### Author(s)

<table>
<thead>
<tr>
<th>First Name</th>
<th>Middle Name</th>
<th>Surname</th>
<th>Role</th>
<th>Email</th>
<th>SMPTE Member?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pierre</td>
<td>Larbier</td>
<td>CTO</td>
<td></td>
<td><a href="mailto:p.larbier@ateme.com">p.larbier@ateme.com</a></td>
<td>yes</td>
</tr>
</tbody>
</table>

### Affiliation

<table>
<thead>
<tr>
<th>Organization</th>
<th>Address</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATEME</td>
<td>26 Burospace, 91570 Bièvres</td>
<td>France</td>
</tr>
</tbody>
</table>

### Author(s) – repeat Author and Affiliation boxes as needed--

<table>
<thead>
<tr>
<th>First Name</th>
<th>Middle Name</th>
<th>Surname</th>
<th>Role</th>
<th>Email</th>
<th>SMPTE Member?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(or initial)</td>
<td>(or initial)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Affiliation

<table>
<thead>
<tr>
<th>Organization</th>
<th>Address</th>
<th>Country</th>
</tr>
</thead>
<tbody>
<tr>
<td>(repeat Author and Affiliation boxes for each author)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Publication Information

<table>
<thead>
<tr>
<th>Pub ID</th>
<th>Pub Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>2011 Fall Conference Paper</td>
<td></td>
</tr>
</tbody>
</table>
AVC-I: Yet another intra codec for broadcast contribution?

Pierre Larbier, CTO
ATEME, Bièvres France, p.larbier@ateme.com

Written for presentation at the
[2011 SMPTE Annual Technical Conference]

Abstract. Large availability of fiber and high-speed IP networks now enable the use of intra video codecs in a wide range of contribution application. Such formats permit low end-to-end delays and are easily editable. In addition, their reduced complexity allows a significant equipment cost reduction over Long GOP equivalents. These well-known advantages have paved the way for a variety of proprietary and standardized solutions like VC-2, VC-3 or JPEG-2K.

AVC-I is simply an MPEG-4 AVC / H.264 profile obtained by capping complexity and removing encoding tools from the standard's broadcast contribution profiles. This simple conception is the promise of a straight-forward compatibility between equipments from various vendors. But it remains to be seen, beyond academic research, if actual AVC-I products are able to meet and exceed the quality levels and the operational advantages of intra codecs currently in use.

Keywords. AVC-I, H.264, AVC, video, compression, intra
Introduction

Broadcast contribution applications like news gathering, event broadcasting or content exchange currently benefit from the large availability of high-speed networks. These high bandwidth links open the way to a higher video quality and distinctive operational requirements such as lower end-to-end delays or the capability to store the content for further edition.

Since a lighter video compression is needed, the complexity of common Long-GOP codecs can be avoided and simpler methods like intra-only compression can be considered. These techniques offer the capability to compress pictures independently, which is highly desirable when low latency and error robustness are of major importance. Several intra-only codecs like JPEG-2000 or MPEG-2 Intra are today available but they might not meet all Broadcaster needs.

AVC-I, which is simply an Intra-only version of H.264/AVC compression, offers a significant bitrate reduction over MPEG-2 Intra while keeping the same advantages in terms of interoperability. AVC-I was standardized in 2005 but broadcast contribution products supporting it were launched only in 2011. Therefore, it may be seen as a brand new technology and studies have to be performed to evaluate if they match currently available technologies in operational use cases.

In order to verify AVC-I performance in actual contribution applications, we conducted a study regarding broadcasters two main criteria: End-to-end latency and video quality. The results presented in this paper were obtained using recently launched products, and using a wide range of bitrates in order to illustrate the advantages of AVC-I within a great variety of contribution use-cases.

Why intra compression?

Video compression uses spatial and temporal redundancies to reduce the bitrate needed to transmit or store video content. These two techniques apply a similar idea: A part of the video sequence is seen as the composition of a predictor, computed from an already decoded portion of the video sequence, and a residual part that has to be compressed. When exploiting temporal redundancies, the predictor is derived from blocks found in adjacent pictures, while spatial predictors are built with pixels found in the same picture. Long-GOP compression makes use of both methods and intra-only compression is restricted to spatial prediction.

