

VIDEO RECORDING: CHIPS, DISKS, OR TAPE?

By - Phil Livingston, Director, Digital Systems Panasonic
and
Richard A. Albert, Manager, DVCPRO Systems, Panasonic

Excerpted from the Panasonic booklet entitled:

DVCPRO and DIGITAL VIDEO RECORDING

A Relatively Unbiased Explanation of the Relationship of Video Recording Formats, Compression, Digital Television, and the Like from a Manufacturer's Perspective.

April 1997

In this digital age, there is a bewildering array of possible recording systems for television signals. Digital video and audio can be recorded in solid-state memory chips; on magnetic, optical, or magneto-optical disks; and on magnetic tape. There is no one single recording system that can be optimal for all applications.

Consider the ubiquitous VHS format, for example. Inexpensive and almost universally available, it's a great format for home video and a good format for television creative staff to use for reviewing rehearsals at home. However, VHS would be a terrible choice for multi-layer post-production.

Consider solid-state memory: There's no better way to store a video frame temporarily for a digital video effects system. Zooms, rotations, flips, page turns, and revolving cubes would all be impossible without solid-state frame stores. However, the capacity of chip-based storage is so low that it would be a terrible choice for archiving even an hour of television programming.

Disks tend to fall between chips and tape in capabilities. Disk-based storage allows rapid access to any point (though not as instantaneous as memory chips), making disk systems ideal for non-linear editing. Arrays of disks allow simultaneous multiple access to the same material, giving them high potential for servers. In fact, disks could be the ideal medium for all forms of television storage if only they could combine the low cost of the stamped optical versions like CD-ROM or DVD with the recording simplicity of the magnetic hard drive varieties. Unfortunately, the right combination of low cost, high capacity, and recording capability does not exist in any disk system at the present time.

At NAB '95, Avid and Ikegami introduced the world's first disk-based camcorder. While a remarkable technological achievement, it is just about to go on sale¹. It certainly has yet to change the industry. The vast majority of the video material edited on Avid systems worldwide is still input from tape. Why? Tape is ideal for archiving and acquisition, offering the best cost/capacity of any recordable television storage medium.

In an ideal world, shooting would be on tape, the material would then magically appear on disk for editing, and, after editing, the edited material would magically reappear --fully edited -- on a convenient archiving tape. No recording format offers such capability yet, but Panasonic's DVCPRO, with high-speed playback and record capability, comes awfully close. However, that's getting a little ahead of the story.

¹ Broadcasting & Cable, March 17, 1997. Page 110

DATA TAPE VS. DIGITAL TAPE

The video industry isn't the only one with a need for data storage. The magnetic disk drives used in non-linear editing systems and video servers were all developed for the computer industry. Couldn't the same signals also be recorded on computer *tape* drives? As long as a tape system has sufficient data-transfer-rate capability, it could indeed store digital video as easily as it can hold credit-card transactions.

With the advent of video compression, recorders with lower data-transfer rates became capable of recording digital audio and video signals. In recent years, such *data* tape recorders have been appearing at NAB with ever increasing frequency. Data tape recorders *do* have a place in the video industry. Graphics devices have long used data tape systems for archiving and interchange. There may also be an appropriate application for data tape recorders in large robotic libraries providing "near-line" and "off-line" archival storage for video file server systems. The video file servers will deal with the necessary audio, video, and control interfaces; all the data recorders need do is record and play undifferentiated data.

A videotape recorder -- even the digital variety -- must do more than that. Typically, DVTRs deal with time code, offer pictures during shuttle modes, and some DVTRs offer broadcastable quality "slow motion." Every format should be controllable by a computer-based editing system and offer editing of anything recorded on a tape -- any number of frames in duration -- to be replaced by new material in a perfectly synchronous insert edit.

CATEGORIZING DIGITAL VIDEOTAPE RECORDERS

Even after excluding chip, disk, data and all analog videotape formats, buyers are still faced with a huge selection of DVTR formats. To make sense of the choices, categorize them. Three formats (D-6, HDCAM, and the HDD-1000) are designed specifically for high-definition television (HDTV) signals. In addition to these, Panasonic's Digital HD Processor turns a standard definition D-5 VTR into an HDTV recorder at lower cost than D-6 or the HDD-1000 and with less compression than Sony's HDCAM. BTS/Philips has a similar device for D-1.

