CinemaSource™

Introduction to Digital Television

CinemaSource ,
18 Denbow Rd., Durham, NH 03824
cinemasource.com
800-483-9778

CinemaSource Technical Bulletins. Copyright 2002 by CinemaSource, Inc.
All rights reserved. Printed in the United States of America.

No part of this bulletin may be used or reproduced in any manner whatsoever without written permission, except in brief quotations embodied in critical reviews.

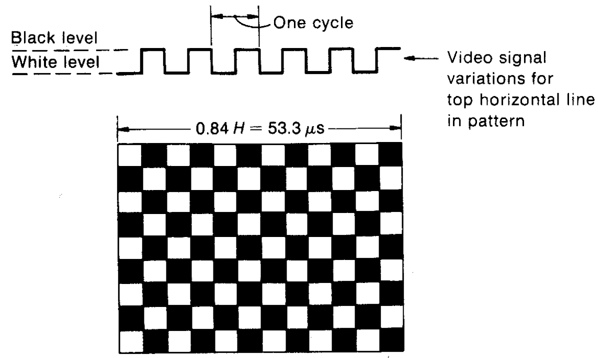
CinemaSource is a registered federal trademark.

For information contact: The CinemaSource Press, 18 Denbow Rd. Durham, NH 03824

Introduction to Digital Television

This guide is a collection of articles related to the rollout of digital television (DTV). The articles have been chosen from various sources and each explains a particular segment of the complex DTV puzzle. We hope that you find this guide useful. Please fell free to email us at CinemaSource if you have any comments or suggestions.

Chapter One: Basic Video Concepts



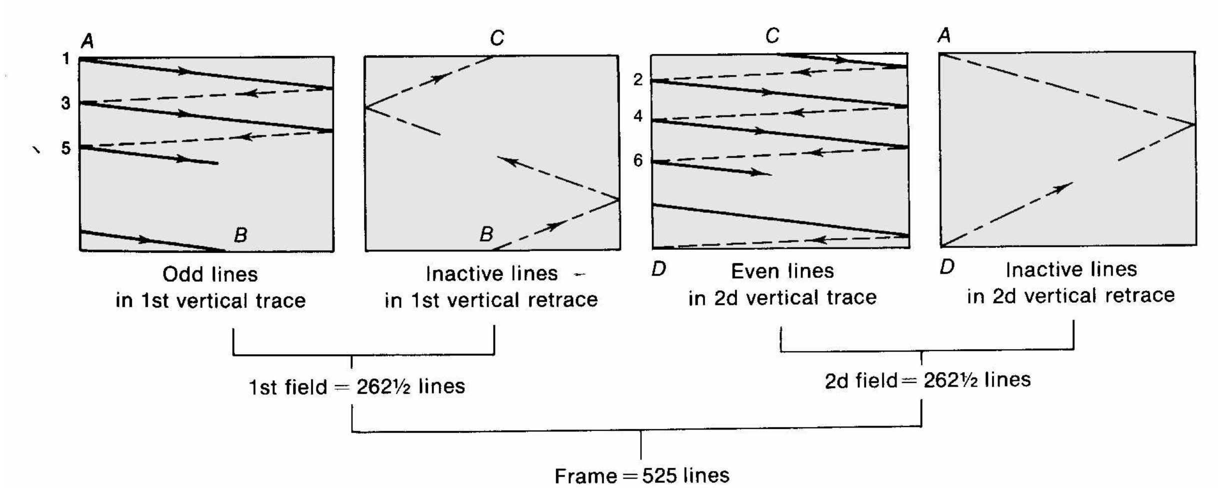
The term resolution is used to quantify a display device's ability to reproduce fine detail in a video image. In a solid state imaging device (LCD, DILA, DLP), the resolution is simply the number of pixels on the imaging elements. In a raster scanned CRT-based device,it is a very different mechanism and there is a significant difference between horizontal and vertical resolution.

Vertical Resolution of a CRT:

Below we have a diagram that shows how an electron beam is "scanned" across a picture tube faceplate to form a NTSC video image. The technique of interlacing the images was developed to minimize that bandwidth of the signal and reduce flicker in the display. The maximum vertical resolution is simply the number of scan lines visible in the display. This number is the number of horizontal scan lines (525) minus the retrace lines (43). Thus the maximum vertical resolution of a NTSC display is 525 -43 = 482 lines. With an HDTV image, the image is swept faster with the result of more lines of resolution.

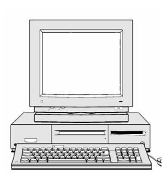
Horizontal Resolution of a CRT:

Horizontal resolution is a completely different mechanism in a CRT-based device. The horizontal resolution is a function of how fast you can turn the electron beam on and off. The image above illustrates this. Here a checkerboard pattern is being displayed by making the electron bean turn on and off very rapidly. Note: By convention, video resolution is measured in picture heights, whether it is vertical resolution or horizontal resolution. So the horizontal resolution is the number of resolvable vertical lines across a width of the display equal to the picture height. For a 4:3 display this is equivalent to 75% of the resolvable lines across the full width of the display.

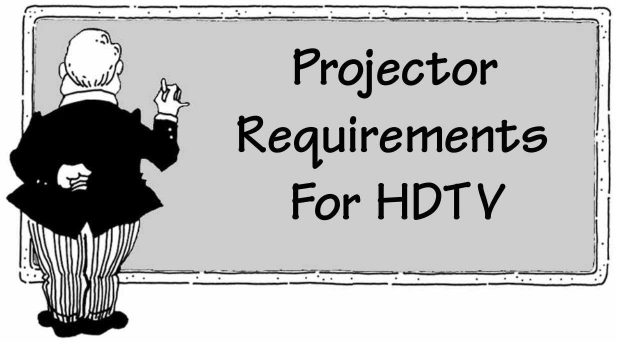


Horizontal Resolution of Various Video Sources

Signal Source	Horiz Resolution
VHS Tape	240
Terrestrial Broadcast	330
Laserdisc	425
DVD	480
Satellite	500

COMPUTER/GRAPHICS ADAPTER		RESOLUTION (H/V)	HORIZONTAL FREQUENCY	VERTICAL FREQUENCY
	VGA	640/480 - 640/400	31.5 Khz	60/72 hz
	SVGA	800/600	35.5 Khz	60/72/76 hz
	XGA	1024x 768	39.4 Khz	60/72/76 hz
	SXGA	1280 x 1024	37.9/48.4/61 Khz	60/72/76 hz

THE GRAND ALLIANCE DIGITAL TELEVISION FORMATS	Horizontal Pixels Across Screen Width (Horizontal Resolution)	Viewable Scan Lines (Vertical Resolution)	Image Aspect Ratio	Picture Refresh Rates
High Definition Television (HDTV) 1080P/1080I	1920	1080	16:9	24/30Hz - Progressive 60Hz - Interlaced
High Definition Television (HDTV) 720P/720I	1280	720	16:9	24/30/60Hz - Progressive
480P/480I and Others	704	480	4:3-16:9	24/30/60 Hz - Progressive 60Hz - Interlaced
Standard Definition Television (SDTV)	640	480	4:3	24/30/60Hz - Progressive 60 Hz - Interlaced



Many thanks to Greg Rogers of The Perfect Vision for his permission to reprint this HDTV material. You can visit his web site at www.cybertheater.com.

Projector requirements can be separated into two categories. First are the minimum requirements to display the HDTV format signals. These include Horizontal Scan Rates, Horizontal Retrace Time and Vertical Retrace Time. If these requirements are not met, then the projector will not sync to the signals and no stable picture will be produced. Second are the requirements necessary to achieve the maximum quality delivered by the HDTV format. Even if the second requirements are not fully met, the picture quality from an HDTV source should still exceed that delivered by SDTV sources.

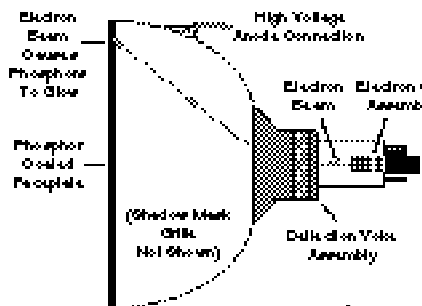
In most cases a video display device's vertical resolution for a 4:3 picture will be approximately the same as the horizontal resolution in TV Lines (referenced to a 4:3 picture height). This assumes that the spot size is approximately round (this will require good adjustment of astigmatism and focus) and that the horizontal resolution is not limited by the RGB bandwidth.

Only a handful of projectors will have sufficiently small spot sizes to truly display the full vertical resolution of the 1080I format. On the positive side, some modest overlapping of the scan lines will help hide the interlacing artifacts from this format. There is a very fine balancing act going on here that makes for the interesting debate between proponents of the 720P and 1080I formats. The 1080I format should produce a 50% improvement in horizontal and vertical resolution over 720P, but CRT and optical limitations in projectors (and aperture-grille and shadow-mask limitations in direct view TVs) will limit that significantly. But those same limitations partially obscure the visibility of the fine line-twitter and other interlace artifacts of the 1080I format. So the debate goes on between proponents of the 1080I and the 720P formats as the networks and others choose up sides for the best HDTV signal.

Feature	NTSC Composite Digital	DVD ITU-R-601	DVD Line Doubled	DVD Line Quadrupled	ATSC 720P	ATSC 1080I	ATSC 1080I Line Doubled
Horizontal Scan Rate (KHz)	15.734	15.734	31.468	62.937	45.0	33.75	67.5
Horizontal Retrace Time (uS)	9.92	10.22	5.11	2.55	4.98	3.77	1.89
Vertical Retrace Time (uS)	1335	1430	1430	1430	667	667	667
Projector Bandwidth (-3 dB) (MHz)	10	10	20	40	50	50	100
Limiting Horiz Resolution per Picture Width (TV Lines)	448 RF 567 LD	667	667	667	1138	1707	1707
Limiting Horiz Resolution/PH (TVL for 4:3 display)	336 RF 424 LD	506	506	506	853	1280	1280
Limiting Horiz Resolution/PH (TVL for 16:9 display)	N/A	380	380	380	640	960	960
Vertical Resolution Lines per Picture Height for a 4:3 Display * 16:9 'Anamorphic' Format	483	480 640 *	480 640 *	960 1280 *	960	1440	1440

Video Projector Requirements For HDTV

Since debuting in the late 1930s, television receivers and the images they display, have evolved continuously and prodigiously. From small, marginally acceptable, B&W affairs television images have morphed into enormous, full color, theater-like displays. And this remarkable change can be attributed to the unrelenting R&D efforts on the parts of hundreds of video technology companies, and individuals, all in pursuit of progress and "competitive advantage". Yet despite the magnitude of this effort, and major advancements in componentry, such as transistors, integrated circuits and microprocessors, the NTSC color signal remains firmly rooted in the past.



How picture tubes produce light

Raster Scanning 101

Raster scanning is the standard process by which CRT-based display devices create video images. There are other ways to derive images from CRT displays, such as vector-based methods (used in some air traffic control displays and military applications), but by far the most common method used is raster scanning. Raster scanning refers to the method by which video images are actually "assembled" on the face of the CRT. But before we dig into the principals of scanning, let's consider how standard picture tubes actually generate light.

It starts with a device located deep in the neck of all picture tubes called an electron gun. Electron guns are assemblies that are designed to emit, focus and control streams of electron particles. They are connected to external high voltage power supplies which generate a tremendous potential (27 to 32 Kilovolts) between the electron gun and shadow mask/face plate assemblies. The result is that electrons fly off the cathode surface of the electron gun, and head straight for individual phosphor patches deposited on the face plate. After impact, the phosphors glow, for a brief moment, and then extinguish.

The key to making a complete video image with this system is to scan all phosphor patches across the face plate repeatedly. And this is where raster scanning comes into the story.

Looking straight at the face of a picture tube, the raster scanning process starts in the upper left hand corner. The electron beam is positioned here, electromagnetically, by the deflection yoke assembly. Scanning starts when the beam is rapidly swept from the left side of the tube over to the right, again, electromagnetically. As it runs across the tube face, the electron beam varies in intensity and causes the phosphors to glow in differing amounts. This first completed sweep becomes one thin slice of a complete video image. Next, the beam is then blanked (turned off) and "flys back" to the left hand side of the tube, and then the whole process begins again. Scan...flyback...scan...flyback... this procedure occurs until the scanning reaches the bottom of the tube and one pass is completed. The electron beam is now blanked again, this time for a longer period, and the vertical section of the deflection yoke lifts the electron beam up to the left-hand top of the tube where the next pass begins.

Now that we have illustrated how one complete pass is completed, let's look at how others are added. This can be accomplished in two ways; either by "interlacing" the scans, or simply writing the entire image at once; "progressively". As it turns out, you have seen both methods in use. Interlaced scanning is the technique utilized by all standard NTSC television receivers. It is called interlacing because incomplete "A fields" are displayed first and then "B fields" come along and interlace between the lines. The diagram on the next page illustrates this. In case you think this is an odd way to create video images, you're right. But there's a good reason for it, and that is to conserve bandwidth. By using scans that interlace, the resultant television signal is half the size (in frequency) as a progressively scanned one, and in the telecommunications world, bandwidth is scarce. There is only so much bandwidth (frequency spectrum) to go around, so engineers are constantly finding ways to maximize the amount of information they can fit into a allotted frequency slots. In the all-analog world of the 1930s, interlacing was the technique chosen to keep the size of the signal manageable, and as a side benefit, it made the receivers less expensive to produce (more on this later).

Progressive scanning is another way to generate and display video images. Instead of transmitting interlacing A & B fields, a complete video image is transmitted all at once. The computer industry long ago decided that progressive scanning was the technique of choice for them. Since they are not constrained to narrow terrestrial broadcast channels, the computer manufacturers went for maximum image quality. Progressive scanning is the technique used on all standard computer monitors.

The Evils of Interlace

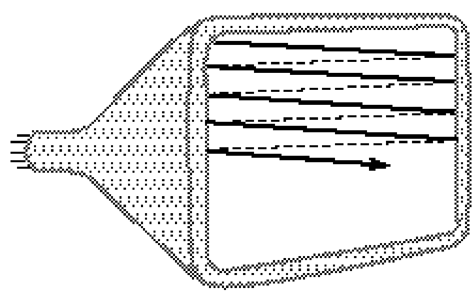
Not only does the concept of interlacing video images seem odd, it also produces odd artifacts. The engineers that designed the system long ago were well aware of these artifacts, but weren't bothered because they were considered imperceptible on the small 5 to 9" B&W displays common at the time. And today? Well, we have displays over ten times that size and, as a result, interlacing artifacts can sometimes be seen. For example:

1) Interline Flicker. Video consists of a rapid series of images or frames displayed one after another. They occur so rapidly that the human visual system integrates them into a continuous moving image. However, if the frequency of frames slows down, you will see the video image flickering, just like in an old B&W movie. This critical "flicker frequency", as measured by countless psychoperceptual studies, occurs somewhere below 50-60 times per second (it depends on the person observing, some people are more perceptible to flicker than others.) Now this is not a problem with larger objects being displayed because both the A and B fields contain sections from the same image. However, if the image is made up of fine horizontal lines, some of the information may not be averaged over different fields. It will show up in specific fields, either all the A fields, or the B fields, and because these are drawn 30 times per second, you are bound to see interline flicker. Engineers sometimes refer to this problem as "venetian blind flutter" because venetian blinds are one of the most common objects demonstrating the phenomena. It occurs when the venetian blind is just at the right size so that each blade of the blind is scanned in the same field. The result is the entire blind pulsates at 30 hz. Our diagram shows how this could happen.

2) Reduction of Vertical Resolution. Another artifact that interlacing brings to us is a reduction in resolution that occurs when fine detailed images move up and down. What happens is that when objects move at exactly the right rate, one video field captures the movement of the object as it scrolls vertically, and the other does not. The effect is to cut the vertical resolution in half because only one field is used to transmit the image. Unfortunately, this often occurs when credits scroll at just the right speed and the result is poor legibility

Can anything be done to help NTSC signals?

On standard NTSC television receivers, not much. Interlacing, and it's attendant artifacts, are simply a way of life. It's been that way since the beginning of television broadcasting. But don't lose sleep over this, interlacing artifacts are rarely perceptible on smaller displays (under 50 inches or so). They really are more of an academic problem, and only occasionally seen in significantly larger images. But you say you want to build a home theater with a 100" front projected display? Then, there is one device that can help: a line doubler.



