# XAVC: An advanced Video Compression System for Broadcast and Cinematographic Production Applications

## 1) Introduction

In 2003 Sony introduced the use of MPEG-2 video compression for broadcast production applications in the family of Optical and memory-based tapeless solutions known as XDCAM. Today, with more than 400,000 devices, encompassing from EFP/ENG camcorders to studio record/play-back decks, servers and complete suites of editing systems and powerful workflows, XDCAM has become the most used solution for HDTV ENG/EFP and program production applications all over the world. Furthermore, as it was predicted in the initial stages of deployment, the pace of advancing signal processing techniques permitted to commoditize the manipulation of the inter-picture video compression techniques of the MPEG-2 codec. These processes were once considered too complex and costly for the hardware and software computing platforms of the day.

Throughout the deployment cycle of XDCAM products, the technological message of the advantages of MPEG-2 video compression became clear: the use of multi-picture compression was essential to establish the high-level of picture quality demanded by the program production marketplace, while at the same time enabling the reduction in data rates and storage requirements for cost-efficient operations.

Presently, rapid developments in high-resolution / high-frame rate imaging technologies (sensors, displays, signal processing technologies), and in high speed storage technology are helping the creation and introduction of a new set of video standards with superior imaging and display characteristics than those of today's HDTV formats. Figure 1 shows the progress in CMOS imager technology, where Sony has developed and commercialized an imaging sensor with data transfer rates exceeding 30Gbps.

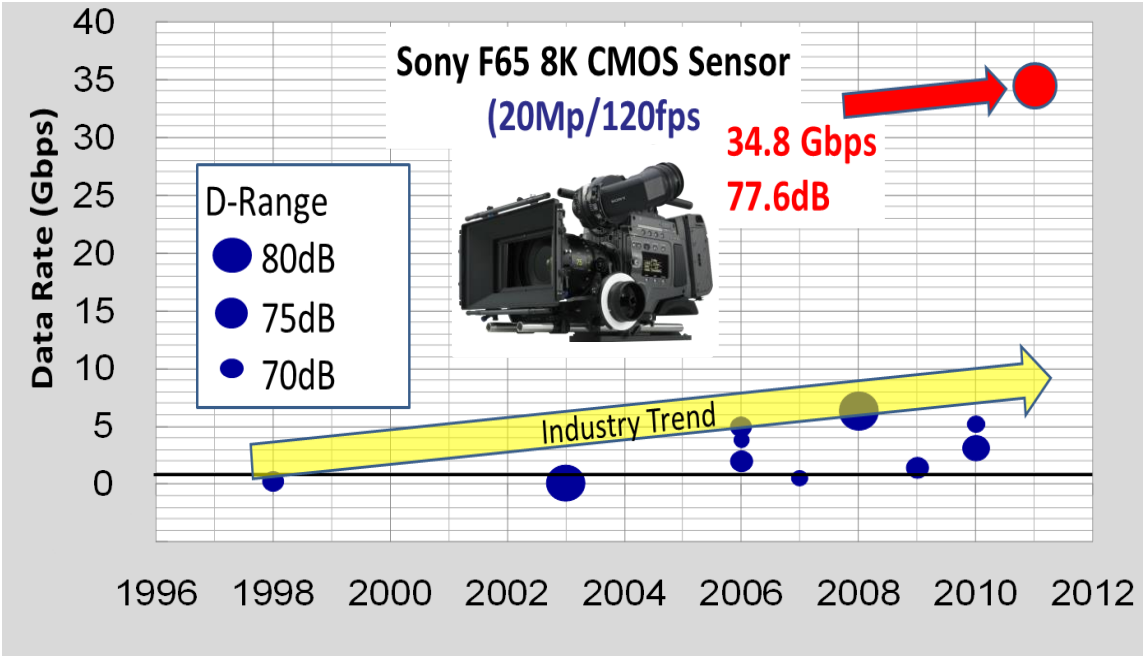



Figure 1. Progress in CMOS technology development

It is expected that high resolution, high frame rates, and high bit rate imaging tools will become more common in the coming years. f SxS Recording Media Technology

and

Figure **2** shows the progress of SxS and XQD memory card technology. Note that the latest cards guarantee real-time recording beyond 1Gbps, and the maximum readout bitrate have exceeded 3Gbps. Also, the recording capacity has significantly increased over the years. And not just maintaining the same compact form factor but to make the size even smaller with the introduction of XQD. The XQD is considered as one of the successors of CompactFlash card at the CFA (Compact Flash Association).

| Media | SxS Pro | SxS Pro+ | SxS Pro+ | SxS-1 | SxS-1 |
|---|---|---|---|---|---|
| |  |  |  |  |  |
| Year of Introduction | 2007 | 2013 | 2014 Nov. | 2009 | 2015 Jan. |
| Model name | SBP-8/ 16<br>SBP-32<br>SBP-64A | SBP-64B<br>SBP-128B | SBP-64C<br>SBP-128C | SBS-32G1<br>SBS-64G1A | SBS-32G1B<br>SBS-64G1B<br>SBS-128G1B |
| Sustained recording data rate | 400Mbps | 1.3Gbps | 1.3Gbps | 240Mbps | 400Mbps |
| Max. readout speed | 1.2Gbps | 1.6Gbps | 3.5Gbps | 1.2Gbps | 3.5Gbps |
| Storage capacity | 8GB/16GB(2007)<br>32GB(2008)<br>64GB(2011) | 64GB/128GB | 64GB/128GB | 32GB(2009)<br>64GB(2010) | 32GB/<br>64GB/128GB |

