

DVCPRO: A Comprehensive Format Overview

By H. Uchida, H. Isaka, T. Yoshida, and J. Safar

Currently, broadcast and ENG camcorders and video editors use mainly analog 1/2-in. (12.7-mm) equipment. However, as news acquisition increases in importance and digitization occurs throughout broadcasting, more compact and cost-effective digital systems with easy editing, improved features, and high picture quality are now in demand. To meet this demand, we have developed DVCPRO, a 6mm videocassette format using digital component recording to assure picture quality that surpasses that of analog component VTRs. Two cassette sizes hold 63 and 123 min, respectively, and have one-quarter the volume, one-sixth the weight, and twice the recording time of conventional analog 1/2-in. cassettes. DVCPRO records 4:1:1 component digital video signals on 6.35mm tape at a video data recording rate of 25 Mbits/sec, using 5:1 intraframe compression. The DVCPRO compression scheme, as well as the transport and cassette design, share technology and development work done to implement the consumer digital video (DV) format. While many aspects are similar and gain the benefit of partial backward compatibility (e.g., recorded consumer tapes can play on some professional models), many aspects are different. In addition, the approach to 625 signals closely follows that of 525, which is also unique. This paper will describe the compression scheme — the differences as well as the similarities — together with the technical rationale for the choices.

As a result of an intensive development program, a compact, lightweight, low-cost ENG camcorder and its studio editor, called DVCPRO, was introduced in 1995 at the NAB show in Las Vegas and in Montreux, Switzerland. DVCPRO has gained wide popularity with a field-editing package, in which two compact VCRs, two monitors, and edit control are packaged together. This field editing package (FEP) can change the way news stories are edited and processed, significantly influencing the way television news is produced and gathered.

DVCPRO uses a 21.7 mm diameter cylinder rotating at 9000 rpm to write the ten 18- μ m helical tracks per video frame specified by the 525 format. The cylinder has six heads (two recording, two playback, and two flying erase) to realize insert editing and

confidence playback. Noiseless slow-motion playback is made possible by reading from both record and play heads in conjunction with digital signal processing. Each track has an audio area, video area, and subcode

area, each separated by a protective space to allow independent editing. Within the track exists a sophisticated block/macroblock/superblock structure to enhance robustness and support "trick play."

The format records two 16-bit audio channels at 48-kHz sampling synchronous with video, allowing DVCPRO to easily interface with other digital VTRs. The broadcast format provides two linear tracks for cue audio and control track, enabling fast servo system lockup and frame-accurate editing, and a third audio channel giving audio in search when the digital audio is of limited use.

The first broadcast-quality digital VTR, component D-1, was announced in 1987, and in 1989, composite D-2 was introduced. These two revolutionary developments represented higher image quality and better operability than conventional analog VTRs. Furthermore, D-3, using 1/2-in. tape, was announced in 1991, and in 1993, the 1/2-in. D-5 and digital Betacam completed the digital development cycle, with new VCRs substantially smaller than conventional analog tape

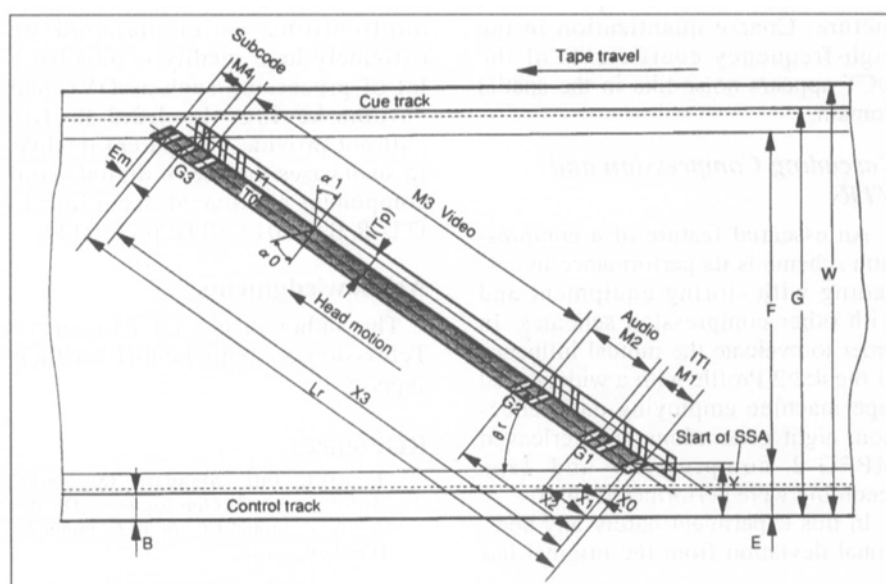


Figure 1. Recorded pattern on tape.

Presented at the 137th SMPTE Technical Conference and World Media Expo in Los Angeles (paper no. 137-30) on September 7, 1995. H. Uchida, H. Isaka, and T. Yoshida are with Matsushita Electric Industrial Co., Ltd., Osaka, Japan 571; J. Safar is with Panasonic Broadcast & Television Systems Co., Secaucus, NJ 07094. Copyright © 1996 by the Society of Motion Picture and Television Engineers, Inc.

Table 1 — I/O Specification of Studio Editing VCR

Input/Output Signals		
Video	Analog	Composite signal Component signal
	Digital	4:2:2 component serial signal (SMPTE 259M)
Audio	Analog:	+4 dBm
	Digital:	AES/EBU
	Embedded audio	Serial digital signal
Time code		Longitudinal time code (LTC) (SMPTE 12M)
CUE		Analog CUE track
Video/Audio Performance		
Video SNR		> 54 dB
Bandwidth		Y: 5.75 MHz C: 1.5 MHz
Audio dynamic range (signal-to-noise)		> 85 dB
Frequency response		20 Hz to 20 kHz
Wow and flutter		Immeasurable

Table 2 — Major Specifications of DVCPRO

Signal structure	4:1:1 component digital	8 bit 13.5-MHz sampling
Compression	5:1 DCT-based	Intraframe DVC standard
Error correction	Reed-Solomon Inner code: (85,77)R-S code over GF(28) Outer code: (14,9)R-S code over GF(28)	
Format	10 tracks/frame (525), 12 tracks/frame (625) 18- μ track pitch	
Azimuth	-19.97, +20.03°	
Tape Speed	33.8539/1.001 mm/sec (525) 33.8539 mm/sec (625)	
Tape	6.35 mm (1/4-in.) metal particle	
Cylinder	21.7 mm diameter	9000 rpm
Video rate	24.948 Mbits/sec	
Recording rate	41.85 Mbits/sec	
Audio	2-channel 16-bit PCM 1-channel analog longitudinal (cue)	
Audio sampling	48 kHz	
Time code	2 independent LTC/VITC	

recorders. This evolution has led to further use of digital technology in broadcasting, ENG, and EFP.

The high mobility in news environment demands development of digital equipment that is more mobile, compact, and lighter than the present portable 1/2-in. VTRs. Video journalists who single-handedly do press coverage, recording, and even tape editing require a truly compact camera/recorder unit with low operating costs and simple operation.

The latest development in recording technology is the ongoing discussion of a unified specification for a home-use digital VCR (DVC). Agreement on such a specification by various manufacturers for home recording systems was reached at the HD Digital VCR Conference by more than 50 companies from around the world; it was finally approved in September 1993. According to this specification, the new digital format employs a compact 1/4-in. cassette with a very small tape transport and high-quality image compression suitable even for use in broadcasting for field acquisition. This new tape format will have far-reaching impact for other applications, such as multimedia or data recording.

Outline of DVCPRO Format

DVCPRO, a broadcast and professional application of new technology, is obtained by extending the full potential of the home-use digital VCR technology. The major modifications to achieve the required professional functionality include use of wider track pitch (18 μ m), in order to ensure repeated editing under adverse conditions, selection of more reliable metal particle (MP) tape, provision for linear control track (CTL), and analog cue-track audio. Such additional information as LTC, VITC, etc., is recorded in the subcode and video auxiliary regions.

