AVC-INTRA: A MODERN APPROACH TO VIDEO COMPRESSION

Sandro Ricci
Panasonic Australia
SMPTE '07

Introduction:

Video compression has been an extremely successful technology that has found application across many areas of television production from content acquisition through to transmission. As is always the case, the challenge faced by any compression scheme is to achieve and maintain the highest standard of picture quality possible within a given bit rate.

The large volumes of data created with today's High Definition video signals have tested traditional schemes and it is now timely that we take advantage of the many advanced and newly developed coding techniques that deliver significantly improved coding efficiencies. Panasonic's desire to continually improve HD picture quality at lower bit rates led it to research the benefits of H.264/AVC.

This new compression scheme has shown great potential for use with high quality, high definition broadcast applications. Delivering on both key objectives and offering superior performance over previous MPEG codecs, it was clear that H.264/AVC offered the best foundation for a new generation compression scheme.

The new compression technology and its related coding standards, offer developers, a broad selection of coding tools that help deliver desirable coding efficiencies. Its rapid adoption across a wide range of applications is evidence of its adaptability and broad appeal. It has been chosen for use in applications that include video telephony, mobile TV, portable video players, satellite TV, HD-DVD and the Blue-ray Disc format.

H.264/AVC which is also part of the MPEG-4 suite of coding standards, is a modern and highly efficient coding scheme offering flexibility to equipment designers. Panasonic's implementation of H.264/AVC, as used in the P2 system, is called **AVC-Intra.**

Brief History of MPEG Compression Technology:

In 1988 the Moving Picture Experts Group, began investigating an encoding system to compress moving pictures and sound. In 1992, the first system was finalized and was designated MPEG-1.

Around the same time, work began on the next generation of encoding which was later called MPEG-2. It was intended for professional broadcast users and found favourable applications with terrestrial broadcasting, DVD-Video, Cable TV, and Satellite TV systems and the consumer HDV format.