D-2 and D-3 were designed for composite digital video signals. Again, with special processing equipment, these formats can be made to record other forms of signals. Panasonic will offer a system with the Zenith developed processor that will record encoded FCC-standard digital television (DTV) data streams on standard D-3 tapes, be those data streams HDTV, SDTV, or multi-program SDTV.

Ten formats record digital *component* video signals based on the sampling rates in the International Telecommunications Union document referred to as Recommendation 601, but only D-1 and D-5 record *uncompressed* component digital video. Panasonic's D-5, in fact, has the highest data recording rate of any standard-definition DVTR, and is suitable for applications demanding perfect transparency. Some applications can utilize smaller, more cost-effective DVTR formats, and eight formats (Betacam SX, DCT, Digital Betacam, Digital-S, DV, DVCAM, DVCPRO, and DVCPRO 50) all utilize digital bit-rate reduction, also known as compression.

A COMPRESSION PRIMER: THE TWO BASIC RULES OF COMPRESSION

Compression "throws away" some portion of the data representing the digital video information. A format with about 2:1 compression, like Digital Betacam, throws away about 50% of the information. A format with 5:1 compression, like DV, throws away 80% of the information. A format with 9:1 compression, like Betacam SX, throws out almost 90% of the original information. Is this bad? The only truthful answer is, "It depends." The results of compression depend in part on the specific source material being compressed, and this is true of all video compression. It also depends on the intended application.

1. It is not possible to compress truly random noise.

Every compression system, from 19th-century Morse Code to the latest circuits and software, tries to take advantage of the fact that real world signals are not truly random. Mr. Morse knew text tends to be weighted heavily towards the letter *e*, which is sent as one “dot.” One frame of a video signal tends to be very similar to the next, and individual picture elements (pixels) tend to be similar to their neighbors. As long as those tendencies hold up, compression becomes efficient. However, when they don't (which is the case with random noise), some compression schemes can become very inefficient or fail.

Since no compression system can compress random noise, there's no such thing as a totally lossless or mathematically lossless compression system for *all* possible input signals. That's also why it's possible to create a “compression breaker” sequence of source material for *any* compression system to show off its flaws, or conversely to select material to show off its virtues.

2. Real-world video signals are not random noise.

In fact, ordinary video is far from random. Again, successive frames and neighboring pixels usually have similarities. This principle of similarity allows video compression systems to function. Often, however, the principle is violated. At a camera cut, for example, one frame is very different from the next. To accommodate these violations of the principle of sameness, compression systems use memories that store excess data until such time as there is room to accommodate it. If that time or space cannot be found, the excess data may be discarded. Compression system design, therefore, often involves compromises of picture quality, data rate, buffer memory size, and issues of latency.

LOSSLESS, MATHEMATICALLY LOSSLESS, AND VISUALLY LOSSLESS

Some manufacturers use terminology like “lossless or mathematically lossless” for their forms of compression. While no compression system can accurately compress random noise, no human being can tell visually whether or not random noise is defective. Improperly compressed random noise can well appear to be visually identical to perfect random noise. Video compression system design, therefore, relies on applications as well as mathematical algorithms. An image that has deteriorated in the compression and decompression process but does not *appear* to have deteriorated may be considered visually lossless. However, the full original quality can never be recovered, and this can be considered a lack of signal “robustness”.

Unlike other forms of compression, video compression differs for different applications within the vast video field. DVD, for example, as a medium not requiring any insert editing, actually allows a variable data rate. Only a fixed amount of data can be recorded on a single disk, but, as long as everything fits on the disk, complex scenes can utilize a much higher data rate than simpler ones.

