Video Signals and Connectors

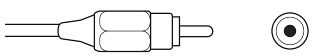
When you capture and output, the type of video signal you use to connect your equipment is a critical factor that goes into determining the quality of your video. Video camcorders, decks, and monitors can use different types of signals, depending on the environment they are intended for. Consumer equipment usually has limited video signal choices; professional equipment gives you the greatest range of options.

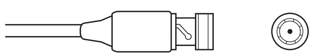
For more information, see Volume IV, Appendix A, "Video Formats."

Composite

Composite is the lowest common denominator of video signals. A composite signal runs all color and brightness information on a single cable, resulting in lower-quality video compared to the quality of other formats. Nearly all video devices have a composite input and output. This format uses a single RCA or BNC connector.

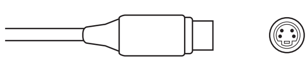
In professional editing environments, composite video signals are most commonly used for troubleshooting, for menu outputs, and for low-quality preview monitoring. For consumer and home use, composite signals are often used to connect VCRs or DVD players to televisions.


RCA connector


BNC connector

S-Video

S-Video, also known as *Y/C*, is a higher-quality video signal used by high-end consumer video equipment. The image looks sharper and has better color than a composite video image because S-Video keeps the color and brightness information separate on two cables. Most low-cost analog-to-digital video interfaces have S-Video as their highest-quality video connector. Use care when working with S-video connectors; the four delicate pins can be bent easily.

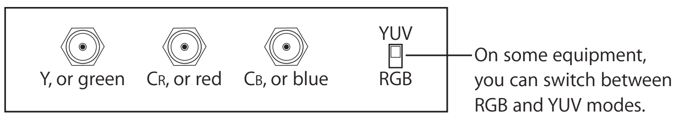

S-Video connector

Component YUV and Component RGB

Professional video equipment, such as Betacam SP decks, has component YUV ($Y'C_BC_R$) video inputs and outputs. Component YUV separates color and brightness information into three signals, which keeps the color quality more accurate than that of other systems. Component YUV is as good as analog video gets. High-end consumer devices, such as DVD players and televisions, have increasingly begun to support component YUV.

Note: Another form of component video, component RGB, is not as widespread on professional equipment as component YUV.

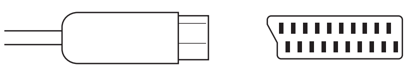
Both component YUV and RGB signals use from three to five connectors. You can use three BNC connectors, plus a fourth (typically labeled "genlock" or "house sync") to send a timing signal. Sync can also be embedded in the Y or G part of the signal (using three connectors), a separate composite signal on a fourth connector, or separate H and V drive signals (using five connectors). See your equipment's documentation for more information.



Component

SCART

Consumer PAL equipment sometimes has a special connector called a *SCART connector.* A SCART connector has multiple pins that run composite, component RGB, and stereo audio in one bundle. SCART input or output can be broken up into individual connections using special adapters available from video and home electronics stores.

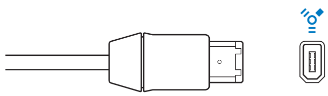


SCART connector

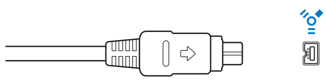
FireWire 400

FireWire 400, also called *IEEE 1394a* or *i.LINK*, is the consumer and professional standard for formats such as DV, DVCAM, DVCPRO, DVCPRO 50, DVCPRO HD, and HDV. FireWire is an inexpensive and easy way to capture and output high-quality digital video using a variety of camcorders and decks and is capable of data rates as high as 400 Mbps. Standard FireWire cables can be up to 4.5 meters long.

There are two kinds of FireWire connectors: a 4-pin connector (typically found on video equipment such as camcorders or decks) and a 6-pin connector (used for computer equipment). However, some newer video equipment uses the 6-pin connector, and some video interfaces use the 4-pin connector. See your equipment's documentation for more information.



FireWire 400 (6-pin)

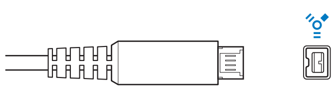


FireWire 400 (4-pin)

FireWire 800

FireWire 800, also called *IEEE 1394b*, is the next generation of FireWire after IEEE 1394a, a higher-bandwidth version capable of data transfer speeds of up to 800 Mbps. FireWire 800 is also capable of supporting cable distances of up to 100 meters.

In addition to the standard 9-pin-to-9-pin FireWire 800 cables, 9-pin-to-4-pin and 9-pin-to-6-pin FireWire 400 to FireWire 800 cables are also available to connect older devices to a FireWire 800 interface.

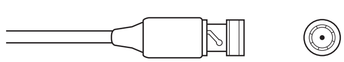


FireWire 800 (9-pin)

Note: FireWire 800 is commonly used to connect hard disks and other data peripherals to your computer, but this connector is rarely used to connect video devices.

SDI

Serial Digital Interface (SDI) is the standard for high-end, uncompressed digital video formats such as D1, D5, and Digital Betacam. Many devices can send both video and audio data through a single SDI connection.



BNC connector

HD-SDI

High Definition Serial Digital Interface (HD-SDI) is a higher-bandwidth version of SDI designed for the extremely high data rates required by uncompressed HD video. Like SDI, HD-SDI is capable of sending both video and audio through a single connection. The following decks have HD-SDI interfaces: DVCPRO HD, D-5 HD, and HDCAM decks.

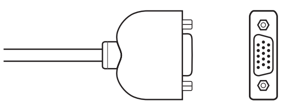
Some devices provide even higher data rates by pairing two HD-SDI channels together (known as *dual-link HD-SDI*). Uncompressed HD RGB video and other digital cinema formats can be transmitted using dual-link HD-SDI.

SDTI

Serial Digital Transport Interface (SDTI) is based on SDI, allowing native video formats to be sent in real time within an SDI video stream. SDTI does not define a specific video signal format but instead uses the structure of SDI to carry any kind of data. This allows video facilities to use their existing SDI patchbays and routers to transfer other native video formats, or transfer any kind of data. For example, some DV decks can transfer DV via SDTI, which means native DV can be transferred long distances over existing coaxial cable instead of the usual FireWire connection. Other formats, such as HDCAM and MPEG, can also be transferred via packets within an SDTI transport stream.

VGA

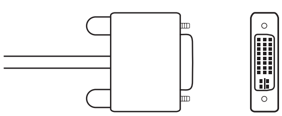
VGA interfaces use a 15-pin D-subminiature connector to transfer analog RGB video and sync information between computers and computer CRT displays or video projectors. This connector and signal format is being replaced by newer display formats such as DVI and HDMI. However, adapters are available to convert between DVI and VGA.



VGA connector

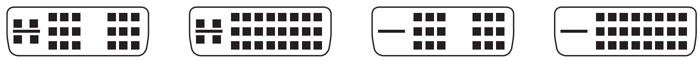
DVI

Digital Visual Interface (DVI) transfers full-resolution analog or digital signals between computers or HD video devices and flat-panel displays or projectors. DVI connectors have up to 24 pins plus four additional pins for analog signals. Not all devices use all pins, so read the documentation included with your equipment before purchasing DVI cables and making connections.



DVI connector

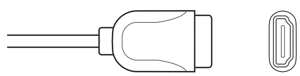
DVI supports single- and dual-link connections. Single-link DVI connections are limited to 2.6 megapixels with a refresh rate of up to 60 Hz. Dual-link DVI connectors extend the number of pixels that can be transferred per second to drive larger displays.



DVI-I (single link) DVI-I (dual link) DVI-D (single link) DVI-D (dual link)

HDMI

High-Definition Multimedia Interface (HDMI) supports both digital television and computer signals and can also include multiple digital audio channels. HDMI devices are compatible with single-link digital DVD signals via an adapter, although no audio or additional metadata can be included. Many HD display devices and digital television set-top boxes include HDMI connectors.



HDMI connector

Connecting Professional Video Devices

Regardless of what format or video interface you use, the same basic steps apply when you connect a VTR or camcorder to your computer.

Connecting Professional SD Video Devices

If you're using a format other than DV, such as Betacam SP or Digital Betacam, you need to install a third-party video interface that supports the proper signal for the format you're using. Some video interfaces can digitize analog video input and output (for analog formats like Betacam SP), while others capture video only if it is already digital. In either case, the video interface encodes the video information using a codec, which may or may not apply compression to the video data to make it smaller while stored on disk. Compression settings used by video interfaces are typically controlled by software.

Unlike DV video devices (which use a single FireWire cable), third-party interfaces send and receive video, audio, and device control data on separate cables. For remote device control connections on professional equipment, 9-pin (DB9) connectors are used.

A Recommended System Using a Third-Party Video Interface

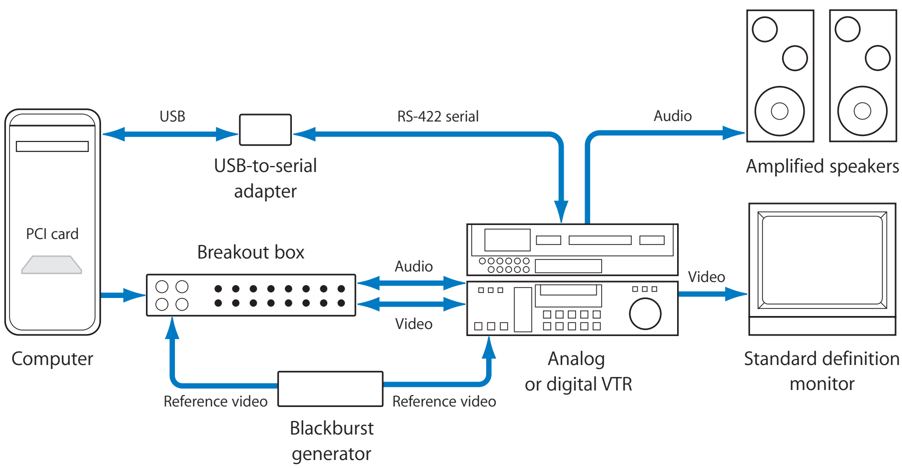
To set up a system using a third-party video interface, you need the following equipment:

- Your computer and display
- A non-DV format video device (a camcorder or deck)
- Audio and video cables for your system
- A third-party capture interface installed in or connected to your computer

  Note: Some third-party video interfaces have a breakout box connected to the card with a special cable, which is included.

