The HDTV Camcorder and the March to Marketplace Reality

By Laurence J. Thorpe, Fumio Nagumo, and Kazuo Ike

A complex 12-year development of high-definition television (HDTV) production standards has vied with consecutive generations of HDTV program origination equipment. While aspiring to the full promise of the ever-evolving standards, equipment designers have grappled with a multiplicity of marketplace realities as separately defined by the world's broadcasters and program producers. HDTV program origination continues to be frustrated by unrealistic equipment costs and a lack of mobile battery-powered HDTV acquisition systems. The latter is central to the proper exploitation of the significant enhancement to television imagery latent within the technical parameters defined by the latest SMPTE 274M HDTV production standard. This paper describes a bold development thrust that sought to simultaneously lower costs and produce a high-performance, highly compact, HD acquisition system. The technological leap from contemporary HDTV cameras and digital recorders to an all-digital one-piece HD camcorder called for the best in engineering innovation, pragmatic design decisions, and a massive commitment to custom-built VLSI. This paper specifically describes the prototype of the 1920 x 1035 (SMPTE 260M) based camcorder. A modified product based on the later SMPTE 274M standard (1920 x 1080 format) will follow.

An extensive 12-year global high-definition television (HDTV) production experience, with no less than three successive generations of Sony's HDVS program origination equipment, has formed the foundation for a unique design approach to a next-generation HDTV production system. Our broad experience has been largely shaped by two somewhat conflicting forces:

- The actual HD shooting experiences of international producers, broadcasters, and film makers over the past decade.
- The continually evolving international HDTV standards, both for program origination (the digital production standard) and for HD broadcasting (the digital transmission standard).

In the U.S., the shaping of the final HDTV standards has seen the emergence of a unique interindustry consensus on the square pixel 1920 x 1080 structure as the highest-resolution format (there is another at 1280 x 720) underlying both production and transmission. The new SMPTE 274M production standard supports this format, and is flanked by the SMPTE 292M HD serial digital interface standard. Indeed, the decade-long SMPTE development of HDTV production standards has presented a continually moving target in terms of spatial resolution (from the original 20/10/10 MHz component set, upon which our original HDVS system design was based) to the present 30/15/15 MHz video component set, defined in 1988 by the SMPTE 240M production standard, to the full digital 1920/1080 x 1080 spatial set of the latest SMPTE 274M production standard.

The Advanced Television System Committee (ATSC) DTV standard for digital broadcasting also supports transmission of this same high resolution format (among others of lesser resolution). These standards, in concert with those separately developed in Japan by the BTA/ARIB, (who did maintain a close working liaison with SMPTE) were all to play a significant role in the development of the new generation HDVS equipment. They did, however, substantially "raise the stakes" in terms of the volume of real-time digital data that must be contend with when originating HD program material.

Marketplace Imperatives

Meanwhile, throughout this protracted standards development era, there has arisen an unrelated but quite universal plea from the many pioneering HD producers, one that more or less separates two primary requirements underlying a next generation HDVS system:

- Portability—to facilitate mobile, battery-operated, HD acquisition on remote location
- Lower cost—to bring HD equipment pricing closer to marketplace realities.

The cost issue, in particular, is critical. The lessons of the past 12 years were palpably clear—the success of HD is squarely predicated upon it being affordable to the majority of program makers and broadcasters. The cost of current generation HDVS cameras and recorders hovers in the neighborhood of $750,000 for a single shoot-and-capture system—and the marketplace quite clearly spoke with its pocketbooks on the viability of such pricing.

The Physical Challenges

More than a decade of HDTV shooting experiences spanning the globe have vividly exposed the sheer logistical impediments to any form of practical HD field acquisition. Many of the world's broadcasters, independent producers, and cinematographers can relate graphic tales of triumphs and misadventures in their early HDTV experiences.

Contemporary HDTV cameras and recorders, in terms of size, weight,
power consumption, are reflective of those equipments supporting standard 525/625 field shooting back in the early 1970s. Sony has to date brought no less than three generations of HDVS production equipment to an embryonic HD marketplace. Even today the camera recorder combination remains woefully unwieldy. Table 1 highlights the current physical state of-the-art HDVS production equipment to illustrate the magnitude of the present challenge to an HD producer attempting a remote HD location shoot.

Clearly, while the present HD camera physically approaches that of contemporary high-end standard definition television (SDTV) portable cameras, the HD recorders are entirely another matter. Handheld mobility is simply not an option. The digital recorder, while technologically formidable in its 1200 Mbit/sec realtime baseband recording capability, can only be operated within a mobile van equipped with AC power, and even then, it necessitates the use of a sizeable CCU with the camera (connected by fiber-optic or multicore cable link).

The arrival of the first cassette-based HD videotape recorder (VTR), the analog UNIH recording format, did at least make HD recorders somewhat transportable and capable of operation from larger car batteries. It needs to be pointed out, however, that significant design compromises were entailed to achieve the reductions in size, weight, and power shown in Table 1.