How Raster Scanning Works: Scan...Flyback...Scan...Flybac

Line doublers are signal processing devices that take standard NTSC video, adds some image enhancement, and converts the signal to progressively scanned 31.5Khz video. Because the output of these devices is progressively scanned, the artifacts we illustrated before are not seen. (It is impossible to get a 30 hz flicker in a 60hz progressively scanned image because every single pixel is refreshed at a 60 hz rate.) But note: because the line-doubled output signal is a higher scan rate than NTSC, it must be displayed by a data or graphics-rate display device, typically a front projection monitor. These are more expensive than standard video-grade monitors.

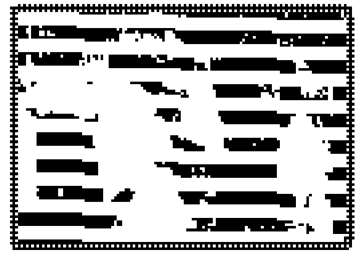
Enter DTV

The reason discussions of interlace vs progressive scanning are becoming so common these days is because of the new high resolution DTV standards. This new standard, DTV (previously referred to as "HDTV" and

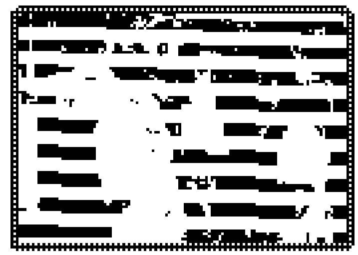
"ATV"), is almost certainly going to incorporate both types of scans. You would think with a new, state-of-the-art, digital television standard about to appear, that interlaced scanning as a technique would be relegated to the video history books. However, this is not the case, and there are several reasons for it.

It starts with the early days of the Grand Alliance. This consortium of key industry groups, including AT&T, General Instruments, MIT, Philips, Sarnoff Labs, Thompson and Zenith, was allowed by the FCC to combine forces and help define the final digital television standard. Incorporating the desires of the television broadcast industry, the computer industry, and international groups, the Grand Alliance suggested four main "modes" and 18 different signal types for the digital television signal format. Today, as you probably know, the list has been widdled down to 480I, 480P, 720P and 1080I The lowest resolution mode 480I is interlaced. The reason for the incorporation of this particular specification is for backward compatibility with existing sets.. This format will be able to be utilized by conventional NTSC television receivers after it is converted from digital to analog composite signal form.

The purpose of the other interlaced scanning mode is more practical. Why would one want to compromise the stellar quality of a 1920 x 1080 (1080I) high resolution mode with antiquated interlacing scanning? The reason is cost. Building interlaced monitors can be significantly cheaper than progressive scanned ones. Interlaced monitors run at slower horizontal scan rates, so deflection circuitry is less expensive and with interlaced monitors, the bandwidth of the video signal channel is less, so video processing and CRT drive boards are less expensive to design and build. And about the artifacts? On smaller displays artifacts are unlikely to be a problem, because they will be minor in nature and difficult to see at high resolutions. So the television broadcast industry has argued that even at the highest resolution mode, the economics of the matter decree that interlacing still has home in digital television displays. Believe it or not, this bandwidth reducing technique from the 1930s is still with us in the new millenium and will be for decades to come.



Interlaced Scanning Field A
(every 1/60th second)

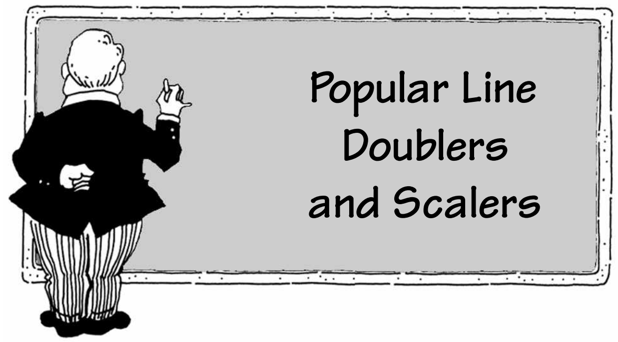


Interlaced Scanning Field B
(every 1/60th second)

Progressive Scanning Frame 1
(every 1/60th second)

Progressive Scanning Frame 2
(every 1/60th second)

Interlacing vs Progressive Scan



DWIN LD-10 Line Doubler

LINE DOUBLERS:

IEV TurboScan 1500 - Converts 480I to 480P

NEC IPS 4000 - Converts 480I to 480P

DVDO -Converts 480I to 480P

DWIN - Converts 480I to 480P

SONY EXB-DS10 -Converts 480I to 480P

QUADRUPLERS:

IEV TurboScan 4000 - Converts 480I to 960P

LINE MULTIPLIERS:

DWIN TranScanner - Converts 480i (composite, s-video and component) to 960P in 200Khz increments

SCALERS:

Communications Specialities Deuce -Converts 480I to 480P, 600P, 960P, 1024P

Faroudja DVP-2200- Converts 480I (composite, s-video and component) to 480P, 600P

Faroudja DVP-3000- Converts 480I (composite, s-video and component) to 480P, 600P, 720P, 960P, 1080i, 1080P

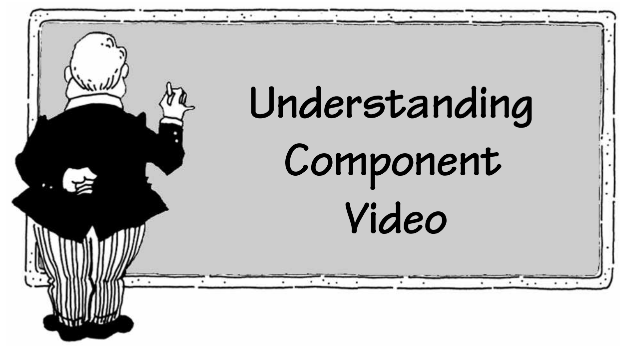
NEC IPS 4000Q - Converts 480I to 480P, 600P, 768P, 960P

QuadScan Elite - Converts 480I (composite, s-video and component) to 480P, 600P, 768P, and 1024P (1365x1024 - for DILA)

NATIVE RATE:

Faroudja NR - Depending on the model, converts 480I (composite, s-video and component) to 480P, 480P plasma,600P, 768P plasma, 1024P (1365x1024 -for DILA)

COMPUTER/GRAPHICS ADAPTER		RESOLUTION (H/V)	HORIZONTAL FREQUENCY	VERTICAL FREQUENCY
	VGA	640/480 - 640/400	31.5 Khz	60/72 hz
	SVGA	900/600	35.5 Khz	60/72/76 hz
	XGA	1024 x 768	39.4 Khz	60/72/76 hz
	SXGA	1280 x 1024	37.9/48.4/61 Khz	60/72/76 hz



Poor NTSC composite video, it's limitations are constantly being revealed and denounced. If you have a smaller TV, you may not have been pulled into the debate because with most smaller images, composite video looks just fine. However, blow the image up big with a projection TV and the situation often changes. For example: large colored graphics can blur beyond their borders. Fine detailed sections of the image often ripple with multicolored rainbows and fine detailed noise can permeate the image in the form of twinkling granulations. And if you are connected to cable, you have undoubtedly seen ghostly images, moving lines and other sundry things floating about. The questions are: are these artifacts endemic to composite video, and if so, is there a better method?

The answer to both questions is yes. But, in order to understand why, let’s look closer at the signals themselves. First, you should know that the artifacts illustrated above have nothing to do with the actual integrity of the basic video signals. They are almost entirely due to the way that the signals are encoded for “composite” distribution. Walter Allen, AmPro's VP of Prosumer Technology, explains: "Standard video delivery methods, such as terrestrial broadcast, cable TV, videotapes, and even laserdiscs, all utilize NTSC composite signals. These are reliable ways to deliver video signals, but because of the encoding/decoding involved, visible image artifacts can be seen in the image. The important part to realize is that these artifacts come from composite encoding/decoding processes, not the fundamental signals themselves. "

How It Happened

In the early 1950s, the desire to add color to the existing B&W television format was very strong. RCA, NBC, and other companies, had each proposed and demonstrated their own color encoding system. The FCC, who desired to keep this selection process under control, formed a working group of industry scientists and engineers to evaluate the proposals. This group, called the National Television and Systems Committee (NTSC), faced a number of weighty considerations, but the foremost was *compatibility.* Would it be possible to design a new color broadcast format to be completely compatible with the existing B&W one? At first, the answer appeared to be no, because the initial favorite was an incompatible one, but after much discussion and lobbying, a completely compatible one, based on the RCAsubcarrier chroma system, was chosen. Color broadcasting began on January, 1st 1954, when the tournament of Roses parade was broadcast live from Pasadena on NBC.

Making the new color format fully compatible with existing B&W receivers required some clever engineering. First, the color information (C) had to be encoded entirely within the B&W (Y) signal. This was done was by taking the "color difference components" (R-Y, B-Y) and using them to modulate a 3.58 Mhz subcarrier (see our diagram). The frequency of this subcarrier was specially chosen because it minimized interference with the B&W signal. However, in order to assure minimal interference, the color signal bandwidth had to be trimmed down dramatically. Down to as little as half a Megahertz of bandwidth, resulting in just over 40 lines of color resolution across the entire screen(!). This is the reason why the color details transmitted in NTSC composite video are not nearly as sharp as the B&W ones, and why some

SONY COMPONENT VIDEO INPUTS

Sony VPH-400Q LCD Video Projector:

Y, R-Y, B-Y: Y= 1Vp-p, R-Y=.7Vp-p, B-Y=.7Vp-p

Y, Pr, Pb: Y= 1Vp-p (Trilevel /Bilevel Sync .3Vp-p), Pr=..35Vp-p, Pb=.35Vp-p

GBR (for the 1125/60 studio format) : G= 1Vp-p (Trilevel /Bilevel Sync .3Vp-p), B=..7Vp-p, R=.7Vp-p

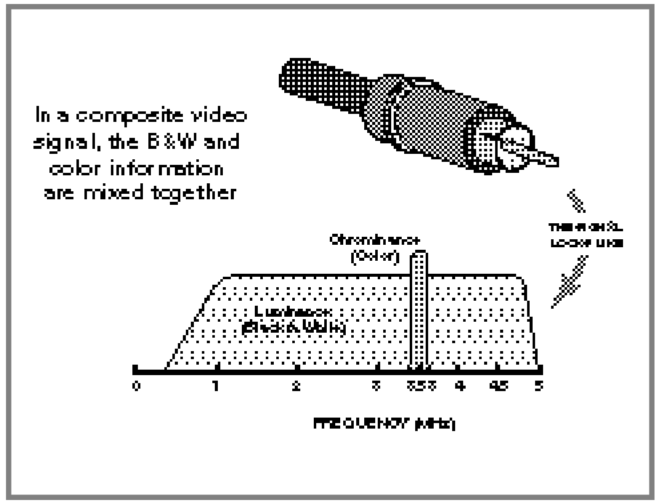
Sony VPH-D50Q CRT Video Projector:

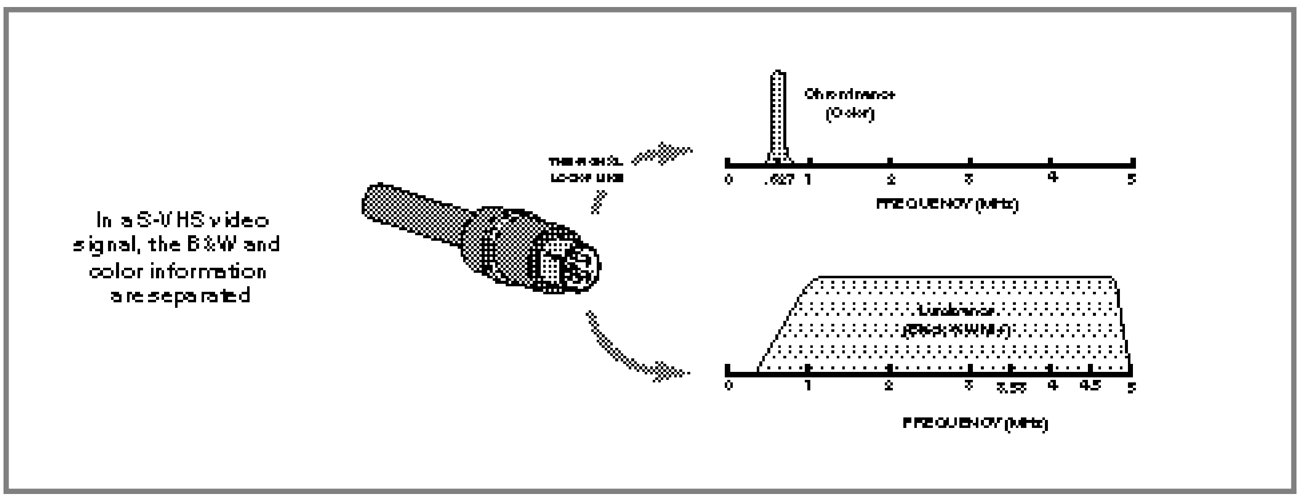
Y, R-Y, B-Y: Y= 1Vp-p, R-Y=.7Vp-p, B-Y=.7Vp-p

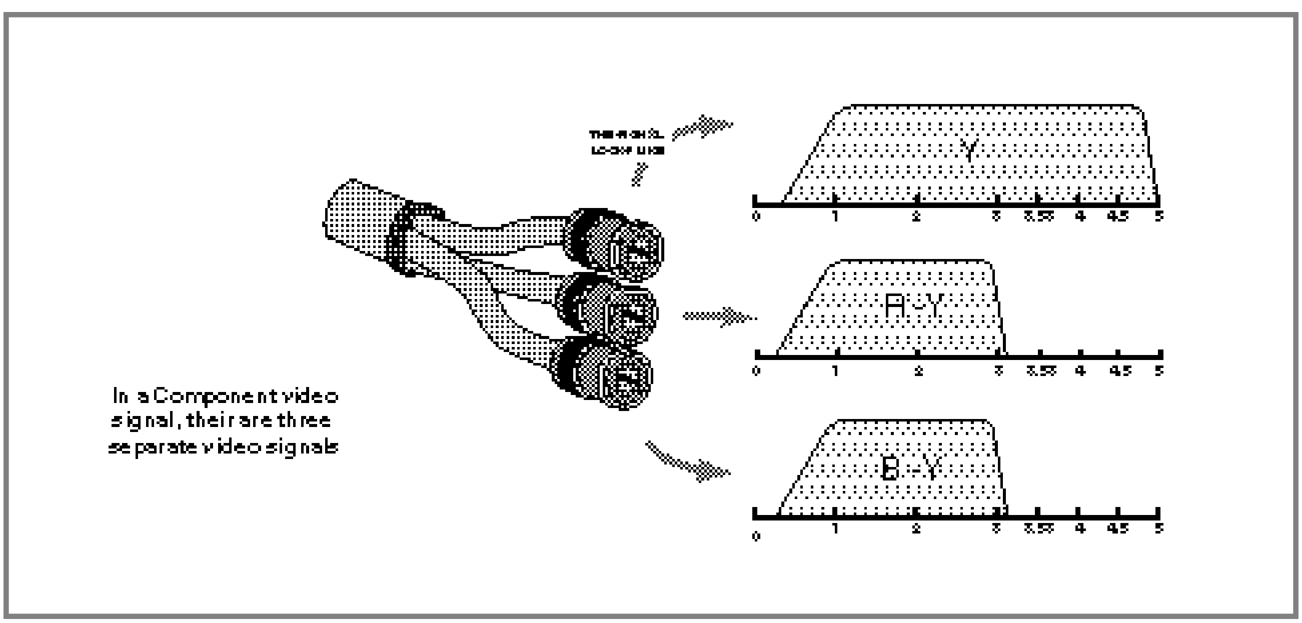
Y, Pr, Pb: Y= 1Vp-p (Trilevel /Bilevel Sync .3Vp-p), Pr=..35Vp-p, Pb=.35Vp-p

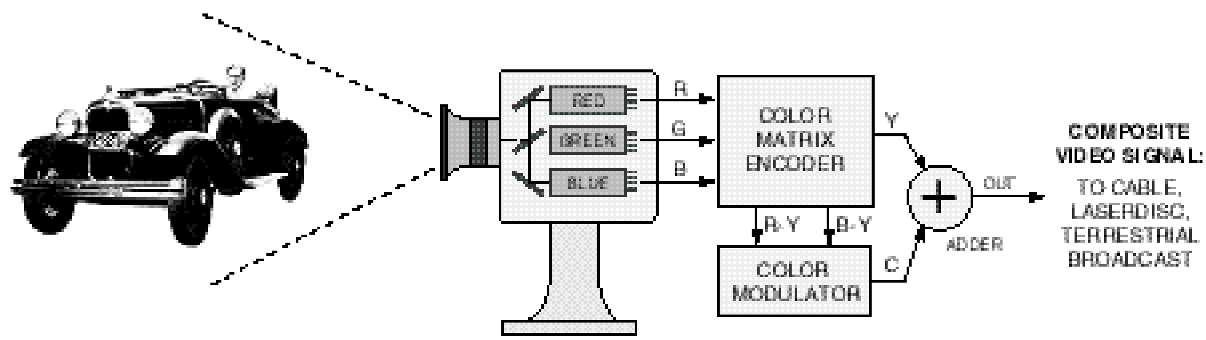
GBR (for the 1125/60 studio format) : G= 1Vp-p (Trilevel /Bilevel Sync .3Vp-p), B=..7Vp-p, R=.7Vp-p colors seem to bleed severely in some formats (VHS tapes, in particular). The color portion of the signal simply lacks the resolution necessary to form sharp edges.