- A USB-to-serial adapter or internal modem serial port adapter
- An RS-422 video deck control cable
- A blackburst generator, with the appropriate cables to connect it to both your third-party video interface and your video and audio devices

The following illustration shows a typical SD setup:



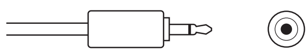
Following are basic instructions for connecting a video device to a third-party video interface in your computer, as well as connecting remote device control.

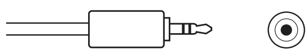
Analog Audio Connectors and Signal Formats

Different audio connectors are suited for different purposes. Audio connectors are often indicative of the kind of signal they transmit. However, there are enough exceptions that it's important to know what kind of audio signal you are connecting, in addition to the connector type. An important distinction is whether an audio connector carries a balanced or an unbalanced signal.

1/8" Mini Connectors

These are very small, unbalanced audio connectors. Many computers have 1/8" mini inputs and outputs at –10 dBV line level, and many portable audio devices such as CD players, iPod digital music players, and MP3 players use these connectors for headphone outputs. Portable MiniDisc and DAT recorders often use 1/8" mini connectors for connecting microphones.

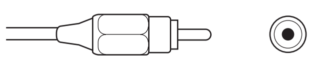




Note: Some Macintosh computers and portable audio recorders also use a connector that combines both a stereo miniplug and a 1/8" optical digital connection (see "S/PDIF" on page 194) in a single jack.

RCA Connectors

Most consumer equipment uses RCA connectors, which are unbalanced connectors that usually handle –10 dBV (consumer) line levels.



1/4" Tip-Sleeve Connectors

1/4" tip-sleeve (TS) connectors (also called *1/4" phone* connectors) with a tip and a sleeve are unbalanced connectors often used for musical instruments like electric guitars, keyboards, amplifiers, and so on.

1/4" tip-sleeve (TS) connector

1/4" Tip-Ring-Sleeve Connectors

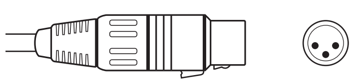
Professional equipment often uses 1/4" tip-ring-sleeve (TRS) audio connectors with +4 dBu line level. 1/4" TRS connectors connect to three wires in an audio cable—hot, neutral, and ground—and usually carry a balanced audio signal. In some situations, the three wires may be used to send left and right (stereo) signals, making the signals unbalanced.

1/4" tip-ring-sleeve (TRS) connector

Note: Tip-sleeve and tip-ring-sleeve connectors look almost identical. Some audio devices (especially mixers) accept a TS connector in a TRS jack, but you should always check the equipment documentation to be sure. Remember that most 1/4" TS connectors connect to –10 dBV line level equipment, whereas 1/4" TRS connectors usually expect a +4 dBu line level.

XLR Connectors

These are the most common professional audio connectors. They almost always carry a balanced signal. Many cables use an XLR connector on one end and a 1/4" TRS connector on the other. The signal may be microphone level (when using a microphone) or +4 dBu/dBm (professional) line level.



XLR connector

Digital Audio Connectors and Signal Formats

Although digital audio signals are completely different from analog signals, the same connectors are often used for convenience. For example, an XLR connector can be used to carry an analog audio signal or an AES/EBU digital audio signal.

AES/EBU

The AES/EBU digital audio specification was jointly developed by the Audio Engineering Society (AES) and the European Broadcasting Union (EBU). AES/EBU audio signals typically use XLR connectors, but 25- or 50-pin D-subminiature connectors can also be used for multiple channels on interfaces or mixers.

Traditionally, AES/EBU sample rates were limited to 44.1 and 48 kHz at up to 24 bits per sample. However, a "dual wire" mode allows some equipment to pair AES/EBU connectors to increase the sample rate. Some newer devices also support "single wire" mode with sample rates up to 192 kHz.

S/PDIF

Sony/Philips Digital Interface Format (S/PDIF) is a consumer-level variation of the AES/EBU digital audio protocol. S/PDIF audio data can be transmitted several ways, including:

- Via coaxial cables with an RCA connector
- Via optical TOSLINK

Connectors for S/PDIF signals are found on most consumer digital equipment, such as DAT recorders, CD players, DVD players, MiniDisc equipment, and some audio interfaces.

Dolby Digital (AC-3)

Dolby Digital (AC-3) is a compressed digital audio signal format used for transmission of 5.1-channel surround sound. Mono and stereo signals can also be carried in this format. Typically, this audio signal is embedded within a S/PDIF signal and carried via TOSLINK or coaxial cables with RCA connectors.

DTS

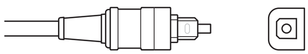
Digital Theater System (DTS) is a compressed digital audio signal format used for transmission of 5.1-channel surround sound. This format is primarily used in movie theaters and on DVD releases. More recent variations of DTS support more than six channels of audio. This audio signal is usually embedded within a S/PDIF signal and carried via TOSLINK or coaxial cables with RCA connectors.

TOSLINK

TOSLINK is an optical digital audio format developed by the Toshiba Corporation. These digital audio formats can be transmitted via TOSLINK optical cables and connectors:

- S/PDIF
- ADAT Lightpipe

Some Macintosh computers have a single interface that combines a TOSLINK connector with an analog stereo miniplug.



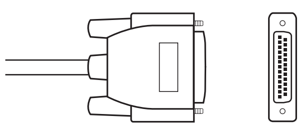
TOSLINK connector

ADAT Lightpipe

ADAT Lightpipe is an eight-channel digital audio format developed by Alesis. This signal format uses TOSLINK optical connectors. Eight channels are supported at sample rates of 44.1 and 48 kHz using 24 bits per sample. Higher sample rates are available by pairing channels (this format is sometimes called *sample multiplexing*, or *S/MUX*). For example, a sample rate of 192 kHz is possible, but the number of channels is reduced to two. However, not all equipment supports channel pairing and increased sample rates.

TDIF

Tascam Digital Interface (TDIF) is a signal format for transferring digital audio between Tascam digital multitrack recorders or digital mixers. A 25-pin D-subminiature connector is used. Eight channels are supported at sample rates of 44.1 and 48 kHz using 24 bits per sample. Higher sample rates are available by pairing channels.



TDIF connector

About Analog Audio Levels

There are six basic kinds of analog audio levels found on most equipment:

- *Microphone level:* Around 50 or 60 dB less than line level. When you use a microphone, the level is very low, requiring a preamplifier to raise the signal to line level before it can be recorded or processed. Most audio mixers, cameras, and professional portable recording devices have built-in preamplifiers.
- *Instrument level:* Between microphone and line level, around –20 dBV or so. Guitars and keyboards usually output at instrument level.
- *Line level (consumer):* Consumer line level is output at –10 dBV.
- *Line level (professional):* Professional line level is output at +4 dBu (or dBm in older equipment).
- *Speaker level:* This signal varies considerably depending on the amplifier used, but it is very strong compared to the others because it is used to drive speakers.
- *Headphone level:* This signal is like speaker level, but much lower. The sole purpose of this signal is to drive stereo headphones.

About Units of Analog Audio Measurement

Professional audio equipment typically uses higher voltage levels than consumer equipment, and it also measures audio on a different scale. Keep the following points in mind when using consumer and professional audio equipment together:

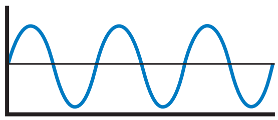
- Professional analog devices measure audio using dBu (or dBm in older equipment). 0 dB on the audio meter is usually set to +4 dBu, which means optimal levels are 4 dB greater than 0 dBu (.775 V), or 1.23 V.
- Consumer audio equipment measures audio using dBV. The optimal recording level on a consumer device is –10 dBV, which means the levels are 10 dB less than 0 dBV (1 V), or 0.316 V.

Therefore, the difference between an optimal professional level (+4 dBu) and consumer level (–10 dBV) is not 14 dB, because they are using different signals. This is not necessarily a problem, but you need to be aware of these level differences when connecting consumer and professional audio equipment together.

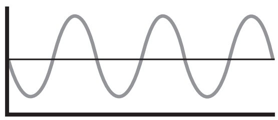
About Balanced Audio Signals

Audio cables can be either *balanced* or *unbalanced,* depending on their intended use. For long cable runs, especially when using relatively low microphone levels, a three-wire balanced audio circuit reduces noise. Balanced audio cables use the principle of phase cancelation to eliminate noise while maintaining the original audio signal.

A balanced audio cable sends the same audio signal on two wires, but inverts the *phase* of one signal by 180 degrees.

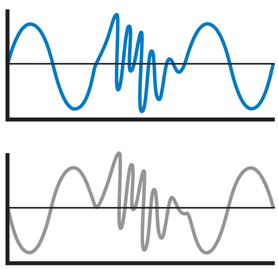


Original signal



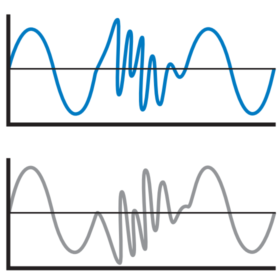
Inverted signal (reverse phase)

When noise is introduced into the cable, it is introduced equally to both the original and the inverted signal.



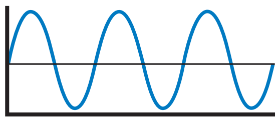
Noise on line
(affects both signals)

When the signal arrives at its destination, the inverted signal is put back in phase and both signals are combined. This puts the original and inverted signals back in phase, but it causes the noise signals on each line to be out of phase.



Inverted signal
(inverted again)

Now, both audio signals are in phase, but the noise is inverted, causing the noise to be canceled. At the same time, the original signal gets a little stronger because it is sent on two wires and combined. This helps compensate for the reduction in signal strength that occurs naturally on a long cable run.



Combined signals
(noise eliminated)

Any noise introduced into the cable across its long run is almost completely eliminated by this process.

Note: Unbalanced cables have no way of eliminating noise and are therefore not as robust for long-distance cable runs, microphone signals, and other professional uses.

Connecting Professional Audio Devices

The steps for connecting audio playback and recording devices are similar to the steps for connecting professional video devices. Many professional audio-only devices such as DAT recorders and Tascam DA-88/DA-98 multitracks support remote device control and audio insert editing.

