XDCAM

Codec Technology for XDCAM Tapeless Products and Systems

By Hugo Gaggioni, Chief Technology Officer, Sony Broadcast & Production Systems Division

Introduction

In deciding on a compression technology for the XDCAM® family of HD tapeless products, various end-to-end operational issues needed to be considered. Of course preserving high quality video was a prime concern. This necessitated using a codec that could preserve the best possible picture quality without using unrealistic amounts of storage space. Since XDCAM products record video and audio as data, the compression method chosen must also be critical as a factor for both conservation of storage capacity and to preserve adequate transfer speed across network connections.

Sony adopted MPEG-2 Long GOP (Group of Pictures) video compression technology, which is used worldwide in DVD and Blu-ray™ video, digital and BS digital (Broadcasting Satellite) broadcasting and as a transmission method in many legacy broadcast operations. Long GOP enables a high compression ratio using video data reduction across multiple frames which helps deliver high quality HD video at lower bit rates.

Some other system considerations related to software processing of compressed streams, comparison with other compression schemes, and complete system efficiencies were also taken into account. A description of this system-wide analysis for the selection of compression technology is given in a later session of this document.

Because the A/V/metadata contents are recorded as data files, it was important to provide advanced devices that could allow media access over network interfaces by equipment remotely located from the recording machine. In the world of VTR's, tapes were the medium used to deliver contents, therefore the compression scheme adopted favored tape technologies. However, some of the above points needed to be reconsidered in Network/File operations. While more efficient compression schemes have emerged, their current implementation is not quite up to broadcast quality, at the modest data rates required in networked systems.

With the overwhelming acceptance of HDV™, most non-linear editing solutions have adopted MPEG-2 Long GOP compression techniques. It's already common to use editing software that can natively manipulate Long GOP compressed data. Sony XDCAM HD tapeless products are ideally suited to take advantage of this compression scheme so edit operations can be performed smoothly using Long GOP compressed content. In addition XDCAM products take advantage of latest advances in VLSI integration which allows for small, low power consumption devices for field equipment.

XDCAM HD Tapeless Solutions – MPEG-2 Long GOP Compression Scheme

1. MPEG-2 Standard

As mentioned above MPEG-2 Long GOP was adopted for XDCAM HD products to allow both high-quality video and audio, and efficient compression. What follows describes the parameters used by MPEG-2 for the compression of HD signals for XDCAM-HD tapeless products.

Profile and Level
MPEG-2 MP@HL (HQ and LP mode)
and MP@High-1440 (SP mode)
MPEG-2 422P@HL (for future 422, full 1920x1080
and 1280x720 signals)

Video Resolution
Horizontal x Vertical = 1440 x 1080
(Next Generation of XDCAM Tapeless products:
1920 x 1080 and 1280 x 720)

Bit Rate

HQ Mode:	35 Mbps	SP Mode:	25 Mbps
LP Mode:	18 Mbps	422 Mode:	50 Mbps

GOP Structure
15 frame/GOP (59.94i, 29.97p)
12 frame/GOP (50i, 25p, 23.98p)

Stream Method
Elementary Stream

2. Multiple Bit Rates

3 bit rates are supported for XDCAM HD, each providing different recording times:

HQ Mode:	35 Mbps	Recording Time: More than 68 minutes**
SP Mode:	25 Mbps	Recording Time: About 90 minutes
LP Mode:	18 Mbps	Recording Time: More than 122 minutes**

** Times are approximate due to variable bit rate.
Recording time will vary based on the complexity of the video material.

HQ mode provides the highest quality picture. LP mode allows for recording times of approximately two hours. It's possible to switch modes depending on the video shooting situation. For HQ and LP mode, XDCAM-HD uses MP@HL. At High Quality mode, there is a larger decoder buffer than SP mode (which uses Main Profile @High-1440). Hence, along with the VBR technique employed, higher picture quality can be guaranteed with more dynamic and flexible contents. In the case of SP mode, consideration was given to using the MPEG ES stream in a compatible manner with the compression used in HDV 1080i/p products. Moreover, installing a TS I/F option, it is possible to directly connect XDCAM-HD products to HDV 1080i/p products over a i.LINK® interface (IEEE-1394) interface.**

** i.LINK is a trademark of Sony used only to designate that a product contains an IEEE 1394 connector. All products with an i.LINK connector may not communicate with each other. Please refer to the documentation that comes with any device having an i.LINK connector for information on compatibility, operating conditions and proper connection.

