Video for Government Applications

SBS Technologies (Canada) Inc
+1(519)880-8228
sales.fpga@sbs.com
Draft 0.5
04.11.2005

Introduction

This paper provides a quick overview of video standards, methods, and system design factors for the use of video in typical government applications.

It provides this information from an historical perspective, starting with the advent of television half a century ago, to today's latest emerging video standards. Because some government applications involve back-fitting new technology to existing systems, standards encountered can be from the earliest mentioned to the latest.

Analog Video

In the Beginning: NTSC/ RS-170

The very earliest video standard was created in the late 1930's for television in the U.S., and this standard still serves as the foundation for all video today. This method, which later became standardized in EIA RS-170, involved sending video a line at a time, from left to right from the top left corner of the image.

Altogether, 525 horizontal lines (of which 480 are active video content) were sent directly from the camera, through the broadcast gear, onto the television's CRT in millions of homes, all synchronized.

To capture motion well, these 525 lines were sent at a rate of 30 frames per second (which is faster than standard motion picture rates of 24 frames/ second).

Because CRT's flicker badly when refreshed at only 30 frames per second, a method called *interlacing* was used, so that first the odd lines, then even lines, on the screen were refreshed from top to bottom. Each of these odd and even fields were sent at a rate of 60 fields per second (resulting in 30 full frames per second), reducing flicker to acceptable levels at normal TV viewing distances.

The overall analog bandwidth required to send this picture with adequate fidelity was about 4.5 MHz, leading to a 6 MHz overall bandwidth (once audio and inter-channel spacing needs were added) for each of today's TV channels.

This remains exactly the same way monochrome signals are sent today, nearly 70 years later.

Analog Color: NTSC, RS-170A

In the 1950's, *color* television emerged, and it required a technique that was back-ward compatible with existing monochrome television sets and stations. This was achieved by adding color information on a separate sub-carrier (signal) within each channel, which monochrome television sets would not detect.

This method was called NTSC (National Television Standards Council), and a draft standard called RS-170A was created based on this. Strangely, this standard was never finalized, even though RS-170A is widely referenced, and simply taken to mean 'NTSC' standard.

European Standards: PAL, SECAM

Similar standards were also created in Europe: PAL in most of Europe and SECAM in France. These standards were necessarily different because in each case, they were designed to be synchronous with the power line frequencies: 60 Hz in the U.S., and 50 Hz in Europe, to reduce the visual impact of power line noise in the picture.

As a result, PAL sends frames at 625 horizontal lines per frame, 50 frames per second, also interlaced. SECAM is the same, but has a different method for encoding the color signal. In all cases, the interlaced method is used.

Government Standards: RS-343, European STANAG Standards

Along the way, other standards came into play.

RS-343

In the 1960's, security applications required higher screen resolution than the standard 525-lines provided by standard television, and in 1969 the new RS-343 standard provided higher resolutions, up to 1,023 lines. Today, the resolution most widely associated with RS-343 is 875 lines, as with standard television. Although the actual RS-343A standard specifies monochrome only, this has been extended to a de facto color standard in the defense industry by providing three RS-343A lines: one for each of R, G, and B. RS-343A and RS-343A again specify *interlaced* video.

STANAG 3350

The STANAG standards evolved in Europe as a means of standardizing more rigorously what was already in common use.

The following table shows which STANAG standard applies to which conventional standard.

	STANAG Stamdard	Basic Method	Based on
1	STANAG 3350 Class A	875 Line @ 60Frames/ Sec	RS-343
2	STANAG 3350 Class B	625 Lines @ 50 Frames/ Sec	PAL
3	STANAG 3350 Class C	525 Lines @ 60 Frames/ Sec	NTSC1 RS-170A

Composite vs. Component Video

All the above standards mentioned above are in the category of ‘composite video’: as is necessary in a television broadcast, all video information is sent on a *single wire*, usually coaxial cable, or over the air. Composite video is sometimes abbreviated as ***CVBS*** – composite video broadcast signal.

However, improved video quality can be achieved if multiple wires can be used, as is quite possible for local, rather than broadcast systems, described next.

S-Video, RGB

In applications where the video is to be transmitted only a short distance, it is economic to provide more than one wire to carry the video, and use higher bandwidths, to provide improved video quality.

S-Video: The first method for improving quality is to separate brightness (also called luminance) information, and color information into two separate wires. This is done in S-Video, and results in higher color resolution since the luminance and color information no longer have to share the same 4.5 MHz bandwidth: each can have a separate 4.5 MHz of bandwidth. S-Video systems still specify interlaced format to be television compatible.

RGB *component* video takes this a step further, and provides three wires, one for each of Red, Green, and Blue color ‘guns’ in a color CRT. This provides still better color quality. RGB component video can involve either interlaced video, or in newer systems ‘progressive scan’ video, where all lines on the screen are sent in order at 60 frames per second. This is used, for example, when DVD players are connected to a nearby television set/ monitor.

PC World: RGB, VGA

With the emergence of PC’s in the 1980’s, there was a strong need for higher resolutions than broadcast television, and the interlaced method that works well for television viewing distances did not work well for closer PC viewing distances.

Although some early, now-gone, standards addressed this (CGA, EGA), the standard that has stood the test of time is VGA, and its descendants. Here is a list of today’s standard VGA resolutions:

	VGA Standards	
	Acronym	**Resolution**
1	VGA	640 x 480
2	SVGA	800 x 600
3	XGA	1024 x 800
4	SXGA	1280 x 1024
5	UXGA	1600 x 1200

All VGA systems are progressive scan, and interlacing is not used.

Also, unlike broadcast systems, the frame rate is variable as agreed by the source (PC) and the monitor, and not fixed by the standard. So many of today's systems run at 75 FPS (Frames per Second) or higher, rather than 60 FPS. Due to characteristics of the human visual perception, flicker on displays is noticeable at 60 Hz progressive, but is not noticeable to most at about 75 FPS and higher.

Analog VGA systems always send the video on three separate lines: R, G and B. As well be seen later, VGA resolutions more recently can be send digitally, rather than in this analog format.

