

# Optimizing the Digital Video Publishing Process: Coupling MPEG Video Encoding and Video Editing

By Mauro Bonomi

*The video distribution industry is going digital, opening new markets and distribution opportunities for content providers. Video compression, also known as encoding, makes digital video distribution viable, reducing video data down to fit bandwidth-constrained media such as compact discs (CDs) and integrated services digital networks (ISDN) lines. Little or no coupling has been developed between video editing and MPEG video compression. Standard edit decision lists (EDLs) feature information very useful to the encoding task. Coupling editing and encoding will yield superior image quality and productivity and will help the industry introduce new and compelling interactive products.*

Digital video distribution via video CD, CD-ROM, and direct broadcast satellite (DBS) is experiencing a vibrant growth. Several million multimedia PCs with MPEG-1 decoding capabilities and over 6 million video CD players have been sold through the end of 1996. The newly introduced digital video disc (DVD) players, thanks to their performance, superior storage capacity, and video quality, will have an immediate impact on the consumer industry. Over 4 million DBS set-top boxes had already been deployed by the end of the year.

Digital video delivery via video servers over wide-area networks (e.g., corporate intranets), cable and telco networks is also growing rapidly. As this trend towards digital video distribution progresses, digital video compression assumes a pivotal role. Video compression makes digital video distribution viable as it reduces video data down to fit bandwidth-constrained media delivery systems.

## Compression Standards

The Moving Picture Experts Group (MPEG), a joint technical committee of the International Standards

Organization (ISO) and the International Electrotechnical Commission (IEC), have defined the MPEG-1 and MPEG-2 compression standards. MPEG-1 provides for VHS quality at bit rates as low as 1.14 Mb/secs. Higher and lower rates are possible with corresponding trade-offs in quality. MPEG-2 provides broadcast-quality video at data rates as low as 6 to 8 Mb/secs. MPEG-2 is a superset of MPEG-1, so MPEG-2 decoding systems will also play MPEG-1 video streams. At a target video data rate of 1.14 Mb/secs, MPEG-1 provides a compression ratio of nearly 150:1, matching the data rate of single-speed CD-ROM drives and enabling 72 min of video and audio to be packed onto one CD.

MPEG standards offer a range of compression ratios suitable for delivering media beyond CDs. Naturally, the greater the compression, the lower the image quality. While the inverse relation of compression ratio to quality is inescapable, the quality versus bit-rate curve as a whole can be raised by applying sophisticated video pre-processing and encoding techniques. Effective encoding provides significant improvements in quality at a given bit rate, or lower bit rates at a given level of quality.

## Broadcast vs. Publishing Encoders

Digital video distribution will have a significant impact on the video production and post-production industry.

With the possible exception of sports, video will be increasingly delivered asynchronously with embedded "headers" to enable filtering and interaction by viewers. More and more video will be "touched" before distribution. This trend will affect the architecture of MPEG encoding systems.

Two different types of MPEG encoders have been developed by several manufacturers: broadcast encoders and publishing encoders. Broadcast encoders are optimized for the compression of "live" feeds, such as sporting or news events. Publishing encoders by contrast are used to convert "recorded" video into compressed video, which is then stored on a CD or on a video server for later distribution via a cable, telephone, or DBS system. Broadcast encoders assume a traditional analog distribution model, while publishing encoders assume the emerging digital video publishing metaphor, with stored video being accessed on demand and interactively by viewers.

Removing the "live" encoding constraint means one can apply a more sophisticated compression process, tightly coupled with the editing process, and achieve superior compression efficiency. A 20 to 30% gain in compression efficiency commands tremendous premiums: more channels per transponder for DBS applications; huge storage savings in near-video-on-demand (NVOD) applications; superior perception of value and product quality in training, education, games, and music video applications; and faster and cheaper access to "video-rich" content on the Intranet and Internet.

## From Publisher to Viewer: The Digital Video Publishing Process

Digital video distribution will ignite the production of new content and the repurposing of old content. Interactive digital video distribution

Presented at the 137th SMPTE Technical Conference (paper no. 137-17) in New Orleans, La., September 6-9, 1995. An unedited version appears in *Moving Images: Meeting the Challenges*, SMPTE, 1995. Mauro Bonomi is with Minerva Systems, Santa Clara, CA 95054. Copyright © 1997 by the Society of Motion Picture and Television Engineers, Inc.