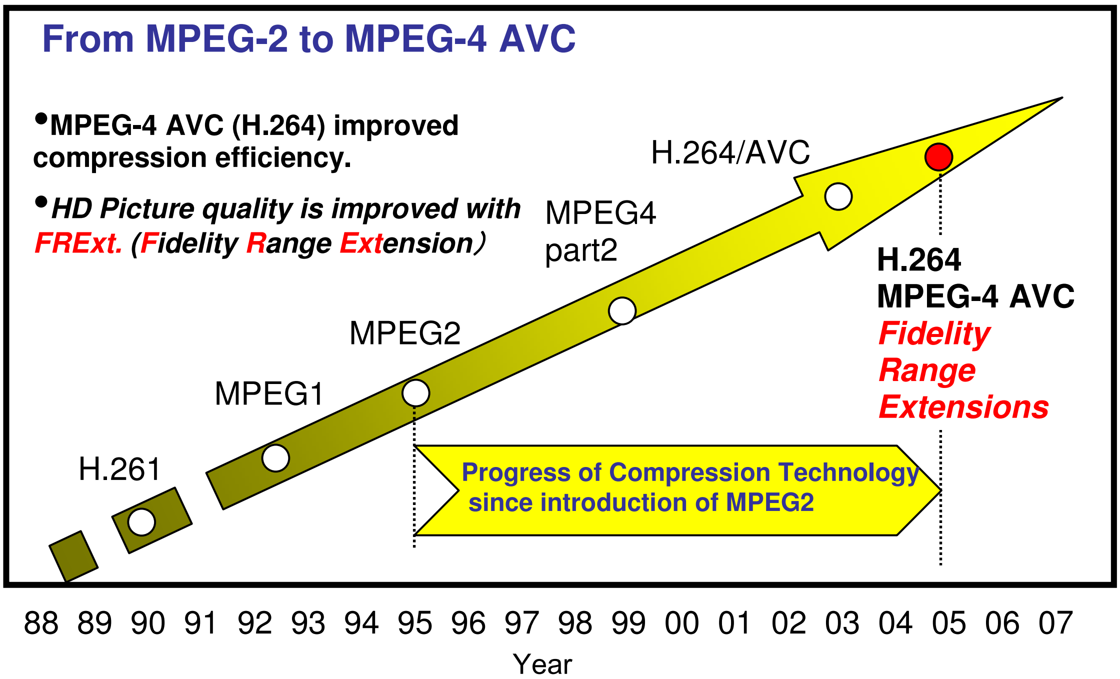


Figure 1: - Evolution of MPEG Compression

In 1999, MPEG-4 Part 2 was first published. It was intended for low bit rate applications within the multimedia field but was later enhanced to accommodate higher quality broadcast applications.

In 2001, video experts from ITU and MPEG collaborated to form the Joint Video Team (JVT). One of their aims was to develop a coding scheme that could achieve a compression efficiency gain of up to 50%. By 2003, the first implementation was finalized and approved by both groups. ITU named it H.264 while MPEG called it MPEG-4 Part 10 Advanced Video Coding (AVC). It is now more commonly referred to as H.264/AVC which offers joint recognition to both organizations.

Further development continued so that the demanding needs of high-end broadcast applications were addressed. These additions were called the Fidelity Range Extensions (FREXT) and were later integrated into the H.264/AVC standard as part of the new HIGH Profiles. Improvements included 10 bits per sample, higher video resolutions and 4:2:2 and 4:4:4 chroma sampling.

AVC-Intra - A New Coding Style:

As with previous compression schemes, a block based approach to coding is used. The picture to be coded is firstly divided into many small areas called macroblocks. With AVC-Intra, each sample within a block is spatially predicted using the neighbouring samples of previously coded blocks. The predicted residual is then transform coded, quantized and entropy coded.

The combination of various coding tools within H.264/AVC, make it possible to achieve a potential bit rate saving of up to 50% over MPEG-2 (depending on the video sequence and implementation). AVC-Intra, delivers bit rate flexibility and high quality pictures through Intra-only coding and ensures that high definition television data is recorded, edited and distributed with little penalty of quality loss.

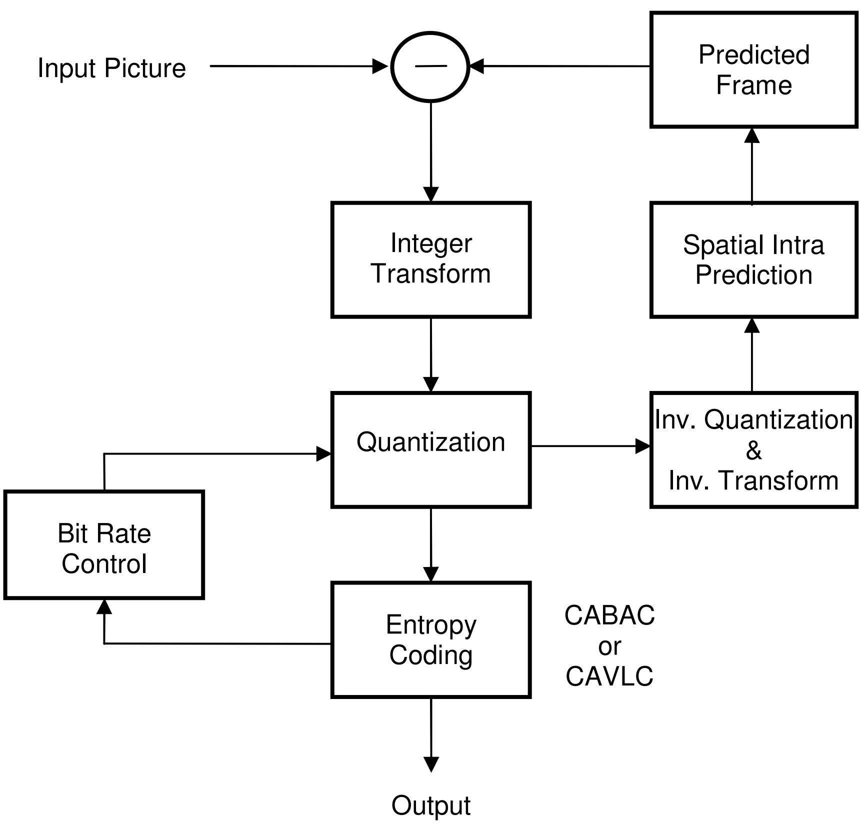


Figure 2: - Simplified view of an AVC-Intra Encoder

AVC-Intra offers users the choice of 2 operating modes:

AVC-Intra 100: High quality, full resolution HD mode

- 100Mbps
- 10 bit precision
- 4:2:2 sampling
- 1920 x 1080 pixel resolution
- 1280 x 720 pixel resolution

AVC-Intra 50: Storage Efficiency HD mode

- 50Mbps,
- 10 bit precision
- 4:2:0 sampling
- 1440 x 1080 (i.e.1920 x 1080 HD format)
- 960 x 720 (i.e.1280 x 720 HD format)

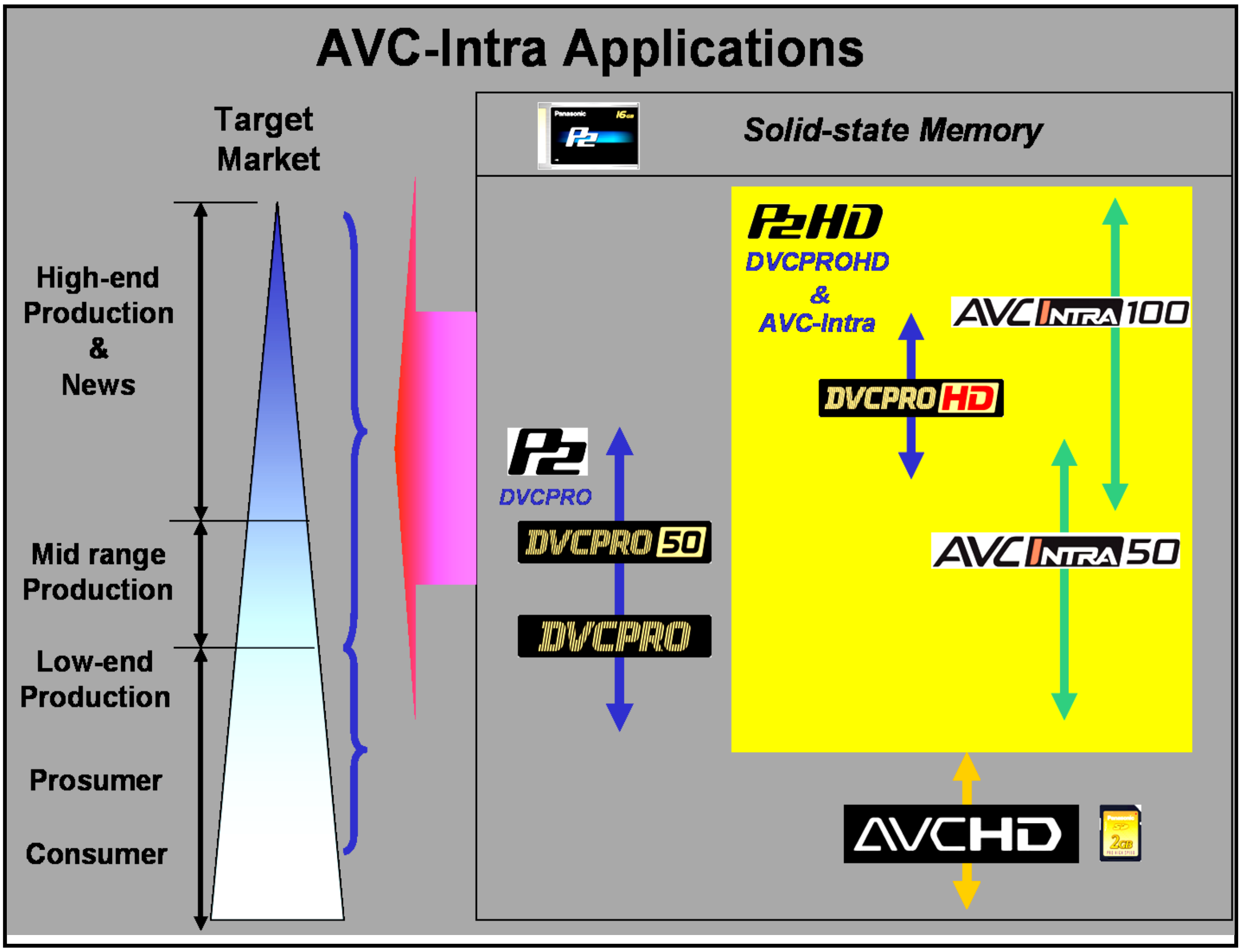


Figure 3: - AVC-Intra Applications

Benefits of Intra-only Compression:

MPEG compression schemes operate in both **Intra** (frame-only) and **Inter** (long GOP) modes. **Intra mode** compression operates totally within frame boundaries, while **Inter mode** achieves compression over multiple frames. It is understood that Inter mode compression using long GOP, is able to achieve increased coding efficiency over Intra mode schemes. Lower bit rates are achieved by taking advantage of temporal redundancy of adjacent picture frames and using Inter-Frame prediction and Bi-Directional interpolation. Under certain conditions, such as fast camera zooms and pans, still camera flash lights and strobe lights as well as other short duration production effects, the correlation of adjacent frames is severely reduced and results in visibly reduced picture quality or at worst, blocking artefacts.

AVC-Intra uses I frame compression only (Intra-only). This makes it easily editable by allowing quick random access to any individual frame. Being I frame only, it is also resilient against undesirable error propagation in the event of a disruption to the data stream. It is unaffected by the changing levels of picture complexity in adjacent frames, so achieves a more consistent level of picture quality. Intra frame coding reduces the complexity of equipment circuitry and also contributes to a lower coding latency in comparison to an Inter frame coding scheme.