DCT - THE TRANSFORM: THE KEY TO COMPRESSION

Small blocks of pixels are converted from the spatial domain to the frequency (detail) domain by a mathematical process known as a discrete cosine transform (DCT). There have been other forms of image compression developed, including wavelet- and fractal-based forms, but DCT is used in all compressed DVTRs from *all* manufacturers. It's also used in most disk-based video compression systems. It is the basis for MPEG, JPEG, M-JPEG, DV, Betacam SX, Digital-S, DVCAM, DVCPRO, DVD, HDCAM, Digital Betacam, and Ampex's identically named DCT.

As is well known, images have the least energy as the detail increases. After the DCT conversion, therefore, those portions of the DCT signal with high detail will tend to be either zero or near zero. Zeroes are compressed efficiently by a technique known as *run-length* encoding, wherein a code simply identifies

how many zeroes follow. If there is sufficient data rate to transmit or record all of the non-zero numbers, the compression will have taken place losslessly.

If there is *not* sufficient data rate to transmit or record all of the non-zero numbers -- and if the buffer memory system can't find some place to hide that additional data (a point to be revisited later) then some of the non-zero numbers will likely be changed into zeroes and *lossy* compression will take place. Those changed numbers will tend to be in the finest detail where the human visual system is least sensitive. Therefore, lossy compression may often be almost visually lossless, but still impair video quality.

MPEG, M-JPEG, AND DV COMPRESSION

JPEG is the oldest common form of DCT-based image compression. It takes advantage of the compressibility of individual images, but it doesn't take advantage of similarities between frames because it was originally developed for still picture transmission. When disk-based non-linear video editing systems began to appear on the market, JPEG was the simplest compression "engine" to use. Since the sequences had motion, the compression was sometimes referred to as "Motion-JPEG" or M-JPEG.

In the meantime, MPEG (Moving Pictures Expert Group) committees were working out the various standards that go by the names MPEG or MPEG-2. This compression was initially designed for mass-market applications not requiring any editing or recording, e.g. transmission to TV sets or encoding of moving sequences on CD-ROM. It is designed to be *asymmetrical*, that is, MPEG uses an expensive encoder to allow inexpensive consumer decoders. In fact, there is not a single MPEG-2 standard, but a large family of standards, consisting of different *profiles* operating at different *levels*. For example, The DVD digital video disk standard utilizes MPEG-2 Main Profile at Main Level (MP@ML).

MPEG is also based on a *group of pictures* (GOP) concept: Since one frame is similar to the next in a typical video sequence, it's possible to make a pretty good prediction of what the next frame in a sequence will look like. It's also possible to bi-directionally interpolate frames based on those that came before and after. MPEG compression allows for three types of frames: I-frames (compressed entirely within a frame), P-frames (based on predictions from a previous frame), and B-frames (bi-directionally interpolated from previous and succeeding frames). These frames are strung into groups of pictures. This is the key to the efficiency of MPEG. Problems arise, however, when a GOP is interrupted.

Suppose the GOP arrives at the decoder missing its initial I-frame. The Predicted frames based on the I-frame have now lost their basis, as have the B-frames based on those P-frames or the I-frame. The result might be a half-second delay before the decoder can present pictures -- a nuisance, but not a tragedy, for the home viewer. However, the MPEG compression efficiency, based on large GOPs, is lost in the video post-production environment. Every edit is likely to destroy the sequence of a GOP. Therefore, *if* MPEG is to be considered for editing, it must be considered with very small GOP sizes -- ideally, "frame-bound" or I-frames only.

To reiterate, the advantages of MPEG -- low-cost decoders (but high-cost encoders) and compression efficiencies brought about by large GOPs utilizing P-frames and B-frames -- fall apart in the professional video post-production environment. Don't take our word for it. Here's what Pluto's Thomas R. Goldberg wrote:

"MPEG, with its 30:1 usable compression ratio, knocks the huge bandwidth of uncompressed video down to where it requires very few drives to accommodate it. Long pieces of media easily fit in small amounts of storage. When scaled back to 20 or 15:1 it can even look good by studio standards, although *it is hard to edit in studio applications*. But, editing aside, MPEG has this real problem --it is asymmetrical

(sometimes called non-complimentary). That is, even though decoders can be built very inexpensively, encoders are still very complex and expensive" (emphasis added)².

