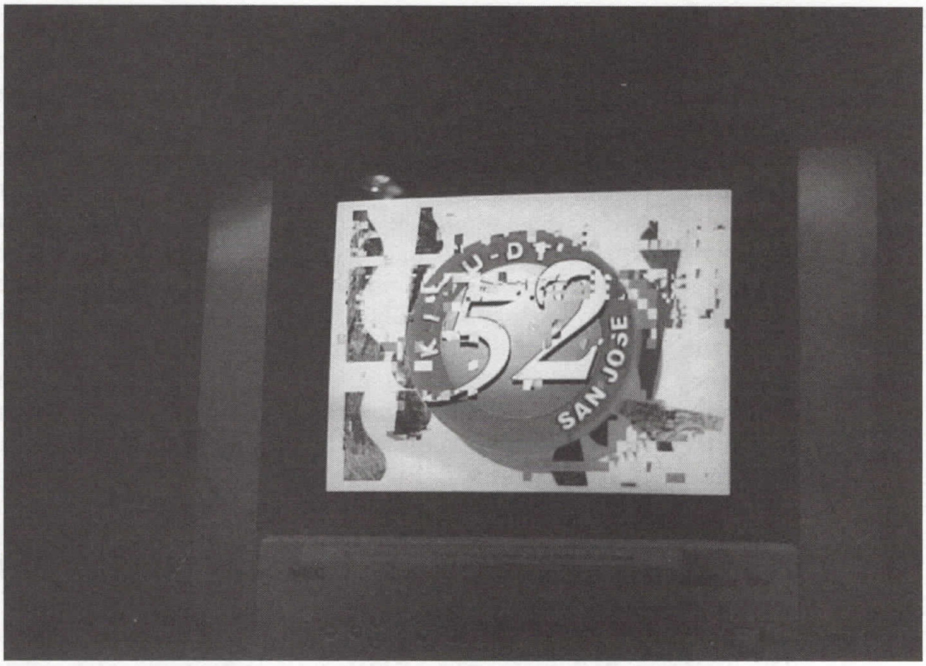

Figure 1.3 Here a received DTV signal is just over the cliff. Notice the video blocks that are out of place. Also notice that there is no increase in noise, no ghosting, just "blockiness." Some encoders and decoders can create block patterns that appear to be noiselike.

1.2.5 Defining HD

Back to HD versus SD. If SD is going to look so much better as DTV, how much better is HD DTV going to look? Well in one flavor of HD (yes, there are a number) the picture content is 6 times more than SD. In another flavor the picture content is only about 3 times greater than SD. Another philosophical argument in the industry is over what exactly is HD. Most agree that more picture content is a requirement. Is 3 times more picture content enough or should it be 6 times? As it turns out, HD is more complicated than that. Two additional ingredients in the HD recipe are how many pictures we present each second and how those pictures are displayed. Thomas Edison discovered that if you present still pictures at a fast enough rate, you create the illusion of motion. Hence, the motion picture. The standard frame rate settled out to be 24 frames/s, which means that every second 24 still pictures are flashed on the movie screen, and our eye and brain make them into a continuous moving picture. Television sends individual pictures to impart a sense of motion as well. Because the frequency of alternating-current (ac) power in this country is 60 Hz, it was easiest in television's infancy to base the rate at which television presented pictures to also be 60 Hz/s.

1.2.6 Resolution

Like any two-dimensional picture, a television picture has a horizontal and a vertical component. The way we draw this picture is one horizontal line at a time. We start at the top of the screen and scan one line from the left all the way to the right of the screen. This is done by an electron beam inside the picture tube. In black-and-white receivers the electron beam is made more intense for bright areas along the scan line than it is for dark areas. For color receivers the process is much more complex. Three scan lines, one for each of the primary colors (green, red, and blue), are swept across the screen. If all is working correctly, each beam will only land on its color phosphors on the face of the tube. These phosphors are what allows this scanning technique to work. Once they have been scanned and made to glow by the electron beam, they will continue to glow long enough so that our eyes (which aren't fast enough) do not notice them dim until the beam finally gets back to this line again. Once the first line has been scanned, the beam is turned off, and the beam-aiming (deflection) circuitry positions the beam back on the left side of the screen, slightly below where it started the first line. The second line is then scanned. This is done 238 more times, at which time the beam is at the bottom of the screen. The beam is turned off and brought back up to the top of the screen to start the process again. Each complete "painting" of the screen is called a *field.* The lines scanned during the first field have enough space left between them that another set of lines could be scanned or interlaced between them. That is exactly what happens in the next field. Both scans together are called a *frame.* These 480 active scan lines, which took two fields to create, is one NTSC picture. Each field happens in 1/60th of a second; therefore, one frame happens in 1/30th of a second. In addition to the 480 lines that make up the picture, there are 45 lines over the two fields in the vertical interval for a total of 525 lines/frame. These lines were needed in television's infancy to allow the television set to stop scanning vertically at the bottom of the picture and return (or retrace) to the top and start a new scan field.