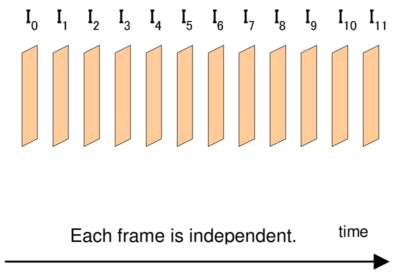
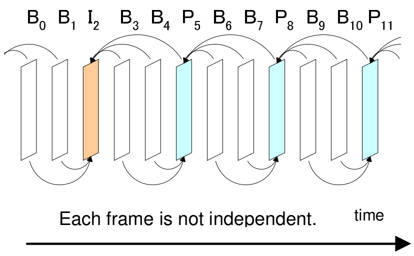
	Intra-only compression		Long GOP compression	
Compression Scheme			Multiple frames（about 15 frames） 	
Lowering bit rate	Smaller	Use spatial correlation only	Greater	Use spatial and temporal correlations
Processing delay	Smaller	1 frame	Greater	Multiple frames
Editing easiness	Greater	Intra structure	Smaller	Long GOP structure
Error spread range	Smaller	Max. 1 frame	Greater	Multiple frames
Dubbing deterioration	Easier	Max. 1 frame	More Difficult	Multiple frames

Table 1 - Comparison between Intra-only compression and Long GOP compression

Coding Efficiency Improvement:

The demanding computational requirements of H.264/AVC and the limitations of past processing hardware, have excluded its use up until now. Continued improvements in processing speed, computing power, reduced IC sizes and lower production costs have been motivating factors that have encouraged the development and introduction of this new coding scheme.

The overall gain in coding efficiency has come as a result of using several new techniques within H.264/AVC. Individually, each technique produces a welcomed lowering of bit rate, but collectively, the resultant bit rate is significantly reduced Some of the new coding tools implemented by AVC-Intra from the H.264/AVC standard that jointly help achieve this new coding efficiency are:

1. **Spatial Intra Prediction:** Exploits the correlation between neighbouring adjacent pixels within a Frame.