➔ *PC World: RGB, VGA*

In order to synchronize the video images, all the above methods employ a 'start of frame' 'vertical' synchronization signal to tell the monitor/ TV to jump to the start of the next field or frame, and then a 'start of line' 'horizontal' synchronization signal to tell it to start the line.

In the single-line composite video methods, these are all embedded in the single wire using different voltage levels and pulses.

For RGB applications, including the VGA standards, three methods are in use:

1. RGB V, H: Vertical and Horizontal Sync on separate wires, so 5 wires (actually coax cables) altogether.
2. RGB Composite Synch: Vertical and Horizontal Sync both on a 'Composite Sync' line, so this involves 4 wires altogether.
3. RGB Synch on Green: Vertical and Horizontal Synch both on the 'G' Green video signal, so 3 lines altogether.

Differential RGB

In some defense applications, rather than sending the RGB active video on three coaxial lines, it is sent on 3 'differential pairs', where the signal is the *difference* in voltage between each half of the pair. This method eliminates the effects of any ground noise in the signal, by canceling it out, and is useful when there can be ground voltage differences between the send and receive ends.

Digital Video

Differential RGB

During the 1990's, much of the video broadcast world moved to digital rather than the above analog techniques, because:

- Digital video does not gradually degrade as it is re-transmitted from system to system as is inevitably true with analog. With digital, the end result is identical to the initial transmission no matter how many stages in the system, for better end-delivered picture quality, with no 'snow', 'ghosts', or other visual degradation.
- Digital video can be highly compressed, for reduced bandwidth (on cable TV and Satellite links) and reduced storage needs (on DVD's) as described later.

Typically, when video is digitized, it is fragmented into individual 'pixels', and the pixels are transmitted sequentially in either serial or parallel streams. Often, color pixels are sent in 24-bit format (8 bits for Red, 8 bits for Green, 8 bits for Blue), or 32-bit format (another 8 bits added to these 24 bits for 'transparency'. In this case, this extra value is called the 'alpha' value.)

As with analog video, pixels are sent sequentially, starting from the top left of a frame, and then left to right, line by line. For digital versions of the interlaced standards, lines are again sent in interlaced fashion.

ITU656 is a standard method of sending NTSC/ RS-170A signals digitally within systems, in an 8-bit parallel format.

Digital Line Sampling

When lines are digitized, the number of digital samples per second can vary. For example, for high quality/ DVD quality television, 720 digital samples per line are often taken.

However, this does not result in ***'square pixels'***. For applications requiring square pixels (so a circle looks circular and Kirsty Allie looks svelte), 640 samples are taken. Because there are 480 active video lines on a standard TV picture, the 640 X 480 format results in square pixels on the standard TV screen, with its 4:3 aspect ratio (width to height ratio).

Broadcast Digital Video: SDI

In the broadcast studio world, video transmission has now largely migrated from analog to digital, using the SDI – Serial data Interface.

For standard definition ('SD') video as used on normal television broadcasts), SDI is a serial data stream of 270 Mbits per second, and the standard is SMPTE 259. SMPTE abbreviates the 'Society of Motion Picture and Television Engineers'.

For high-definition ('HD') television, the SDI stream (called HD-SDI) is sent at a rate of 1.485 Gbits per second, and the standard is SMPTE 292M.

In each case, the all data (video, audio, synchronization, and ancillary data such as closed-captioning) is sent on a single 75 ohm coax cable, in the form of 10-bit packets.

This method is now in wide use in broadcast studios and transmission facilities worldwide.

Display Video: DVI

Another progression occurring is that display and television monitors are converting from analog (composite or component RGB) to digital interfaces, using the DVI (Digital Video Interface) standard.

This standard, being digital, ensures no picture degradation between the video output and the display input.

The video is sent on 4 shielded differential wire pairs, at the rate dictated by the video/ graphics frame rate and resolution. This rate is often well over 1 Gbit per second on each pair, and the cabling has length limitations making it suitable for close-by connections.

Digital Video

Video Compression Standards Overview

One of the big advantages of digital video is that it can be compressed for reduced bandwidth applications transmitted over satellite, cable TV and Internet-based networks. Compressed video is particularly useful for reducing storage requirements for video recorder (DVR) applications, especially in the government market.

There are "lossless" and "lossy" forms of data compression. Lossless data compression is used when the data must be restored exactly as it was before compression. Since losing a single character can make restored numbers or text misleading, number and text files are stored using lossless techniques, such as Huffman Coding, Lempel-Ziv-Welch (LZW). A lossless compression technique for images, Portable Network Graphics (PNG) is a recommendation of the World Wide Web Consortium and now an ISO standard, and is especially useful for images displayed and stored on Web sites.

There are limits, though, to the amount of compression that can be obtained with lossless compression techniques. Lossless compression ratios are generally in the range of 2:1 to 8:1. Lossy compression, on the other hand, works on the assumption that the data doesn't have to be restored perfectly. A good deal of redundant information can be simply thrown away from images, video data, and audio data, and when uncompressed such data will still be of acceptable quality. Compression ratios can be an order of magnitude greater than those available from lossless methods.

Over the years, there have been two main standards bodies doing parallel development of video compression standards. The first widely-used standards for video compression were developed by the Moving Picture Experts Group (MPEG). Another standards group, the International Telecommunication Union (ITU), has developed the H series of compression standards, mainly for the telecommunications industry.

As a standards organization, the MPEG and ITU committees do not specify end-user product or equipment. MPEG does, however, standardize Profiles. A Profile is a selection of tools that a group of participating companies within the standard have selected as a basis for deploying products to meet specified application areas. For example, MPEG-4 Simple Profile (SP) and Advanced Simple Profile (ASP) were developed for streaming video over Internet connections.

To become standardized, Profiles pass through a requirements process where the tools and applications are reviewed and voted on as being an interoperable profile for the industry. Within each Profile there can be one or more Levels. Levels allow for increasing complexity of the tools to allow some diversity within a Profile in addressing devices of varying performance. Levels may thus restrict bit-rates, size, number of nodes etc.

➔ *MPEG-1*

The first lossy compression scheme developed by the MPEG committee, MPEG-1, is still in use today for CD-ROM video compression and as part of early Windows Media players. The MPEG-1 algorithm uses a combination of techniques to achieve compression, including use of the Discrete Cosine Transform (DCT) algorithm to first convert each image into the frequency domain, and then process the frequency coefficients to optimally reduce a video stream to the required bandwidth.