MPEG 4:2:2 - STUDIO MPEG

MPEG-2 MP@ML uses a 4:2:0 structure allowing only half as much detail in color as in brightness information in *both* the vertical and horizontal directions and a maximum data rate of just 15 Mb/s. Most people do not think that's good enough for professional video processing applications. Avid's Katie Cornog wrote:

"This is generally insufficient to code a video stream with only I-pictures and maintain a high level of video quality."³

In fact, no original MPEG profile was suited in any way to professional studio applications, so a new profile was created, the MPEG 4:2:2 Profile, sometimes called the Studio profile. It allows data rates up to 50 Mb/s, sufficient to allow I-frame-only coding with good quality. At 50 Mb/s, however, MPEG I-frame-only encoding offers no significant advantages over any other form of intra-frame coding, such as JPEG or DV. And the only DVTR currently offering MPEG 4:2:2 Profile, Sony's Betacam SX, encodes at just 18 Mb/s, a rate found to be insufficient for "visually lossless" quality in the Avid article, even with the IB GOP used in the format. To quote Ms. Cornog again:

"To produce very high quality results using all intra-frames on 4:2:2 based video requires from 20 to 40 Mb/s, depending on the complexity of the images."⁴

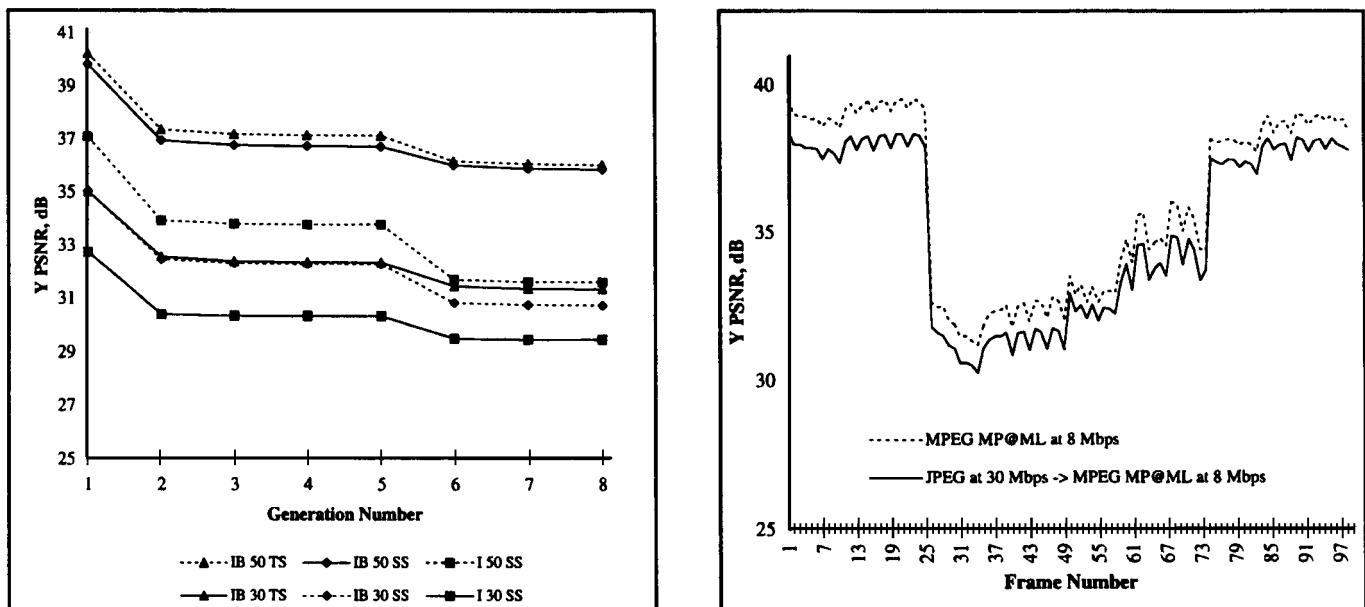


Figure 4. Bit rate needed for various GOP's to produce "visually lossless" quality. Four different clips are shown. Each clip was encoded using five different GOP's: I-only, IP, IB, IBPB, and N=15, M=3. But rate was adjusted to achieve an average PSNR of 40 dB on each clip. "Mobile" and "Flower Garden" are complex, while "Billboard" and "Swanboats" are moderately complex. The source material was Betacam SP.