3. Adoption of VBR

For HQ and LP mode, a Variable Bit Rate compression method was adopted in order to take advantage of the random access nature of tapeless media. It is different from the 2 pass* VBR compression technique used in DVD authoring. This codec was specifically developed for MPEG to perform high-speed efficient picture coding.

The XDCAM product line is based on the MXF file format and MPEG-2 Long GOP for exchanging and transmitting Video, Audio and Metadata. In the VBR (Variable Bit Rate) encoding method adopted for XDCAM HD products, the maximum bit rate that can be processed in real-time is fixed. This results in the high level of picture quality exhibited by XDCAM HD products over different picture material while minimizing the file size and maximizing transfer speed.

* Encoding with multiple passes lowers the complexity of input resolutions and assigns the most appropriate bit-rate depending on the picture. This technique is mostly used when processing video content using software tools for DVD or Blu-ray authoring.

4. Development of MPEG-2 HD codec LSI

Sony developed an MPEG-2 HD Codec LSI that uses extremely small amounts of power for use with HDV and XDCAM HD technologies. This MPEG-2 HD codec LSI can be found widely deployed in the consumer and professional HDV product lines manufactured by Sony. The same codec has been adopted for XDCAM HD tapeless products, with the addition of extended "bit rate" range, as well as its own VBR rate control algorithm, thus delivering high speed and quality MPEG compression. In case of the XDCAM HD camcorder, the use of this highly integrated LSI allows lower levels of power consumption – much lower power levels as compared to the family of XDCAM SD products.

Figure 1:
MPEG-2 HD codec LSI
(13 mm per side)

Selection of Video Compression Codec for XDCAM-HD

In the implementation of XDCAM products for SD applications, the 23.3 GB capacity of the Professional Disc™ media was used for the storage of the following video compressed data:

- DVCAM (4:1:1/4:2:0) at 25 Mbps ......approx. 85 minutes
- MPEG-2 IMX (422P@ML) at 30 Mbps ....approx. 68 minutes
- MPEG-2 IMX (422P@ML) at 40 Mbps ....approx. 55 minutes
- MPEG-2 IMX (422P@ML) at 50 Mbps ....approx. 45 minutes

Additional capacity is also reserved for the storage of the associated audio and metadata information (i.e., proxy audio/video, clip and user-definable metadata).

In evolving to XDCAM-HD tapeless solutions, Professional Disc and Solid State Express Card™ media are being used for the storage of HD signals. In the first generation of HD products, XDCAM is storing HD signals in the 1080i/p format. XDCAM EX™ (with its use of Express Card media) and MPEG HD 422 tapeless products process multi-format HD signals. As is readily apparent, the storage of HD signals necessarily requires use of higher data compression ratios than those used for standard definition signals. In the selection of a video compression system for XDCAM-HD products, a number of algorithms exist which could be considered for the production and professional applications envisioned for the family of XDCAM-HD tapeless products.

Sony carried out extensive analysis of a number of contemporary video compression schemes, taking into account not only compression efficiencies, but many performance parameters which are fundamental to establishing a commercially practical system.

The algorithms studied were:

- MPEG-2 Long GOP (4:2:0 and 4:2:2)
- AVC-Intra and Inter-frame (Long-GOP)
- JPEG-2000 (intra-picture by definition)

(AVC: Advanced Video coding is also known as H.264 or MPEG-4 part 10).

Before describing the results of the technical studies it is important to emphasize an examination of data rates suitable for broadcast production operations.

A number of HD, tape-based formats are already in existence for high-end acquisition and program production. In particular, HDCAM-SR™ has generally been accepted as the highest picture quality for the most demanding applications such as electronic cinematography and digital intermediate masters. However, this high performance is achieved by the use of not only a state-of-the-art, high picture quality, compression algorithm (MPEG-4 Studio Profile), but also high data rates of 440 Mbps and 880 Mbps for 1080 60P RGB applications.

