What is a Codec?

Codec is a portmanteau of either "Compressor-Decompressor" or "Coder-Decoder," which describes a device or program capable of performing transformations on a data stream or signal.

Codecs encode a stream or signal for transmission, storage or encryption and decode it for viewing or editing. Codecs are often used in videoconferencing and streaming media solutions. A video codec converts analog video signals from a video camera into digital signals for transmission. It then converts the digital signals back to analog for display. An audio codec converts analog audio signals from a microphone into digital signals for transmission. It then converts the digital signals back to analog for playing.

The raw encoded form of audio and video data is often called essence, to distinguish it from the metadata information that together make up the information content of the stream and any "wrapper" data that is then added to aid access to or improve the robustness of the stream.

Most codecs are lossy, in order to get a reasonably small file size. There are lossless codecs as well, but for most purposes the almost imperceptible increase in quality is not worth the considerable increase in data size. The main exception is if the data will undergo more processing in the future, in which case the repeated lossy encoding would damage the eventual quality too much.

Many multimedia data streams need to contain both audio and video data, and often some form of metadata that permits synchronization of the audio and video. Each of these three streams may be handled by different programs, processes, or hardware; but for the multimedia data stream to be useful in stored or transmitted form, they must be encapsulated together in a container format.

An endec is a similar (but not identical) concept for hardware.

Audio and Video Codecs (A/V codecs)

REDCODE RAW (.R3D)

RAW is a proprietary file format that efficiently encodes measurements from a camera's digital sensor in a way that maximizes post-production capabilities. It achieves this in part by storing each of the sensor's color channels separately, prior to conversion into a full color image. Similar to the advantages that RAW files brought to stills photography, this improves control over white balance, exposure and grading in post-production. Furthermore, since such settings are appended to the file as metadata only, grading is completely

ARRIRAW

Arriraw is a raw codec similar to CinemaDNG that contains unaltered Bayer sensor information, the data stream from the camera can be recorded via T-link with certified recorders like those from Codex Digital or Cineflow.

The ArriRaw format (along with the other recordable formats) contains static and dynamic metadata. These are stored in the header of the file and can be extracted with the free web tool metavisor[12] or with the application Meta Extract provided by Arri. Of particular importance for visual effects are the lens metadata, which are stored only when Arri's lens data system (LDS) is supported by the lens used.

CinemaDNG

CinemaDNG is the result of an Adobe-led initiative to define an industry-wide open file format for digital cinema files. CinemaDNG caters for sets of movie clips, each of which is a sequence of raw video images, accompanied by audio and metadata. CinemaDNG supports stereoscopic cameras and multiple audio channels. CinemaDNG specifies directory structures containing one or more video clips, and specifies requirements and constraints for the open format files, (DNG, TIFF, XMP, and/or MXF), within those directories, that contain the content of those clips.[2]

CinemaDNG is different from the Adobe DNG (Digital Negative) format that is primarily used as a raw image format for still cameras. However, each CinemaDNG image is encoded using that DNG image format. The image stream can then be stored in one of two formats: either as video essence using frame-based wrapping in an MXF file, or as a sequence of DNG image files in a specified file directory. Each clip uses just one of these formats, but the set of clips in a movie may use both.

S-LOG and LOG- C

S- LOG (Sony Log) or Log-C (Arri, but Arri also can shoot S-log)
S-Log or Log-C mode to capture a cinema quality image and accomplish negative film quality latitude. S-Log and Log-C mode captures an image with the greatest amount of range in tonal reproduction and the lowest noise floor and highest ceiling. What this means is you can shoot an image using S-Log or Log-C and have the greatest amount of flexibility when entering the digital intermediate phase of post-production. Using S-Log or Log-C shooting RAW files using 4:4:4 compression creates a digital image that emulates the latitude of negative film.

The image of negative film without color correction looks flat with only midtone colors and little black and white within it. This is what the image will look like with S-Log or Log-C applied because the greatest amount of tonal information is in the midtone range of an image.

Rec. 709

is an ITU Recommendation, first introduced in 1990, that sets out the standards for HDTV. Included in these standards is the Rec. 709 Color Space which is an RGB color space that is identical to the sRGB color space.

LUTS

That is where a Look Up Tables, or LUTs, come into play. An LUT is a digital process, either in the camera, in a mobile recording device, or in a digital intermediate phase that creates a desired image using different exposures charts and digital knees. Like an ENG or prosumer camera's scene file, a LUT can change the look of an image instantaneously. However, scene files will embed into the image, and LUT acts as an overlay to an image where the raw S-Log/Log-C is recorded and not the altered imaged.

AVI

Older codec, an acronym for Audio Video Interleave, is a multimedia container format introduced by Microsoft in November 1992, as part of the Video for Windows technology. AVI files contain both audio and video data in a standard container that allows simultaneous playback. Most AVI files also use the file format extensions developed by the Matrox OpenDML group in February 1996. These files are supported by Microsoft, and are known unofficially as "AVI 2.0".