The UNIH analog recording format incorporates necessary expediencies to bring the total recorded bandwidth within the realm of contemporary, reliable 1/2-in. analog recording technology. Luminance bandwidth is preferrited to 20 MHz (from the full 30 MHz called for in the SMPTE 240M production standard), the color difference signals are curtailed to 7 MHz each (from the 15 MHz of that same SMPTE standard), and the latter are recorded line-sequentially. To put an approximate digital perspective on these restrictions, if the full SMPTE production standard is symbolized as a superset of the well-known 4:2:2 structure, then the UNIH signal format is symbolized by a 2:7:1:0 structure.

Nevertheless, it must be said that a considerable body of remarkable HDTV program origination from remote land, sea, and air locations has been captured with this, the only presently viable HD field recording option (generally, post-production is done on a full-baseband digital VTR). But a far more radical advance is clearly required to meet the growing frustration of the HD production community.

Basic Design Principles

A Technological Leap Forward

Pressured by the veterans of these early HD remote productions, Sony elected to bypass the more historic approach to a paced evolution of the camera-VTR shoot and capture system (Fig. 1). It was felt that industry confidence in the viability of HDTV could only be restored with the development of an HD field acquisition system contemporary with that of vintage 525/625-line portable SDTV equipment. Accordingly, in 1994 an ambitious design program was launched to develop a compact one-piece HDVS camcorder. A high priority was assigned also to achieving a cost that would be as close as was technologically possible to that of today's Digital Betacam camcorder—the DYM-700.

In every respect—especially in size, weight, power, and cost—the digital recorder section of the proposed camcorder posed daunting obstacles. An early decision was made to deal aggressively with the substantial cost issue by capitalizing on the economies of scale offered by the overt utilization of the Digital Betacam transport. Thus 1/2-in. metal particle tape recording was defined early.
This seminal decision was followed by the attendant decision to utilize the DVW-700 chassis itself—again, to achieve an important economy of scale. This in turn quickly dictated a necessary commitment to the 2/3-in. image format for the 16:9 aspect ratio CCD imager in the camera section, and the necessary utilization of digital signal processing (DSP) camera video processing to meet the imposed cost and size goals. The genesis of an all-digital one-piece HDVS camcorder thus was born.

**Strategic Design Approach**

But some crucial design decisions still confronted the design team. Criteria of reliable digital recording on 1/2-in. tape were by now well established; Digital Betacam had done that. Somewhere in the neighborhood of 120 to 150 Mbits/sec total video payload would ensure the requisite integrity of component video capture, even within the most hostile of remote location environments (including extremes of temperature, humidity, dust, smoke, etc.). But full origination according to the SMPTE 260M/ARIB Standard BTA S-002A digital production standard (1920 x 1085) would entail HD acquisition at an awesome data rate in the neighborhood of a 1000 Mbits/sec of Y, Pb, Pr component video. Clearly, digital data reduction techniques would be mandatory, entailing a bit-rate-reduction ratio of approximately 10:1.

A number of technological options were possible:

1. Data reduction in both camera and recorder.
   - Subsampling in the imaging section itself, as had already been done by other workers in HDTV, would reduce the total data rate delivered by the camera to the VTR.
   - Also subsequent, more modest bit-rate reduction within the digital VTR.
   - Alternatively, data reduction within the recording section only.
     - Full 1920 x 1085 (1080) 3-CCD imaging, followed by full baseband DSP video processing in the camera section.
     - Bit-rate reduction exclusively within the VTR.
   - The latter approach was favored by the design team. It offered important advantages, especially in facilitating a more orderly progression toward higher-performance systems as both imaging and recording technologies improved:
     - Higher MTF in both the horizontal and vertical domains would be recorded in the first-generation capture.
     - Possibility of full SMPTE specification baseband parallel recording (via separate high-definition serial digital interface (HD SDI) output from the camcorder system, according to the recently developed SMPTE 292M standard).

This pivotal design decision put the full onus of captured picture performance on the VTR section. Considerations of HD data rate reduction, of a level necessary to achieve the 120 to 150 Mbits/sec on tape, opened a careful examination of two critical (and somewhat opposing) imperatives:

1. Capture of the highest possible subjective HD picture quality on the first generation tape.
2. Preservation of a very high subjective picture quality over a practical number of generations of recording (five minimum) in post-production.

**Support System**

The development of a camcorder utilizing a new recording signal format requires a minimum support system to allow easy integration into an existing HDTV infrastructure and to facilitate post-production work on the captured HD program material (Fig.2). Recognition of the important coexistence of HDTV and 525-line SDTV program origination infrastructure was also factored into the support system.

**The HDVS Support Studio VTR**

Central to the new HDVS system developed around the fledgling camcorder was the flanking studio VTR for editing and post-production. Once again, Digital Betacam became the center of a design program that kept cost control as a high priority. The Digital Betacam DVW-500 studio editing VTR was utilized to the highest degree possible in terms of the physical chassis, tape transport, and recording heads. The electronics were, of course, unique to the new HDVS recording format.