Figure 1. Progress of SxS Recording Media Technology

| Media | XQD N series | XQD S series | XQD G series |
|---|---|---|---|
| |  |  |  |
| Year of Introduction | 2013 | 2013 | 2014 Oct. |
| Model name | QD-N32<br>QD-N64 | QD-S32E<br>QD-S32E | QD-G32A<br>QD-G64A<br>QD-G128A |
| Sustained recording data rate | - | - | - |
| Max. readout speed | 1.0Gbps | 1.44Gbps | 3.2Gbps |
| Storage capacity | 32GB/64GB | 32GB/64GB | 32GB/64GB/<br>128GB |

Figure 2. Progress of XQD Recording Media Technology

The bandwidth and storage demands imposed by new video formats - such as, 1080P60, QFHD, 4K, etc - far exceeds the system parameters of today's HDTV systems and cannot be met by the established MPEG-2 compression scheme. Hence, a more advanced compression algorithm is required for the new era of "Beyond HD" formats with their accompanying higher resolutions, higher frame rates and pixel bit-depths.

A high performance yet efficient compression technology such as H.264 (MPEG-4) plays a critical role in encoding the vast amount of imaging data generated by modern sensors into a modest file size, so that those images can be recorded on affordable memory cards, and edited/viewed on computers/editing software packages that are readily available.

The H.264/MPEG-4 Part-10 Advanced Video Coding technology is predominantly being used in HDTV distribution systems, such as Blu-ray, digital broadcasting (terrestrial/cable/satellite), and web browsers. Initial standardization documents date back to year 2003, and it has been extended over the years until 2009 to cover much more than HDTV distribution applications. Today the family of operational levels covers an extremely wide range of compressed image data, starting from several Kilobits per

second up to 1.2Gbps, with advanced performance characteristics capable of supporting 4K, 3D, 14bit sampling, and way over 120 frames per second. Sony was one of the active members of the JVT (Joint Video Team) that completed the standard, and has made significant efforts in establishing its use in high-end professional applications by expanding the levels/profiles of H.264. Sony uses the commercial name of XAVC for its implementation of the H.264/MPEG-4 video compression system at the highest level of performance permitted by the official ITU/ISO standard.

The following sections illustrate the attributes and benefits of the XAVC format. It also explains how XAVC fits into the current production workflow, along with well-established compression formats such as MPEG-2, MPEG-4 SStP (SR-Codec), and various camera RAW files.

## 2) The XAVC Format

The Sony XAVC format complies with H.264 Advance Video Coding, level 5.2 which is the highest level of performance determined by the H.264 standard. The video essence is encapsulated in an industry standard MXF OP-1a wrapper, accompanied by audio and meta-data elements. The compression algorithm is completely flexible in establishing compression tools for Intra-frame as well and Inter-frame compression (Long Group of Pictures or Long GOP), for interlace as well as progressive video formats, and with color sampling structures of 4:2:0, 4:2:2 and 4:4:4 at up to 12 bits of color pixel bit-depth.

The primary objective in adopting the XAVC format is to develop a family of professional production tools that can economically handle High-Frame-Rate (HFR) HD and 4K imaging formats.

Figure 4 shows the scope of the XAVC format. Please note that this format table describes the global scope of the XAVC format, and actual product implementation may be restricted to a certain portion or range of values of this table. Also, this chart

excludes any off-speed recording capability (over/under cranking) that some products may offer.

| Range | Resolution | Frame Rate | Color | Max Bitrate | Intra / Long | Audio |
|---|---|---|---|---|---|---|
| 4K | 4096x2160<br>3840x2160 | 23.98p to 59.94p | 4:2:0/8bit to 4:4:4/12bit | 960Mbps | Intra<br>Long | Uncompressed, 48k 24bit x 16ch (max) |
| HD | 2048x1080<br>1920x1080<br>1440x1080<br>1280x720 | 23.98p to 59.94p<br>50i/59.94i | 4:2:0/8bit to 4:4:4/12bit | 440Mbps | Intra<br>Long | Uncompressed, 48k 24bit x 16ch (max) |
| Proxy | 1920x1080<br>1280x720<br>640x360<br>480x270 | 23.98p to 59.94p | 4:2:0/8bit | 28Mbps | Long | AAC LC<br>256 kbps for stereo |

Figure 4. XAVC Format overview

In addition, MP4 wrapping format with 4:2:0 color coding and the use of only Long GOP has also been introduced and branded as XAVC-S to serve the consumer mass applications. This expansion will encourage the growth of 1080/60P and 4K content in the consumer market.

## 3) XAVC and HDTV

It is often described that newly developed compression algorithms are more efficient than their predecessors. Although this is true in the sense that a given image quality can be achieved with less amount of image data (or bit rate), the increased complexity of modern coding schemes do demand more computational power, which could be a major challenge when migrating the production infrastructure and associated workflow from one generation to the next. The amount of processing power requirements to decode a certain compressed bit stream is extremely critical when multiple files are simultaneously used in an edit session.

Today, the vast majority of broadcast and reality-TV industry is operating on the MPEG-2 HD Long GOP format (50Mbps or 35Mbps) for HDTV production due to its small file size, high picture quality, and economy in computational requirements. From breaking news to reality shows, and prime sports events, the data rate of 35 to 50Mbps has become the sweet spot to operate a file based HDTV infrastructure.

Figure 5 shows how different compressed video streams can be decoded on a given computer platform, without resorting to any hardware accelerators or GPU`s. The horizontal scale represents Frames-per-Second, and it is obvious that the maturity in the implementation of MPEG-2 50Mbps makes it the fastest (or most computationally efficient) amongst all of the contenders. But XAVC Intra also performs quite well for HD, and runs just above realtime for 4K. XAVC Long GOP also marks x2 to x3 of realtime that will enable sufficient performance in the production area.