Other artifacts occur when NTSC composite video is decoded. Because the color information and the B&W information were mixed together for transmission; they must be separated at the television receiver for display. This is usually done via electronic filters. The problem is that electronic filters are far from perfect devices and some remnants from one signal often remains in the other after separation. In other words, some of the color information remains in the B&W signal, and some of the B&W information remains in the color signal. These additional tidbits of information often fool the decoding circuitry causing odd things occur. For example: in fine









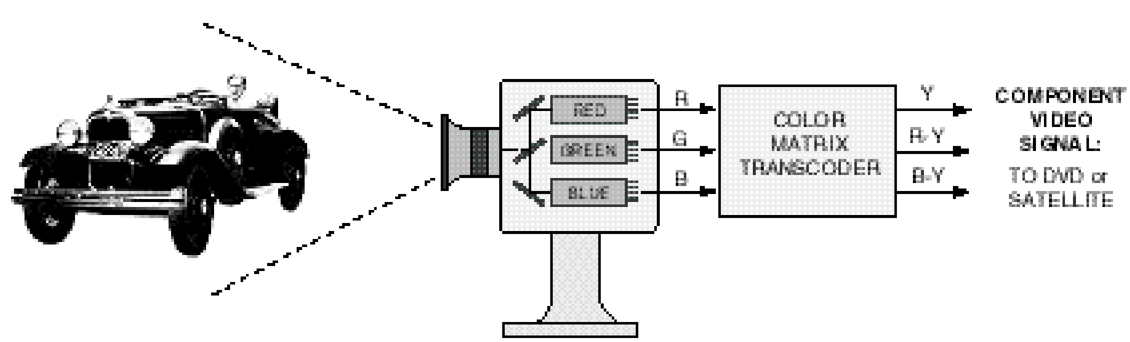
HOW A VIDEOCAMERA CREATES A COMPOSITE VIDEO SIGNAL

detailed areas of the image, the B&W details trick the color circuitry and rippling color rainbows result (Engineers call this cross-color interference.) A reverse effect occurs when the color subcarrier remains in the B&W information. This manifests itself in running dots on the edges of large colored areas (cross-luminance interference). Today, we have sophisticated 2D and 3D digital comb filters, which do a much better job of separating the color and B&W signals than older ordinary comb filters, but these are expensive and they are still not perfect in operation.

Enter Component Video

Instead of stuffing all three color signals into a space that was designed for just one B&W one, why not keep them separate and discrete right from the start? That's exactly the premise behind component video. All three of the color signals are kept separate, and are distributed to display devices that way. This technique has been used in professional broadcast studios for years.

Component video is now available from DVD players, DBS satellite and DTV decoders. So, you can count on component video to breath new life into NTSC format. The elimination of composite encoding/decoding artifacts alone yields a significant improvement in picture quality. Combine that with a much higher color signal bandwidth and you've got a pretty smashing picture. It is especially exciting for large screen projection televisions



HOW A VIDEOCAMERA CREATES A COMPONENT VIDEO SIGNAL

Popular Component Video Formats For Professional Use

Format	Tape Width	Tape Type	Signal Type	Luminance S/N Ratio	Chrome S/N Ratio	Horizontal Resolution	Total Number of Audio Tracks
VHS	1/2 in.	Oxide	Composite	46 dB	45 dB	240 lines	1 linear; 1 stereo AFM (Hi-Fi)
8mm	8mm	Metal Particle	Composite	46 dB	45 dB	260 lines	1 stereo AFM; 1 stereo PCM
S-VHS	1/2 in.	Oxide	Y/C	47 dB	46 dB	400 lines	1 linear; 1 AFM (Hi-Fi)
Hi-8	8mm	Metal Particle	Y/C	47 dB	46 dB	400 lines	1 stereo AFM; 1 stereo PCM
3/4 in. U-matic	3/4 in.	Oxide	Composite	46 dB	46 dB	280 lines	2 linear
3/4 in. SP	3/4 in.	Metal Particle	Composite	47 dB	48 dB	340 lines	2 linear
Betacam	1/2 in.	Oxide	Component Analog	48 dB	50 dB	300 lines	2 linear
Betacam SP	1/2 in.	Metal Particle	Component Analog	51 dB	53 dB	344 lines	2 linear; 2 AFM
M-II	1/2 in.	Metal Evaporated	Component Analog	52 dB	50 dB	344 lines	2 linear, 2 AFM
1-inch Type C	1 in.	Oxide	Composite	52 dB	52 dB	360 lines	3 linear
D-1	3/4 in.	Metal Particle	Component Digital	N/A	N/A	460 lines	4 digital, 48kHz
D-2	3/4 in.	Metal Particle	Composite Digital	N/A	N/A	450 lines	(4) 20-bit
D-3	1/2 in.	Metal Particle	Composite Digital	N/A	N/A	450 lines	4 digital, 48kHz
D-5	1/2 in.	Metal Particle	Component Digital	N/A	N/A	N/A	4 digital, 48kHz
Digital Betacam	1/2 in.	Metal Particle	Component Digital	N/A	N/A	< 500 lines	(4) 20-bit
DV or DVC	6mm	Metal Evaporated	Component Digital	N/A	N/A	< 500 lines	(2) 16-bit; (4) 12-bit PCM
DVCPRO (D-7)	6mm	Metal Particle	Component Digital	N/A	N/A	< 500 lines	(2) 16-bit; (4) 12-bit PCM
DVCAM	6mm	Metal Evaporated	Component Digital	N/A	N/A	< 500 lines	(2) 16-bit; (4) 12-bit PCM
Digital S	1/2 in.	Metal Particle	Component Digital	N/A	N/A	< 500 lines	(4) 16-bit 48kHz PCM
Betacam SX	1/2 in.	Metal Particle	Component Digital	N/A	N/A	< 500 lines	(4) 20-bit

VHS - VHS was originally designed for home use where it still remains dominant. Standard VHS is commonly used in non-broadcast facilities and for offline screening. The format is terrible for television production because of its low resolution, extremely limited dubbing capabilities, and poor color reproduction.

8mm - The original 8mm camcorders were marketed by Kodak in the 1980s. The format's resolution is a little better than VHS. Although 8mm is not used for TV production because of its low quality, it uses a high quality metal tape making it suitable for other formats.

Hi-8 - Introduced in 1989 by Sony, the Hi-8 format was developed specifically for the "prosumer" in mind. It was originally designed as an acquisition format, but was adopted into editing. Sony expected users to bump up footage recorded on Hi-8 to Betacam SP or 1-inch for editing. It has grown popular among professionals who wanted a solid format, but were unable to fork up the big bucks for Betacam or M-II.

S-VHS - S-VHS evolved from VHS. The format has a much higher resolution compared to VHS. Therefore, S-VHS can survive multiple dubbings without being degraded. S-VHS is an appealing format to the "prosumer" market because S-VHS machines can play back VHS tapes. However, S-VHS (as well as Hi-8's) S/N ratio is only slightly better than their "lower-end" counterparts. Therefore, videographers who want to take advantage of S-VHS and Hi-8 should also invest in high quality, industrial cameras that can produce a strong signal (one that exceeds the maximum S/N ratio of the video format).

3/4 inch and 3/4 inch SP - The 3/4 inch format was invented in the early 1970s and was very popular throughout that decade as well as most of the 1980s. It has been replaced by Betacam SP and M-II as broadcast formats. The major downside to the format is the size of the tape (larger than 1/2 inch and 8mm formats) and its quality. However, it is still being used in cable facilities and post production houses. The SP format is an enhancement to the original 3/4 inch. SPhas better resolution as well as a better S/N ratio.

Betacam and Betacam SP - Betacam SP, introduced by Sony, has become the most popular video format used in the professional and broadcast industry. The video is processed componently meaning that the color and brightness information are recorded on separate video tracks. The format provides outstanding color reproduction as well as resolution. Because the video is processed componently, additional generations for layering/effects are possible . Compared to the original Betacam format (introduced early 1980s), Betacam SP uses higher quality metal tape and delivers hi-fi sound.

M-II - The M-II format was designed by Panasonic for NHK, Japan's major television network. Although it does not have the same market share as Betacam SP, M-II is both cheaper and has similar quality. Compared to Betacam SP, M-II has a slower recording speed but uses better tape (metal evaporated compared to metal particle).

1-inch Type C - 1-inch Type C was one of the first video formats introduced into the industry. It has been used greatly throughout the 70s, 80s, and 90s. The 1-inch format has outstanding slow-motion and freeze-frame capabilities since a complete video field is written on every head scan of the tape. However, because of the introduction of better component analog and digital formats, it is no longer used for day to day broadcasts. Instead, 1-inch is mostly used in sporting events for playing back slow-motion instant replays. It is also used for broadcasting movies.

D-1 - D-1 is a component-digital format introduced by Sony in 1986. The width of the tape is 3/4 of an inch, with the resolution of about 460 lines. It is considered as a quality reference because of its supreme quality. However, D-1 recorders and tape stock are very expensive and extremely impractical for industry use. (Some D-1 recorders cost an excess of $100,000).

D-2 - D-2 was developed by Ampex around the same time that D-1 was introduced. It is a composite-digital format, meaning a composite signal, instead of a component signal is recorded. The width of the tape is 3/4 of an inch and the resolution is about 450 lines. Again the format is superseded by other formats because of tape cost, size, and impracticality.

D-3 - D-3 was introduced by Panasonic in 1991. It has been considered as one of the first successful digital formats in the industry. The tape width is 1/2 inch making it possible to build D-3 camcorders. D-3 was used in the Barcelona Olympic games in 1992 and in the 1996 Atlantic Olympic games. (Panasonic was the official video equipment sponsor for both games.)

D-5 - D-5 was introduced in 1993-1994 by Panasonic. It is a component-digital format meaning that the overall picture quality is better than the older D-3. D-5 has also been successful because of its ability to play back D-3 tapes. Currently, NBC, NHK (Japan), and PBS are the big networks using D-5.

Digital Betacam - Digital Betacam was introduced by Sony in 1993 as a successor to Betacam SP. It is a component digital format using 10-bit 4:2:2 sampling. The format has been popular in film transfer because of its excellent quality and its ability to record video with a 16:9 aspect ratio.

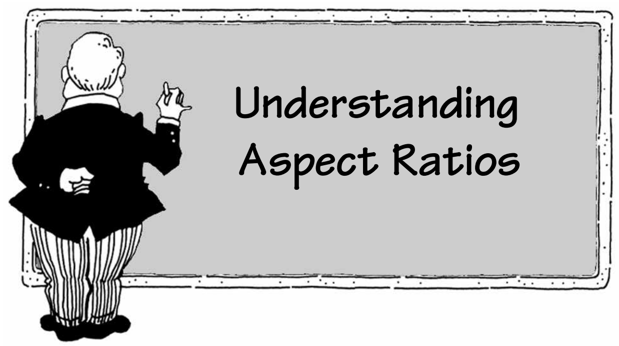
DV or DVC - This new format introduced in 1995 is the first major, high quality video format to be introduced into the consumer market. The format uses a 5:1 compression, M-JPEG algorithm. Some popular camcorders that utilize the DV format include the Sony VX-1000 and Canon XL-1.

DVCPRO - The DVCPRO format was introduced by Panasonic simultaneously when the regular DV format was introduced. Panasonic has pushed the marketing for DVCPRO since it is much more affordable and possesses a quality, meeting or exceeding Betacam SP. DVCPRO is different from the regular DV format because of increased tape speed and wider track pitch. DVCPRO also uses metal particle tape compared to the metal evaporated used on regular DV.

DVCAM - DVCAM was introduced by Sony as their professional DV format. The DVCAM recording format incorporates a higher tape speed compared to regular DV, but it is slower than DVCPRO. To compensate for the slower tape speed, DVCAM uses metal evaporated tape.

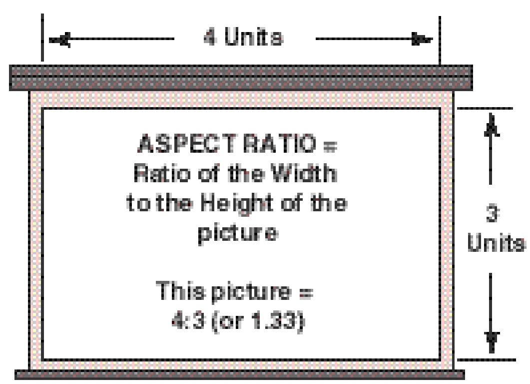
Digital S - Digital S was a format created by JVC. Compared to DV, DVCPRO, and DVCAM, Digital S has two advantages: (1) it uses 4:2:2 sampling to record digital video (like D-1), (2) Digital S VTRs can playback S-VHS tapes. JVC claims that the Digital S format is more robust than DVC,DVCPRO, and DVCAM. Technically, Digital S is better than the DV formats which only use 4:1:1 sampling. As a result, DV does not produce sharp chroma keys. However 4:2:2 allows better color sampling and hence better keys. If tape size contributes to "robustness", then JVC takes the cake, because the format uses 1/2 inch tapes looking similar to S-VHS tapes. In addition, Digital-S is the only deck in the industry that has pre-read capabilities (the ability to record and playback at the same point on the tape track - useful for A/B rolling with only two decks) in the same price class as a high-end Beta SP deck. Currently, the FOX network and its affiliates have begun using Digital S.

Betacam SX - Betacam SX was developed by Sony and introduced in 1996. When Digital Betacam was introduced in 1993, Sony believed that it would replace Betacam SPas a new digital video format. Because of forbidding tape costs, Digital Betacam was not accepted as a successor for Beta SP. As the years progressed and with the introduction of new digital formats, Sony took another stab at introducing a successor for Beta SP. Betacam SX, unlike the DV formats, uses 4:4:2 MPEG 2 sampling and 10:1 compression making the image quality close to Digital Betacam. Unlike Digital Betacam, Betacam SX allows the videomaker to playback and record on analog Betacam SP cassettes. (However, the deck can only record the digital signal on the analog cassettes.) Sony also claims that Betacam SX equipment costs much less to buy and run than analog Beta SP.



The first thing we want to do is demystify this phrase. An aspect ratio is simply a numerical way of describing a rectangular shape. The aspect ratio of your standard television, for example, is 4:3. This means that the picture is 4 “units” wide and 3 “units” high. Interestingly, professional cinematographers tend to prefer a single number to describe screen shapes and reduce the familiar 4:3 television ratio down to 1.33:1, or just 1.33. This is most likely because they deal with a vastly larger number of screen shapes than television people do and out of necessity, long ago, jettisoned bulky fractional descriptions.