In terms of input/output (I/O) interface, the same principles of Digital Betacam (in itself a digital recording format utilizing bit-rate reduction) prevailed, namely all external interfaces would be at baseband. The SMPTE 292M/ARIB BTA S-004A high-definition serial digital interface (HD SDI) standards (both are fully compatible), were chosen as the primary digital I/O. Component analog interfaces were included as an optional (HD SDI) to analog Y, Pb, Pr board. The four channels of digital audio can be employed embedded within the SDI format, while separate digital and analog standard interfaces are also provided.

A digital down-converter has been incorporated into this studio VTR to deliver a parallel digital output feed of 525-line SDTV; selectable as the same full 16:9 widescreen image as the original HD, a letterbox 4:3 version of the same, or a derived standard 4:3 center-cropped version with a remote pan/scan control. These SDTV output choices are delivered with a further selectable choice of a serial component or serial composite version of the standard SMPTE 259M SDI digital feed.
THE HDTV CAMCORDER AND THE MARCH TO MARKETPLACE REALITY

Table 2 — Comparison of CCD Imagers

<table>
<thead>
<tr>
<th>CCD Device (16:9 Format)</th>
<th>520K pixel (2/3-in.)</th>
<th>2 million pixel (1-in.)</th>
<th>2 million pixel (2/3-in.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital Betacam</td>
<td>1038 (H) ×</td>
<td>1920 (H) ×</td>
<td>1920 (H) ×</td>
</tr>
<tr>
<td></td>
<td>504 (V) ×</td>
<td>1036 (V) ×</td>
<td>1036 (V) ×</td>
</tr>
<tr>
<td>Unit cell size</td>
<td>9.0 (H) ×</td>
<td>7.3 (H) ×</td>
<td>5.0 (H) ×</td>
</tr>
<tr>
<td></td>
<td>9.5 (V) ×</td>
<td>7.6 (V) ×</td>
<td>5.2 (V) ×</td>
</tr>
</tbody>
</table>

Table 3 — Comparison of 2/3-inch and 1-Inch HD CCD Imagers

<table>
<thead>
<tr>
<th>Performance</th>
<th>2/3-In CCD</th>
<th>1-In CCD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (V/m/Lux)</td>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td>Saturation Voltage (V)</td>
<td>500</td>
<td>500</td>
</tr>
<tr>
<td>Dynamic Range (dB)</td>
<td>70</td>
<td>72</td>
</tr>
<tr>
<td>Vertical Smear (dB)</td>
<td>-126</td>
<td>-100</td>
</tr>
</tbody>
</table>

Camcorder SDI Output

An important system feature first developed for Digital Betacam has been planned as part of this new HD acquisition system. The ability to output a separate HD SDI feed of the full-camera baseband 10-bit component video set can be important in many EFP applications. While recording in the camcorder itself according to the compressed signal format, this separate realtime parallel SDI output from the camera facilitates simultaneous recording on another VTR, which might be the same recording format or, alternatively, could be a baseband digital HD/HDV-1000 recorder or the HD recording format of another manufacturer (Fig. 3).

The same SDI output can be used to feed the camcorder playback video and audio to an offline editing system. Because this feature is one that is used only for special EFP system applications, the basic camcorder itself was not burdened with the required SDI circuitry. Instead, a compact adaptor, which mounts on the rear of the HDVS camcorder and interfaces with the internal electronics via a special multipin interface, has been developed. This same adaptor also carries two additional audio-input interfaces to facilitate full utilization of the four digital audio recording capability of the camcorder.

Portable Standalone HDVS VTR

Just as the marketplace had previously dictated in the case of both Betacam SP and Digital Betacam, complete flexibility in handling all forms of field acquisition often demands the use of a very mobile, standalone, battery-operated VTR. The need to do parallel recording for certain types of program origination, or to provide long-form recording (beyond the capability of the camcorder’s 40 min maximum), or to facilitate the use of another HD camera (such as the many HDC-500 cameras and those of other manufacturers presently in use all over the world) spurred the development of a sister portable VTR to the camcorder. And again, in the interest of controlling manufacturing costs, the popular Digital Betacam DVW-250 portable VTR formed the nucleus of that design program.

DSP Camera Design Details

The decision to utilize the 2/3-in. image format for the new camcorder was primarily determined by the prior choice of the DVW-700 Digital Betacam chassis, but it also served to ensure an attractive compatibility (such as lens interface) with the new family of HDVS studio/OB/EFP cameras recently developed. Finally, the use of the same 2/3-in. HD CCD imager across all of these HDVS acquisition systems gained a further economy of scale in terms of manufacturing costs on this high-performance imager.

CCD Technology

The CCD is a full 1920 x 1035 sensor (2 million pixel) device using the very latest in Hyper HAD technology first introduced in 1992. It uses frame interline transfer (FIT) technology. This CCD structure is in conformance with the SMPTE 240/260M HD production standards and also with the BTA S-001A/S-002A standards used widely in Japan. North America's recent move to a preferred square, pixel-based standard, reflected in the new SMPTE 274 M production standard, has launched a development program to design a commensurate 1920 x 1080 CCD for a camcorder to service its marketplace (this will arrive

Figure 3. Camcorder with dockable HD SDI adaptor provides full-baseband camera video output via SMPTE 292M serial interface.