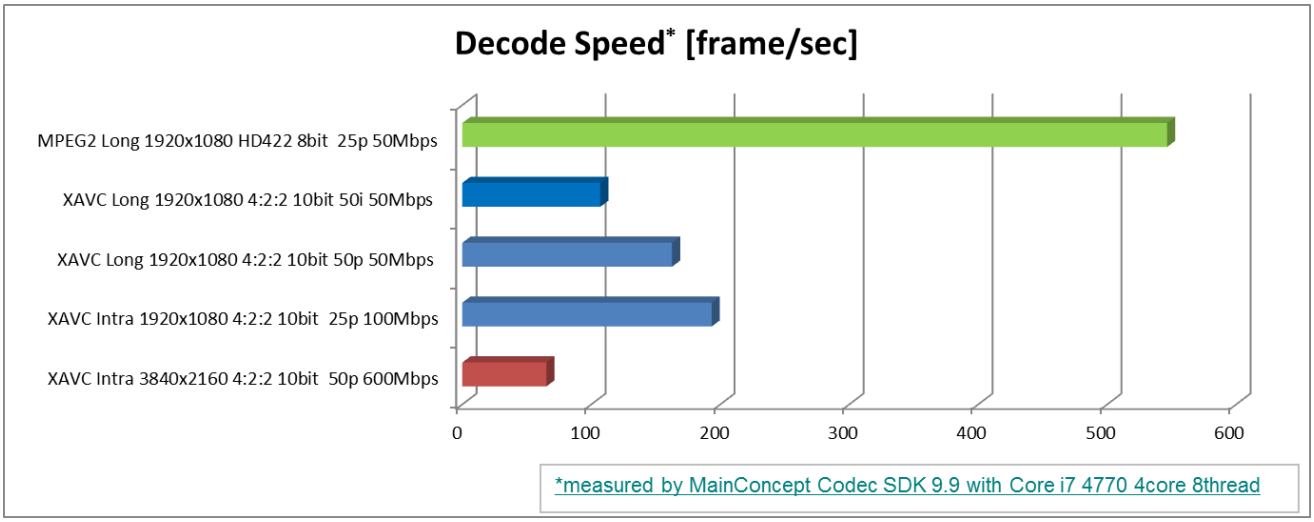



Figure 5. Software decode speed comparison Unit: Frames-per-second

6. 6. below shows the performance of software encoding process. Again, MPEG-2 is quite attractive in this area as well, but XAVC Intra marks over 60fps which is not so far behind, and XAVC Long GOP performs just around realtime for 1080i. Encoding 4K requires a little more computational power and the performance is around 17fps, which is not surprising as the bitrate is six (6) times higher than HD.

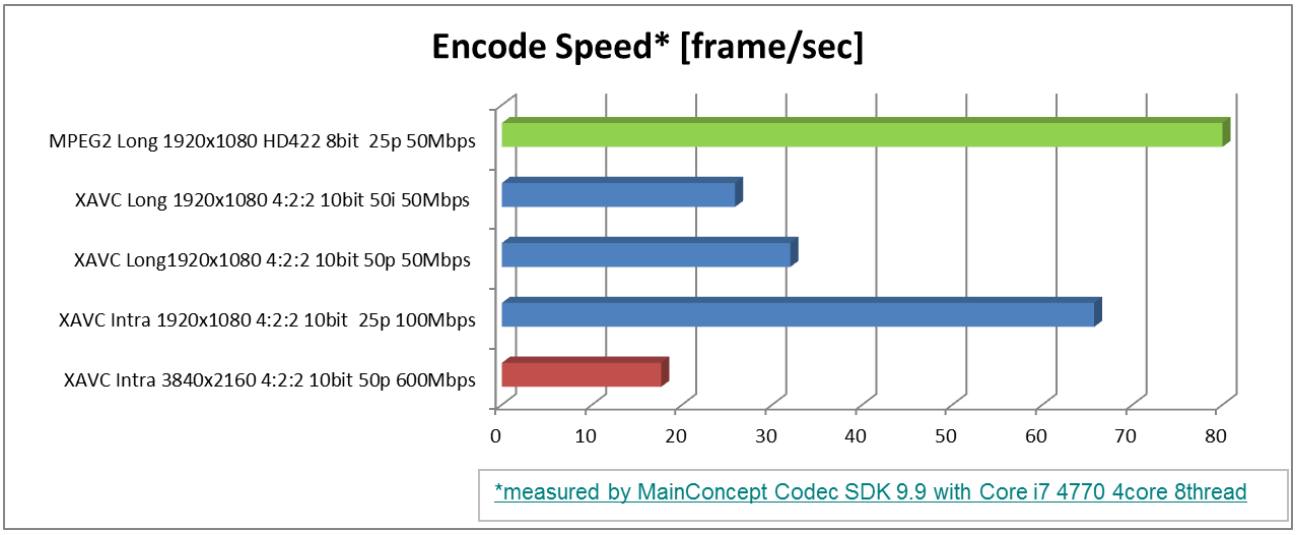



Figure 6. Software encode speed comparison Unit: Frames-per-second

Recently several broadcasters have started showing interest in adopting the H.264 format as their main in-house format due to the following reasons.

- Consolidate all program files, from prime time programming to news, into one single codec contained in a common, industry-standard wrapper.
- 10 bit sampling as opposed to 8 bit on MPEG-2.
- MPEG-2 50Mbps image quality perceived as not enough to replace current tape formats such as HDCAM.
- Storage space, network bandwidth, and processing power are becoming less of an issue to handle multiple high bit rate streams.