The History Of Cinema Aspect Ratios



The original aspect ratio utilized by the motion picture industry was 4:3 and according to historical accounts, was decided in the late 19th century by Thomas Edison while he was working with one of his chief assistants, William L.K. Dickson. As the story goes, Dickson was working with a new 70MM celluloid-based film stock supplied by photographic entrepreneur George Eastman. Because the 70MM format was considered unnecessarily wasteful by Edison, he asked Dickson to cut it down into smaller strips. When Dickson asked Edison what shape he wanted imaged on these strips, Edison replied, "about like this" and held his fingers apart in the shape of a rectangle with approximately a 4:3 aspect ratio. Over the years there has been quite a bit of conjecture about what Edison had in mind when he dictated this shape. Theories vary from from Euclid's famous Greek "Golden Section", a shape of approximately 1.6 to 1, to a shape that simply saved money by cutting the existing 70MM Eastman film stock in half. Whatever the true story may be, Edison's 4:3 aspect ratio was officially adopted in 1917 by the Society Of Motion Picture Engineers as their first engineering standard, and the film industry used it almost exclusively for the next 35 years.

Because of the early precedent set by the motion picture industry with the 4:3 aspect ratio, the television industry adopted the same when television broadcasting began in the 1930s, and today the 4:3 aspect ratio is still the standard for virtually all television monitor and receiver designs. The same situation applies to video programming and software. Only until recently has there been any software available except in 4:3 format (letterboxed videos are the same thing electronically). There simply wasn't any reason to shoot or transfer in any other aspect ratio because of the standard 4:3 shape of the television displays. For the home theater owner, this situation means that compatibility issues are essentially nonexistent with standard 4:3 television receivers and standard 4:3 programming. They are all "plug and play", so to speak, at least when it comes to the shape of image.

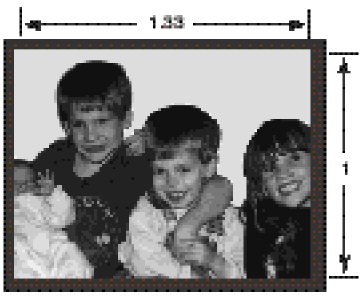
Getting Wide

Back to our history lesson. After many years of experimentation, television broadcasting formally began on April 30, 1939 when NBC broadcasted Franklin Roosevelt's opening of the 1939 World's Fair. As you might imagine, the availability of a device that delivered sound and pictures in the home immediately concerned the Hollywood studios. After all, this medium had the potential to erode their lifeblood; their vital paying customer base. When color was introduced in late 1953, the studios stopped wringing their hands and sprang into action. The result was the rapid development of a multitude of new widescreen projection ratios and several multichannel sound formats. Today, just a few of these widescreen formats survive, but a permanent parting of the ways had occurred: film was now a wide aspect ratio medium, and television remained at the academy standard 4:3 aspect ratio.

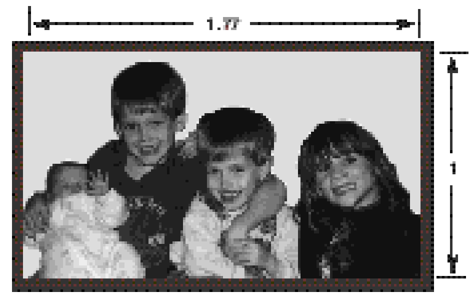
As we mentioned, the fact that film formats went “wide” in the 1950s never really impacted the production end of television. Everything stayed at 4:3 for them because of the uniformity of 4:3 television design. However, the transfer of motion pictures to video...that was another story. The question is: How do you make a wide shape fit into a narrow one? One way you've undoubtedly heard about "panning and scanning". This technique of transferring film to video requires that a telecine (video transfer) operator crop a smaller 4:3 section out of a widescreen movie while panning around following the movie's action. This technique, when properly done, actually works pretty well, but not everyone likes the artistic compromise of "throwing away part of the director's vision". Not the least of which is the film directors themselves, and one of the first to really object to this process was Woody Allen. In 1979, when his film Manhattan was being transferred for television release, he steadfastly refused to have it panned and scanned. He insisted that the feature be shown with the widescreen aspect ratio intact, and this lead to the technique of "letterboxing". Letterboxing, a method where the middle of a 4:3 image is filled with a smaller, but wider, aspect ratio image, may have had the blessing of Hollywood directors but was originally shunned by the viewing public. The objection was the black bars on the top and the bottom of the picture, people just didn't like them. Today, letterboxing has gained much broader acceptance and you can find it available from sources such as prerecorded tapes (occasionally), broadcast television (occasionally), on cable and DSS (AMC and other movie channels broadcast in letterbox), on laserdisc (fairly common), and DVD releases (very common).

So, what about displaying letterbox material with a projection display? On a standard 4:3 display, the situation is pretty simple, letterboxed software can be seen basically one way: as a stripe across the center of the display with black bars top and bottom. On a widescreen display, you can do something different. The letterbox section of the frame can be "zoomed into" so that the image fills the wider screen essentially eliminating the black bars. What is interesting about this technique is that it is conceptually similar to what is done in professional cinemas with standard widescreen releases with "matting". Our diagrams following at the end of this chapter illustrate this. By zooming the letterbox section in to fill the screen, the audience simply sees a widescreen image. The main difference between video display and film display, however, is the way the zooming is done. In a movie theater, an optical zoom lens is used. In a CRT-based video display, it is done by increasing the size of the picture electronically in the picture tube, but with an LCD/DLP-based device it is again done with an optical lens. (Note: some solid state projectors does "zoom" electronically, the SONY VPL-VW10HT is one.)

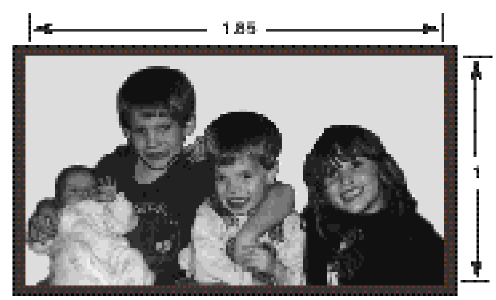
Are there any drawbacks to letterboxing on a 4:3 display as a general technique ? As we mentioned, with the right equipment, letterboxed software can be zoomed to fill a wide screen, but you should know that this comes at a certain price. The issue is loss of vertical resolution. Let's take a matted widescreen film frame, as an example. There is finite amount of resolution in a 35MM frame and, unfortunately, a great deal is taken up with matting. In video, the same principle applies. In a standard video frame there is some 480 lines of vertical resolution



• Television = 1.33 (4:3) Aspect Ratio



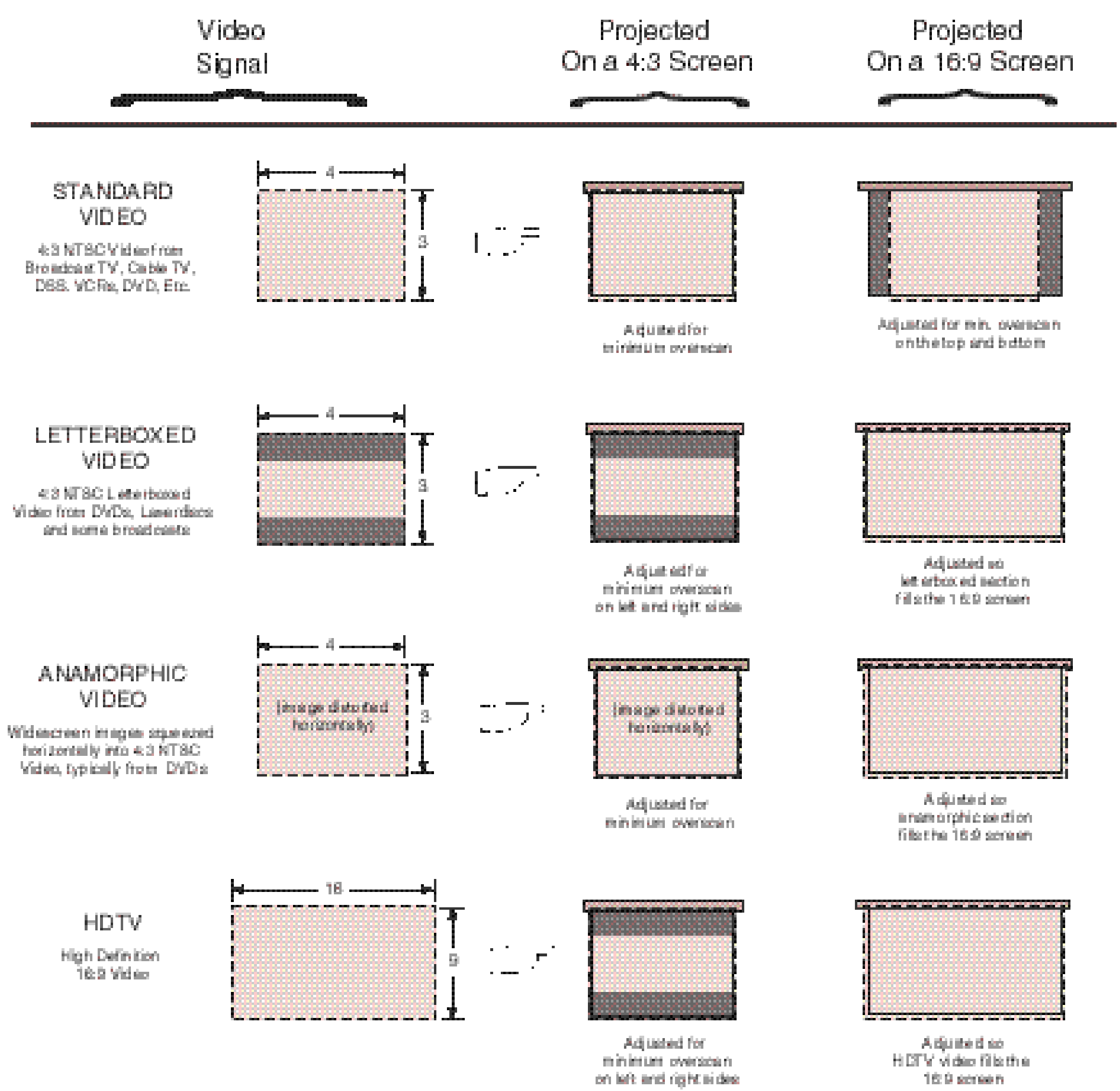
• HDTV, Many Letterboxed movies = 1.77 (16:9) Aspect ratio



• Standard Cinema Widescreen = 1.85 Aspect Ratio

• CinemaScope Widescreen = 2.35 Aspect Ratio

The Four Most Common Aspect Ratios



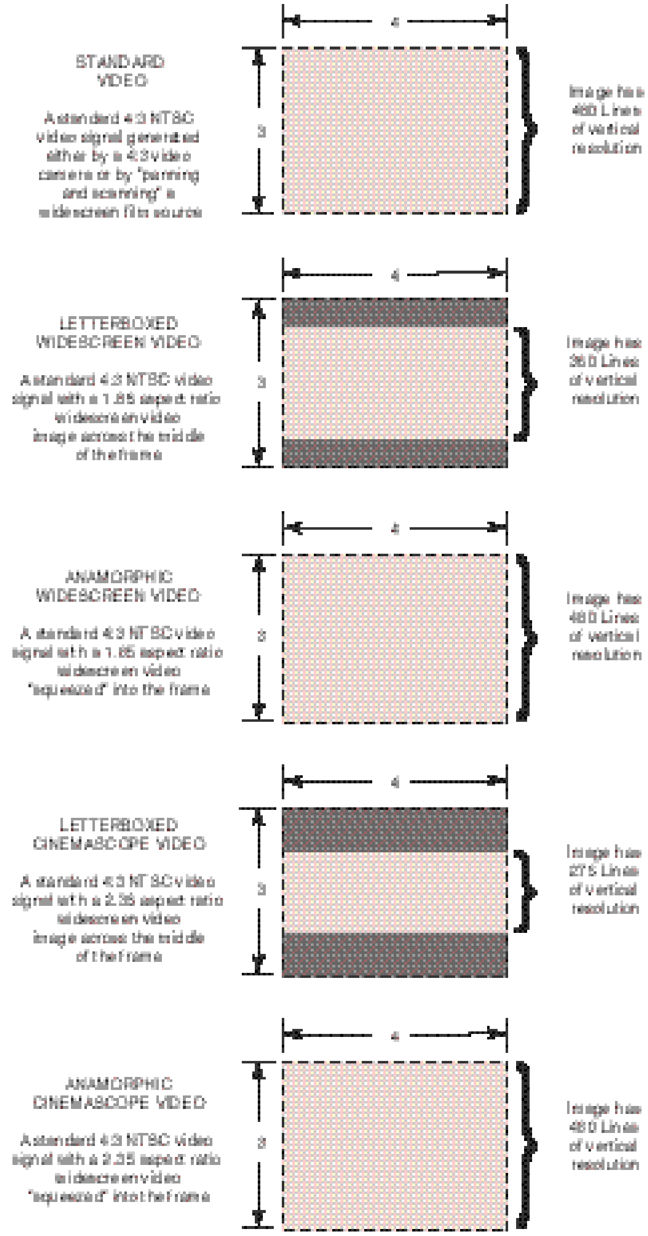
Different Video Standards Displayed on 4:3 and 16:9 Front Projection

available and in the letterbox section, this number is reduced to about 320 to 360 lines (depending on the degree of letterboxing). True, this doesn't have to affect the size of the letterbox section, depending on the size of your television, you have it as wide as what your display allows. However, regardless of the size of your display, the resolution will be less than a full video frame's capability.

A Bit Of A Stretch

Back in the 1950s, the people at CinemaScope came up with a novel solution to the resolution problem outlined above. The solution was to optically squeeze a full widescreen image into a 4:3 film frame via a special device called an "anamorphic lens". The genius of this idea was that no major change was necessary in the camera equipment or the theater projection equipment, all that was necessary was to place an anamorphic lens on the filming cameras to squeeze the image, and a reverse one in the theaters to unsqueeze it. At first, it was said, the Hollywood film community didn't care much for this odd technique, but after using it awhile embraced it hardily. The reason: it was an undeniably elegant solution to the problem of producing and delivering widescreen movies with equipment basically designed for 4:3 format. What is particularly interesting about this 40 year old technique is that a similar concept is now being applied to widescreen electronic video releases. As we mentioned before, the black bars in a letterbox video release also represent lost resolution, just like in the cinema, and the letterbox section is thus lower resolution. Again our concept of anamorphic compression can be used to squeeze more picture into a 4:3 space, but instead of lens, it is done electronically. Some of the first anamorphic video programs were pressed on laserdiscs but with the DVD format, the concept is catching on big.

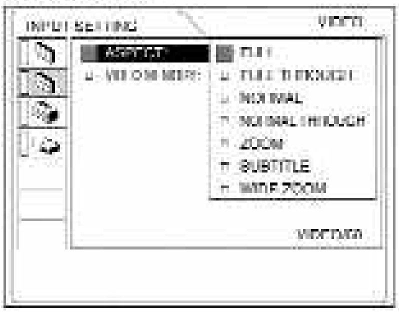
Displaying anamorphic images in a home theater requires a display device with the capability of stretching out the anamorphic image horizontally. Most CRT-based projectors with digital convergence and picture memories (typically graphics-grade projectors) should be able to unsqueeze anamorphic material. With LCD/DLP-based front projectors, the situation concerning anamorphic software is more clear cut than CRT projectors. Most *do not* unsqueeze anamorphic material because picture size changes are accomplished optically via a zoom lens. One LCD projector that we know of that does unsqueeze anamorphic material is the SONY VPL-VW10HT. It has a "full" mode that is designed specifically for this.



Displayed Resolution Of Different Video

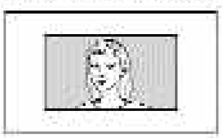
ASPECT RATIO FLEXIBILITY: Switching picture modes with the Sony VPL-VW10HT LCD Projector

4:3 NORMAL, NORMAL THROUGH
16:9 FULL, FULL THROUGH, ZOOM, SUBTITLE, WIDE ZOOM.



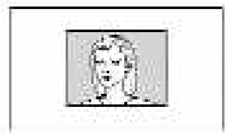
FULL: The 16:9 squeezed image is displayed with the correct aspect. The 4:3 image is enlarged horizontally.

FULL THROUGH: One-to-one mapping is done on a squeezed 16:9 image. The image is displayed at the center of the screen.



NORMAL: The picture with normal ratio 4:3 is displayed.