SMPTE Journal, March 1998
in 1998). This paper, however, describes the camcorder shortly expected to undergo first field trials with Japanese broadcasters. A perspective on this new CCD is illustrated in Table 2.

Three of these high-resolution imagers are used within a classic RGB optical system. The new CCD design succeeded in pushing the frontiers of solid state imaging even further than the predecessor 1-in. format HD imager described at the NAB HDTV World Conference. Two million sensors in a relatively small 2/3-in. 16:9 image format represents a significant challenge in optimizing all aspects of an individual sensor that is now very small indeed, being a mere 5.0 μm x 5.2 μm (only 47% of the pixel area of its 1-in. format predecessor). Each of these minuscule sensors directly determines the opto-electronic sensitivity, the dynamic range, signal to noise, and immunity to artifacts such as the vertical smear associated with severe highlights in the scene being imaged, and contributes significantly to the overall camera colorimetry. Table 3 compares the imaging performance specifications of the new 2/3-in. CCD with the larger format device used in the older HDC-500 HDVS camera (first introduced in 1992).

The smaller format imager inherently incurs a penalty in sensitivity compared to its larger-image format counterpart. This loss was almost fully recovered by incremental improvements in both the sensor quantum efficiency and, separately, in the optical efficiency of the micro lens (Fig. 4). In particular, a further improvement in controlling the various CCD noise sources was achieved by Sony physicists, and this produced an operational sensitivity and signal-to-noise (SN) performance very close to the earlier-developed 1-in. format imager. A significant breakthrough in reducing vertical smear was achieved (by a factor of 10), which greatly enhances the capabilities of the camera in outdoor shooting. Dynamic range was also incrementally improved and comes very close to that of the larger format device.12

In combination, these performance attributes make this camcorder a highly viable new acquisition system for digital electronic cinematography.13,14 The excellent imaging performance that was ultimately achieved was a consequence of painstaking optimization of process control gained on a broad scale with the very small pixel CCD used in Sony's contemporary consumer 1/4-in. imager, having 380K pixels (Fig. 5). In the latter case, the image diameter is more than 2.5 times smaller than the 2/3-in. HD device, with a pixel size of only 4.74 x 5.55 μm compared to the 5.0 x 5.2 μm.

Sampling Lattice

The sought-after highest modulation transfer function (MTF) in the spatial domain called for no subsampling in the basic image sampling lattice, which was retained at full 1920 x 1035 (1080). The small 2/3 in. image format called for highly refined manufacturing techniques in the precision-machining of the three CCDs to the optical block. Horizontal spatial offset of the two R and B imagers relative to the G imager...
was employed to maintain a high-luminance horizontal MTF while simultaneously minimizing aliasing.\textsuperscript{13} The imaging system made use of a very innovative expedient to accomplish this—one determined by the special challenges posed by implementation of the RGB video processing with high-speed DSP (Fig. 6).

**Resolution Performance**

The resolution of the camera is separately determined in the horizontal and vertical domains by the spatial sampling and associated filtering techniques. These techniques are quite different in the two dimensions. Good overall picture sharpness in HDTV imaging is achieved by seeking both a maximum MTF in the horizontal and vertical dimensions and also in achieving a more isotropic resolution closer to the natural capability of the eye-brain combination than is possible with 525/625 SDTV. Referring to Figure 7, in the horizontal dimension the effective luminance resolution is a highly complex convolution of:

- Lens MTF.
- Optical low-pass prefilter (to reduce aliasing).
- CCD aperture (determined by horizontal sampling and opening ratio of the individual pixel).
- Spatial offset (G relative to R/B to reduce Y aliasing).
- Sample and hold circuit on CCD output.
- Output electrical low-pass filter.

Figure 8 illustrates this convolution and shows how the final luminance MTF response is produced. At 800 TVL/ph the depth of modulation is 45%. This tops to 35 to 40% at the band edge defined by the SMPTE 274M production standard, namely 875 TVL/ph (or 30 MHz). Both the optical input filter and the electrical output filter combine to roll off the response as rapidly as possible following the 30-MHz point to reduce horizontal aliasing to an acceptable minimum. The 37.125-MHz low-pass filter defines the horizontal limiting resolution in the neighborhood of 1000 TVL/ph.

Figure 9 illustrates the convolution mechanism producing the vertical MTF response, determined by the following:

- Lens MTF.
- Optical low-pass filter.
- CCD vertical aperture (determined by 1035 samples).
- CCD vertical FIR filter (to attenuate interlace related artifacts).

Figure 10 shows the good match achieved between the horizontal and vertical MTF responses, the combination of which produces a very high spatial sharpness enhanced by better subjective acceptance of the eye-brain of this approximately isotropic overall resolution.

The shortfall in displayed vertical...
resolution due to the display Kell Effect can be easily made up with a modest degree of vertical enhancement as shown in Figure 11. This issue will be re-examined later following the discussion on the hard limiting effects of the associated digital VTR.