The extended specification of the DVCPRO recorder system has the following functionality and advantages:

- Broadcast-quality image recording using 8-bit quantization and 13.5-MHz sampling frequency, including component video signal structure with 5:1 data reduction for PAL and NTSC and intraframe compression system with editing capability as unlimited as all current recording systems.

- Lower equipment costs realized by using common major components (compression system and cassettes) between the home-use and the professional-use VCRs.

- Long recording time realized by use of ultracompact cassette (medium cassette: 63 min, large cassette: 123 min).

- Compatible playback of tapes recorded in the home-use DV format.

Tables 1 and 2 show the basic specification of DVCPRO tape and signal format.

Tape Format

Figure 1 shows the recorded pattern on the 1/4-in. tape, and Table 3 shows the detailed specifications of the DVCPRO format.

Each television frame is recorded on 10 tracks in the 525/60 system and 12 tracks in the 625/50 system. The cue track on the upper tape edge and the CTL on the lower tape edge are linear tracks.

Figure 2 shows the control track pattern recorded on the CTL track. Each frame is identified by rising edge waveform following the 60/40 duty cycle pulse, and the color frame is identified by the 40/60 duty cycle pattern. Since the color frames are identified in each of the frames and the detection scheme is common to the 525/60 and 625/50 systems, the color frame detection time is much shorter, resulting in very short synchronization time to a specific color frame. Because the DVCPRO recorder operates in the NTSC domain as well as in a component environment, color frame identification is necessary for correct composite video editing.

Figure 3 shows the allocation of individual sectors on each of the recorded helical tracks. The insert and track information (ITI) sector is first, followed by an audio sector, video sector, and subcode sector, all separated by editing gaps.

Cassette

The new DVCPRO format employs two types of cassettes, large and medium, as shown in Fig. 4. The large (L), 123-min cassette is the same size as the home-use DV standard cassette, and a second, medium (M) cassette holds 63 min. The total data capacity

Table 3 — Definitions of Tape Pattern

		Dimensions	Nominal
			525/60 625/60
A	Control track lower edge	0	←
B	Control track upper edge	0.4	←
E	Program area lower edge	0.56	←
F	Program area width	5.24	←
G	Cue audio track lower edge	6.00	←
H	Cue audio track upper edge	6.35	←
I	Helical track pitch	0.018	←
M3	Video sector length	27.548	27.576
Lr	Helical tack total length	32.842	←
M1	ITI sector length	0.876	0.877
M2	Audio sector length	2.810	2.813
M4	Subcode tack length	0.906	0.877
P1	Control track	67.500	
P2	Cue track	P1 + 2.5 frame	
X0	Location of ITI signal (SSA)	0	←
X1	Location of audio sector	0.809	0.810
X2	Location of video sector	3.790	3.793
X3	Location of subcode sector	31.885	31.917
Y	Helical track reference	0.615	←
W	Tape width	6.350	←
θ	Track angle	9.1784	←
α_0	Azimuth angle (track A)	19.97	←
α_1	Azimuth angle (track B)	20.03	←
Ts	Tape speed	33.8539/1.001	33.8539

for the L cassette is approximately 23 Gbytes of video.

To simplify the loading and transport mechanism, the sections of both cassettes into which tape loading components are inserted, including the optical paths provided for detection of the beginning and the end of tape, are identical in both cassettes.

Tape

The tape employed by DVCPRO is high-reliability MP tape with a width of 6.35 mm (1/4 in.) and a total thickness of about 9 μ m.

Scanner Structure

Figure 5 shows the scanner head arrangement employed by the

DVCPRO recorder. The diameter of a scanner is 21.7 mm, very close to the diameter of a quarter, and its rotating speed is about 9000 rpm (9000/1.001 rpm at 525/60 or 9000 rpm at 625/50). Compared to a conventional 1/2-in. VCR, the diameter of the DVCPRO cylinder is 3.5 times smaller, so not only is the total mechanism smaller, but power consumption and rotating noise are substantially reduced as well. This is of significant importance for an ENG camcorder and the fundamental basis for unique portable equipment such as the field editing package (FEP).

When an insert edit is performed, the recording and reproduction heads are reading the ITI sector positioned at the

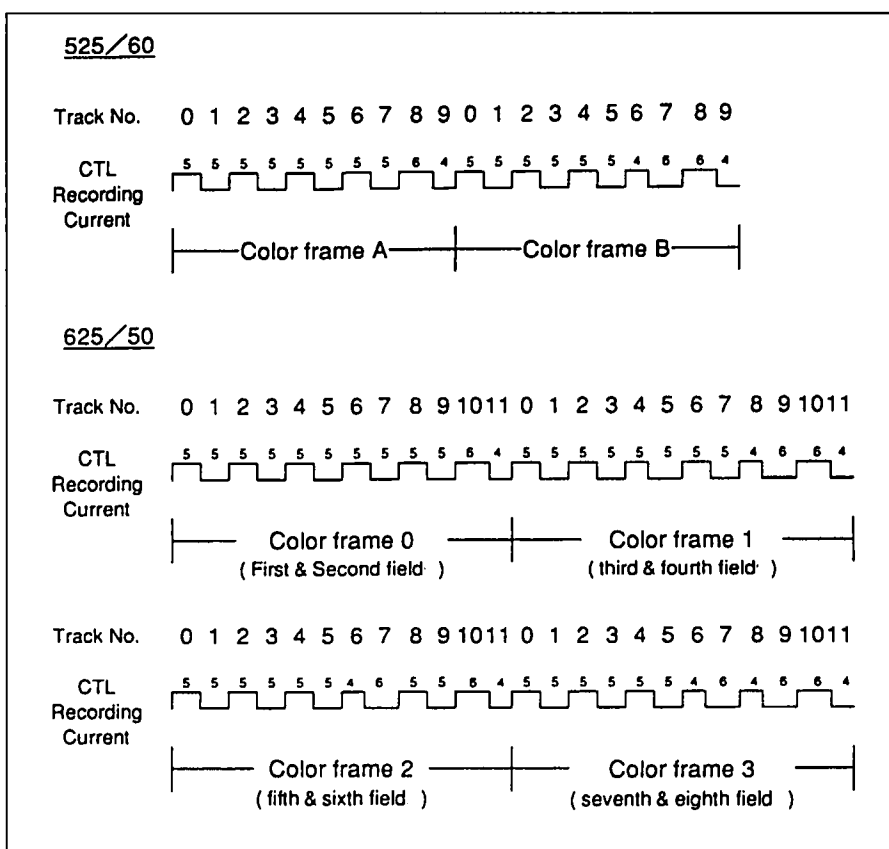


Figure 2. Control pattern.

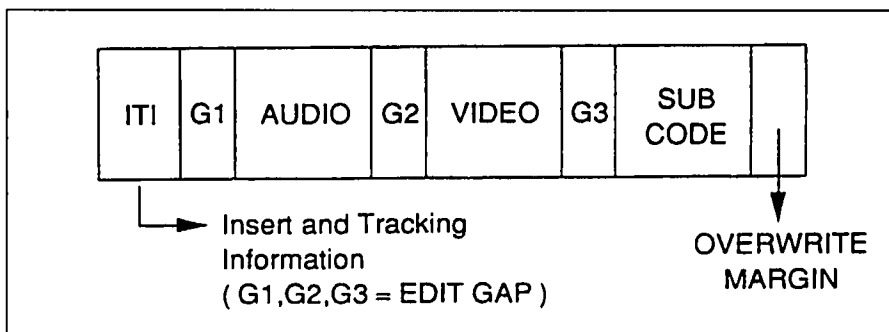


Figure 3. Allocation of sectors on a helical track.

start of a helical track. This is necessary for proper tracking control, and it allows precise determination of the editing gap timing to compensate for minor deviations of the recorded track pattern that was recorded on another tape deck.

Figure 6 shows the relative position between individual heads when writing or reading to a tape. The erase head is positioned two tracks ahead of the recording/reproducing heads, thereby preventing a reduction of a track width at points other than the editing point. Moreover, the erase head current is controlled in such a

way that erasure is not possible during the ITI sector scan; therefore, the crosstalk into recording/playback heads is eliminated during the reading of the ITI sector.