NORMAL THROUGH: One-to-one mapping is done on the picture with a normal ratio of 4:3. The picture is displayed at the center of the screen.



ZOOM: The picture with normal ratio 4:3 is enlarged vertically and horizontally (with same ratio) to the screen size. This mode is ideal for capturing the full-screen drama of wide format movies.

SUBTITLE: The subtitle area is compressed. This mode leaves the subtitles on the lower part of the screen.

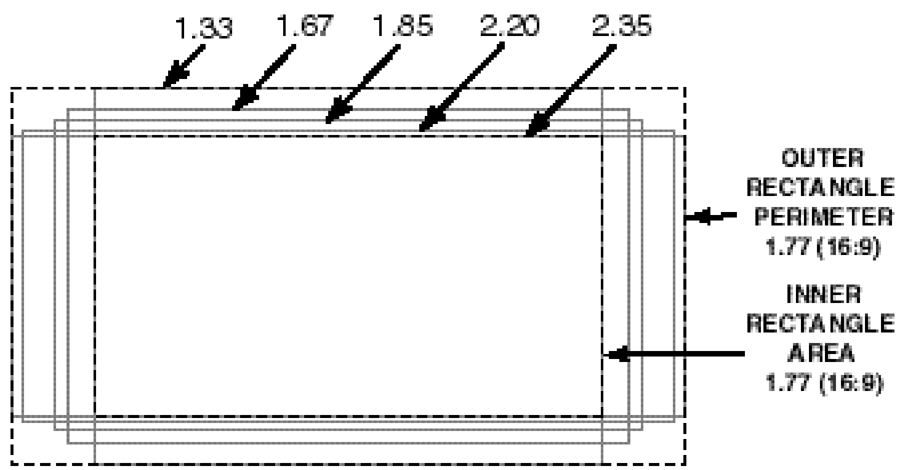
WIDE ZOOM: The picture with normal ratio 4:3 is enlarged and the upper and lower portions of the picture are compressed. This is ideal for general programs, such as news or variety shows.

The Sony VPL-VW10HT LCD projector has 16:9 imaging panels and is capable of displaying images in 16:9 or 4:3 aspect ratios. Sony built in a number of picture expansion modes to allow the user to expand 4:3 images into 16:9. They assumed that most owners woould be using the projector on a 16:9 screen. One of the most useful modes is the FULL mode which allows one to expand anamorphic DVDs.

The Father Of 16:9

The most prevalent aspect ratios filmmakers deal with today are: 1.33 (The Academy standard aspect ratio), 1.67 (The European widescreen aspect ratio), 1.85 (The American widescreen aspect ratio), 2.20 (Panavision), and 2.35 (CinemaScope). Attentive videophiles may note that 1.77 (16:9) isn't on this list and may ask: "If 16:9 isn't a film format, then just exactly where did this ratio come from". The answer to this question is: "Kerns Powers".

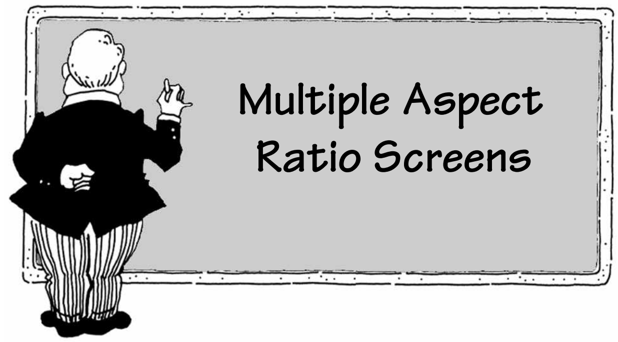
The story begins in the early 1980s when the issue of high definition video as a replacement for film in movie theaters first began to arise. During this time, the Society Of Motion Picture And Television Engineers (SMPTE) formed a committee, the Working Group On High-Definition Electronic Production, to look into standards for this emerging technology. Kerns H. Powers was then research manager for the Television Communications Division at the David Sarnoff Research Center. As a prominent member of the television industry, he was asked to join the working group, and immediately became embroiled in the issue of aspect ratios and HDTV. The problem was simple to define. The film community for decades has been used to the flexibility of many aspect ratios, but the television community had just one. Obviously a compromise was needed.



KERNS POWERS' SOLUTION

As the story goes, using a pencil and a piece of paper, Powers drew the rectangles of all the popular film aspect ratios (each normalized for equal area) and dropped them on top of each other. When he finished, he discovered an amazing thing. Not only did all the rectangles fall within a 1.77 shape, the edges of all the rectangles also fell outside an inner rectangle which also had a 1.77 shape. Powers realized that he had the makings of a "Shoot and Protect" scheme that with the proper masks would permit motion pictures to be released in any aspect ratio. In 1984, this concept was unanimously accepted by the SMPTE working group and soon became the standard for HDTV production worldwide.

Ironically, it should be noted, the High-Definition Electronic Production Committee wasn't looking for a display aspect ratio for HDTV monitors, but that's what the 16:9 ratio is used for today. "It was about the electronic production of movies," Kerns Powers states, "that's where the emphasis was". Interestingly, today, there is little talk today about the extinction of film as a motion picture technology, but there is a lot of talk about delivering HDTV into the home. And, as a testament to Kern H. Powers clever solution, it's all going to be on monitors with a 16:9 aspect ratio.

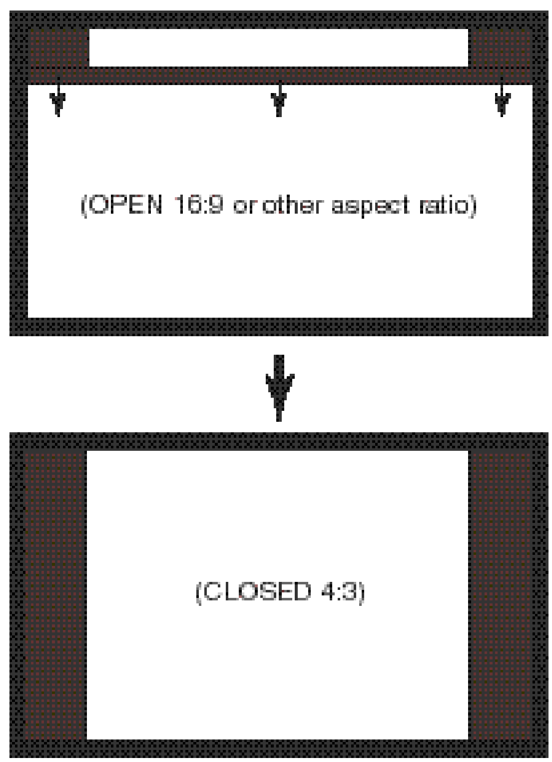


Variable aspect ratio screen systems are a convenient way to add professional looking screen masking to home theater rooms. Each of the products we describe here are available in many sizes and configurations. This page is simply to illustrate the different types of variable aspect ratio screen systems that you can chose from. For further information, visit the manufacturers web sites.

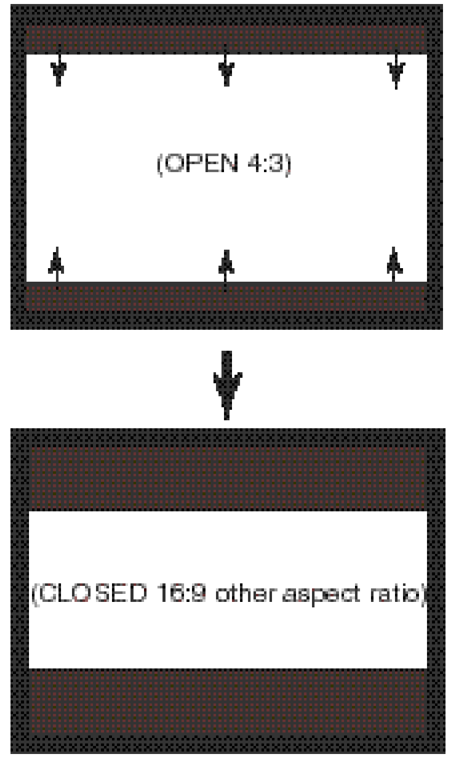
Flat Screen with Motorized Left and Right Masking Panel Assembly

These masking systems consist of a fixed frame assembly that mounts over a stretched flat screen. It has motorized panels that lower on the left and right sides changing a 4:3 screen to a 16:9 (or other) aspect ratio. They are sold under the following brand names:

- DRAPER Eclipse H™ system
- STEWART Vertical Screenwall Electrimask™
- VUTEC Vision XFV™
- DA-LITE Pro Imager Horizontal Masking System



Flat Screen with Motorized Top and Bottom Masking Panel Assembly



These masking systems consist of a fixed frame assembly that mounts over a stretched flat screen. It has motorized panels that lower on the top and bottom changing a 4:3 screen to a 16:9 (or other) aspect ratio.They are sold under the following brand names:

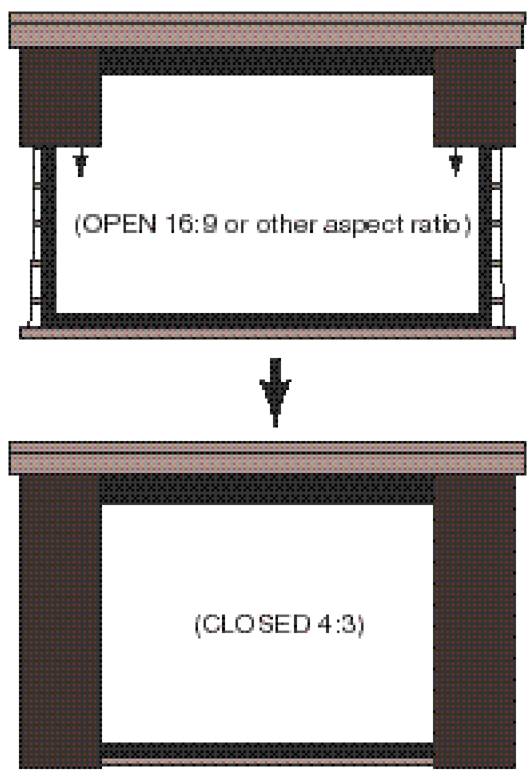
• DRAPER Eclipse V™

• STEWART Horizontal Screenwall Electrimask™

• VUTEC Vision XFH™

• DA-LITE Pro Imager Horizontal Masking System

Electric Roll-Down Screen with Motorized Left and Right Masking Panels



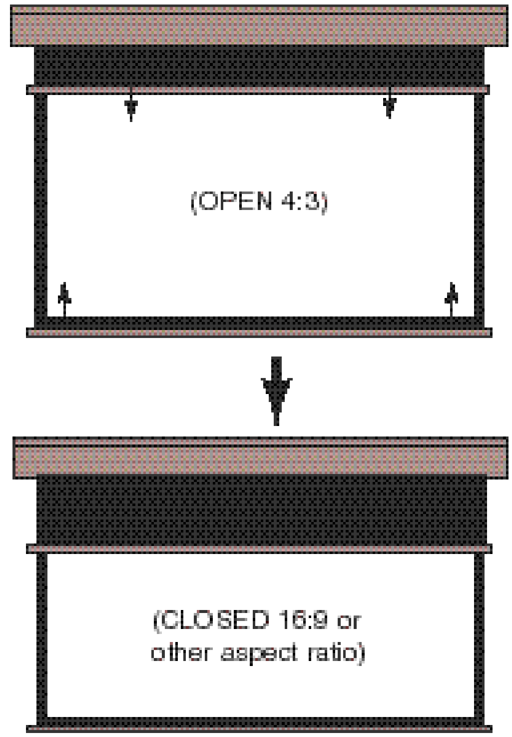
These masking systems consist of a regular rolldown screen assembly with left and right masking panels built into the same housing. When lowered they convert a 16:9 (or other) aspect ratio screen into a 4:3. They are sold under the following brand names:

• STEWART Vertical ElectriScreen ElectriMask™

• DA-LITE Dual Masking Electrol™

• DRAPER Access Multiview™

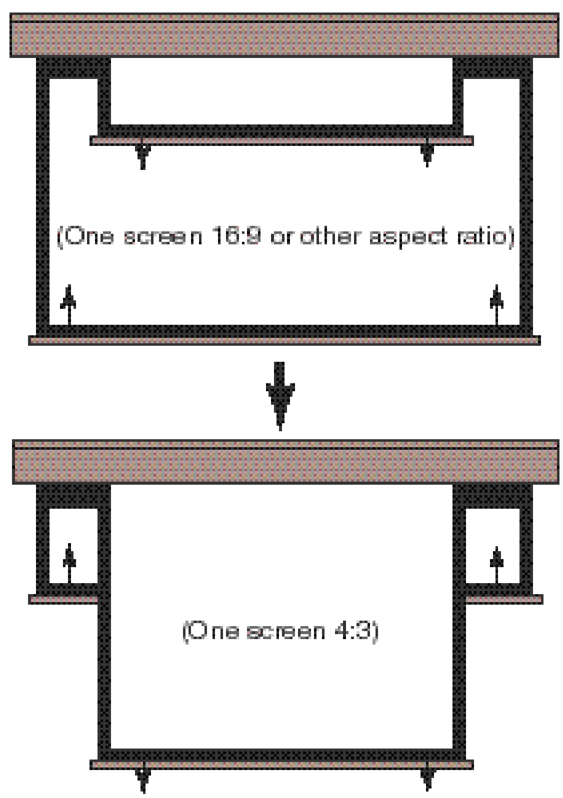
Electric Roll-Down Screen with Motorized Top and Bottom Masking Panels



This system consists of one screen surface (Typically 4:3) and one upper masking panel. The 4:3 surface is lowered for 4:3 sources and when 16:9 sources are viewed, the 4:3 screen moves up several inches and the black upper masking panel rolls down. The result is a 16:9 viewing surface. These screens are sold under the following brand names:

- DA-LITE Horizontal Electrol™
- VUTEC Vision XM™
- DRAPER Access Sonata™

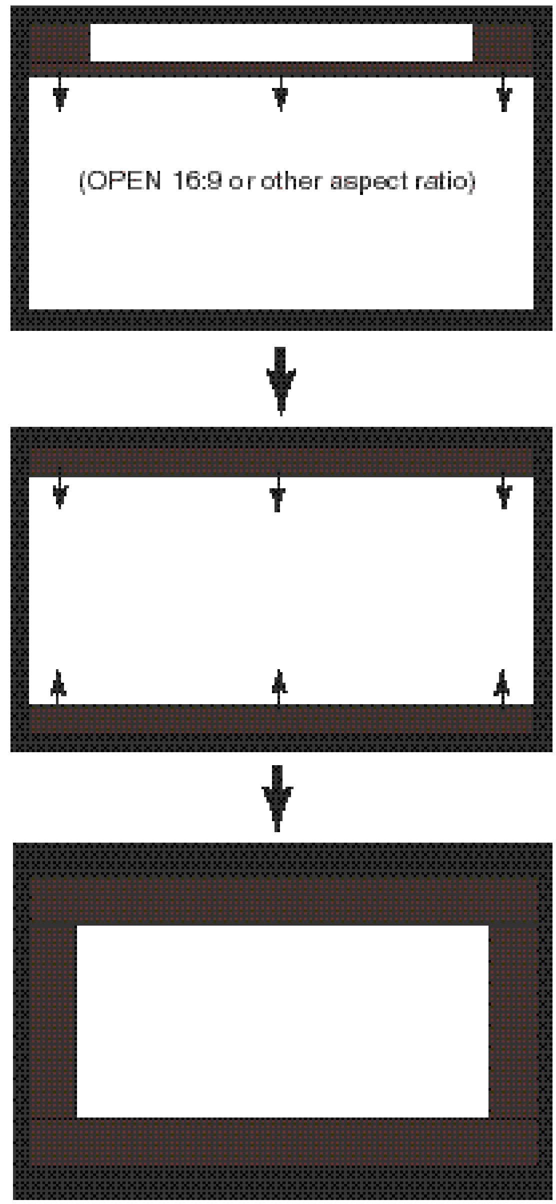
Dual Aspect Ratio Screen Assembly



Offered as VUTEC Vu-Flex Pro Duplex™. This system consists of two separate screen surfaces housed in the same assembly. One surface is used at a time and both roll down in the same plane so image focus is constant. Typically these screens are ordered with a 4:3 surface and a 16:9 (or other ratio) surface.