If you plan to capture footage using separate video and audio interfaces, you may need to set up additional synchronization between your audio device and audio interface. For more information, see "Synchronizing Equipment with a Blackburst Generator" on page 200.

Connecting Professional Digital Audio Devices

Professional digital audio devices often use balanced XLR connectors. Each XLR connector carries two AES/EBU digital audio channels. Connect the digital audio outputs of your video or audio device to your audio interface (or its breakout box). If your video interface has the appropriate connectors, you can also connect the audio outputs of your device to the audio inputs on the video interface.

Note: Professional analog audio devices also use XLR connectors, but the signal is incompatible with AES/EBU digital audio.

Determining Your Hard Disk Storage Options

13

To make the most of your Final Cut Pro editing system, you need to make appropriate choices about hard disk selection and maintenance.

This chapter covers the following:

Working with Scratch Disks and Hard Disk Drives

By default, Final Cut Pro uses the hard disk on which the application is installed as your scratch disk to store captured and render files. Ideally, you should use a hard disk other than your main system disk as your scratch disk. Depending on how much space you need for your media, you can have up to twelve scratch disks in your Final Cut Pro editing system.

Important: If you have multiple hard disks and partitions, make sure they do not have similar names, or you could encounter problems during capture. For more information, see "Using Multiple Hard Disks" on page 39.

Data Rates and Storage Devices

The data rate of the video you capture depends on the format of the source video and the codec you use for capture. If you are capturing low data rate video, chances are you can use more inexpensive storage devices. If you need to capture extremely high data rate video, you may need a faster hard disk. Here are some examples of data rates for common capture formats:

Format	Typical data rate
OfflineRT (using Photo JPEG)	Varies between 300 and 500 KB/sec.
25:1 Motion JPEG (M-JPEG)	1 MB/sec.
DV (25) HDV (1080i)	3.6 MB/sec.
DVCPRO 50	7.2 MB/sec.
DVCPRO HD (1080i60) DVCPRO HD 720p60)	11.75 MB/sec.
DVCPRO HD (720p24)	5 MB/sec.
2:1 Motion JPEG (M-JPEG)	12 MB/sec.
Uncompressed SD video	24 MB/sec.
Uncompressed 8-bit 1080 29.97i HD video	121.5 MB/sec.
Uncompressed 10-bit 1080 29.97i HD video	182.3 MB/sec.

Whatever disk drive technology you decide to use, your storage disk's sustained transfer speed must be fast enough to keep up with the data rate. Depending on the data rate of the video you're capturing, a single drive may or may not be enough.

For example, if you plan to capture uncompressed SD video at 24 megabytes per second (MB/sec.), it's unlikely that a single hard disk will be able to record the data fast enough. Even if you somehow successfully get the data on disk, Final Cut Pro may drop frames during playback or output.

If your hard disk or its connection to your computer does not support the data rate of your video format, you need to consider three factors:

- *Sustained transfer speed* is a measurement of how fast data can be written to a disk in MB/sec. When you use a video interface that utilizes M-JPEG compression, the sustained transfer speed of your hard disk determines the maximum quality of the video you can capture. Disks with a higher sustained transfer speed allow you to capture video media files with a higher data rate, which results in higher visual quality.
- *Seek time* is a measurement of how quickly data stored on the disk can be accessed in milliseconds (ms). Low seek times are important when playing back an edited sequence of clips, because the disk must spend a lot of time searching for the next clip to play.
- A faster *spindle speed* increases a disk's sustained transfer rate (typical multimedia disks run at 7200 revolutions per minute, or rpm). However, the faster a hard disk runs the more it heats up, so ventilation is important when you install disks internally or in external enclosures.

Note: Removable media drives such as Jaz, Zip, and CD-RW drives are not suitable for video capture and playback because of their low data transfer rates.

Determining How Much Space You Need

The amount of disk space you need depends on the specifications of the video format you are using for editing. In some cases, you can capture video at a lower quality (which saves disk space) for rough editing and then recapture only what you need at higher quality to create the finished movie. This process is known as *offline/online editing*. For more information, see Volume IV, Chapter 5, "Offline and Online Editing."

Know Your Shooting Ratio

Remember that when you start editing a movie, you need to capture much more media than you will use in the final movie. The ratio between the amount of footage you begin with and the final duration of the movie is called the *shooting ratio*. When you are estimating how much disk space you need for a project, calculate it based on the total amount of media you plan to capture and use during editing, not the intended duration of the final movie.

Planning for Additional Media Files

In addition to space for captured files and project files, you need extra space for render files, graphics, movie files created in other applications (such as animations), additional audio files, and so on. A loose rule of thumb to determine how much space you need is to multiply the amount of space needed for your finished program by five.

Ultimately, the amount of extra space you reserve depends on how much additional media you create during editing. For example, if you use hardly any effects, additional render files may not be a factor. If you are using only a few graphics files and little additional audio, these may not be a concern, either.

Keep in mind that although real-time effects don't require additional drive space for rendering, you still need to render the effects at high quality for final output, so at that point you need enough disk space for render files.

Calculating Hard Disk Space Requirements

You can use the table below to estimate how much disk space you need for your project.

Video data transfer rates	30 sec.	1 min.	5 min.	10 min.	30 min.	60 min.
500 KB/sec. OfflineRT (using Photo JPEG)	15 MB	30 MB	150 MB	300 MB	900 MB	1.8 GB
1 MB/sec. Offline-quality M-JPEG	30 MB	60 MB	300 MB	600 MB	1.8 GB	3.6 GB
3.6 MB/sec. DV-format video HDV (1080i)	108 MB	216 MB	1.08 GB	2.16 GB	6.5 GB	13 GB
6 MB/sec. Medium-quality M-JPEG	180 MB	360 MB	1.8 GB	3.6 GB	10.8 GB	21.6 GB
11.75 MB/sec. DVCPRO HD 1080i	352.5 MB	705 MB	3.4 GB	6.9 GB	20.7 GB	41.3 GB
12 MB/sec. High-quality 2:1 M-JPEG	360 MB	720 MB	3.6 GB	7.2 GB	21.6 GB	43.2 GB
24 MB/sec. Uncompressed SD video	720 MB	1.4 GB	7.2 GB	14.4 GB	43.2 GB	86.4 GB
121.5 MB/sec. 8-bit uncompressed 1080 29.97i HD video	3.6 GB	7.3 GB	36.5 GB	72.9 GB	218.7 GB	434.4 GB

Sample Calculation for Disk Space Requirements

Suppose you want to create a music video that's approximately 4 minutes long using DV video for capture, editing, and output. Consider a shooting ratio of 15:1, meaning you shot 15 times more footage than you will use in the final movie.

Total duration of media captured to disk:

- 15 x 4 minutes = 60 minutes

Data rate requirements for DV media:

- 3.6 MB/sec. video data rate x 60 seconds = 216 MB/min.

Calculated disk space requirements for media:

- 60 minutes x 216 MB/min. = 12,960 MB
- 12,960 MB ÷1024 MB per GB = 12.66 GB

Multiply the final movie length by a safety margin of 5 for extra files:

- 4 minutes x 216 MB/min. = 864 MB x 5 = 4320 MB
- 4320 MB ÷ 1024 MB per GB = 4.22 GB

Total disk space requirements:

- 12.66 GB + 4.22 GB = 16.88 GB

Round your calculation up to 17 GB to be safe. This is the amount of disk space you'll need for this one project. If you plan to work on multiple projects at the same time, estimate the amount for each project and add these numbers together.

Note: These calculations are also important when planning how to archive your projects when they are finished, though many people choose to archive only the project file and not back up their media files (since the original footage is stored on tape, you can always recapture the footage if necessary).

Video Formats

A

Appendix

This appendix covers the following:

Characteristics of Video Formats

All video formats achieve the same basic goal: they store black-and-white or color information as electronic *lines* that make up a *video frame*. The number of video frames recorded per second varies depending on the video standard the format supports (for example, NTSC formats are recorded at 29.97 fps; PAL formats are recorded at 25 fps).

Video formats can be characterized by the following factors:

- *The medium used to store the video information:* This is primarily videotape, but can also be optical disc, solid-state memory, or a hard disk.
- *The size of the media and the shape of the shell:* For example, videotape may be 1", 1/2", 3/4", or 8 mm. Many video formats have different shell sizes for portable and studio use, such as mini-DV (portable) and the larger DV cassettes for studio decks.
- *The video standard supported:* For example, NTSC, PAL, ATSC (HDTV 1080i or 720p), and so on.
- *The type of electronic signal recorded on tape:* In other words, the way luma (black-and-white) and chroma (color) information are combined and recorded.
- *The aspect ratio of the video frame:* The ratio of the frame width to the frame height.
- *The dimensions of the video frame:* The number of pixels per line, and the number of lines per frame.
- *The aspect ratio of the pixels:* This is a subtle factor that is explained in more detail below.
- *The frame rate:* The number of frames recorded per second.
- *The scanning method:* Interlaced fields (two fields per frame) or progressive (one complete frame at a time).
- *Color recording method:* RGB, component (YUV), S-Video (Y/C), or composite.

- *Color sampling:* For component digital formats, the ratio of color samples to black-and-white (or luma) samples (for example, 4:4:4, 4:2:2, and 4:1:1).
- *Sample rate:* The number of samples per second of each video line. This is just like the sample rate for audio, except the signals sampled are video lines, where each sample represents light intensity instead of sound intensity.
- *Bit depth:* The number of bits used to store each video sample, which determines the ability of the format to capture each sample's (or pixel's) light intensity precisely, and how well subtle differences in intensity can be stored.
- *Compressor (or codec):* A video compressor attempts to reduce the amount of digital data required to store each frame without compromising the quality of the image.

Storage Medium

Video—especially digital video—can be stored on more than just videotape. The characteristics of the storage medium determine playback and recording capabilities. For example, magnetic and optical disc media (such as CDs, DVDs, and hard disks) are capable of nonlinear reading and writing, whereas videotape is inherently linear. Videotape is still a very efficient means of storing large amounts of digital data in a small space, but other types of media are quickly gaining ground.