The A/D Converter
Three analog-to-digital (A/D) converters, each sampling their respective CCD analog video outputs at 74.25 MHz (called for in the SMPTE/BTA HDTV production standards), are used to digitize three wideband RGB video signals directly out of their CCD imaging subsystems (Fig. 6). They are 10-bit A/D converters implemented by high-speed CMOS circuits operating at 3.3 V and implemented with 0.3-μm semiconductor technology. The prefilter to the A/D converters are in fact the output filtering systems described in the last section for the CCD imaging subsystems.

The DSP Processing System
Figure 12 shows a block schematic of the video processing channel. All the primary linear functions of linear matrixing (for color correction according to SMPTE 240M colorimetry), and black shading and white modulation shading are performed in the digital domain. A fixed analog pre-knee circuit is employed before the A/D conversion to reduce the total dynamic range (the CCD has a linear transfer characteristic that can deliver up to 600% of nominally exposed white level when imaging overexposed scene content) to about 225%. The nonlinear processing of gamma precorrection, aperture correction, detail enhancement (both horizontal and vertical), skin-tone detail, variable knee compression (for overexposed signals), and black and white clipping are also all digitally implemented.

The requisite calculations are done at 74.25 MHz, with the linear functions processed at 10-bit resolution and the nonlinear functions at a maximum word length of 22 bits. The camera processing is implemented in one VLSI microchip, utilizing 0.35-μm technology operating at 2.5 V.

HDVS Camera Performance
The operational performance of the camera section of the HD camcorder is compared in Table 4 with that of the earlier 1-in. format CCD camera HDC-500 and also with its current 2/3-in. SDTV equivalent, the DVW-700WS Digital Betacam one-piece camcorder.

This operational sensitivity facilitates shooting with the same impunity as is possible with current Digital Betacam camcorders. Full luminance signal levels can be achieved with

![Figure 10. The basic HD camera output has almost isotropic H/V resolution, but the viewed resolution incurs a vertical loss due to the display's Kell Effect.](image)

![Figure 11. A modest application of vertical enhancement restores a reasonably isotropic H/V resolution balance on the final HD display.](image)

### Table 4 — Performance Comparison with HDC-500 and DVW-700WS

<table>
<thead>
<tr>
<th>Camera Performance</th>
<th>New HD Camera</th>
<th>HDC-500</th>
<th>DVW-700WS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (2000 Lux @ 86.9% R)</td>
<td>f-8.0</td>
<td>f-8.0</td>
<td>f-8.0</td>
</tr>
<tr>
<td>Signal-to-noise ratio (30 MHz)</td>
<td>54</td>
<td>54</td>
<td>62</td>
</tr>
<tr>
<td>Unweighted (dB) (30 MHz)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
scene lighting of only 7.5 lux, using an f/1.4 lens at full open aperture and 18 dB of master gain. In motion picture film terms this means that the camcorder has an Exposure Index of about 580 (3200K) in Tungsten illumination. This increases to almost 650 in Daylight illumination (about 6500K).

As shown in Figure 13, the low noise floor defined by the 54-dB SN performance coupled with the high CCD saturation ceiling endows the camcorder with an effective Exposure Latitude in excess of 10 f-stops, actually exceeding that of contemporary 35mm motion picture film.

The HDVS VTR Design Details

The most daunting challenge of all lies in the crucial area of the digital HD VTR compression scheme to be mobilized. The ground rules had already been established with the choice of the Digital Betacam tape and transport; reliable digital recording would require somewhere in the neighborhood of 120 to 150 Mbits/sec total video payload on tape. The SMPTE 274M production standard involves about 1000 Mbits/sec of baseband video origination. Clearly, somewhere in the neighborhood of an 8:1 compression ratio would be involved.

Considerations of the larger recording system requirements to support contemporary production techniques, such as long recording time, field accurate editability, multigeneration capability, read-modify-write, high-speed shuttle, four channels of uncompressed digital audio, and the requisite error correction and error concealment strategies all entered into the multifaceted quest for an appropriate digital compression scheme. The final choice of this scheme would be judged on the basis of extensive subjective testing of a wide range of picture content. Computer simulation was extensively used to search for basic design principles in the compression scheme.

Selection of the Compression Algorithm—the Initial Decisions

Anticipating the full 1920 x 1080 HD sampling format of the North American SMPTE 274M HDTV production standard produces the following realtime total digital data rate presented in Table 5; the final digital payload on tape must be shared between the compressed video components, the four channels of digital uncompressed audio, and the necessary error correction data (for both video and audio).

The need for close vigilance to camcorder size, weight, and power consumption mitigated against the employment of complex motion estimation and motion compensation techniques. Initial simulation began with a
video payload of 124 Mbits/sec, which in turn suggested a required compression ratio of 8:1. The traditional deployment of both error correction (to manage the smaller burst errors and head clogs) and error concealment (to cope with larger burst errors, tape scratches, or a damaged head) was also decided upon in the interests of minimizing data overhead, and thus the total recorded data.