2. **New Entropy Coding:** CAVLC (Context Adaptive Variable Length Coding) and CABAC (Context Adaptive Binary Arithmetic Coding):

Spatial Intra Prediction:

Spatial Intra prediction basically predicts all samples in a block based on its surrounding adjacent samples. AVC-Intra supports three different types of Spatial Intra prediction modes to take advantage of the spatial correlation that exists within many areas of the picture. These include Intra_4x4, Intra_8x8 and Intra_16x16 modes.

When using either Intra_4x4 or Intra_8x8 mode, the 16x16 pixel macroblocks are divided into many 4x4 sub-blocks. These are then individually spatially predicted using one of nine possible prediction modes. When using the Intra_16x16 mode, the whole macroblock is predicted using one of four available prediction modes. Intra_16x16 mode is more suited to large picture areas of same or similar content.

The key to improving coding efficiency with spatial prediction is by selecting the most suitable prediction mode for each macroblock. Each spatial prediction mode used, is stored along with the compressed residual image in the bit stream for use by the decoder.

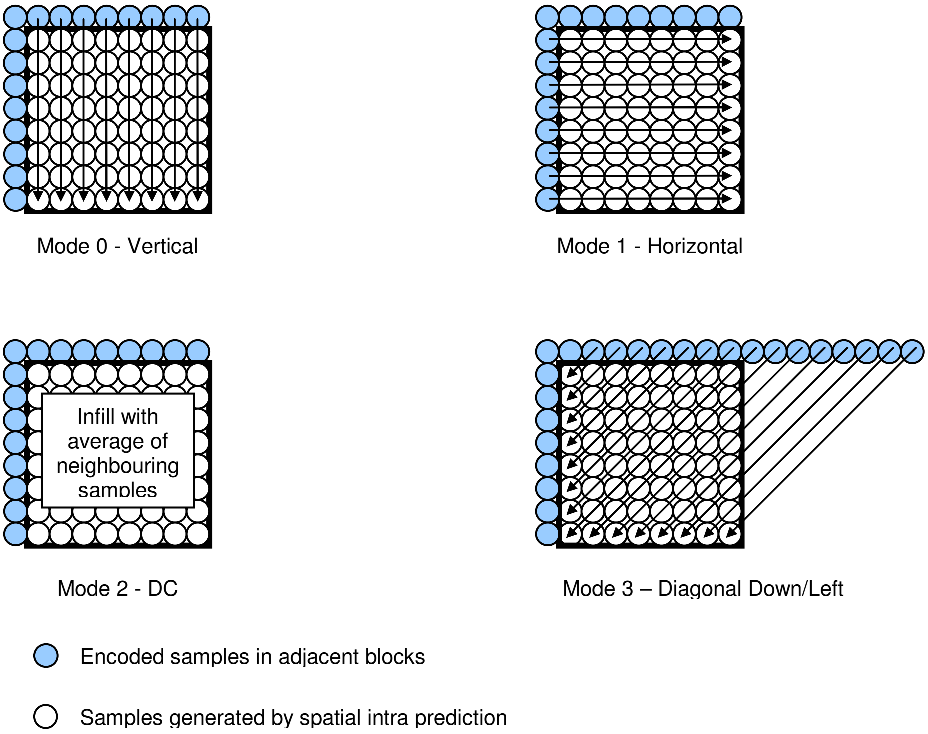


Fig. 4
Examples of Spatial Intra prediction modes

Figure 5 shows a comparison between an original encoder input image and its Intra predicted image. As can be seen, the high accuracy of spatial prediction produces an intra predicted image that is very similar to the original.