Logical Format

Insert and Track Information Sector

The insert and track information (ITI) sector provides optimum tracking control during insert editing and determination of the exact timing of the editing gap. This type of tracking

control is used during the playback of home-use DV recordings that lack the DVCPRO control track.

The ITI sector comprises the ITI preamble, start sync block area (SSA) section, track information area (TIA) section, and the ITI postamble (Fig. 7). In order to detect the precise head position on the track, an identification code is recorded in the SSA section; in the TIA section, an application ID provides additional information about the track construction.

Audio Sector

Figure 8 shows the audio sector structure consisting of an audio preamble, 14 data sync blocks, and an audio postamble.

The two-byte sync identification is modulated into a pattern of 17 bits, and the three-byte identification code consists of a two-byte data ID and a one-byte data ID parity. Information such as the sampling frequency is recorded in an audio auxiliary (AAUX) block while the actual audio information is stored in the audio data block. The inner parity is defined as a codeword of an inner error-correction code, which is a (85,77) Reed-Solomon code in GF(256), and the outer parity is defined as a codeword of an outer error-correction code, a (14,9) Reed-Solomon code in GF(256).

Video Sector

Structure of the video sector consisting of the video preamble, 149 data sync blocks, and the video postamble is shown in Fig. 9. The closed caption data, a VITC replication, and other information are stored in a video auxiliary (VAUX) block in order to implement the functionality desired for DVCPRO. The video signals are stored in the video data block, and the inner parity is same as that of the audio sector. The outer parity is defined as a codeword of an outer error-correction code, which is a (149,138) Reed-Solomon code in GF(256).

Subcode Sector

The subcode sector, shown in Fig. 10, consists of the subcode preamble, 12 data sync blocks, and the subcode postamble. The LTC replication and other information are stored in the

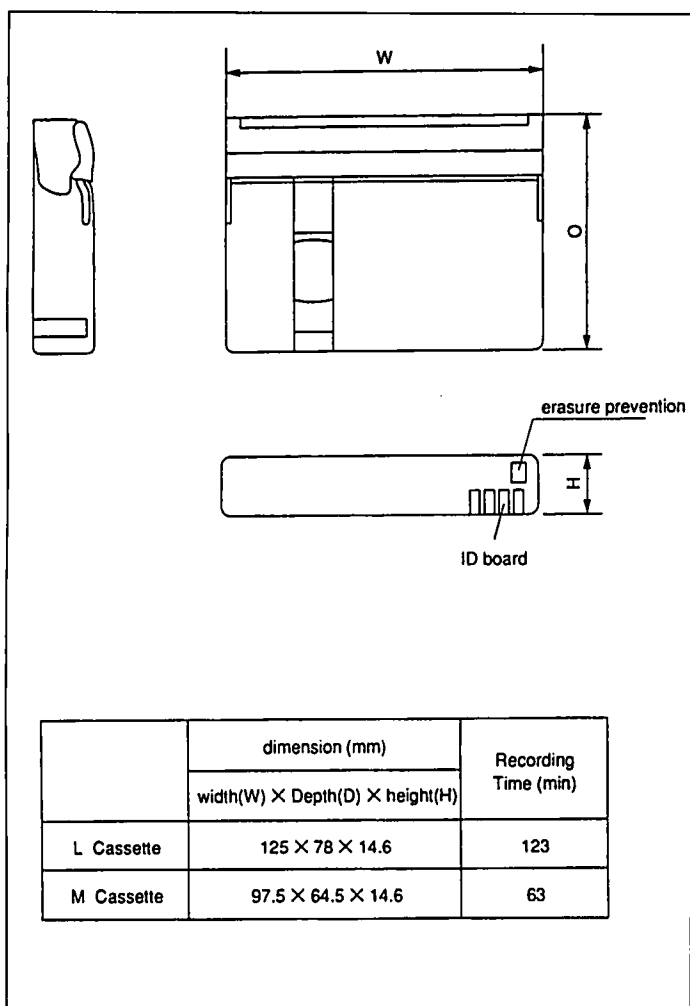


Figure 4. General view of cassettes.

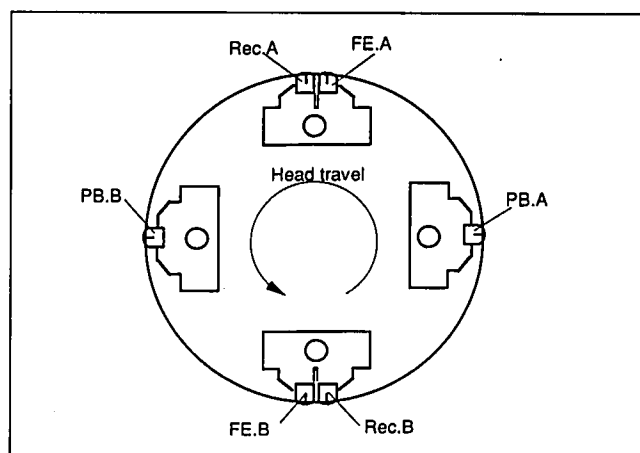


Figure 5. Arrangement of scanner heads.

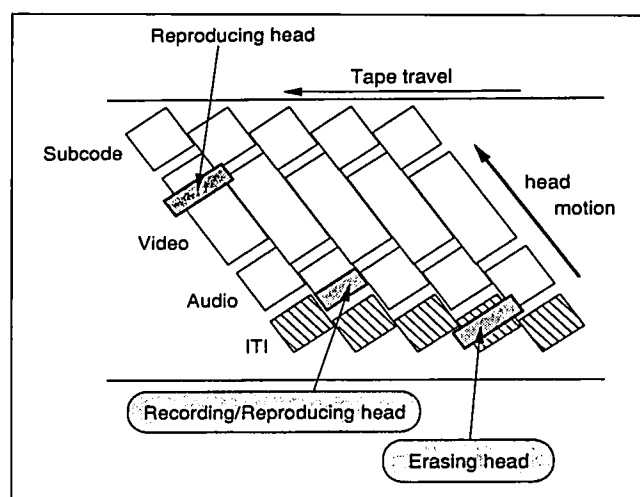


Figure 6. Arrangement of heads on tape.

subcode data block. In order to reproduce the subcode data at high speed, the length of the sync block in the subcode sector is shorter than that of the audio sector or video sector, and the data multiplexed during recording. Subcode parity is defined as a code-word of a subcode error-correction code, which is a (14,10) Reed-Solomon code GF(16).

Audio Signal Processing

Audio Encoding

The audio signal is sampled with 16-bit precision and at 48-kHz sampling frequency (locked to video), as described in the AES/EBU standard. The audio block recording consists of continuous recording of five tracks in the 525/60 system or continuous recording of six tracks in the 625/50 system. Two independent audio sig-

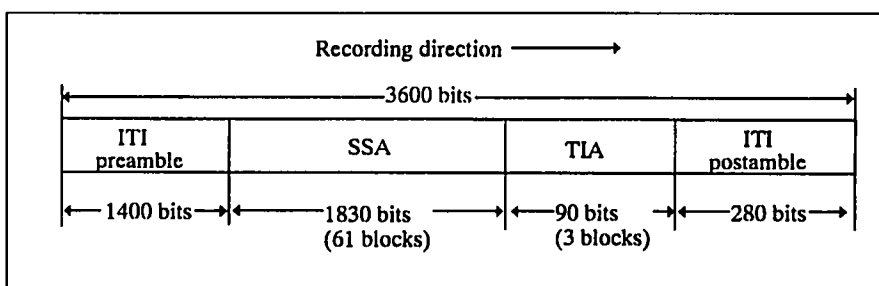


Figure 7. Structure of ITI sector.

nals from one stereo audio channel are recorded on the tape.

Audio Frame

Although audio signals are recorded separately from video, the duration of one audio frame is the same duration as one video frame. In order to prevent a time discrepancy between the audio and video signals caused by the recording and reproducing processes,

the audio frame begins with the audio sample acquired within the period from -50 samples to sample zero from the first pre-equalizing pulse of the vertical blanking period of the input video signal.