Long-GOP approaches are more efficient than Intra-only compression, but they have also distinctive disadvantages:

- Decoding a picture requires having previously decoded several pictures and these picture dependencies may be complex to handle when seeking in a file. This makes editing a Long-GOP file a complex task.

- Any information loss during transmission or decoding error might spread from a picture to the following ones and span a full GOP. This means that a single transmission error can affect decoding for several hundred milliseconds of video and therefore be very noticeable.

- Encoding and decoding delay might be increased using Long-GOP techniques mainly because the compression tools are much more complex. The net consequence is a higher end-to-end latency compared to intra-only compression.

Another problem inherent to Long-GOP compression relates to video quality which varies significantly from picture to picture. This can introduce difficulties when editing a file as it contains “good pictures” and “bad pictures”. For example, Figure 1 depicts the PSNR along the
sequence *ParkJoy* when encoding it in Long-GOP and in Intra-only. While the quality of the Long-GOP pictures is always higher than the one of their Intra-only counterparts, it varies considerably. On the other hand, the quality of consecutive Intra-only coded pictures is much more stable.

![AVC Long GOP vs AVC-I](image)

**Figure 1.** Long-GOP vs Intra-only compression.

For these reasons, Intra-only compression might be a better choice than Long-GOP when:

- Enough bitrate is available
- Low end-to-end latency is a decisive requirement
- Streams have to be edited
- Application is sensitive to transmission errors

Several intra-only codecs are currently available to broadcasters to serve the needs of contribution applications:

- **MPEG-2 Intra**: This version of MPEG-2 [1] compression is restricted to the use I-frames, removing P-frames and B-frames. Its coding efficiency is just slightly better than JPEG. But due to its simplicity and the large availability of interoperable MPEG-2 equipments, it has been widely deployed for more than 15 years.

- **JPEG-2000**: This codec [2] is a significantly more efficient successor to JPEG that was standardized in 2000. The availability of low-cost compression chips enabled the launch of JPEG-2000 products targeting contribution applications in 2006. But while the video codec is well defined, the transport of compressed streams has still to be clarified. Therefore, JPEG-2000 equipments are not interoperable.

- **VC-2**: Also known as Dirac-Pro, this codec has been designed by BBC Research and was standardized by SMPTE in 2009 [3]. Like JPEG-2000 it uses wavelet compression. Some compression products are available but up to this day, VC-2 has not been widely deployed for contribution applications.

Older codecs like MPEG-2 Intra benefit from a large base of interoperable equipments but lack coding efficiency. On the other hand, more recent formats like JPEG-2000 are more efficient but are not interoperable. Consequently, there is a need for a codec that could be at the same time efficient and that would ensure interoperability between equipments from various vendors.
What is AVC-I?

AVC-I designates a fully compliant variant of the H.264/AVC video codec restricted to the intra toolset [4]. In other words, it is just plain H.264/AVC using only I-frames. But some form of uniformity is needed in order to ensure interoperability between equipments provided by various vendors. Therefore ISO/ITU introduced a precise definition in the form of profiles (compression toolsets) in the H.264/AVC standard.

### H.264/AVC profiles

While the first edition of the H.264/AVC standard was targeting low-bitrate video delivery, subsequent revisions added the support for high quality/high bitrate professional applications. In this regard, provision to using only I-frame coding was introduced in the second edition of the H.264/AVC standard with the inclusion of four specific profiles\(^1\): **High10 Intra profile**, **High 4:2:2 Intra profile**, **High 4:4:4 Intra profile** and **CAVLC 4:4:4 Intra profile**. They can be described as simple sets of constraints over profiles dedicated to professional applications. Table 1 gives an overview of the main limitations introduced to define those four intra profiles.