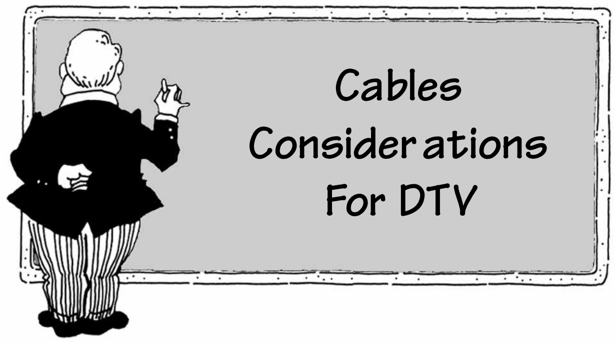
- VUTEC Vu-Flex Pro Duplex™

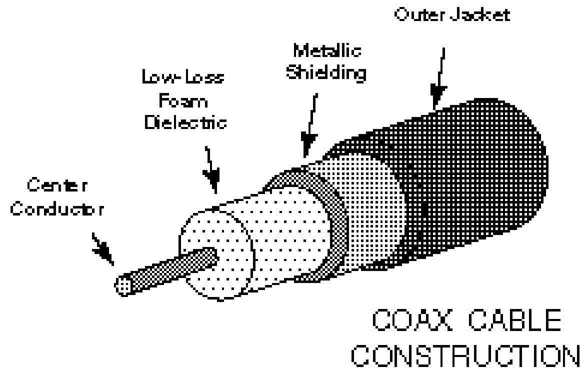
Flat Screen with Motorized Top and Bottom and Left and Right Masking Panel Assembly



These masking systems consist of a fixed frame assembly that mounts over a stretched flat screen. It has motorized panels that lower on the left and right sides changing a 16:9 (or other) aspect ratio screen to a 4:3 .

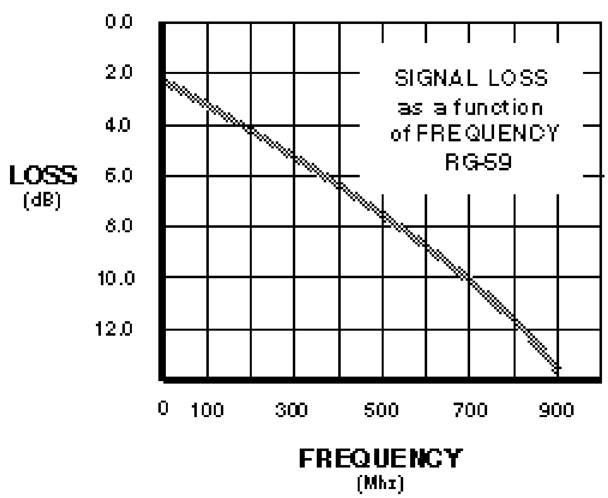
• STEWART Ultimate 4-Way Electrimask-Screenwall™ system





Uncompressed, high definition video signals run at a data rate of 1.485 Gbps and a bandwidth of 750 MHz. It is no surprise, therefore, that cables designed to operate at 4.2 MHz for analog video have a much harder time at 750 MHz. These high frequencies require greater precision and lower loss than analog. Where effective cable distances were thousands of feet for analog, the distance limitations are greatly reduced for HD.

When SMPTE first addressed this problem, they looked at the bit error rate at the output of various cables. Their purpose was to identify the "digital cliff", the point where the signal on a cable goes from "zero" bit errors to unacceptable bit errors. This can occur in as little as 50 feet.



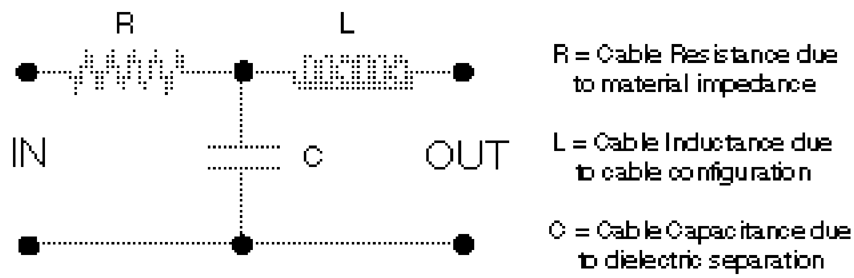
The SMPTE 292M committee cut cables until they established the location of this cliff, cut that distance in half, and measured the level on the cable. From there they came up with the standard: where the signal level has fallen 20 dB, that is as far as your cable can go for HD video. It should be apparent, therefore, that these cables can go up to twice as far as their 'recommended' distance, especially if your receiving device is good at resolving bit errors. Of course, you could look at bit errors yourself, and that would determine whether a particular cable, or series of cables, would work or not.

There is one other way to test HD cable and that is by measuring return loss. Return loss shows a number of cable faults with a single measurement, such as flaws in the design, flaws in the manufacturing, or even errors or mishandling during installation of a cable. Ultimately, return loss shows the variations in impedance in a cable, which lead to signal reflection, which is the "return" in return loss.

A return loss graph can show things as varied as the wrong impedance plugs attached to the cable, or wrong jacks or plugs in a patch panel. It can also reveal abuse during installation, such as stepping on a cable or bending a cable too tightly, or exceeding the pull strength of the cable. Return loss can even reveal manufacturing errors. Broadcasters are familiar with VSWR--Voltage Standing Wave Ratio, which is a cousin to return loss. For instance, SMPTE recommends a return loss of 15 dB up to the third harmonic of 750 MHz (2.25 GHz), this is equivalent to a VSWR of 1.43:1. If you know VSWR, you will recognize this as a very large amount of return. Others have suggested that 15 dB return loss is insufficient to show many circuit flaws.

It is suggested that a two-band approach be taken, since return loss becomes progressively more difficult as frequencies increase. In the band of 5 to 850 MHz, a minimum of 23 dB would be acceptable (equivalent to a VSWR of 1.15:1) and from 850 to 2.25 GHz a minimum 21 dB (equivalent to a VSWR of 1.2:1). Some manufacturers are sweeping cables and showing 21 dB return loss out to 3 GHz, which is even better.

So what cables should you use and what cables should



ELECTRICAL MODEL OF A CABLE

you avoid? Certainly, the standard video RG-59 cables, with solid insulations and single braid shields lack a number of requirements. First their center conductors are often tin-plated to help prevent oxidation and corrosion. While admirable at analog video frequencies, these features can cause severe loss at HD frequencies. Above 50 MHz, the majority of the signal runs along the surface of the conductor, called "skin effect". What you need is a bare copper conductor, since any tinned wire will have that tin right where the high-frequency signal wants to flow. And tin is a poor conductor compared to copper.

Around the conductor is the insulation, called the "dielectric." The performance of the dielectric is indicated by the "velocity of propagation," as listed in manufacturer's catalogs. Older cables use solid polyethylene, with a velocity of propagation of 66 percent. This can easily be surpassed by newer gas-injected foam polyethylene, with velocities in the +80 percent range. The high velocity provides lower high-frequency attenuation.

However, foam is inherently softer than a solid dielectric, so foam dielectrics will allow the center conductors to "migrate" when the cable is bent, or otherwise deformed. This can lead to greater impedance variations, with a resultant increase in return loss. Therefore, it is essential that these foam cables have high-density hard-cell foam. The best of these cables exhibit about double the variation of solid cables (±3‡ foamed versus ±1-1/2‡ solid), but with much better high frequency response.

This is truly cutting-edge technology for cables, and can be easily determined by stripping the jacked and removing the braid and foil from short samples of cables that you are considering. Just squeeze the dielectric of each sample. The high-density hard cell one should be immediately apparent.

Over the dielectric is the shield. Where a single braid was sufficient coverage for analog video, it is not for HD. Older double braid cables have improved shielding, but the ideal is a combination of foil and braid. Foil is superior at high frequencies, since it offers 100 percent coverage at "skin effect" frequencies. Braid is superior at lower frequencies, so a combination is ideal. Braid coverage should be as high as possible. Maximum braid coverage is around 95 percent for a single braid.

The jacket has little effect on the performance of a cable, but a choice of color, and consistency and appearance, will be of concern. There are no standards for color codes (other than red/green/blue indicating RGB-analog video), so you can have any color indicate whatever you want.

From Chapter Nine of
The Guide To Digital Television

Published by
United Entertainment Media
460 Park Avenue South, 9th Floor
New York, NY 10016
(212) 378-0449

For more information:
www.digitaltelevision.com

Chapter Two: Understanding Digital Signals

From Chapter Two of
The Guide To Digital Television

Published by
United Entertainment Media
460 Park Avenue South, 9th Floor
New York, NY 10016
(212) 378-0449

For more information:
www.digitaltelevision.com

In order to understand digital, you must first understand that everything in nature, including the sounds and images you wish to record or transmit, was originally analog. The second thing you must understand is that analog works very well. In fact, because of what analog and digital are, a first-generation analog recording can be a better representation of the original images than a first-generation digital recording. This is because digital is a coded approximation of analog. With enough bandwidth, a first-generation analog VTR can record the more "perfect" copy.

Digital is a binary language represented by zeros (an "off" state) and ones (an "on" state). Because of this, the signal either exists (on) or does not exist (off). Even with low signal power, if the transmitted digital signal is higher that the background noise level, a perfect picture and sound can be obtained--on is on no matter what the signal strength.

The Language Of Digital: Bits & Bytes

Bit is short for Binary digit and is the smallest data unit in a digital system. A bit is a single one or zero. Typically 8-bits make up a byte (although byte "words" can be 10-bit, 16-bit, 24-bit, or 32-bit).

In an 8-bit system there are 256 discrete values. The mathematics is simple: It is the number two (as in binary) raised to the power of the number of bits. In this case 28=256. A 10-bit system has 1,024 discrete values (210=1,024). Notice that each additional bit is a doubling of the number of discrete values.

Here's how this works, as each bit in the 8-bit word represents a distinct value: The more bits, the more distinct the value. For example, a gray-scale can be represented by 1-bit which would give the scale two values (21=2): 0 or 1 (a gray-scale consisting of white and black). Increase the number of bits to two-bits and the gray-scale has four values (22=4): 0, 1, 2, and 3, where 0=0 percent white (black), 1=33 percent white, 2=67 percent white, and 3=100 percent white. As we increase the number of bits, we get more accurate with our gray-scale.

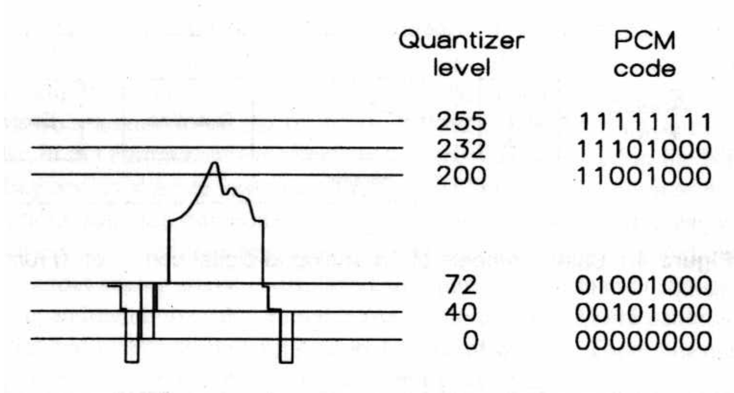
In digital video, black is not at value 0 and white is neither at value 255 for 8-bit nor 1,023 for 10-bit. To add some buffer space and to allow for "superblack" (which is at 0 IRE while regular black is at 7.5 IRE), black is at value 16 while white is at value 235 for 8-bit video. For 10-bit video, we basically multiply the 8-bit numbers by four, yielding black at a value of 64 and white at a value of 940.

Also keep in mind that while digital is an approximation of the analog world--the actual analog value is assigned to its closest digital value--human perception has a hard time recognizing the fact that it is being cheated. While very few expert observers might be able to tell that something didn't look right in 8-bit video, 10-bit video looks perfect to the human eye. But as you'll see in Chapter 4: Audio, human ears are not as forgiving as human eyes--in audio most of us require at least 16-bit resolution--while experts argue that 20-bit, or ultimately even 24-bit technology needs to become standard before we have recordings that match the sensitivity of human hearing.

Digitizing: Analog To Digital

To transform a signal from analog to digital, the analog signal must go through the processes of sampling and quantization. The better the sampling and quantization, the better the digital image will represent the analog image.

Sampling is how often a device (like an analog-to-digital converter) samples a signal. This is usually given in a figure like 48 kHz for audio and 13.5 MHz for video. It is usually at least twice the highest analog signal frequency (known as the Nyquist criteria). The official sampling standard for standard definition television is ITU-R 601 (short for ITU-R BT.601-2, also known as "601").
For television pictures, eight or 10-bits are normally used; for sound, 16 or 20-bits are common, and 24-bits are



Video waveform quantized into 8-bit words.

being introduced. The ITU-R 601 standard defines the sampling of video components based on 13.5 MHz, and AES/EBU defines sampling of 44.1 and 48 kHz for audio. Quantization can occur either before or after the signal has been sampled, but usually after. It is how many levels (bits per sample) the analog signal will have to force itself into. As noted earlier, a 10-bit signal has more levels (resolution) than an 8-bit signal. Errors occur because quantizing a signal results in a digital approximation of that signal.

When Things Go Wrong: The LSB & MSB

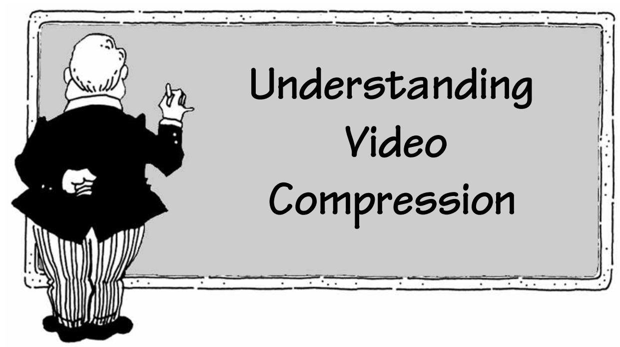
Things always go wrong. Just how wrong is determined by when that "wrongness" occurred and the length of time of that "wrongness." Let's take an 8-bit byte as an example: The "1" on the far right that represents the value 1 is called the least significant bit (LSB). If there is an error that changes this bit from "1" (on) to "0" (off), the value of the byte changes from 163 to 162--a very minor difference. But the error increases as problems occur with bits more towards the left.

The "1" on the left that represents the value 128 is called the most significant bit (MSB). An error that changes this bit from "1" (on) to "0" (off) changes the value of the byte from 163 to 35--a very major difference. If this represented our gray-scale, our sample has changed from 64 percent white to only 14 percent white.

An error can last short enough to not even affect one bit, or long enough to affect a number of bits, entire bytes, or even seconds of video and audio.

If our error from above lasted in duration the amount of time to transmit two bits, the error can be anywhere from minor (if it is the LSB and the bit to its left) to major (if it is the MSB and the bit to its right).

Where and how long errors occur is anyone's guess, but as you'll see below in Error Management, digital gives us a way to handle large errors invisibly to the viewer.



From Chapter Two of
The Guide To Digital Television

Published by
United Entertainment Media
460 Park Avenue South, 9th Floor
New York, NY 10016
(212) 378-0449

For more information:
www.digitaltelevision.com

Some people say that compressing video is a little like making orange juice concentrate or freeze-dried back-packing food. You throw something away (like water) that you think you can replace later. In doing so, you gain significant advantages in storage and transportation and you accept the food-like result because it's priced right and good enough for the application. Unfortunately, while orange juice molecules are all the same, the pixels used in digital video might all be different. Video compression is more like an ad that used to appear in the New York City subway which said something like: "If u cn rd ths, u cn get a gd pying jb" or personalized license plates that don't use vowels (nmbr-1). You understand what the message is without having to receive the entire message--your brain acts as a decoder. Email is taking on this characteristic with words such as l8r (later) and ltns (long time no see).

Why Compress?

There is a quip making the rounds that proclaims "compression has never been shown to improve video quality." It's popular with folks who think compression is a bad compromise. If storage costs are dropping and communication bandwidth is rapidly increasing, they reason, why would we want to bother with anything less than "real" video? Surely compression will fall by the wayside once we've reached digital perfection.