For 50p/60p HDTV operation, XAVC Intra supports up to 440Mbps and can be considered as the mezzanine level format that fills the gap between the mastering quality format (MPEG4 SStP, or HDCAM-SR), and MPEG2 as illustrated in **Error! Reference source not found.**.

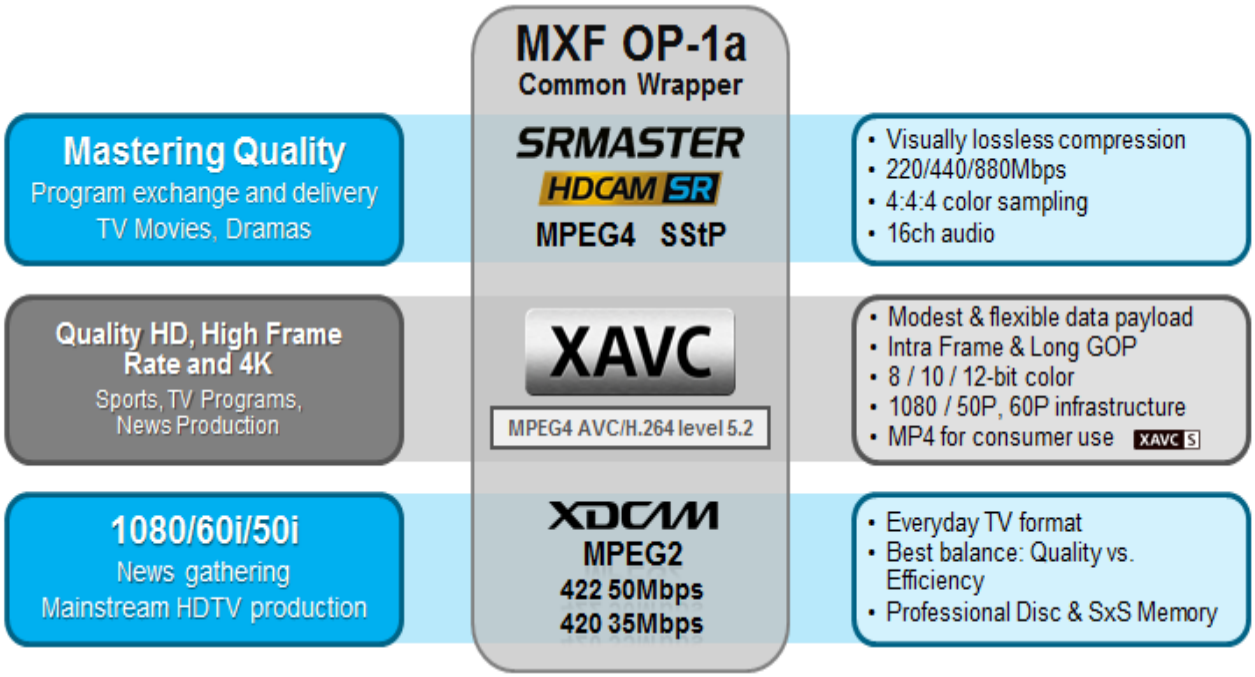



Figure 7. HDTV Compression format overview

However, demand for keeping the file size small enough to run through the current 35-50Mbps pipelines of broadcast stations and production facilities still remains strong, even with 1080-50p/60p operational workflows, - which is a growing trend as the next step in the transition to UHDTV.

Utilizing XAVC Long GOP is the preferable technique to resolve this difficult technical requirement, which makes file size smaller without affecting the picture quality. The high compression efficiency of XAVC permits to achieve the high levels of picture quality for the new progressive formats while maintaining low bit-rate payloads for significant savings in storage requirements.

Figure  shows the cost required for 100H of archiving content on Optical Disc Archive cartridge.  As obvious, Long GOP reduces the cost significantly.