Since the audio sampling frequency is synchronized with the video signal, the number of audio samples per frame is predetermined and repeats at a fixed frequency.

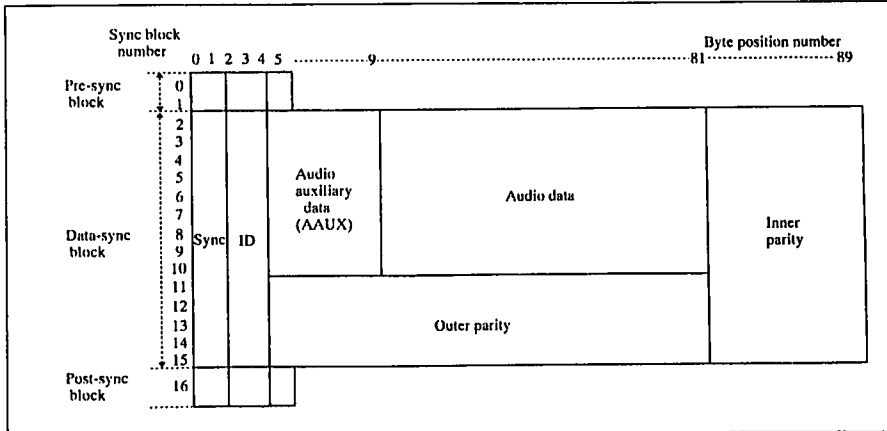


Figure 8. Structure of sync blocks in audio sector.

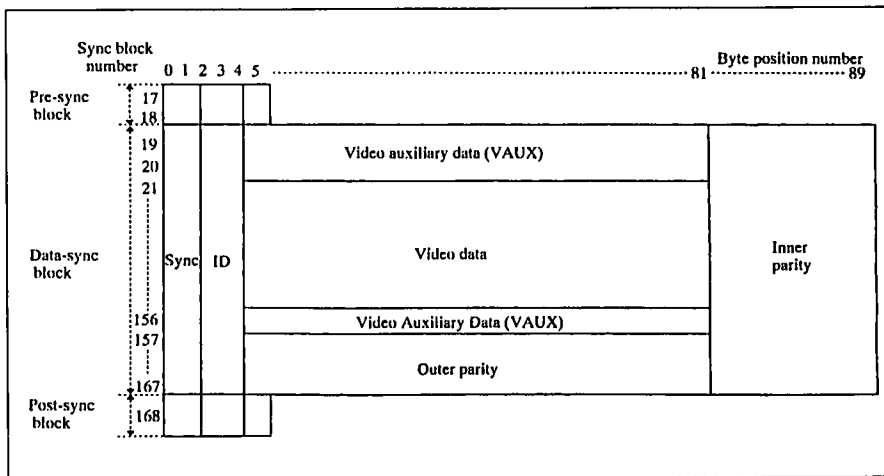


Figure 9. Structure of sync blocks in video sector.

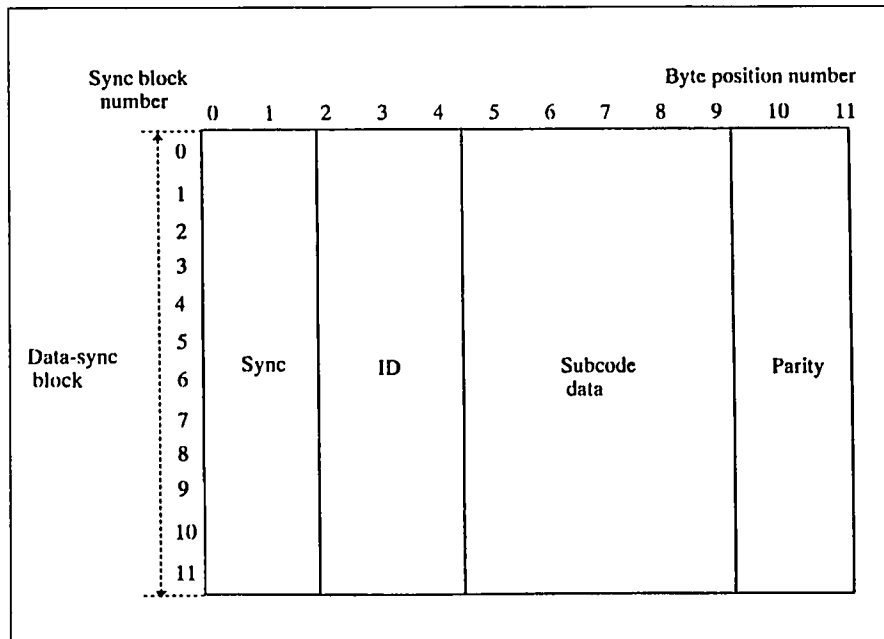


Figure 10. Structure of sync blocks in subcode sector.

Audio Shuffling

The audio signals are shuffled for each of the audio blocks, taking the value shown in Fig. 11 as an initial value. The audio samples are separated into three series ($3n$, $3n+1$, $3n+2$) across the tape width; they are separated longitudinally into two tracks as well. In this way, samples are dispersed on the tape with a predetermined separation distance in order to improve the correction capability.

Video Signal Processing Sampling Structure

The video sampling structure is the same as the sampling structure described in ITU-R Rec. 601 for 4:2:2 component television signals. Each color difference signal is filtered to reduce the number of color difference pixels by half. In DVCPRO, the 525/60 system and 625/50 system therefore employ the 4:1:1 processing system. The 720 luminance pixels and 180 pixels of each color difference signal per line are processed as shown in Fig. 12. This sampling structure has proven to be more than adequate for such field acquisition as news, and it avoids smearing of the color picture elements.

DCT Block

The luminance and color difference pixels are grouped into DCT blocks for the purpose of compression, and these DCT blocks are further organized into macroblocks and superblocks.

The DCT blocks are constructed from 8 horizontal pixels and 8 vertical pixels, with the exception of color difference signals positioned at the right end of the screen. At that part of the screen, the color difference DCT blocks are constructed from 4 horizontal pixels and 16 vertical pixels.

Macroblock

Each macroblock consists of four luminance DCT blocks and two DCT blocks of color difference signal that existed within one DCT block for each of color difference signal. Figure 13 shows a macroblock and DCT blocks organization.

Superblock

Each superblock consists of 27 macroblocks. Overall, pixels in one

television frame are divided into 50 superblocks, 10 vertical rows, and 5 vertical columns. The superblock arrangement for a 525/60 system frame is shown in Fig. 14a; Fig. 14b shows the arrangement for the 625/50 system. In this case, the pixels are divided into 60 superblocks per frame.

Video Segment Shuffling

A video segment recorded on the tape consists of five macroblocks gathered from different areas of the image: S0,0; S6,1; S2,2; S8,3; and S4,4. Doing this allows the shuffling to be performed in macroblock units. Moreover, the original data can be compressed into the predetermined value in the five-macroblock video segment to achieve higher compression efficiency.

DCT Processing

DCT processing has two distinctive operational modes called 8-8-DCT and 2-4-8-DCT. These different modes are used to improve the picture quality after bit-rate reduction. If the compressed material is a still picture, the 8-8-DCT mode processing is performed; the 2-4-8-DCT mode is used for moving pictures. During this process, the data in each DCT block are converted into 1 DC component and 63 AC components. Furthermore, an additional weighing process is performed on the DCT block coefficients.