Tape Size, Cassette Shape, and Tape Coating

The width of a videotape is directly related to how much information can be stored. In analog formats, wider tape usually yields higher quality, but other factors can help reduce tape size with minimal loss of quality. For example, Betacam SP and VHS both use 1/2"-wide tape, but Betacam SP uses a high-quality component video recording method that keeps luma and chroma information separate, whereas VHS uses a composite method that mixes these signals into one, causing interference between the two. The physical formulations of these two kinds of tape are also different, which accounts for some of the quality difference.

The size of the cassette itself can vary as well. For example, the Betacam SP format comes in both small and large sizes. The small tapes are used for camcorders, while the large format is used in studio VTRs.

Aspects of the physical composition of magnetic tape, such as density of magnetic particles, limit the data rate and track size that can be recorded on the tape. The magnetic coating on a videotape is formulated to work with particular camcorders and VTRs. If you choose the wrong tape coating, the tape can actually clog the video record heads of your video equipment, leading to video signal *dropouts* during recording and playback. Always read the documentation that comes with your video equipment before purchasing a particular brand of videotape stock.

About File-Based Media

Historically, video footage has been recorded on videotape. As digital acquisition quickly replaces analog technology, camcorders are starting to record footage as files on non-tape-based media such as hard disks, solid-state cards, and optical discs.

Today, some common file-based media formats include:

- Panasonic P2 cards (solid state)
- Sony Video Disk Unit devices (hard disk)
- Sony XDCAM and XDCAM HD (optical disc)

Video Standards

For the last 50 years, there have been two major signal types recorded on videotape: NTSC and PAL. With the emergence of new high definition (HD) video formats, NTSC and PAL formats are now referred to as *standard definition* (SD) video formats.

Standard Definition Video

National Television Systems Committee (NTSC) is the television and video standard used in most of the Americas, Taiwan, Japan, and Korea. Phase Alternating Line (PAL) is the television and video standard used in most of Europe, Australia, India, Brazil, China, and many African countries. There are several variations of both NTSC and PAL used in different parts of the world, but these variations are not described here.

Standard	Lines per frame	Frame rates	Scanning method
NTSC	525	29.97 fps	Interlaced
PAL	625	25 fps	Interlaced

SECAM is a video standard based on PAL. It is used in France, Poland, Haiti, and Vietnam. SECAM is strictly an analog composite video standard, so it is not used in digital video editing. Post-production work for SECAM broadcast is usually done in PAL and later converted to SECAM.

Note: SECAM is not supported by Final Cut Pro.

SD formats almost always have an aspect ratio of 4:3 (1.33:1).

High Definition Video

In the late 1990s, HD video formats were standardized in the United States by the Advanced Television Standards Committee (ATSC). These HD video formats are the next generation of broadcast and recording video formats. Unlike SD formats, which are restricted to fixed frame rates and numbers of lines per frame, HD video provides several options per format. While the increased flexibility is convenient, it also makes format interchange more complicated. Simply saying "HD video" is not enough; you need to define the frame size, frame rate, and scanning method of your HD format.

Standard	Frame size	Frame rates	Scanning method
720p	1280 x 720	23.98, 29.97, 59.94 24, 30, 60[1] 25, 50	Progressive
1080p	1920 x 1080	23.98, 29.97 24, 30 25	Progressive
1080i	1920 x 1080	25 (50i), 29.97 (59.94i), 30 (60i)	Interlaced

[1] 720p footage recorded at 24, 30, and 60 fps is rare. The 29.97 fps rates are more common because they are compatible with NTSC equipment.

There are an increasing number of HD tape and file-based formats available. Most HD formats support only a subset of the options shown in the table above, and most camcorders and video decks do not support every combination.

Types of Video Signals

Video signals are separated into several channels for recording and transmission. There are different methods of color channel separation, depending on the video format and its historical origins. For example, broadcast video devices were originally designed for black-and-white video, and color was added later. This is still evident in today's video formats that break image information into separate black-and-white and color information. On the other hand, video and image processing on computers is more flexible and developed later, so a three-color RGB model was adopted instead of a luma-chroma model.

The luma (black-and-white channel) and chroma (color channels) information can be recorded and transmitted several different ways in a video signal.

- *RGB (Red, Green, Blue):* This is the native format for most computer graphics and video files. This signal is also used inside traditional color CRTs, video cameras, flat-panel displays, and video projectors. Red, green, and blue signals can be combined to make any color, as well as grayscale images ranging from black (no signal on any channel) to white (full signal on every channel). RGB signals do not have a separate luma channel, because black-and-white signals can be represented by equal amounts of R, G, and B signals.
- *Component YUV, or $Y'C_BC_R$:* This three-channel signal has a luma (Y') signal and two *color difference* channels (C_B and C_R)[1]. Component video was invented in the 1950s as a way of making color television signals compatible with existing black-and-white televisions. Black-and-white televisions could use the luma signal, while color televisions could convert Y', C_B, and C_R back to RGB for display.

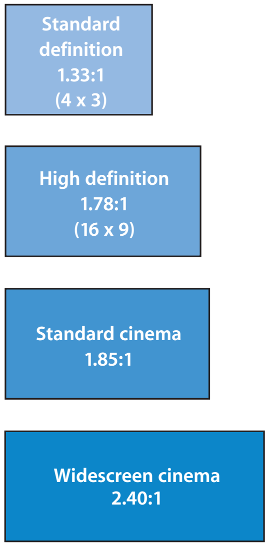
  The luma signal is derived by combining R, G, and B signals in proportions similar to the way human vision perceives those three colors. Therefore, the luma signal approximates the way the human eye sees brightness in color images. Humans are most sensitive to the green portion of the visible spectrum, and therefore the luma channel mostly consists of the green channel. The color difference channels are so named because they are derived from RGB by subtracting signals from the luma channel for each of the color channels (for example, R-Y or B-Y).
- *S-Video (Y/C):* An S-Video signal is also considered a component video signal because the luma and chroma signals are separate. However, the C signal is derived by combining the C_B and C_R component signals, which reduces the quality of the color channel compared to $Y'C_BC_R$.
- *Composite:* The luma (Y′) and chroma (C) signals are combined into a single composite video signal for broadcast. The chroma signal is placed on a *color subcarrier* frequency related to the main luma frequency. This method of superimposing color information on top of the black-and-white information indicates that this format originated in the early days of color television, when black-and-white TV compatibility was critical for widespread adoption.

  Black-and-white televisions are unaware of the color subcarrier, and so only the luma (Y′) channel is shown. Color televisions reverse the composite process, re-creating the $Y'C_BC_R$ component signal and then the RGB signal for display. Because the chroma and luma channels are superimposed, they do not separate perfectly, causing artifacts in the resulting image.

[1] The pair of color difference channels has different names depending on the particular format but serves a similar function in all formats. Some common names for color difference channels include C_B, C_R; R-Y, B-Y; and U,V.

Aspect Ratio of the Video Frame

The ratio of horizontal to vertical dimensions of a film or video frame is called the *aspect ratio*. Aspect ratio is independent of absolute image size or resolution.



Aspect ratio can be expressed as absolute dimensions (4 x 3), a ratio (4:3), a fraction (4/3), or as the decimal equivalent of a ratio (1.33:1, or simply 1.33).

- Video aspect ratios are often written as ratios, such as 4:3 for SD video or 16:9 for HD video.
- Film aspect ratios are often written as decimal equivalents, such as 1.33, 1.85, and 2.40. The higher the decimal number, the wider the image. An aspect ratio of 2.40 is wider than 1.85, and 1.85 is wider than 1.33.
- Digital video resolutions are usually written as absolute pixel dimensions, such as 720 x 480, 1280 x 720, 1920 x 1080, and so on.

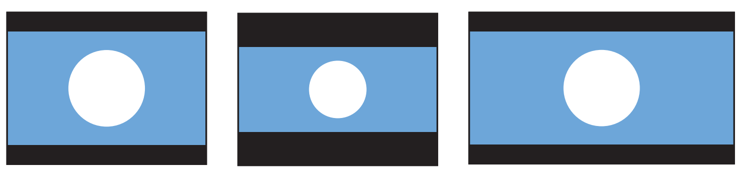
Below is a list of commonly used aspect ratios, mostly from the film and television industry, plus a few others for comparison.

Aspect ratio	Medium
1.33 (4:3)	Early 35 mm film and SD television
1.37	4-perforation 35 mm camera footage (prior to projection)—also known as "Academy" aspect ratio
1.66 (15:9)	Standard European film; masked in projector
1.78 (16:9)	HD television
1.85	Standard North American and UK film; masked in projector
2.40 (also referred to as *2.35* and *2.39*)	Widescreen (anamorphic) film projection

Footage with different aspect ratios can be combined using a variety of techniques.

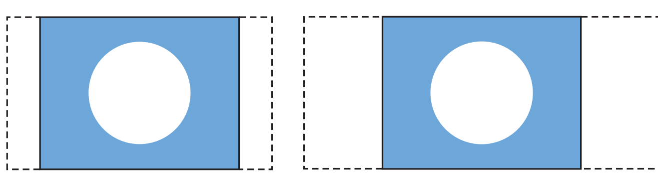
Letterboxing

Letterboxing preserves the aspect ratio of widescreen movies on a narrower screen. The movie is scaled until it fits within the width of the screen, resulting in blacks bars at the top and bottom of the frame.



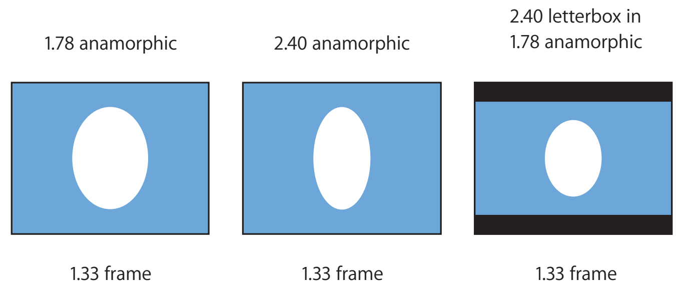
Pan and Scan

The pan and scan technique crops widescreen movies to fit on a narrower screen. In some cases, artificial camera moves may even be added to show the entire content of a widescreen frame. Pan and scan does not preserve the aspect ratio of widescreen movies.