<table>
<thead>
<tr>
<th>H.264/AVC Intra profiles</th>
<th>Based on:</th>
<th>Summary of the restrictions to the base profile</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High10 Intra</strong></td>
<td><strong>High 4:2:2 profile</strong> (targets mainly Contribution applications with up to 4:2:2 10-bit pixels)</td>
<td>• All pictures are IDR (no P or B pictures) &lt;br&gt; • Limited to 4:2:0 chroma format (no 4:2:2 chroma format)</td>
</tr>
<tr>
<td><strong>High 4:2:2 Intra</strong></td>
<td><strong>High 4:4:4 Predictive profile</strong> (targets mainly Archiving applications with up to 4:4:4 14-bit pixels)</td>
<td>• All pictures are IDR (no P or B pictures)</td>
</tr>
<tr>
<td><strong>High 4:4:4 Intra</strong></td>
<td><strong>High 4:4:4 Predictive profile</strong> (targets mainly Archiving applications with up to 4:4:4 14-bit pixels)</td>
<td>• All pictures are IDR (no P or B pictures) &lt;br&gt; • Only CAVLC entropy coding</td>
</tr>
<tr>
<td><strong>CAVLC 4:4:4 Intra</strong></td>
<td><strong>High 4:4:4 Predictive profile</strong> (targets mainly Archiving applications with up to 4:4:4 14-bit pixels)</td>
<td>• All pictures are IDR (no P or B pictures)</td>
</tr>
</tbody>
</table>

Table 1. H.264/AVC Intra profiles.

Since the intra profiles are defined as reduced subsets over commonly used H.264/AVC profiles, they don’t introduce new features, technologies or even stream syntax. Therefore, AVC-I video streams can be used within systems that already support plain H.264/AVC video streams. This enables the usage of file containers like MPEG files or MXF, transport like MPEG-2 Transport Stream or RTP, audio codecs like MPEG Audio or Dolby Digital, and many metadata standards.

Another advantage of its very simple design is that AVC-I streams can be encoded, transported and decoded by most, if not all, equipments or software that already support contribution profiles like **High 4:2:2 profile**. In addition, the compatibility with a much broader range of H.264/AVC products can be obtained by configuring the encoder to restrict the compression toolset, for instance using only a 4:2:0 8-bit format. Hence interoperability between systems provided by various vendors is easy to achieve. And products not designed specifically for intra-only application could be compatible from the start.

\(^1\) An intra-only profile was also added to the Scalable Video Coding (SVC) extension of H.264/AVC: **Scalable High Intra profile**
The latency issue

Latency is defined as the end-to-end delay between the acquisition of the video signal on the encoder input to its restitution on the decoder output, not considering transmission time. This parameter is of great importance to Broadcasters because a low delay is a prerequisite to major contribution applications like interviews.

The world first contribution pair supporting AVC-I is ATEME CM4101/DR8400. This codec was added as a configuration preset on top of the existing H.264/AVC High422 profile, without deep changes to the internal architecture. This was done to guarantee interoperability and maximize video quality. But the drawback is that system latency is not specifically optimized. It can be found around 250ms which is about the latency measured with JPEG-2000 encoder/decoder pairs when they were first introduced.

But much better delays can be achieved on systems designed explicitly for AVC-I. We will start by examining the theoretical lower bound of the latency that could be attained by an interoperable and conformant AVC-I system. We will then detail the various stages of an actual contribution system to propose a realistic figure.

**Theoretical lower bound**

The first factor that impacts latency is the time required to produce compressed samples following their acquisition. This delay depends on the inner algorithm used by the video codec:

- Block-based compression, as used in AVC-I, is performed while traversing the picture in scan order. Therefore, the first compressed bits may be produced as soon as the first block of the picture has been acquired, which is achieved after receiving 16 lines of the source picture. If all other delays were negligible, the absolute minimum latency would then be around 500µs.

- Scalable compression, as used in JPEG-2000, involves the decomposition of source pictures into multiple resolution layers. This process requires to wait until a full picture is acquired on the encoder input. Consequently, the minimum latency for an ideal system is equal to the picture period, around 17ms.