Other people, like Avid Technology VP Eric Peters, contend that compression is integral to the very nature of media. The word "media," he points out, comes from the fact that a technology, a medium, stands between the originator and the recipient of a message. Frequently that message is a representation of the real world. But no matter how much bandwidth we have, we will never be able to transmit all of the richness of reality. There is, he argues, much more detail in any source than can possibly be communicated. Unless the message is very simple, our representation of it will always be an imperfect reduction of the original. Even as we near the limits of our senses (as we may have with frequency response in digital sound) we still find there is a long way to go. People perceive many spatial and other subtle clues in the real world that are distorted or lost in even the best digital stereo recordings.

Furthermore, the notion of quality in any medium is inherently a moving target. We've added color and stereo sound to television. Just as we start to get a handle on compressing standard definition signals, high definition and widescreen loom on the horizon. There will never be enough bandwidth. There is even a Super High Definition format that is 2048x2048 pixels--14 times as large as NTSC. Perhaps former Tektronix design engineer Bruce Penny countered the quip best when he said, "Compression does improve picture quality. It improves the picture you can achieve in the bandwidth you have."

Compression Basics

Compression comes in a number of flavors, each tailored for a specific application or set of applications. An understanding of the compression process will help you decide which compression method or group of methods are right for you.

The essence of all compression is throwing data away. The effectiveness of a compression scheme is indicated by its "compression ratio," which is determined by dividing the amount of data you started with by what's left when you're through. Assuming a high definition camera spits out around one billion video bits a second, and this is ultimately reduced to something around 18 million bits for broadcast in the ATSC system, the compression ratio is roughly 55:1.

However, don't put too much stock in compression ratios alone. On a scale of meaningful measures, they rank down somewhere with promised savings on long distance phone calls. To interpret a compression ratio, you need to know what the starting point was. For a compression system that puts out a 25 megabit per second (Mbps) video stream, the compression ratio would be about 8.5:1

if the starting point was 485x740 pixels, 4:2:2, 10-bit sampled, 30 frames per second (fps) pictures. If, however, the starting video was 480x640, 4:1:1, 8-bit, 30 fps, the ratio would be about 4.5:1.

Lossless Versus Lossy

There are two general types of compression algorithms: lossless and lossy. As the name suggests, a lossless algorithm gives back the original data bit-for-bit on decompression.

One common lossless technique is "run length encoding," in which long runs of the same data value are compressed by transmitting a prearranged code for "string of ones" or "string of zeros" followed by a number for the length of the string. Another lossless scheme is similar to Morse Code, where the most frequently occurring letters have the shortest codes. Huffman or entropy coding computes the probability that certain data values will occur and then assigns short codes to those with the highest probability and longer codes to the ones that don't show up very often. Everyday examples of lossless compression can be found in the Macintosh Stuffit program and WinZip for Windows.

Lossless processes can be applied safely to your checkbook accounting program, but their compression ratios are usually low--on the order of 2:1. In practice these ratios are unpredictable and depend heavily on the type of data in the files. Alas, pictures are not as predictable as text and bank records, and lossless techniques have only limited effectiveness with video. Work continues on lossless video compression. Increased processing power and new algorithms may eventually make it practical, but for now, virtually all video compression is lossy.

Lossy video compression systems use lossless techniques where they can, but the really big savings come from throwing things away. To do this, the image is processed or "transformed" into two groups of data. One group will, ideally, contain all the important information. The other gets all the unimportant information. Only the important stuff needs to be kept and transmitted.

Perceptual Coding

Lossy compression systems take the performance of our eyes into account as they decide what information to place in the important pile and which to discard in the unimportant pile. They throw away things the eye doesn't notice or won't be too upset about losing. Since our perception of fine color details is limited, chroma resolution can be reduced by factors of two, four, eight or more, depending on the application.

Lossy schemes also exploit our lessened ability to see detail immediately after a picture change, on the diagonal or in moving objects. Unfortunately, the latter doesn't yield as much of a savings as one might first think, because we often track moving objects on a screen with our eyes.

Predictive Coding

Video compression also relies heavily on the correlation between adjacent picture elements. If television pictures consisted entirely of randomly valued pixels (noise), compression wouldn't be possible (some music video producers and directors are going to find this out the hard way--as encoders lock-up). Fortunately, adjoining picture elements are a lot like the weather. Tomorrow's weather is very likely to be just like today's, and odds are that nearby pixels in the same or adjacent fields and frames are more likely to be the same than they are to be different. Predictive coding relies on making an estimate of the value of the current pixel based on previous values for that location and other neighboring areas. The rules of the estimating game are stored in the decoder and, for any new pixel, the encoder need only send the difference or error value between what the rules would have predicted and the actual value of the new element. The more accurate the prediction, the less data needs to be sent.

Motion Compensation

The motion of objects or the camera from one frame to the next complicates predictive coding, but it also opens up new compression possibilities. Fortunately, moving objects in the real world are somewhat predictable. They tend to move with inertia and in a continuous fashion. In MPEG, where picture elements are processed in blocks, you can save quite a few bits if you can predict how a given block of pixels has moved from one frame to the next. By sending commands (motion vectors) that simply tell the decoder how to move a block of pixels already in its memory, you avoid resending all the data associated with that block.

Inter- Versus Intra-frame Compression

As long as compressed pictures are only going to be transmitted and viewed, compression encoders can assign lots of bits into the unimportant pile by exploiting the redundancy in successive frames. It's called "inter-frame" coding. If, on the other hand, the video is destined to undergo further processing such as enlargement, rotation and/or chromakey, some of those otherwise unimportant details may suddenly become important, and it may be necessary to spend more bits to accommodate what post production equipment can "see."

To facilitate editing and other post processing, compression schemes intended for post usually confine their efforts within a single frame and are called "intra-frame." It takes more bits, but it's worth it.The Ampex DCT videocassette format, Digital Betacam, D9 (formerly Digital-S), DVCPRO50, and various implementations of Motion-JPEG are examples of post production gear using intra-frame compression. The MPEG 4:2:2 Profile can also be implemented in an intra-frame fashion.

Symmetrical Versus Asymmetrical

Compression systems are described as symmetrical if the complexity (and therefore cost) of their encoders and decoders are similar. This is usually the case with recording and professional point-to-point transmission systems. With point-to-multipoint transmission applications, such as broadcasting or mass program distribution where there are few encoders but millions of decoders, an asymmetrical design may be desirable. By increasing complexity in the encoder, you may be able to significantly reduce complexity in the decoders and thus reduce the cost of the consumer reception or playback device.

Transforms

Transforms manipulate image data in ways that make it easier to separate the important from the unimportant. Three types are currently used for video compression: Wavelets, Fractals, and the Discrete Cosine Transform or DCT.

1) Wavelets--The Wavelet transform employs a succession of mathematical operations that can be thought of as filters that decompose an image into a series of frequency bands. Each band can then be treated differently depending on its visual impact. Since the most visually important information is typically concentrated in the lowest frequencies in the image or in a particular band, they can be coded with more bits than the higher ones. For a given application, data can be reduced by selecting how many bands will be transmitted, how coarsely each will be coded and how much error protection each will receive.The wavelet technique has advantages in that it is computationally simpler than DCT and easily scalable. The same compressed data file can be scaled to different compression ratios simply by discarding some of it prior to transmission.

The study of wavelets has lagged about 10 years behind that of DTC, but it is now the subject of intensive research and development. A Wavelet algorithm has been chosen for coding still images and textures in MPEG-4, and another is the basis for the new JPEG-2000 still image standard for which final approval is expected in 2001 (ISO 15444). More applications are likely in the future.

2) Fractals--The fractal transform is also an intra-frame method. It is based on a set of two dimensional patterns discovered by Benoit Mandelbrot at IBM. The idea is that you can recreate any image simply by selecting patterns from the set and then appropriately sizing, rotating and fitting them into the frame (see figure 1). Rather than transmitting all the data necessary to recreate an image, a fractal coder relies on the pattern set stored in the decoder and sends only information on which patterns to use and how to size and position them.

The fractal transform can achieve very high compression ratios and is used extensively for sending images on the Internet. Unfortunately, the process of analyzing original images requires so much computing power that fractals aren't feasible for realtime video. The technique also has difficulties with hard-edged artificial shapes such as character graphics and buildings. It works best with natural objects like leaves, faces and landscapes.

3) DCT--The discrete cosine transform is by far the most used transform in video compression. It's found in both intra-frame and inter-frame systems, and it's the basis for JPEG, MPEG, DV and the H.xxx videoconferencing standards.

Like wavelets, DCT is based on the theory that the eye is most sensitive to certain two-dimensional frequencies in an image and much less sensitive to others.With DCT, the picture is divided into small blocks, usually 8 pixels by 8 pixels. The DCT algorithm converts the 64 values that represent the amplitude of each of the pixels in a block into 64 new values (coefficients) that represent how much of each of the 64 frequencies are present.

At this point, no compression has taken place. We've traded one batch of 64 numbers for another and we can losslessly reverse the process and get back to our amplitude numbers if we choose--all we did was call those numbers something else. Since most of the information in a scene is concentrated in a few of the lower-frequency coefficients, there will be a large number of coefficients that have a zero value or are very close to zero. These can be rounded off to zero with little visual effect when pixel values are reconstituted by an inverse DCT process in the decoder.

The Importance Of Standards

The almost universal popularity of DCT illustrates the power of a standard. DCT may not be the best transform, but once a standard (either de facto or de jure) is in wide use, it will be around for a long time. Both equipment-makers and their customers need stability in the technologies they use, mainly so they can reap the benefits of their investments. The presence of a widely accepted standard provides that stability and raises the performance bar for other technologies that would like to compete. To displace an accepted standard, the competitor can't just be better, it must be several orders of magnitude better (and less expensive won't hurt either).

The incorporation of DCT techniques in the JPEG and MPEG standards and subsequent investment in and deployment of DCT--based compression systems have ensured its dominance in the compression field for a long time to come.

M-JPEG--JPEG, named for the Joint Photographic Experts Group, was developed as a standard for compressing still photographic images. Since JPEG chips were readily available before other compression chip sets, designers who wanted to squeeze moving pictures into products such as computer-based nonlinear editing systems adapted the JPEG standard to compress strings of video frames. Motion-JPEG was born. Unfortunately, the JPEG standard had no provision for storing the data related to motion, and designers developed their own proprietary ways of dealing with it. Consequently, it's often difficult to exchange M-JPEG files between systems.

Not long after the JPEG committee demonstrated success with still images, the Motion Picture Experts Group (MPEG) and DV standardization committees developed compression standards specifically for moving images. The trend has been for these newer motion standards to replace proprietary M-JPEG approaches.

A new JPEG-2000 still image standard using wavelet compression is being finalized. An extension of this standard (expected in 2001) may include a place to store data specifying the order and speed at which JPEG-2000 frames can be sequenced for display. This feature is designed to accommodate rapid sequence, digital still cameras and is not intended to compete with MPEG, however, it's conceivable that a new, standardized motion JPEG could emerge.

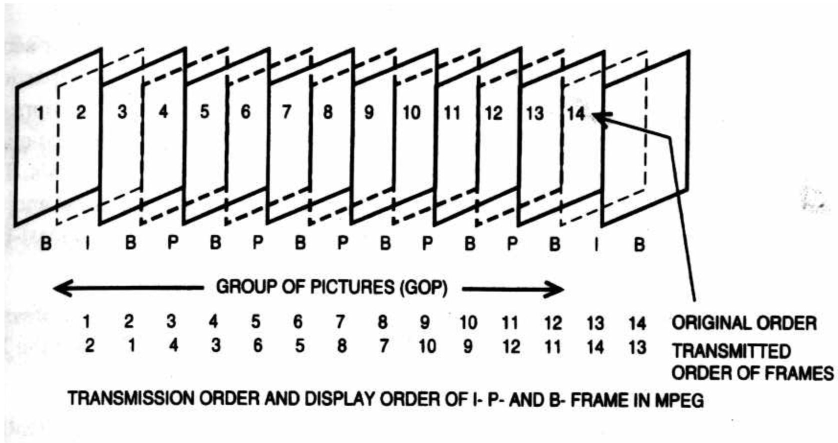
DV--The DV compression format was developed by a consortium of more than 50 equipment manufacturers as a consumer digital video cassette recording format (DVC) for both standard and high definition home recording. It is an intra-frame, DCT-based, symmetrical system. Although designed originally for home use, the inexpensive DV compression engine chip set (which can function as either encoder or decoder) has proved itself versatile enough to form the basis for a number of professional products including D9, DVCAM and DVCPRO. Both D9 and DVCPRO have taken advantage of the chipset's scalability to increase quality beyond that available in the consumer product. At 25 Mbps, the consumer compression ratio is about 5:1 with 4:1:1 color sampling. D9 and DVCPRO50 use two of the mass-market compression circuits running in parallel to achieve a 3.3:1 compression ratio with 4:2:2 color sampling at 50 Mbps. DVCPROHD and D9HD (scheduled to debut in 2000) are technically capable of recording progressive scan standard definition or interlaced and progressive HDTV at 100 Mbps. Similar extensions are possible beyond 100 Mbps and DV compression is not limited to video cassette recording, but can be applied to a range of compressed digital video storage and transmission applications.

MPEG--MPEG has become the 800--pound gorilla of compression techniques. It is the accepted compression scheme for all sorts of new products and services, from satellite broadcasting to DVD to the new ATSC digital television transmission standard, which includes HDTV. MPEG is an asymmetrical, DCT compression scheme which makes use of both intra- and inter-frame, motion compensated techniques. One of the important things to note about MPEG is that it's not the kind of rigidly defined, single entity we've been used to with NTSC or PAL, or the ITU-R 601 digital component standard. MPEG only defines bit streams and how those streams are to be recognized by decoders and reconstituted into video, audio and other usable information. How the MPEG bit streams are encoded is undefined and left open for continuous innovation and improvement. You'll notice we've been referring to MPEG bit streams in the plural. MPEG isn't a single standard, but rather a collection of standardized compression tools that can be combined as needs dictate. MPEG-1 provided a set of tools designed to record video on CDs at a data rate around 1.5 Mbps. While that work was underway, researchers recognized that similar compression techniques would be useful in all sorts of other applications.

The MPEG-2 committee was formed to expand the idea. They understood that a universal compression system capable of meeting the requirements of every application was an unrealistic goal. Not every use needed or could afford all the compression tools that were available. The solution was to provide a series of Profiles and Levels (see figure 2) with an arranged degree of commonality and compatibility between them.

Profiles And Levels--The six MPEG-2 Profiles gather together different sets of compression tools into toolkits for different applications. The Levels accommodate four different grades of input video ranging from a limited definition similar to today's consumer equipment all the way to high definition. Though they organized the options better, the levels and profiles still provided too many possible combinations to be practical. So, the choices were further constrained to specific "compliance points" within the overall matrix. So far, 12 compliance points have been defined ranging from the Simple Profile at Main Level (SP@ML) to the High Profile at High Level (HP@HL). The Main Profile at Main Level (MP@ML) is supposed to approximate today's broadcast video quality.

Any decoder that is certified at a given compliance point must be able to recognize and decode not only that point's set of tools and video resolutions, but also the tools and resolutions used at other compliance points below it and to the left. Therefore, an MP@ML decoder must also decode SP@ML and MP@LL. Likewise, a compliant



TRANSMISSION ORDER AND DISPLAY ORDER OF I- P- AND B- FRAME IN MPEG

MP@HL decoder would have to decode MP@H14L (a compromise 1440x1080 pixel HDTV format), MP@ML, MP@LL and SP@ML. As with MP@H14L, not all of the defined compliance points have found practical use. By far the most common is MP@ML. The proposed broadcast HDTV systems fall within the MP@HL point.

Group Of Pictures--MPEG achieves both good quality and high compression ratios at least in part through its unique frame structure referred to as the "Group of Pictures" or Gop (see figure 3). Three types of frames are employed: 1) intra-coded or "I" frames; 2) predicted "P" frames which are forecast from the previous I or P frame; and 3) "B" frames, which are predicted bidirectionally from both the previous and succeeding I or P frames. A GoP may consist of a single I frame, an I frame followed by a number of P frames, or an I frame followed by a mixture of B and P frames. A GoP ends when the next I frame comes along and starts a new GoP.