The DCT algorithm is well known and widely used for data compression. Similar to Fast Fourier Transform, DCT converts data, such as the pixels in an image, into sets of frequencies. To compress data, the least meaningful frequencies are stripped away based on allowable resolution loss—generally user defined. This loss of resolution results in a lossy compressed image. Images that are converted to frequencies.

Rather than fully encoding and compressing every video frame, MPEG-1 compression processes a 'Group of Pictures' where it:

- Fully encodes an 'I' (independent) frame.
- Encodes only the differences on subsequent 'P' (progressive) frames, and
- In the most complex case, also provides 'B' (bi-directional) frames, which look both ahead and back in time to determine how to best compress the signal.

In slow–moving scenes, the image differences between successive frames are small, resulting in higher compression rates without great loss of detail achieving great bandwidth savings. With fast-moving sequences, image differences are greater and bandwidth savings are far less as exemplified by sports channels on satellite television which consume more bandwidth than talk shows, for example.

MPEG-1 found usage on CD-ROM Videos, in early versions of Windows Media player, and other PC applications, but does not support higher-quality video such as today's DVD standards.

Interestingly, the current wildly popular MP3 (MPEG-1, layer three) audio standard is actually the audio compression portion of the MPEG-1 standard and provides about 10:1 compression of audio files at reasonable quality.

MPEG-2

The MPEG-2 compression standard evolved to meet the needs of compressing higher-quality video. MPEG-2 is used in today's video DVD's and digital broadcasts via satellite and cable and uses bit rates typically ranging from 5 to 8 Mbps, although MPEG-2 is not really limited to a bit rate range. MPEG-2's basic compression techniques are very similar to MPEG-1, using DCT transforms, I and P frames, but also provides support for interlaced video (the format used by broadcast TV systems).

Interlaced video causes less visible display flickering on a CRT monitor than non-interlaced methods by alternating between drawing the even-numbered lines and the odd-numbered lines of each picture. In contrast, a non-interlaced raster display draws every line of a picture, or frame, in sequence from top to bottom. This takes a certain amount of time, during which time the image on the CRT begins to decay, resulting in flicker.

An interlaced display reduces this flicker effect by drawing first all the even-numbered lines (forming the even field), leaving spaces between them for all the odd-numbered lines (forming the odd field) which it fills in afterwards to complete the frame. This results in the display being refreshed from top to bottom twice as frequently as in the non-interlaced case.

MPEG-2 video is not optimized for low bit-rates (less than 1 Mbit/s), but outperforms MPEG-1 at 3 Mbit/s and above. MPEG-2 also introduces and defines Transport Streams, which are designed to carry digital video and audio over unreliable media, and are used in broadcast applications. With some enhancements, MPEG-2 is also the current standard for High Definition Television (HDTV) transmission. MPEG-2 also includes additional color subsampling, improved compression, error correction and multichannel extensions for surround sound.

Although MPEG-2 excels at full broadcast television, and can be used to retrieve and control streams from a server, just like MPEG-1 compression, MPEG-2 audio and video compression are still essentially linear and interactivity is limited to operations such as slow motion, frame-by frame or fast forward.

MPEG-3

MPEG-3 is the compression standard that never was. While it was originally intended by the MPEG committee that an MPEG-3 standard would evolve to support HDTV, it turned out that this could be done with minor changes to MPEG-2. So MPEG-3 never happened and there are now 'profiles' of MPEG-2 that support HDTV as well as Standard Definition television.

MPEG-4 (also called H.263)

Although a full explanation of the MPEG-4 standard is well beyond the scope of this whitepaper, MPEG-4 has emerged as much more than a video and audio compression and decompression standard. The MPEG committee designed MPEG-4 to be a single standard covering the entire digital media workflow from capture, authoring, and editing to encoding, distribution, playback, and archiving. It is a container for all types of items—called "media objects"—beyond audio and video. Media objects can be text, still images, graphic animation, buttons, web links, and so on. These media objects can be combined to create polished interactive presentations.

The MPEG-4 file format, based on Apple Computer's QuickTime technology, was developed by the MPEG committee as a standard designed to deliver interactive multimedia and graphics applications over networks and to guarantee seamless delivery of high-quality audio and video over IP-based networks and the Internet.

A major goal of the MPEG-4 standard was to try to solve two video transport problems:

1. Sending video over low-bandwidth channels such as the Internet and video cell phones.
2. Achieving better compression than MPEG-2 for broadcast signals.

MPEG-4 functions well in terms of compression and it is used in a wide range of bit rates, from 64 Kbps to 1,800 Mbps. However, it had limited success in achieving dramatically better compression than MPEG-2 for broadcast signals, and although it is in the range of 15% better at compressing video data than MPEG-2, this has not been enough of an advantage to convert the whole broadcast industry to the MPEG-4 format.

So, MPEG-4's role will likely remain in lower-bandwidth applications in the desktop computer, Internet, and cell phone worlds, and also in new applications where a 15% compression improvement over MPEG-2 is desired and MPEG-2 compliance is not an issue.

MPEG-4 is actually a super-set of MPEG-2, so MPEG-4 players which decompress the video stream can theoretically play both MPEG-2 and MPEG-4 formats. However, for secondary reasons, this may not be true in practice.

This initial MPEG4 standard is also called MPEG4 Part 2, which is also often confused with MPEG4 Part 10 below.

➔ *H.264/ AVC (also called MPEG4 Part 10)*

With MPEG-4 failing to considerably improve compression performance for full broadcast signaling, another effort was initiated late in the 1990's. This new effort, H.264, is able to achieve a 2:1 improvement over MPEG-2 on full quality SDTV and HDTV, and is expected to come into wide use in satellite and cable TV over the next decade.

H.264/MPEG4-AVC is a jointly developed standard by the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG) and has been standardized by the ITU under the H.264 name, and is also called MPEG-4 Part 10 AVC (Advanced Video Compression) even though it is unrelated in operation to MPEG-4.