² March 13, 1997 Pluto, Inc. White Paper, Page 5

³ SMPTE Journal, February, 1997. Pages 95-102

⁴ Ibid.

One important conclusion can be quickly drawn from Ms. Cornog's article quoted above. The sacrifice of editability by using a GOP greater than 1 comes at a small improvement in data rate - about 20%! The requirement for 20% more bandwidth or storage space in an era when processors are getting ever faster and disk drives ever larger is minuscule, while editing is the crux of post production. Secondly, even long, complex GOP structures only gain 40 to 50% - good, but still not good enough for transmission if future post production is contemplated. Why? The compressed stream can be decoded and restored to baseband video to fix the GOP editability "problem," but the quality will never equal the original!

The MPEG-2 4:2:2 profile was conceived to allow commonality between studio equipment and such distribution media as DVD, direct satellite broadcasts (DBS), and the recently approved terrestrial digital television (DTV) broadcasting scheme. Unfortunately, there is no way to convert from an MPEG-2 4:2:2 profile data stream to any other MPEG or MPEG-2 data stream without first returning to uncompressed digital video, at least at the present time, giving MPEG-2 4:2:2 profile no advantage in relation to subsequent MPEG processing. Equally unfortunate is the mis-information in our industry. Read, for example what Matthias Zahn of Fast, Inc. wrote:

"Consider that this new standard provides ... backward compatibility with Main Profile. and... While early subsets of this standard such as ... DVCPRO are intended to bring this technology to market early, it is clear that the parent technology should ultimately prevail."⁵

While this article has many good points, such as the use of SDI to carry compressed video, it implies there is a seamless way to convert or inter-operate between MPEG-2 Main Profile and MPEG 4:2:2 Profile. No such process currently exists - transcoding by returning to raster-based video, albeit digital video, is still required.

DV COMPRESSION: EQUIPPED FOR THE FUTURE

Like JPEG and MPEG, DV compression is also DCT-based. As the newest form of DCT-based compression, however, it was able to improve on JPEG and MPEG techniques, utilizing *feed forward* instead of *feed back* techniques, for example. Most importantly, DV was designed from the very beginning as an extensible format, able to accommodate standard- definition and HDTV. Of the compressed standard-definition DVTR formats being sold today for professional purposes, five (including Sony's DVCAM) use DV compression, and only one, Sony's Betacam SX, uses any form of inter-frame MPEG. The latest HDTV format, Sony HDCAM, actually uses a form of compression more similar to DV than to MPEG.

DV compression was designed for consumer camcorders and is low power and fully symmetrical. Intended for a mass market, its chip set is relatively inexpensive, and functions as both decoder *and* encoder. In addition, DV compression is exclusively intra-frame and was designed to be easily editable. There are simply no inter-frame editing problems. When one edits, every DV frame stands on its own, unlike MPEG where any stream splice requires "reasserting" the PES header and other complex housekeeping.

Lastly, the DV compressed data structure consists of 77-byte packets of information, a figure precisely matching the payload of the IEEE 1394 "Firewire." DV compression and 1394 interconnection were made for each other. This simplified access to the compressed data stream means DV-based

⁵ Television Broadcast, January, 1997. Page 36

recorders could be used as "bit buckets" (data recorders), and the design of third-party editing systems is simplified to allow transparent interconnection and transfer.