(a) Original input image (b) Intra prediction image

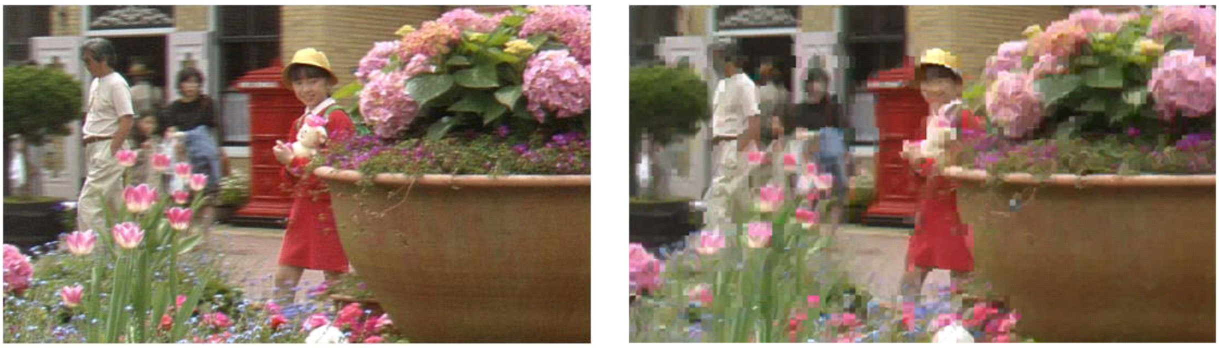


Figure 5: - Original and Intra Predicted Image

The Intra Frame predicted image is subtracted from the original, leaving only a prediction error or “residual”. Transform coding is then applied to this residual image in the form of an Integer Transform which is conceptually similar to DCT. The role of this transform coding is to reduce the spatial redundancy of the residual image and code it into a form that allows easier quantization by concentrating the energy into low frequency coefficients. The resulting coefficients are then transform coded, adaptively quantized and then entropy coded.

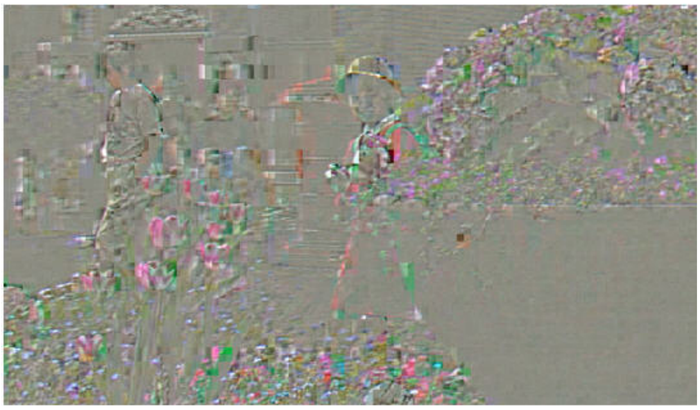


Figure 6: - Residual image

The highly accurate intra prediction process is able to reduce the amount of data required in the residual image and results in a more efficient compression even when using an Intra-only coding scheme. In contrast to Inter-frame prediction, spatial intra prediction is performed within the confines of a single frame, which means that the prediction accuracy and subsequent compression efficiency does not deteriorate with motion as it would with long GOP schemes.

New Advanced Entropy Encoding:

Entropy encoding is a lossless compression scheme. Entropy defines the statistical probability of data symbols, that is, the higher the probability, the lower the entropy. Efficient compression of data occurs when its entropy is low. Variable Length Coding (VLC) is an example of Entropy encoding and has been successfully used in previous compression schemes.

This form of coding is based on a fixed table of variable length codes, but it does not adapt to the actual symbol statistics present in the transform coefficients. These statistics can and do vary over time and can depend largely on the video content.

The VLC process produces short code words for frequently occurring symbols and longer code words for those that are less frequent. The output data rate and compression efficiency after VLC is not optimal and depends on scene content and complexity. Using a scheme that can adapt its coding in relation to the changing symbol statistics can offer a significant improvement in coding efficiency.

The Entropy encoding choices for H.264/AVC can be selected from either CABAC (Context Adaptive Binary Arithmetic Coding) or CAVLC (Context Adaptive Variable Length Coding). Both modes are able to continually adjust their entropy encoding parameters to better suit the content of the image.