The main goals of the H.264/ MPEG4-AVC standardization effort are to provide significantly enhanced compression performance and provision of a "network-friendly" packet-based video representation addressing "conversational" (i.e., video telephony) and "non-conversational" (i.e., storage, broadcast, or streaming) applications.

H.264 uses techniques fairly different from MPEG-2 and can match the best MPEG-2 quality at up to half the data rate. H.264 also delivers excellent video quality across the entire bandwidth spectrum — from 3G to HDTV and everything in between (from 40 Kbps to upwards of 10 Mbps). Efficient encoders and decoders for H.264 are just coming into use in 2005.

The H.264 design incorporates a Video Coding Layer (VCL), which provides the core high-compression of the video content, and a Network Abstraction Layer (NAL), which packages that compressed content for delivery over networks. The VCL design has achieved a significant improvement in rate-distortion efficiency - providing nearly a factor of two in bit-rate savings against existing standards. The NAL designs are being developed to transport the coded video data over existing and future networks such as circuit-switched wired networks, MPEG-2/ H.222.0 transport streams, IP networks, and 3G wireless systems.

H.264 contains a number of features that allow it to compress video much more effectively than older codecs over a wide variety of network environments. Key H.264 features include:

- Multi-picture motion compensation using previously-encoded pictures as references in a much more flexible way than in past standards, thus allowing up to 32 reference pictures to be used in some cases (unlike prior MPEG standards, where the limit was typically one or two in the case of conventional B pictures.
- Variable block-size motion compensation (VBSMC) with block sizes as large as 16×16 and as small as 4×4, enabling very precise segmentation of moving regions.
- An in-loop deblocking filter which helps to prevent the blocking artifacts common to other DCT-based image compression techniques used in MPEG standards.
- A secondary Hadamard transform performed on Discreet Cosine coefficients of the primary spatial transform to obtain even more compression in smooth regions.
- A network abstraction layer (NAL) definition allowing the same video syntax to be used in many network environments, including features such as sequence parameter sets (SPSs) and picture parameter sets (PPSs) that provide more robustness and flexibility than provided in prior standards.
- Frame numbering, a feature that allows the creation of "sub-sequences" (enabling temporal scalability by optional inclusion of extra pictures between other pictures), and the detection and concealment of losses of entire pictures (which can occur due to network packet losses or channel errors).
- Picture order count, a feature that serves to keep the ordering of the pictures and the values of samples in the decoded pictures isolated from timing information allowing timing information to be carried, controlled and changed separately by a system without affecting decoded picture content.

JPEG 2000

The MPEG and H.264 standards all relate primarily to motion video. For still pictures, the familiar JPEG standard developed by the Joint Picture and Editors Group committee has been in use for some years, and it is just now gradually being replaced by the JPEG committee's improved JPEG 2000 standard, which was released in the year 2000.

JPEG 2000 is mentioned in this video overview because, even though JPEG 2000 is designed for still picture use, Part 3 of the JPEG 2000 standard–Motion JPEG 2000–also provides for motion video. Motion JPEG 2000 adds a mechanism to the JPEG 2000 standard for sending JPEG 2000 images in a video stream with support for associated audio. Since the MJ2 (Motion JPEG 2000) format does not involve inter-frame coding and each frame is coded independently, high-quality frame-based video recording and editing based on the high quality of JPEG 2000 compression is possible.

JPEG2000 uses ‘wavelet’ compression technology rather the DCT technology used in the MPEG and JPEG standards. DCT compresses an image into 8x8 pixel blocks and places them consecutively in the file. The blocks are compressed individually, without reference to the adjoining blocks, resulting in the blocky look associated with compressed JPEG files. With high levels of compression, only the most important information is used to convey the essentials of the image and much of the detail is lost, lowering the dynamic range of an image.

In contrast, JPEG 2000 wavelet compression converts the image into a series of wavelets that can be stored more efficiently than pixel blocks. Wavelet algorithms compress the entire image with ratios of up to 300:1 for color and 50:1 for gray scale. Wavelet compression also supports non-uniform compression, where specified parts of the image can be compressed more than others.

JPEG 2000 is able to render pictures better by eliminating the blockiness that is a common feature of DCT compression. Not only does JPEG 2000 exhibit smoother color toning and clearer edges where there are sharp changes of color, JPEG 2000 also produces smaller image file sizes than JPEG image files with the same level of compression.

The advantage of JPEG 2000 becomes apparent when high compression ratios are required; when 2 bits per pixel are available, both standards provide a comparable image quality. However, when this reduces to 0.5 bits per pixel available, JPEG 2000 still produces a usable image, whereas JPEG does not.

The advantages of using Motion JPEG2000 for video are:

- Low latency compared to MPEG streams which use ‘P’ and ‘I’ frames.
- For DVR applications, every image is self-contained and complete; no need to reconstitute from P and I frames.

The disadvantages are:

- Lower compression ratios than MPEG algorithms.
- Requires more computing power for decoding, so hardware-assisted decoding via FPGAs, DSPs and ASICs are required. Software-only decoders are not currently acceptable for full broadcast quality.

Digital Video Transmission Formats

When digital video is transmitted, a number of different formats are in use, as follows:

Broadcast Studios

SDI as mentioned above

- SMPTE259M for SD (Standard Definition) Video at 270 Mbits/ second.
- SMPTE292M for HD (High Definition) at 1.54 Gbits/ second.

Between Remote Computers, Over Ethernet

- Compressed formats: MPEG-2, MPEG-4, MotionJPEG2000.
- These sent on RTP (Real Time Protocol) over UDP (User Datagram Protocol) over IP over Ethernet.
- Typical data rates for compressed video are in the range of 0.1 Mbits/ second (low resolution teleconferencing and webcasts) to 4 mbits per second (DVD quality video) on up to 19 Mbits/ second (HD TV).
- With these rates, it is possible to send multiple streams of video over Ethernet links (even 100BT).
- It is also possible to send lower-end video over Internet links down to 1 Mbit/ second rates.

Within Computer Systems

PCI and VME bus rates in embedded systems are usually in a range of 50 to 500 MBytes per second. Uncompressed video streams are about 27 Mbytes/ second for SD television and 150 MBytes/ second for HD television quality. So it is possible to send modest amounts of video over PCI and VME buses if they are not already busy. Compressed video reduces video rates by factors of about 30:1 to 100:1, so it is possible to send multiple streams of video over back-plane PCI and VME buses.

