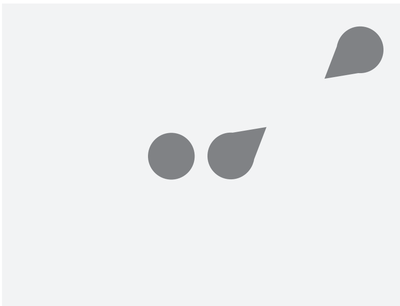
STRUCTURAL FORCE OF INDEX VECTORS

Take a look at the following figure. **SEE 8.11** You will inevitably perceive the two center discs as belonging together (attraction of mass, see figure 8.8) and the upper-right disc as the isolated one (pull of screen corner, see figure 7.12).

But vectors can easily override these relatively subtle structural forces and cause you to perceive a different pattern. **SEE 8.12** The discs are in the same positions as in figure 8.11, but we have put "noses" on two of them. They have now



8.11 Structure Through Attraction of Mass
The two center discs attract each other and are seen together. The disc in the upper-right corner is isolated.



8.12 Structure Through Index Vectors
The converging index vectors now combine the right-center disc and the disc in the upper-right corner. The left-center disc is isolated.

become high-magnitude index vectors that establish a new relationship. Now the right-center disc and the corner disc are strongly connected through their converging index vectors, putting the left-center disc in isolation.[2] The increasing magnitude of index vectors of somebody's turning from a straight-on (z-axis) shot to a profile shot has similar structural consequences.

NOSEROOM AND LEADROOM

Imagine the following figures as a shot sequence. **SEE 8.13** In 8.13a the woman is looking directly into the camera (at you), generating a z-axis index vector. Because this z-axis index vector points directly at you, its magnitude within the screen is practically zero. You can therefore ignore the force of this index vector and stabilize the picture strictly by graphic weight and magnetism of the frame. Putting the subject in the center of the frame (maximum stability of graphic mass) with adequate headroom (neutralizing the magnetism of the upper edge) is the most logical thing to do.

8.13 Force of Index Vector: Lack of Noseroom

The increasing magnitude of the index vector when someone gradually shifts from looking directly into the camera (z-axis vector) to looking to the side will cause the index vector to crash into the screen edge.

8.14 Force of Index Vector: Proper Noseroom

If you compensate for the increasing vector force by pulling the subject back to provide enough noseroom, the vector has enough space to have its force absorbed or comfortably guided through the screen border and into the next shot.

But watch what happens when the woman turns and starts looking at the screen edge (figures 8.13b–d). The more she looks to the side, the greater the magnitude of the index vector and the more the structural force of the vector comes into play. This index vector reaches its maximum magnitude in the profile shot (figure 8.13d). Although she has moved only her head and shoulders and has not changed her basic screen position, in the 4 × 3 aspect ratio the shot looks strangely out of balance. She seems cramped into the screen-left space, with her index vector crashing into the left edge of the frame. Assuming that no other person will walk up behind her to balance the picture through graphic weight (see figure 8.2), you will have to shift the subject more to screen-right to give the index vector enough space to run its course and travel relatively unhindered to, and even through, the screen edge. This space is often called ***noseroom*** for index vectors and ***leadroom*** for motion vectors.[3]

Note that you must leave more noseroom the higher the magnitude of the index vector becomes. **SEE 8.14** The index vector is at its maximum magnitude in the profile shot; therefore, the noseroom has also reached its maximum length (figure 8.14d). This noseroom principle also applies to the 16 × 9 aspect ratio, especially when you frame tighter shots. **SEE 8.15**

But wouldn't the index vector have a better chance of penetrating the screen edge if it originates as close to the edge as possible? Apparently not. When the index vector operates too close to the screen edge, it draws undue attention to the edge itself. As part of a picture frame whose major function is to contain the event, the screen edge acts as the final barrier to the index vector. This "short" vector and the graphic mass of the subject's head fall victim to the magnetism of the frame; both graphic forces are firmly glued to the edge and can no longer penetrate it.

A wider noseroom not only neutralizes the magnetism of the frame but also creates enough space around it to divert our attention away from the edge. Rather than deplete the vector's energy, the noseroom seems to signal clear sailing for the vector, not only through on-screen space but also—through the edge—into ***off-screen space***.[4]

You treat motion vectors in the same way. When panning the camera with a moving object, always pan enough ahead to maintain leadroom for the motion vector to play out. If you do not sufficiently lead the moving object, the screen edge toward which the object moves will appear as a formidable barrier. All you see is where the object has been but not, as you should, where it is going. A lack of leadroom will make the motion look cramped and hampered.

Similar to the noseroom of lateral index vectors, lateral motion vectors need the most leadroom. Objects that move at an oblique screen angle need less leadroom, and z-axis vectors don't need any.

8.15 Lack of Noseroom in 16 × 9 Aspect Ratio

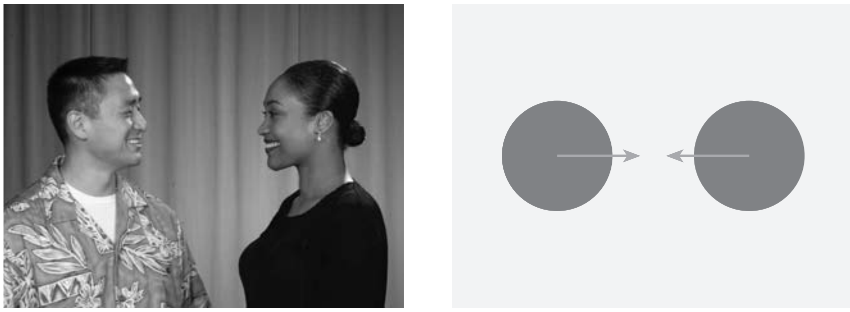
The lack of proper noseroom shows up prominently even in the wider 16 × 9 aspect ratio.

CONVERGING VECTORS

You can also balance an index vector with a converging one within the same frame. **SEE 8.16** With two people looking at each other, you achieve balance through the converging index vectors and the almost symmetrical placement of the subjects, whose graphic mass translates into just about equal graphic weight.

GRAPHIC VECTORS

Although graphic vectors as a category have relatively low magnitude, you can nevertheless use them to stabilize the two-dimensional field. **SEE 8.17** The perspective lines of the buildings in the figure produce fairly strong graphic vectors that lead our eyes naturally, though not as directly as an index vector, along the downhill diagonal from screen-left to screen-right. If you want to prevent the vector from plunging through the right screen edge—to contain it properly within the frame—you can do it with other graphic vectors or with graphic mass. In figure 8.17 the downhill vector of the entablature (on top of the columns) is arrested by the horizontal block of the back building.



8.16 Converging Index Vectors

Converging index vectors of equal magnitude balance each other.

8.17 Distribution of Graphic Vectors and Graphic Mass

You can use graphic vectors and graphic mass to contain other graphic vectors. Here the horizontal and vertical graphic vectors of the back building and the resulting graphic mass block the sloping vectors of the columns.

Stages of Balance

Look back at figures 7.1 through 7.9. You will probably notice that the various structural arrangements in these illustrations do not have the same degree of balance, that is, the same degree of structural stability. Some look more at rest, whereas others appear to have more internal tension—they look more dynamic.

Because our organism strives to obtain a maximum of potential energy and to apply the best possible equilibrium to it, as Rudolf Arnheim points out, balance does not necessarily mean maximum stability within the screen.[5] Rather, ***balance*** can range from, or fluctuate between, static (stable) and dynamic (unstable) field structures.[6]

STATIC BALANCE

A ***static balance*** is solid, extremely stable, and not about to move. One extreme form is a symmetrical structuring of visual elements. This means that more or less identical picture elements appear on the left and right sides of the screen. The forces of graphic mass, frame magnetism, and vectors are the same, or at least almost the same, on both sides of the screen. SEE 8.18

DYNAMIC BALANCE

In a ***dynamic balance***, the graphic elements are asymmetrically distributed. This means that the graphic weight and the various vectors are no longer equal on both sides of the screen. Instead they are engaged in a sort of tug of war with one another that increases dynamic energy. SEE 8.19

Golden section One of the classical ways of creating a dynamic balance is to use the proportions of the ***golden section***—a division of the screen (or any other linear dimension) into roughly 3 × 5 units (or, more accurately, 0.616:1). To achieve a golden section, you divide a given horizontal dimension, such as the width of this page, into a larger part (approximately three-fifths of the total width) and a smaller part (approximately two-fifths of the total width). The point where these portions meet is the golden section.

In such a division, the small and large screen areas are competing with each other, with the larger part not quite able or even eager to outweigh the smaller one.

8.18 Static Balance

Symmetrical balance is among the most stable forms of balance. In this picture both sides have identical graphic weights, frame magnetism, and vector distributions. The tension is low.

8.19 Dynamic Balance

In a dynamic balance, the aesthetic energy is increased because the asymmetrical distribution of graphic elements and vectors causes some tension.

Also, the dividing line (actual or imaginary) has not given in to the magnetic pull of one or the other screen edge, although one is definitely pulling harder than the other. The result is a less stable structure with increased graphic energy, yet the picture is still balanced. For many centuries this proportion was considered ideal and, at times, even divine. (See more details about the golden section on the next page.)

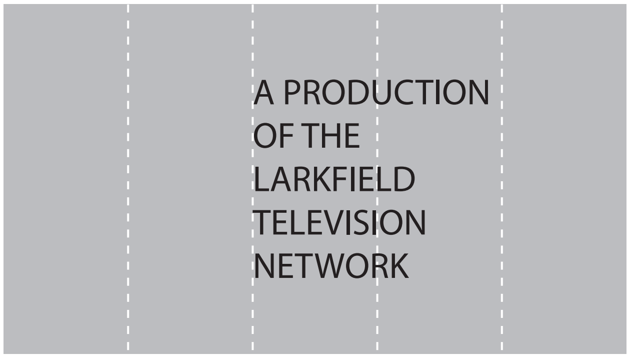
Although the golden section is rarely applicable when dealing with moving images, it is nevertheless valuable when framing relatively static shots and designing titles and still images for the screen. For example, a title that divides the screen into the golden section proportions often gains dynamism and visual interest compared with one that is centered. **SEE 8.20** Even when using illustrated titles, you may well arrange the major picture elements in proportions according to the golden section. This way the title gains in graphic energy without threatening the overall balance of the screen image.

The golden section division of screen width is especially effective when a vertical element (vertical graphic vector) divides a clean, horizontal vista (horizontal graphic vector). **SEE 8.21**

In a similar way, you can use a relatively uncluttered horizon line (horizontal graphic vector) to divide the screen at approximately three-fifths or two-fifths of its height. Thus, you can place the horizon line in the vertical golden section of the frame. **SEE 8.22**

The golden section division is especially appropriate in arranging a Web page that contains multiple visual elements. By placing the visuals in the golden sections on the screen, you avoid the all-too-common scattered look of the pages without sacrificing their dynamism.

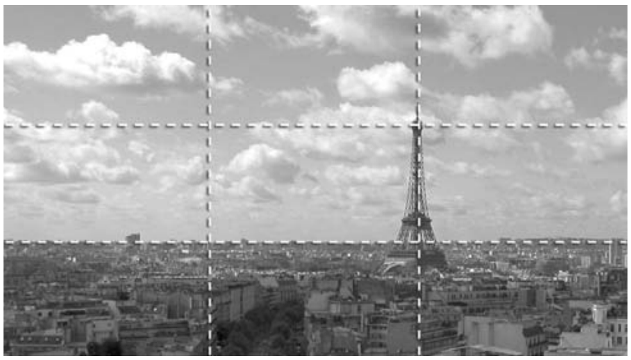
Rule of thirds A variation of the golden section is the ***rule of thirds***, which suggests dividing the screen into three horizontal and three vertical fields. **SEE 8.23** You can always



8.20 Golden Section Applied: Titles
The left edge of the text coincides with the golden section, providing increased visual interest. The figure balances the titles through graphic weight.

8.21 Golden Section Applied: Horizontal Division
It is especially effective to have a single vertical element divide a clean horizontal graphic vector at the golden section.

8.22 Golden Section Applied: Vertical Division
You can also use horizontal graphic vectors (such as the horizon line) to divide the screen vertically at the golden section.



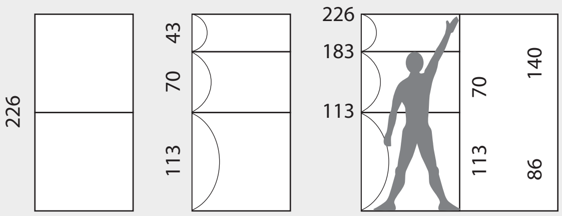
8.23 Rule of Thirds
To achieve a pleasing composition within a frame and placement continuity from shot to shot, divide the screen into thirds. Position the principal subjects where the horizontal and vertical lines intersect.

achieve a fail-safe composition by placing subjects where a vertical and a horizontal line intersect. These fields can help you maintain continuity of subject and object placement when shooting out of sequence and help you, as pointed out earlier, achieve and maintain proper headroom.

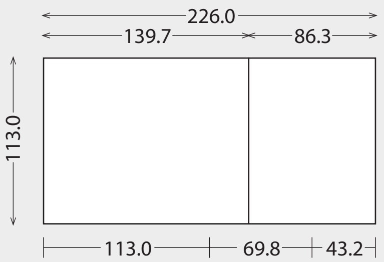
Although you may now presume that the golden section or the rule of thirds can provide area proportions that are, indeed, divine, avoid going overboard with them. Think twice before placing the newscaster in the golden section simply

The Modulor Well-known contemporary Swiss-French architect **Le Corbusier** (Charles E. Jeanneret-Gris, 1887–1965) developed a proportional system that is essentially a refined version of the golden section.[7] His system, which he called *the Modulor,* is also based on the proportions of the human figure, specifically the proportions of a 6-foot man.

All the Modulor proportions are presented in a gradually diminishing scale of numbers. Here, in Corbusier's diagram, all numbers are in centimeters (one one-hundredth of a meter).



The more exact Modulor dimensions are:

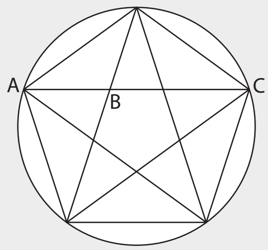


Harmonious Proportions The question of discovering proper proportions and using them with consistency in the various forms of art has been of major concern to artists for centuries. The Egyptian temples and wall paintings; the Greek and Roman buildings and sculptures; the churches, palaces, and paintings of the Renaissance; the modern skyscraper; magazine layouts; and automobile design—all reveal the human preoccupation with proportional harmony. Amazingly enough, the proportions as revealed by the Egyptian and Greek temples, a Gothic cathedral, or a Renaissance palace still seem harmonious to us today.

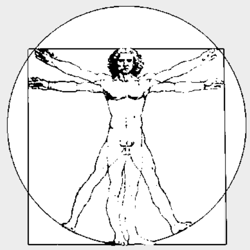
Obviously, we have, and always have had, a built-in feeling for what proportion constitutes. But because we are never satisfied with just feeling but also want to know *why* we feel a particular way and what makes us feel that way—mainly to make emotional responses more predictable—people have tried to rationalize about proportional ratios and develop proportional systems. Mathematics, especially geometry, was of great help to people who tried to find the perfect, divine proportional ratio.

One charming illustration of this point is a statement by **Albrecht Dürer**, the famous German Renaissance painter (1471–1528). In the third book of his *Proportionslehre* (Teachings on Proportions), he writes: "And, indeed, art is within nature, and he who can tear it out, possesses it....And through *geometrica* you can prove much about your works."[8]

Golden Section The most well-known proportional ratio is the golden section, often called the "divine proportion" or the "golden mean." The familiar pentagram, or five-pointed star, contains a series of golden sections, each line dividing the other into a golden section proportion.

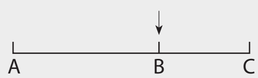


This proportion was produced by calculating minds from ancient Egypt to the Renaissance. Although the Egyptians knew about the golden section proportion and used it extensively in their architecture, sculpture, and painting, the Greeks are usually credited with working out the mathematics of this proportion and relating it to the proportions of the human figure. Later, **Leonardo da Vinci**, the great Italian Renaissance painter, scientist, and inventor, spent much time proving and making public the validity of the mathematical formula of the golden section (0.616:1), worked out by Greek philosopher and mathematician **Pythagoras** as early as 530 B.C.

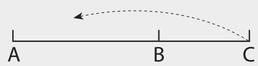


In the golden section, the smaller section is to the greater as the greater is to the whole. Thus:

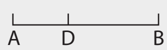
$$\frac{BC}{AB} = \frac{AB}{AC}$$



This proportion continues ad infinitum. If you fold the BC section (on the right) into the AB section (on the left), you will again have created a golden section.



$$\frac{AC}{BC} = \frac{AB}{AD}$$



The quickest way to divide a line into golden section proportions is to multiply its length by 0.616. Mark off the line with the resulting length. This point will divide the line into the larger and smaller parts of the golden section.

because you don't like the lost-in-video-space feeling when he or she is centered in the large 16 × 9 screen. Such a maneuver would most likely fail to make the shot more dynamic and keep the viewers watching. If the newscaster lacks dynamism in personality as well as message, even a divine screen placement will fail to improve the communication. Unless the newscaster's shift to the side of the screen is done to accommodate additional visual material, the newscaster will merely look off-center.[9]

Modular units Architects and scene designers have modified golden section proportions into a modular concept. This means that a piece of scenery or a prefabricated wall can be used in a variety of configurations, with, for example, two or three widths of one scenic or building unit fitting the length of another. This makes the units easily interchangeable. With modular units you can create a great variety of scenic structures without having to build custom sets each time a new environment is required.

PUSHING DYNAMICS

The distribution of graphic weight, frame magnetism, and vectors can be pushed to their structural limit. As viewers we sense that even the slightest change in the field structure would cause a total loss of balance and stability. This instability makes the graphic tension and energy quite high. Note that in video and film, such extreme dynamics are usually temporary, which means that this instability in composition is done to intensify an especially energetic or precarious moment, after which the composition reverts to a more comfortable stasis.

You can achieve such pictorial tension by overloading one or the other side of the screen with graphic weight, by not providing the vectors with enough room to play out, or by constantly having high-energy vectors converge within the shot or in a shot series. **SEE 8.24**

The easiest way to achieve more tension and move from a static and stable balance to a highly dynamic and ***labile***, or unstable, one is to tilt the horizon line. In this way the customary horizontal/vertical equilibrium is disturbed enough to create tension without changing the balance of the other structural forces of graphic mass, frame magnetism, and vectors. **SEE 8.25 AND 8.26**

Our inborn sense of equilibrium—our desire to see things stand upright on level ground—reacts so strongly to this labile balance that we try almost physically

8.24 Labile Dynamics

An extremely dynamic balance may even start out as a temporary imbalance. With the slightest change in the distribution of graphic elements, the vectors would lead to an unbalanced picture field.

8.25 Static Balance
The straight horizon gives this picture a highly stable balance.

8.26 Dynamic Balance
A heavily tilted horizon gives this picture a labile (unstable) balance.

to keep the objects in the picture from slipping out of the frame and to bring the horizon line back to its normal, level position. Hence we perceive such labile, extremely dynamic types of balance as high-energy.

Whether the balance should be static or dynamic is largely a matter of communication intent and context. As pointed out earlier, if you want to reflect the extreme excitement, tension, or insecurity of an event, the pictorial equilibrium should reflect this instability. You may do well to choose an extremely dynamic, unstable picture balance.

On the other hand, if you want to communicate authority, permanence, and stability, the pictorial arrangement should reinforce and intensify this by means of a static balance. Note that *static* in this context does not mean dull or uninteresting but rather a balance that is solid and stable.

8.27 Imbalance
If the graphic weight and the vectors create a totally lopsided picture, the extreme dynamism has changed into an unbalanced composition. If temporary, it can still serve as a perceptual attention-getter.

UNBALANCED SCREEN SPACE

In an unbalanced composition, there is no longer any aesthetic structure in screen space. Usually, the placement of pictorial elements is arbitrary and uncontrolled. The elements lack deliberate structure and look and feel chaotic. SEE 8.27 Why talk about unbalanced picture compositions when our goal is to achieve balance in structuring the two-dimensional field? Because in the temporary and transitory moving image of video and film, an unbalanced structure can be so startling as to be an attention-getter (see figures 8.37 to 8.39). You cannot afford to stay with such bad compositions for too long, however, or your intended perceptual jolt will become a perceptual rejection. An unbalanced shot that is not corrected is like a dissonant musical chord that is not resolved.

When looking through the viewfinder of a camera, you must learn to *feel* the proper balance. Especially when working in television news, the constantly shifting events do not tolerate long and careful structural contemplations. Like a ballet dancer or a champion athlete, you should be able to react to the structural demands of the moment both immediately and intuitively. But this type of intuition must grow out of a solid grounding in how to structure the two-dimensional field. Before your reflexes become optimal, you must practice them.

