SECTION TWO: VIDEO ENCODING STANDARDS

The critical first step in the digital video process is the encoding of analog video to digital form. As mentioned in the last section, this usually requires compression of some information to create a usable digital file. There are many encoding technologies that create digital files that can be stored, served and played back on PC and Mac desktops. The only international standards--developed and ratified according to established procedures by an authorized maintenance agency--are those which have emerged from committees of the Moving Pictures Experts Groups: MPEG-1, MPEG-2 and MPEG-4.

MPEG-1 (ISO/IEC 11172)

The first digital video and audio encoding standard, MPEG-1, was adopted as an international standard in 1992 to provide digital video at bit rates up to 1.5 Mb/sec.* The impetus for the standard was to provide encoding and playback of VHS-quality digital video for CD-ROM playback. MPEG-1 is a progressive-video-sequence encoding standard. The standard implementation for MPEG-1 (known as "constrained bit stream") supports 352 pixels x 240 lines/sec at 30 frames/sec and requires 1.5 Mbps bandwidth for transport. MPEG-1 compression relies on the considerable redundancy of information within and between frames to compress a video object without significantly compromising the integrity of the information it contains.

Video contains spatial, spectral and temporal redundancies, which may be compressed without significant sacrifice in meaning. The encoding techniques in MPEG-1 involve compression based on statistical redundancies in temporal and spatial directions. Spatial redundancy is based on the similarity in color values shared by adjacent pixels. A red sweater in a video frame will generally possess a uniform color value, with little or no perceptual variation from one pixel to the next. MPEG-1 employs intraframe spatial compression on redundant color values using DCT (discrete cosine transform).

Spectral redundancy in video is the similarity between color spectra or "brightness." MPEG-1 operates in the YUB color space. RGB data is converted to YUB. 24-bit RGB is subsampled at 4:2:0 YCrCB, where Y = luminance (brightness) and CrCB = chrominance (color difference). The human eye distinguishes difference in brightness more readily than difference in pure color value.

.

* The standard actually scales higher than 1.5 Mb, but 1.5 Mb is the accepted "sweet spot" for MPEG-1

Temporal redundancy is the sameness in temporal motion between video frames. If frames were not redundant, there would be no perception of smooth, realistic motion in video. MPEG-1 relies on prediction--more precisely, motion-compensated prediction--for temporal compression between frames. MPEG-1 utilizes three frames to create temporal compression—I-Frames, B-frames and P-frames. An I-frame is an intra-coded frame, a single image heading a sequence, with no reference to past or future frames. MPEG-1 compresses only within the frame with no reference to previous or subsequent frames. P-frames are forward-predicted frames, encoded with reference to a past I- or P-frame, with pointers to information in a past frame. B-frames are encoded with reference to a past reference frame, a future reference frame or both. The motion vectors employed may be forward, backward, or both. B-frames are also sometimes known as digital video "spackle."

The MPEG-1 coding standard is a generic standard, intended to be independent of a specific application, serving as a toolbox to be adapted to different applications and their associated hardware and software.

MPEG-2 (ISO/IEC 13818- 2

MPEG-2, published as a standard in 1994, is a high-bandwidth encoding standard, supporting a bandwidth range of approximately 2Mbps to more than 20 Mbps. It was originally designed for coding of television broadcast video with CCIR Rec. 601 resolution at data rates below 10 Mbps, but was expanded to encompass HDTV requirements at app. 12-20 Mbps.

MPEG-2 was designed to encompass, and be backward compatible with, MPEG-1 encoding techniques but was also enhanced to support interlaced video, as provided by television input sources. The MPEG-2 standard was designed for scalability and flexibility, supporting many levels of service depending on the needs of the application. It was expected that an MPEG-3 standard would be developed for HDTV (high definition television), but the MPEG-2 standard scaled to encompass the bandwidth requirements of HDTV.

The most common MPEG-2 compression is main level ("CCIR 601") at 720 pixels x 480 lines, 30 frames/second. The sweet points for MPEG-2 support the bandwidth bit rates of 2-6 Mbps, scaling up to 40 Mbps for very high-level HDTV applications.

The MPEG-2 encoding standard builds on, and is backward compatible with, the statistical redundancy compression of MPEG-1. The most important difference between MPEG-1 and MPEG-2 is the encoding of interlaced frames for broadcast TV. MPEG-1 supports only progressive frame encoding, while MPEG-2 provides both progressive frame and interlaced frame encoding. Video movies, originally in a film format, are a progressive frame format.

Television broadcasts are an interlaced format. A broadcast frame is created with two separate fields, a top and bottom interlaced field, with the first line of the bottom field appearing immediately after the first line of the top field. MPEG-2 splits frames into two fields for interlacing, so that 30 frames/sec becomes 60 fields/sec.

In addition, MPEG-2 includes the ability to multiplex video streams, additional color subsampling, improved compression and error correction and improved audio, including "low sample rate" and multichannel extension for surround sound. The many profiles and levels of service include NTSC (app. 3Mbit/sec), PAL (app. 4 Mbps) and Broadcast HDTV (12-20 Mbps).