Compressed Video to Home

Cable and Satellite Modems:

Both of the following standards involve converting digital video back to analog to fit into standard analog 6 MHz channels as available on cable TV and satellite TV systems.

Digital TV set-top boxes implement one of these standards:

ATSC (Advanced Television Systems Committee)

- Actually a set of standards from A/52 to A/90.
- For HD, includes Dolby Digital Audio Compression.
- Sends 19.4 Mbps of digital video data in a standard 6 MHz channel, allowing one HD channel with full audio in standard 6 MHz channel.
- Can send multiple SD digital channels on standard 6 MHz channel.
- U.S. origin.

DVB (Digital Video Broadcast)

- Actually a set of standards from ETSI EN 300 421 to ETSI EN 300 775.
- Similar to ATSC, but European origin.

Compressed Video in Broadcast Studio

The standard ***SMPTE310M*** is in very common use, and applies to sending MPEG-2 information digitally in a broadcast studio.

Digital Video Transmission Formats

The most prevalent file format for multimedia (video + audio) files on PC's are:

- The Above MPEG Standards (.mpg or .mpeg).
- The AVI (Audio Video Interleave) standard from Microsoft (.avi). This is a very flexible standard, allowing for virtually any compression method (it simply searches for the specified decoder as referenced in the video stream), and also provides similar high flexibility in the audio stream.
- Apple's Quicktime Standard (.mov), which uses MPEG-4 internally.

Glossary of Commonly Used Video Terms

Alpha Mixing	A method for combining two where an “alpha channel” is used to define where the location of an over-lay image.
Anti-Alias Filter	Low-pass filtering used to lower a signal’s bandwidth to less than half the sampling rate.
ARIB	Association of Radio Industries and Business – standards organization in Japan.
Artifacts	In video context they are blemishes, noise white spots, or other undesirable added or missing content on an image. Artifacts are not related to display setup issues.
ARIG-B	Military standard for supplying time information for time-stamping of video capture.
Aspect Ratio	The ratio of picture width to height. Displays commonly have 4:3 or 16:9 aspect ratios.
Asynchronous	Circuitry or transmission of data without a common clock or timing signal.
ATSC	Advanced Television Systems Committee – standards committee in the United States for HDTV.
Baseband	An audio or video signal which is not modulated onto another carrier signal.
BITC	Burned In Time Code.
Bit Rate	Number of bits (logical “1s”or “0s”) per second a transmission network can supply or data stream requires.
Blanking	On a CRT, can lines move from left to right and jump back to the left to start again. As the scan line is reset from right to left the signal to the CRT is turned off – blanked – so the retrace is not seen.
Blooming	When video becomes whiter-than-white and greats a distortion or halo around light colored objects.
Brightness	Refers to how much light is emitted from the display.
BT.470	ITU specification for various NTSC, PAL and SECAM video standards.
BT.656	An ITU recommended specification for transmission of parallel and serial data for transmission of YCbCr data. Also see SMPTE125.
B-Y	The Blue- Luma signal, also called color difference.
Camera Link	Industrial digital camera standard for high speed digital data up to 850 MBytes/ sec.
CBR	Constant Bit Rate.
CCIR	International Radio Consultive Committee which is now part of ITU standards organization.

Chroma	The color component of NTSC, PAL and SECAM signal.
Chroma Bandpass	In NTSC, PAL and SECAM images have the luma (brightness) and Chroma (color) information combined. The Chroma Bandpass filter will separate this information.
Chroma Key	A method to combine two video images. An example is a weather person standing in front of computer generated map – the weather person image is in front of a blue/green background. A chroma key system will review the video data and display the map where the blue/ green data is present and the person elsewhere. Also known as Color Key.
CIF	Common Interface Format or Common Image Format. Originally designed to support video conferencing with an active resolution of 352 x 288 with a refresh rate of 29.97 frames per second.
Closed Caption	A service which decodes text information which is transmitted with the video signal, then displays it with the video. NTSC has the caption signal on lines 21 and 284, for PAL the information is lines 22 and 334. For MPEG-2 closed caption data is multiplexed as a separate data stream within MPEG-2.
CMYK	Cyan, Magenta, Yellow and Black – Color space primarily used in color printing.
Comb Filter	A method of performing Y/C separation of Luma and Chroma information.
Common Image Format	See CIF.
Component Video	Video using three separate color components such as YCbCr, YPbPr, or RGB.
Composite Video	A single video signal that contains all the necessary brightness, color, and timing information to re-create and video image. NTSC, PAL are examples of composite video.
Compression Ratio	The rate at which the image size is reduced. For example and 1 Mbyte image could be reduced to 128 Kbyte image through and MPEG-2 compression alogrithim – resulting in an 8 to 1 (8:1) compression ratio.
Constant Bit Rate	CBR – means the bitstream of a compressed or uncompressed has the same number of bits each second. This can be parameter in an algorithm with variable compression rate (i.e. MPEG motion estimation) to maximize image quality with a defined transmission bandwidth.
Contouring	An image artifact caused by not having enough bits to represent the image.
Contrast	A video term to describe how far the whitest whites are from the blackest blacks.
CVBS	Composite Video and Baseband Signal or Composite Video, Blanking and Synchronization.
Decoder	Algorithm used to return a compressed signal to it uncompressed state for display.
Decimation	When a digital signal is sampled, but all the data is used. An example would be storing every other sample of signal – would be a 2:1 decimation. Commonly used for converting images for different size displays.