Object Framing

Take a breather from reading for a moment and look around. You'll probably notice that there is no object that you can see in its entirety unless you pick it up and look at it from all sides. When you have to portray this environment within the restricted space of a video or even a film screen, the viewer will see even less. The small video and computer screen, and especially the tiny mobile media display, favor close-ups and close-up sequences, all of which require, as you recall, a great deal of psychological closure, a mental filling in of missing information.

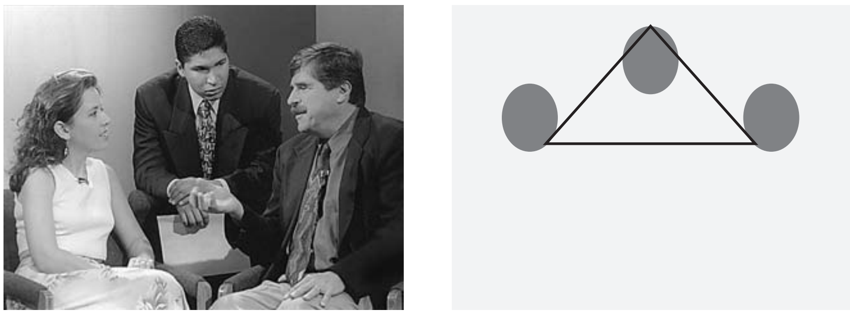
Although closure is an automatic perceptual response, you can nevertheless facilitate it by applying some basic compositional principles when framing a shot; you can inhibit it by ignoring them.

This section provides some answers by discussing closure within a single screen and in the extended two-dimensional field of multiple screens.

FACILITATING CLOSURE

To help you structure the two-dimensional field, try to arrange its visual content so that viewers can group and organize it into easily recognizable patterns of simple geometrical figures. **SEE 8.28–8.30** Even if you encounter visual material that is more complicated to compose than in these illustrations, you can always look for and feature dominant graphic vectors that may give viewers some orientation in the visual jungle and help them apply closure and form a relatively stable gestalt.

8.28 Closure into Rectangle
The graphic vectors in this picture help form the pattern of a rectangle.



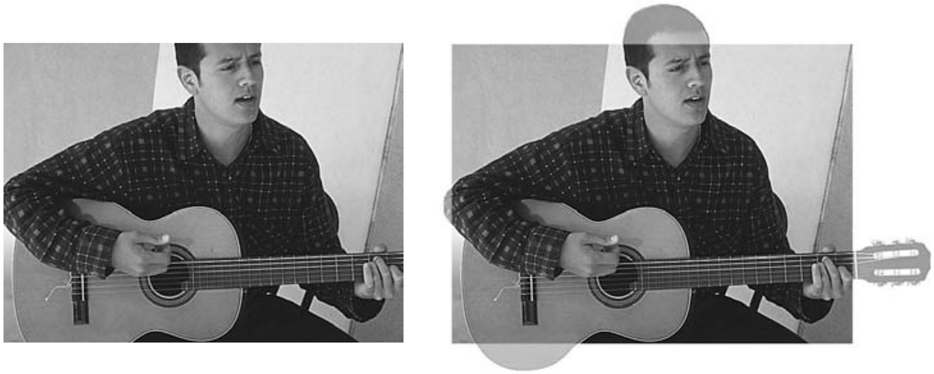
8.29 Closure into Triangle
This composition facilitates closure into a triangle.

8.30 Closure into Diagonals and Triangles
The strong graphic vectors in this Frank Lloyd Wright building divide the picture into strong diagonals and triangles.

GRAPHIC CUES

When showing only part of an object or a person on-screen, you must frame the subject so that the viewer can easily fill in the missing parts and perceive the whole. In vector terminology this means that you facilitate psychological closure by arranging the vector field within the screen area so that all the vectors (graphic, index, and motion) contribute to the intended stage of balance and especially that they extend easily beyond the screen into off-screen space. **SEE 8.31** In this waist shot, you don't see a picture of a person with the top of the head and the body cut off, or parts of the guitar missing. Instead you automatically perceive a complete person who is seated, playing his guitar. This is primarily because the shot contains enough graphic cues (graphic vectors) to provide psychological closure in off-screen space.

8.31 Closure in Off-screen Space
In this waist shot, we perceive the whole person playing the guitar, although only part of his body and the guitar are visible. Sufficient graphic cues lead us into off-screen space.



PREMATURE CLOSURE

There are instances, however, in which improper framing can lead to ***premature closure***, which occurs when the vector field within the frame entertains such easy psychological closure that the image no longer compels us to extend it beyond the screen. This can happen even if only parts of an object or a person are shown. **SEE 8.32** The head of the subject in figure 8.32 (and in 7.26) is framed in such a way that all necessary conditions for closure exist within the frame; you automatically reduce the head to a simple, self-contained, highly stable gestalt—an oval. Practically no graphic cues lead us beyond the frame. As a result, the head has become an independent unit and so it appears disconnected from its body, resting—like John the Baptist's head on Salome's platter—on the lower part of the screen edge.

8.32 Premature Closure
The shape of the head inevitably leads to closure into an oval. Because an oval is a highly stable and self-contained configuration, you are no longer inclined to project beyond the frame. Hence the premature closure can cause you to perceive the head as disconnected from its body.

NATURAL DIVIDING LINES

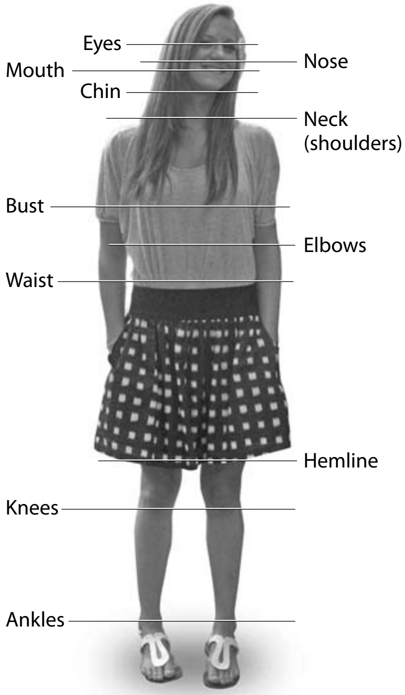
Similar problems of premature closure occur whenever the frame cuts off a person at any of the ***natural dividing lines,*** such as the eyes, mouth, chin, shoulders, elbows, hemline, and so forth. **SEE 8.33** When framing a shot, do not have these lines coincide with the top or bottom edge of the frame. Always try to frame a person so that the natural dividing lines fall either within or outside the screen edges. That way you give viewers cues to project the image into off-screen space and apply closure to the whole person.

Objects too have natural dividing lines. The principal graphic vectors in the houses in figure 8.34 are the vertical edges of their sides, the V-shaped roof lines, the horizontal belly bands that divide the houses into upper and lower floors, and the fences. If you frame the row of houses so that the side edges of the screen coincide with the outside-wall vectors of two houses, we have no need to project into off-screen space, so we perceive the two houses as a self-contained unit. **SEE 8.34**

But if you crop the picture as in the next figure, we are forced to extend the graphic vectors into off-screen space for proper closure. The outside walls now fall beyond the sides of the screen, and the graphic vectors of the roof lines are added cues that lead into off-screen space. When this happens we are inclined to extend the two houses into a row of houses. **SEE 8.35**

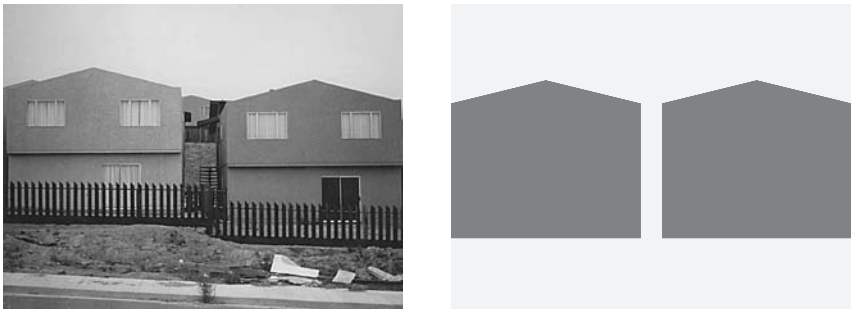
ILLOGICAL CLOSURE

Our need to make sense of the chaotic environmental and perceptual stimulus overload is so strong that we form patterns even when the visual elements that make up those patterns are illogical and obviously do not belong together. Thus we tend to group together in a single structure those visual elements that seem to provide an easy continuation of graphic vectors.[10] Because of the attraction of



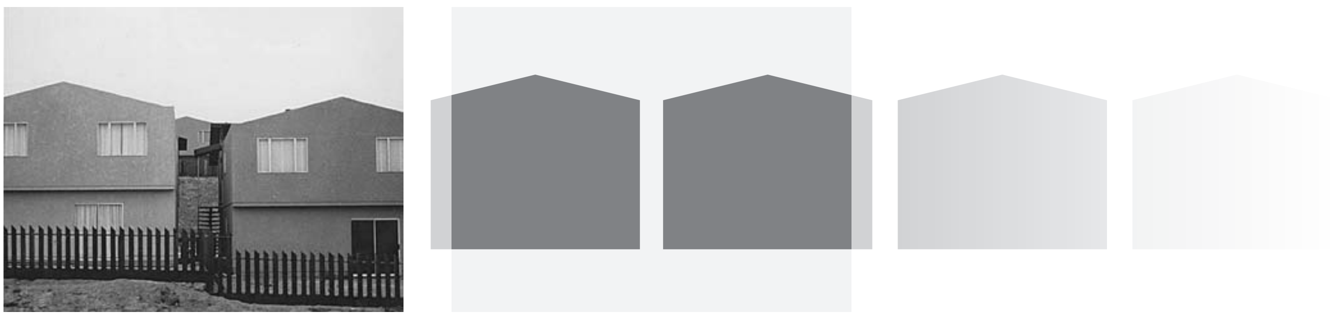
8.33 Natural Dividing Lines in Persons

Premature closure occurs when the upper or lower screen edge coincides with natural dividing lines. You should therefore frame a person so that these lines fall either inside or outside the upper and lower screen edges.



8.34 Dividing Lines Coinciding with Screen Edges

These houses are cut off at their natural dividing lines by the sides of the screen. We cannot tell whether there are only two houses or a whole row.



8.35 Dividing Lines Falling Outside the Screen Edges

Now that the edges of the natural dividing lines fall outside the screen edges, you are forced to apply closure in off-screen space; consequently, you are more apt to perceive a row of houses.

8.36 Illogical Closure
If facilitated by the smooth continuity of vectors, we tend to group objects into stable perceptual patterns regardless of whether they belong together. In this case, the field reporter seems to balance a street sign on his head.

mass and the close proximity of the street sign, we tend to perceive and form a stable configuration even when we know that the field reporter does not balance a street sign on his head. **SEE 8.36**

To avoid this problem when framing a shot, you must learn to look *behind* the main subject to see if possible closure anomalies lurk in the background. Be especially alert when the scene designer puts potted plants behind interview chairs. When the camera is turned on the scene, some of the stems will almost inevitably appear to grow out of people's heads, ears, or shoulders. Similar precautions apply when shooting outside. Look behind the immediate scene you are shooting to see whether any background objects, such as telephone poles or trees, might lead to illogical closure.

The Aesthetic Edge: Unusual Compositions

Now that you have learned the basic rules of composition, it's time to break them. Sometimes, under certain circumstances, unusual compositions will not only startle viewers and make them pay renewed attention but also sharpen the message in a subtle yet compelling aesthetic way. **SEE 8.37**

The woman is almost pulled through the left screen edge by the magnetism of the frame. The bench that extends to screen-right does not carry enough graphic weight or have a strong enough graphic vector to balance the shot. The woman's head is cut off by the upper screen edge just below the nose, preventing us from seeing her eyes—the windows to her personality. But this "wrong" framing increases the weight of the left side of the screen and directs our attention on the dog. Besides, this composition almost begs for a man to enter from screen-right and to stop and pay more attention to the dog than to the woman. You can continue this plot to your liking. Notice, however, that we were setting up this situation aesthetically, through an unusual composition rather than with a wide establishing shot or dialogue.

Other ways of drawing attention to a specific screen area are to place an object way off screen-center or show only part of it on-screen. **SEE 8.38 AND 8.39**

You can see some striking unusual compositions in design books, fashion magazines, and, of course, television commercials and some of the more stylish videos and movies. There is ample room to experiment with your camcorder and still camera. Whatever you do, the guiding principle for such unusual framing must be the intensification of the message, giving it an aesthetic edge, rather than

8.37 Emphasis on Dog

Placing the woman at extreme screen-left and having the upper frame cut below her nose puts the emphasis on the dog. We also set up the scene for somebody to enter from screen-right.

8.38 Emphasis Through Off-center Placement

By placing the vase near the right screen edge, the magnetism of the frame creates an aesthetic discrepancy that draws our attention to the flowers.

8.39 Emphasis Through Partial On-screen Placement

By showing only part of the car on-screen, we inevitably want the car to drive into on-screen space so that we can get a closer look at it.

making it simply look odd. If overdone, the unusual framing is perceived not as an intensifier but as a mistake; the initial attention-getter becomes an irritant.

Multiple Screens

We live in a world of ever-increasing complexity. Print and electronic media bombard us day and night with significant and not-so-significant information about what is going on in our neighborhood and in the far corners of the world. Instant access to information is no longer a luxury; it has become a necessity and a means of power. We have become information-dependent; that is, we need quick access to a large amount of varied information to function properly in our society. Whether you are shopping in a supermarket or going to the voting booth, you are asked to make choices. To make the right decisions, you need to know about the products or political candidates and be subjected to multiple viewpoints.

Despite its somewhat chaotic nature, the Internet makes information on practically any subject available instantly. Our predicament is not the availability of the information but rather locating what is relevant. Other problems are that much information is drastically streamlined to remain manageable—headline news and sound bites are but two examples—and that the information is still principally delivered linearly, line by line, picture after picture. As a result, the mediated events lack the complexity that you actually experience in everyday life. What is needed is a rethinking of the conventional television, film, and computer display techniques.

There are various ways of coping with this problem. As early as the beginning of the twentieth century, cubist painters, such as Pablo Picasso and Georges Braque, were well aware of the ever-increasing complexity of our existence and let us simultaneously see various points of view of the same event (see figure 12.1). Just a decade later, French film director Abel Gance pioneered the showing of simultaneous ***multiple screens*** in his film *Napoléon Bonaparte.*[11] His efforts, however, were and remain largely ignored by the film industry.

Today's media are well aware of the need for more efficient and effective communication—the increase of information density while keeping it lucid and understandable.

INCREASED INFORMATION DENSITY

Increased information density is especially noticeable in commercials and newscasts. The average running time of a standard commercial is 30 seconds, in which you may be subjected to 60 different shots. Some newscasts reduce their stories to headlines and sound bites. The more important, and often overlooked, aspect of heightened information density, however, is the increased amount of simultaneous information crammed into a single screen display. Well-known examples are the newscaster who is usually surrounded by a secondary frame—the box—and a variety of unrelated pictures and written information, ranging from the latest acts of inhumanity to weather forecasts, sports, and market results.[12] SEE 8.40

The idea behind such a display is to make a variety of information available to viewers and to let them pick and choose. But a newscast is not a newspaper or a magazine or even a Web page, where you are free to scan the information and then go back to the item that interests you. The temporal medium of television

8.40 High-density News Display
This frame from a news presentation shows a high-density information display. The information is not well structured, however, and the individual news items are hard to comprehend.

does not allow such scrutiny (unless you have recorded the news). Most often all you can do is take in as much of the information as you can, realizing that you will inevitably miss out on some if not most of it.

One of the problems is that once you have zeroed in on a particular item in a high-density display, your mental operating system tries to save your sanity by blocking out most of the information that is peripheral at that moment. For example, while you are reading a printed news item that is scrolling along the bottom of the screen, you are apt to miss what the anchor is saying and vice versa. Information density on-screen does not guarantee increased information processing by the viewer. In fact, much of the bombardment of simultaneously displayed images not only remains ineffective but often makes the viewer lose interest and tune out completely.

But can't you look at the pictures that relate to the news story while listening to the anchor's commentary? Yes, you can. This indicates that the degree of information density is not a problem so long as it is structured properly. We are now confronted with learning to parlay potential information overload into effective high-density communication. The following points give some guidance on dealing with simultaneously displayed multiple images. You may recall some of the multiple-screen principles from the discussion of aspect ratios in chapter 6.

Dividing the Screen: Graphic Blocks

If you do not have distinct, clearly marked secondary frames within the screen, you must organize the information into discrete information areas, or graphic blocks. Such blocks help the viewer comprehend the information quickly and easily. Once you have assigned the blocks a specific place on the video or computer screen, you should keep the same type of information in the same area in subsequent presentations. For example, if the weather information is in the lower-right corner of the screen, don't move it to the upper-left in the next newscast. Likewise, if you have placed specific navigation instructions for the first part of an interactive module on screen-left, we expect them to be there again for the second part. Such consistency in screen position helps the viewer locate the desired information quickly without having to scan the whole screen. **SEE 8.41**

8.41 Use of Information Blocks
Grouping information into multiple frames and text blocks makes it relatively easy to seek out and comprehend the desired information.

Dividing the Screen: Screens Within the Screen

As discussed in chapter 6, we can place a number of secondary frames of various aspect ratios within the primary video screen. To help the viewer assimilate the content of the secondary frames without scrutinizing each individually, you need to be aware of the major aesthetic principles of screens within the screen.

Secondary frame If you insert a secondary frame on the primary screen, such as the well-known box over the newscaster's shoulder, you need to be especially cognizant of the principal graphic elements operating in this new frame. SEE 8.42

Although the field reporter in the secondary frame has proper noseroom according to single-screen standards, her index vector seems blocked by the left edge of the secondary frame and also by the left border of the primary screen. Because her gaze is directed off-screen, she seems oddly disconnected, if not isolated, from the event in the primary screen space (the anchor). Let's reverse the index vector so that the reporter in the box looks toward the inside edge of the secondary frame and see whether this solves the problem. SEE 8.43 With the woman looking toward the inside (right) edge of the secondary frame, her index vector seems to flow uninhibited through the box and connect with the primary screen event. The anchor's story is now properly supported by the visual structure of the screen display.

Z-axis vectors in split screen We are accustomed to accepting the sequential z-axis shots of host and guest as converging index vectors. When we see a close-up of the host looking at the camera (and, by extension, at us) followed by a similar close-up of the guest, we have no problem believing that the two people are talking to each other. When they are placed in individual side-by-side boxes, however, such an effect is more difficult to achieve. Unless their original locations (studio for the host and remote for the guest) are clearly established, we may be unsure whether the boxed people are looking at us or at each other.[13] SEE 8.44

8.42 Isolation Through Index Vector Pointing Outward

In this secondary frame, the index vector points toward the outside (left) edge of the primary screen. Because there is no off-screen space on this side, the woman looks isolated from the news event.

8.43 Connection Through Index Vector Pointing Inward

When her index vector points toward screen-center, the field reporter looks connected to the news event.

8.44 Split-screen Z-axis Vectors
When two people are isolated in secondary frames, the convergence of their z-axis index vectors depends on context. They can be perceived as talking to each other, to an off-screen person, or to the viewers.

Temporal and spatial contexts When you use several secondary frames, their structural requirements vary, depending on whether they are used in a temporal or a spatial context.

If you use multiple screens to deal with temporal complexity, for example to indicate simultaneous events that occur at different locations, the graphic weight and the vectors within the secondary frames are of little structural concern. Because the various boxes represent events that occur in widely different locations, their structural separation within the confines of the primary screen may be an asset rather than a handicap. The connection among the frames is strictly temporal—all events take place at the same time. Still, the screen showing the boy running looks better with the motion vector (in this illustration, the index vector) pointing at screen center rather than at the left screen edge. SEE 8.45

8.45 Distribution of Secondary-frame Vectors in Time Context
When several simultaneous events are shown in separate, isolated secondary frames, the direction of the index and motion vectors within such frames is relatively unimportant.

If the secondary frames are used in a spatial context, such as people in different boxes talking to one another, the placement of the frames in the primary screen space and the direction of the index and motion vectors within each secondary frame become significant structural considerations. Let's assume that you have three equal-sized secondary frames placed side-by-side inside the primary video screen.

In such a three-frame arrangement, only the center frame can project lateral index and motion vectors into the adjacent frames. The two adjacent, or wing, frames on either side of the center frame can project their lateral vectors only toward the center frame but not in the opposite direction. Contrary to a single primary video screen, which operates with off-screen space on all sides, the secondary three-frame arrangement has only limited off-screen space. For example, if a person in the left wing frame looks left, and the person in the right wing frame looks right, we do not project their index vectors into off-screen space. Rather than imply that they are looking at something off-screen, their glance seems to be blocked by the outside frame edges. SEE 8.46 As soon as the people in the wing frames turn around and look toward the inside edges, however, they connect with the center frame—and even through the center frame with each other. SEE 8.47

8.46 Blocked Off-screen Space on Wing Frames
In this three-frame arrangement, we do not project the index vectors of the two outside people into off-screen space. The outside edges of the wing frames seem to block off-screen space to index and motion vectors.