In Fig. 1.4b each pulse (or dot along the bottom) represents the start of a horizontal line. These are called *sync pulses.* The first four and one-half horizontal lines at the left represent the last four and one-half active scan lines of a field. Then, notice the six half lines. These are called *equalizing pulses.* Next come six wide pulses followed by another six equalizing pulses. Together, these are called the *vertical interval.* The equalizing pulses and wide vertical interval pulses were needed by early television sets so the scan circuitry could stay in lock and to indicate when to start the vertical scan retrace. The blank horizontal lines after the vertical interval were originally there to allow time for the television set's deflection circuitry (which controls horizontal and vertical positioning of the electron beam or scan) to settle after the retrace. Now, these lines often carry data such as closed captioning, ghost-canceling signals, program information (for both broadcaster and viewer use), and test signals. The first three lines that are not blank are actually test signals used by broadcasters for quality control. The short pulses right

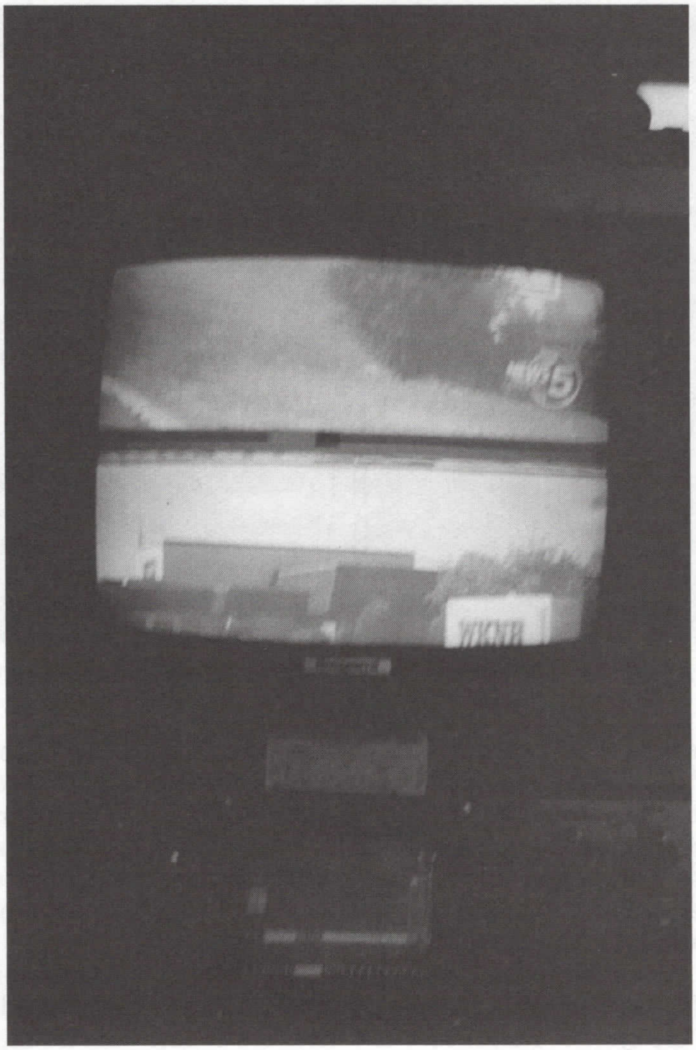

Figure 1.4a Video monitor able to be time-shifted to display the vertical interval.

after the sync pulses at midscale (amplitudewise) are the color reference (also known as color burst) information. These are the horizontal pulses referred to in Fig. 1.2. Since there are two fields in a frame and the total number of lines is an odd number (525), one field has an additional horizontal line. This is achieved by starting and ending one field with half lines. These half lines ensure that the proper interlace will take place. *Interlace* occurs when the scan lines of one field fall between the scan lines of the next field. More on this in Sec. 1.2.10.