Demodulation	The process of recovering an original signal from a modulated carrier.
Digital Component Video	Digital Video using three separate color components – such as RGB.
Digital Composite Video	Digital video that is essentially the digitized waveform of NTSC or PAL.
Digital Video Recorder	A system used to control, store and playback digital video. Often combined with compression encoding and decoding.
Discrete Cosine Transform (DCT)	Discrete Cosine Transform is another way of representing an image in frequency domain – rather than in time domain, as we normally do. It is often used in compression algorithms as a more efficient way of representing the image. It is also used in image analysis and filtering.
Downlink	A frequency satellites use to transmit data to Earth stations.
DVI	Digital Visual Interface – a digital interface to a video display, designed to replace the analog RGB or YPbPr interface.
EIA	Electronics Industry Alliance standards committee
Field	An interlaced display is made of two “fields”, each one containing half of the scan lines needed to make up one frame of video. Each of these two fields of “even or odd” scan lines are displayed in their entirety to create a full frame of video.
Firewire	IEEE 1394 A/B high speed serial digital
Flicker	Flicker occurs when the video frame rate of the video is too low and each individual frame can be “recognized” rather than continuous motion.
FM	See Frequency Modulation.
Frame	A single picture or image from a video stream. By playing these individual frames fast enough, it appears images are “moving” on the screen.
Frame Buffer	A frame buffer is a memory used to hold an image for display or processing. If we use as an example a 640 x 480 color RGB image, the frame buffer size could be calculated as: $640 \times 480 \times 3 = 921{,}600$ bytes or 900 kB
Frame Rate	The frame rate of a video source is how fast a new still image “frames” are streamed to create the effect of motion. In NTSC system, a new still image is re-displayed every 30th of a second, for a frame rate of about 30 frames per second. For PAL, the frame rate is 25 frames per second. For computer displays, the frame rate is usually about 75 frames per second.
Frame Rate Conversion	Frame rate conversion is the act of converting one frame rate to another.

Frame Grabber	Terminology used for a electronic board which provides the external camera interface and acquires an image or video stream and outputs this image to the host computer in digital form. No image or video processing is typically done on frame grabber – only image conversion and formatting.
Graphic Overlay	The mixing or adding of graphic images or alphanumeric data to a live video stream for a more content rich display. Often the graphics images are generated and mixed on a graphics processor chip, such as ATI or NVVIDA device and in real time prior to display.
Gamma	The characteristics of most cameras and displays are nonlinear. For a display, a small change in amplitude when the signal level is small produces a large change in the display brightness level, but the same change in amplitude at a high level will not produce the same magnitude of change. This nonlinearity is known as gamma.
Genlock	A synchronization signal which supports the video decoder's capability to "lock onto" the timing information embedded in the video signal. The circuitry within the decoder which allows this synchronization of video is called the Genlock circuitry.
Gray Scale	In some cases, it means the luma component of color video signals. In other cases, it means a black-and-white video signal.
H.263, H.261	The ITU-T H.261 and H.263 video compression standards were developed to implement video conferencing over ISDN, LANs, regular phone lines, etc. H.261 supports video resolutions of 352 × 288 and 176 × 144 at up to 29.97 frames per second. H.263 supports video resolutions of 1408 × 1152, 704 × 576, 352 × 288, 176 × 144, and 128 × 96 at up to 29.97 frames per second.
H.264	The "next-generation" video codec. Previously known as "H.26L", "JVT", and "AVC" (advanced video codec), it is now also a MPEG-4 Part 10 standard. ITU-T H.264 offers bit rates up to 50% less than the MPEG-4 advanced simple profile (ASP) video codec for the same video quality. It is designed to compete with the SMPTE VC-9 video codec in bit rate and quality.
HD-SDI	High definition television resolution and frame rate (HD) transmitted over Serial Data Transport Interface (SDI) protocol.
HDTV	See High Definition Television.
High Definition Television	HDTV is video capable of displaying at least 720 progressive or 1080 interlaced active scan lines. It must be capable of displaying a 16:9 image format using at least 540 progressive or 810 interlaced active scan lines. A typical HD TV standards include SMPTE.292.
Horizontal Blanking	During the horizontal blanking interval, the video signal is at the back level so as not to display the electron beam when it sweeps back from the right to the left side of the CRT screen.
Horizontal resolution	The number of combined RGB pixels of video information across the horizontal axis of the display.

Horizontal Scan rate	The speed with which the scanning beam in a display sweeps from side to side. In the NTSC system, this rate is 63.556 ms, or 15.734 kHz. That means the scanning beam moves from side to side 15,734 times a second.
Horizontal Sync	The part of the video signal that determines where to place the image in the left-to-right dimension. The horizontal sync pulse signals the beginning of the new scan line.
HSI	Hue, Saturation and Intensity -HSI is based on polar coordinates, while the RGB color space is based on a three-dimensional Cartesian coordinate system. The intensity, analogous to luma, is the vertical axis of the polar system. The hue is the angle and the saturation is the distance out from the axis. HSI is more intuitive to the human eye and easier to manipulate colors as opposed to the RGB space. For example, in the HSI space, if you want to change red to pink, you decrease the saturation. In the RGB space – it is not as straight forward. In the HSI space, if you wanted to change the color from purple to green, you would adjust the hue.
HSYNC	See Horizontal Sync definition.
Huffman Coding	Huffman coding is a method of data compression used in JPEG, MPEG, H.261, and H.263 to help with the compression. A table is created depending that lists how many times each piece of unique data occurs. A small code word is used to represent each data. The most frequently occurring data is assigned the smallest unique code word.
Interlaced	An interlaced video system is one where two interleaved fields are used to generate one video frame. Therefore, the number of lines in a field is one-half of the number of lines in a frame. In NTSC, there are 262.5 lines per field (525 lines per frame), while there are 312.5 lines per field (625 lines per frame) in PAL. Each field is drawn on the screen consecutively—first one field, then the other.
JPEG	Joint Photographic Experts Group (JPEG) - JPEG is used a generic name for the compression algorithm developed by the JPEG users group. JPEG compresses still images which can be streamed together into video – called Motion JPEG or MJPEG. JPEG is considered to have lower compression ratios, however less loss due to compression than past MPEG-1 and MPEG-2 standards.
JPEG2000	JPEG2000 is an updated motion-based MJPEG. JPEG2000 is often used in medical and digital cinema applications where lossless reproduction is required.
kbps	Kilobits per second.
kBps	Kilobytes per second.
Latency	The delay caused by processing and transmission of data, in this case digital video, typically measured in frames or milliseconds. Typical maximum latency for interactive video control of a vehicle or robot is the order of 100 Milliseconds