Quantization

Each of the DCT coefficient blocks is classified into one of four classes, as shown in Table 4. Several different parameters determine the selection of the quantization step. AC coefficients within the DCT block are classified

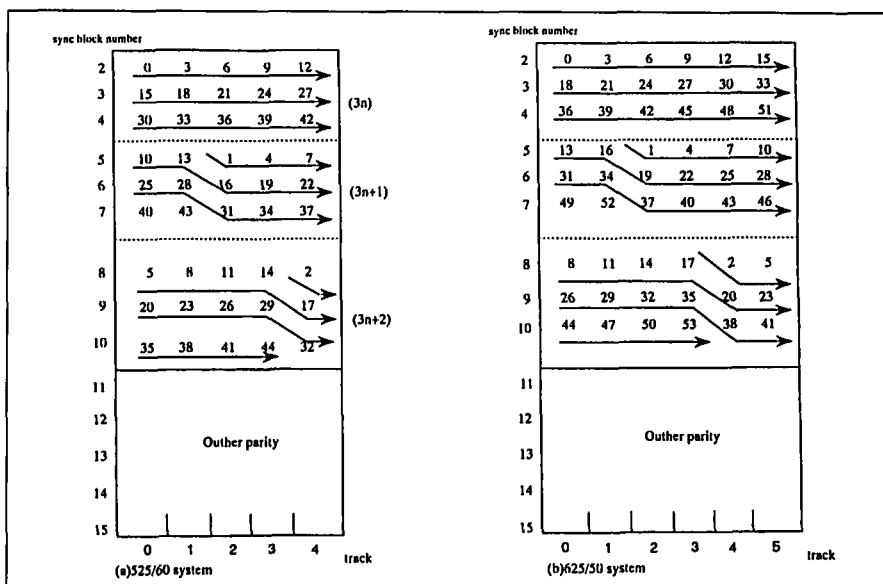


Figure 11. Audio shuffling method.

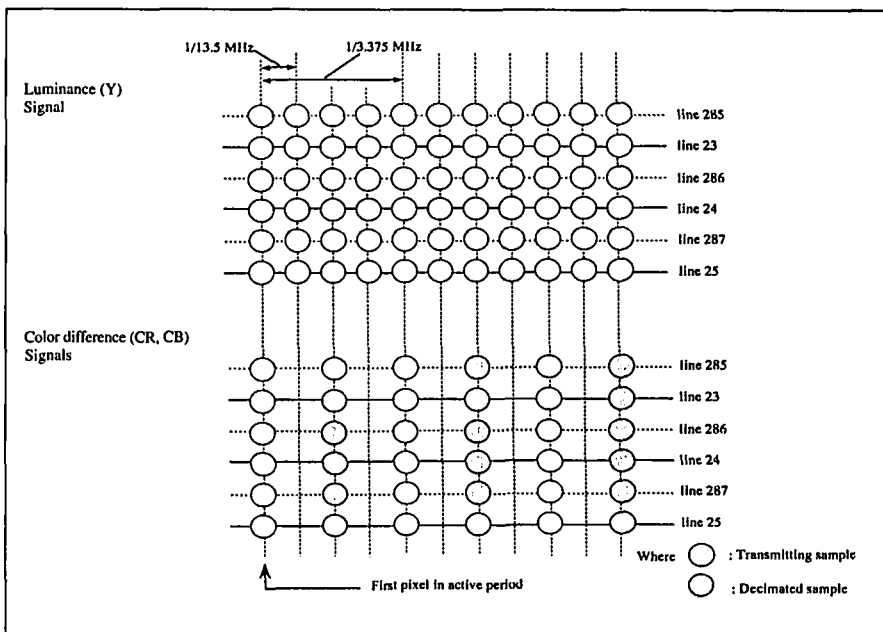


Figure 12. Sampling structure.

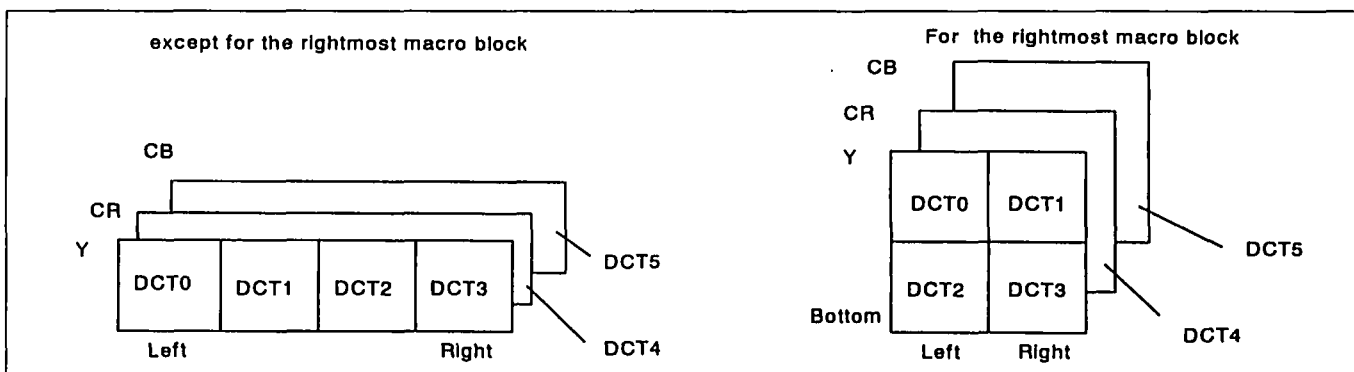


Figure 13. Macroblock and DCT blocks.

TECHNICAL PAPER

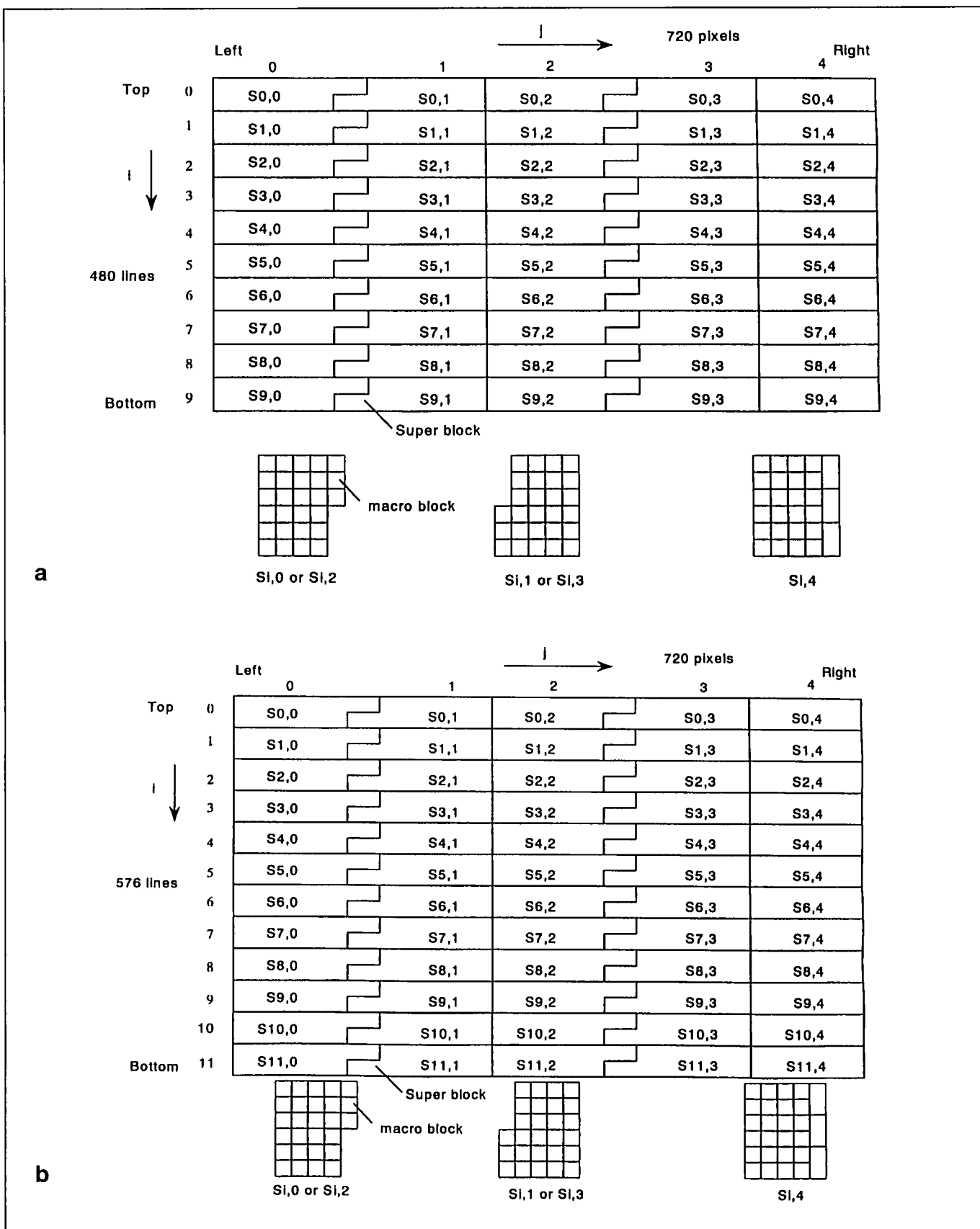


Figure 14. Superblocks and macroblocks in a frame on TV screen: (a) for 525/60 system; (b) for 625/50 system.

into four areas and each area is defined as well (Fig. 15). The quantization step will be decided by the class number, area number, and quantization number in order to limit the amount of data to five compressed macroblocks in each video segment.