DVCPRO: DV COMPRESSION ON PROFESSIONAL MEDIA

As good as consumer DV is, video professionals rightly expect something more robust. Bringing the consumer DV format compression algorithm, chips and small 1/4 inch tapes to broadcast television DVTR functionality required modifications to the format. The most significant modification was increasing the DV 10 mm track pitch to improve editing. Panasonic DVCPRO increased the track pitch to 18 mm for absolutely reliable editing while retaining compatibility. Sony also recognized this in DVCAM and increased their track pitch to 15 mm.. Incidentally, all DVCPRO editing decks can play back ordinary consumer DV, Sony DVCAM, and DVCPRO tapes. Even with the larger track pitch, DVCPRO's ultra compact cassette still allows ENG camcorders to capture more than an hour on a single field tape; the larger cassette (still very small in comparison to a VHS or Betacam cassette) captures over two hours for studio VTR and general purpose camcorders.

Another significant change was from DV's metal-evaporated tape to metal-particle. Metal- particle tape – the type used in such formats as D-3, D-5, and Digital Betacam -- is not only more reliable than metal-evaporated tape, but also allowed Panasonic to give DVCPRO two longitudinal tracks. One provides an additional audio track for audio during high-speed transport modes (digital audio is available during jogging and variable-speed modes) and the other provides a control track for near-instantaneous color-frame lock-up. While DVCPRO is a component video format, it may well be used in a composite video environment, making color framing a significant issue.

4:1:1 VS. 4:2:2 SIGNAL STRUCTURE

Certainly, many think DVCPRO is a great format - in about a year, over 10,000 DVCPRO units have been sold. However, as was noted at the beginning, no one single format -- tape-, disk-, or chip-based -- can be ideal for all applications. The original DVCPRO format (as well as DV and DVCAM) is a 4:1:1 format. In other words, it offers 33% more overall detail than component analog, not counting the freedom from degradation that digital recording brings, and no less color detail.⁶ That means better-than-first-generation-Betacam-SP quality - certainly adequate for most broadcast applications, especially ENG.

Although there are no restrictions on 4:1:1 from a DTV or production standpoint, each user must answer the question of whether 4:1:1 is *good enough* based on a cost/benefit analysis. While DVCPRO is very good and very inexpensive, is 4:2:2 better than 4:1:1?

There is a clear answer: All things being equal, 4:2:2 is better than 4:1:1 at the same compression ratio. If the 4:2:2 compression ratio is greater than that of the 4:1:1 system (as is the case when comparing Betacam SX to DVCPRO), the 4:1:1 system may well be equal or superior.

DVCPRO 50: VIRTUALLY LOSSLESS PERFORMANCE

The superiority of 4:2:2 to 4:1:1 at the same or better compression ratio is the reason DVCPRO 50 is being introduced. DVCPRO 50 offers full 4:2:2 detail with an even *lower* compression ratio than DVCPRO (3.3:1 instead of 5:1) by doubling the recorded bit rate to 50 Mb/s, while still using the same mass-market compression chips, and still playing DVCPRO. This is possible because the DV format was

⁶ In many professional applications, the industry camcorder standard has been Betacam SP and MII. Both are *analog* component formats, so they are not directly comparable to DVTR's, but if Rec.601-style numbers had to be assigned to them, however, they would be about 3:1:1.

designed with a 50 Mb/s mode for HDTV, and the chip set can allow two compressor chips to run in parallel, at 2:1:1 each, to create a 4:2:2 recording. This compression technique has been named "DV 422."

The most marvelous part, however, is not the additional color detail (*twice* that of Betacam SP or M-II) but the effectively lossless compression. Remember those close-to-zero portions of the DCT converted blocks? When they are truncated to zero that compression becomes "lossy"; if they can be preserved, the compression is "lossless". In 25 Mb/s DVCPRO, the compression creates four adjacent luminance DCT blocks and two chrominance blocks in each macroblock. Any excess data must be squeezed into already filled blocks or be truncated. The results are usually very good, but they aren't perfect.

The macroblocks in the DV 422 dual 2:1:1 compression streams of DVCPRO 50 actually consist of a luminance block, a dummy block, another luminance block, and another dummy block (plus the same two chrominance blocks). This time, any excess data from a luminance block has a whole, empty dummy block for overflow. For almost any video signal short of random noise, the compression results *are virtually lossless*. While DVCPRO works fine with the standard-definition versions of DTV, DVCPRO 50 using DV 422 will work even better. This means DVCPRO 50 works better than the only MPEG-based DVTR (e.g. Betacam SX) for feeding MPEG transmission and disk systems!