Another well accepted format has been the HDCAM™ format, serving as a mainstream production format for HD television production, and even adopted as a news gathering format in some countries. But despite its success as a high performance, cost efficient HD production format, its high data rate (144 Mbps for the video information alone) and tape-based characteristics have prevented it from becoming a ubiquitous broadcast format.

It has been apparent to many manufacturers of professional broadcast equipment that there are many economic efficiencies at lower data rates (between 25-50 Mbps), and that today most broadcast plants will favor the use of lower data rates for HD program production applications.

During the design phases of XDCAM-SD products, particular attention was paid to the use of MXF file-storage and the use of IT network interfaces, such as Gigabit Ethernet and i.LINK (IEEE-1394). The reasonable data rates assigned for DVCAM™ and MPEG IMX® compression schemes for the various performance levels of picture quality facilitated significant reductions in storage requirements, as well as faster-than-real-time, network-based, exchange of program material. When examining the use of XDCAM technologies for HD storage and processing, a new set of technical questions arose which needed to be addressed to reach sensible conclusions for data rates for broadcast operations. Referring to figure 2, one can easily formulate the following questions:

- What data rates would be required for adequate storage times of HD signals in the tapeless media of XDCAM products?
- Would the quality of the compressed HD video signals be sufficient for multi-generation production work?
- Are the data rates efficient enough to satisfy storage and transmission costs?
- Are the data rates of the compressed HD signals manageable with present and near-future technologies for lap-top and desktop computers?
- Are the data rates of the compressed HD signals low enough to be carried by the new generation of broadcast digital microwave devices?
- Is the picture quality at these low rates satisfactory for archiving applications and future repurposing of archived material?

Sony's analysis of the aforementioned questions (and many other considerations) led to the decision to select compressed data rates for HD tapeless storage in a range approximately between 18 Mbps to 50 Mbps. Also, it was decided to make use of the mature and efficient compression technology of MPEG-2 Long GOP for all operations involved in the broadcast production environment. What follows describes some of the conclusions reached during the comparison and analysis of the candidate video compression schemes mentioned above.

New video compression techniques vs. matured, complete system solutions.

MPEG-2 was first specified in 1992, with its international publication in 1994. This compression scheme was designed to produce very high picture quality at moderate to high data rates. It has taken more than 10 years and approximately 6 generations of newer and more powerful VLSI implementations of the original algorithm to reach significant lower bit-rates, while still maintaining good levels of picture quality. (see to figure 3)

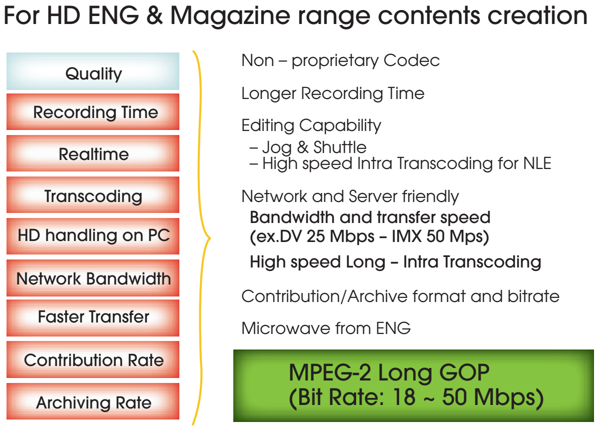


Figure 2:
System considerations for selection of codec and bit-rates for XDCAM-HD tapeless products

The need for more channels of compressed video on distribution channels, or longer times of storage capacity has fueled the development of more advanced compression schemes such as the Advanced Video Compression system (AVC, also known as H.264 or MPEG-4, part 10). The AVC Long GOP technique is relatively new (standardized in 2003), hence, there is a necessary learning curve required to implement the more complex mathematical calculations in the form of VLSI circuitry, as well as to gain availability of processing speeds and newer semiconductor design processes to maintain the power consumption levels as low as possible.