The probability of a symbol is based on the "current context", i.e. symbols that precede it at the time of coding. This allows prediction modeling to change, which makes the coding "adaptive". MPEG-2 does not support this adaptive function and performs entropy encoding as a simple VLC, i.e. by means of a fixed code table. CAVLC can adaptively switch between many different VLC tables, compared to the fixed table use, of past schemes. The choice of each table is determined by the statistical probability of past symbols.

Table 2 shows the comparison of the various methods of entropy encoding.

	MPEG-2 VLC	**H.264/AVC**	
		CAVLC	**CABAC**
Encoding method	Non-adaptive VLC	Context Adaptive VLC	Context Adaptive Arithmetic Coding
Context Adjustment process	No	Transformed coefficient based	bit by bit
Encoding efficiency	Moderate	Very good	Excellent

Table 2: Comparison of Entropy Encoding Methods.

CABAC, on the other hand, performs its function on words instead of their symbol components. Firstly context modeling is performed and an appropriate data probability model is chosen based on past analysis of the data symbols. Binary Arithmetic Coding is then performed in a final efficient step. Whole words are coded using the selected data model and are then represented as a single non-integer number i.e. binary fraction. The resultant number of bits required to represent the words are less than that required if encoding were done at symbol level.

Even though both modes perform context adaptive entropy encoding, it is CABAC that delivers the most significant gains in efficiency through its use of binary arithmetic coding. Using collected data from numerous test sequences, CABAC has been shown to reduce the bit rate by 10-20% more than CAVLC.

Choice of Bit Rate Delivers Flexibility:

It can be said that that one bit rate is not ideally suited to all production applications. Production requirements will influence the choice of bit rate with the chosen bit rate being a measured balance between recording times and picture quality. AVC-Intra provides a choice of two efficient bit-rates in two HD picture resolutions. Each mode produces high quality HD images, and allows users to choose a bit rate and picture format that delivers the desired quality required for the production application.

HD Format	Bit Rate	Bit Depth and Sampling	Resolution
720p	**AVC-Intra 50Mbps**	10 bit 4:2:0	960x720
	AVC-Intra 100Mbps	10 bit 4:2:2	1280x720
1080i	**AVC-Intra 50Mbps**	10 bit 4:2:0	1440x1080
	AVC-Intra 100Mbps	10 bit 4:2:2	1920x1080

Table 3: AVC-Intra Bit Rate and HD Resolution choices.

MXF File format and AVC-Intra:

The Material Exchange Format (MXF) was developed to allow easy interchange of video and audio files as well as metadata. The design of the MXF file format allows it to be applied and used in many different stages of the production cycle such as Acquisition, Editing and Archiving.

Panasonic chose to use the Operational Pattern Atom as the essence MXF file format for the P2 System. Each clip consists of separate video and audio MXF files that include a Generic Container that has been designed to transport many different forms of data.

The video data can be either uncompressed or compressed and includes the use of MPEG bit streams as specified in SMPTE 381M. The AVC-Intra compressed bit stream can therefore be mapped into the MXF container as an MPEG data file.