8.47 Connected Frames
When index and motion vectors in the wing frames point toward the center frame, they have no difficulty penetrating the inside edges and even the center frame. The three-frame unit represents a unified space. The focus here is on the center frame.

If the persons in the left and center frames look at each other and the person in the right frame looks toward the outside screen edge, the converging vectors connect the left and center screens, but the right screen remains isolated. **SEE 8.48**

When all secondary frames show the people facing the camera and thus projecting z-axis index vectors, the viewer can be persuaded that they are talking to one another if the context suggests a conversation among these people. Such an arrangement is often used when several people report from various locations that would make switching live between action and reaction shots of everybody else extremely difficult. But they also may well talk to the viewer. Such a switch occurs when the newscaster first introduces the field reporters in their remote-location boxes and then, off-camera, asks each individual for specific comments. Their z-axis index vectors are no longer directed toward one another but are intended to converge on the off-screen host. **SEE 8.49**

With individual motion vectors, however, such a convergence switch will not transpire. Just imagine that all three secondary frames show cars speeding along the z-axis toward the camera. You would have a hard time conceiving their motion vectors as converging and that the cars are on a collision course. Despite some credible context, the cars will remain clearly separated.

8.48 Isolation Through Diverging Index Vectors

The diverging vectors in the center and right wing frames cause the right frame to become isolated from the other two.

8.49 Triple Z-axis Vectors

When all three frames display z-axis vectors, the target depends on context. If the people in the three frames are in a three-way conversation, it is the audio track that connects the people rather than the index vectors. Their z-axis vectors are more appropriate if they are talking to the viewer or somebody else off-screen.

SUMMARY

Structuring the two-dimensional field means making the interplay of screen forces work for rather than against you. The relevant processes and topics include stabilizing the field through distribution of graphic mass and magnetic force, stabilizing the field through vector distribution, stages of balance, object framing, unusual compositions, multiple screens, and dividing the screen.

One of the most basic ways of stabilizing the two-dimensional field is to balance the forces of graphic mass and magnetism of the frame. Each graphic mass (object occupying a certain amount of screen area) has a graphic weight determined by the size of the object, its basic shape and orientation, its location within the frame, and its color.

The magnetism of the frame becomes more powerful the closer an object is to the screen edges. The screen edges attract objects regardless of their size.

The most stable position of an object is screen-center, where graphic weight and frame magnetism are symmetrically distributed. If an object is on one side of the screen, you can balance it with an object of similar weight on the other side.

One of the most important structuring processes is the distribution of vectors. High-magnitude index and motion vectors generally override such forces as graphic weight and frame magnetism. High-magnitude index vectors, as well as motion vectors, require deliberate framing to cope with their structural force and to stabilize the field. You usually do this by giving the vector-producing object enough noseroom (for index vectors) or leadroom (for motion vectors). Although graphic vectors are of relatively low magnitude, you must arrange them so that they do not cause an imbalance in the shot.

There are two basic structural stages of balance: static and dynamic. The most stable form of a static balance is a symmetrical arrangement within the frame. In a dynamic balance, the graphic elements are asymmetrically distributed so that they lead to a higher-energy interplay of forces.

An unbalanced picture lacks aesthetic structure. The placement of all elements within the frame looks arbitrary.

When framing an object so that only part of it can be seen on-screen, you need to provide enough graphic cues to facilitate psychological closure in off-screen space. This enables viewers to mentally fill in the missing parts of the object outside the frame. Premature closure occurs when viewers apply closure to the partial image within the frame without extending the object into off-screen space. This is why you should avoid having the natural dividing lines of a person or an object coincide with the screen edges.

Illogical closure occurs when the visual elements of two unrelated objects provide enough continuation of graphic vectors to be perceived as a single configuration. This visual grouping of continuing vectors is so strong that we tend to group objects that in fact do not belong together, such as seeing a telephone pole as an extension of the person standing in front of it.

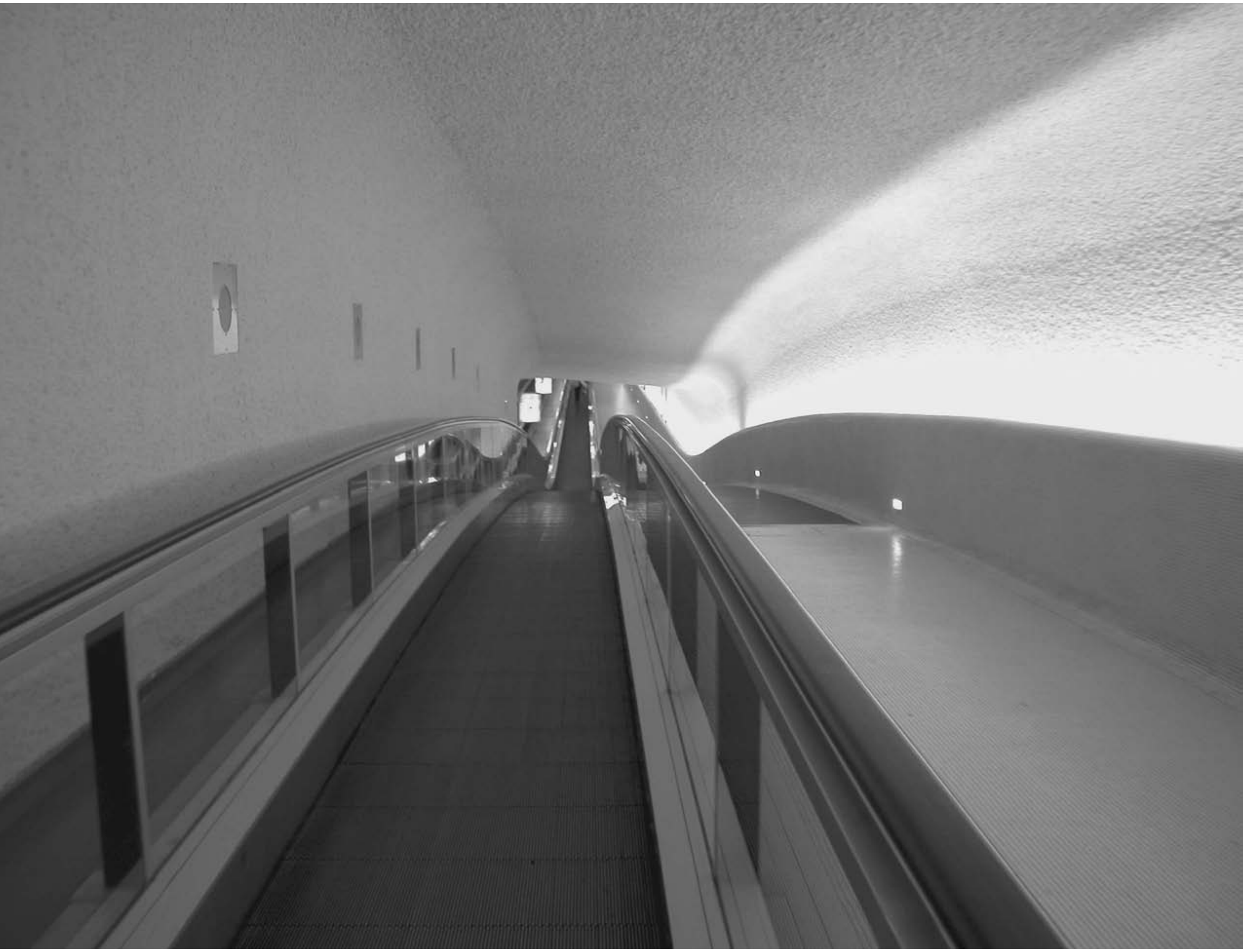
Unusual compositions that seem to ignore, or even flaunt, the established rules can be used to emphasize a particular screen area or intensify an event.

The use of multiple screens—or secondary frames within a primary screen—requires additional structural techniques. In contrast to a single screen, which has off-screen space in all directions, the outside edges of secondary frames seem to block index and motion vectors that do not point toward screen-center.

When two people appear in side-by-side boxes, you need to establish that they are talking to each other before the viewer accepts their z-axis index vectors as converging. If you want to show several events that are happening simultaneously in different locations, you can use variously shaped secondary frames within the primary screen. Careful vector continuity is not essential. When multiple screens show people who talk to one another, however, their index vectors must connect so that none of the screens remain isolated.

NOTES

1. The specific structural principles and demands of a shot series are discussed in more detail in chapters 14 and 17.
2. Credit for the "nose" index vector example goes to my colleague Mike Woal, who used it in his media aesthetics class in the Broadcast and Electronic Communication Arts Department at San Francisco State University.
3. *Leadroom* seems to be a more flexible term than *noseroom.* See Lynne S. Gross, James Foust, and Thomas Burrows, *Video Production: Disciplines and Techniques* (New York: McGraw-Hill, 2005).
4. Vector travel toward or through the screen edge was analyzed by Mark Borden in "On the Problem of Vector Penetration" (Media Aesthetics I, Broadcast and Electronic Communication Arts Department, San Francisco State University, March 1996, unpublished).
5. Rudolf Arnheim, *Toward a Psychology of Art* (Berkeley: University of California Press, 1966), p. 45.
6. The simplification of stages of balance was kindly suggested by William Deering, University of Wisconsin at Stevens Point, Wisconsin.
7. For more information on the Le Corbusier Modulor, see Le Corbusier [Charles E. Jeanneret-Gris], *Modulor,* 2nd ed., trans. by Peter de Francia and Anna Bostock (Cambridge, Mass.: Harvard University Press, 1954); and their *Modular 2* (London: Faber and Faber, 1958).
8. Translation by the author. The original reads: "Dann wahrhaftig steckt die Kunst in der Natur, wer sie herausreissen kann, hat sie.…Und durch die Geometrika magst du deine Werks viel beweisen." In Johnannes Beer, *Albrecht Dürer als Maler* (Königstein i.T, Germany: Karl Robert Langewiesche Verlag, 1953), p. 20.
9. Whereas the golden section or the rule of thirds may be useful for basic shot compositions, both Gerald Millerson and Peter Ward warn against its overuse because such "sameness" of composition may tire the viewer and defeat its intended dynamic tension. See Gerald Millerson, *The Technique of Television Production,* 13th ed. (Boston and London: Focal Press, 1999), pp. 141–42; and Peter Ward, *Picture Composition for Film and Television,* 2nd ed. (Boston and London: Focal Press, 2003). Bruce Block defends it because it helps maintain continuity in relative screen positions. See Bruce Block, *The Visual Story* (Boston: Focal Press, 2001), pp. 141–45.
10. See Irvin Rock, *Perception* (New York: Scientific American Library, 1984), p. 116. Also see Max Wertheimer, "Experimentelle Studien über das Sehen von Bewegung" (Experimental Studies About the Seeing of Motion), *Zeitschrift für Psychologie* 61 (1912): 161–265; and Kurt Koffka, *Principles of Gestalt Psychology* (New York: Harcourt, Brace, and World, 1935), p. 110.
11. In his 1934 film *Napoléon Bonaparte,* Abel Gance (1889–1981) used three 4 × 3 screens side-by-side to simulate an aspect ratio similar to today's wide-screen format—a process he called polyvision. See William Phillips, *Film: An Introduction* (Boston: Bedford/St. Martin's, 1999), p. 46.
12. Because multiple screens are so rarely seen in film presentations, this discussion is limited to the use of multiple screens in television, although similar multiscreen principles apply to the separate frames in Web page displays.
13. In a study that had three- to five-year-old children react to such second-order frame events, Georgette Communtziz-Page reports that the children were confused with whom the boxed people were communicating. See her "Comprehension of Visual Images in Television" in *Handbook of Visual Communication Research: Theory, Methods, and Media,* ed. by Ken Smith, Sandra Moriarty, Gretchen Barbatsis, and Keith Kenney (Mahwah, N.J.: Lawrence Erlbaum, 2005), pp. 211–23.



9

The Three-dimensional Field: Depth and Volume

In video, film, and computer displays, as in painting and still photography, we must project the three-dimensional (3D) world onto a two-dimensional surface. Fortunately, the camera and its single-lens optical system transact such a projection automatically. It is also fortunate that we are willing to accept such a projection as a true representation of all three dimensions: height, width, and depth. SEE 9.1

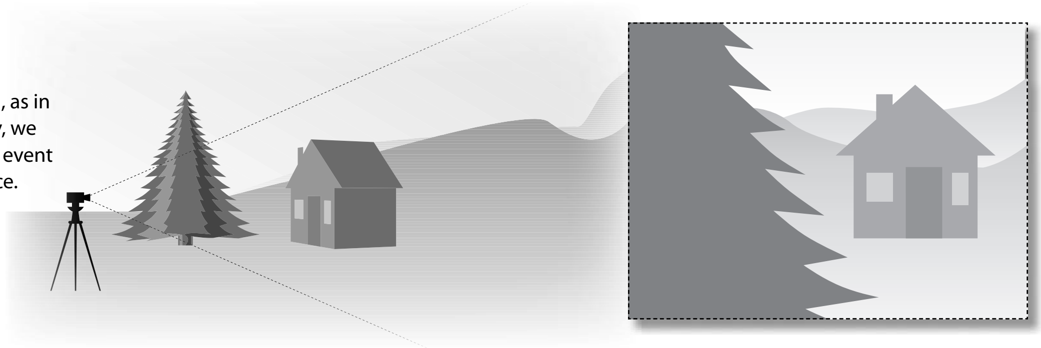
The synthetically constructed images of paintings and computer displays often copy the single-lens-generated depth cues to create the illusion of a third dimension. Although in video and film the camera and the event are often in motion, we can nevertheless draw on the firmly established and proven aesthetic rules of painting to help structure and manipulate this process of projecting a three-dimensional world onto a two-dimensional plane. Note, however, that these techniques do not show us a true stereoscopic projection of the world. Stereoscopy, which is an old technique of displaying the 3D world on a 2D plane, uses two-lens photography or computer programs that simulate our actual two-eyed perception process. Stereoscopic projection is making a comeback in film and has spawned a flurry of activity in video and digital animation.

In this chapter we examine four specific three-dimensional field areas: the z-axis, graphic depth factors, depth characteristics of lenses, and stereoscopic projection.

The Z-axis

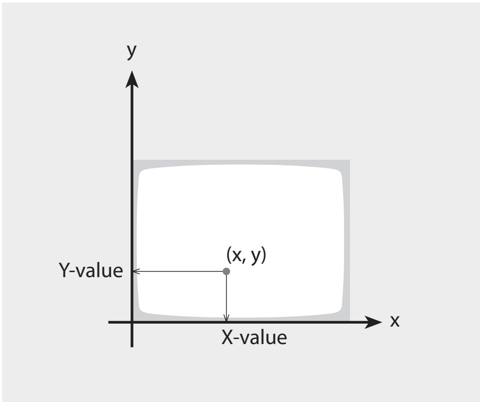
As you probably remember from geometry class, the x and y coordinates precisely locate a point in a two-dimensional plane such as the frame of the picture opposite this page, or the video screen. You can describe the width of the screen as the x-axis and the height of the screen as the y-axis. A point within the screen can be assigned an x-value, indicating its relative position along screen width, and a y-value, indicating its position relative to screen height. SEE 9.2

In the three-dimensional model, the z-axis is added, which describes depth. The ***z-axis*** value describes a point located away from the frontal plane—in our case, how far an object seems to be from the camera.[1] You learned about the z-axis in chapter 7 in the context of index and motion vectors that extend from the camera to the horizon and vice versa. SEE 9.3



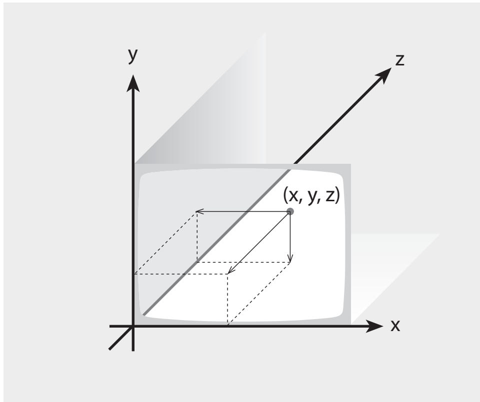
9.1 Projection of 3D Space onto a 2D Plane

In conventional video and film, as in painting and still photography, we project the three-dimensional event onto a two-dimensional surface.



9.2 X and Y Coordinates

The x and y coordinates locate a point precisely within an area, such as the screen. A point within the screen can be assigned an x-value, indicating where it is located on the x-axis (screen width), and a y-value, indicating its position on the y-axis (screen height).



9.3 Three-dimensional Model

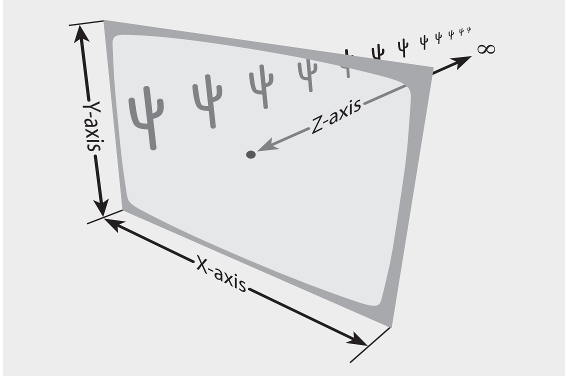
To locate a point precisely within a described volume, the z-axis, describing depth, becomes an essential dimension. The z-value describes how far a point is located away from the frontal plane (the screen).

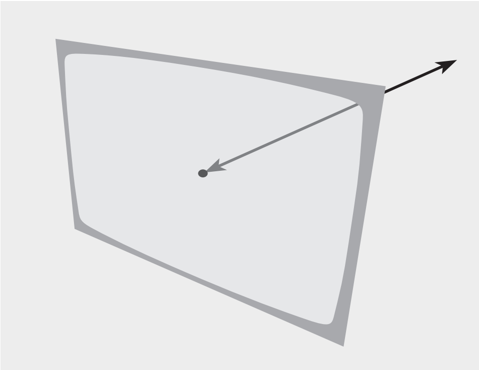
Amazingly enough, the illusory third dimension—depth—proves to be the most flexible screen dimension in film and especially in video. Whereas the screen width (x-axis) and height (y-axis) have definite spatial limits, screen depth (z-axis) is virtually infinite. **SEE 9.4**

Notice that without stereovision or hologram projection (as is the case with all single-lens films, video, and computer displays), we perceive the z-axis

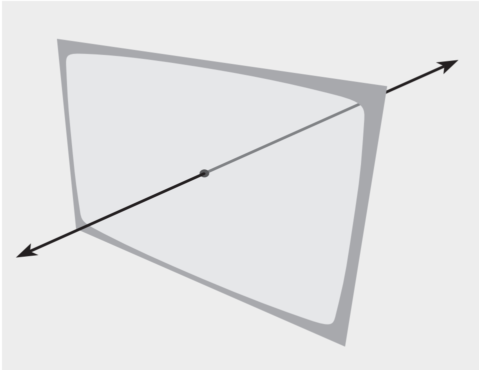
9.4 Z-axis Dimension

Although the z-axis—the depth dimension—is illusory in television and film, it is aesthetically the most flexible screen dimension.





9.5 Z-axis Directions
In normal lens-generated images, the z-axis stretches from the screen (camera lens) to the horizon. The z-axis is bidirectional: movement can occur from camera to horizon or from horizon to camera.



9.6 Z-axis Directions in Stereovision and Hologram Projections
In stereovision or a hologram, the z-axis extends not only from the screen surface back toward the horizon but also through the screen toward the viewer.

as originating from the screen and going backward, from the camera lens to the horizon. The closest object seems to lie on the screen surface; it does not extend toward the viewer. **SEE 9.5**

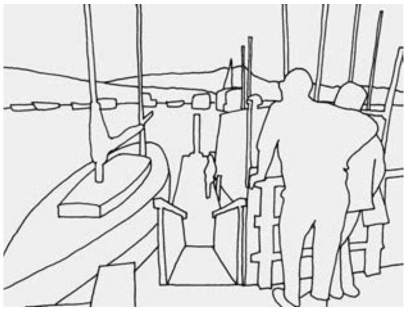
In stereovision or a hologram, the z-axis extends not only to the horizon but also to the viewer: objects appear to extend out from the screen toward the viewer. We judge their perceived distance relative to ourselves and not to the screen's surface. **SEE 9.6**

As explored in later chapters, the z-axis becomes an important element in structuring screen space and motion.