Anamorphic

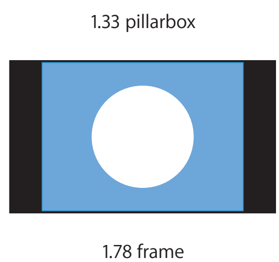
Anamorphic techniques use special lenses or electronics to squeeze a widescreen image to fit in a narrower aspect ratio. During projection or playback, the squeezed image is stretched back to its original widescreen aspect ratio.



Anamorphic video can also be letterboxed. For example, 16:9 anamorphic DVDs may contain letterboxed 2.40 aspect ratio footage.

Pillarboxing

Pillarboxing displays movies with a small aspect ratio on a wide screen. Black bars appear on the left and right sides of the frame.



Frame Dimensions, Number of Lines, and Resolution

A video frame is composed of lines. In digital video, each line is sampled to create a number of pixels (samples) per line. The more lines per frame, the higher the image resolution. The more pixels per line, the higher the resolution of each line.

Number of Lines

NTSC uses 525 lines, whereas PAL uses 625. In analog video, many lines are not actually used for picture information, so the total numbers relevant for the picture are somewhat smaller: 486 lines for NTSC and 576 lines for PAL. HD formats defined by the ATSC have either 1080 or 720 active picture lines per frame.

Pixels (Samples) per Line

In digital video formats, each line is sampled a number of times. In an attempt to create a single digital VTR that could digitize and record both NTSC and PAL signals, the ITU-R BT. 601 specification uses 720 samples per line for both NTSC and PAL video. Therefore, a digital NTSC video frame is 720 pixels x 486 lines, and a PAL video frame is 720 pixels x 576 lines.

HD video with 1080 lines uses 1920 pixels per line (1920 x 1080). 720-line HD video uses 1280 pixels per line (1280 x 720). Both of these formats have an aspect ratio of 16:9.

Common video frame sizes are shown in the table below.

Width	Height	Pixel aspect ratio	Screen aspect ratio	Description
320	240	1:1	4:3	Used for web distribution or offline video editing.
640	480	1:1	4:3	An early standard for analog-to-digital video editing, and an ATSC video specification.
720[1]	480	Height greater than width	4:3	NTSC DV and DVD image dimensions. Also part of the ATSC video specification.
720[1]	486	Height greater than width	4:3	NTSC SD video dimensions used for professional digital formats such as Digital Betacam, D-1, and D-5.
720[1]	576	Width greater than height	4:3	PAL SD video dimensions used for digital formats such as Digital Betacam, D-1, and D-5, as well as DVD and DV.
1280	720	1:1	16:9	An HD video format, capable of higher frame rates in exchange for smaller image dimensions.
1920	1080	1:1	16:9	An HD video format with very high resolution.
960	720	4:3	16:9	Some 720p formats (such as DVCPRO HD and HDV) subsample 1280 pixels to 960 to minimize the data rate.
1440 1280	1080	4:3 3:2	16:9	Some 1080-line formats (such as HDV and DVCPRO HD) subsample 1920 pixels to 1440 or even 1280 to minimize the data rate.

[1] In most video devices, only 704 or 708 pixels are actually used for picture information.

720 x 486 Versus 720 x 480

One issue that comes up during post-production is the subtle difference between NTSC SD formats that use 486 lines per frame (such as Digital Betacam, D-1, and D-5) and formats that use 480 lines per frame (such as DV, DVCPRO, and DVD). Why is there this subtle difference? The reason is simple: 480 is divisible by 16, and 486 isn't. Divisibility by 16 is convenient for certain block-based codecs because each frame is broken into 16 x 16 pixel blocks (known as *macroblocks*) during compression.

The only time this should be a concern is when you are converting between a 486-line format like Digital Betacam and a 480-line format like DVD. However, the extra six lines are typically not visible on an analog television.

Pixel Aspect Ratio

A *pixel* usually refers to a physical picture element that emanates light on a video display. But a pixel is also a term for a *sample* of light intensity—a piece of data for storing luma or chroma values. When stored on tape or on hard disk, the intensity of a pixel has no inherent shape, height, or width; it is merely a data value. For example, one pixel may have a value of 255, while another may have a value of 150. The value of each pixel determines the intensity of a corresponding point on a video display. In an ideal world, all pixels would be captured and displayed as squares (equal height and width), but this is not always the case.

The ITU-R BT. 601 specification makes it possible to transmit either NTSC or PAL information in a single signal. To achieve this goal, both NTSC and PAL video lines are sampled 720 times. In both NTSC and PAL, the frame displayed has an aspect ratio of 4:3, yet neither 720 x 486 nor 720 x 576 constitutes a 4:3 ratio. The solution to this problem is to display the pixels (the samples of light intensity) taller-than-wide, or wider-than-tall, so that they fit into a 4:3 frame. This results in the concept of "rectangular pixels"—pixels that must be stretched or squeezed to fit in the 4:3 frame. Most SD video devices actually use 704 or 708 pixels for picture information but stretch these pixels to 720 when recording to tape.

None of this was obvious in the days of linear editing, when video was simply copied from one tape to another, because the video equipment always compensated automatically. However, as people began using computers to work with video, digital video captured to the computer looked distorted (squashed vertically or stretched horizontally) because the computer displayed the pixels as squares, without compensating.

Some video formats use rectangular pixels to reduce the amount of information stored on tape. For example, DVCPRO HD effectively records 1280 pixels per line (when using the 720p format), but to save space on tape, the intensity of every 1.33 pixels is averaged together (a process known as *subsampling*) and only 960 pixels are recorded. These pixels are not representing a square area, but a wider, rectangular portion of each video line. This results in a 4-to-3 reduction in the amount of information recorded on tape.

Video and image editing programs like Final Cut Pro and Photoshop must compensate for these rectangular pixels so they appear correctly on a computer display. However, there are several different pixel aspect ratios in use, and there is unfortunately no single accepted standard in the industry. The exact aspect ratio used may vary slightly from one software application to another, as well as among different third-party video interfaces.

These days, the biggest challenge comes when exchanging graphics between applications that use different pixel aspect ratios, or when using an application that does not support rectangular pixels with one that does. The key to a simple workflow is to use applications that can work at the native, non-square pixel image dimensions and compensate on the computer display. Fortunately, major video and graphics applications such as Photoshop, After Effects, Final Cut Pro, and DVD Studio Pro can work with graphics and video at native resolutions. This way, you are always working with the exact pixel dimensions that you will eventually output to videotape or DVD.

Frame Rate

The frame rate of any motion picture, whether film or video, defines how often pictures are taken per second. The higher the frame rate, the more accurately you capture moments in time and reduce flicker during playback. To achieve double the perceived frame rate (flicker), film projectors actually double or triple the shutter speed, even though the same frame is repeated two or three times, respectively. This is because a faster flicker creates more convincing motion. Video uses a similar, although more complicated, technique called *interlacing*. For more information about interlacing, see the next section, "Scanning Method." For more details about frame rate, see Appendix B, "Frame Rate and Timecode," on page 405.

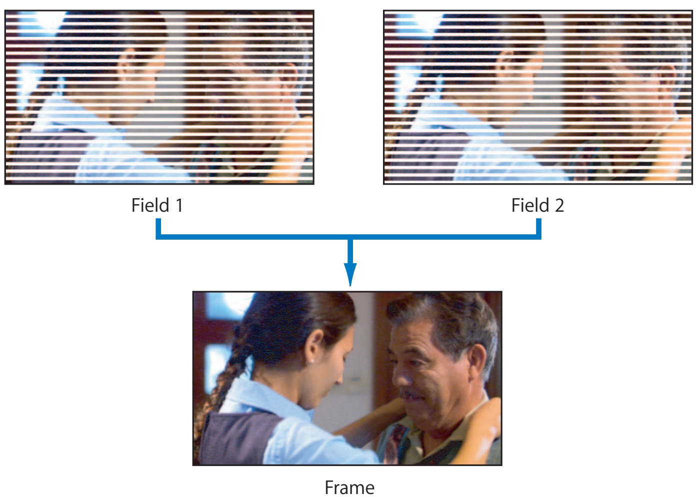
Scanning Method

A video frame is made of horizontal lines that are scanned from one side of a display to the other. *Progressive* video scanning happens when each line of a video frame is scanned, one after another. *Interlaced* scanning fills the entire frame with only half the lines, which requires half the time, thus doubling the perceived frame rate and reducing flicker.

About Interlaced Scanning

Frame rates lower than 40 fps can cause noticeable flicker. When NTSC and PAL were invented, faster frame rates were not practical to implement. Interlaced scanning was devised to create a perceived frame rate of 60 fps (NTSC) or 50 fps (PAL). Interlaced video scans the display twice, using two *fields*, to complete a single frame. A single field contains only the odd lines (1, 3, 5, 7, and so on) or the even lines (2, 4, 6, 8, and so on) of the frame. If you could stop the video scanning process to observe a single video field, you would see that every other line is missing, like venetian blinds or a comb.

Because the fields are changing at twice the frame rate, there is less perceived flicker than if each frame was scanned progressively. For example, with NTSC, a field of odd lines is scanned in 1/60 of a second and a field of even lines follows in the next 1/60 of a second, resulting in a complete frame every 1/30 of a second.



About Progressive Scanning

Progressive scanning is much simpler than interlaced scanning: each line is scanned consecutively until a complete frame is drawn. Computer displays and many recent HD televisions use progressive scanning.

Here are some significant facts about interlaced and progressive scanning methods:

- Interlacing provides twice the perceived frame rate with only half the recording or transmission requirements.
- Progressive scanning is preferred when interlacing artifacts (such as thin flickering horizontal lines) would be unacceptable. Progressive images are often considered more film-like because there are no flickering interlacing artifacts.
- Computer displays are almost always scanned progressively.
- NTSC and PAL televisions always use interlaced scanning.
- Many HD video cameras can record progressive frames.
- Video destined for computer-only use, such as web video, should always be made progressive.

About Field Dominance

Field dominance is an issue when recording and playing back interlaced video material. With progressive video, there is only one way to play back a video frame: start at line 1 and scan until the end of the last line. With interlaced video, the video player must know whether to scan the odd lines first, or the even lines. In other words, each time a frame is displayed, which field should be displayed first, field 1 or 2? The field displayed first is totally dependent on which field was captured by the camera and recorded first.