It is a special case of the Resource Interchange File Format (RIFF), which divides the file's data up into data blocks called "chunks". Each "chunk" is identified by a FourCC tag. An AVI file takes the form of a single chunk in an RIFF formatted file, which is then subdivided into two mandatory "chunks" and one optional "chunk".

The first sub-chunk is identified by the "hdrl" tag. This chunk is the file header and contains metadata about the video such as the width, height and the number of frames. The second sub-chunk is identified by the "movi" tag. This chunk contains the actual audio/visual data that makes up the AVI movie. The third optional sub-chunk is identified by the "idx1" tag and indexes the location of the data chunks within the file.

By way of the RIFF format, the audio/visual data contained in the "movi" chunk can be encoded or decoded by a software module called a codec. The codec translates between raw data and the data format inside the chunk. An AVI file may therefore carry audio/visual data inside the chunks in almost any compression scheme, including: Full Frames (Uncompressed), Intel Real Time Video, Indeo, Cinepak, Motion JPEG, Editable MPEG, VDOWave, ClearVideo / RealVideo, QPEG, MPEG-4, XviD, DivX and others.

Cinepak

Cinepak is a video codec, developed by Radius Inc to accommodate 1x (150 kbyte/s) CD-ROM transfer rates.

It was the primary video codec of early versions of QuickTime and Microsoft Video for Windows, but was later superseded by Sorenson Video, Intel Indeo, and most recently MPEG-4 and h.264. However, movies compressed with Cinepak are generally still playable in most media players.

Cinepak is based on vector quantization, which is a significantly different algorithm from the discrete cosine transform (DCT) algorithm used by most current codecs (in particular the MPEG family, as well as JPEG). This permitted implementation on relatively slow CPUs, but tended to result in blocky artifacting at low bitrates.

DivX®

DivX is a video codec created by DivX, Inc. (formerly DivXNetworks, Inc.), regarded for its ability to compress lengthy video segments into small sizes while maintaining relatively high visual quality. DivX uses lossy MPEG-4 Part 2 compression, where quality is balanced against file size for utility. It is one of several codecs commonly associated with ripping, where audio and video multimedia are transferred to a hard disk and transcoded. As a result, DivX has been a center of controversy because of its use in the replication and distribution of copyrighted DVDs.

Many newer "DivX Certified" DVD players are able to play DivX encoded movies, however, "DivX" is not to be confused with "DIVX", an unrelated attempt at a new DVD rental system employed by the US retailer Circuit City. Early versions of DivX included only a codec, and were named "DivX ;-)", where the winking emoticon was a tongue-in-cheek reference to the failed DIVX system.

DivX, XviD and 3ivx: Video codec packages basically using MPEG-4 Part 2 video codec, with the *.avi, *.mp4, *.ogm or *.mkv file container formats.

DV

Digital Video (DV) is a video format launched in 1996, and, in its smaller tape form factor MiniDV, has since become one of the standards for consumer and semiprofessional video production. The DV specification (originally known as the Blue Book, current official name IEC 61834) defines both the codec and the tape format. Features include intraframe compression for uncomplicated editing, a standard interface for transfer to non-linear editing systems (FireWire also known as IEEE 1394), and good video quality, especially compared to earlier consumer analog formats such as 8 mm, Hi-8 and VHS-C. DV now enables filmmakers to produce movies inexpensively, associated with no-budget cinema.

There have been some variants on the DV standard, most notably the more professional DVCAM and DVCPRO standards by Sony and Panasonic, respectively. Also, there is a recent high-definition version called HDV, which is rather different on a technical level since it only uses the DV and MiniDV tape form factor, but MPEG-2 for compression.

Video compression- DV uses DCT intraframe compression at a fixed bitrate of 25 megabits per second (25.146 Mbit/s), which, when added to the sound data (1.536 Mbit/s), the subcode data, error detection, and error correction (approx 8.7 Mbit/s) amounts in all to roughly 3.6 megabytes per second (approx 35.382 Mbit/s) or one Gigabyte every four minutes. At equal bitrates, DV performs somewhat better than the older MJPEG codec, and is comparable to intraframe MPEG-2. (Note that many MPEG-2 encoders for real-time acquisition applications do not use intraframe compression.)

Chroma subsampling - The chroma subsampling is 4:1:1 for NTSC or 4:2:0 for PAL, which reduces the amount of color resolution stored. Therefore, not all analog formats are outperformed by DV. The Betacam SP format, for example, can still be desirable because it has similar color fidelity and no digital artifacts. The lower sampling of the color space is also a reason why DV is sometimes avoided in applications where chroma-key will be used. However, a large contingent feel the benefits (no generation loss, small format, digital audio) are an acceptable tradeoff given the compromise in color sampling rate.