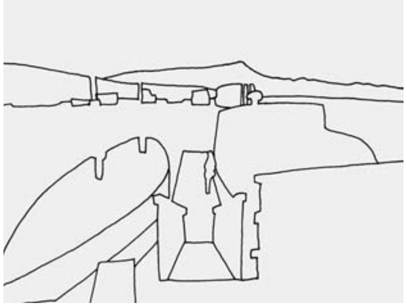
Graphic Depth Factors

But how, exactly, can we create the illusion of depth on the two-dimensional plane of the screen? **SEE 9.7** Examine the figure and try to identify the many factors that contribute to the illusion of depth on the 2D surface of the page.

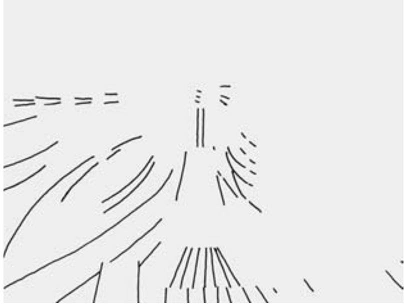
9.7 Graphic Depth Factors
What factors contribute to the illusion of depth in this picture?



9.8

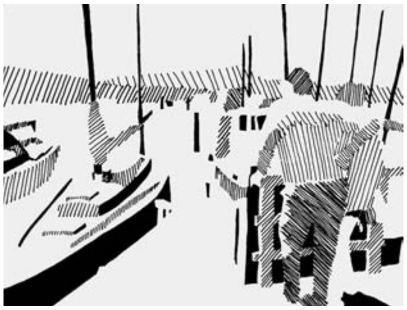


9.9



9.10

9.11



9.12

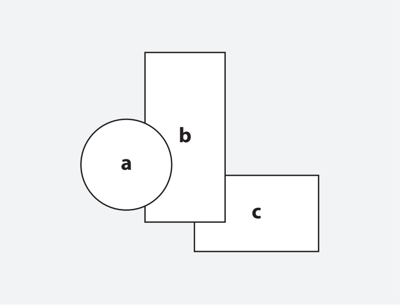
You probably noticed that some objects are partially hidden by other objects. **SEE 9.8** Also the farther away some objects are (boats, people, hills), the smaller they appear and the higher they seem to be positioned in the picture field. **SEE 9.9** Parallel lines, such as the edges of the boardwalk, appear to converge in the distance. **SEE 9.10** Objects in the foreground are more clearly defined than those in the background. **SEE 9.11** And, finally, the light and the shadows indicate volume, that is, the presence of a third dimension. **SEE 9.12** This section discusses five such ***graphic depth factors***: overlapping planes, relative size, height in plane, linear perspective, and aerial perspective.

OVERLAPPING PLANES

The most direct graphic depth cue is an ***overlapping plane***. When you see one object partially covered by another, you know that the one doing the covering must be in front of the one that is partially covered. **SEE 9.13 AND 9.14**

Courtesy of the Museo Civico di Padova.

Medieval painters relied heavily on overlapping planes to indicate depth. In this detail from *The Meeting at the Golden Gate,* one of the many excellent frescoes in the Arena Chapel in Padua, Italy, by Florentine painter and architect **Giotto di Bondone** (ca. 1267–1337), we can see how effectively overlapping planes were used to indicate depth. The only depth confusion arises from the merging halos of Joachim and Anna—caused by the fading of the contour of Joachim's halo at the point of overlap.



9.13 Overlapping-planes Principle
Object (a) is partially covering object (b), which is partially covering object (c). Although all three figures obviously lie on the same plane (this page), (a) seems to be in front of (b), which seems to lie in front of (c) but behind (a).

9.14 Depth Through Overlapping Planes
Any object that is partially blocked from our view by another object must lie behind that object. Even with other depth cues missing, we perceive a third dimension by readily assigning partially overlapping objects a "behind" or an "in front of" position.

RELATIVE SIZE

If you know how big an object is or can guess its size by contextual clues (such as other known objects), you can tell approximately how far the object is from the camera by the ***relative size*** of the screen image. The larger a subject or an object appears relative to the screen borders, the closer it seems to the viewer. **SEE 9.15** The smaller a subject or an object appears relative to the screen borders, the farther away it seems. This seems to be a hardwired response. **SEE 9.16**

If you know that two objects are similar or identical in size, you perceive the smaller screen image as being farther away and the larger screen image as being closer. **SEE 9.17** In the absence of contradicting contextual clues, we automatically interpret the smaller screen image of the man as being relatively far away from the woman, rather than being unusually small. The more comparable the head sizes, the closer the subjects seem to stand to each other along the z-axis. **SEE 9.18**

9.15 Relative Size: Close-up
The larger the object or subject appears within the screen—that is, the more area it takes up relative to the screen borders—the closer it seems to us. Appropriately, we call this framing a close-up.

9.16 Relative Size: Long Shot
The smaller the object or subject appears within the screen, the farther away it seems. We call this framing a long shot.

In this sixteenth-century Persian painting, overlapping planes and especially height in plane serve as major depth cues.

Courtesy of Editions d'Art Albert Skira.

9.17 Interpreting Object Size as Distance: Far

The man seems farther away from us than the woman because his screen image is considerably smaller than hers.

9.18 Interpreting Object Size as Distance: Close

The man seems much closer to the woman now because his screen image is almost as large as hers.

HEIGHT IN PLANE

Assuming that no contradictory distance cues are evident and that the camera is shooting parallel to the ground, you will perceive people and objects as being more and more distant the higher they move up in the picture field. This distance cue operates only until they have reached the horizon line. This is known as ***height in plane***. **SEE 9.19** Because of the mobility of the camera, however, which causes the horizon line to shift constantly within a shot or from shot to shot, the height-in-plane distance cue is not always reliable. **SEE 9.20** The fire escape near the top edge of the picture in figure 9.20 seems closer than the high-rise buildings at the bottom edge of the screen. Obviously, the camera did not shoot parallel to the ground. Especially when shooting up or down a large object, the height-in-plane cue is no longer a valid depth indicator.

9.19 Height in Plane: Camera Parallel to Ground

In the absence of contradictory distance cues and with the camera shooting parallel to the ground, the people seem farther away the higher up they move toward the horizon in the picture plane. As they line up along the x-axis, they do not move up in the picture plane any longer.

9.20 Height in Plane: Camera from Below

When the camera does not shoot parallel to the ground, height-in-plane distance cues are no longer valid.

LINEAR PERSPECTIVE

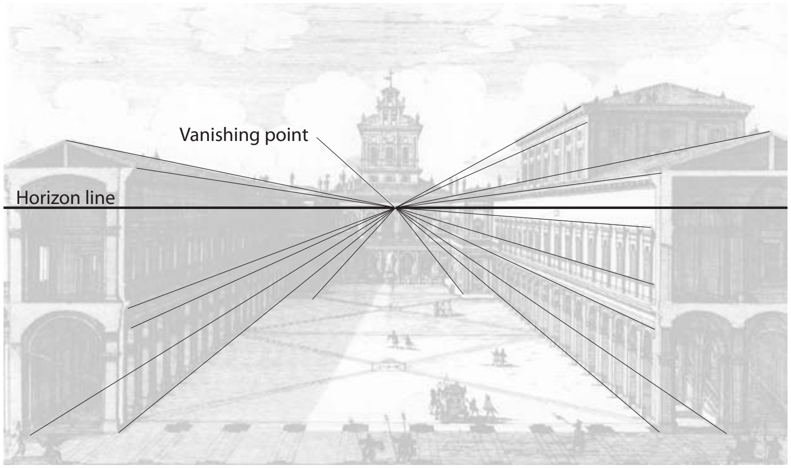
This is among the more powerful and convincing graphic depth factors. In a ***linear perspective***, all objects look progressively smaller the farther away they are, and parallel lines converge in the distance, with the vertical and horizontal lines becoming more crowded as they move away from the observer (camera). **SEE 9.21**

All parallel lines converge and stop or disappear at the ***vanishing point***, which always lies at eye level or camera level on the horizon line. **SEE 9.22**

9.21 Linear Perspective

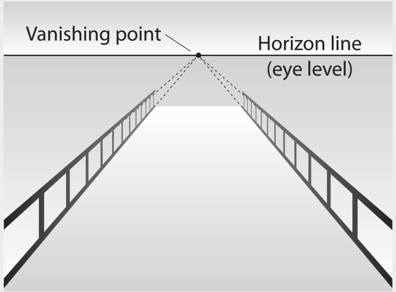
In this architect's drawing of an Italian palazzo, all the prominent horizontal lines (graphic vectors) converge at one point. We call this perceptual phenomenon linear perspective.

Courtesy of the DeBellis Collection, San Francisco State University.

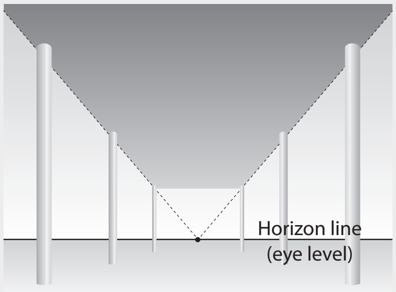


9.22 Vanishing Point

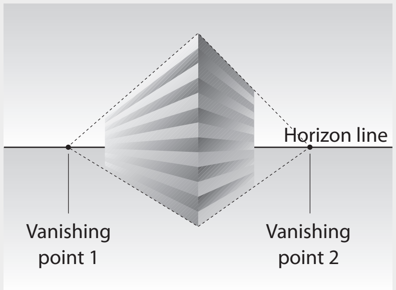
The point at which all parallel lines converge and discontinue (vanish) is aptly called the vanishing point. The vanishing point always lies at eye (or camera) level on the horizon line.



If we look down on an object, the eye (camera) level is above the object. Therefore the horizon line, and with it the vanishing point, lies above the object. We see the object—in this case, the bridge—from above.



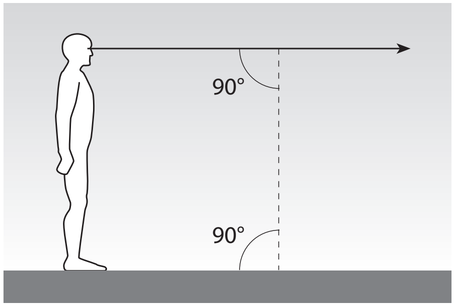
If we are below the object and look up, the eye (camera) level, the horizon line, and the vanishing point lie below the object. We see the object—the bridge—from below.



If we see two sides of a building, with one corner closest to us, we perceive two sets of converging lines going in opposite directions. We have, therefore, two vanishing points. But note that the two vanishing points lie on the horizon line; after all, we look at both sides of the building from the same eye level. This is called two-point perspective.

Attempts at using linear perspective to create the illusion of depth on a two-dimensional surface were made by many painters long before the Renaissance, but it was not until the first half of the fifteenth century that Italian artists established scientific laws of linear perspective, such as the horizon line and the vanishing point. With this woodcut one of the masters of the Renaissance, German painter **Albrecht Dürer**, illustrated some of the techniques used by the artist to ensure correct foreshortening (linear perspective).

Taken from *The Complete Woodcuts of Albrecht Dürer,* edited by Dr. Willi Kurth, republished in 1963 by Dover.



9.23 Horizon Line

The horizon line is an imaginary line parallel to the ground at eye level. More technically, it is the plane at right angles to the direction of gravity that emanates from the eye of the observer at a given place. If you want to find the eye level and the actual horizon line, simply stand erect and look straight forward. The eye level and the horizon line are in the distance where you do not have to look up or down.

To find the ***horizon line*** and ***eye level***, simply stand erect and look straight forward or point the camera parallel to the ground. Assuming your index vector runs parallel to the ground, the horizon line moves up or down with your eyes (camera) regardless of whether you are kneeling on the ground, standing on a ladder, or pointing the camera out a helicopter window. **SEE 9.23**

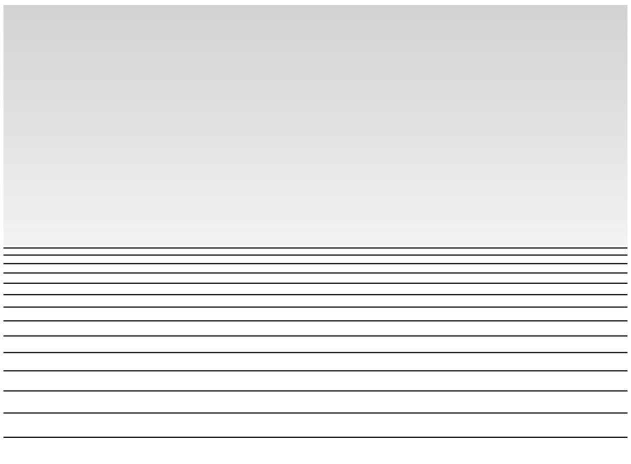
Now take another look at figure 9.21. Can you tell from which height the artist looked at the palazzo? Was he looking at it from the street level? Sitting in a chair? Standing up? Perhaps from the balcony or window of an unseen building opposite the clock tower?

If you chose the window or balcony, you estimated the artist's correct position. As you can clearly see, the parallel lines converge at a vanishing point that lies above the palazzo near the roofline of the clock tower building (see figure 9.22). The artist must therefore have looked at the building from that position.

Also note that the arches and the windows of the building seem to lie closer together the farther away they are from the observer (see figure 9.21). Many painters have used this crowding effect, or texture, to simulate depth. **SEE 9.24** You can apply this principle just as effectively with computer-generated graphics. **SEE 9.25**

9.24 Crowding Effect Through Texture

Notice how the sunflowers appear more and more crowded the farther away they are from the camera. This crowding effect is an important depth cue.



9.25 Depth Through Crowding

In this computer-generated image, we perceive depth through the crowding effect of distant objects.

Forced perspective Because we tend to interpret image size and convergence of lines with relative distance, we can generate the impression of distance by having parallel lines converge "faster"—more readily—and make distant objects appear smaller than we would ordinarily perceive. Such an artificial forcing of linear perspective is called, appropriately enough, ***forced perspective***. One of the more striking applications of such a forced perspective is the grand staircase in one of Hong Kong's luxury hotels. The wide staircase seems to curve up to the mezzanine in a long, impressive sweep. But when you actually climb the stairs, you will notice that they gradually narrow to less than half of their original width about a third of the way up. As you will discover in this chapter, we can achieve the same effect with the proper choice of lenses.

AERIAL PERSPECTIVE

A certain amount of moisture and dust is always in the atmosphere. We therefore see objects that are close to us somewhat more sharply than those farther away, a phenomenon known as ***aerial perspective***. **SEE 9.26** In fog this difference in sharpness and image density between foreground and background is especially pronounced. Colors also lose their density and become less saturated the farther away they are from the observer (camera). Outdoors, distant colors take on a slightly blue tint.[2]

When creating scenery, for example, you should paint the background objects slightly bluer and less sharp than the foreground objects. This greatly enhances the illusion of depth.

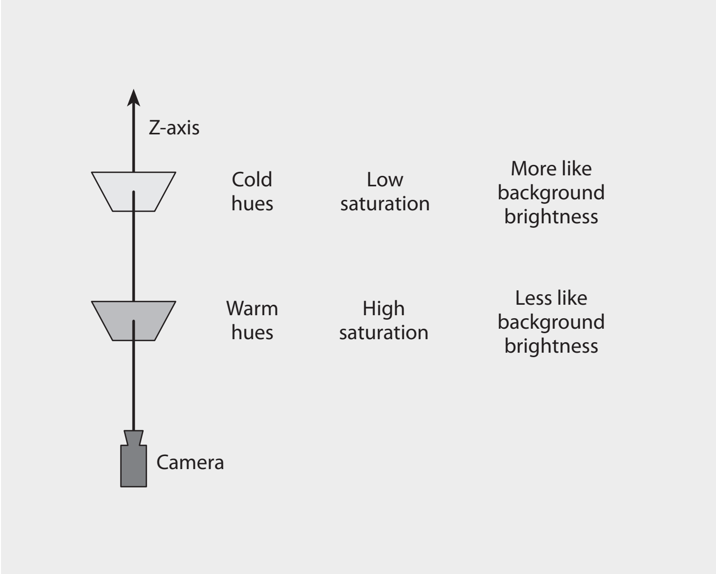
Generally, warm hues seem to advance and cold hues seem to recede. Highly saturated colors seem closer than less saturated colors. Assuming a fairly dark background, the brighter colors (higher position on the brightness scale) seem closer, and the less bright colors (lower position on the brightness scale) seem farther away. We can say that the more the brightness of an object assumes the

9.26 Aerial Perspective

Notice how the foreground objects in this picture are relatively sharp and dense and that the background objects become progressively less clear and less textured the farther away they are.

9.27 Aerial Perspective and Color

So far as aerial perspective is concerned, objects with warm, highly saturated colors that are the opposite of the background brightness seem closer to the viewer than objects with cold, less saturated colors that are similar to the background brightness.



brightness of the background, the farther away from the observer (the camera) it appears. SEE 9.27

Depth Characteristics of Lenses

The optical characteristics of lenses can greatly enhance or hinder the illusion of a third dimension on the video or movie screen. Moreover, your choice of lens is important in achieving the certain "feel" of a screen event—whether buildings or objects look squeezed or stretched or whether the z-axis looks compressed or elongated. Synthetic computer images generally simulate the depth characteristics of lenses in their manipulation of the third dimension.

Before discussing the psychological impact of such space manipulation, we will first review the basic depth characteristics of wide-angle and narrow-angle lenses.

Note that our use of the terms *wide-angle* (or *short-focal-length*) and *narrow-angle* (or *long-focal-length* or *telephoto*) lenses includes the wide- and narrow-angle positions of the zoom lens. To put the zoom lens in the extreme wide-angle position, you zoom all the way out. To put it in the extreme narrow-angle, or telephoto, position, you zoom all the way in. The so-called normal lens position is in the middle of the zoom range. For emphasis we concentrate here on the extreme wide-angle and narrow-angle lens positions.[3]

The focal length of lenses influences four principal graphic depth factors: overlapping planes, relative size, linear perspective, and aerial perspective.

OVERLAPPING PLANES: WIDE-ANGLE LENS

Although the wide-angle lens does not rely on overlapping planes as much as the narrow-angle lens does, it can't avoid showing them. Because the objects along the z-axis look more stretched out with the wide-angle lens, it renders overlapping planes less essential as a depth indicator. SEE 9.28

OVERLAPPING PLANES: NARROW-ANGLE LENS

The narrow-angle lens does just the opposite of the wide-angle lens, making objects appear closer together along the z-axis than they really are. Because the

9.28 Overlapping Planes: Wide-angle Lens
Overlapping planes are reduced in prominence, but are not eliminated, with the wide-angle lens.

9.29 Overlapping Planes: Narrow-angle Lens
With a narrow-angle lens, overlapping planes are a major depth cue.

narrow-angle lens enlarges the background objects where things look crowded, foreground and background objects look similar in size. Consequently, they appear closer together than they really are. Objects positioned along the z-axis look squeezed, and the z-axis itself appears shorter than you would ordinarily see.

Because of the similarity in size of foreground and background objects, overlapping planes become a major depth cue for separating one object from another and, ultimately, separating foreground, middleground, and background objects. **SEE 9.29**

RELATIVE SIZE: WIDE-ANGLE LENS

The wide-angle lens greatly exaggerates relative size. Objects that lie close to the camera photograph as relatively large, whereas similar objects positioned on the z-axis only a short distance behind the close object show up in a dramatically reduced image size. **SEE 9.30** The image size of the foreground tugboat is relatively large, and the one just a short distance behind it is relatively small. This great difference in relative size is lessened only at the far end of the z-axis.[4] Because image size is an important distance cue, we interpret this difference as meaning that the background object is farther behind the foreground object than it really is. Thus the wide-angle lens stretches the virtual z-axis.

9.30 Relative Size: Wide-angle Lens
The wide-angle lens makes objects close to the camera look relatively large and those just a short distance farther away on the z-axis look relatively small. Because relative size is an important distance cue, the tugboats look farther apart than they really are.

9.31 Relative Size: Narrow-angle Lens
The narrow-angle lens enlarges the background image so drastically that the tugboats as well as the background hills seem much closer together than in figure 9.30, although their actual positions along the z-axis have not changed.

RELATIVE SIZE: NARROW-ANGLE LENS

When the same scene is photographed with a narrow-angle lens, the two boats seem much closer to each other. **SEE 9.31** This is despite the fact that their actual distance is identical to that in figure 9.30. You now know why. The narrow-angle lens enlarges the background, making the second object appear close in size to that of the foreground object. We translate this similarity in size as relative proximity. The narrow-angle lens shows objects placed along the z-axis squeezed; the z-axis therefore appears shortened.