Each field is a snapshot in time, so if field 1 is recorded earlier in time than field 2, field 1 must be played back before field 2. If the wrong field order is chosen, each frame's fields play backward in time, even though each frame as a whole still moves forward. The effect is a very noticeable stutter happening 60 (NTSC) or 50 (PAL) times a second.

Each piece of video equipment and each video format has a preferred field dominance. This prevents you from, say, editing two field 2s back to back, and makes sure that each field is played back in the right order.

Setting Field Dominance in Final Cut Pro

In Final Cut Pro, the field dominance of clips must match the sequence field dominance. Otherwise, the fields stutter during playback because each pair of fields plays back in the wrong order. For example, DV NTSC and DV PAL always have a field dominance of Lower (Even). If you're working in a sequence and you see that imported clips are flickering, check to make sure the field dominance of those additional clips matches the field dominance of your edited sequence.

Important: You need to change the Field Dominance setting of your projects and sequences only if you change your video hardware setup.

In Final Cut Pro, there are two options for field dominance:

- Upper (field 2 is dominant, so the second field is drawn first)
- Lower (field 1 is dominant, so the first field is drawn first)

Generally, Upper is used by 640 x 480 systems, while Lower is most common in professional 720 x 486 and DV 720 x 480 systems.

Color Recording Method

The color recording method of a video format may be either RGB, $Y'C_BC_R$ (component), Y/C (S-Video), or composite. The more discrete channels a format has, the higher the quality of the image, but the more data required to store and transmit that information.

Color recording method	Video formats
Composite	1", 3/4" U-matic, 1/2",VHS, D-2, D-3
Y/C (S-Video)	Hi8, S-VHS
$Y'C_BC_R$ (component)	Betacam SP, Digital Betacam, DVD, DV, D-1, D-5
RGB	Computer graphics and digital cinema acquisition. Although video originates in and is displayed in this format, it is rare for tape formats ($Y'C_BC_R$ is used instead).

Today, almost all digital video formats are $Y'C_BC_R$ (component). Computers typically store image data using RGB, although many $Y'C_BC_R$ (component) formats can now be processed natively on the computer (such as DV).

Video Sample Rate and Bit Depth

The video sample rate of a digital video format determines how often the light intensity of each video line is sampled.

Sample rate	Description
74.25 MHz	HD video luma (Y´) sample rate.
37.125 MHz	HD video chroma (C_BC_R) sample rate. This is half of the luma sample rate, used for 4:2:2 HD video.
14.3 MHz	Early NTSC digital video recorders sampled video at exactly four times the frequency of the color subcarrier signal (3.58 MHz x 4). This is the origin of the 4 in color sample ratios such as 4:2:2.
13.5 MHz	This is the sample rate for the luma (Y´) channel for SD digital video. This sample rate was chosen to work with both NTSC and PAL digital video. The 4 in 4:2:2 is now represented by this sample rate.
6.75 MHz	This is the sample rate for the color difference channels in 4:2:2 video. This is half of 13.5 MHz.

Color Sample Ratio

Color sample ratio refers to the ratio of luma (Y´) samples to each color difference sample (C_B and C_R). For example, 4:2:2 color sampling means that for every four pixels of Y´ data stored, only two C_R samples and two C_B samples are stored. By reducing the number of chroma samples, less color detail is recorded and less bandwidth is required for storage and transmission. Because we are less sensitive to color detail than we are to luma detail, subsampling the chroma signal can be considered perceptually lossless. In absolute terms, chroma subsampling can make processes like chroma keying much more difficult.

Sampling ratio	Description
4:4:4	Each R, G, and B channel, or each Y´, C_B, and C_R channel, is sampled at the same rate. Maximum color detail is maintained.
4:4:4:4	Full sample rate for each color channel, plus a fourth alpha channel at the full sample rate.
4:2:2	The color channels are subsampled so that the color resolution is halved. For example, the first pixel in a line contains Y´, C_B, and C_R samples. The next pixel contains only a Y´ sample. This pattern repeats. Most professional video formats use 4:2:2 color subsampling.
4:2:2:4	4:2:2 sample rate for each color channel, plus an alpha channel at the full sample rate.

Sampling ratio	Description
4:1:1	The color is subsampled so that the color resolution is quartered. The first pixel in a line contains Y´, C_B, and C_R samples. The next three pixels only contain Y´ samples. This pattern repeats.
4:2:0	This ratio indicates that the C_B and C_R channels are subsampled both horizontally (as in 4:2:2) and vertically. This reduces color resolution in both the horizontal and vertical dimensions compared to 4:2:2, which only reduces horizontal chroma resolution. There are several methods for locating C_B and C_R samples relative to Y´ samples, yielding several different 4:2:0 formats.

The following table shows a list of color sample ratios used in various digital video formats:

Sample ratio	Video formats
4:4:4	HDCAM SR Most RGB computer graphics files (implicit)
4:2:2	Digital Betacam, D-1, D-5, DVCPRO HD, DVCPRO 50, and HDCAM SR
3:1:1	HDCAM
4:1:1	NTSC DV, NTSC DVCAM, and DVCPRO
4:2:0	PAL DV, PAL DVCAM, DVD, and HDV

Bit Depth

The number of bits used per sample determines how accurately the sample is stored, as well as how much intensity variation is possible within the signal. For example, a video signal with a bit depth of only 1 bit can have either a value of 0 or 1, resulting in only black or white pixels. Two bits per sample results in four possible values: 00, 01, 10, or 11, or any of four shades of gray (or some other color) per sample.

Most digital video formats use a minimum of 8 bits per color channel, or 256 gradations of intensity. RGB images are traditionally described by the total bits used per pixel (8 bits per channel x 3 channels = 24 bits). 32-bit RGB images usually have 24-bit color plus 8 more bits for an alpha channel.

Note: Still images using 16 bits per color channel, or 48 bits per RGB pixel, are becoming more common. However, most video formats use 8 or 10 bits per color channel.

Video signal bit depth is usually described per channel. For example, DV and DVCPRO HD use 8 bits per color component (in other words, 8 bits for Y´, 8 bits for C_B, and 8 bits for C_R). Other formats, such as D-5, use 10 bits per component. This provides 1024 possible gradations instead of 256, which means much more subtle variations in intensity can be recorded.

In fact, 8-bit $Y'C_BC_R$ video does not use all 256 codes to represent picture information. Black is stored as code 16 and white is code 235. Codes 1–15 and 236–254 are retained for footroom and headroom, respectively. These areas allow for quick spikes in the signal caused by filtering in analog-to-digital conversions and, in the case of white levels, can prevent clipping for highlights that may exceed 235 (white). Levels above 235 are sometimes referred to as *super-white levels*. For more information about super-white levels, see Volume III, Chapter 29, "Rendering and Video Processing Settings."

Internally, Final Cut Pro can do pixel calculations using 32-bit floating-point precision, which allows for very accurate calculations without rounding errors. This leads to much more accurate color reproduction when applying filters and compositing layers of video. This is especially important when you are going to show your movie on film or broadcast-quality video monitors. In Final Cut Pro, the Video Processing tab in the Sequence Settings window allows you to choose the rendering bit depth for a sequence. For more information, see Volume III, Chapter 29, "Rendering and Video Processing Settings."

Perceptual Coding and Gamma

The limited number of brightness steps in 8-bit digital video requires efficient use of the 256 available codes. Because perception of brightness follows a power law function, humans are more sensitive to absolute intensity changes in dark areas. In other words, the amount of light required to make a perceptual shift in brightness increases exponentially. Therefore, a gamma correction is applied to video so that the step between each code represents a perceptual shift in brightness. Without this gamma correction, the darker areas would appear to abruptly jump from one brightness level to the next ("banding") and white levels would waste many codes with imperceptible brightness shifts. This gamma correction is reversed by video monitors so that the viewer sees the original light intensity of the image. For more information about gamma, see Volume III, Chapter 29, "Rendering and Video Processing Settings."

Video Compression

Once a video signal is digital, it requires a large amount of storage space and transmission bandwidth. To reduce the amount of data, several strategies are employed that compress the information without negatively affecting the quality of the image. Some methods are *lossless*, meaning that no data is lost, but most are *lossy*, meaning that information is thrown away that can't be retrieved.

Some simple methods of data compression are:

- *Throw away pixels at regular intervals:* This essentially scales the image, or makes it more blocky.
- *Average several pixel values together (subsampling):* This involves taking several adjacent pixel values and averaging them together, resulting in a single rectangular pixel that approximates the value of several. For more information, see "Pixel Aspect Ratio" on page 384.
- *Throw away color channel information at regular intervals:* This results in color sample ratios like 4:2:2 and 4:1:1. Ideally, throwing away this color information is not noticeable to the viewer, but it may be a problem if you are trying to do detailed color correction or chroma keying that requires a lot of color information to start with.

Lossless Codecs

Once these basic methods have been employed, much more intensive algorithms can be employed to reduce the amount of transmitted and stored image data. Mathematical algorithms can be used to encode and decode each video frame. These codecs (such as enCode and Decode) must be installed in the VTR or software you are using to play back your video. For example, QuickTime supports many different video codecs for video export and playback.

The simplest encoding algorithm, called *run-length encoding*, represents strings of redundant values as a single value and a multiplier. For example, consider the following bit values:

0000000000000000000000001111111111111111000000000000000000000000

Using run-length encoding on the bit values above can reduce the amount of information to:

0 x 24, 1 x 16, 0 x 24

Or in binary:

0 [11000], 1 [10000], 0 [11000]

In the example above, the original 64 bits can be transmitted using only 18 bits.

Run-length encoding is lossless, because all the information is retained after decoding. This technique is particularly useful for computer graphics applications, because there are often large fields of identical colors.

Note: If each bit in the original image were to alternate between 0 and 1, run-length encoding would not only be ineffective, it could actually make the overall data rate higher! Each codec is designed to anticipate and compress different kinds of data patterns. For example, a codec designed for audio compression is not useful for video compression, which has very different data patterns.

Lossy Codecs

Most video codecs are necessarily lossy, because it is usually impractical to store and transmit uncompressed video signals. Even though most codecs lose some information in the video signal, the goal is to make this information loss visually imperceptible. When codec algorithms are developed, they are fine-tuned based on analyses of human vision and perception. For example, if the human eye cannot differentiate between lots of subtle variation in the red channel, a codec may throw away some of that information and viewers may never notice.