Audio - DV allows either 2 digital audio channels (usually stereo) at 16 bit resolution and 48 kHz sampling rate, or 4 digital audio channels at 12 bit resolution and 32 kHz sampling rate. For professional or broadcast applications, 48 kHz is used almost exclusively. In addition, the DV spec includes the ability to record audio at 44.1 kHz (the same sampling rate used for CD audio), although in practice this option is rarely used. DVCAM and DVCPRO both use locked audio while standard DV does not. This means that at any one point on a DV tape the audio may be +/- 1/3 frame out of sync with the video. This is the maximum drift of the audio/video sync though it is not compounded throughout the recording. In DVCAM and DVCPRO recordings the audio sync is permanently linked to the video sync.

DVCAM

Sony's DVCAM is a semiprofessional variant of the DV standard that uses the same cassettes as DV and MiniDV, but transports the tape 50% faster, leading to a higher track width of 15 micrometres. The codec used is the same as DV, but because of the greater track width available to the recorder the data are much more robust, producing 50% less errors known as dropouts. The LP mode of DV is not supported. All DVCAM recorders and cameras can play back DV material, but DVCPRO support was only recently added to some models. DVCAM tapes (or DV tapes recorded in DVCAM mode) have their recording time reduced by one third. DVCAM is now also available in HD mode.

DVCPRO

Panasonic specifically created the DVCPRO family for ENG use (NBC's newsgathering division was a major customer), with better linear editing capabilities and robustness. It has an even greater track width of 18 micrometres and uses another tape type (Metal Particle instead of Metal Evaporated). Additionally, the tape has a longitudinal analog audio cue track. Audio is only available in the 16 bit/48 kHz variant, there is no EP mode, and DVCPRO always uses 4:1:1 color subsampling (even in PAL mode). Apart from that, standard DVCPRO (also known as DVCPRO25) is otherwise identical to DV at a bitstream level. However, unlike Sony, Panasonic chose to promote its DV variant for professional high-end applications.

DVCPRO50 is often described as two DV-codecs in parallel. The DVCPRO50 standard doubles the coded video bitrate from 25 Mbit/s to 50 Mbit/s, and improves color-sampling resolution by using a 4:2:2 structure. DVCPRO50 was created for high-value ENG compatibility. The higher datarate cuts recording-time in half (compared to DVCPRO25), but the resulting picture-quality is reputed to rival Digital Betacam, a more expensive studio format.

DVCPRO HD, also known as DVCPRO100, uses four parallel codecs and a coded video bitrate of 100 Mbit/s. Despite HD in its name, DVCPROHD downsamples native 720p/1080i signals to a lower resolution. 720p is downsampled from 1280x720 to 960x720, and 1080i is downsampled from 1920x1080 to 1280x1080 for 59.94i and 1440x1080 for 50i. Compression ratio is approximately 7:1. To maintain compatibility with HDSDI, DVCPRO100 equipment internally downsamples video during recording, and subsequently upsamples video during playback. A camcorder using as special variable-framerate (from 4 to 60 frame/s) variant of DVCPRO HD called VariCam is also available. All these variants are backward compatible but not forward compatible.

Other variants

Sony's XDCAM format allows recording of MPEG IMX, DVCAM and low resolution streams in an MXF wrapper on an optical medium similar to a Blu-Ray Disc, while Panasonic's P2 system uses recording of DV/ DVCPRO/ DVCPRO50/ DVCPROHD streams in an MXF wrapper on PCMCIA-compatible flash memory cards. Ikegami's Editcam System can record in DVCPRO or DVCPRO50 format on a removable hard disk. Note that most of these distinctions are for marketing purposes only - since DVCPRO and DVCAM only differ in the method in which they write the DV stream to tape, all these non-tape formats are virtually identical in regard to the video data.

JVC's D-9 format (also known as Digital-S) is very similar to DVCPRO50, but records on videocassettes in the S-VHS form factor. (NOTE: D-9 is not to be confused with D-VHS, which uses MPEG-2 compression at a significantly lower bitrate)

Digital8 standard uses the DV codec, but replaces the recording medium with the venerable Hi8 videocassette. Digital8 offers DV's digital quality, without sacrificing playback of existing analog Video8/Hi8 recordings.

H.261: Used primarily in older videoconferencing and videotelephony products. H.261, developed by the ITU-T, was the first practical digital video compression standard. Essentially all subsequent standard video codec designs are based on it. It included such well-established concepts as YCbCr color representation, the 4:2:0 sampling format, 8-bit sample precision, 16x16 macroblocks, block-wise motion compensation, 8x8 block-wise discrete cosine transformation, zig-zag coefficient scanning, scalar quantization, run+value symbol mapping, and variable-length coding. H.261 supported only progressive scan video.

H.263: Used primarily for videoconferencing, videotelephony, and internet video. H.263 represented a significant step forward in standardized compression capability for progressive scan video. Especially at low bit rates, it could provide a substantial improvement in the bit rate needed to reach a given level of fidelity.