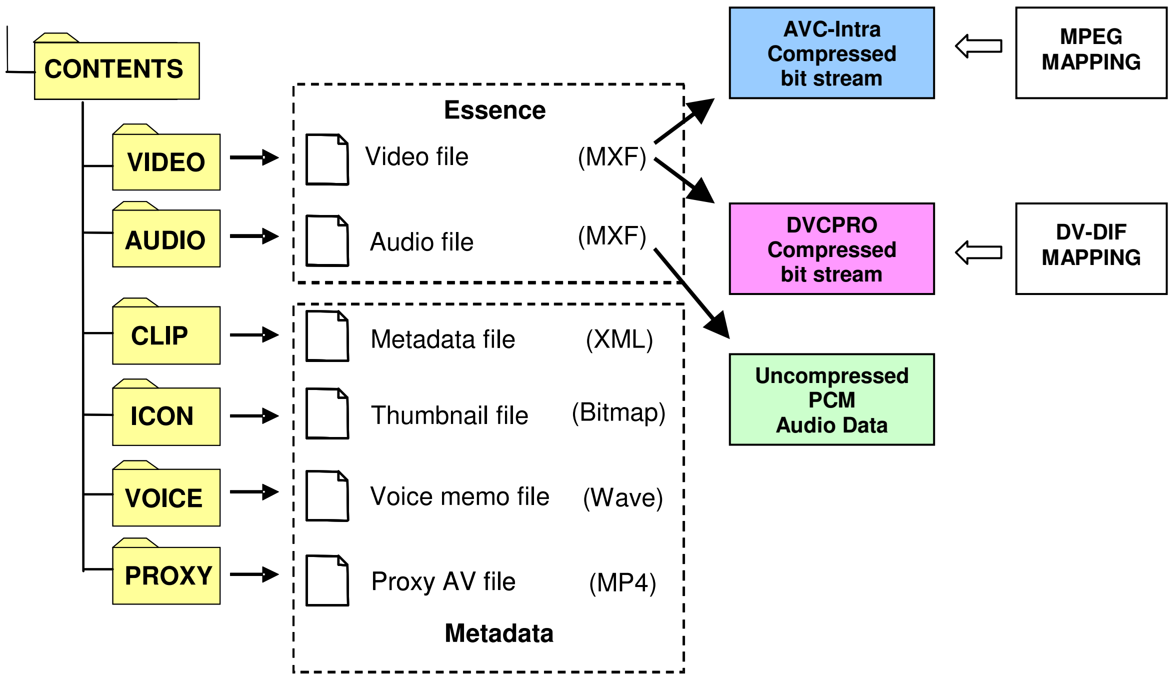


Figure 7: - P2 Folder and Content Structure

Figure 8 is a simple illustration of how the AVC-Intra compressed bit stream is encapsulated in the MXF OP Atom file structure making its integration into the P2 family a straightforward task.

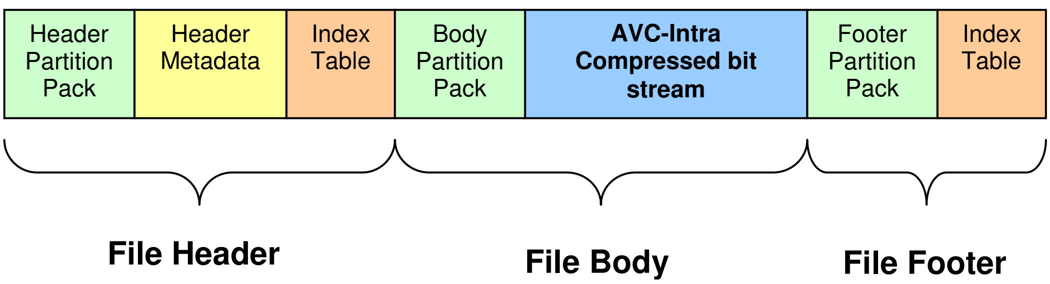


Figure 8: - MXF OP-Atom File Structure (simplified view)

Conclusion:

The versatility and adaptability of H.264/AVC has seen it successfully implemented across a broad range of product applications. This widespread adoption is mainly due to the many features and coding tools available within the standard and has allowed equipment designers to implement it at either low or high bit rates and across a wide range of video resolutions and sampling structures.

Panasonic has leveraged the IT benefits provided by its successful P2 Card solid state media system to allow a seamless integration of AVC-Intra files. Along with DVCPRO, end users now have the flexibility of choice to select a compression scheme that best meets their quality and storage cost needs.

Panasonic's implementation of H.264/AVC for the P2 system has for the first time, realized, full resolution, high definition picture acquisition using 10 bit precision sampling in a camcorder.

The undeniable advantages of the AVC-Intra system and its outstanding HD picture quality make it an obvious and ideal choice for inclusion into Panasonic's current and future P2-HD professional video products.

Sandro Ricci
Panasonic Australia
July 2007