This early simulation, according to the model shown in Figure 14, soon revealed that compression distortions on certain difficult pictures were unacceptably visible. Unlike 525/625 SDTV, where the imagery is heavily arbitrated by zoom lenses that continually seek to frame pictures having captured higher frequency detail content (commensurate with the very limited 4 or 5 MHz of transmission bandwidth), the six-fold increase in spatial resolution supported by HDTV allows zoom lenses to more dramatically open and capture imagery having much wider angles of view. Thus, such HD imagery can have a far greater complexity in terms of horizontal and vertical spatial frequency content than SDTV, and this stresses the dynamic operation of the DCT and associated quantizer. In other words, HDTV picture content is far more taxing on the bit-rate-reduction system than the comparatively pre-filtered 525-line SDTV video.

The extensive computer simulation employed further exposed a performance degradation on certain complex pictures in multigeneration recording. This did not meet the established criteria of handling a wide variety of high-detail image content with minimum subjective degradation over at least five generations of recording. The necessary intraframe DCT scheme simply could not sustain this. A two-step approach was thus decided upon: digital spatial prefiltering of the three component video signals to prepare them for a more modest degree of compression, followed by the actual data compression algorithm.

**First Level of Data Reduction—Digital Spatial Prefiltering Prior to Compression**

Accordingly, the simulation was diverted to a search for an appropriate prefiltering of the three video com-
nent signals prior to their being compressed. Digital spatial preprocessing of the camera-processed 1920 x 1085 images was resorted to, with careful assessments of a variety of taxing HD image content guiding both the final combination of this video data preprocessing and the subsequent degree of compression. Recognizing that a departure from the full spatial resolution of the HDTV production standard was unavoidable, a high priority was given to choosing a spatial filtering scheme that still maintained high subjective pictorial sharpness in those pictures inherently containing a considerable degree of image detail.

The final decisions on the digital preprocessing yielded the pragmatic choice of retaining of full vertical information in all three components that employment of horizontal bandlimiting in the form of digital preprocessing. Extensive computer simulation of the concatenation of spatial filtering and DCT-based intrafield/frame compression, which also included a careful examination of multigeneration recording, finally converged on a 3/8 horizontal band limitation in combination with the now more modest degree of bit-rate reduction.

The 1920 horizontal samples from the camera are resampled at 1440 samples for luminance. The 960 sampled color difference signals P_r and P_b are also resampled at 480 samples. This introduced an 8/5:1 reduction in the total video data payload, producing a net data rate of 622 Mbits/sec prior to compression proper.

**Second Level of Data Reduction—Compression Algorithm**

Meanwhile, ongoing refinements to basic digital recording technologies now permitted very reliable recording of 140 Mbits/sec of video in the field. Hence, a mere 4:4:1 compression ratio was needed. Use of the DCT within the compression algorithm offered the best choice. The fact that a great deal of real-world imagery is either stationary or slow moving argued for the deployment of intraframe processing to realize a higher energy compression within the DCT. Against this, the need to properly manage inevitable fast-moving scene material in sports coverage called for consideration of intrafield processing. Accordingly, an adaptive field-frame process was decided upon. This suggested a potential restriction of frame rather than field editing in this recording format, a limitation considered unacceptable today. Happily, this was overcome by the incorporation of preread heads and appropriate signal processing so that field-accurate editing would be fully preserved in this recording format.

Following the basic decision making on the overall bit-rate-reduction scheme, work continued on the optimization of subjective picture quality. Considerable exploration of techniques relating to picture activity information were conducted. This scheme relies on the fact that areas of pictures having high activity are visually masked and can thus withstand a higher degree of compression than the rest of the picture content. More bits are dynamically assigned to those more quiescent areas of the pictures to which the eye-brain is more visually sensitive. For example, on a sharp picture transmission the luminance edge generally masks the color information so bits can be dynamically allocated from the color difference channels to augment the luminance edge.

Additional improvements were incorporated by capitalizing on further visual limitations related to the fact that the eye-brain sees compression distortions more in certain colors than in others. So once again, more bits were assigned to those colors. Final optimizations to these subjectively based refinement techniques were done on multigeneration recordings with and without picture shift.

**The HD VTR Digital Processing Technique**

Data rate control is an essential part of any digital VTR system. Similar techniques to those developed for the new Betacam SX recording format, based upon handling the MPEG-4/2 Profile @ Main Level, were employed to ensure a constant quantity of digital data for each television frame. This in turn ensured that field-accurate editing is preserved. While ensuring this constant data rate, it was also desirable to minimize the overall system delay in terms of the encode + record-to-confidence + decode cycle. This expedient both facilitated the popular read-modify-write feature within a single VTR and ensured overall ease of use of the VTRs within a post-production system.