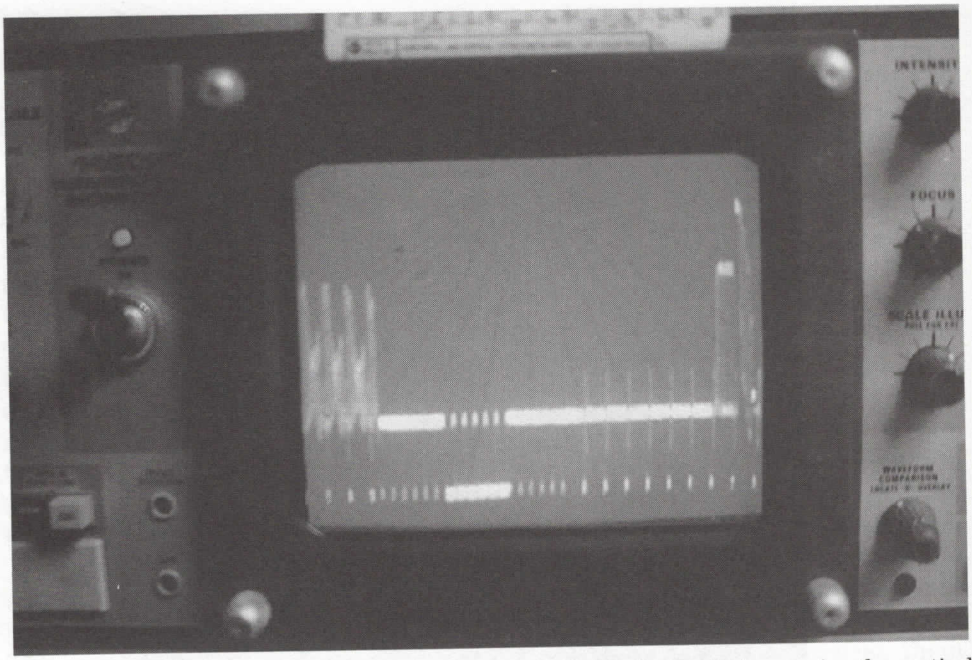

Figure 1.4b Waveform monitor below displays the varying voltages that comprise the vertical interval.

This means that if a camera shot a tall ladder, we could, in theory, see 240 rungs of that ladder. No, not 480, because to separately see each of the ladder's rungs, we would need one line for a rung, and at least one line of the space between the rungs. We would say that NTSC has 240 vertical lines of resolution. Now many things limit that resolution. One is how well the television set can interlace the scan lines for each field into a frame. This issue will resurface again very soon in our discussion.

What about horizontal resolution? Horizontal resolution is even less straightforward. Initially, horizontal resolution was mainly based on the bandwidth of the system. *Bandwidth* is a measure of how wide a range of frequencies can be sent. Your telephone sounds "tinny" because only 3000 Hz of voice, or audio, information can get through. This frequency range has been kept low to keep phone rates from rising higher. Modems play many modulation tricks to get data throughput much higher. Your AM (amplitude-modulation) radio sounds slightly better because back when the FCC set up AM broadcasting, each AM station was given only 10 kHz of spectrum, which equates to an audio bandwidth of only 5 kHz, as the audio information is duplicated on the upper and lower part of the signal (called dual sideband). Although the signal could be made to occupy only the upper or lower (called single sideband) side, it would result in much more expensive receivers. FM (frequency modulation) was given much more bandwidth, and thus it can broadcast audio frequencies up to 15 kHz, which is near the upper hearing limit of most people. That is why most music is on FM, and AM has gravitated to news and talk shows, where the fidelity doesn't have to be as good. So television must have very wide bandwidth compared to other RF services to have good horizontal resolution.