Many formats, including JPEG and all varieties of DV, use a fairly complicated algorithm called *DCT encoding*. Another method, called *wavelet compression*, is starting to be used for popular codecs, such as the Apple Pixlet video codec. DVDs, modern digital television, and formats such as HDV use MPEG-2 compression, which not only encodes single frames (intraframe, or spatial compression) but encodes multiple frames at once (interframe, or temporal compression) by throwing away data that is visually redundant over time.

About Uncompressed Video

Video that has no compression applied can be unwieldy, so it is only used for the highest-quality video work, such as special effects and color correction at the last stage of a project. Most professional projects have an *offline* phase that uses compressed video and then an *online*, finishing phase that uses uncompressed video recaptured at full resolution. Uncompressed video requires expensive VTRs and large, high-speed hard disks.

About MPEG Compression

MPEG encoding is based on eliminating redundant video information, not only within a frame but over a period of time. In a shot where there is little motion, such as an interview, most of the video content does not change from frame to frame, and MPEG encoding can compress the video by a huge ratio with little or no perceptible quality loss.

MPEG compression reduces video data rates in two ways:

- *Spatial (intraframe) compression:* Compresses individual frames.
- *Temporal (interframe) compression:* Compresses groups of frames together by eliminating redundant visual data across multiple frames.

Intraframe Compression

Within a single frame, areas of similar color and texture can be coded with fewer bits than the original, thus reducing the data rate with minimal loss in noticeable visual quality. JPEG compression works in a similar way to compress still images. Intraframe compression is used to create standalone video frames called *I-frames* (short for *intraframe*).

Interframe Compression

Instead of storing complete frames, temporal compression stores only what has changed from one frame to the next, which dramatically reduces the amount of data that needs to be stored while still achieving high-quality images.

Groups of Pictures

MPEG formats use three types of compressed frames, organized in a *group of pictures*, or *GOP*, to achieve interframe compression:

- *I-frames:* Intra (I) frames, also known as *reference* or *key frames*, contain all the necessary data to re-create a complete image. An I-frame stands by itself without requiring data from other frames in the GOP. Every GOP contains one I-frame, although it does not have to be the first frame of the GOP. I-frames are the largest type of MPEG frame, but they are faster to decompress than other kinds of MPEG frames.
- *P-frames:* Predicted (P) frames are encoded from a "predicted" picture based on the closest preceding I- or P-frame. P-frames are also known as *reference frames*, because neighboring B- and P-frames can refer to them. P-frames are typically much smaller than I-frames.
- *B-frames:* Bi-directional (B) frames are encoded based on an interpolation from I- and P-frames that come before and after them. B-frames require very little space, but they can take longer to decompress because they are reliant on frames that may be reliant on other frames. A GOP can begin with a B-frame, but it cannot end with one.

GOPs are defined by three factors: their pattern of I-, P-, and B-frames, their length, and whether the GOP is "open" or "closed."

GOP Pattern

A GOP pattern is defined by the ratio of P- to B-frames within a GOP. Common patterns used for DVD are IBP and IBBP. All three frame types do not have to be used in a pattern. For example, an IP pattern can be used. IBP and IBBP GOP patterns, in conjunction with longer GOP lengths, encode video very efficiently. Smaller GOP patterns with shorter GOP lengths work better with video that has quick movements, but they don't compress the data rate as much.

Some encoders can force I-frames to be added sporadically throughout a stream's GOPs. These I-frames can be placed manually during editing or automatically by an encoder detecting abrupt visual changes such as cuts, transitions, and fast camera movements.

GOP Length

Longer GOP lengths encode video more efficiently by reducing the number of I-frames but are less desirable during short-duration effects such as fast transitions or quick camera pans. MPEG video may be classified as *long-GOP* or *short-GOP.* The term *long-GOP* refers to the fact that several P- and B-frames are used between I-frame intervals. At the other end of the spectrum, short-GOP MPEG is synonymous with I-frame–only MPEG. Formats such as IMX use I-frame–only MPEG-2, which reduces temporal artifacts and improves editing performance. However, I-frame–only formats have a significantly higher data rate because each frame must store enough data to be completely self-contained. Therefore, although the decoding demands on your computer are decreased, there is a greater demand for scratch disk speed and capacity.

Maximum GOP length depends on the specifications of the playback device. The minimum GOP length depends on the GOP pattern. For example, an IP pattern can have a length as short as two frames.

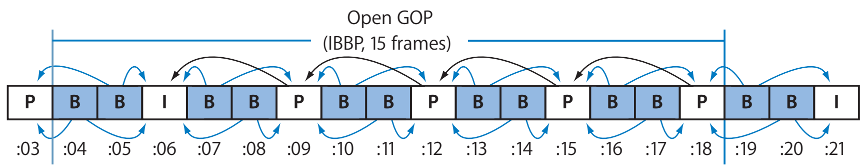
Here are several examples of GOP length used in common MPEG formats:

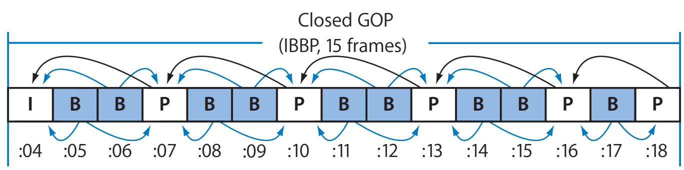
- *MPEG-2 for DVD:* Maximum GOP length is 18 frames for NTSC or 15 frames for PAL. These GOP lengths can be doubled for progressive footage.
- *1080-line HDV:* Uses a long-GOP structure that is 15 frames in length.
- *720-line HDV:* Uses a six-frame GOP structure.
- *IMX:* Uses only I-frames.

Open and Closed GOPs

An open GOP allows the B-frames from one GOP to refer to an I- or P-frame in an adjacent GOP. Open GOPs are very efficient but cannot be used for features such as multiplexed multi-angle DVD video. A closed GOP format uses only self-contained GOPs that do not rely on frames outside the GOP.

The same GOP pattern can produce different results when used with an open or closed GOP. For example, a closed GOP would start an IBBP pattern with an I-frame, whereas an open GOP with the same pattern might start with a B-frame. In this example, starting with a B-frame is a little more efficient because starting with an I-frame means that an extra P-frame must be added to the end (a GOP cannot end with a B-frame).





MPEG Containers and Streams

MPEG video and audio data are packaged into discrete data containers known as *streams.* Keeping video and audio streams discrete makes it possible for playback applications to easily switch between streams on the fly. For example, DVDs that use MPEG-2 video can switch between multiple audio tracks and video angles as the DVD plays.

Each MPEG standard has variations, but in general, MPEG formats support two basic kinds of streams:

- *Elementary streams:* These are individual video and audio data streams.
- *System streams:* These streams combine, or multiplex, video and audio elementary streams together. They are also known as *multiplexed streams.* To play back these streams, applications must be able to demultiplex the streams back into their elementary streams. Some applications only have the ability to play elementary streams.

MPEG-1

MPEG-1 is the earliest format specification in the family of MPEG formats. Because of its low bit rate, MPEG-1 has been popular for online distribution and in formats such as Video CD (VCD). DVDs can also store MPEG-1 video, though MPEG-2 is more commonly used. Although the MPEG-1 standard actually allows high resolutions, almost all applications use NTSC- or PAL-compatible image dimensions at quarter resolution or lower.

Common MPEG-1 formats include 320 x 240, 352 x 240 at 29.97 fps (NTSC), and 352 x 288 at 25 fps (PAL). Maximum data rates are often limited to around 1.5 Mbps. MPEG-1 only supports progressive-scan video.

MPEG-1 supports three layers of audio compression, called *MPEG-1 Layers 1, 2, and 3*. MPEG-1 Layer 2 audio is used in some formats such as HDV and DVD, but MPEG-1 Layer 3 (also known as *MP3*) is by far the most ubiquitous. In fact, MP3 audio compression has become so popular that it is usually used independently of video.

MPEG-1 elementary stream files often have extensions such as .m1v and .m1a, for video and audio, respectively.

MPEG-2

The MPEG-2 standard made many improvements to the MPEG-1 standard, including:

- Support for interlaced video
- Higher data rates and larger frame sizes, including internationally accepted standard definition and high definition profiles
- Two kinds of multiplexed system streams—Transport Streams (TS) for unreliable network transmission such as broadcast digital television, and Program Streams (PS) for local, reliable media access (such as DVD playback)

MPEG-2 categorizes video standards into MPEG-2 Profiles and MPEG-2 Levels. Profiles define the type of MPEG encoding supported (I-, P-, and B-frames) and the color sampling method used (4:2:0 or 4:2:2 $Y'C_BC_R$). For example, the MPEG-2 Simple Profile (SP) supports only I and P progressive frames using 4:2:0 color sampling, whereas the High Profile (HP) supports I, P, and B interlaced frames with 4:2:2 color sampling.

Levels define the resolution, frame rate, and bit rate of MPEG-2 video. For example, MPEG-2 Low Level (LL) is limited to MPEG-1 resolution, whereas High Level (HL) supports 1920 x 1080 HD video.

MPEG-2 formats are often described as a combination of Profiles and Levels. For example, DVD video uses Main Profile at Main Level (MP @ ML), which defines SD NTSC and PAL video at a maximum bit rate of 15 (though DVD limits this to 9.8 Mbps).

MPEG-2 supports the same audio layers as MPEG-1 but also includes support for multichannel audio. MPEG-2 Part 7 also supports a more efficient audio compression algorithm called *Advanced Audio Coding*, or *AAC*.

MPEG-2 elementary stream files often have extensions such as .m2v and .m2a, for video and audio, respectively.

MPEG-4

MPEG-4 inherited many of the features in MPEG-1 and MPEG-2 and then added a rich set of multimedia features such as discrete object encoding, scene description, rich metadata, and digital rights management (DRM). Most applications support only a subset of all the features available in MPEG-4.

Compared to MPEG-1 and MPEG-2, MPEG-4 video compression (known as *MPEG-4 Part 2*) provides superior quality at low bit rates. However, MPEG-4 also supports high-resolution video as well. For example, Sony HDCAM SR uses a form of MPEG-4 compression.

MPEG-4 Part 3 defines and enhances AAC audio originally defined in MPEG-2 Part 7. Most applications today use the terms *AAC audio* and *MPEG-4 audio* interchangeably.