The Moving Picture Experts Group or MPEG is a working group of ISO/IEC charged with the development of video and audio encoding standards. Its first meeting was in 1988 in Hanover. As of late 2005, MPEG has grown to include approximately 350 members from various industries and universities. MPEG's official designation is ISO/IEC JTC1/SC29 WG11.

MPEG (pronounced EM-peg) has standardized the following compression formats and ancillary standards:

* MPEG-1: Initial video and audio compression standard. Later used as the standard for Video CD, and includes the popular Layer 3 (MP3) audio compression format.
* MPEG-2: Transport, video and audio standards for broadcast-quality television. Used for over-the-air digital television ATSC, DVB and ISDB, digital satellite TV services like DirecTV, digital cable television signals, and (with slight modifications) for DVD video discs.

* MPEG-3: Originally designed for HDTV, but abandoned when it was discovered that MPEG-2 was sufficient for HDTV.
* MPEG-4: Expands MPEG-1 to support video/audio "objects", 3D content, low bitrate encoding and support for Digital Rights Management. Several new (newer than MPEG-2 Video) higher efficiency video standards are included (an alternative to MPEG-2 Video), notably, Advanced Simple Profile and H.264/MPEG-4 AVC.
* MPEG-7: A formal system for describing multimedia content.
* MPEG-21: MPEG describes this future standard as a multimedia framework.

MPEG-1 Part 2: Used for Video CDs, and also sometimes for online video. The quality is roughly comparable to that of VHS. If the source video quality is good and the bitrate is high enough, VCD can look better than VHS, and all in all very good, but VCD requires high bitrates for this. However, to get a fully compliant VCD file, bitrates higher than 1150 kbit/s and resolutions higher than 352 x 288 should not be used. Includes the *.mp3 standard. When it comes to compatibility, VCD has the highest compatibility of any digital video/audio system. Almost every computer in the world can play this codec, and very few DVD players do not support it. In terms of technical design, the most significant enhancements in MPEG-1 relative to H.261 were half-pel and bi-predictive motion compensation support. MPEG-1 supported only progressive scan video.

MPEG-2 Part 2 (a common-text standard with H.262): Used on DVD and in another form for SVCD and used in most digital video broadcasting and cable distribution systems. When used on a standard DVD, it offers good picture quality and supports widescreen. When used on SVCD, it is not as good but is certainly better than VCD. Unfortunately, SVCD will only fit around 40 minutes of video on a CD, whereas VCD can fit an hour. Will also be used on HD-DVD and Blu-Ray. In terms of technical design, the most significant enhancement in MPEG-2 relative to MPEG-1 was the addition of support for interlaced video. MPEG-2 is now considered an aging codec, but has tremendous market acceptance and a very large installed base.

MPEG-4 Part 2: An MPEG standard that can be used for internet, broadcast, and on storage media. It offers improved quality relative to MPEG-2 and the first version of H.263. Its major technical features beyond prior codec standards consisted of object-oriented coding features and a variety of other such features not necessarily intended for improvement of ordinary video coding compression capability. It also included some enhancements of compression capability, both by embracing capabilities developed in H.263 and by adding new ones such as quarter-pel motion compensation. Like MPEG-2, it supports both progressive scan and interlaced video.

MPEG-4 Part 10 (a technically aligned standard with the ITU-T's H.264 and often also referred to as AVC). This emerging new standard is the current state of the art of ITU-T and MPEG standardized compression technology, and is rapidly gaining adoption into a wide variety of applications. It contains a number of significant advances in compression capability, and it has recently been adopted into a number of company products, including for example the PlayStation Portable, the Nero Digital product suite, Mac OS X v10.4, as well as HD-DVD/Blu-Ray.

Theora: Developed by the Xiph.org Foundation as part of their Ogg project, based upon On2 Technologies' VP3 codec, and christened by On2 as the successor in VP3's lineage, Theora is targeted at competing with MPEG-4 video and similar lower-bitrate video compression schemes.

WMV (Windows Media Video): Microsoft's family of video codec designs including WMV 7, WMV 8, and WMV 9. It can do anything from low resolution video for dial up internet users to HDTV. Files can be burnt to CD and DVD or output to any number of devices. It is also useful for Media Centre PCs. WMV can be viewed as a version of the MPEG-4 codec design. The latest generation of WMV is now in the process of being standardized in SMPTE as the draft VC-1 standard.

QuickTime

QuickTime is a multimedia technology developed by Apple Computer, capable of handling various formats of digital video, sound, text, animation, music, and immersive panoramic (and sphere panoramic) images.

The most recent versions are available for the Macintosh and Windows platforms.

QuickTime file format

A QuickTime file (*.mov) functions as a multimedia container file that contains one or more tracks, each of which store a particular type of data, such as audio, video, effects, or text (for subtitles, for example). Each track in turn contains track media, either the digitally encoded media stream (using a specific codec such as Cinepak, Sorenson codec, MP3, JPEG, DivX, or PNG) or a data reference to the media stored in another file or elsewhere on a network. It also has an "edit list" that indicates what parts of the media to use.