1.2.7 Bandwidth

When it came time to allocate spectrum for NTSC television, it was decided that a 6-MHz-wide television channel would be enough. However, that channel had to hold not only the video or picture but also the aural or sound. Television stations are really two broadcasting stations in one. The picture, or visual half, uses the same transmission technique as AM radio, except that a bandwidth of just under 5 MHz is available for pictures, unlike the 5 kHz for AM radio. The aural or audio part of the signal is FM just like any FM radio station. Like FM radio, 150 kHz of the NTSC television channel is allocated to the audio. In AM the amplitude of the carrier is modulated (varied) to convey information to the receiver, and in FM it is the frequency that is modulated. Now why doesn't it take 10 MHz to transmit 5 MHz of video you might ask? Because television is not transmitted double sideband. It uses a technique called *vestigial sideband,* which means that the entire upper sideband is transmitted and only part of the lower. The gains at various frequencies are controlled so that any frequencies below the carrier are combined to make the response across the whole channel flat. This means that, in theory, all frequencies across the channel are treated equally (Fig. 1.5).

So now that we've explained how much bandwidth is available for video in the NTSC television channel, we can get back to horizontal resolution. A signal with a frequency of 5 MHz would have a period of 200 ns (0.0000002 s). This period would be represented by a sine wave with a positive and negative excursion. It takes approximately 52 μs (0.000052 s) for the visible portion of a horizontal line to be swept across the screen, which means that 260 occurrences of our 5-MHz sine wave could occur across the screen. If instead

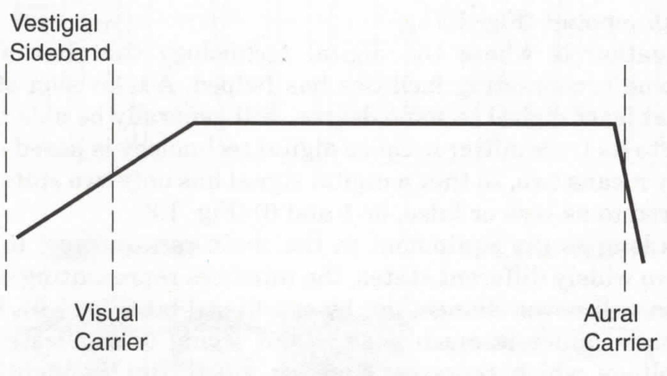

Figure 1.5 Response across NTSC channel.

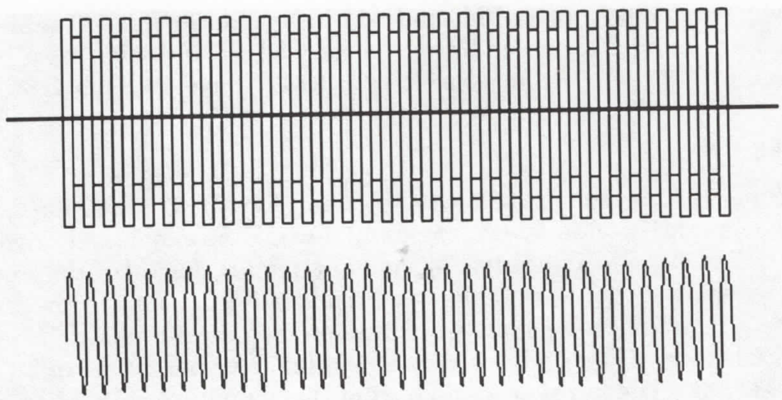

Figure 1.6 Picket fence scanned by camera and resultant electronic signal.

of displaying a vertical ladder on the screen, we shot a picket fence extending horizontally across the screen, we would find that the maximum pickets that could be displayed would be 260 (Fig. 1.6). The reason is that because each picket, and the corresponding space between pickets, could be represented by our 5-MHz sine wave. Of course, not all picture content is at 5 MHz; picture content can be a cacophony of many frequencies from 0 to 5 MHz.

1.2.8 Maintaining quality

The NTSC television signal we receive today has a theoretical resolution of 260 (horizontal) by 240 (vertical). But, and that is a big "but," that resolution is only obtained if the NTSC television station you are watching has the best equipment with no engineering or equipment problems, the transmission process and your receiving antenna/cable system are perfect, and finally your set is in perfect working order. Care to place your bets? Every time an analog signal passes through cable, or a piece of equipment, some small nip is taken off the high-frequency information, and all the lower frequencies are scrambled a small amount (Fig. 1.7).