Today, Sony makes use of a single MPEG-2 HD encoding/decoding VLSI device which is state-of-the-art from the algorithmic, speed, power consumption, and manufacturing viewpoints. The VLSI manufacturing technology has not advanced to the level necessary to enable implementation

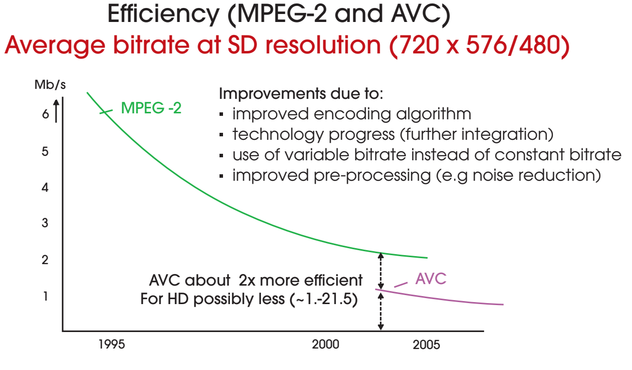


Figure 3:
Improvements in compression efficiency of MPEG-2

of an AVC Long GOP codec for HD applications on a single device, and under the same processing speed and power consumption conditions as in use with MPEG-2 Long GOP today. The current implementations of real-time compression, and decompression of video signals, using the techniques defined by the AVC Long GOP algorithm, have only been implemented, to date, in consumer quality applications.

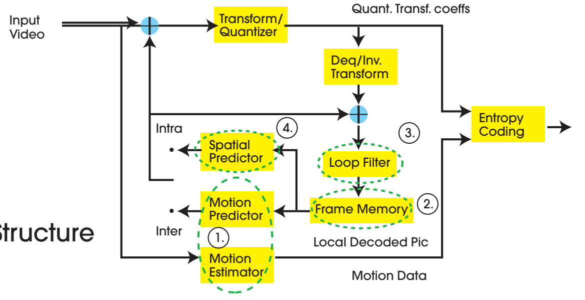
While it is undeniable that the VLSI industry will make progress towards faster processing devices, which will eventually give rise to single devices for AVC Long GOP compression at HD resolutions, there are still some fundamental considerations about whether the AVC algorithm – and all of its mathematical tools – is truly suitable for high-quality HD program production. More specifically, the touted 50% increase in compression efficiency of AVC over MPEG-2 has been realized not by the discovery of a single, new compression technique, but by the enhancements of many of the techniques already in use in MPEG-2. (see figure 4)

While small improvements, taken altogether, they amount o significant compression gains, under the assumed design objectives: i.e., the best possible quality at minimal data rates. Since delivery of compressed streams is the original premise of AVC, some of its algorithmic tools could produce very poor results when presented with a different set of criteria, such as the delivery of high-picture quality during multi-generation operations.

Indeed, the use of the internal "Loop Filter" has been highly questionable for multi-generation performance. The "Loop Filter" is a great tool in reducing blocking artifacts in the reconstructed picture, which in turn helps in making better predictions of the incoming macroblocks, and overall reduction of the compressed data rate. This technique is indeed very useful if the objective is a single pass compression for video distribution application. However, the same filter will soften the picture repeatedly as the material is subjected to multi-passes, as is typical of program production operations. Hence, the "Loop Filter" is, – along with some other techniques such as the "intra-frame prediction tools" – responsible for the poor multi-generation performance of AVC, in either its Intra – or Inter-frame implementations.

Key Elements for Improvements
in Compression Efficiency and/or Picture Quality

AVC Basic Structure



Features of Video Processing

AVC is super set of MPEG-2

1. Seven kinds of motion estimation
2. Expands picture selection availability for motion estimation
3. Loop filter for reduction of blocking artifacts
4. Intra frame Prediction

Figure 4

Furthermore, initial tests have shown that the computational load of using the "Loop Filter" can be as much as 50% of all the mathematical calculations carried out, which adversely affects software implementations of AVC.

It is not only the lack of dedicated, power efficient VLSI hardware that affects the adoption of AVC Long GOP compression for high-end applications: contemporary levels of software computation (in laptops or desktop devices) cannot achieve anywhere near real-time encoding or decoding of high definition video signals using the AVC algorithm.