The same parallel processing that allows DVCPRO 50 to offer low-compression-ratio 4:2:2 will also allow it to offer a 525 progressive-scanning in the studio DVTR. The DVCPRO family will continue to grow -- *without* obsoleting the initial investment.

DVCPRO 525 PROGRESSIVE

The development of the 50 Mb/s transport and processing for high-end production makes for more expensive equipment, but with the additional processing comes flexibility. The new AJ-D950 VTR will accommodate both 525 and 625 interlace signals, and 525 progressive as well. Just as the external Time Base Correctors of yesteryear first became integrated into component analog VTRs, and then became an inherent part of the processing of digital VTRs, we can foresee the integration of standards conversion technology into digital VTRs a year from now. This means users will be able to play back current tapes into future systems. Of course, camcorders will still be needed to be developed for the new signal standards, and the AJ-D900 is a prime example.

DVCPRO, DISKS, AND CONNECTIVITY

At the beginning it was noted that, in an ideal world, acquisition would take place on tape, material would magically appear on disks for editing, and, when the editing was done, the material would magically return to tape for archiving. In a way, DVCPRO has almost achieved that ideal situation. The "magic" is the four-times-normal speed VTR that can feed a 20 minute field tape to a non-linear editing system in just 5 minutes. This rapid process can be contained within the NLE system itself, or it can be an external VTR in a traditional system, or it can be part of an archive.

However, one of the most marvelous developments in DVCPRO has been a series of demonstrations conducted for standardization committees and being shown at NAB. It is actually possible to switch seamlessly between 25 Mb/s DVCPRO compressed video signals and 50 Mb/s DVCPRO 50 (or JVC Digital-S) compressed video signals. That means that if one recorded compressed digital video on a disk-based server, some segments could be captured at 25 Mb/s, and some at 50 Mb/s.

In the demonstrations, switching between 25 Mb/s and 50 Mb/s compressed signals was done using CSDI over industry-standard (SMPTE 259) serial digital devices and interfaces. Thus, IEEE 1394 "Firewire" and SMPTE 259 SDI are two ways DVCPRO DVTRs can be interconnected, and Fibre Channel and Toshiba's DVCS can not be far behind.

It has been rumored that DVCPRO cannot make use of DS-3 (45 Mb/s) high-speed digital carrier connections. That is simply false. There is no reason DVCPRO cannot be immediately connected to the audio and video connections of telecommunications-standard DS-3 video codecs. Direct transmission of compressed DVCPRO signals would require a formatter to adapt the data stream to the requirements of the DS-3 standard, as would be true of any other sub-45 Mb/s compressed DVTR format, including Betacam SX. The DVC's modem being developed by Toshiba also addresses this issue for satellite transmission.

THE BOTTOM LINE: DVCPRO IS GOOD FOR YOUR BOTTOM LINE

The television industry is in the midst of a great transition from analog to digital technology. Equipment purchases are trickier than ever. While even the President of the United States seems to be trying to speed video equipment purchase decisions, whatever is purchased now must continue to work in the digital future.

We at Panasonic believe DVCPRO is the wise choice. It offers digital quality at lower-than- analog prices. It offers size advantages for acquisition and archiving. It offers high-speed connections to disk-based editing systems. It has been submitted for industry standardization (SMPTE D-7). It utilizes low-cost, mass-market compression engines. It utilizes robust, industry-standard, metal-particle tape, allowing longitudinal tracks for audio and rapid color- frame lock. It offers video-industry-standard features and connections -- with a bridge to the computer industry through IEEE 1394 connections and data formatting. It has been *proven* to be extensible with the introduction of DVCPRO 50, which works better than MPEG for feeding MPEG. It is fully compatible with DS-3 and other forms of video transmission, and will soon be available in multi-standard, progressive-scanning and (encoded) HDTV versions. It is the sure investment.

All names and trademarks used herein are the property of their respective owners.