## OPTIMIZING THE DIGITAL VIDEO PUBLISHING PROCESS: COUPLING MPEG VIDEO ENCODING AND VIDEO EDITING

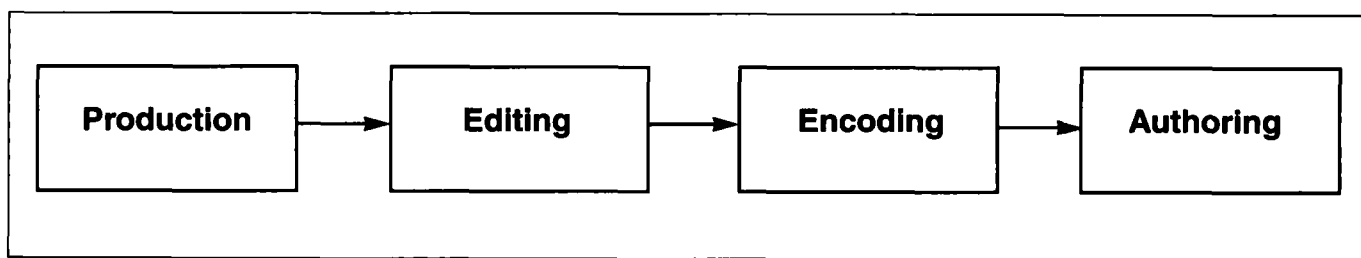


Figure 1. The digital video publishing process.

will enable us to look at old content in new ways. Repurposing goes hand-in-hand with the birth of any new medium. Film reused plays, and TV recycled movies. The true mastering of a new medium can take a long time, but multimedia productions are growing up quickly. Digital video publishing implies new steps in the post-production process, such as encoding and authoring (Fig. 1).

Little or no coupling has been developed between video editing and MPEG video compression. Encoding is performed without leveraging the information on the video sequence gained during the editing process. Coupling the two processes will yield superior image quality and productivity.

### Efficient MPEG Encoding

MPEG performs lossy compression; that is, it discards data. To achieve a compression ratio of 150:1, for example, MPEG-1 must eliminate more than 99% of the data in a video stream. The data loss can cause visible by-products, known as artifacts, which may include poorer image definition and in some cases blocking, ringing, and jerky motion.

The goal of efficient encoding is to minimize the impact of such artifacts. "Human-assisted" encoding techniques have been developed to maximize compression efficiency. In human-assisted encoding, a compression specialist, the "compressionist," applies techniques such as spatial and temporal filtering, inverse telecine, and I-frame placement to reduce artifacts and boost quality.

Human-assist may involve subjective aesthetic judgment. The compressionist sets an encoding parameter, views the impact on the video scene, and further modifies the parameter until the desired result is achieved. When the compressed video achieves

the desired quality, the encoding decision is entered on an encoder control list (ECL). Similar to an edit decision list (EDL), the encoder control list spells out all of the encoding decisions and the time codes at which they are to be executed. The ECL serves as a script for the encoding system, which automatically carries out the desired functions at the time of encoding. The ECL contains both global and time-code-specific parameters.

Global parameters include the source video resolution and format (NTSC or PAL), picture size, compressed video bit rate, audio sampling

---

*"Broadcast encoders assume a traditional analog distribution model, while publishing encoders assume the emerging digital video publishing metaphor, with stored video being accessed on demand and interactively by viewers."*

---

rate, compressed audio bit rate, and the desired multiplexed bitstream format. Time-code-specific parameters include In and Out points, telecine phases (for NTSC film-originated content), I-frame placement, entrypoints into the compressed stream, and filtering parameters. Time-code-specific parameters require the compressionist to evaluate scene by scene, and at times frame by frame, the video being

encoded. Standard EDLs, supported by mainstream video editing equipment, feature time-code-specific parameters (e.g., scene cuts) that could guide the compressionist in generating the ECL. Therefore, a close coupling of editing and encoding systems would significantly boost the throughput of the MPEG compression process.

### Coupling MPEG Encoding and Video Editing: Examples

Optimum encoding parameters can be selected manually through a human-assisted process, or they can be partially derived from the video sequence EDL. Standard EDLs contain information on scene cuts, fades, and wipes, which can be used to guide the encoding process. Tight coupling between the editing and the encoding process would increase compression quality and throughput, as demonstrated in the following examples.