The adopted approach achieves an effective delay of just one frame despite the employment of a 2-frame RAM in both the encoder and the

---

**Table 7 — A Comparison between the HD Camcorder and Studio Editing Recorder**

<table>
<thead>
<tr>
<th></th>
<th>HD Camcorder</th>
<th>HD Studio Deck</th>
</tr>
</thead>
<tbody>
<tr>
<td>Drum diameter (mm)</td>
<td>61.7</td>
<td>81.4</td>
</tr>
<tr>
<td>No. record heads</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Drum speed (Hz)</td>
<td>45</td>
<td>90</td>
</tr>
<tr>
<td>Writing speed (mm/sec)</td>
<td>11.5</td>
<td>22.9</td>
</tr>
<tr>
<td>Recording time (minutes)</td>
<td>40</td>
<td>124</td>
</tr>
<tr>
<td>Audio quantization (bits)</td>
<td>(S-cassette)</td>
<td>(L-cassette)</td>
</tr>
<tr>
<td>No. of audio channels</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>4</td>
</tr>
</tbody>
</table>
decoder, as indicated in Figure 15. Each television frame is made up of 540 code blocks (CB). Each CB contains 45 luminance-compressed DCT blocks and 15 compressed DCT blocks each for the two color difference components \( P_r \) and \( P_b \). The desired fixed bit rate is achieved by setting the number of bits for each CB to be: \( \text{Bits/CB} = \text{Bits per Frame}/540 \). 

A shuffling algorithm was used to ensure that each CB contained a reasonable distribution of DCT blocks encompassing the entire television frame. The 2-frame RAM was used to guarantee this and also format the data into the CB structure. The shuffle pattern ensured an optimum distribution of DCT blocks from the entire frame and, in turn, proper distribution among the multiple record tracks. This considerably improved shuffle picture performance.

The choice of both the number and the size of the DCT blocks within each CB was an integral part of the extensive experimentation on recorded picture quality. The chosen block size was a compromise based upon careful examination of a wide array of pictures. Too many blocks would soon reach diminishing returns in terms of the performance increase versus design complexity, while too few blocks would result in each CB within the frame having too great a disparity in compression ratios.

Each CB comprises 77 x 64 original video samples. The chosen compression ratio of 4:4:1 produces 1080 bytes per CB. To support standard byte-based Reed-Solomon error correction, and to provide good picture-in-shuttle performance, the block size recorded on tape needs to be less than 256 bytes. To implement this, each CB was subdivided into 5 sync blocks. The sync block structure is shown in Figure 16. Figures 17(a) and 17(b) show the inner and outer Reed-Solomon error correction surrounding both the video and the audio macroblocks.

Maximum error protection for the recorded compressed video data was supported by further division of the sync block into cells. Each cell is the size of one compressed DCT block (if all the luminance DCT blocks are compressed to the same degree and if the color difference component DCT blocks are compressed to half the num-
Table 8 — Recording Performance Comparison of HD Digital Camcorder and UNIH Analog VTR

<table>
<thead>
<tr>
<th>Performance Specification</th>
<th>New HDVS Camcorder</th>
<th>HDV-10 UNIH Recorder</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bandwidth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Luminance (Y)</td>
<td>24 MHz +0/-1.5 dB</td>
<td>20 MHz +1/-6 dB</td>
</tr>
<tr>
<td>Color difference (R-Y and B-Y)</td>
<td>6 MHz +0/-1.5 dB</td>
<td>7 MHz +0/-6 dB</td>
</tr>
<tr>
<td>(Full vertical)</td>
<td>(Line sequential)</td>
<td></td>
</tr>
</tbody>
</table>

Figure 19. Final image resolution as recorded on tape (Note the high-horizontal MTF despite VTR bandlimiting; note also the quite isotropic portrayal on an HD display which accounts for the observed high picture sharpness).

Recording Performance

Clearly, some decisive design initiatives were resorted to in the VTR in order to achieve the quite radical leap forward sought by producers and broadcasters for high definition using a highly mobile, battery-powered acquisition system. The final picture captured on tape is largely determined by a high-performance front-end HD camera imaging system (fulfilling the promise of the SMPTE 240/260M HDTV Studio Production Standard). It is, however, subsequently curtailed (in terms of total picture information) by some necessary bandlimiting expediences in the associated digital recorder.

In practical terms, the compromises struck in the digital VTR design achieve a significant improvement in performance over the present-generation analog UNIH recorder (the only practical VTR option currently available in terms of any form of field acquisition—and even then, mobile battery operation is not practical) as indicated in Table 8.

Subjective Picture Sharpness of Recorded Imagery

In terms of picture spatial resolution, the subsequent testing of the recordings yielded a surprisingly high subjective picture sharpness. One explanation for this is that the original HD camera sampling does create images having a high MTF (especially horizontally), and this high MTF is fully preserved right up to the new band-edge imposed by the digital horizontal prefiltering, namely 24 MHz for the luminance Y and 8 MHz each for the two components, P, and P

As described in the seminal work on picture sharpness by Schade,16 the visual sharpness of a picture is proportional to the square of the area under the MTF curve. Thus, while the very high spatial frequencies captured by the camera (24 to 30 MHz) are lost in the recording process, the spectral energy here is typically very low, and the pictures overall still retain full contrast over the VTR-captured frequency domain (0 to 24 MHz) (Fig. 19).