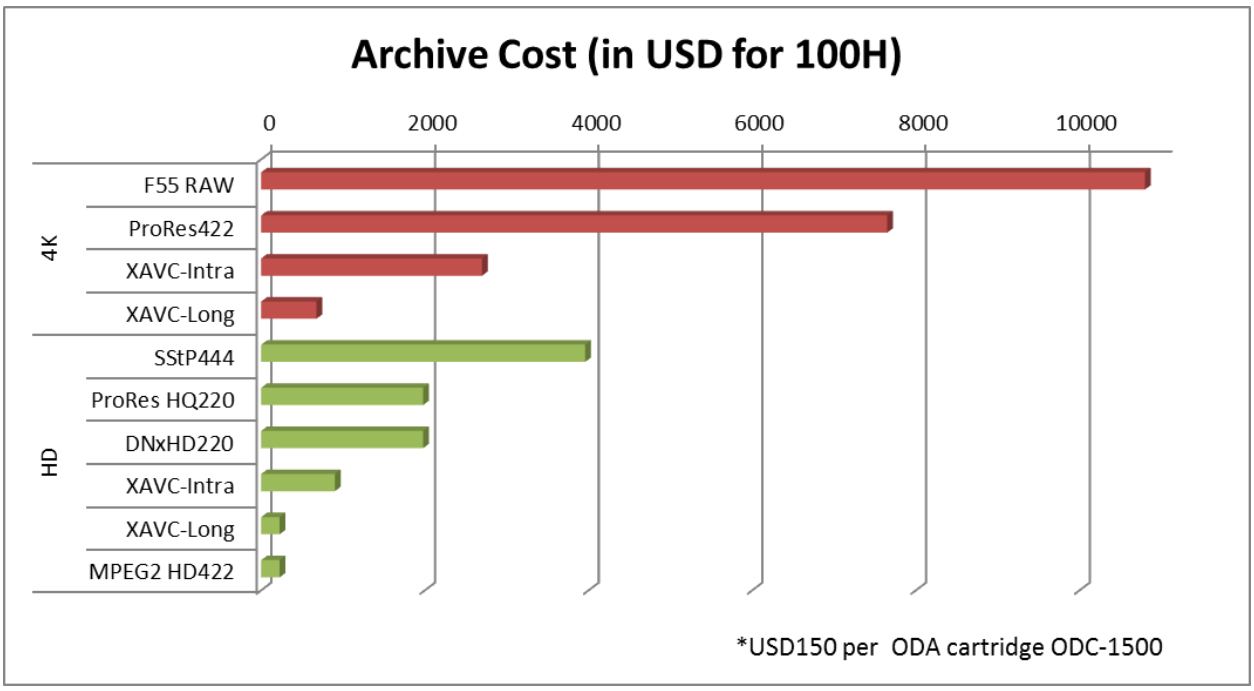



Figure 8. Archive cost

As shown in Figure 9 below, the coding performance of XAVC Long GOP at 35 Mbps is almost indistinguishable to today's 50 Mbps MPEG-2, in addition to employing 10 bit depth per color pixel. XAVC Long GOP 25 Mbps is also close to MPEG-2 50Mbps in picture quality (due to the 4:2:2 color coding of XAVC), even at half the data rate, although with more computational efforts required for coding operations.

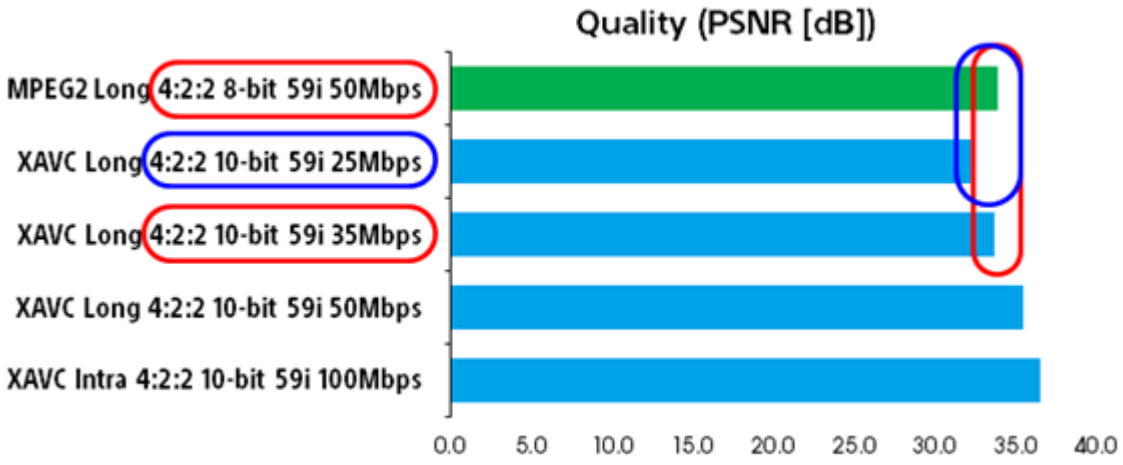
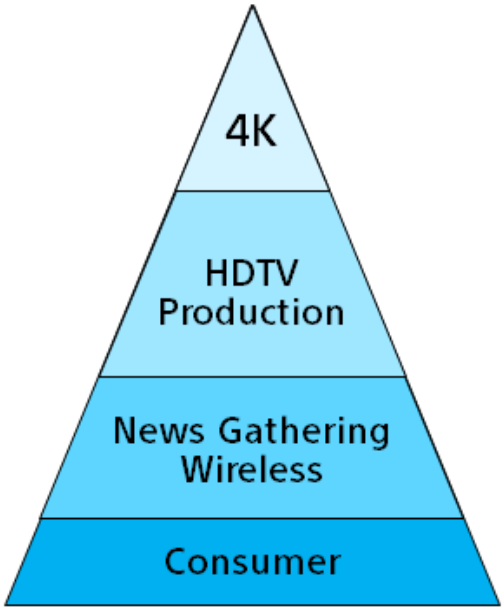



Figure 9. XAVC Higher Coding Efficiency: high picture quality at low data rate

Sony has developed a comprehensive list of products for broadcast and high-end production applications employing the XAVC compression system. The preferred range of compressed bit rates and coding structures implemented by XAVC for these applications is shown in figure 10.




| | Format | Intra Frame | Long GOP |
|---|---|---|---|
| 4K / UHD | 4096 x 2160<br>3840 x 2160<br>50, 59.94 Progressive | 600Mbps | 150Mbps |
| HD 1080p | 1920 x 1080<br>50, 59.94 Progressive | 200Mbps | 50Mbps |
| HD 1080i | 1920 x 1080<br>50, 59.94 Interlace | 100Mbps | 25, 35, 50Mbps |
| Proxy | 480 x 270<br>to 1280 x 720 | N/A | 384 kbps<br>to 9Mbps |

Figure 10. XAVC Coding parameters for various applications

XAVC also plays a critical role in enabling a portable camcorder with extreme high-frame rate capability in HDTV resolution. The PMW-F55 records full resolution 1920x1080 10bit 4:2:2 images up to 180 Frames-per-second onto the internal SxS Pro+ memory cards. By filling the two memory card slots with 128GB cards, continuous recoding time is extended to approx. 40 minutes at 180Fps.

