SBS Technologies

Guide to Video Standards

SBS Technologies
05.20.2005

Introduction

This overview of video standards, methods, and system design factors
provides an historical perspective, starting with the advent of television half a
century ago, and continuing up to today's latest emerging video standards.
Because some applications involve retrofitting 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 1930s 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 one 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. 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. 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.

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.

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 1950s, *color* television emerged, and it required a technique that was backward 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 1960s, 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-343 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 Standard	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	NTSC RS-170A

➡ *Composite vs. Component Video*

All the 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 economical 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 Red, Green, and Blue color 'gun' 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 PCs in the 1980s, 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 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. Unlike broadcast systems, the frame rate is variable as agreed by the source (PC) and the monitor, and not fixed by the standard. 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 will be seen later, VGA resolutions can also be sent 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 wires 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, 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 megabits per second, and the standard is SMPTE 259. (SMPTE abbreviates 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 gigabits 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 coaxial 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 current development is that displays 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 gigabit 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.

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 standards organizations, 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 standards organizations 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 Microsoft® 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. 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 Microsoft® Windows® Media player, and other PC applications, but does not support higher-quality video such as today's DVD standards. Interestingly, the currently 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).

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 paper, 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.

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 1990s. 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 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 relate primarily to motion video. For still pictures, the familiar JPEG standard developed by the Joint Photographic Experts 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 JPEG 2000 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 megabits/ second.
- SMPTE292M for HD (High Definition) at 1.54 gigabits/ second.
- SMPTE310M for sending MPEG-2 digitally in a broadcast studio.

➔ *Between Remote Computers, Over Ethernet*

- Compressed formats: MPEG-2, MPEG-4, Motion JPEG 2000.
- 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 megabits/ second (low resolution teleconferencing and webcasts) to 4 mbits per second (DVD quality video) on up to 19 megabits/ second (HDTV).
- With these rates, it is possible to send multiple streams of video over Ethernet links (even 100BaseT).

It is also possible to send lower-end video over Internet links down to 1 megabit/ second rates.

➔ *Compressed Video to Home*

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 of European origin.

→ *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.