This situation is where the digital technology that has slowly spread through some broadcasting facilities has helped. A television station that is digital, or at least digital to some degree, will generally be able to deliver better quality to its transmitter because digital technology is based on binary values. Binary means two, so that a digital signal has only two states, high or low (also referred to as true or false, or 1 and 0) (Fig. 1.8).

Thus, as long as the equipment in the chain can continue to differentiate between two widely different states, the numbers representing pieces of video information will never change, not by one literal bit (Fig. 1.9). Working with these number values is much easier than trying to duplicate a constantly varying voltage which represents analog video. But as mentioned earlier, even if the signal is digital in the television station, it must be converted back

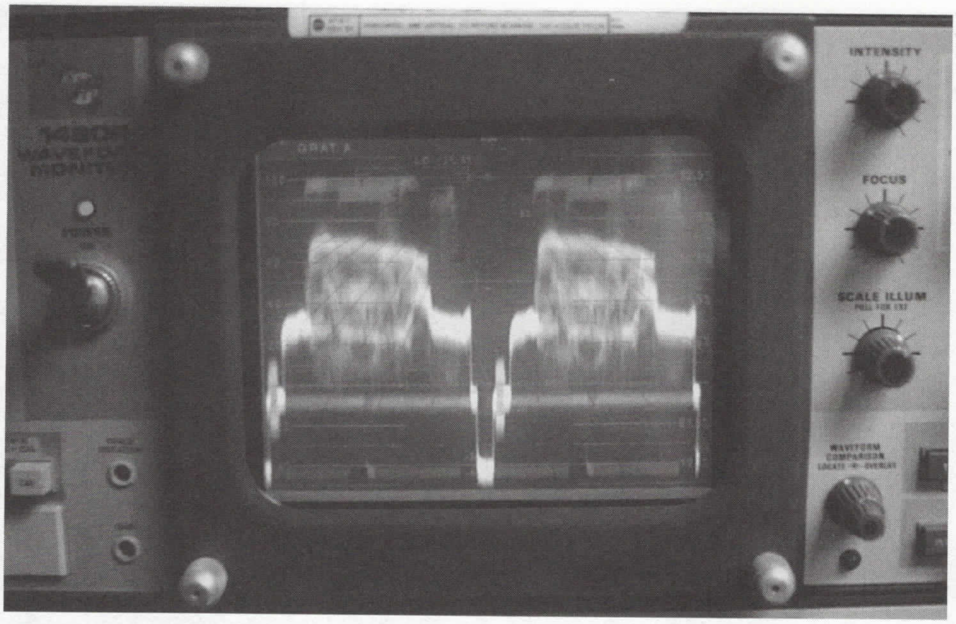

Figure 1.7 Received NTSC video waveform.