This compression effect of the long (or narrow-angle) lens is very apparent when you shoot the same row of columns with both a wide-angle and a narrow-angle lens. Using the wide-angle lens, the columns quickly diminish in size the farther away they are from the camera; they seem comfortably stretched out. **SEE 9.32** But when you shoot the same scene using a narrow-angle lens, the image size of the background columns is almost the same as that of the foreground columns. They now seem closer together than they really are; they no longer feel graceful but instead look massive and crowded. **SEE 9.33**

LINEAR PERSPECTIVE: WIDE-ANGLE LENS

The wide-angle lens accelerates the convergence of parallel lines; that is, they seem to converge more quickly than when seen normally, thereby giving the illusion of stretching an object or a building. The z-axis space appears elongated. **SEE 9.34**

9.32 Stretching with a Wide-angle Lens
This row of columns seems quite long, and the columns seem to be a comfortable distance from one another.

9.33 Compressing with a Narrow-angle Lens
With a narrow-angle lens, the columns appear very close together. The space between them seems squeezed.

LINEAR PERSPECTIVE: NARROW-ANGLE LENS

In contrast, the narrow-angle lens inhibits the normal convergence of parallel lines and thus reduces the illusion of depth through linear perspective. It also squeezes space and makes the doors appear narrower and closer together than they really are. **SEE 9.35**

By now you should have no problem distinguishing between the wide-angle and narrow-angle shots of a piano keyboard. **SEE 9.36 AND 9.37** The wide-angle lens makes the graphic vectors of the keyboard converge much more drastically than when shot with a narrow-angle lens. The "wide-angle" keyboard looks longer, with the keys farther away from the camera looking distinctly smaller. The "narrow-angle" keyboard, on the other hand, does not seem to converge much toward a vanishing point. In fact, the keys farthest from the camera look almost as big as the ones closest to it. This makes the keyboard look short and squeezed. Were you to watch somebody playing quick runs up and down the keyboard, the wide-angle lens would exaggerate such dexterity; the narrow-angle lens would reduce such motion and, with it, the pianist's virtuosity.

9.34 Linear Perspective: Wide-angle Lens
The wide-angle lens makes parallel lines converge much "faster" (more drastically) than when seen normally.

9.35 Linear Perspective: Narrow-angle Lens
The narrow-angle lens "retards" our normal expectations of parallel lines converging. The horizontal lines do not converge as readily as with a normal or a wide-angle lens.

9.36 Piano Keys: Wide-angle Lens
When shot with a wide-angle lens, the piano keys reduce drastically in size the farther they are from the camera.

9.37 Piano Keys: Narrow-angle Lens
When shot with a narrow-angle lens, the same keys look squeezed. The piano keys at the far side of the z-axis look almost as big as the ones that are close to the camera.

WORKING WITH AERIAL PERSPECTIVE

You can achieve aerial perspective by manipulating the ***depth of field***—the area along the z-axis that appears in focus—and by making use of selective focus, that is, focusing on only a specific area along the z-axis.

When objects are placed at different distances from the camera along the z-axis, some of them will appear in focus and some will be out of focus. The portion of the z-axis in which the objects appear in focus—depth of field—can be shallow or great. In a shallow depth of field, only a relatively small portion of the z-axis shows objects in focus. In a great depth of field, a large portion of the z-axis shows objects in focus. The depth of field depends on the focal length of the lens, the lens aperture (iris opening), and the distance from the camera to the object.

Assuming that you shoot with a wide-angle lens under normal light levels and do not move the camera extremely close to the target object, the depth of field will be great. A narrow-angle lens gives a shallow depth of field. Generally, wide shots have a great depth of field; close-ups have a shallow one. Take another look at figures 9.36 and 9.37. The wide-angle lens shows the whole keyboard in focus. When shot with the narrow-angle lens, the keys closest to the camera are out of focus because the camera was focused on the middle part of the keyboard.

Aerial perspective: wide-angle lens Because the wide-angle lens generates a great depth of field, it de-emphasizes aerial perspective. In a great depth of field, most of the articulated z-axis appears in focus. This means you cannot easily focus on only one spot along the z-axis while keeping the rest of the z-axis out of focus (see figure 9.36). A great depth of field is obviously advantageous when covering news, where you normally have little time to achieve optimal focus. Although a misnomer, a great depth of field is also called deep focus.

Aerial perspective: narrow-angle lens The narrow-angle lens has a shallow depth of field and thus emphasizes aerial perspective. Once you focus on an object using a narrow-angle lens, the areas immediately in front and in back of the object are out of focus. Even a slight position change of camera or object along the z-axis will necessitate refocusing (see figure 9.37).

Although it is difficult to keep a moving object or camera in focus in a shallow depth of field, the advantage of this critical focal plane is that you can use selective focus to emphasize events. You have probably noticed that shooting in a shallow depth of field has become stylish in video and film production. For example, you may see two out-of-focus people walking along the z-axis toward the camera until their images become focused in the depth of field. Such aerial-perspective maneuvers are often accompanied by a similar audio manipulation: when the people are out of focus, you can barely make out what they are saying; but once they are in focus, their dialogue becomes loud and clear. In an over-the-shoulder shot, you may initially see in focus the shoulder and the head of the camera-near person, but the camera-far person, who is facing the camera, is out of focus. The camera will then switch the focus to the camera-far person, with the camera-near person being out of focus.

Selective focus The technique of ***selective focus*** allows you to choose the precise portion (plane) of the z-axis that you want to be in focus, with the areas immediately in front of or behind the focused object being out of focus. Contrary to a natural aerial perspective that occurs on a foggy day—where only the foreground object is "in focus," that is, more clearly visible than the background objects—the optically induced aerial perspective using selective focus allows you to move the focused plane from the foreground to the middleground or background or the other way around. In the next figure, we start out with the focus on the foreground

9.38 Selective Focus: Person in Front
A narrow-angle lens is used to create a shallow depth of field that allows selective focus. Note how the focus is on the person closest to the camera, with the people behind out of focus.

9.39 Selective Focus: Person in Middle
The middleground person is in focus, with the foreground and background persons out of focus.

9.40 Selective Focus: Person in Back
The person farthest from the camera is in focus, with the middle and front persons out of focus.

person, with the people in the middleground and the background out of focus. **SEE 9.38** Or you can feature the middleground and leave the foreground and the background out of focus. **SEE 9.39** You can also focus on the background, while the middleground and the foreground remain out of focus. **SEE 9.40**

Rack focus The ***rack focus*** effect involves changing the focus from one location on the z-axis to another. If, for example, you want to shift the emphasis from a spray can to the person holding it without changing the shot (through a dolly, zoom, cut, or dissolve), you can first focus on the spray can with the person out of focus and then "rack through" the focus range of the lens until the person's face comes into focus, throwing the spray can out of focus. **SEE 9.41 AND 9.42** Obviously, you need a relatively shallow depth of field to accomplish such a rack focus effect, which means that you must use a narrow-angle lens.

If you had a great depth of field (wide-angle lens with a small lens aperture), you could just about rack through the entire focus range without noticeably affecting the focus. A rack focus effect is therefore not possible in this case. With a narrow-angle lens, on the other hand, the depth of field becomes so shallow that even a slight racking of focus shifts the focal plane from one point along the

9.41 Rack Focus Effect: Object Emphasized
In this shot the focus is on the spray can. The shallow depth of field renders out of focus the person holding the can.

9.42 Rack Focus Effect: Person Emphasized
Emphasis has shifted from one z-axis location (the spray can) to another (the person). Because the depth of field is shallow, we can shift focus from the spray can to the person by changing (racking through) the camera's focus.

9.43 Lens Characteristics and Depth Cues

! A + sign indicates that the lens characteristic is facilitating the illusion of depth; a – sign indicates that it inhibits the illusion of depth.

Depth Effects	Lens Position	
	Wide-angle	Narrow-angle
Overlapping planes	–	+
Relative size	+	–
Linear perspective	+	–
Aerial perspective	–	+

z-axis to another. This means that a little adjustment of the focus control shifts the focus from one object to the other, even if they are only a short distance from each other along the z-axis. The table above summarizes how lenses influence our perception of depth. SEE 9.43

You can also achieve a type of aerial perspective by using "fog filters" that render portions of the picture out of focus while keeping other portions sharp. Though the filter does not actually distinguish among different z-axis locations, but rather among picture areas that are in and out of focus, we still perceptually interpret this as changes in the picture depth.

3D Stereoscopic Projection

Stereoscopic displays differ technically and aesthetically from the standard representation of a third dimension we have just explored. Whereas the traditional (single-lens) depth cues are contained in the visual representation of a scene, the stereoscopic projection is strictly illusory and exists only in your mind. We can trace and measure the elements of standard 3D clues on a photo (see 9.22), but not the ones of a stereoscopic projection. The stereoscopic effect is truly virtual.

STEREOSCOPIC 3D VERSUS STANDARD 3D: TECHNICAL DIFFERENCE

The basic technical difference between standard 3D and stereoscopic 3D is that standard 3D is based on a single-lens, or one-eyed, view of an event and stereoscopic 3D is based on a dual-lens, or two-eyed, view. Because our eyes are some distance apart, we see an object from slightly different angles; this is called binocular disparity. The easiest way to achieve this disparity with some consistency is to use a two-camera rig, with the two lenses set apart much like our eyes. Some devices use two lenses but only one camera. In this case, the two images are recorded side by side on the same medium. This disparity gives us important distance cues and enables us to engage in stereo vision, provided we have a suitable mechanism that directs one of the two pictures into one eye and the slightly offset picture into the other eye.

Such mechanisms can take various forms and are constantly being refined for film, video, and computer projections. The most basic one is the stereoscope, which was developed in the early 1830s. You have most likely seen or even used

9.44 Stereoscope

In the traditional stereoscope, the viewer directs the left eye to the left-eye photograph point of view of a scene, and the right eye to the right-eye point of view.

one of these contraptions, whereby you look through two different openings dedicated to each of the two slightly differing scenes (usually tinted with monochrome sepia to make them look old). Each eye is blocked from seeing the other's point of view.[5] **SEE 9.44**

This principle of making one of your eyes see an event from one point of view and the other eye from a slightly offset one is still used in the anaglyph process, in which the left-eye image of a scene is given a specific color (usually red), and the right-eye image is given its complementary color (cyan). Both images are slightly offset to simulate the different points of view of each eye. When looking at this double red/cyan image with stereoscope glasses, the left red lens lets you see the red part of the double image with your left eye, but not with the right one, which is blocked by the cyan-filtered image. The right cyan lens, on the other hand, lets you see the cyan-colored image with your right eye, but blocks most of the red image.

You probably noticed that we are dealing here with basic subtractive color mixing: the red filter lets no light of the complementary color pass through and renders the entire cyan part of the stereoscopic image black. The complementary cyan filter does the same thing to the red light, rendering the red part of the double image almost invisible to the left eye. Whatever colors are used in the stereoscopic glasses, they must be complementary to effect subtractive color mixing and, thereby, prevent both images from simultaneously entering either eye. Careful color correction is required for all techniques so that what you see through your rose-colored glasses represents the normally colored world.

There are many other techniques of displaying the two offset images—such as using polarized rather than colored filters, slightly offset fields in interlaced video scanning, or pixels that achieve the binocular disparity through a slight pixel shift[6]—but they all work on the same stereographic principle: to present the same scene in two slightly offset images: one as our left eye would see and the other as our right eye would see. The stereoscopic effect happens not on the screen but in the optical system of your brain.

STEREOSCOPIC 3D VERSUS STANDARD 3D: AESTHETIC DIFFERENCE

When you view a stereoscopic 3D projection, the event is not contained on the flat screen, as with video or film, but comes *toward you*—it seems to float in front of you. As mentioned at the beginning of this chapter, the major aesthetic difference between the simulated 3D space of the standard one-lens projection of video and film and the two-lens stereoscopic 3D projection is how we perceive the z-axis of the projected event (see figures 9.5 and 9.6). In the standard single-lens 3D, the z-axis, and with it the 3D articulation, points from the screen surface back toward the event horizon. In the ***stereoscopic projection***, the z-axis extends both ways, from the screen surface to the horizon but also from the screen surface toward the viewer. In fact, the screen as such no longer functions to separate the audience from the projected event, and the stereoscopic event is virtually dumped into the viewer's lap.

This phenomenon is, of course, a great advantage, if not a necessity, when simulating an event for students who want to learn how to fly, operate a certain piece of machinery, or perform surgery. With stereoscopic 3D the learner doesn't merely look at a 3D picture of an event but actually shares the 3D space. Coupled with interactive controls, this sharing facilitates his or her participation in the event, which comes about as close to the real thing as currently possible.

But this space sharing of audience and stereoscopic event may also be the deeper cause for the apathy the audience has so far shown toward stereoscopic movies, video, and the holograms of still shots. Even if we can eventually do away with the colored glasses and improve the resolution and the color fidelity of the stereoscopic image, we may still resist, however subconsciously, having our comfort zone invaded by stereoscopic film or video events. The extended z-axis, which hurls the event not only from horizon to the screen but *through* the screen and into our face leaves us psychologically, if not psychophysically, unprotected. We are coerced to confront the event literally head-on and somehow participate in it. Bereft of any aesthetic distance between ourselves and the simulated event, we are no longer comfortable watching violent or intimate scenes from a distance while eating popcorn but have the hero or villain step off the stage and sit next to us in the audience. Even in a holographic still image, such as a postcard from Hawaii, the event is no longer contained within its borders or a picture frame but invades our personal space, clamoring for attention.

The renewed and combined efforts of the entertainment industry and the equipment manufacturers to resurrect stereoscopic movies and video will, at least initially, be most successful if the content is nonthreatening, such as cartoons, or predisposed to participation, such as sports.[7] We will then see whether the new superhigh-definition stereo camera systems can remove "old boundaries, freeing us to take audiences on fantastic adventures and tell incredible stories in a whole new way."[8] But we will have a hard time accepting a character's escaping a highly charged emotional scene by stepping through the screen and walking toward us. SEE 9.45

There is also some serious concern about the inevitable eye (and brain) strain resulting from making the two slightly offset images appear as a single 3D image.[9]

What deserves special attention in such stereographic presentations is the audio. So far, the spatial manipulation of stereo audio occurs mostly sideways, along the x-axis, and to a limited extent along the z-axis that stretches from the screen to the horizon. Even 5.1 surround sound has only a limited reach when it comes to extending the z-axis from the screen to the viewer. The sound technique that most successfully defines the z-axis space from horizon to screen and from screen to viewer/listener is binaural audio.

9.45 Invasion of Viewer Space
The stereoscopic 3D event extends from the screen toward the viewer, invading the viewer's space.

Binaural audio is recorded with two microphones stuck in a dummy head, approximately where the ears would be, and is reproduced through headphones (see "Outer Orientation Functions of Sound" in chapter 15). Yes, we used a similar method by using two lenses, separated by the approximate distance of our eyes, when recording video for stereoscopic images. But don't worry. By the time stereoscopic digital cinema and television are in full swing, we probably won't need to wear earbuds in addition to the two-colored glasses every time we switch to a stereoscopic program.

SUMMARY

In video, film, and computer displays, the three-dimensional world must be projected onto the two-dimensional surface of the screen. Although the third dimension (the z-axis) is illusionary, it proves to be aesthetically the most flexible screen dimension.

Three concepts describe the aesthetics of the three-dimensional field: the z-axis, graphic depth factors, and the depth characteristics of lenses.

Whereas the screen width (x-axis) and height (y-axis) have definite spatial limits, screen depth (z-axis) is virtually limitless. The camera is therefore much less restricted in its view and movement along the z-axis than it is either horizontally or vertically. We perceive the z-axis as originating from the screen, extending toward the horizon. In stereovision and hologram displays, the z-axis extends toward the viewer as well.

Graphic depth factors include overlapping planes (objects that partially overlap each other); relative size (an object positioned close to the camera projects a larger screen image than one of similar size that lies farther down on the z-axis); height in plane (assuming that the camera shoots parallel to the ground, we perceive an object that is higher on the screen as farther away from the camera); linear perspective (horizontal parallel lines converge toward a vanishing point at the eye-level horizon line; equally spaced objects appear to lie closer together the farther away they are from the camera); and aerial perspective (the foreground in focus with the background out of focus). More generally, aerial perspective means selected focus on a spot along the z-axis in a shallow depth of field with the surrounding areas out of focus.

The depth characteristics of lenses are significant in the manipulation of the third dimension of a screen image. Wide-angle lenses exaggerate relative size and linear perspective; they de-emphasize overlapping planes and aerial perspective. Narrow-angle (telephoto) lenses exaggerate overlapping planes and aerial perspective, de-emphasizing relative size and linear perspective. Selective focus and rack focus effects are powerful means of articulating the z-axis and drawing attention to a specific plane along the z-axis. They are possible only in a shallow depth of field and therefore with narrow-angle (telephoto) lenses.

A stereoscopic 3D effect differs from the standard 3D representation on a screen in how we perceive the z-axis of picture space. In the standard 3D simulation, the graphic depth factors define the z-axis from the screen back to the horizon; in the stereoscopic projection, the articulated z-axis extends through the screen space toward the viewer. It is generated by a two-lens recording system, whereby the lenses are set apart similar to our eyes; we get a left-eye point of view and a right-eye point of view of the same scene. To show both slightly offset points of view on a single screen, one is colored red and the other is cyan. When the viewer wears glasses with a red lens for the left eye and a cyan lens for the right eye, each eye sees its dedicated point of view. The optical system in the viewer's brain melds the two images into a single three-dimensional projection, which extends through the screen toward the viewer.

NOTES

1. The x-, y-, and z-axes are used here in the traditional sense of the Cartesian model that quantifies Euclidian space.

2. Although we call both phenomena "aerial perspective," there is a difference between aerial perspective and detail perspective. Detail perspective refers to the gradual diminishing of detail in the distance, whereas aerial perspective refers to the more bluish tint the farther away the object is from the observer.

   Leonardo da Vinci described vividly what happens in aerial perspective:

   > There is another kind of perspective which I call aerial, because by the difference in the atmosphere one is able to distinguish the various distances of different buildings when their bases appear to end on a single line, for this would be the appearance presented by a group of buildings on the far side of a wall, all of which as seen above the top of the wall look to be the same size; and if in painting you wish to make one seem farther away than another you must make the atmosphere somewhat heavy. You know that in an atmosphere of uniform density the most distant things seen through it, such as the mountains, in consequence of the great quantity of atmosphere which is between your eye and them, will appear blue, almost of the same colour as the atmosphere when the sun is in the east. Therefore you should make the building which is nearest above the wall of its natural colour, and that which is more distant make less defined and bluer; and one which you wish should seem as far away again make of double the depth of blue, and one you desire should seem five times as far away make five times as blue. And as a consequence of this rule it will come about that the buildings which above a given line appear to be of the same size will be plainly distinguished as to which are the more distant and which larger than the others.

   Leonardo da Vinci, *The Notebooks of Leonardo da Vinci*, ed. and trans. by Edward McCurdy (Old Saybrook, Conn.: Konecky and Konecky, 2003), p. 880.

3. Herbert Zettl, *Television Production Handbook,* 10th ed. (Belmont, Calif.: Wadsworth, 2009), pp. 97–98.

4. Actually, the narrow-angle lens simply enlarges the end of the z-axis, where even with a wide-angle lens objects look crowded and space is squeezed. This crowding effect is entirely in line with increased density of texture at the far end of the z-axis.

5. The first commercial use of stereoscopic 3D is credited to Sir Charles Wheatstone, who introduced binocular vision and his stereoscope to the Royal Scottish Society of Arts in 1853. During the second half of the nineteenth century, stereoscopy rapidly gained in popularity and became a huge commercial success. Producing stereo photos and viewing them with the stereoscope become a huge commercial success and a popular pastime. For more information about stereoscopy photography, see Ray Zone, *Stereoscopic Cinema and the Origins of 3-D Film 1838-1952* (Lexington: University Press of Kentucky, 2008); and his *3-D Filmmakers: Conversations with Creators of Stereoscopic Motion Pictures* (Lanham, Maryland: Scarecrow Press, 2005), part of the Scarecrow Filmmakers series.

6. Trevor Boyer, "Stereo Hype," *Digital Content Producer* 35, no. 2 (February 2009): 15–18.

7. One of the more obvious examples of rekindling the public's interest in stereoscopic 3D movies was that the 2009 Cannes Film Festival showed as opening film the animated story *Up* by Pixar. There are many consumer products on the market now that encourage the production of stereoscopic video.

8. D. J. Roller, "World's First 4K 3D Cinema Camera System for Studio, Location, and Underwater Filming with RED camera," PRNewswire (March 3, 2009). Available at *http://news.prnewswire.com/DisplayReleaseContent.aspx?ACCT=104&STORY=/www/story/03-03-2009/0004982328&EDATE=.*

9. Daniel Engber, "The Problem with 3-D," *Slate* (April 2, 2009). Available at *www.slate.com/id/2215265.* See also Bryant Frazer, "How Big Will 3D Be?" *Studio Daily* (April 23, 2009). Available at *www.studiodaily.com/blog/?p=1371.*

10

Structuring the Three-dimensional Field: Screen Volume and Effects

So far we have discussed some of the basic principles and factors that help you project the three-dimensional world onto the two-dimensional video, film, and computer screens and create the illusion of depth. We now examine how to structure the three-dimensional field. We do this principally by manipulating lens-generated space, building computer-generated space, and a combination of both.