Internally, QuickTime files maintain this format as a tree-structure of "atoms", each of which uses a 4-byte OSType identifier to determine its structure. An atom can be a parent to other atoms or it can contain data, but it cannot do both.

The ability to contain abstract data references for the media data, and the separation of the media data from the media offsets and the track edit lists means that QuickTime is particularly suited for editing, as it is capable of importing and editing in place (without data copying) other formats such as AIFF DV, MP3, MPEG-1, and AVI. Other later-developed media container formats such as Microsoft's Advanced Streaming Format or the open source Ogg and Matroska containers lack this abstraction, and require all media data to be rewritten after editing.

QuickTime and MPEG-4

On February 11, 1998 the ISO approved the QuickTime file format as the basis of the MPEG-4 *.mp4 container standard. Supporters of the move noted that QuickTime provided a good "life-cycle" format, well suited to capture, editing, archiving, distribution, and playback (as opposed to the simple file-as-stream approach of MPEG-1 and MPEG-2, which does not mesh well with editing). Developers added MPEG-4 compatibility to QuickTime 6 in 2002. However, Apple delayed the release of this version for months in a dispute with the MPEG-4 licensing body, claiming that proposed license fees would constrain many users and content providers. Following a compromise, Apple released QuickTime 6 on 15 July 2002.

Apple Pro Res Family

Pro Res 422

422 is a standard-definition and high-definition lossy video compression format developed by Apple Inc. for use in post production. It was introduced in 2007 with Final Cut Studio 2 [1] and is comparable to Avid's DNxHD codec which has the same purpose and uses similar bit rates. Both are DCT based[2] intra-frame-only codecs, and are therefore simpler to decode than distribution oriented formats like H.264.
Target data rate of approximately 145 Mbps (1920 x 1080 at 60i)
Higher quality than Apple ProRes 422 (LT)

Full-width 1920x1080 and 1280x720
4:2:2 chroma sampling
10-bit sample depth
I frame-only encoding
Variable bit-rate (VBR) encoding
Normal 147 Mbit/s and High-Quality 220 Mbit/s for HD resolution at 60i
Normal 42 Mbit/s and High-Quality 63 Mbit/s for SD resolution at 29.97
Fast encoding and decoding (both at full size and half size)

Pro Res 422 (HQ)

The Apple ProRes 422 (HQ) codec offers the utmost possible quality for 4:2:2 or 4:2:0 sources (without an alpha channel) and provides the following:
Target data rate of approximately 220 Mbps (1920 x 1080 at 60i)
Higher quality than Apple ProRes 422

ProRes 4444

4444 is a lossy video compression format developed by Apple Inc. for use in post production that can handle standard definition, high definition, and 2K material. It was introduced with Final Cut Studio Pro 7 [1] as another in their line of intermediate codecs for editing material but not for final delivery. It shares many features with Apple's ProRes family of codecs but provides better quality than its predecessors particularly in the area of color.[2] ProRes 4444. For compositing and digital workflows that require the highest-possible image fidelity.

Full-resolution, mastering-quality 4:4:4:4 RGBA color (an online-quality codec for editing and finishing 4:4:4 material, such as that originating from Sony HDCAM SR or digital cinema cameras such as RED ONE, Thomson Viper FilmStream, and Panavision Genesis cameras). The R, G, and B channels are lightly compressed, with an emphasis on being perceptually indistinguishable from the original material.
Lossless alpha channel with real-time playback
High-quality solution for storing and exchanging motion graphics and composites
For 4:4:4 sources, a data rate that is roughly 50 percent higher than the data rate of Apple ProRes 422 (HQ)

Direct encoding of, and decoding to, RGB pixel formats
Support for any resolution, including SD, HD, 2K, 4K, and other resolutions
A Gamma Correction setting in the codec’s advanced compression settings pane, which allows you to disable the 1.8 to 2.2 gamma adjustment that can occur if RGB material at 2.2 gamma is misinterpreted as 1.8. This setting is also available with the Apple ProRes 422 codec.

2K, HD (up to 1920x1080), & SD resoultions
4:4:4 chroma sampling
up to 12-bit pixel depth
Variable Bit Rate (VBR)
Alpha Channel Support

ProRes 422 (Proxy)

For craft editing or offline editing on a MacBook or MacBook Pro.
The Apple ProRes 422 (Proxy) codec is intended for use in offline workflows and provides the following: Roughly 30 percent of the data rate of Apple ProRes 422
High-quality offline editing at the original frame size, frame rate, and aspect ratio
High-quality edit proxy for Final Cut Server

ProRes 422 (LT)

For projects such as news, sports, and multicam events that require reduced file sizes at broadcast quality. Roughly 70 percent of the data rate of Apple ProRes 422 (thus, smaller file sizes than Apple ProRes 422)
Higher quality than Apple ProRes 422 (Proxy)

DNxHD codec

Avid DNxHD, which stands for "Digital Nonlinear Extensible High Definition", is a lossy high-definition video post-production codec engineered for multi-generation compositing with reduced storage and bandwidth requirements. It is an implementation of SMPTE VC-3 standard.[1] DNxHD codec was developed by Avid Technology, Inc. It is comparable with Apple's ProRes 422 which uses similar bit rates and has the same purpose.