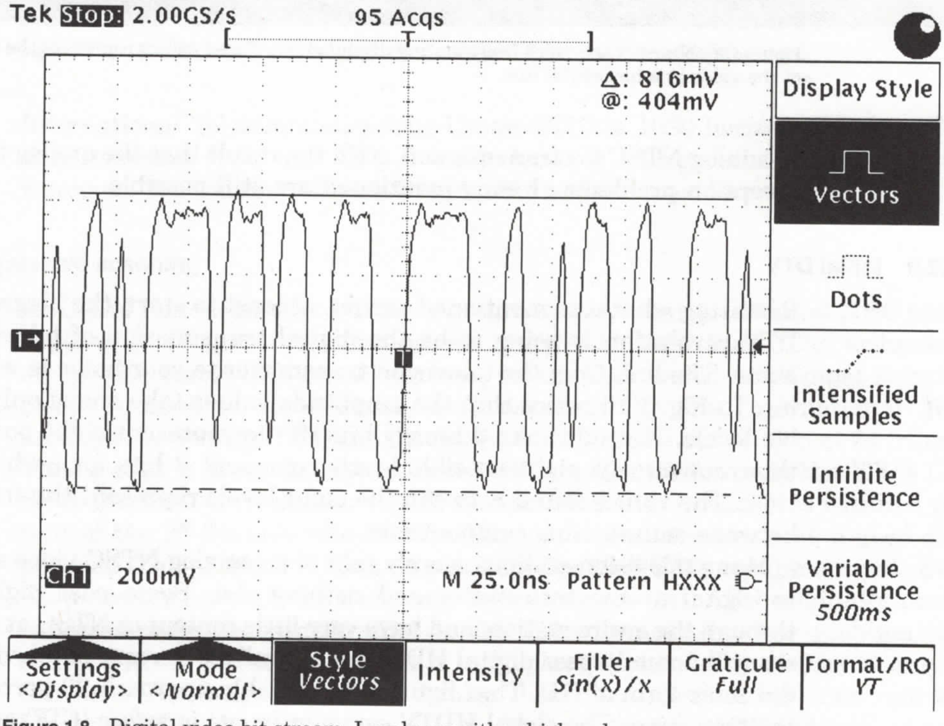

Figure 1.8 Digital video bit stream. Levels are either high or low.

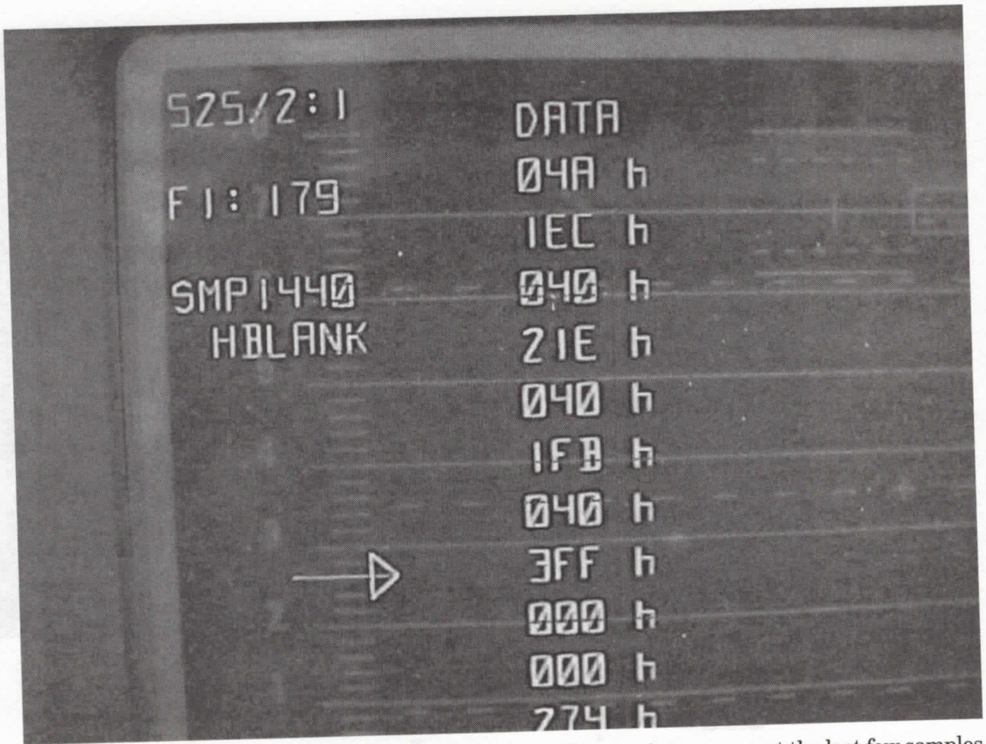

Figure 1.9 Number sequence representing digital video. These values represent the last few samples of the video in a horizontal line.

to analog NTSC for transmission, with the result that the analog transmission/reception problems already mentioned are still possible.

1.2.9 Initial DTV

Restating what was mentioned earlier, at least to start, the biggest part of the DTV revolution is going to be the digital transmission of television to your home. The link from the television transmitter to your home is what is changing. In Fig. 1.10 notice that the amplitude values take one of only eight possible levels. It would take 3 binary bits (2^3) to represent eight possible values; thus, conversely, eight possible levels represent 3 bits for each DTV sample time. The values found between the eight levels represent the transition time between actual value sample times.