Regardless of whether you capture part of the real world through the camera lens or construct an image with 3D computer software, you need to provide distinct depth planes along the z-axis.

The most basic structure of the three-dimensional field consists of a ***foreground*** (the depth plane closest to the camera, marking the beginning of the z-axis), a ***middleground*** (the depth plane marking the approximate middle of the z-axis), and a ***background*** (the depth plane farthest from the camera, marking the end of the z-axis). SEE 10.1

When structuring the three-dimensional field, you must—as with most other aesthetic fields—take into account the element of change, that is, the movement of the event itself, of the camera, and of the sequence of shots. A camera that dollies past a row of columns; people dancing; cars moving along the z-axis; a zoom; or a cut from one camera to another—all create a changing structural pattern, a changing three-dimensional field. For example, when the camera is zooming in, the foreground may disappear, the middleground takes on the role of the foreground, and the background becomes the new middleground. Or when you cut from one camera angle to another view, viewers essentially see a new z-axis with its own spatial articulation.

To structure the lens- or computer-generated 3D field, you need to consider three important aspects: volume duality, z-axis articulation, and z-axis blocking. Special effects operate strictly in a virtual volume; that is, they are purposely displayed not as a simulation of real space but as synthetic, constructed ***screen space***. Realize that in ***stereoscopic projection***, the foreground plane extends toward the viewer and is no longer perceived as part of screen space.

10.1 Depth Planes
This picture is clearly divided into a foreground plane (the tree trunks and the camera-near ground), a middleground plane (the creek bank and the bushes), and a background plane (the mountain). This threefold breakdown represents the most basic structure of the three-dimensional field.

Although the most basic articulation of the z-axis is its division into a foreground plane, a middleground plane, and a background plane, you can divide the z-axis into as many planes as you wish, but such divisions must fulfill some aesthetic purpose; otherwise the z-axis space simply looks cluttered. **Josef von Sternberg** (1894–1969), who directed Marlene Dietrich in the famous 1930 film *The Blue Angel,* was obsessed with articulating the z-axis space in his films with appropriate space modulators and meaningful action to keep “dead” z-axis planes to a minimum. Look for z-axis articulations in the latest dramatic television series.

Volume Duality

When you look at the cloudless sky, you do not experience any depth. Empty space gives no clues as to distance because nothing is either near to or far from you. An extreme long shot looks exactly the same as when you are zoomed all the way in. As soon as you look ahead of you and see the houses, trees, and people, however, you can readily tell which things are near you and which are farther away. The objects help define the space around you and make it possible to perceive the third dimension—depth. We can say that positive volumes (houses, trees, and people) have articulated empty space with a series of negative volumes.

A ***positive volume*** has substance; it can be touched and has a clearly described mass. Positive volumes include such objects as cars, pillars, desks, and chairs as well as people. But a positive volume is also any screen image that has the appearance of substance.[1]

A ***negative volume*** is empty space that is somehow delineated by positive volumes. Unlimited negative space, such as the cloudless sky, constitutes negative space but not negative volume. The interior of a room is a negative volume because it is clearly described by the positive volumes of the walls, ceiling, and floor. The hole in a doughnut is also a negative volume, but the space surrounding the doughnut is not (assuming that we ignore the larger negative volume of the room).

The interplay between positive and negative volumes is called ***volume duality***. The control of volume duality—how the positive volumes articulate the negative space—is an essential factor in the manipulation of three-dimensional space and the illusion of screen depth. Designing scenery for video or film and blocking talent and camera action—all are careful and deliberate manipulations of volume duality. For example, the empty studio represents a clearly defined (by the studio walls) yet unarticulated negative volume—the empty studio. As soon as you put things into the studio, such as scenery and set pieces, you begin to modulate the negative volume of the empty studio with positive volumes, giving each scenic element its specific place and dividing the large negative volume into smaller, organically related negative volumes. Thus the positive volumes act as space modulators.

10.2 Creating Volume Duality
The negative volume of the studio space is articulated by the positive volumes of the set pieces, which remain relatively stationary, and the dancers who move around and thereby change the volume duality. The dynamic changing relationships between negative and positive volume is called volume duality.

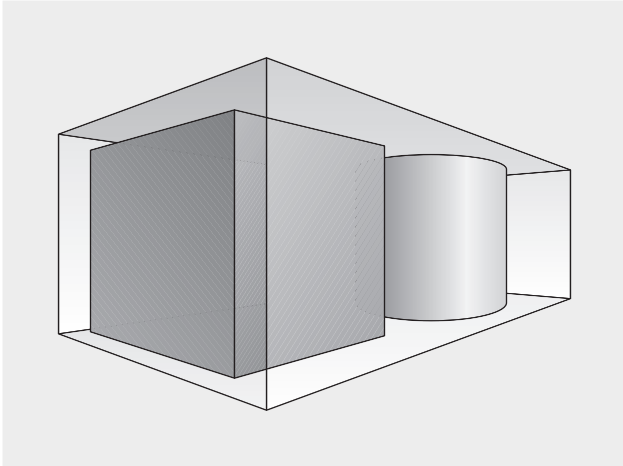
Volume duality can vary in degree. When people or objects move through the negative volume (such as the studio, for example), they continuously redefine the volume duality. SEE 10.2

DOMINANT POSITIVE VOLUME

When you stuff many objects into the studio or any other room, the positive volumes outweigh the negative ones, making the room look crowded. Too much positive volume can confine movement and make you feel restricted or boxed in.

In this medieval chamber, positive volumes overpower the interior's cramped negative volume. Thick walls and solid furniture give the room a heavy, restricted look. One feels boxed in, if not imprisoned.

Taken from *The Complete Woodcuts of Albrecht Dürer,* edited by Dr. Willi Kurth, republished in 1963 by Dover.



10.3 Preponderant Positive Volume

Volume duality can vary in degree. When using many and/or large space modulators, the positive volume can outweigh the negative volume.

A medieval castle is a good example of a preponderance of positive volume with relatively little negative volume. **SEE 10.3**

Like modern bunkers, the heavy positive volumes protect the restricted living space inside but also inhibit freedom. It's no wonder that medieval knights broke out of this spatial confinement sporadically to beat each other up for no apparent reason! Some modern concrete structures reflect the same confined negative space even if their function is to provide as much negative volume as possible, like a parking garage. **SEE 10.4**

If you want to re-create for the camera this sense of confined space, you need to crowd the negative volume with positive volumes—that is, stuff things into a relatively small space. If, for example, you want to show that a crowded office is difficult to work in, confine the action area—the area in which people can move

10.4 Restricted Operating Space

The parking garage on the left of this picture has so much positive volume that it looks cramped even if there is ample space inside.

about—by placing lots of desks and file cabinets in close proximity. By using a narrow-angle (long-focal-length) lens and chiaroscuro lighting, you can further intensify such a dominance of positive volumes. If you want to emphasize the potentially explosive situation of a mass political protest, wedge the protesters into a narrow street or intersection rather than show them marching through the wide-open space of the city park. In a similar way, you can intensify the sheer bulk and power of a big person by having him sit on a small chair in a tiny, low-ceilinged room.[2]

DOMINANT NEGATIVE VOLUME

A large, well-articulated negative volume invites mobility. We feel less restricted and can breathe freely and move about easily. **SEE 10.5**

A dominant negative volume isolates us and can make us feel insignificant and humble in the presence of so much emptiness. Much like the huge positive volume of a mountain, a large negative volume, such as an open plaza, an empty stadium, or the interior of a Gothic cathedral, can fill us with awe.[3] **SEE 10.6**

Too much negative volume, however, can promote a certain emptiness wherein we feel alone, cold, isolated, and lost.[4] **SEE 10.7** It's no wonder that people who work in the large, unarticulated space of modern offices put up screens

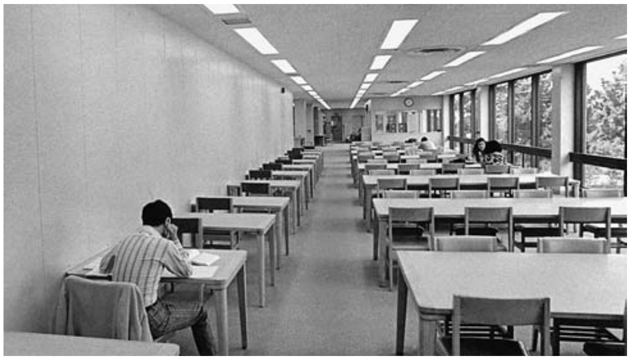
Gothic cathedrals were deliberately built with a maximum of negative volume. The positive space modulators, such as the walls and the pillars, were kept to a bare structural minimum to emphasize the wide expanse of the church interior. Such a vast, vertically oriented interior space contributes greatly to making us feel both reverent and subservient.

10.5 Dominant Negative Volume

With only a minimum of space modulators, the negative volume remains large though highly articulated. This interior has an open, spacious feeling.

10.6 Feeling Vastness

This plaza has a preponderance of negative volume, which is nevertheless articulated by the people and the monument in the center. Such vast negative volumes can inspire awe. At the very least, we feel dwarfed by the space.



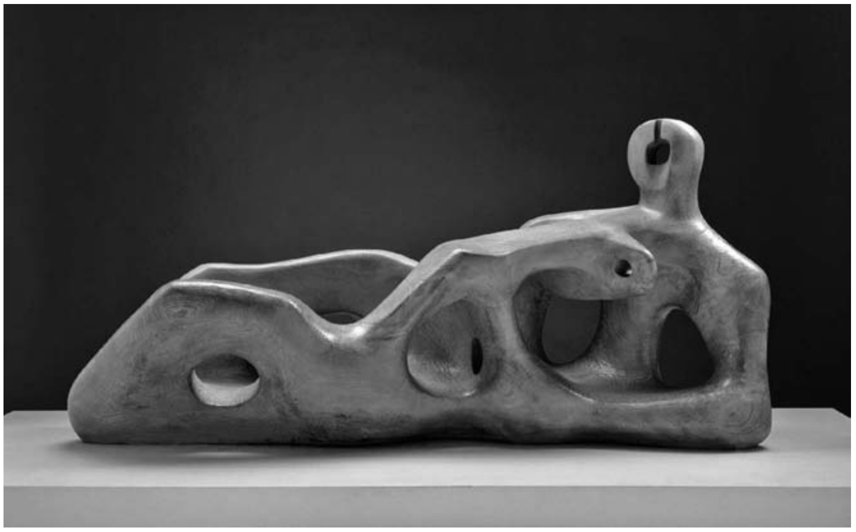
10.7 Lost in Negative Volume

An interior with too much negative space can make us feel isolated and uncomfortable. This library room is anything but cozy.

and partitions or use space modulators such as file cabinets to create a less public and more personal space for themselves.

APPLICATIONS OF VOLUME DUALITY

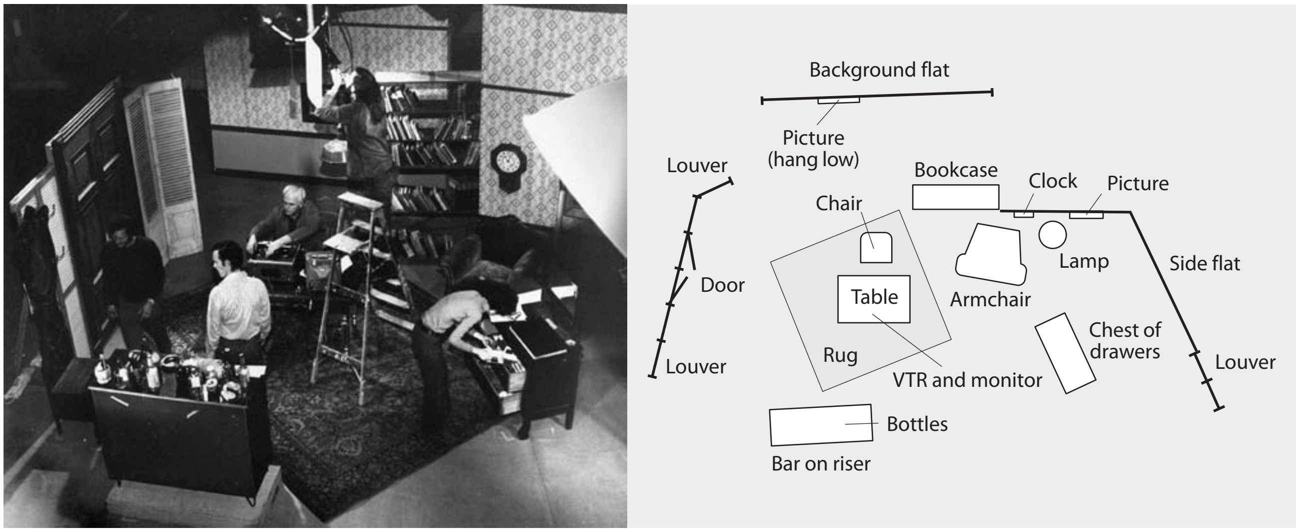
You can see the most obvious application of volume duality for its own sake in sculptures done primarily to explore the interrelationship of positive and negative volumes. SEE 10.8



10.8 Applied Volume Duality in Sculpture

This sculpture by the late English artist Henry Moore shows a masterful handling of volume duality. The holes in the sculpture are so well placed and well defined by the positive volumes that they attain positive spatial characteristics. As in a figure/ground reversal, the holes no longer represent nothingness—an absence of something—but rather define a variety of negative volumes that counteract and supplement the positive volume.

Henry Spencer Moore (1898–1986) *Reclining Figure* (1939), elm. The Detroit Institute of Arts, USA/Founders Society purchase with funds from Mr. and Mrs. Charles Theron Van Dusen and Beatrice W. Rogers Fund/ The Bridgeman Art Library/Reproduced by permission of The Henry Moore Foundation.



10.9 Open Set

In an open set, we modulate rather than merely describe the negative volume of the studio by the positive volumes of scenery and set properties. Modulating the negative volume means transposing it into an important scenic element. The spaces among the flats are no longer empty holes but rather appear as solid walls much as the positive flats themselves do. Proper lighting is essential to transpose negative space into positive scenic elements.

In scene design volume duality is applied in an open set. ***Open set*** refers not to a set that is open to the public but rather to scenery that is not continuous; the open set is not closed or boxed in by connected walls but instead consists of only the most important parts of a room, for example, a window, some furniture, and a few separate single flats. **SEE 10.9**

The open-set method is particularly effective for a single-camera video production that builds its screen events inductively, bit by bit, close-up by close-up, in a mosaic fashion. So long as your visualization approach is inductive, the open set is even advantageous in multicamera productions; it allows for optimal camera points of view and a more fluid shot sequencing. This is different when shooting "landscape-style" for digital movies or large-screen video. The frequent long shots of large vistas or interiors require a continuous interior and make an open set impractical.

Z-axis Articulation

The z-axis is especially important in structuring video space because the other principal spatial dimensions of the standard video screen—height and width—are limited compared with the larger and especially much wider HDTV or motion picture screen. The camera, very much like the human eye, has no trouble looking along the z-axis all the way to the horizon. It can therefore take in a great number of objects stationed along the z-axis. It can cope successfully with even extremely fast movement without panning (horizontal camera movement) or tilting (camera looking up and down) so long as the objects are placed or move along the z-axis. Vertical and horizontal object motion require a great deal of camera panning and tilting, and this becomes especially difficult for the camera if the object motion is rapid. The articulation of the z-axis is therefore one of the principal factors of spatial and motion control in standard video.

10.10 Wide-angle View of Traffic
This z-axis shot, taken with a wide-angle lens, shows a preponderance of negative volume. The z-axis seems elongated, and traffic appears relatively light.

10.11 Normal View of Traffic
Taken with a normal lens (medium zoom lens position), the z-axis now looks slightly shorter than in figure 10.10, and traffic appears a little more congested.

10.12 Narrow-angle View of Traffic
With a narrow-angle lens, the cars look closer together, making the z-axis look compressed.

Articulating the z-axis means to place positive volumes along the z-axis to help the camera distinguish among the depth planes. By placing objects or people at various z-axis locations and by choosing a specific lens (wide-angle, normal, or narrow-angle zoom lens positions), you can make the viewer perceive restricted or open space, with objects being crowded or else comfortably or agonizingly far apart.

Take a look at the three traffic situations in the figures above. **SEE 10.10–10.12** Figure 10.10 shows a long stretch of road with apparently light traffic. In figure 10.11 the traffic seems slightly heavier; in figure 10.12 the traffic appears completely jammed. Actually, all three shots were taken from the same position within seconds of one another, but each was taken with a different lens, that is, with a different zoom lens position. You probably recognized that the shot in figure 10.10 was taken with a wide-angle lens. The resulting volume duality shows preponderant negative volume. The exaggerated convergence of parallel lines and relative size make the cars appear much farther apart than they actually are. The aesthetic effect is that we perceive traffic to be light. The shot in figure 10.11 was taken with a normal lens. The perspective, relative-size factors, and volume duality in this shot appear approximately as you would normally see them. This means that traffic conditions are reflected accurately in this shot. The shot in figure 10.12 was taken with a narrow-angle (telephoto) lens. Here the volume duality shifts to a predominant positive volume. The linear perspective and the relative-size differences are minimized, and the parallel lines do not converge as rapidly as in the other two shots. The cars are reduced in size much less toward the background than in the wide-angle shot. Because they appear similar in size, the z-axis space seems to have shrunk, and the cars appear much more crowded than they actually are; traffic seems heavy.

NARROW-ANGLE LENS DISTORTION

The crowding of objects through a narrow-angle lens can cause a variety of perceptions. **SEE 10.13** By reducing the negative space to a minimum, the signs in figure 10.13 seem right on top of one another. The signs have lost their effectiveness. Instead of being individual carriers of specific information, they have become mass elements of visual pollution.

10.13 Sign Pollution

The narrow-angle (long-focal-length) lens not only crowds these signs but also emphasizes the visual pollution.

When spatially condensed by the narrow-angle lens, a row of separate columns becomes a single, massive support. With the reduction of negative space, the volume duality has given way to a mass of positive volume. **SEE 10.14** Similarly, when squeezed by a narrow-angle lens, a row of houses connotes certain psychological and social conditions, such as closeness or crowdedness with little room for expansion. Depending on the overall context, you might also think of fire danger. **SEE 10.15**

Like a military uniform, the narrow-angle lens can rob people of their individuality and suggest sameness of behavior, collective goals, a high degree

10.14 Massive Columns

When shot with a narrow-angle lens, the columns lose their individuality. We perceive only positive volumes.

10.15 Cramped Housing

The narrow-angle lens crowds these homes even more than they really are, intensifying their close proximity and suggesting cramped living conditions or fire danger.

10.16 Tailgating
By reducing the space between the cars to a minimum, the narrow-angle lens intensifies the danger of tailgating.

10.17 Rush-hour Traffic
The crowding of the vehicles by the narrow-angle lens intensifies the density of rush-hour traffic.

10.18 Undesirable Distortion
Even if we know the actual distance between pitcher and batter (60 feet 6 inches), the long-focal-length lens (extreme narrow-angle zoom lens position) shows the players as standing fairly close together.

of persuadability, or simply raw, irrational power. Extreme closeness can suggest danger, especially when applied to fast-moving vehicles. The precariousness of tailgating is highly intensified by the narrow-angle field of view. **SEE 10.16**

Shooting heavy traffic along the z-axis with an extremely long-focal-length lens (zoomed in all the way to its narrowest-angle position) crowds the vehicles even more than they really are. Such a shot readily communicates the frustration of the people stuck in rush-hour traffic. **SEE 10.17**

As with every aesthetic effect, the depth distortion through a narrow-angle lens can, of course, also work to your disadvantage. You are no doubt familiar with the deceiving proximity of the pitcher to home plate. Because cameras must remain at a considerable distance from the action, the extreme narrow-angle lens used in the reverse-angle shot from pitcher to batter drastically shrinks the apparent distance between the two. In such a shot, we may wonder how the batter could ever hit the ball when the pitcher fires at him from such a close range. **SEE 10.18**

WIDE-ANGLE LENS DISTORTION

The wide-angle lens also exaggerates size relationships. An object close to the camera appears much larger than a similarly sized object placed just a short z-axis distance away (see figure 9.17 in the previous chapter). We automatically interpret this size difference as increased z-axis distance. Depending on the context, however, we can also interpret such a size discrepancy as an exaggeration of object size. Through what Sergei Eisenstein called "conflict of volumes and spatial conflict," such distortions carry not only aesthetic but also psychological messages.[5]

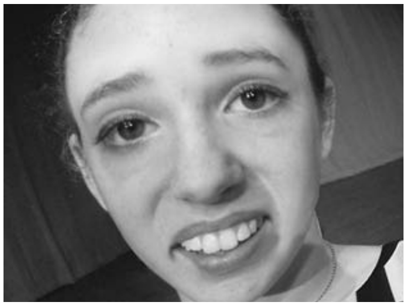
Ordinary shots can become highly dramatic through a wide-angle lens. **SEE 10.19 AND 10.20** For example, the

10.19 Implied Message: Importance
The distortion of relative size through the wide-angle (short-focal-length) lens implies that the box contains a precious gift.