Figure 19 also shows that the images captured on tape by the camcorder are close to isotropic in terms of their visual portrayal on an HD display. The shaded area in the top curve is the recorded horizontal MTF (with no image enhance-
Camcorder Operational Features

Setup Card

The HD camcorder includes a powerful operational capability first introduced into the Digital Betacam camcorder, namely, the miniature plug-in setup card that facilitates prealignment of the camera to produce a desired look to the captured imagery. This is a small integrated circuit (IC) electronic memory card, similar to a smart card, that is preprogrammed to store desired digital data settings of those many operational video adjustments that, in combination, exercise an aesthetic control over the camera’s picture reproduction. These can include colorimetry, transfer characteristic, horizontal and vertical image enhancement, skin-tone detail, electronic shutter, etc.

Data can be readily written to and read from the setup cards. The data writing is accomplished by inserting the card into the small receptacle designed into the side of the camcorder and actuating an associated up-down switch that alters the relevant digital control settings (the viewfinder portrays a menu guide). Write protection and ID editing are provided.

In addition to allowing a camcorder to be prepared for a particular type of shooting (a function of scene content and lighting, and the artistic desires of the producer/director) in a manner that frees the videographer from considerations of in-the-field video adjustments, the setup card has a special relevance in digital electronic cinematography.

Just as the insertion of a chosen motion picture film stock preprograms the film camera for a specific shooting environment and aesthetic look, the setup card can also be preprogrammed to emulate a given film stock in terms of colorimetry, tonal reproduction, etc. This empowers the film cinematographer to accurately achieve an HD image having attributes similar to those that might otherwise be created from a film origination followed by telecine transfer to HD video. In fact, because of the precision inherent to the setup card setting of the DSP camera, the cinematographer can be endowed with a higher degree of image capture predictability than in the case of film shooting.

Field Playback from the Camcorder

The digital recorder of the camcorder contains all the requisite digital processing to support full-component video playback in the field without the use of any externally connected playback adapter box. Three BNC connectors on the camcorder facilitate baseband component analog HD video connection to an HD monitor. This adds an important convenience when shooting on remote location.

Conclusion

An important breakthrough in HD portable field acquisition has been made. While somewhat unorthodox in approach, the technical choices made in both the shoot side (the CCD/DSP camera section) and the capture side (the bit-rate-reduced digital component VTR section) did allow an unprecedented leap forward to the world’s first all-digi-
The HDTV camcorder and the March to Marketplace Reality

The goals of achieving a far more mobile, power-conservative, and lower-cost HD shoot-and-capture system have been realized, as testified by the summary in Table 9. To put a practical perspective on this, as might be assessed by the pragmatic director of photography or a videographer, the physical enhancements achieved over current HD shooting techniques are graphically depicted in Table 10.

The use of extensive computer simulation and a team of experts in the field of compression was flanked by considerable subjective testing, all of which contributed to a final optimization of picture quality that holds high promise for high-definition electronic field production (EFP) and electronic newsgathering (ENG) program origination.

High-definition television's greatest visual impact lies in its ability to capture images having angles of view considerably beyond those that can be imaged within the constraints of the 525/625-line television systems. Such wide-angle imagery can impart a heightened sensation of reality to television pictures and nowhere more so than in images of the great outdoors.

The arrival of an HD camcorder has long been awaited by those who produce wildlife and natural history programs, underwater photography, aerial photography, coverage of mountaineering feats, major sporting events, and drama location shooting. It is to be expected that the debut of this engineering tour de force will spurn broad adoption of HDTV shooting over a wide variety of television program origination.

The evolution of electronic cinematography for moviemaking has been hampered by the absence of HD acquisition equipment that could in any way rival the convenience and mobility of the motion picture film camera. The new popularity of the DVW-700 all-digital camcorder (transferring to 35mm for theatrical release by electronic beam recorder [EBR] techniques) as an alternative among low-budget moviemakers for the traditional Super 16mm/16mm origins blown up to 35mm film, is surely a portent of the future. The imaging prowess of an HDVS camcorder, in cost-effectively originating programming for subsequent transfer to a far greater resolution 35mm motion picture film, is certain to flank the spectacular work already well under way using digital imaging workstations for important elements of major movies.

References

The Authors

Larry Thorpe is a renowned industry expert in the field of video acquisition and is generally considered to be one of the leaders of the HDTV movement. His impressive list of accomplishments includes pioneering HDTV market development in the U.S. and holding ten patents in the broadcast development field. Thorpe was promoted to his current position as vice-president, acquisition systems, business and professional group, Sony Electronics, in 1995. He has business responsibility for all of the broadcast studios and portable cameras (including HDTV) for the Sony BPS Division.

Prior to his current position, Thorpe was vice-president, production technology for Sony Advanced Systems. He was responsible for HDTV market development and represented Sony on ATSC technology groups, as well as various SMPTU working groups dealing with high-definition electronic production.

Fumio Nagumo is general manager of Atsugi Broadcast Camera Division, Sony Corp. In the position he has been responsible for the development of Sony's first HDVS CCD camera, the and its successors. Nagumo has also been responsible for the company's SDT camera family. His developmental responsibilities also include the DSP/CCD camera for the HDW-700 described in this paper. Joining Sony in 1969 after graduating from the Tokyo Institute of Technology with a masters degree in electronics, he was primarily engaged in the research and development of image sensing technologies.