Many television stations merely take their analog NTSC video and convert it to digital at the transmitter and nothing else. Some pass digital SD video through the entire station, and have very little content in NTSC at all. A few others will actually pass digital HD video. Initially, 5 percent of television facilities did some form of HD. That figure will probably approach 20 percent early into the transition. The global HDTV *common image interface* (CIF) adopted by the

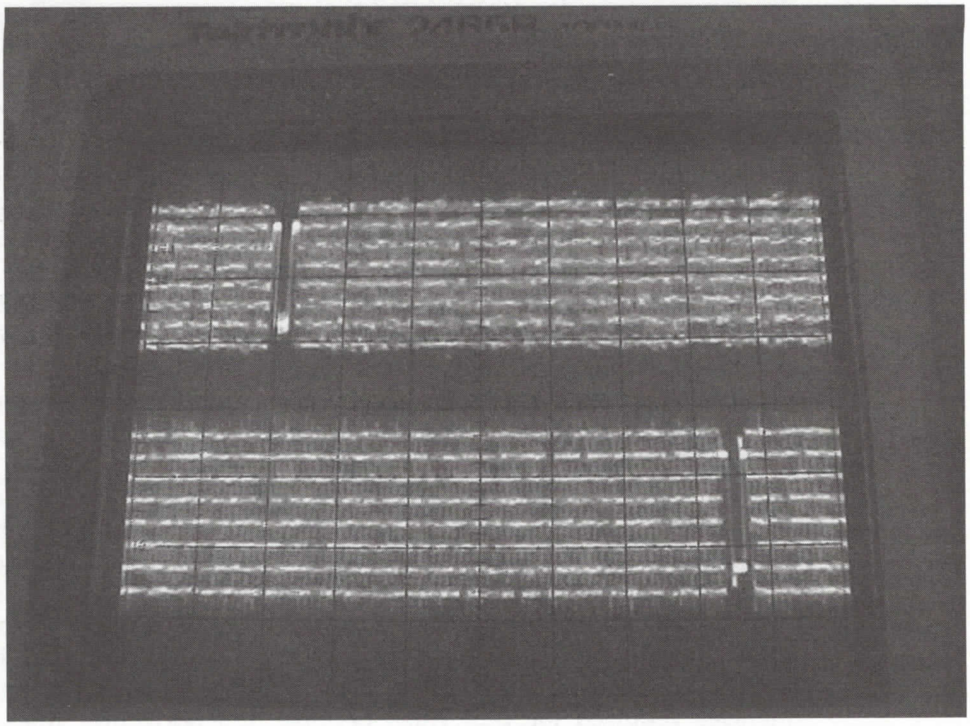

Figure 1.10 Transmitted DTV signal. Preprocessed (top), postprocessed (bottom).

International Telecommunications Union (ITU) is 1080 horizontal lines. ATSC only considers 1080i and 720p (i = interlaced, p = progressive) as HD, but the European Telecommunications Standards Institute considers 480p to be HD.

1.2.10 Progressive scanning

ABC and Fox insist that HD have progressive scanning. CBS and NBC aren't opposed to progressive scanning but they want 1080 lines. The technology, along with bandwidth considerations, means that, today, 1080 must be interlaced. A possible compromise, first advanced by the International Teleproductions Society (ITS) and discussed within SMPTE, is 1080p at 24 frames. This is even a lower data rate than 1080i. It also conforms to the ITU's CIF. Since most network shows are still shot on film, and because it is within 4 percent of the 25-frame/s rate used in Europe, it would seem to be a good compromise. The problem is that no one supplies 1080@24p equipment yet. Even 720p doesn't have all the necessary infrastructure yet. ABC originally used a supercomputer to transfer film to its 720p format. Another problem that occurs with 1080p at 24 frames/s is in displays that *actually* display at only 24 frames/s. The fatigue from the display flicker with such a low scan rate will hinder production monitoring at the 24-frame/s rate. 1080p24 will require monitors that display at a much higher rate even though the actual frame scan rate is only at 24 frames/s.

We covered the resolution part of what makes an HD picture, but only partly. When we talk of horizontal and vertical resolution, we are referring to *spatial resolution,* that is, resolution in two-dimensional space. There is another type of resolution called *temporal,* or resolution that occurs over time. As already mentioned, NTSC provides new complete pictures at 30 times/s, which is the frame rate, but remember that it takes two scans, or fields, to make up that frame. It is the interlacing of those two fields that results in the frame. There are proponents who claim that 60 frames/s is a better, maybe higher, definition picture, but interlace was originally adopted to limit the bandwidth needed. Intuitively, it should be obvious that to send all the scan lines on each vertical sweep should double the information or frequency bandwidth required. That is one reason why interlace was adopted early on. Another reason is that the phosphors used in early picture tubes, or CRTs, didn't glow long enough between scans if all the horizontal lines were sent each scan but at a rate that didn't increase the bandwidth. That is, one vertical sweep would take 1/30th of a second instead of 1/60th. This problem is what is generally called *flicker.*