## 4) XAVC and 4K

The use of 4K digital motion imaging systems is currently restricted to the domain of digital cinema, where the DCI (Digital Cinema Initiative) theatrical presentation standard is set to 4096 x 2160 @24Fps with compressed JPEG2000 MXF files. As history repeats itself, multiple consumer display manufacturers have started marketing 4K flat panels and projection systems for home entertainment. Although it may take a few more years for terrestrial 4K broadcasting to begin operation, various cable/satellite/network

operators and OTT internet content providers consider 4K as a new business opportunity for content delivery and have begun the distribution of multiple programming services. Current HDTV content can benefit from 4K displays, as the internal up-conversion process will help reduce line and pixel structure effects in large displays or multiple HDTV streams can be displayed simultaneously at full resolution. Today, most digital still cameras, even those built into cell phones boast a native resolution above HDTV. The 4K home panel can serve as a pristine quality digital photo frame.

Due to the wide array of operating points that the XAVC format offers, 4K imaging bandwidth can be reduced to under 100Mbps depending on GOP structure, frame rate, and color sampling. Such efficient selection of operating points is expected to significantly augment the home entertainment experience and applied to certain B2B applications where high resolution imaging is of prime interest.

The active pixel count of most 4K home displays is restricted to 3840 x 2160, a quadruple of 1920x1080. (Quad HD or QFHD), which is different from the Cinema presentation standard that has 4096 pixels across the image plane. The XAVC format covers both 4096 & 3840 horizontal sampling formats, allowing the XAVC production tools to be used in both cinema and television applications.

Sony's PMW-F55 camera records 4K XAVC Intra frames at operating points between 240Mbps (@24P) to 600Mbps (@60P) within the camera. Sony's PWS-4400 production server supports these operating points as well, to be suitable for 4K Live production platform. At much lower data rates, for example, the Sony's PXW-X70 Palm Camcorder can record and playback QFHD signals with 4:2:0 encoding and up to 30 fps at 60 Mbps with the use of XAVC Long GOP compression techniques.

Bitrate for 4K was the major discussion point throughout the development. 100Mbps for 1080i is regarded as a reasonable bitrate for intra-frame video compression and has been used in the broadcast industry. Theoretically, to satisfy the required bit rate for 4K 60p, four times in resolution and double the frame rate would require eight times of

data. This sums up to 800Mbps, nevertheless, 25% of the data can be reduced based on encoding efficiency for progressive frame compared to interlace.

Figure 11 PMW-F55
CineAlta 4K Camera

Figure 12 PWS-4400
4K/HD multi-port AV storage unit

In order to securely record such high data rate by a cost efficient yet compact media, Sony have developed the SxS Pro+ memory card family. The SxS Pro+ memory cards are compatible with all devices that have an SxS card slot, and achieve a sustained recording data rate up to 1.3Gbps. On a single SxS Pro+ 128GB memory card, the PMW-F55 records up to 50 minutes in 4K/24P or approx. 20 minutes in 4K/60P.

## 5) XAVC, RAW, and ACES workflow

The digital cinema production community has standardized a common set of image parameters that encompasses images created on film, digital cameras, and those generated by computers. The ACES (Academy Color Encoding System) standard is expected to establish a common playing field for images of different origins, offer maximum headroom for image manipulation (color grading), and achieve a consistent appearance across different tool sets and service providers. Sony is an active participant in the ACES initiative, and has created various IDTs (Input Device Transforms) to allow high-end cameras to fit into the ACES workflow. The 16bit half-float bit depth of ACES files achieves the best from high-end camera images and CGI elements, and offers maximum flexibility for color grading in the DI (Digital Intermediate) suite. This is why recent cameras from Sony (F65+SR-R4, PMW-F55+AXS-R5, PMW-

F5+AXS-R5) have the ability to record 16 bit linear RAW files via compact on-board recorders.

While the PMW-F5/F55 cameras are recording camera RAW files via the AXS-R5 on-board recorder, the in-camera SxS card slots can simultaneously record XAVC files at HD resolution, which perfectly matches the RAW files in respect to recording In/Out points, audio, timecode, and other metadata including file names. The XAVC files complement the RAW files as editorial files, which allow the editing process to start immediately when the SxS cards are removed from of the cameras.

In order to color grade and edit images derived from camera RAW files, an extra image processing step (typically referred as De-Bayering, or De-mosaicking) is required. Although the camera RAW files offer the maximum creative freedom which is imperative for sophisticated post production work, there can be time/budget constraints. 4K XAVC files can be considered as a cost efficient, high-quality, alternative to camera RAW files. As shown on Figure 13 the 4K XAVC file size is similar to those of HD resolution files that are commonly used today. It is expected that 4K XAVC files will be one of the major driving forces in expanding 4K production, especially for today's HDTV program consumption, which are to be archived in a master 4K resolution for future repurposing of the program.