The compression system selects quantization for each DCT block from 16 quantization tables. The quantization table is selected such that the amount of data after the quantization and the subsequent variable-length coding that occurs later in the processing chain is closest to a predetermined value. This selection of quantization

tables forms a feed forward compression system.

Variable-Length Coding

Data upon which quantization has been performed is further reduced in volume by applying a variable-length coding process. The variable-length coding employs two-dimensional Huffman coding, so that one or more successive AC coefficients within a DCT block are coded into single-length code. During this process, information contained in the AC coefficients of the DCT block is coded into a sequence of two numbers; the first

number indicates a number of successive AC coefficients equal to zero and the second number, the absolute value of the following AC coefficient. Following this step, the resulting string of dual numbers is further converted to achieve a constant bit rate at the output of the codec.

Formatting

The actual data obtained after variable-length coding and the parameters necessary for decoding (the data class definition, the number of the selected quantization table, etc.) are formatted in the video data area of a video sector.

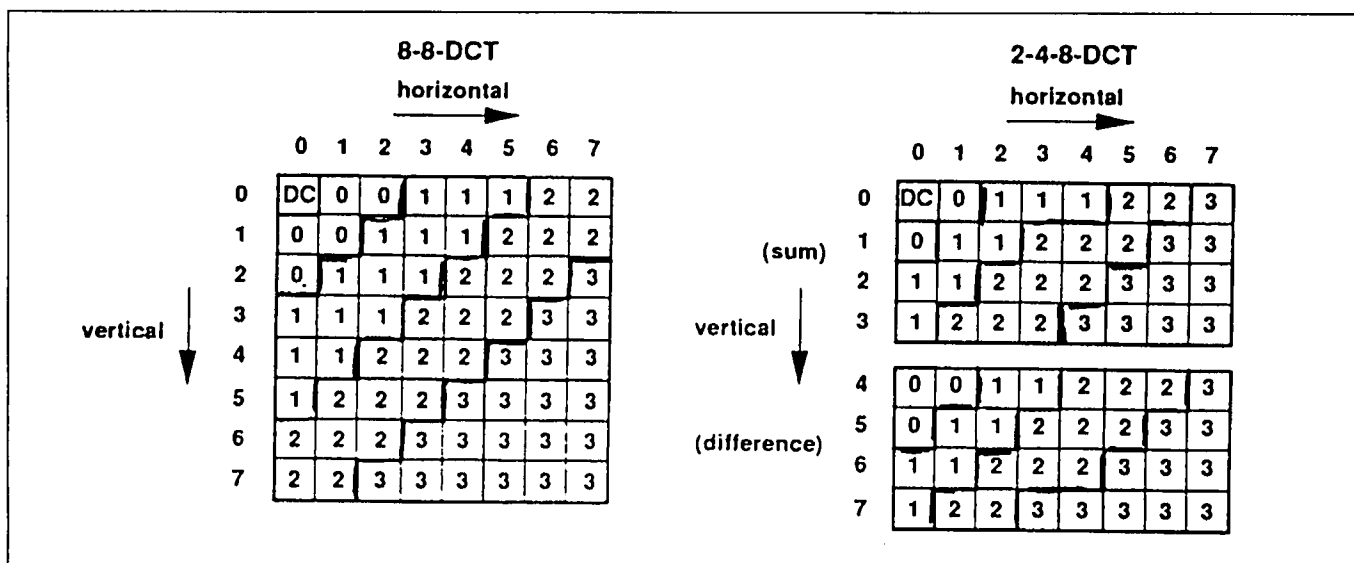


Figure 15. AC coefficient areas.

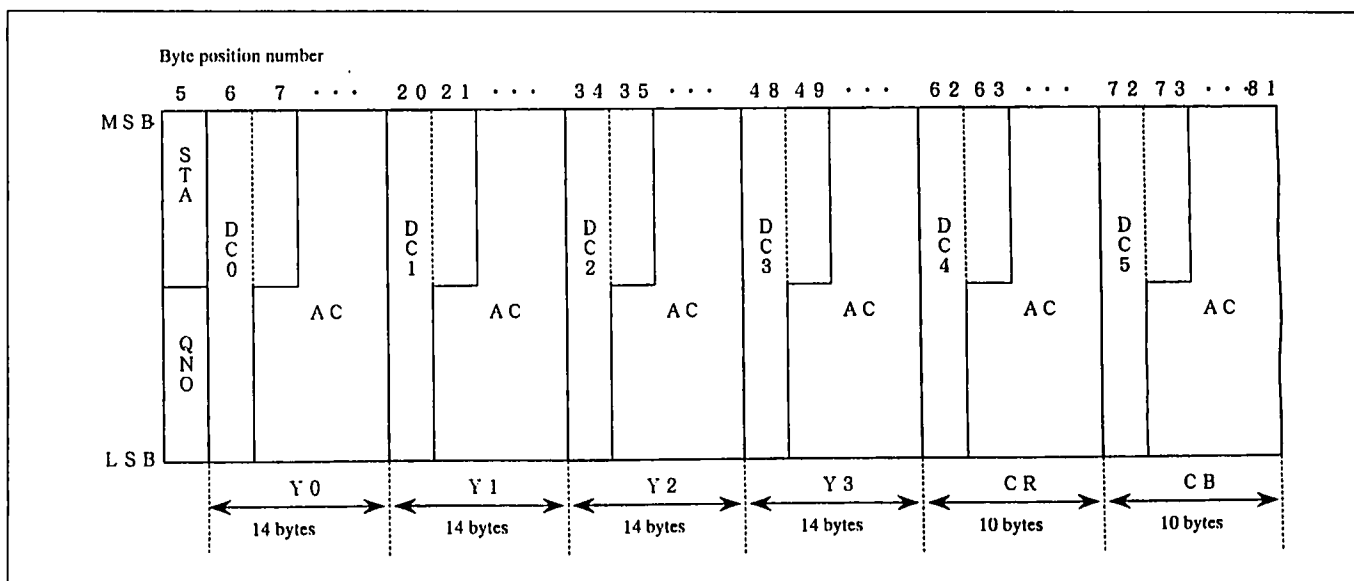


Figure 16. The arrangement of a compressed macroblock.