MPEG-4 Part 10, or H.264

MPEG-4 Part 10 defines a high-quality video compression algorithm called *Advanced Video Coding* (AVC). This is more commonly referred to as *H.264*. H.264 video compression works similarly to MPEG-1 and MPEG-2 encoding but adds many additional features to decrease data rate while maintaining quality. Compared to MPEG-1 and MPEG-2, H.264 compression and decompression require significant processing overhead, so this format may tax older computer systems.

Video Formats Supported by Final Cut Pro

Final Cut Pro supports any video format that uses an installed QuickTime codec. QuickTime natively supports codecs used by a number of video devices, such as DV, DVCPRO 50, DVCPRO HD, HDV, and IMX devices. With these formats, the distinction between file format and tape format is blurred, and transferring from tape to hard disk or other media is essentially a file transfer, allowing you to edit footage natively. For more information, choose HD and Broadcast Formats from the Final Cut Pro Help menu.

When you work with videotape formats such as Digital Betacam, D-5, Betacam SP, and so on, you need a third-party video interface to connect to the SDI or analog component video connectors on the deck. In this case, the video interface must convert the incoming or outgoing video signal to or from a QuickTime codec. Many video interfaces come with codecs for high-quality compressed and uncompressed editing.

DV Formats

You can easily capture and output any DV-format video via the FireWire port on your computer. Video, audio, timecode, and device control data are all transferred via a single FireWire cable. FireWire (also referred to as *IEEE 1394* or *i.LINK*) is a high-speed technology for connecting and transmitting data to and from various external devices, such as video and audio interfaces, hard disks, and digital cameras. FireWire is supported by many professional and consumer-level DV camcorders and decks.

Digital format	Maker	Color sample ratio	Compression ratio	Recorded bit rate
DV (25)	Multiple manufacturers	4:1:1 4:2:0 (PAL)	5:1	25 Mbps
DVCAM	Sony	4:1:1 4:2:0 (PAL)	5:1	25 Mbps
DVCPRO (D-7)	Panasonic	4:1:1 (NTSC and PAL)	5:1	25 Mbps
DVCPRO 50	Panasonic	4:2:2	3.3:1	50 Mbps
DVCPRO HD	Panasonic	4:2:2	6.7:1	100 Mbps

24p Video

Formats that capture complete (progressive) video frames at 24 frames per second have received a lot of attention lately. This is because 24p video uses the same frame rate as film, and it scans images progressively. For example, a 24 fps, 1920 x 1080, progressively scanned video format closely matches the resolution of a 35 mm film theater distribution print. For the first time since the invention of television, moviemakers can choose video instead of film without suffering significant resolution loss or having to cope with frame-rate conversions.

There are many ways to record 24p video within other frame rates. For more information, see Appendix C, "Working with 24p Video," on page 417.

High Definition Video Formats

Final Cut Pro has native support for HD formats such as HDV, DVCPRO HD, and XDCAM HD. For other HD formats, you need an appropriate third-party capture interface and hard disks with sufficient speed and capacity. HD formats are often defined by their vertical resolutions (number of lines), scanning method (interlaced versus progressive), and frame or field rate. For example, 1080i60 HD video has 1080 lines per frame, uses interlaced scanning (indicated by the *i*), and scans 60 fields per second.

Scanning Methods

Most HD equipment can record both progressive and interlaced video. Typically, 1080-line video is interlaced (1080i) and 720-line video is progressive (720p). Several 1080p formats exist, such as 1080p24, but there are no 720-line interlaced formats. For more information, see "About Interlaced Scanning" and "About Progressive Scanning" on page 387.

Compressed High Definition Formats

Because of the high data rate generated by HD video cameras, most HD formats compress the image data to fit on tape. For example:

- DVCPRO HD; also generally referred to as *DV-100* (in reference to its bit-rate of 100 Mbps)
- D-9 HD, an extension of the Digital S format (Digital S is designated SMPTE D9)
- D-5 HD, an extension of the D-5 format
- HDCAM and HDCAM SR

Format	Manufacturer	Color sample ratio	Bit depth	Recorded data rate
D-5 HD	Panasonic	4:2:2	8-bit 10-bit	235 Mbps
D-6	Philips, Toshiba	4:2:2	10-bit	1.2 Gbps
HDCAM	Sony	3:1:1	8-bit (internal) 10-bit (in/out)	143 Mbps
HDCAM SR	Sony	4:2:2 4:4:4	10-bit log 10-bit linear	440 Mbps (SQ) 880 Mbps (HQ)
DVCPRO HD	Panasonic	4:2:2	8-bit	100 Mbps
XDCAM HD	Sony	4:2:2	8-bit	35 Mbps (LP) 25 Mbps (SP) 18 Mbps (HQ)
HDV	Sony, JVC, Canon	4:2:0	8-bit	19 Mbps (720) 25 Mbps (1080)
RGB video • 1080p30 • 720p60	n/a (computer graphics)	4:4:4	8 bits per color channel	1.39 Gbps 1.24 Gbps

Note: The data rates shown here are approximate. For purposes of determining hard disk capacity for capture, carefully research the details of the format you are using.

Uncompressed High Definition Formats

HD requires extremely high data rates (around 1.4 Gbps). There are no camcorder formats currently available for recording uncompressed HD video. High-capacity, general-purpose digital tape formats like D-6 can be used in combination with camera heads and digital telecine machines capable of outputting uncompressed RGB and component HD video data. High-speed disk arrays can also be used to record uncompressed HD video.

Data Rate Comparisons

The following table is useful when preparing to capture video to a particular codec on your hard disk.

Format	Typical data rate
OfflineRT (using Photo JPEG)	Varies from 300–500 KB/sec.
25:1 compressed M-JPEG	1 MB/sec.
DV-25	3.6 MB/sec.
DVCPRO 50	7.2 MB/sec.
2:1 compressed M-JPEG	12 MB/sec.
Uncompressed SD video	24 MB/sec.
Uncompressed 8-bit 1080i 29.97 fps HD video	121.5 MB/sec.
Uncompressed 10-bit 1080i 29.97 fps HD video	182.3 MB/sec.

A Brief History of Film, Television, and Audio Formats

The timeline below helps to illustrate the constantly evolving list of media formats as well as developmental peaks and valleys.

Year	Event
1826	First photograph is taken.
1877	Thomas Edison makes the first sound recording of "Mary Had a Little Lamb."
1879	Thomas Edison invents commercially viable incandescent light bulbs.
1888	Heinrich Hertz shows that electricity can travel through space and that radio waves are physically identical to light.
1889	35 mm film is invented by splitting Eastman Kodak 70 mm in half (1.33 aspect ratio).
1895	Marconi develops radio transmitter and receiver.

Year	Event
1895	Lumière brothers demonstrate combination camera/projector (16 fps).
1918	First color motion picture appears.
1920	Commercial radio broadcasts begin.
1923	16 mm film is introduced.
1927	First major motion picture with sound is released (1.37 aspect ratio), ending the silent movie era.
1932	BBC begins official monochrome, 30-line video broadcast.
1934	RCA experiments with 343-line, 30 fps television format, removing flicker by introducing interlacing.
1936	BBC begins broadcasting a high definition, monochrome, 405-line, 25 fps interlaced signal tied to European 50 Hz electrical frequency.
1939	NBC begins regularly scheduled broadcasts of electronic television, 441 lines and 30 fps.
1941	National Television Systems Committee (NTSC) standardizes U.S. commercial television format, 525 lines, 30 fps tied to U.S. 60 Hz electrical frequency.
1945	FCC allocates 13 channels for television broadcasting and moves existing radio channels to 88–108 MHz.
1946	ENIAC, the first electronic computer, using 18,000 vacuum tubes, is unveiled.
1948	Long-playing (LP) phonograph records are introduced.
1948	Hollywood switches to nonflammable film.
1948	Ampex introduces its first professional audio tape recorder.
1948	The transistor is invented.
1951	The first commercially available computer, UNIVAC I, goes on sale.
1952	The FCC provides UHF channels 14 through 83.
1953	Second NTSC adopts RCA color TV standard, 525 lines, 29.97 fps, interlaced.
1953	First CinemaScope, anamorphic film is released with 2.66 aspect ratio (1.33 x 2).
1955	Stereo tape recording is introduced by EMI Stereosonic Tapes.
1956	Ampex introduces its first video recorder using 2-inch reel-to-reel tape.
1961	Stereo radio broadcasts begin.
1963	Philips introduces audio cassette tapes.
1967	BBC TWO becomes the first British color broadcast network, using the PAL system, 625 lines, 25 fps interlaced.
1967	France introduces SECAM, 625 lines, 25 fps, interlaced.
1967	The Society of Motion Picture and Television Engineers (SMPTE) standardizes timecode.

Year	Event
1968	The computer mouse is invented.
1970	3/4-inch U-Matic video format is introduced.
1970	Computer floppy disk is introduced.
1971	First permanent IMAX film system is installed.
1972	FCC establishes rules for cable TV.
1972	The first computer editing system, the CMX-300, is introduced.
1975	JVC introduces the Video Home System (VHS).
1977	First preassembled personal computer, the Apple II, is introduced.
1982	Sony, Fujitsu, and Philips introduce audio compact discs (CDs).
1984	Apple introduces the Macintosh computer.
1986	Betacam SP is introduced.
1987	The first commercial digital videotape format, D-1, is introduced.
1990	General Instrument proposes an all-digital HDTV system in the U.S.
1991	Japan adopts Hi-Vision/MUSE as the national HDTV standard, 16:9 aspect ratio,1,125 scanning lines, 30 fps, interlaced.
1991	QuickTime 1.0 is introduced, including the Apple Video codec and Animation codec.
1993	Digital Betacam is introduced.
1996	DV format is introduced.
1997	DVD format is introduced.
1997	Advanced Television Systems Committee (ATSC) digital television standards are adopted by FCC, including 18 formats, 6 of which are HDTV.
1999	Final Cut Pro 1.0 is introduced.
2000	DVCPRO HD equipment begins shipping.
2000	First IMX VTRs begin shipping.
2003	First HDV camcorder is introduced.
2005	QuickTime 7 is released, including support for H.264.
2006	XDCAM HD format introduced.