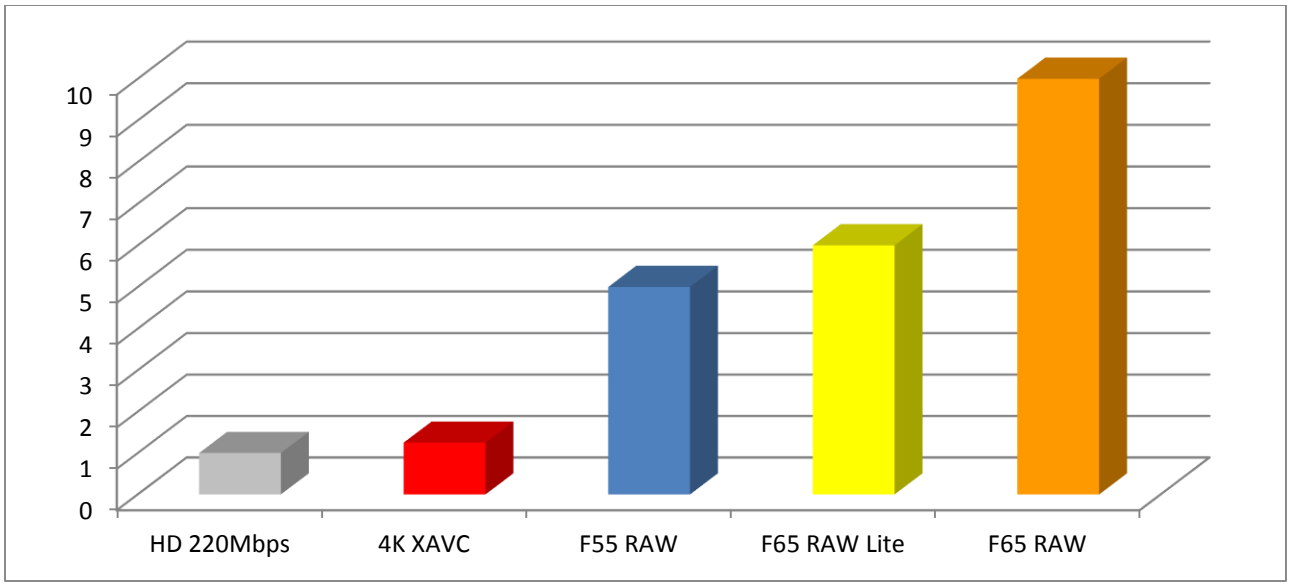



Figure 13. Image file size comparison: Unit Giga Byte/Frame

## 6) XAVC enabling technology

To make possible the use of XAVC compression at its highest quality levels in portable camcorders with modest power consumption, Sony has developed a custom hardware LSI codec that handles the XAVC encoding & decoding process (see figure 14). Moreover, the LSI codec has the ability to encode/decode MPEG-2 as well. The multi codec capability is expected to significantly increase the product life value. It will allow facility/equipment owners to create a service infrastructure that can easily convert between MPEG-2 and XAVC. The first product to make full use of this multi codec capability was the PMW-F5/F55 production camcorder, which handles MPEG-4 SStP and RAW in addition to XAVC and MPEG-2. Presently, all new camcorders from Sony for professional and consumer applications make use of this advanced XAVC/MPEG-2 LSI codec.

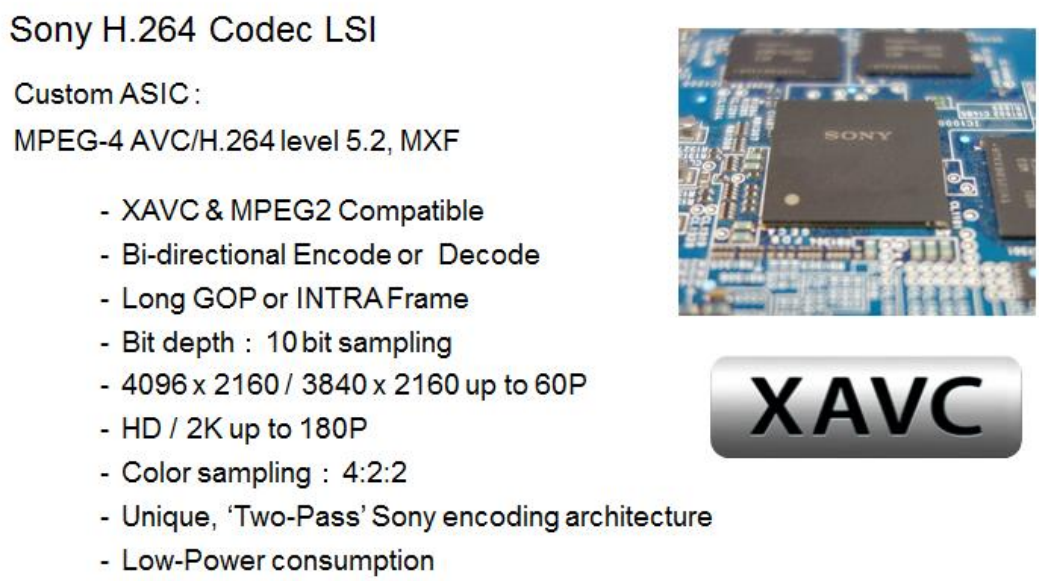



Figure 14. XAVC/MPEG-2 codec LSI

The XAVC encoder has a unique pre-coding mechanism that helps to make optimum use of the data-rate allocated for each image frame in the compressed stream. The pre-coding mechanism is integrated into both software and hardware encoders and is implemented even during high-frame-rate and 4K recordings. See figure 15 below.