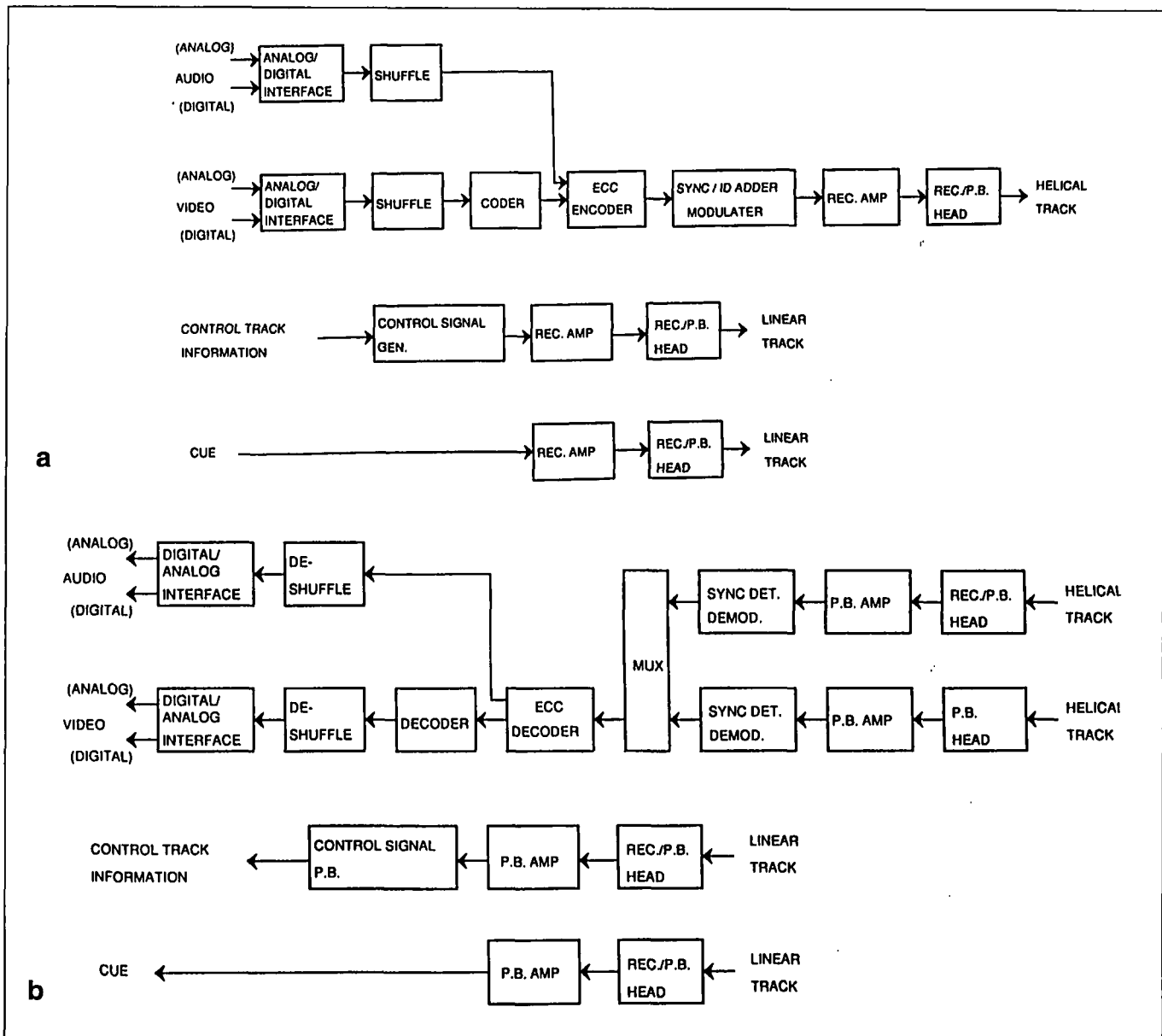


Figure 19. (a) Record block diagram of VCR; (b) playback block diagram of VCR.

tion blocks will be changed. The blocks are reshuffled so that the macroblocks corresponding to the synchronization blocks are arranged in such a way that information recovered during high-speed operations such as fast forward or rewind will provide sufficient information to make a viewable (recognizable) picture.

Modulation

The randomizing 24-25 modulation and interleaved NRZI processes are performed on the data stream to be recorded. Figure 17 shows a 24-25 modulation. The "0" or "1" in the

"extra-bit" position is selected such that the maximum run length of "0" or "1" is less than 10. Additionally, a second criteria helps to obtain a predetermined frequency spectrum from among three spectrum types shown in Fig. 18.

The three types of frequency spectra are specified and shown in Fig. 18 as F0, F1, and F2. The recording is performed in a repetitive order: F0, F1, F0, F2. The ITI sector with recorded pilot signal of high signal-to-noise ratio (SNR) can maintain precise tracking at the insert point. At playback, tracking can be performed by

monitoring the pilot signal component introduced from adjacent tracks.

Record and Playback Apparatus

Panasonic has developed a studio editing VTR conforming to the DVCPRO format. The block diagrams of the recording and reproduction processes are shown in Figs. 19a and 19b. As was shown in Fig. 6, confidence playback can be performed by positioning the reproduction head at a position delayed by a 1/2-track relative to the recording/reproduction head.

For variable-speed playback, the

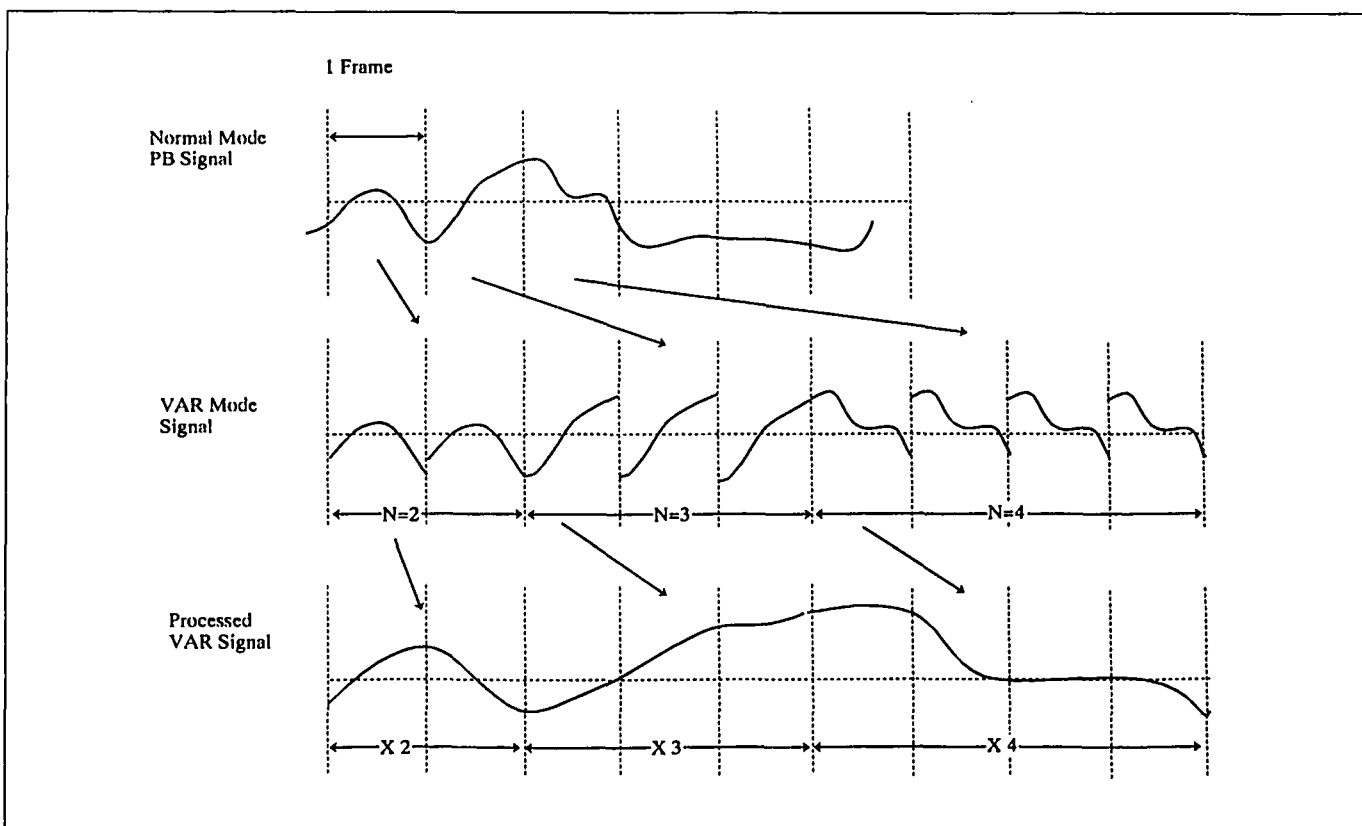


Figure 20. Slow playback waveform.