#### I-frame Placement

An MPEG stream consists of three kinds of pictures. Intra pictures (I frames) are encoded using only information internal to the picture (intra-coding). I frames, in turn, serve as a reference for predicted pictures (P frames) and bidirectionally predicted pictures (B frames). In predicted and bidirectionally predicted frames, compression is attained by encoding only the difference between the current picture and the reference frame. I, P, and B frames alternate periodically.

Prediction takes advantage of the fact that picture content often varies little from frame to frame. When a scene cut occurs, however, the I, B, P pattern should ideally be reset, and the first frame of the new scene should be encoded as an I frame. An I frame ensures that the new scene starts with a high-quality image. Scene cuts are

---

*“Little or no coupling has been developed between video editing and MPEG video compression. Encoding is performed without leveraging the information on the video sequence gained during the editing process. Coupling the two processes,, will yield superior image quality and productivity.”*

---

recorded in standard EDLs. Therefore, compression efficiency would benefit from importing the time codes for scene cuts from the EDL and turning them into time codes for I-frame placement in the ECL.

#### ***Fades, Wipes, and Graphics Overlay***

Within fades and wipes the correlation among adjacent frames is very limited, and prediction does not work effectively. Picture quality tends to suffer during fades and wipes. Similarly, graphic overlays can create sharp edges, which may result in ringing artifacts. Standard EDLs carry wipes, fades, and graphic overlay information that will help the compressionist identify areas where special prefiltering should be used to maximize image quality.

Filtering prior to encoding can be used to screen out image data that, if presented to the encoding engine, would result in artifacts. Spatial filtering applies within a frame, and can be used to soften sharp edges (e.g., graphics, scrolling titles, etc). The resulting picture may have a softer appearance, but this is often preferable to ringing artifacts. Similarly, temporal filtering can be used to smooth fast transitions (e.g., fades). Motion may be slightly blurred as a consequence, but this is nevertheless better than visible blocking artifacts.

#### ***Inverse Telecine***

Encoding of video NTSC material derived from film can be greatly improved if the duplicate fields inserted during the telecine process are first removed. This removal, called inverse telecine, has several benefits: it does

away with redundant fields, allowing the compression system to allocate more bits to the remaining unique fields; and, most important for MPEG-1 encoding, it avoids motion artifacts that might otherwise occur if the 3:2 sequence is left intact.

Some encoding systems do feature the ability to perform 3:2 extraction, automatically removing fields in a fixed pattern. This works fine as long as the extraction matches the rhythm of the original telecined video. The problem is that nearly every video derived from film contains disruptions in the telecine cadence; places where film reels were joined or where edits were made subsequent to video transfers. Spotting these phase changes, and readjusting the 3:2 extraction accordingly, is very challenging.

A superior encoding system can accomplish this task in several ways. It could provide some form of automated 3:2 pattern detection. Scene cut information from the video EDL, when available, should be used to pinpoint the time code where the phase change occurred. EDL-assisted phase detection would drastically improve the accuracy of the inverse telecine process, thereby reducing manual intervention. When varispeed is used in the editing process, it is very difficult even for the most sophisticated telecine detection unit to determine the 3:2 cadence. Access to the EDL varispeed parameters is then essential to performing the best possible inverse telecine extraction.

#### **Conclusion**

MPEG specifies the encoding syntax, therefore ensuring standardization at the decoder (viewer) end. MPEG

does not specify the means by which encoding is performed. The implementation details are left up to the encoding system designer, allowing for a great deal of differentiation among MPEG encoding systems. A powerful and flexible encoding engine, coupled with human-assist application software tools, is necessary to deliver good-quality encoding. As digital video distribution becomes pervasive, high encoding throughput is required. A streamlined digital video publishing process requires close coupling between video editing and video encoding systems. Tight coupling between the encoding and the editing processes will deliver both compression efficiency and high encoding throughput.

---

#### **THE AUTHOR**



**Mauro Bonomi** is chief executive officer of Minerva Systems. Prior to founding that company, Bonomi served on the executive staff of C-Cube Systems, where he was instrumental in the definition and marketing of leading JPEG and MPEG very large-scale integration (VLSI) encoding solutions.

Bonomi holds master's degrees in electrical engineering and management from Stanford University. He also received bachelor's and master's degrees in engineering and applied mathematics from Pavia University, Italy.