10.20 Implied Message: Danger
The wide-angle-lens distortion of this cup implies that the liquid it contains might be dangerous to drink.

10.21 Intensified Message: Stop
The wide-angle-lens distortion of her hand intensifies the guard's stop command.

10.22 Implied Message: Power
The power of this truck is intensified by the wide-angle-lens distortion.



10.23 Implied Message: Emotional Stress
Although wide-angle-lens distortions of a face are usually avoided, you can nevertheless use them to intensify extreme emotional stress or unbalance in a person.

wide-angle-lens distortion of the little gift box underscores the importance of the occasion (figure 10.19). On the other hand, the distortion of the cup suggests that it may contain something other than the aromatic coffee that wakes you in the morning (figure 10.20). A gesture signaling a stop is made more forceful and authoritative when shot with a wide-angle lens. **SEE 10.21** The power of huge things, such as trucks, jet planes, or heavy machinery, is also aptly dramatized by the wide-angle shot. **SEE 10.22**

You can also use wide-angle-lens distortion to communicate intense emotional stress in a person. **SEE 10.23** Extreme facial distortions as in figure 10.23 suggest that the woman may be no longer stable or rational. Obviously, such a distortion must operate in concert with other contextual media aesthetic clues.

When you shoot through prominent foreground pieces, the foreground acts as a secondary frame (the primary frame being the video screen), which focuses our attention on the middleground or background objects. **SEE 10.24** Because of the volume conflict (large foreground objects on screen-left) and the relatively small background objects (people), this shot inevitably directs our attention to the main subject—the two boys sitting in chairs. Such implied secondary frames are especially important when composing shots for the wider HDTV or movie screens.

10.24 Secondary Frame
You can focus attention on a scene by shooting through a prominent foreground piece, creating a secondary frame.

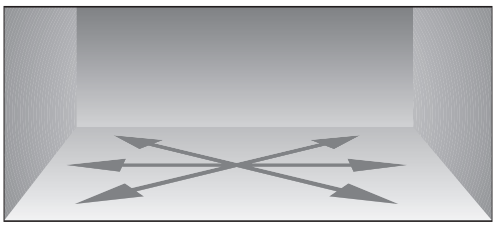
Z-axis Blocking

Z-axis blocking refers to placing people and their movements primarily along the z-axis—toward and away from the camera. Such blocking is one of the major devices for effectively articulating the z-axis, creating a dynamic volume duality and intensifying the illusion of a third dimension on the two-dimensional screen.

When blocking action for the theater stage, we usually rely heavily on lateral or diagonal rather than upstage/downstage (z-axis) motion. In fact, lateral action is generally preferred in theater because the stage is usually wider than it is deep. **SEE 10.25**

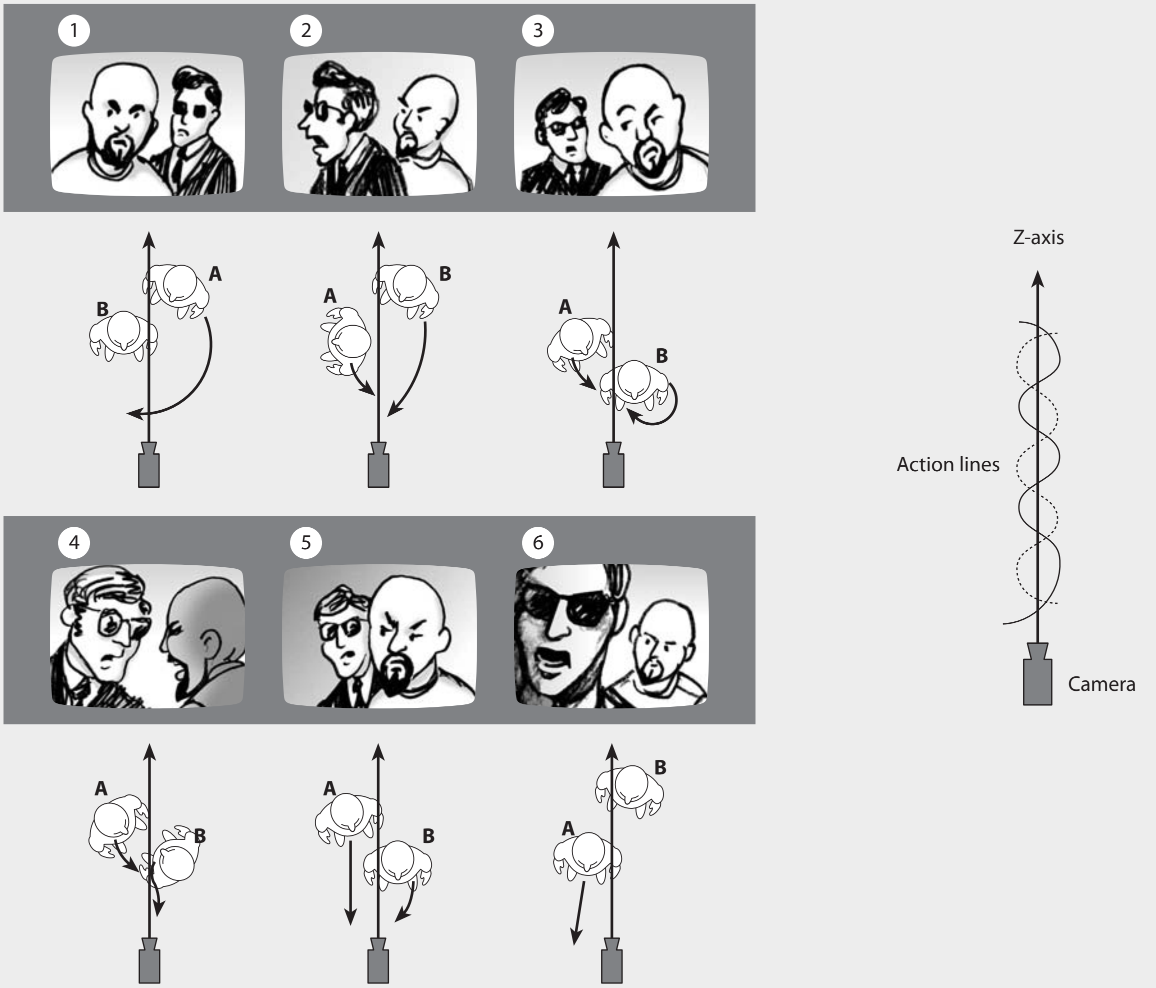
The blocking for the wider HDTV screen can be more lateral so long as the video is projected onto a large screen. If a projection is limited to a relatively small HDTV screen, however, you should still block primarily along the z-axis, much as you would for the standard 4 × 3 video screen. The small screen cannot tolerate much lateral action without having the camera pan or truck along with it. Aside from the technical problems of keeping a fast-moving object properly framed, too much lateral action can become distracting and disorienting.

Proper blocking along the z-axis minimizes camera movement and emphasizes object (and people) motion, not camera motion. In fact, high-impact action is often blocked along the z-axis, even in large-screen movies. A person running or an object hurtling along the z-axis toward the camera is usually more dramatic and carries more aesthetic energy than when staged sideways along the x-axis.



10.25 Lateral Action in Stage Blocking
The relatively large lateral space of the theater stage and the wide motion picture screen make lateral action an effective blocking technique in theater and film.

Scene: Two people (A and B) arguing; person A finally gives up and exits past the camera.



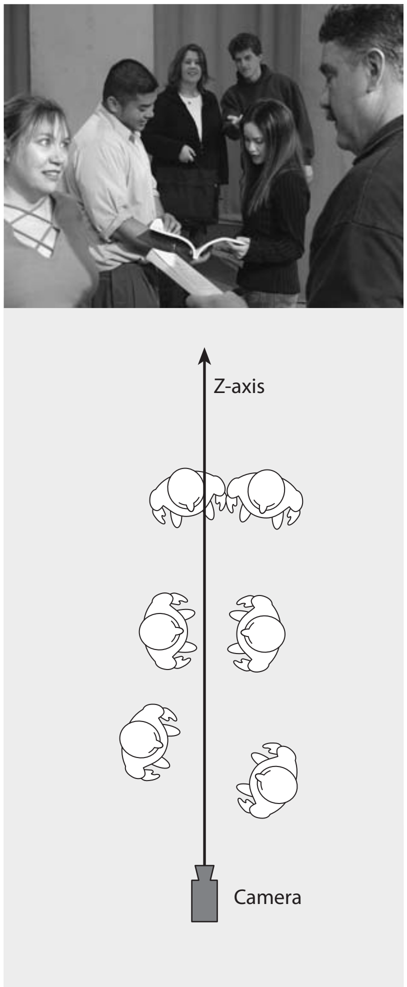
10.26 Z-axis Blocking

Because of the limited height and width of the standard video screen or mobile media display, action is most appropriately staged and blocked along the z-axis.

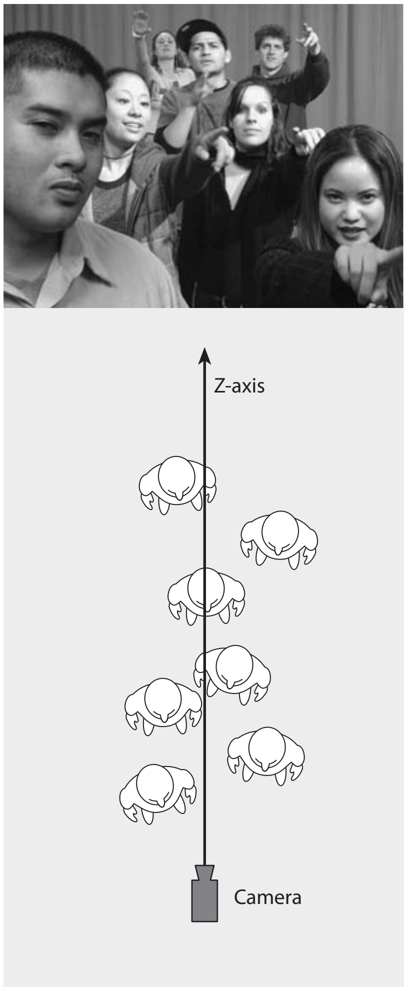
When blocking for television, you can have the action weave toward and/or away from the camera along the z-axis. **SEE 10.26**

Such z-axis blocking for the standard video screen allows you to show a relatively large number of people interacting with one another from a single camera's point of view. **SEE 10.27** For example, you can use z-axis blocking to simulate a crowd scene with just a few people. **SEE 10.28** The same scene shot from another angle reveals how sparse the "crowd" really is. **SEE 10.29**

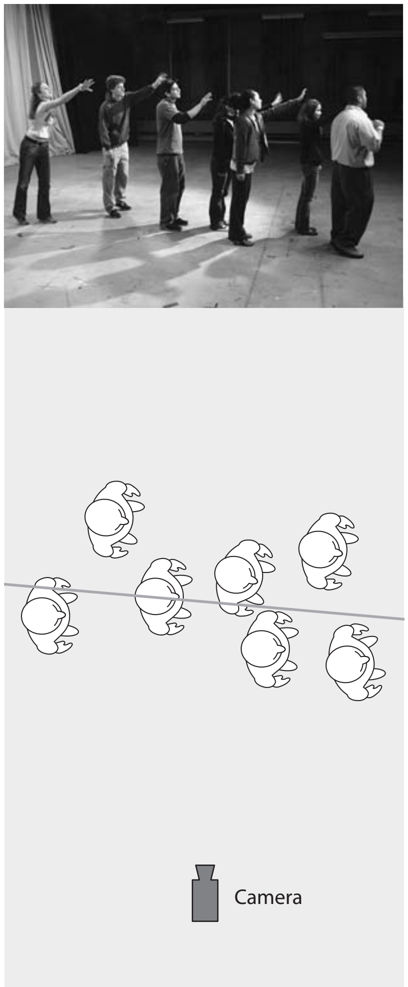
If, in a multicamera production, you intend to use a second camera for close-ups in such a "crowd" scene, you would need to have the second camera fairly close to the first so that the z-axes of both cameras are practically identical or at least run pretty much parallel to each other. Otherwise the first camera would see the crowd along the z-axis, and the close-up camera would view them along the x-axis (see figures 10.28 and 10.29).



10.27 Z-axis Blocking of Several People
A relatively large number of people can be included in a single shot so long as they are blocked along the z-axis.



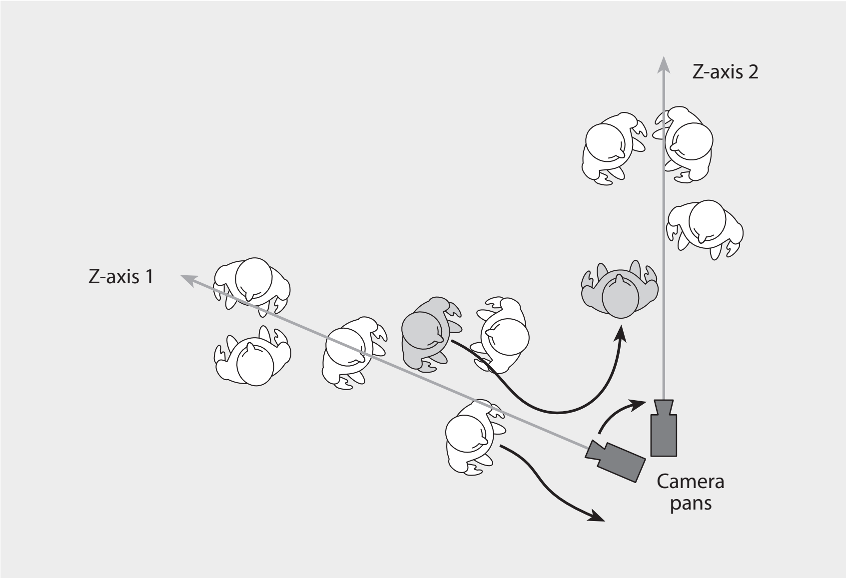
10.28 Z-axis Blocking of Crowd
A crowd scene is easily simulated in video by blocking a few people along the z-axis. A wide-angle lens is used to stretch the z-axis.



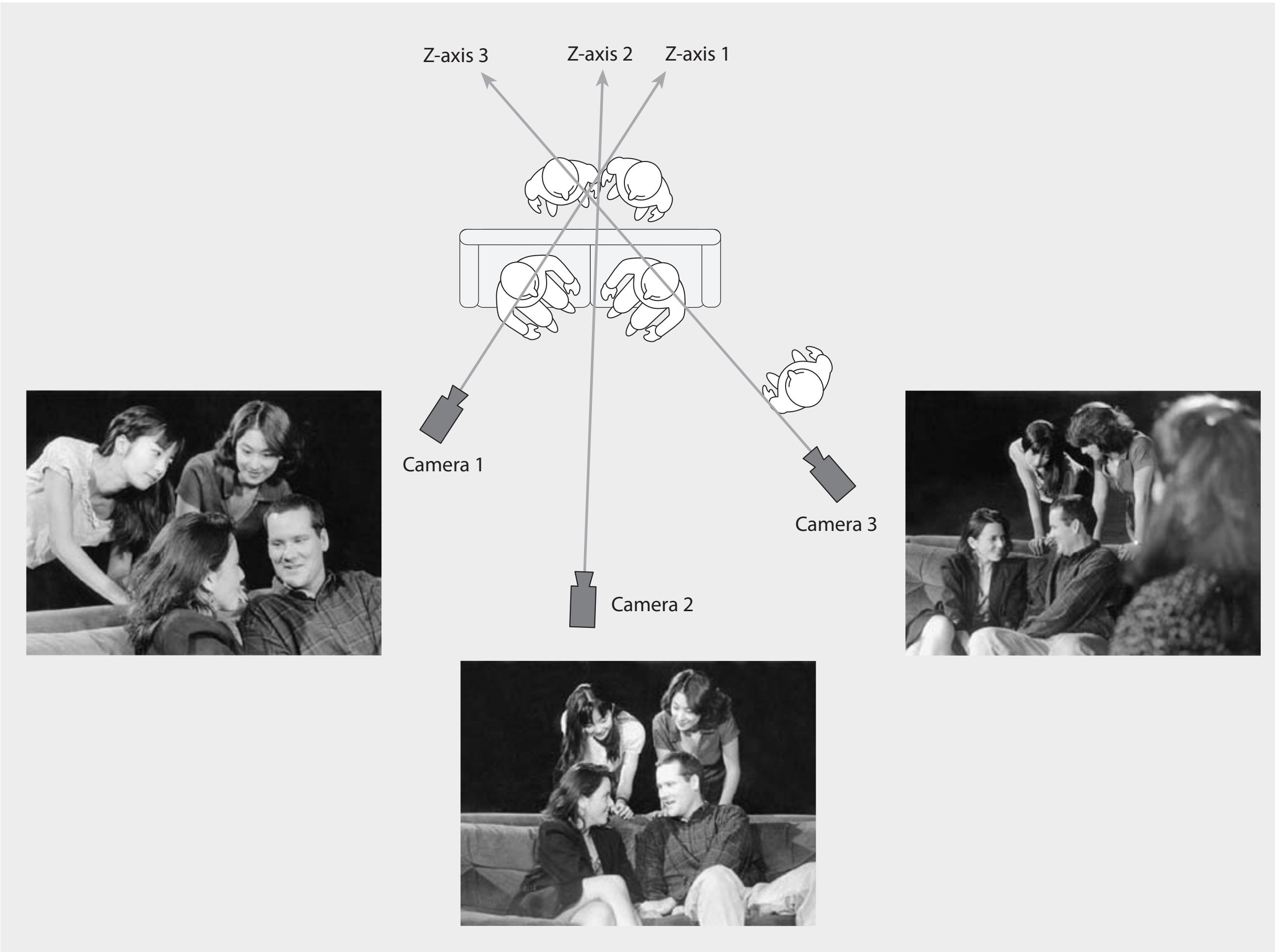
10.29 Crowd Scene Shot Along the X-axis
When seen from another angle (along the x-axis), these few people certainly don't simulate a crowd. This is one of the reasons why we use so many extras in crowd scenes for the horizontally stretched motion picture screen.

Good blocking for the small screen means staging an event along the z-axis for each camera; this is called ***multiple z-axis blocking***. For example, you can have the camera look down a hallway, such as a hospital corridor, and block the action from the back of the hallway toward the camera (z-axis 1). Have most of the people exit by walking past the camera. Now have the camera pan left and follow someone (perhaps the chief surgeon) walking down another articulated z-axis (z-axis 2). **SEE 10.30** You probably recognize this technique as being used in many serial dramas.

In a typical situation comedy setup with the inevitable couch, you can easily block the action along three articulated z-axes—one for each camera. If done properly, such blocking requires little or no camera movement. **SEE 10.31** Such blocking is desirable regardless of whether the production calls for a single camera or multiple cameras.



10.30 Double Z-axis
You can block action so that the camera can pan from one articulated z-axis to another in a single shot.



10.31 Z-axis Blocking for Three Cameras
In this setup each camera has its own properly articulated z-axis.

Special Effects and Spatial Paradoxes

The ready availability of ***digital video effects (DVE)*** software may tempt you to use such effects just to liven things up a bit—to interject some motion and excitement into an otherwise dull and slow-moving show.

You can see such digital wizardry every time you turn on the television, regardless of program content. These effects are most prominent in news presentations. Brief videos of events from around the world are sometimes frozen into still images, making them look like still shots in a photo album or a picture that hangs over the news anchor's shoulder. Sometimes various background effects reinforce the story content. When the news item is about war, tanks in low-saturated colors roll through the background; if it is about an election issue, the inevitable flag will undulate in slow motion behind the candidate.

Of course, we all know that even the most inventive effects will not make a boring story interesting. These effects may also unintentionally or intentionally communicate fairly powerful metamessages—messages that carry latent meanings—that often go unnoticed by less attentive viewers. In fact, such seemingly harmless messages are sucked up by your unconscious without your ever knowing and may well prime your reaction in a predictable way.[6] Knowing some of the aesthetic codes and subcodes that underlie such effects can enable you to identify their semantic significance and both use them to enhance the intended communication objective and avoid their irresponsible use.

This section examines five such effects: graphication, first- and second-order space, personification, topological and structural changes, and spatial paradoxes.