For example, employing a single Pentium-4® 3GHz CPU (Dual-Core, Hyper-Treaded) processor, typical decoding of HD AVC compressed streams can be 3-4 times slower than real-time (compared to 2 times faster than real-time performance using MPEG-2 Long GOP). Encoding of HD signals can take between 15 to 50 times longer than real-time (compared to 2 to 5 times slower than real time for MPEG-2 Long GOP).

Hence, the computational demands imposed by AVC on software platforms using single CPU engines generally prevented its commercial use for production applications. One possible solution to this problem would be to apply distributed processing and to employ multiple workstations in parallel to handle the tasks of encoding and/or decoding of AVC streams for real-time performance. But again, this becomes not only very complex (from the viewpoint of software design) but also very expensive because of the large number of PCs involved.

As stated, the motivation for creating advanced compression techniques such as AVC, was to exploit, primarily, Long GOP (motion estimation and compensation across multiple frames) to be able to maintain acceptable picture quality at very low bit-rates (as compared to existing algorithms). However, the efficiency gain exhibited at low bit-rates begins to diminish and become closer to the performance levels of MPEG-2 as the bit-rates and picture quality levels – acceptable for broadcast applications – increase.

For Long GOP use, the differences in picture quality between MPEG-2 and AVC – for HD pictures – become minimal at bit-rates above 20 Mbps. At the higher rates of 25, 35 and 50 Mbps, not only, does MPEG-2 Long GOP perform extremely well compared to AVC Long GOP, but has the added advantages of multi-generation robustness and lower computational software load which swings the selection balance, once again, towards the MPEG-2 techniques.

What about some of the recent proposals for AVC Intra-frame codecs?

The advantages in compression efficiencies offered by AVC are almost entirely based on its better exploitation of frame-to-frame redundancies. By reducing only redundant information within a picture, AVC-I can only afford to exhibit modest compression ratios while maintaining acceptable levels of picture quality.

When compressing HD images with the use of intra-frame only techniques, the AVC algorithm is forced to operate at high compression factors – relative to the lower values expected for high quality broadcast applications (see figure 5). These high compression ratios not only produce unacceptable results (even on a single pass), especially with complex pictures, but can also render the "acceptable" pictures very "weak" for multi-generation work.

It has been said that "AVC is more efficient than MPEG-2 Long GOP." This statement would be accurate when comparing an "AVC Long GOP" scheme (with all of its compression techniques at their fullest performance) against its equivalent implemented with MPEG-2 Long GOP. However, an AVC Intra-only compression discards most of its advanced techniques resulting in an intra-picture scheme with just 15-20% more efficiency as compared to other intra-only techniques (e.g., DV, JPEG, MPEG-2 Intra, etc). Therefore, the long GOP (multi-frame) techniques of MPEG-2 create significantly higher picture quality than an AVC-Intra scheme of similar data rate. Or, viewed from a different perspective, MPEG-2 Long GOP produces approximately from 1/3 to 1/2 the data rate required by an AVC-Intra scheme at similar levels of picture quality. Figure 5 depicts the relative picture quality of various contemporary video compression algorithms, along with their compression ratios and use of single or multiple frame techniques.

Even with the use of only single frame information, AVC is still very complex to compute – due to "Intra-Picture" prediction modes, and mostly by newer mathematical techniques for entropy encoding (such as CABAC and CAVLC). Based on Sony's preliminary tests, more than 150 Mbps would be necessary to produce acceptable levels of picture quality for multi-generation work on HD signals. At these high bit-rates, and with the prohibition of software solutions for real-time encoding/decoding (at the present time), all the possible benefits of the new algorithm are negated, when compared to existing intra-frame techniques and to the MPEG-2 Long GOP system.

With the existing generation of Dual – and Quad-Core CPUs it is now possible to encode/decode multiple streams of HD MPEG-2 Long GOP signals in real-time with software-only tools.