MPEG-4 ISO/IEC 14496

MPEG-4, the latest encoding standard from MPEG, was finalized in October 1998 and should be ratified as a standard in the first half of 1999. MPEG-4 arose from a need to have a scalable standard supporting a wide bandwidth range from streaming video at <64 Kbps, suitable for Internet applications, to app. 4 Mbps for higher-bandwidth video needs. MPEG-4 also arose from a desire, as digital encoding matures, to advance beyond simple conversion and compression to object recognition and encoding, as well as the provision of synchronized text and metadata tracks, to create a digital file that carries a meaning greater than the sum of its individual parts.

MPEG-4 supports both progressive and interlaced video encoding. The standard is object-based, coding multiple video object planes into images of arbitrary shape. Successive video object planes (VOPs) belonging to the same object in the same scene are encoded as video objects. MPEG-4 supports both natural ("analog") and synthetic ("computer-generated") data coding. Some VRML technology is incorporated to encode dimensionality.

MPEG-4 compression provides temporal scalability utilizing object recognition, providing higher compression for background objects, such as trees and scenery, and lower compression for foreground objects, such as an actor or speaker—much as the human eye filters information by focusing on the most significant object in view, such as the other party in a conversation. Object encoding provides great potential for object or visual recognition indexing, based on discrete objects within a frame rather than requiring a separate text-based or storyboard indexing database. In addition, MPEG-4 provides a synchronized text tract for courseware development and a synchronized metadata track for indexing and access at the frame level.

Proprietary Formats

MPEG-1 and MPEG-2 were developed for reliable moderate to high-bandwidth transport. Neither standard successfully supports streaming over the Internet, particularly at the common modem speeds of 28.8 to 56 Kbps available for personal Internet use. Encoding formats proprietary either to a microcomputer platform or a specific manufacturer arose to provide streaming digital video to Internet users at low-bandwidth ranges. MPEG-4 is expected, over time, to displace many proprietary formats but at present, proprietary "de facto" standards are well-established on the Internet.

QuickTime began as a Mac-based video encoding, file management and playback system but with version 3.0 became a cross-platform encoding format, supporting digital video on Mac and Windows. QuickTime video files have the file extension .mov. QuickTime is a versatile digital video encoding format, supported by a range of commercial and shareware software products, including encoding, editing, and client plug-ins. QuickTime 4.0, released in April, 1999, supports timecode tracks and Web transport and streaming protocols, including HTTP, RTP and RTSP. QuickTime 4.0 provides built-in support for digital video, including MiniDV, DVCPro, and DVCam camcorder. The QuickTime digital video file format was selected as the basis for MPEG-4. Although the multi-track, object-based MPEG-4 goes beyond QuickTime functionality, the QuickTime wrapper will be supported by MPEG-4-based service and streaming, making this a safe interim choice for low-bandwidth videos over Internet.

Microsoft ActiveMovie (AKA Video for Windows), provides digital video technology, including encoding, file naming and playback, on a Windows platform. The file format is identified by the .avi (audio & video interleave) extension. Video and multimedia creation and editing packages for use on Windows platforms, such as Adobe Premiere® must support the .avi file format

Microsoft ASF (Advanced Streaming Format) is an open streaming format developed collaboratively by Microsoft, Progressive Networks, Inc., Intel Corp., AdobeSystems Inc. and Vivo Software Inc., as well as from the feedback and suggestions of other companies. It is currently available in beta in version 1.0 for use with Microsoft's NetShow streaming media client/server software and in version 2.0 as a preliminary developers' toolkit. The ASF format is intended for streaming synchronized audio, video and multimedia for use over the Internet.

Progressive Networks' RealVideo (file format extension .rm or .ram) is a robust low-bandwidth format intended for Internet streaming. RealVideo is supported by a range of commercial and shareware products including encoding software, multimedia authoring tools, server software and client plug-ins. The G2 Real Media client supports SMIL synchronized text files, audio only, and plug-in extensions for MPEG. RealVideo is in widespread use and is supported by a complete client/server suite, including publishing, synchronized multimedia, and streaming server software, available in shareware and inexpensive commercial versions.

Other encoding formats proprietary to different manufacturers include Intel's Indeo® format and **Cinepak**, first developed by SuperMac Technologies and now owned by Radius.

M-JPEG is a quasi-standard provided by many video encoding cards. M-JPEG consists of sequential JPEG-encoded frames. JPEG stands for Joint Photographic Experts' Group and is the popular name for the still image encoding standard JFIF (JPEG File Interchange Format). JPEG is an intraframe compression standard intended for still images only. Video encoding cards provide M-JPEG compression so that the resulting digital video file may be edited. MPEG-1 and MPEG-2 files include I-frames, which are compressed intraframe, but also P-frames and B-frames, which do not include essential information, such as color or movement, but instead reference that information in a forward or backward frame. Since frame types are not eye-readable, video editing can result in the removal of critical reference frames. Some video cards now provide editable MPEG-1, eliminating the need for M-JPEG encoding. Editable MPEG-2 was introduced by Hewlett-Packard in 1997 for its HP MediaStream broadcast server.