Line Store	A line store is memory used to hold one scan line of video. If the horizontal resolution of the active display is 640 samples and RGB is used as the color space, the line store would have to be 640 locations long by 3 bytes (R, G & B) wide.
Lossless	Lossless is a term used relative to compression. Lossless compression is when the decompressed video is exactly the same as the original video data.
Lossy	Lossy compression is the exact opposite of lossless. The regenerated data is different from the original data. The differences may or may not be noticeable, but if the two images are not identical, the compression was lossy.
Luma	As mentioned in the definition of chroma, the NTSC and PAL video systems use a signal that has two pieces: the black and white part, and the color part. The black and white part is the luma.
Mbps	Megabits per second.
MBps	Megabytes per second.
M-JPEG	See Motion JPEG.
Modulator	A modulator is a circuit that combines two different signals in a way that they can be pulled apart later. Often used in signal processing to combine, data video or audio signals onto a carrier frequency for transmission.
Moiré	A moiré effect is an image distortion. A moiré pattern is typically generated when two different frequencies interfere to create a new, unwanted frequency.
Monochrome	A monochrome signal is a video source having only one color component – and example is luminance of a black/ white video image.
Motion Estimation	Motion estimation is attempting to predict where an object has moved to from one video frame to the other. It is often used in video compression algorithms for better compression rates and picture quality. Motion estimation is an integral part of MPEG, H.261, and H.263.
Motion JPEG	JPEG compression or decompression of image frames which are streamed into real-time video.
MJPEG	JPEG compression or decompression that is applied real-time to video. Each field or frame of video is individually processed. See Motion JPEG above.
MPEG-1	This is the first video compression standard by the Motion Picture Experts Group - MPEG defining the compression format for real-time audio and video. The video resolution is typically 352 × 240 or 352 ×288, although higher resolutions are supported. The maximum bit rate is about 1.5 Mbps. MPEG-1 is used for the Video CD format.
MPEG-2	MPEG-2 extends the MPEG-1 standard to cover a wider range of applications. Higher video resolutions are supported to allow for HDTV applications, and both progressive and interlaced video. MPEG-2 is often used for DVD–Video.

MPEG-3	MPEG-3 was originally targeted for HDTV applications. This was incorporated into MPEG-2, so there is no MPEG-3 standard for video.
MPEG-4 Part 2	Also called H.263 or MPEG4 Short Header is an enhancement of MPEG2 for optimized compression and video quality. Optimization was designed for small format video phone type applications.
MPEG-4 Part 10	MPEG-4 (ISO/IEC 14496) supports an object-based approach, where scenes are modeled as compositions of objects, both natural and synthetic. Visual objects in a scene can be described mathematically and given a position in a two- or three-dimensional space. Similarly, audio objects can be placed in a sound space. Thus, the video or audio object need only be defined once; the viewer can change his viewing position, and the calculations to update the audio and video are done locally. H.264, a "next-generation" video codec, has been included in the MPEG-4 standard as Part 10.
MPEG-7	MPEG-7 standardizes the description of multimedia material (referred to as metadata), such as still pictures, audio, and video, regardless if locally stored, in a remote database, or broadcast.
Noninterlaced	Scanning of a video display where all of the lines in the frame are scanned out sequentially, one right after the other. Another term for a noninterlaced system is progressive scan.
NTSC	NTSC is a color modulation scheme defined by the National Television Standards Committee.
OGL	See Open Graphics Language.
Open Graphics Language	A cross-platform standard for 3D rendering and 3D hardware acceleration developed by Silicon Graphics Inc
Overlay	The placement or combining of fixed images or video data to create a more content rich display. Overlaying is typically done by a graphics processor.
PAL	PAL stands for Phase Alternation Line. PAL is a color modulation scheme. Many forms of PAL video format exist, referred to as (B, D, G, H, I, M, N, or NC) PAL. (B, D, G, H, I) PAL is the color video standard used in Europe and many other countries.
Pixel	A pixel or picture element is the smallest element which makes up a single scan line of a digital image. For example, when the horizontal resolution is defined as 640 pixels, which indicates there are 640 individual locations, or picture elements, that make up the image. Pixels may be square or rectangular.
Pixel Clock	The pixel clock is used to divide the horizontal line of video into samples of individual pixels.
PMC	Also known as PCI PMC - Peripheral Component Interconnect Mezzanine Card.
Progressive Scan	See Noninterlaced
QAM	See Quadrature Amplitude Modulation.

QCIF	Quarter Common Interface Format. This video format was developed to allow the implementation of cheaper video phones. The QCIF format has a resolution of 176 × 144 active pixels and a refresh rate of 29.97 frames per second.
QSIF	Quarter Standard Interface Format. The computer industry, which uses square pixels, has defined QSIF to be 160 × 120 active pixels, with a refresh rate of whatever the computer is capable of supporting.
Quadrature Amplitude Modulation	A method of encoding digital data onto a carrier for RF transmission. QAM is typically used for cable transmission of digital SDTV and HDTV signals. DVB-C supports 16-QAM, 32-QAM, 64-QAM, 128-QAM, and 256-QAM, although receivers need only support up to 64-QAM.
Quantization	The process of converting a continuous analog signal into a set of discrete levels (digitizing).
Raster	A raster is the series of scan lines that make up a picture. All of the scan lines that make up a frame of video form a raster.
Resolution	Describes the number of picture elements/ pixels in a displayed image. Described as “h” × “v.” The “h” is the horizontal resolution (across the display) and the “v” is the vertical resolution (down the display).
RGB	Abbreviation for red, green, blue.
RS170/ RS170A	RS-170 is the United States standard that was used for black-and-white TV, and defines voltage levels, blanking times, the width of the sync pulses, etc. SMPTE 170M is essentially the same specification, modified for color TV by adding the color components. RS-170 is a common analog video format for security and surveillance cameras.
RS-343	RS-343 is a high resolution variation of RS-170 using 875 lines of video at 1v p-p signaling. RS343 is typically only used in military video and imaging sensors.
RTSP	RTSP (Real-Time Streaming Protocol) is a client-server protocol to enable controlled delivery of streaming audio and video over an IP network. It provides “VCR-style” remote control capabilities such as play, pause, fast forward, and reverse.
Sample Rate	Sample rate is how often a sample of a signal is taken. The sample rate is determined by the sample clock.
Saturation	Saturation is the amount of color present in an image. Saturation is the quality of an image that separates pink (low saturation) and red (high saturation). Saturation does not mean the brightness of the color, rather the “pigment” used to make the color.
Scaling	Scaling is the act of changing the resolution of an image. For example, scaling a 640 × 480 image by one-half results in a 320 × 240 image. Scaling by 2×results in an image that is 1280 × 960.
Scan Line	A scan line is an individual line across a display. It takes 525 scan lines to make up a NTSC TV picture and 625 scan lines to make up a PAL TV picture.
SDI	Serial Digital I/O (SDI) is a 270 Mbps or 360 Mbps serial interface defined by BT.656. This describes the image transport protocol, but not resolution. There both standard definition and high definition images can be broadcast over SDI protocol.