All the information necessary to reconstruct a single frame of video is contained in an I frame. It uses the most bits and can be decoded on its own without reference to any other frames. There is a limit to the number of frames that can be predicted from another. The inevitable transmission errors and small prediction errors will add up and eventually become intolerable. The arrival of a new I frame refreshes the process, terminates any accumulated errors and allows a new string of predictions to begin. P frames require far fewer bits because they are predicted from the previous I frame. They depend on the decoder having the I frame in memory for reference. Even fewer bits are needed for B frames because they are predicted from both the preceding and following I or P frames, both of which must be in memory in the decoder. The bidirectional prediction of B frames not only saves lots of bits, it also makes it possible to simulate VCR search modes.

The Simple Profile does not include B frames in its toolkit, thus reducing memory requirements and cost in the decoder. All other profiles include B frames as a possibility. As with all MPEG tools, the use, number and order of I, B and P frames is up to the designer of the encoder. The only requirement is that a compliant decoder be able to recognize and decode them if they are used. In practice, other standards that incorporate MPEG such as DVB and ATSC may place further constraints on the possibilities within a particular MPEG compliance point to lower the cost of consumer products.

Compression Ratio Versus Picture Quality

Because of its unique and flexible arrangement of I, P and B frames, there is little correlation between compression ratio and picture quality in MPEG. High quality can be achieved at low bit rates with a long GoP (usually on the order of 12 to 16 frames). Conversely, the same bit rate with a shorter GoP and/or no B frames will produce a lower quality image. Knowing only one or two parameters is never enough when you're trying to guess the relative performance of two different flavors of MPEG.

4:2:2 Profile

As MPEG-2 field experience began to accumulate, it became apparent that, while MP@ML was very good for distributing video, it had shortcomings for post production. The 720x480 and 720x526 sampling structures defined for the Main Level ignored the fact that there are usually 486 active picture lines in 525-line NTSC video and 575 in 625-line PAL. With the possible exception of cut transitions and limited overlays, lossy compressed video cannot be post-processed (resized, zoomed, rotated) in its compressed state. It must first be decoded to some baseband form such as ITU-R 601. Without specialized decoders and encoders designed to exchange information about previous compression operations, the quality of MP@ML deteriorates rapidly when its 4:2:0 color sampling structure is repeatedly decoded and re-encoded during post production. Long GoPs, with each frame heavily dependent on others in the group, make editing complex and difficult. And, the MP@ML 15 Mbps upper data rate limit makes it impossible to achieve good quality with a short GoP of one or two frames. Alternative intra-frame compression techniques such as DV and Motion-JPEG were available. But many people thought that if the MPEG MP@ML shortcomings could be corrected, the basic MPEG tools would be very useful for compressing contribution-quality video down to bit rates compatible with standard telecom circuits and inexpensive disk stores. And so they created a new Profile.

As its name suggests, the 4:2:2 Profile (422P@ML) uses 4:2:2 color sampling which more readily survives re-encoding. The maximum number of video lines is raised to 608. And the maximum data rate is increased to 50 Mbps. Noting the success of the new profile for standard definition images, the Society of Motion Picture and Television Engineers used MPEG's 422P@ML as a foundation for SMPTE-308M, a compression standard for contribution quality high definition. It uses the MPEG tools and syntax to compress HDTV at data rates up to 300 Mbps.

SMPTE submitted 308M to MPEG to help guide their work on a high level version of 422P. The documents for MPEG 422P@HL have been completed. The two standards are independent, but fully interoperable. The principal difference is that SMPTE 308M specifies an encoder constraint, requiring a staircase relationship between GoP and bitrate. Longer GoPs are permitted only at lower bitrates. MPEG places no restrictions on encoders and any combination of bitrate and GoP is permissible.

MPEG-4

With work on MPEG-1 and MPEG-2 complete, the Experts Group turned its attention to the problems posed by interactive multimedia creation and distribution. MPEG-4 is the result. It is not intended to replace MPEG 1 or 2, but, rather, builds on them to foster interactivity. Like MPEG-2, it is a collection of tools that can be grouped into profiles and levels for different applications. Version one of the MPEG-4 standard is already complete, and the ink is drying fast on version two. In committee jargon, MPEG-4 provides a Delivery Multimedia Integration Framework (DMIF) for "universal access" and "content-based interactivity." Translated, that means the new toolkit will let multimedia authors and users store, access, manipulate and present audio/visual data in ways that suit their individual needs at the moment, without concern for the underlying technicalities. It's a tall order. If accepted in practice, MPEG-4 could resolve the potentially unmanageable tangle of proprietary approaches we've seen for audio and video coding in computing, on the internet and in emerging wireless multimedia applications. Toward that end, it borrows from videoconferencing standards and expands on the previous MPEG work to enhance performance in low bitrate environments and provide the tools necessary for interactivity and intellectual property management.

What really sets MPEG-4 apart are its tools for interactivity. Central to these is the ability to separately code visual and aural "objects." Not only does it code conventional rectangular images and mono or multi-channel sound, but it has an extended set of tools to code separate audio objects and arbitrarily shaped video objects. A news anchor might be coded separately from the static background set. Game pieces can be coded independently from their backgrounds. Sounds can be interactively located in space. Once video, graphic, text or audio objects have been discretely coded, users can interact with them individually. Objects can be added and subtracted, moved around and re-sized within the scene. All these features are organized by a DIMF that manages the multiple data streams, two-way communication and control necessary for interaction.

Both real and synthetic objects are supported. There are MPEG-4 tools for coding 2D and 3D animations and mapping synthetic and/or real textures onto them. Special tools facilitate facial and body animation. Elsewhere in the toolkit are methods for text-to-speech conversion and several levels of synthesized sound. A coordinate system is provided to position objects in relation to each other, their backgrounds and the viewer/listener. MPEG-4's scene composition capabilities have been heavily influenced by prior work done in the Internet community on the Virtual Reality Modeling Language (VRML), and there is formal coordination between MPEG and the Web3d Consortium to insure that VRML and MPEG-4 evolve in a consistent manner.

Unlike VRML, which relies on text-based instructions, MPEG- 4's scene description language, Binary Format for Scenes (BIFS), is designed for real-time streaming. Its binary code is 10 to 15 times more compact than VRML's, and images can be constructed on the fly without waiting for the full scene to download.

Coding and manipulating arbitrarily shaped objects is one thing. Extracting them from natural scenes is quite another. Thus far, MPEG-4 demonstrations have depended on chromakey and a lot of hand work.
In version 2, programming capabilities will be added with MPEG-J, a subset of the Java programming language. Java interfaces to MPEG-4 objects will allow decoders to intelligently and automatically scale content to fit their particular capabilities.

The standard supports scalability in many ways. Less important objects can be omitted or transmitted with less error protection. Visual and aural objects can be created with a simple layer that contains enough basic information for low resolution decoders and one or more enhancement layers that, when added to that base layer, provide more resolution, wider frequency range, surround sound or 3D.

MPEG-4's basic transform is still DCT and quite similar to MPEG 1 and 2, but improvements have been made in coding efficiency and transmission ruggedness. A wavelet algorithm is included for efficient coding of textures and still images. MPEG-4 coding starts with a Very Low Bitrate Video (VLBV) core, which includes algorithms and tools for data rates between 5 kbps and 64 kbps. To make things work at these very low bit rates, motion compensation, error correction and concealment have been improved, refresh rates are kept low (between 0 and 15 fps) and resolution ranges from a few pixels per line up to CIF (352x288).

MPEG-4 doesn't concern itself directly with the error protection needed in specific channels such as cellular radio, but it has made improvements in the way payload bits are arranged so that recovery will be more robust. There are more frequent resynchronization markers. New, reversible variable length codes can be read forward or backward like a palindrome so decoders can recover all the data between an error and the next sync marker.

For better channels (something between 64 kbps and 2 Mbps), a High Bitrate Video (HBV) mode supports resolutions and frame rates up to Rec.601. The tools and algorithms are essentially the same as VLBV, plus a few additional ones to handle interlaced sources.

While MPEG-4 has many obvious advantages for interactive media production and dissemination, it's not clear what effect it will have on conventional video broadcasting and distribution. MPEG-2 standards are well established in these areas. For the advanced functions, both MPEG-4 encoders and decoders will be more complex and, presumably, more expensive than those for MPEG-1 and 2. However, the Studio Profile of MPEG-4 is expected to have an impact on high-end, high-resolution production for film and video.

MPEG-4 Studio Profile

At first glance, MPEG-4's bandwidth efficiency, interactivity and synthetic coding seem to have little to do with high resolution, high performance studio imaging. The MPEG-4 committee structure did, however, provide a venue for interested companies and individuals to address some of the problems of high-end image compression.
When you consider realtime electronic manipulation of high resolution moving images, the baseband numbers are enormous. A 4000 pixel by 4000 pixel, 4:4:4, YUV/RGB, 10-bit, 24 fps image with a key channel requires a data rate in excess of 16 Gbps. Even the current HDTV goal (just out of reach) of 1920x1080 pixels, 60 progressive frames and 4:2:2, 10-bit sampling requires just under 2.5 Gbps. Upgrade that to 4:4:4 RGB, add a key channel and you're up to about 5 Gbps. It's easy to see why standards for compressing this stuff might be useful.

The MPEG-4 committee was receptive to the idea of a Studio Profile, and their structure provided an opportunity to break the MPEG-2 upper limits of 8-bit sampling and 100 Mbps data rate. The project gathered momentum as numerous participants from throughout the imaging community joined in the work. Final standards documents are expected by the end of 2000.

A look at the accompanying table shows three levels in the proposed new profile. Compressed data rates range between 300 Mbps and 2.5 Gbps. With the exception of 10-bit sampling, the Low Level is compatible with and roughly equivalent to the current MPEG-2 Studio Profile at High Level. The Main Level accommodates up to 60 frames progressive, 4:4:4 sampling, and 2048x2048 pixels. The High Level pushes things to 12-bit sampling, 4096x4096 pixels and up to 120 frames per second. The draft standard is expected to include provisions for key channels, although the number of bits for them were still in question as of this writing.

Although you can't have everything at once (a 12-bit, 120 fps, 4:4:4:4, 4096x4096 image isn't in the cards), within a level's compressed data rate limitations, you can trade resolution, frame rate, quantizing and sampling strategies to accomplish the task at hand. Like all MPEG standards, this one defines a bitstream syntax and sets parameters for decoder performance. For instance, a compliant High Level decoder could reproduce a 4096x4096 image at 24 frames per second or a 1920x1080 one at 120 fps. At the Main Level, a 1920x1080 image could have as many as 60 fames per second where a 2048x2048 one would be limited to a maximum of 30 fps.

As a part of MPEG-4, the Studio Profile could use all the scene composition and interactive tools that are included in the lower profiles. But high-end production already has a large number of sophisticated tools for image composition and manipulation, and it's not clear how or if similar components of the MPEG-4 toolkit will be applied to the Studio Profile.

One side benefit of a Studio Profile in the MPEG-4 standard is that basic elements such as colorimetry, macroblock alignments and other parameters will be maintained all the way up and down the chain. That should help maintain quality as the material passes from the highest levels of production all the way down to those Dick Tracy wrist receivers.

The Other MPEGs

MPEG 7 and 21 are, thankfully, not new compression standards, but rather attempts to manage motion imaging and multimedia technology.

MPEG-7 is described as a Multimedia Content Description Interface (MCDI). It's an attempt to provide a standard means of describing multimedia content. Its quest is to build a standard set of descriptors, description schemes and a standardized language that can be used to describe multimedia information. Unlike today's text-based approaches, such a language might let you search for scenes by the colors and textures they contain or the action that occurs in them. You could play a few notes on a keyboard or enter a sample of a singer's voice and get back a list of similar musical pieces and performances. If the MPEG-7 committee is successful, search engines will have at least a fighting chance of finding the needles we want in the haystack of audio visual material we're creating. A completed standard is expected in September 2000.

MPEG-21 is the Group's attempt to get a handle on the overall topic of content delivery. By defining a Multimedia Framework from the viewpoint of the consumer, they hope to understand how various components relate to each other and where gaps in the infrastructure might benefit from new standards.The subjects being investigated overlap and interact. There are network issues like speed, reliability, delay, cost performance and so on. Content quality issues include things such as authenticity (is it what it pretends to be?) and timeliness (can you have it when you want it?), as well as technical and artistic attributes. Ease of use, payment models, search techniques and storage options are all part of the study, as are the areas of consumer rights and privacy. What rights do consumers have to use, copy and pass on content to others? Can they understand those rights? How will consumers protect personal data and can they negotiate privacy with content providers? A technical report on the MPEG-21 framework is scheduled for mid-2000.

The Missing MPEGs

Since we've discussed MPEG 1, 2, 4, 7 and 21, you might wonder what happened to 3, 5, 6 and the rest of the numbers. MPEG-3 was going to be the standard for HDTV. But early on, it became obvious that MPEG-2 would be capable of handling high definition and MPEG-3 was scrapped. When it came time to pick a number for some new work to follow MPEG-4, there was much speculation about what it would be. (Numbering discussions in standards work are like debates about table shape in diplomacy. They give you something to do while you're trying to get a handle on the serious business.) With one, two and four already in the works, the MPEG folks were on their way to a nice binary sequence. Should the next one be eight, or should it just be five? In the end, they threw logic to the winds and called it seven. Don't even ask where 21 came from (the century perhaps?).

Some Final Thoughts

Use clean sources. Compression systems work best with clean source material. Noisy signals, film grain, poorly decoded composite video--all give poor results. Preprocessing that reduces noise, shapes the video bandwidth and corrects other problems can improve compression results, but the best bet is a clean source to begin with. Noisy and degraded images can require a premium of 20 to 50 percent more bits.

Milder is better. Video compression has always been with us. (Interlace is a compression technique. 4:2:2 color sampling is a compression technique.) It will always be with us. Nonetheless, you should choose the mildest compression you can afford in any application, particularly in post production where video will go through multiple processing generations.
Compression schemes using low bit rates and extensive inter-frame processing are best suited to final program distribution.

More is better. Despite the fact that there is only a tenuous relationship between data rate and picture quality, more bits are usually better. Lab results suggest that if you acquire material at a low rate such as 25 Mbps and you'll be posting it on a nonlinear system using the same type of compression, the multigeneration performance will be much better if your posting data rate is higher, say 50 Mbps, than if you stay at the 25 Mbps rate.

Avoid compression cascades. When compressed video is decoded, small errors in the form of unwanted high frequencies are introduced where no high frequencies were present in the original. If that video is re-encoded without processing (level changes, zooming, rotation, repositioning) and with the same compression scheme, the coding will usually mask these errors and the effect will be minimal. But if the video is processed or re-encoded with a different compression scheme, those high frequencies end up in new locations and the coding system will treat them as new information. The result is an additional loss in quality roughly equal to that experienced when the video was first compressed. Re-coding quality can be significantly improved by passing original coding parameters (motion vectors, quantization tables, frame sequences, etc.) between the decoder and subsequent encoder. Cascades between different transforms (i.e. from DCT based compression to Wavelets and vice versa) seem to be more destructive than cascades using the same transform. Since Murphy's Law is always in effect, these losses never seem to cancel each other, but add rapidly as post production generations accumulate.

Quality is subjective. Despite recent advances in objective measures, video quality in any given compression system is highly dependent on the source material. Beware of demonstrations that use carefully selected material to achieve low bit rates. Be sure to see what things look like with your own test material covering the range of difficulty you expect in daily operation.

Bandwidth based on format. The total ATSC bandwidth is 19.39 Mbps, which includes audio, video and other data. As the image quality is increased, more bandwidth is needed to send the image, even though it is compressed. Below is a list of popular distribution formats and the approximate bandwidth they will require (30 fps for interlace, 60 fps for progressive).

• 1080i: 10 to 18 Mbps (10 with easy clean film material, easy clean video material may be a little higher, sports will require 18, all material will require 18 on some of the earlier encoders).
• 720p: 6 to 16 Mbps (low numbers with talking heads and films, sports may be acceptable under 16 Mbps).
480p: 4 to 10 Mbps (low number highly dependent on customer expectation that this a very high quality 16:9 image).
• 480i: 2 to 6 Mbps (could average under 3 Mbps with good statistical multiplexing).

FROM: Video Compression,
an article by Glen Pensinger