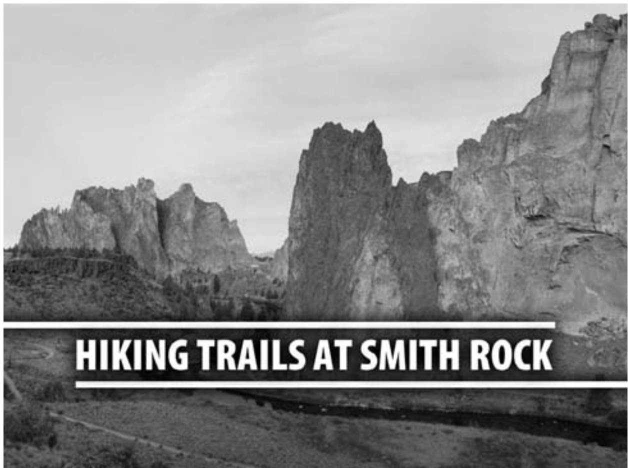
GRAPHICATION

No sooner have we learned how to create the illusion of depth on the two-dimensional plane of the video screen than we are confronted with digital techniques used to render a three-dimensional scene deliberately 2D and graphiclike again. Such a process, in which the three-dimensional lens-generated screen image is deliberately rendered in a two-dimensional, graphic- or picturelike format, is called ***graphication***.[7] Graphication can take on many forms, of which the most common are lines and lettering, secondary frames within the video screen, and a combination of the two.

Lines and lettering One of the simplest graphication devices is lines and lettering used for titles or keyed over the actual full-screen image. Like we perceive the text on this page, we readily perceive the on-screen lines and lettering keyed over the scene as the figure, and the images behind the letters (the actual scene) as the ground. This figure/ground relationship persists whether the ground is a static or a moving three-dimensional scene. The lines and lettering prevent the scene from pushing itself into the foreground and remind us that the scene, moving or not, is merely a picture. SEE 10.32

People are not immune to graphication through keyed letters and lines. Regardless of whether people on-screen appear in a studio or a field setting, the overlapping graphics inevitably remove them from the foreground position and relegate them to a lesser space farther back along the z-axis. Much like the velvet ropes in banks, theaters, and hotels that keep us in a queue, the on-screen lines and lettering prevent people on-screen from assuming foreground prominence. SEE 10.33

Secondary frames One of the more popular ways of graphicating a scene is to put it into a secondary frame, such as windowboxing. The digitally generated



10.32 Graphication Through Lines and Lettering
Lines and lettering superimposed over a video scene become the figure, relegating the actual scene to a ground position. The scene has become picturelike, very much like a photo in a magazine. We call this process graphication.

10.33 Graphication of People
Lines and lettering can also have a graphication effect on people, regardless of whether they are moving.

10.34 Graphication Through Secondary Frame
A common windowbox graphication device is the secondary frame within the primary video screen. It contains a scene in much the same way as a picture that hangs on the wall.

box clearly delineates an additional picture area within the borders of the actual primary space of the video screen. This technique renders the event displayed in the secondary frame as a picture similar to the pictures in newspapers and magazines. In fact, we often make the secondary frame look like a picture by giving it a border and setting it off from the background of the primary screen with drop shadows. **SEE 10.34**

FIRST- AND SECOND-ORDER SPACE

When a secondary frame is placed over the newscaster's shoulder, there are two types of screen space: the primary space as defined by the video screen's borders, and the space of the digitally created secondary frame. We call the total screen area ***first-order space*** and the secondary frame ***second-order space***. The anchor

10.35 First- and Second-order Space
The newscaster occupies nongraphicated first-order space. The field reporter appears in a graphicated secondary frame, the so-called box. This box represents second-order space.

10.36 Aesthetic Personification Factor
Persons occupying nongraphicated first-order space experience a higher degree of personification from viewers than people appearing in graphicated second-order space. The host operates in first-order space; his guest appears in the graphicated second-order space.

occupies first-order space; the person or event in the box is in second-order space.[8] **SEE 10.35**

Because the second-order frame is such a strong graphication device, we tend to perceive its content as picturelike, regardless of whether the pictures are static iconic motifs or lens-generated events that move. In contrast to this abstracted space, first-order space seems to be more "real."[9] A strong possibility exists that, under certain circumstances, we may even extend first-order space into our own living space and share our environment. This extension is quite different from a stereoscopic projection because it is a pure psychological phenomenon. At the very least, events in first-order space seem to attain a certain degree of verisimilitude and believability. An example of an extension of first-order space into our living space is when a newscaster, operating in nongraphicated first-order space, first talks to us—the viewers—introducing the guest who is confined in a second-order box. When the anchor then turns toward the secondary frame for the actual interview, we may perceive, at least temporarily, the newscaster sharing our actual environment, interviewing the guest from our—the viewers'—position. **SEE 10.36**

PERSONIFICATION

Such close personal contact, however much imagined, fosters familiarity and trust.[10] Subconsciously, we attribute to the people operating in the extended space the flesh-and-blood qualities of real people—a certain degree of ***personification***.

This personification effect does not seem to take place in second-order space. The abstraction through graphication is so great that viewers inevitably consider the people appearing in second-order space as video images or pictures. Even when the second-order people are occasionally and temporarily "let out of the box" to occupy the full primary screen, we still consider them as occupying second-order space so long as the full-screen display is brief. To make sure that the people appearing in second-order space do not invade first-order space, additional graphication devices in the second-order space, such as name, place, and time superimpositions, are generally used.

10.37 Host and Guest in Second-order Space
When both the anchor (screen-left) and the guest (screen-right) are confined to secondary frames, the personification factor for both parties is eliminated.

But don't we often see the anchor and the field reporter or guest appear in identical side-by-side boxes? Aren't they now both graphicated and operating in second-order space? The answer to both questions is yes. But, then, aren't we making a mistake by graphicating the anchor in this way? From a personification point of view, yes, we are. With both host and guest confined to second-order space, we inevitably perceive them to be picturelike and equally removed from us. For all practical purposes, the personification effect is lost. **SEE 10.37**

As you recall, this is also a questionable practice from an aesthetic point of view. Despite the context of a two-way conversation, the continuing z-axis index vectors are not necessarily perceived as converging; they may well be seen as continuing, being directed at us, the viewers.

From a production point of view, however, it is a convenient way to show two people talking to each other without the need for frequent cutting between the two close-ups. Such a side-by-side arrangement is especially convenient when the anchor is conversing with four or five guests simultaneously.

In any case, a better solution would be to have the host remain in first-order space, looking with the viewers, at the guest framed in second-order space (see figure 10.36). You can then switch to the secondary frame (or frames, if there are several guests) and have the host ask questions from an off-screen position. The z-axis index vectors of the guests are now directed toward the off-screen host and, at the same time, toward us, the viewers. We are then able to identify with the host, who is temporarily operating from an extended first-order space.

TOPOLOGICAL AND STRUCTURAL CHANGES

Digital video effects enable you to shrink the full-screen video and change its topology—its outer shape—at will. For example, you may see an especially tragic event briefly appear on-screen, freeze, peel off a large stack of snapshots, and flip unceremoniously through video space to make room for a new series of disasters. **SEE 10.38**

10.38 Flips

One of the many digital video effects is a topological change—a manipulation of the outer shape of the video image. In this case, the video image is made to look like a photo that flips through first-order space. Such special effects can carry negative metamessages.

At first glance such effects are relatively harmless attention-getters. On closer examination, however, we discover that such visual acrobatics carry metamessages that, although not consciously perceived by most viewers, can readily change their attitude toward the event itself. The freezing and careless flips of disturbing scenes may convey an I-don't-really-care attitude toward the calamities displayed, making it all too easy for us to remain emotionally uninvolved. What we perceive, however subconsciously, is that we are no longer watching a slice of reality, a documentation of an actual occurrence, but merely some easily deposable pictures. Although we do indeed see only fleeting pictures when watching television, such topological changes to the regular video image become more anomalous the more we believe that the television-mediated event is the real thing. The combination of negative graphication and the often simultaneously displayed irrelevant information, such as sports scores and the weather, provide an easy escape from getting emotionally involved—or even from social responsibility to do something about the displayed human suffering.

The large palette of possible DVE also invites you to change the structure of the image. The mosaic effect as well as images with heightened contrast and color distortions are but two well-known examples. SEE 10.39 They may be less weighty than the graphication flips, but they nevertheless transform what we normally perceive as a three-dimensional "real" video scene into picturelike abstractions. Most often they simply provide visual interest but can, if appropriate, also intensify the message. Before using any of these effects, however, think carefully about why you would choose a specific effect and what possible metamessages it may carry.

10.39 Structural Change: Mosaic

We can also graphicate an image through a structural change. One popular DVE technique of this sort is the mosaic effect, in which an image is made to break up and look as though it were composed of a series of mosaic tiles.

SPATIAL PARADOXES

Some special effects not only alter the topology or structure of a particular image but also combine pictures to create

10.40 Spatial Paradox: Superimposition
The super works against our hardwired organization of a figure that lies in front of a background. Its overlapping transparent images confuse the figure/ground principle and suggest event complexity.

spatial paradoxes—arrangements that go against our perceptual habits. Some of these effects help structure and communicate complex messages; others violate our perceptual principles and work against effective communication.

Superimposition As mentioned in chapter 7, a superimposition leads to an image in which the usual figure/ground relationship and the overlapping planes are largely dissolved into a complex array of intersecting images. By "supering" one image over another, the objects seem to become transparent, eliminating the illusion of depth and volume. The collapsed separate viewpoints or events into a single two-dimensional picture plane change the viewer's normal perceptual expectations and give not only a more complex view of things but particularly deeper insight into the event's underlying complexity. Because of this new structural bond, a *super* can suggest a strong relationship between seemingly unrelated events. Thus we often use the super to create a surrealistic or dreamlike feeling. **SEE 10.40**

Figure/ground paradox Never mind what the newscaster is saying—all we care about right now is what the newscast looks like or, specifically, how the three-dimensional field is structured in this particular shot. When you see a close-up of the anchor in the foreground and the keyed-in graphicated box over his shoulder, you automatically presume that the anchor is in the foreground, the box behind his shoulder is in the middleground, and parts of the newsroom or news set are in the background. **SEE 10.41**

At first glance we seem to have a classic structuring of the three-dimensional field, with a prominent foreground (the anchor), a somewhat ambiguous middleground (the secondary frame seems to float somewhere between the anchor and the back wall), and a common background (the back wall of the news set). But perceptual problems occur when the anchor shifts his position or the box key is slightly off. In either case, the box may temporarily overlap the anchor's shoulder.

10.41 Figure/Ground Organization in Newscast
In this shot the three-dimensional field is clearly organized into a foreground plane (the anchor), a middleground plane (the graphicated box), and a background plane (the back wall of the news set).

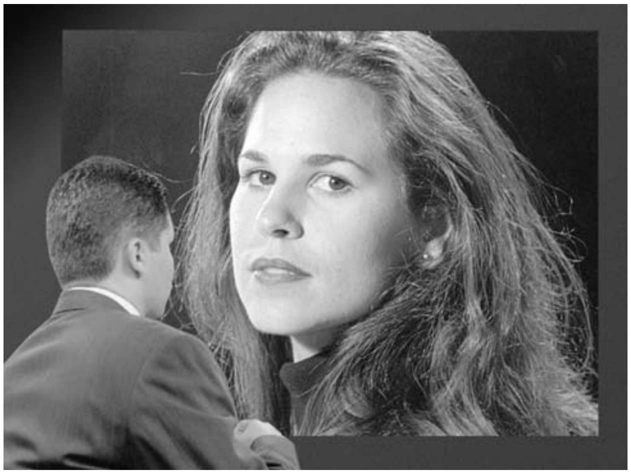
10.42 Figure/Ground Paradox
We experience a structural paradox when the middleground plane suddenly overlaps the foreground plane.

SEE 10.42 According to our normal perceptual expectations, we would expect the box (middleground) to be overlapped by the anchor (foreground). But this basic spatial organization is paradoxically upset: the middleground overlaps the foreground. Will we now perceive the box as being the foreground? Not really. Despite the strong overlapping-planes cue, which clearly puts the box in front of the news anchor, our mental map holds on to the basic organization and tells us that the anchor is still in the foreground.

This problem persists even if the overlap is intentional. In the attempt to make the second-order box a foreground piece, some graphic designers attach it to one of the edges of the first-order primary screen. At times the anchor is overlapped by the box and thus relegated to the middleground of the video space. SEE 10.43 Despite this deliberate figure/ground reversal, it is difficult for us to go against our hardwired organizational figure/ground principles. First, we still expect the larger image of the anchor to be the foreground figure and the smaller, graphicated second-order box to be somewhere behind the anchor. Second, such an arrangement also works against our psychological organization principle: because the anchor is definitely the primary information source, we expect her to occupy the foreground, and not the middle- or background. A similar problem arises when we reverse the image size between host and guest.

10.43 Intentional Overlap
Even if the second-order frame is purposely placed in the foreground plane, the occasional overlap still represents a spatial paradox.

Relative-size paradox Most likely you have seen interviews in which the studio host (usually one of the anchors) talks to a guest who appears from a remote location in an oversized second-order box on-screen. SEE 10.44 Considering that we perceive objects as smaller the farther away from us they are, you can probably spot the structural problem right away. According to the relative-size principle, the image of the guest, who appears in the background, should be smaller than that of the host, who is located in the foreground of the first-order space of the news set. This relative-size paradox makes the guest appear oversized if not overpowering. But why would you perceive the guest as too big instead of the host as too small?



10.44 Relative-size Paradox
In this reversal of the relative-size principle, the background figure is larger than the foreground figure. We perceive the background figure as exceptionally large. This arrangement diminishes the personification of the anchor.

First, as relatively similar-sized objects (in this case, heads), the one in the background (the guest's) should look smaller than the one in the foreground (the host's), even if it contains a bigger brain. Second, you have probably appointed the host rather than the guest as the size standard. Now you need to recall the reality aspect of first-order space compared with the graphicated second-order space. Because the host operates in the more "real" first-order space, you will use the host as the size standard rather than the guest, who appears in the picturelike, graphicated second-order space. Compared with the host, who has become your size standard, the guest in the background looks too large. The sheer graphic mass of the guest's image inevitably overpowers the host even if he is functioning in first-order space. Subconsciously, you are inclined to shift the host's authority and credibility to the guest.

STEREOSCOPIC PROJECTIONS

Even when the z-axis extends in both directions from the screen, the stereoscopic projection does not alter the basic principles of screen volume and their effects. You still need an articulated z-axis to emphasize depth, regardless of whether you look at it with one eye (single-lens 3D volume) or with both eyes (two-lens 3D volume). The difference is, of course, that in stereoscopic projection the z-axis extends from the screen to the horizon as well as through the screen toward the viewer.

SUMMARY

We basically structure the three-dimensional field into various planes, or grounds, indicating z-axis placement or depth: a foreground, a middleground, and a background. To establish such depth planes, we must articulate the z-axis with positive space modulators—people or objects that are placed along the z-axis.

Structuring the three-dimensional field includes creating and controlling volume duality, articulating the z-axis, z-axis blocking, and dealing with spatial paradoxes.

As soon as we put things into an empty studio, we are modulating the negative volume of the studio with positive volumes. Positive volumes define space. We create a volume duality—the interplay of positive and negative volumes. When positive volumes prevail, the space looks and feels crowded. Large, well-articulated negative volumes suggest mobility. They can inspire awe because of their vastness but also suggest isolation and loneliness.

Various zoom lens positions (narrow- and wide-angle) create depth planes that the viewer perceives in specific ways. Generally, narrow-angle (long-focal-

length) lenses squeeze space between objects; they shorten the z-axis and make objects look crowded. Wide-angle (short-focal-length) lenses exaggerate the relative size of objects close to the camera and those farther away, creating the impression of increased depth and a lengthened z-axis. Extreme wide-angle-lens distortions carry powerful psychological associations of intense feelings and actions.

Because of the limited width and height (x and y axes) of the standard video screen, action is best blocked along the depth dimension (z-axis). Such z-axis blocking increases the feeling of depth, is relatively easy for the camera to follow, and intensifies the drama and the aesthetic energy of the shots. It is also a powerful blocking device for large HDTV projections and motion picture screens.

Digital video effects (DVE) can create spatial paradoxes that, when used in a proper context, can convey specific meanings. Unintentional spatial paradoxes, however, such as those that violate figure/ground and relative-size principles, can disturb our mental map and upset our perception of the three-dimensional structure, thus reducing the impact of the message.

Special effects often carry powerful metamessages that need to be considered before the effects are used. Graphication is the deliberate rendering of television-mediated events, which appear three-dimensional on the video screen, as two-dimensional images that look like pictures in magazines or newspapers. The major graphication devices are lines and lettering and the creation of secondary frames within the primary screen, such as the box over the newscaster's shoulder.

Such secondary fames divide the screen into first-order space (the actual illusory three-dimensional screen space) and second-order space (the picturelike secondary frame). Personification means that we assign the person operating in nongraphicated first-order space more prominence and, to some extent, the flesh-and-blood qualities of real people compared with those appearing in the highly graphicated second-order space.

Topological changes are images whose shape is digitally manipulated. Structural changes affect the way images are built. Such manipulation may result intentionally or inadvertently in powerful metamessages and therefore must be used with great care.

NOTES

1. László Moholy-Nagy, *Vision in Motion* (Chicago: Paul Theobald, 1947, 1965). Moholy-Nagy discusses in detail the forms and the applications of volume duality, especially in sculpture. He distinguishes between actual positive volumes that have mass and are thinglike, and virtual positive volumes that are created by moving elements, such as a rotating flashlight. In a similar way, the screen images of a person or object can be considered a positive volume.

2. See Steven D. Katz, *Film Directing: Shot by Shot* (Studio City, Calif.: Michael Wiese Productions, 1991), pp. 239–40. Katz refers to the use of a dominating positive volume in the early scenes of Orson Welles's classic film *Citizen Kane:* "In most of these scenes the ceilings are low, and Welles, who actually was tall, seems to dominate his surroundings."

3. A plaza becomes a negative volume because it is bordered by surrounding buildings and/or landscaping. Big plazas were, and still are, a sign of wealth and power. The huge expanse of a plaza is often defined by a positive volume in the middle of it—a fountain, monument, column, or arch. If the positive volume is too big or obtrusive, however, such as the pyramid in the middle of the plaza of the Musée du Louvre in Paris, it defeats rather than accentuates the imposing nature of the plaza's large negative volume.

4. Katz continues to comment on the influence of negative volume: "By the end of the film, Kane, an old man, who has lost much of his influence, is still photographed from low angles, but now the huge spaces of the rooms and halls of Xanadu dominate him and he seems small by comparison." *Film Directing,* p. 240.

5. Sergei Eisenstein, *Film Form and The Film Sense,* ed. and trans. by Jay Leyda (New York: World, 1957), *Film Form,* p. 54.

6. Malcolm Gladwell cites persuasive evidence of how our unconscious can be influenced to prime us for specific behavior patterns and attitudes. See especially the section "Primed for Action" in his book *Blink* (New York: Little, Brown, 2005), pp. 52–61.

7. Herbert Zettl, "The Graphication and Personification of Television News," in *Television Studies: Textual Analysis,* ed. by Gary Burns and Robert J. Thompson (New York: Praeger, 1989), pp. 137–63.

8. Zettl, "Graphication and Personification," pp. 157–61. Compare alienation through graphication with the idea of Bertold Brecht's *Verfremdungseffekte* [alienation effects]. See Bertold Brecht, *Brecht on Theatre: The Development of an Aesthetic,* ed. and trans. by John Willett (New York: Hill and Wang, 1964), pp. 90–99, 101. See also Boris Uspensky, *A Poetics of Composition,* trans. by Valentina Zavarin and Susan Wittig (Berkeley: University of California Press, 1973). Uspensky advances the idea of *ostranenie*—a clear separation of the stage event and the perception of such an event by the audience. This concept was first introduced by Russian theater critic Victor Shklovskij, who called it *priem ostrannenija*—a device for making strange.

9. Belgian painter René Magritte (1898–1967) tried to make us aware of such a switch between first- and second-order space in his witty painting *Les deux mystères* (the two mysteries), in which he shows a large pipe, which Sherlock Holmes would have liked, floating in the upper part of the painting, and an easel with a framed painting of the same pipe on the right half of the primary painting. The small painting on the easel has this caption: *Çeci n'est pas une pipe* (this is not a pipe). Apparently, the mystery is why the larger pipe is more readily perceived as the real thing than the small pipe in the painting on the easel, which he has labeled as not a pipe. If Magritte had read *Sight Sound Motion,* the two mysteries would have been readily solved: the big pipe is like the news anchor operating in first-order space, being "personified" (made more real), and the painting of the pipe on the easel is like the second-order box over the newscaster's shoulder. The graphicated pipe is definitely not the real thing. The caption that the little pipe is only a picture is no longer necessary.

   See John Willats, *Art and Representation: New Principles in the Analysis of Pictures* (Princeton, N.J.: Princeton University Press, 1997), pp. 274–275.

10. Erik Bucy and John Newhagen, "The Micro- and Macrodrama of Politics on Television: Effects of Media Format on Candidate Evaluation," *Journal of Broadcasting and Electronic Communication* 43, no. 2 (1999): 193–210.