Uncompressed high definition digital video has a substantially higher bitrate than standard definition and can require powerful computers to process and edit. Other codecs such as HDV, DVCPRO HD, AVC-Intra, AVCHD, and HDCAM use compression techniques that limit the spatial and temporal resolution of the image. While suitable for acquisition, these codecs will tend to degrade the image over the multiple encode-decode cycles that are typically required during the post-production of complex layered imagery. DNxHD offers a choice of three user-selectable bit rates: 220 Mbit/s with a bit depth of 10 or 8 bits, and 145 or 36 Mbit/s with a bit depth of 8 bits.
DNxHD data is typically stored in an MXF container, although it can also be stored in a Quicktime container. A standalone Quicktime codec for both Windows XP and Mac OS X is available to create and play Quicktime files containing DNxHD material. There is also an experimental support for DNxHD in open source FFMPEG project.

DNxHD is intended to be an open standard, but as of March 2008, has remained effectively a proprietary Avid format. Ikegami's Editcam camera system is unique in its support for DNxHD, and records directly to DNxHD encoded video. Such material is immediately accessible by editing platforms that directly support the DNxHD codec. The source code for the Avid DNxHD codec is freely available from Avid for internal evaluation and review, although commercial use requires Avid licensing approval. It has been commercially licensed to a number of companies including Ikegami, FilmLight, Harris, JVC, Seachange and EVS[2].

DNxHD was first supported in Avid DS Nitris (Sept 2004), then Avid Media Composer Adrenaline with the DNxcel option (Dec 2004) and finally by Avid Symphony Nitris (Dec 2005). Xpress Pro is limited to using DNxHD 8-bit compression, which is either imported from file or captured using a Media Composer with Adrenaline hardware. Media Composer 2.5 also allows editing of fully uncompressed HD material that was either imported or captured on a Symphony Nitris or DS Nitris system. On February 13, 2008 Avid reported that DNxHD was approved as compliant with the SMPTE VC3 standard.[1] In 2007, Apple unveiled ProRes 422, a codec matching many of the features of DNxHD. ProRes lacked a low bandwidth offline resolution like DNxHD 36 until the 2009 release of Final Cut Pro 7. With that release Apple added Pro Res 422 (Proxy) which runs around 45 Mbps, among other additions to ProRes. ProRes is supported for playback on Apple Macintosh and Windows computers, and is supplied and licensed for use when purchased as part of Apple's professional video editing software package, Final Cut Studio, (version 2 or later). DNxHD is available in 8 and 10 bit formats on any system which supports Quicktime. Unlike DNxHD, ProRes 422 provides full functionality at advanced resolutions (2K and 4K cinema) and SD.

Since September 2007 FFmpeg is providing 8-bit (but not 10-bit) VC-3/DNxHD encoding and decoding features thanks to BBC Research who sponsored the project and Baptiste Coudurier who implemented it. It is included in stable version 0.5 of FFmpeg, released on March 10, 2009.[3][4] (ffmpeg -i <input_file> -vcodec dnxhd -b <bitrate> -an output.mov). This allows Linux non-linear video editors Cinelerra and Kdenlive to use DNxHD.

DNxHD is very similar to JPEG. Every frame is independent and consists of VLC-coded DCT coefficients.

Header consists of many parts and may include quantization tables and 2048 bits of user data. Also each frame has two GUIDs and timestamp. The frame header is packed into big-endian dwords. Actual frame data consists of packed macroblocks using a technique almost identical to JPEG: DC prediction and variable-length codes with run length encoding for other 63 coefficients. DC coefficient is not quantized.
The codec supports alpha channel information.

RealVideo

is a proprietary video codec developed by RealNetworks. It was first released in 1997 and as of 2004 is at version 10. RealVideo is widely used by content owners because of its reach to desktops (Windows, Mac, Linux, Solaris) and mobile phones (Nokia Series 60, Motorola Linux, Samsung, Sony-Ericcson, and LG).

RealVideo has historically been used to deliver streaming video across IP networks at low bit rates to desktop personal computers. Today's prevalence of broadband and use of bigger pipes allow video to be encoded at higher bitrates resulting in increased quality and clarity. With mobile carriers, such as Cingular Wireless, starting to offer data services to customers with enabled handsets, video streaming enables consumers to watch video on their mobile phones, be it today's news highlights or even live television.