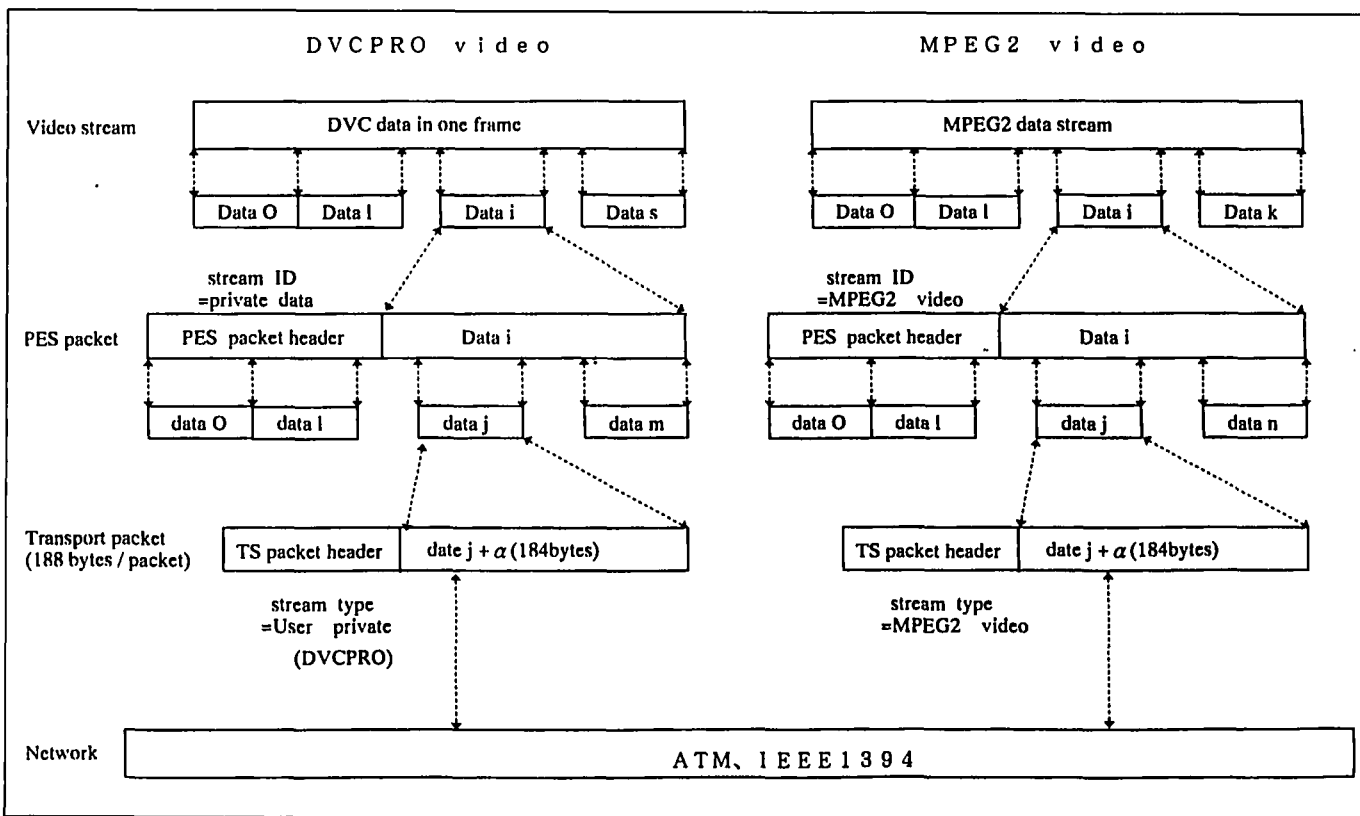


Figure 21. Transmission of DVCPRO video over MPEG system.

data acquisition at high and slow-speed playback is improved by performing playback using not only the playback head, but the recording/reproduction head as well.

The DVCPRO format was developed so that slow reproduction of audio signals recorded on the helical track is possible. Figure 20 shows the slow playback of an audio signal. A waveform repeated at a frame unit frequency is obtained as compared to the normally reproduced waveform. Since the repeated waveform contains all of the information of the original, applying an extension or interpolation process along the time axis, according to reproduction speed, yields reproducible audio signals in the JOG/VAR mode.

DVCPRO Bit Streams

The DVCPRO bit streams (both the video and audio bit streams) can be carried in the MPEG-2 transport streams in a very similar way to other MPEG-1 or MPEG-2 video and audio bit streams. Figure 21 shows a method for carrying the DVCPRO bit streams in the MPEG-2 transport stream.

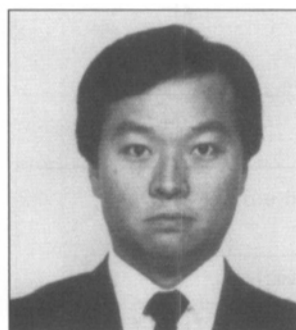
Conclusion

As described above, DVCPRO is a tape format utilizing the basic technology of the home-use DV tape format. Compared to a conventional 1/2-in. VCR, DVCPRO data storage volumetric density is increased by more than a factor of four, and its recording capability is doubled at the same time.

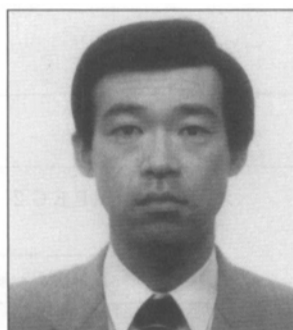
Use of the new technology leads to the realization of a DVCPRO camcorder with excellent mobility — a camcorder with body weight of about 6 lb and a total operational weight of about 11 lb — and total recording time of over 1 hr.

Since a digital VTR employing DVCPRO technology can be connected to various other digital equipment such as a video server, nonlinear equipment, digital optical disc recorder, and quadruple-speed transfer equipment, a practical digital broadcasting environment can be created to cope with the coming multimedia age. Moreover, it is possible to carry the DVCPRO bit streams in the MPEG-2 systems transport stream.

THE AUTHORS



Hirofumi Uchida



Haruo Isaka



Takayasu Yoshida



Johann Safar

Hirofumi Uchida received his B.S. and M.S. degrees in electric engineering from Kyoto University in 1977 and 1979, respectively. He joined Matsushita Electric Ind. Co., Ltd., in 1988, and since then he has been engaged in the development of digital recording technology and signal processing. He has also participated in development of high-definition videocassette recorders. He is a member of the Institute of Television Engineers of Japan.

Haruo Isaka received his B.S. and M.S. degrees in electric engineering from Nagoya University in 1979 and 1981, respectively. He joined Matsushita Electric Ind. Co., Ltd., in 1981, and since then he has been engaged in the development of servo

systems for audio-visual equipment. He is a member of the Institute of Television Engineers of Japan.

Takayasu Yoshida received his B.S. and M.S. degrees in electric engineering from Nagoya Institute of Technology in 1980 and 1982, respectively. He joined Matsushita Electric Ind. Co., Ltd., in 1982, and since then he has been engaged in the development of VCR for professional use. Yoshida specializes in many different aspects of videotape recording technology including analog and digital signal processing.

Johann Safar is with Panasonic Broadcast and Television Systems Co. as manager, Strategic Product

Development Team.

Safar has been working in the broadcast industry since 1972, when he started at Philips Broadcast Equipment Corp. as a video camera design engineer. Seven years later, he joined Thomson CMF Broadcast, where he worked on the development of digital signal processors, color correctors, and other processing equipment. He later became director of engineering and remained in that position until he joined Panasonic in 1988.

Safar is a member of the SMPTE, the AES, and the IEEE. He is currently a participating member of SMPTE Technology Committees V16, S17, P18, and PT20, and chairs SMPTE Working Group S17.10 on Ancillary Data.