The computer people are big proponents of sending all the horizontal lines each vertical scan. This is called *progressive scanning,* and is what your computer monitor does. However, to eliminate flicker, vertical scanning is usually done at 72 times/s instead of 60, which results in a very wide band of frequencies required for this video because 480 horizontal lines are sent 72 times/s. This makes your computer monitor display 28,800 active horizontal lines/s versus half that for NTSC video. However, since it only has to be sent a few feet and not through a television channel, it doesn't matter.

Computer people who feel production, or broadcast, facilities should do 480p, or 720p, because it is computer friendly and because cameras in that format are available have not grasped the entire scope of even a small facility. While TV people seem to add complexity in their minds when it comes to computers, most computer people simplify the infrastructure in a television plant to cameras, maybe routers, transmitters, and monitors. And most computer people think all storage is on NT servers. To be fair, 720p switchers and VTRs are now available. But computer types have claimed that the broadcasters' problem is due to long-standing regulation by the FCC so that the future path is only charted by managers who have no grasp of new technology. Some of that is undoubtedly true. However, creating specifications for the whole DTV pipeline based on what is easy for a PC to display will most likely optimize PC display at the expense of real-time television display. This fact might be true only because the computer industry is adamant that optimizing the bit stream for display on a television receiver will stunt PC use for DTV.

NHK had pressed Panasonic to stop 720p demos and production in 1998. But the U.S. Department of Defense's (DOD) complaints were forwarded to Japan, which stopped NHK's efforts. The National Imagery and Mapping Agency (NIMA), a part of the Pentagon, has embraced 720p.

So the argument goes, is it better to refresh all the lines each vertical scan and eat up bandwidth that way (the progressive scanning approach) or is it better to send only half the lines each vertical scan and use the bandwidth for spatial (horizontal and vertical) resolution? Temporal resolution, the number of complete pictures we send a second, and interlace versus progressive scanning are linked together.

Besides pushing for progressive scanning, the PC camp is worried about interactivity. The Advanced TV Enhancement Forum, a coalition formed by Microsoft, Intel, Disney, Direct TV, CNN, and others, recently agreed on a standard for enhanced TV content. This standard applies to over-the-air (terrestrial) cable and satellite transmission for the enhancement of interactive TV.

1.2.11 Aspect ratios

There is one other issue clouding the SD/HD discussion: aspect ratios. The television screen you watch today is one-third wider than it is tall. In other words, for every 3 in high it is 4 in wide. Hence, the term *4:3 aspect ratio* is used to describe the dimensions of the screen. This was the original film format, but cinematographers have realized that our peripheral vision extends wider than taller, so they started making films with wider aspect ratios. Many films are shot with aspect ratios as wide as 21:9. With the advent of HDTV, it was decided that television should have a wider aspect ratio also; 16:9 was settled on after early pioneer NHK adopted it for their HDTV experiments. HD formats 720p and 1080i are both 16:9. Once the aspect ratio is settled on horizontal, pixel count and vertical line counts fall out. The H/V ratios for the two HD formats are as follows:

1920 pixels $\div$ 16 = 120	120 $\times$ 9 = 1080 lines
1280 pixels $\div$ 16 = 80	80 $\times$ 9 = 720 lines

Many program producers are starting to acquire SD material in 16:9 format also. Many professional cameras offer options that allow acquisition in 16:9. Some mistakenly believe that any 16:9 programming they see must be HD, just as all DTV programming is HD. Neither is true.

1.3 Why Compression?

SD versus HD is only a part of the DTV equation. Digital television ushered another transforming technology into prominence—compression—more exactly, lossy compression. As already stated, digital has only two states. When digital changes states, it does so abruptly. When it goes from one state to the next, it does it almost instantaneously. What's instantaneous? A few nanoseconds, which results in lots of sharp edges. What does it take to make lots of sharp edges? Very high frequencies. This means that any coax carrying digital signals better be able to handle extremely high frequencies. How high? Up into the gigahertz range. As these high frequencies are rolled off, or attenuated, the nice