MPEG - 2 Long GOP:
High Quality, File - Size & Speed Efficiencies

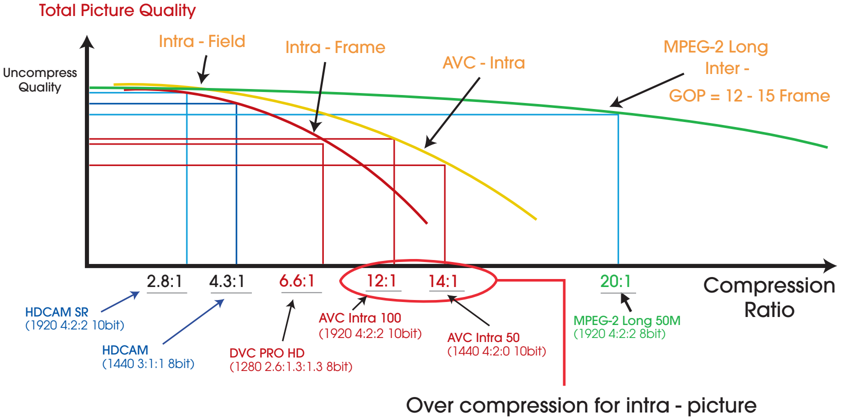


Figure 5

However, the complexity needed to compute the new entropy coding techniques of CABAC and CAVLC at HD resolutions prevents AVC-Intra streams from being encoded or decoded at any speed closer to real-time. Figure 6 depicts an estimated time-line of CPU computational capabilities based on current and future deployment of multi-core devices and the expected availability of software applications for compression/ decompression of HD streams for MPEG-2 Long GOP and AVC Intra GOP techniques. From this data one can conclude that the exchange and processing of AVC-Intra files for practical broadcast operations is all but impossible in the near term, unless it is handled by hardware-based solutions at all points of decompression and re-compression of the HD material.

Sony's selection of MPEG-2 Long GOP

After an extensive series of tests, and what is believed to be a thorough evaluation of the real capabilities of several contemporary video compression schemes, Sony selected the matured and highly integrated MPEG-2 Long GOP compression system. The level of picture quality and overall practical feasibility achieved with this algorithm surpasses all other examined techniques at the present time. Hence, it is envisioned that the use of MPEG-2 Long GOP, at bit-rates of up to 50 Mbps, will become most common for broadcast operations.

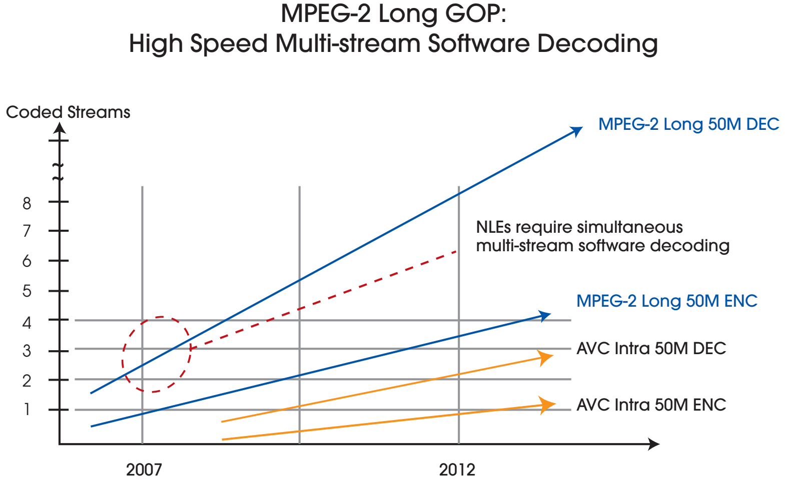


Figure 6

Recent developments of real-time, compressed stream processing, indicate that the past obstacles of native stream manipulation have been overcome, both in software and hardware implementations. This further presents opportunities for continuing the use of Long GOP techniques at all levels of content production – from consumer, to professional and broadcast applications.

As the LSI and signal processing technologies continue to advance, it is expected that in the future the great attributes of the AVC Long GOP compression scheme will become available in cost efficient, practical implementations. Only the further judicious selection of the extensive set of compression tools for AVC will still be necessary to satisfy the requirements of high quality HD productions at somewhat higher levels of compression efficiency.

Sony Electronics Inc.
1 Sony Drive
Park Ridge, NJ 07656