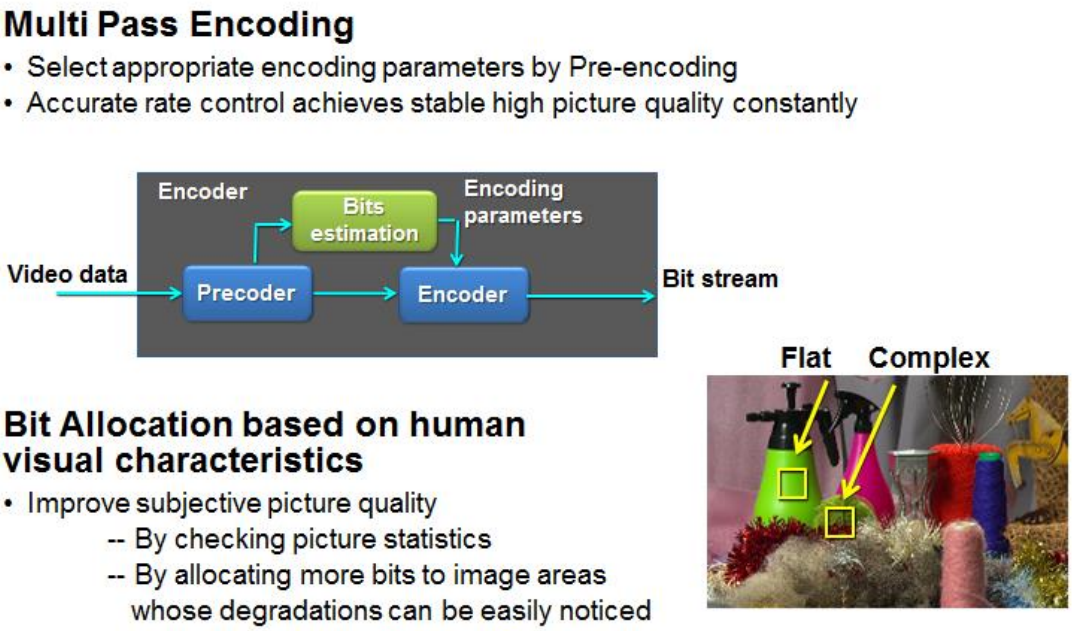



Figure 15. Pre-coding compression technology

The XAVC format fully complies with the H.264 format specifications, and Sony has strived to improve the image quality, while maintaining file interoperability with products provided by other manufacturers. Figure 16 shows the XAVC bit stream based on a traditional KLV structure. One key element on this bit stream is the frame-by-frame integration of SPS (Sequence Parameter Set) and PPS (Picture Parameter Set) into the bit stream. This allows the recording device to dynamically optimize the image quality per picture frame, and the optimized image setting value will remain attached to the picture file after editing. It also helps to maximize the image quality during random access playback.

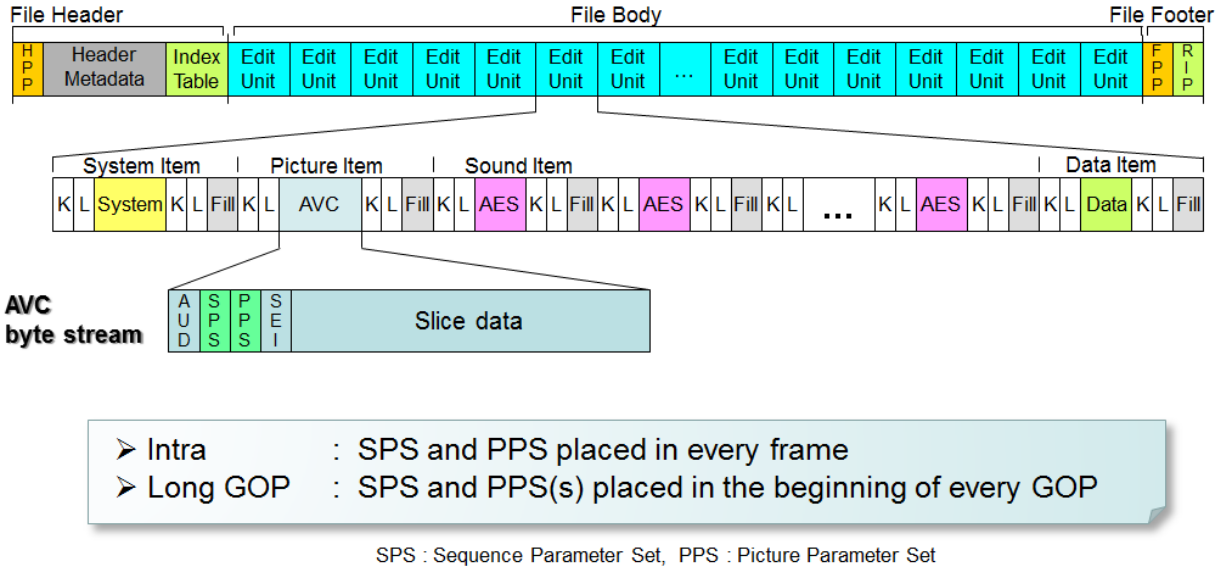



Figure 16. XAVC bit stream structure (example)

Sony has been adopting MXF OP1a from the 1st stage of file-based format in the professional domain, and has been continuously evolving it. MPEG-2 Long GOP implementation known as XDCAM or SMPTE-RDD9 has been widely adopted as HD file base format. XAVC adopts this well established and widely implemented MXF OP1a throughout the format, from HD to 4K, Intra and Long GOP. Moreover, multi-partition OP1a is being considered for edit-while-ingest type of application in the near future.

In order to efficiently cope with ever escalating production demands and yet meet the

project budget, it is extremely important to select the optimum format and operating point. The XAVC format elevates the creative possibilities to a new level, and offer quality and efficiency at the same time.

## 7) XAVC supporters

As of June 2015, 73 media companies have decided to support the use of XAVC compression in a variety of applications covering acquisition, production, post-production, archive, and streaming for both broadcasting and cinematographic environments.

Figure 17 depicts the commercial logos of the XAVC supporting companies.


Figure 17. Companies supporting XAVC Coding Technology