SDTV	See Standard Definition Television.
SECAM	A color video format similar to PAL. The major differences between PAL and SECAM are the chroma information is FM modulated and the R–Y and B–Y signals are transmitted line sequentially.
SIF	Standard Input Format is a video format developed to allow the storage and transmission of digital video. The 625/ 50 SIF format has a resolution of 352 × 288 active pixels and a refresh rate of 25 frames per second. The 525/ 60 SIF format has a resolution of 352 × 240 active pixels and a refresh rate of 30 frames per second.
Signal-to-Noise Ratio	Signal-to-noise ratio is the magnitude of the signal divided by the amount of unwanted signal that is interfering with the data (the noise). SNR is usually described in decibels, or “dB,” for short; the bigger the number the less noise.
SMPTE	Society of Motion Picture and Television Engineers
SMPTE 170M	NTSC video specification for the United States. See RS-170A and BT.470.
SMPTE 259M	SMPTE definition for transmission of standard definition television signals.
SMPTE 274M	1920 × 1080 pro-video interlaced and progressive standards (29.97, 30, 59.94 and 60 Hz). The digital serial interface is defined by SMPTE 292M.
SMPTE 292M	1.485 Gbps HDTV serial interface standard.
SMPTE 296M	1280 × 720 video standards, progressive scanned. This includes the digital representation and the analog RGB and YPbPr interfaces. Digital parallel interface uses SMPTE 274M. Digital serial interface is defined by SMPTE 292M.
Subtitles	Text that is added below or over a picture that usually reflects what is being said, possibly in another language. Open subtitles are transmitted as video that already has the subtitles present. Closed subtitles are transmitted during the video blanking interval, and rely on the TV to decode it and position it below or over the picture.
S-Video	Separate video, also called Y/C video. Separate luma (Y) and chroma (C) video signals are used, rather than a single composite video signal. By simply adding together the Y and C signals, you generate a composite video signal. The IEC 60933-5 standard specifies the s-video connector, including signal levels.
STANAG 3350	STANAG 3350 is the analogue video standard for aircraft system applications.
SVGA	Super Video Graphics Array - video standard one step above VGA. SVGA monitors are capable of displaying up to 16 million colors with a resolution of 800 x 600 on 14" monitors or up to a 1200 x 1600 resolution on a 20" monitor.
SXVGA	Super eXtended Graphics Array - a video resolution that supports a max resolution of 1400 horizontal pixels by 1050 vertical pixels.

Sync	The sync signal determines where to put the picture on the display. The horizontal sync, or HSYNC, determines where to put the picture in the left-to-right dimension. The vertical sync, or VSYNC, tells the display where to put the picture from top-to-bottom.
Sync on Green	The sync signal for all three colors Red, Green & Blue is derived from Green signal. Typically this used in military display applications.
Sync Generator	A sync generator is a circuit that provides sync signals. A sync generator may have genlock capability.
Synchronous	Refers to two or more events that happen in a system or circuit at the same time.
Uplink	The carrier used by Earth stations to transmit information to a satellite.
Variable Bit Rate	Variable bit rate (VBR) is a bitstream (compressed or uncompressed) which has a changing number of bits each second. Simple scenes can be assigned a low bit rate, with complex scenes using a higher bit rate.
Vertical Resolution	See Resolution.
Vertical Sync	This is the portion of the video signal that tells the video decoder where the top of the picture is.
VGA	Video Graphics Array. Provides 640 x 480 resolution color display screens with a refresh rate of 60Hz and 16 colors displayed at a time
Video Carrier	This is the portion of the video signal that tells the video decoder where the top of the picture is.
Video Interface Port	A digital video interface designed to simplify interfacing of video ICs together.
Video-on-demand	Video-on-demand, or VOD, allows a user to select which program to view at their convenience and playing starts almost immediately. It is commonly called “streaming video” over the internet. For broadcast, satellite and cable networks, it is commonly called “pay-per-view” and is usually confined to specific start times.
VIP	See Video Interface Port.
VSync	See Vertical Sync.
VX Works	A real-time multitasking operating system from Wind River Systems.
X-Windows	Windowing and graphics system standard for video and graphics display.
XVGA	eXtended Video Graphics Array – Provides maximum resolution of 1024 x 768 pixels (non-interlaced) and a 70 Hz refresh rate.
Y/C Video	See S-Video

YCbCr	YCbCr is the color space originally defined by BT.601, and now used for all digital component video formats. Y is the luma component and the Cb and Cr components are color difference signals. The technically correct notation is Y'Cb'Cr' since all three components are derived from R'G'B'. Many people use the YCbCr notation rather than Y'CbCr or Y'Cb'Cr'. 4:4:4 Y'CbCr means that for every Y sample, there is one sample each of Cb and Cr. 4:2:2 Y'CbCr means that for every two horizontal Y samples, there is one sample each of Cb and Cr. 4:1:1 Y'CbCr means that for every four horizontal Y samples, there is one sample each of Cb and Cr. 4:2:0 Y'CbCr means that for every block of 2 × 2 Y samples, there is one sample each of Cb and Cr. There are three variations of 4:2:0 YCbCr, with the difference being the position of Cb and Cr sampling relative to Y.
YUV	YUV is the color space used by the NTSC and PAL video systems. The Y is the luma component. The U and V are the color difference components.

SBS Technologies, Inc., 2400 Louisiana Blvd. NE #5-600, Albuquerque, NM 87110