RealVideo differs from standard video codecs in that it is a proprietory codec that is optimized only for streaming via the proprietary PNA protocol or the Real Time Streaming Protocol. It can be used for download and play (dubbed on-demand) or for live streaming.

RealVideo is often paired with RealAudio and packaged in a RealMedia (.rm) container. The only licensed desktop media player for RealMedia content is RealNetworks' RealPlayer, currently at version 10.5. Unofficial players include MPlayer and Real Alternative.

RealPlayer does not record RealVideo streams, and RealNetworks has advertised this feature to content owners such as broadcasters, film studios, and music labels, as a means of discouraging users from illegally copying content. However, due to the open nature of the Real Time Streaming Protocol, other software exists which can save the streams to files for later viewing.

Sorenson codec

The Sorenson codec (also known as Sorenson Video Codec, Sorenson Video Quantizer or SVQ) is a digital video codec devised by the company Sorenson Media and used by Apple's QuickTime and, in the newest version of Macromedia Flash, a special version called Sorenson Spark.

The Sorenson codec first appeared in QuickTime 3. With QuickTime 4 it was widely used for the first time at the release of the teaser trailer for Star Wars Episode I: The Phantom Menace on March 11, 1999.

The specifications of the codec were not public, and for a long time the only way to play back Sorenson video was to use Apple's QuickTime player, or the MPlayer for Unix/Linux, which in turn piggy-backed Microsoft Windows DLL-files extracted from Apple's player.

According to an anonymous developer1 of FFmpeg, reverse engineering of the SVQ3 codec revealed it as a tweaked version of H.264. The same developer also added support for this codec to FFmpeg, making native playback on all platforms supported by FFmpeg possible.

Sorenson 3: A codec that is popularly used by Apple's QuickTime, basically the ancestor of H.264. Many of the Quicktime Movie trailers found on the web use this codec.

Audio Codecs

AIFF

Audio Interchange File Format (AIFF) is an audio file format standard used for storing sound data on personal computers. The format was co-developed by Apple Computer based on Electronic Arts Interchange File Format (IFF) and is most commonly used on Apple Macintosh computer systems. AIFF is also used by Silicon Graphics Incorporated.

The audio data in an AIFF file is uncompressed big-endian pulse-code modulation (PCM) so the files tend to be much larger than files that use lossless or lossy compression formats such as Ogg and MP3. The AIFF-Compressed (AIFF-C or AIFC) format supports compression ratios as high as 6:1.

WAV

WAV (or WAVE), short for WAVE form audio format, is a Microsoft and IBM audio file format standard for storing audio on PCs. It is a variant of the RIFF bitstream format method for storing data in "chunks", and thus also close to the IFF and the AIFF format used on Macintosh computers. It takes into account some differences of the Intel CPU such as little-endian byte order. The RIFF format acts as a "wrapper" for various audio compression codecs. It is the main format used on Windows systems for raw audio.

Though a WAV file can hold audio compressed with any codec, by far the most common format is pulse-code modulation (PCM) audio data. Since PCM uses an uncompressed, lossless storage method, which keeps all the samples of an audio track, professional users or audio experts may use the WAV format for maximum audio quality. WAV audio can also be edited and manipulated with relative ease using software.

Popularity

As file sharing over the Internet has become popular, the WAV format has declined in popularity, primarily because uncompressed WAV files are quite large in size. More frequently, compressed but lossy formats such as MP3, Ogg Vorbis and AAC are used to store and transfer audio, since their smaller file sizes allow for faster transfers over the Internet, and large collections of files consume only a conservative amount of disk space. There are also more efficient, lossless codecs available, such as Monkey's Audio, TTA, WavPack, FLAC, Shorten, Apple Lossless and WMA Lossless.

Limitations

The WAV format is limited to files that are less than 2 gigabytes in size, due to the way its 32-bit file size header is read by most programs. Although this is equivalent to more than 3 hours of CD-quality audio (44.1 kHz, 16-bit stereo), it is sometimes necessary to go over this limit. The W64 format was created for use in Sound Forge. Its 64-bit header allows for much longer recording times. This format can be converted using the libsndfile library.
[edit]

Audio CDs

Audio CDs do not use WAV as their storage format. The commonality is that both audio CDs and WAV files have the audio data encoded in PCM. WAV is a data file format for computer use. If one were to transfer an audio CD bit stream to WAV files and record them onto a CD-R as a data disc (in ISO format), the CD could not be played in a player that was only designed to play audio CDs.

M-law algorithm

In telecommunication, a **mu-law algorithm** (μ-law) is a standard analog signal compression or companding algorithm, used in digital communications systems of the North American and Japanese digital hierarchies, to optimize (in other words, modify) the dynamic range of an audio analog signal prior to digitizing. It is similar to the A-law algorithm used in Europe.

For a given input *x*, the equation for μ-law encoding is as follows,

$$F(x) = \operatorname{sgn}(x)\frac{\ln(1+\mu|x|)}{\ln(1+\mu)} \qquad -1 \leq x \leq 1,$$

where μ = 255 (8 bits) in the North American and Japanese standards.

μ-law expansion is then given by the inverse equation:

$$F^{-1}(y) = \operatorname{sgn}(y)(1/\mu)[(1+\mu)^{|y|} - 1] \qquad -1 \leq y \leq 1$$

This encoding is used because speech has a wide dynamic range that does not lend itself well to efficient linear digital encoding. Moreover, perceived intensity (loudness) is logarithmic. Mu-law encoding effectively reduces the dynamic range of the signal, thereby increasing the coding efficiency and resulting in a signal-to-distortion ratio that is greater than that obtained by linear encoding for a given number of bits.

The mu-law algorithm is also used in some rather standard programming language approaches for storing and creating sound (such as the classes in the sun.audio package in Java 1.1, in the .au format, and in some C# methods).

Pulse-code modulation (Encoding)

Pulse-code modulation (**PCM**) is a digital representation of an analog signal where the magnitude of the signal is sampled regularly at uniform intervals, then quantized to a series of symbols in a digital (usually binary) code. PCM is used in digital telephone systems and is also the standard form for digital audio in computers and various compact disc formats. It is also standard in digital video.

Several Pulse Code Modulation streams may be multiplexed into a larger aggregate data stream. This technique is called time-division multiplexing, or TDM.

Digitization as part of the PCM process

In conventional PCM, the analog signal may be processed (e.g. by amplitude compression) before being digitized. Once the signal is digitized, the PCM signal is not subjected to further processing (e.g. digital data compression).

Some forms of PCM combine signal processing with coding. Older versions of these systems applied the processing in the analog domain as part of the A/D process, newer implementations do so in the digital domain. These simple techniques have been largely rendered obsolete by modern transform-based signal compression techniques.

- **Differential** (or **Delta**) **pulse-code modulation** (**DPCM**) encodes the PCM values as differences between the current and the previous value. For audio this type of encoding reduces the number of bits required per sample by about 25% compared to PCM.
- **Adaptive DPCM** (ADPCM) is a variant of DPCM that varies the size of the quantization step, to allow further reduction of the required bandwidth for a given signal-to-noise ratio (SNR or S/N).

In telephony, a standard audio signal for a single phone call is encoded as 8000 analog samples per second, of 8 bits each, giving a 64 kbit/s digital signal known as DS0. The default encoding on a DS0 is either μ-law (mu-law) PCM (North America) or a-law PCM (Europe and most of the rest of the world). These are logarithmic compression systems where a 12 or 13 bit linear PCM sample number is mapped into an 8 bit value. This system is described by international standard G.711.

Where circuit costs are high and loss of voice quality is acceptable, it sometimes makes sense to compress the voice signal even further. An ADPCM algorithm is used to map a series of 8 bit PCM samples into a series of 4 bit ADPCM samples. In this way, the capacity of the line is doubled. The technique is detailed in the G.726 standard.

Later it was found that even further compression was possible and additional standards were published. Some of these international standards describe systems and ideas which are covered by privately owned patents and thus use of these standards requires payments to the patent holders.

Some ADPCM techniques are used in Voice over IP communications.

Encoding the bitstream as a signal

Pulse-code modulation can be either return-to-zero (RZ) or non-return-to-zero (NRZ). For a NRZ system to be synchronized using in-band information, there must not be long sequences of identical symbols, such as ones or zeroes. For binary PCM systems, the density of 1-symbols is called 'ones-density'.

Ones-density is often controlled using precoding techniques such as Run Length Limited encoding, where the PCM code is expanded into a slightly longer code with a guaranteed bound on ones-density before modulation into the channel. In other cases, extra 'framing' bits are added into the stream which guarantee at least occasional symbol transitions.

Another technique used to control ones-density is the use of a 'scrambler' polynomial on the raw data which will tend to turn the raw data stream into a stream that looks pseudo-random, but where the raw stream can be recovered exactly by reversing the effect of the polynomial. In this case, long runs of zeroes or ones are still possible on the output, but are considered unlikely enough to be within normal engineering tolerance.

In other cases, the long term DC value of the modulated signal is important, as building up a DC offset will tend to bias detector circuits out of their operating range. In this case special measures are taken to keep a count of the cumulative DC offset, and to modify the codes if necessary to make the DC offset always tend back to zero.

Many of these codes are bipolar codes, where the pulses can be positive, negative or absent. Typically, non-zero pulses alternate between being positive and negative. These rules may be violated to generate special symbols used for framing or other special purposes.

History of PCM

PCM was invented by the British engineer Alec Reeves in 1937 while working for the International Telephone and Telegraph in France.

The first transmission of speech by pulse code modulation was the SIGSALY voice encryption equipment used for high-level Allied communications during World War II from 1943.