These two methods, depicted in Figure 2, could therefore yield to systems exhibiting very low latencies, with a clear advantage toward block-based compression as used in AVC-I.

![Figure 2. Block-based vs spatially scalable compression.](image)

---

2 JPEG-2000 allows some optimizations like tiling or using a scan-based mode, but the implementation of such techniques is extremely uncommon in Broadcast products.
But such low latencies are rarely seen in real systems because an important aspect limits the ability to reduce the latency of block-based codecs: a smooth video on the decoder’s output can be achieved only if it can decode a continuous flow of pixels. If the network between the encoder and the receiver has a limited bandwidth, then the transmission of a compressed block could take an excessive duration and prevent the decoder from having received the required data on time to process them.

To avoid this problem, the encoder has adjust the compression process so that the decoder will always have received the compressed data on time. But this requires also that the encoder has a model of the receiver processing behavior. In AVC-I, it is standardized in a model called HRD (Hypothetical Reference Decoder). A conformant encoder will use this model to guarantee that a conformant receiver is capable of decoding and displaying a smooth video.

The HRD is designed around the concept of leaky-bucket which implies that a decoder cannot start processing the stream before having received at least a complete compressed picture. This implies that a compliant AVC-I system will exhibit a minimum latency corresponding to the transmission delay of at least a compressed picture. This correspond to at least a picture period, around 17ms for HD formats.

AVC-I systems with lower latencies can be constructed but they require a different receiver model, block-based instead of picture based for instance. Such systems are not conformant to the H.264/AVC standard, which prevents system vendors to guarantee interoperability.

**Latency in actual AVC-I systems**

As shown before the minimum theoretical latency of a conformant and interoperable AVC-I system is therefore similar to scalable codecs like JPEG-2000, around a single picture period. But beyond theoretical considerations, it remains to be seen what kind of latencies could be achieved by actual AVC-I encoder-decoder pairs.

Figure 3 depicts an overview of the processing chain of an actual professional AVC-I encoder-decoder pair regarding all involved delays.

In fact most of the building blocks described in the following paragraphs are not specific to an AVC-I system. They are required to achieve simultaneously a high video quality and the conformance to standards that guarantee interoperability. It should be noticed that some currently available intra systems lack some of those blocks. This is one of the reasons why they are not interoperable.

**Time-base correction**

To ease operations, it is highly desirable that a contribution pair produces at all time a continuous video stream on the output of the receiver. Therefore, the stream has to remain continuous and compliant on the output of the encoder when the source video is unplugged, replaced or even exhibit glitches. This requires to be able to detect source losses in order to
substitute it gracefully with an alternate pattern like a color bar. Source loss detection and its replacement without losing system lock is performed by a time base corrector that is usually integrated as an encoder front-end. This module is optional but, when activated, it introduces an additional delay to the chain of up to 33ms in 1080i30, and up to 17ms in progressive modes like 720p60.

Rate-control

Even in an intra-only system with fixed-sized compressed pictures, the rate-control algorithm is a crucial part of an encoder. It is responsible for choosing the quantizers required to reach the target size and at the same time optimize the perceived video quality. This function can be performed on-the-fly during picture acquisition but has to be done before actually encoding pictures. Otherwise, severe video quality variations could be observed between the top-left and the bottom-right of the pictures. Consequently the rate-control algorithm induces an additional picture period delay, around 17ms in 30Hz systems.

Video compression

The time needed to actually compress the acquired pictures is highly implementation dependent. But in order to reach real-time operation, it has just to be lower than a picture period, or 17ms in 30Hz systems. As an example, ATEME encoders are actually able to encode HD pictures in about half this time.

Decoder buffer

As we’ve seen it before, compliancy and interoperability impose the management of a bit reservoir that adds a picture period delay. In actual operations, this buffer is jointly managed by the encoder and the decoder in close relations with the multiplexing/demultiplexing processes.

Decoding delay

Finally, the time required by the decoder to process the compressed pictures has to be added. The HRD assumes an instantaneous decoding process, which is obviously impossible to achieve. Like compression time, the actual decoding time is implementation dependent. The limiting factor is mostly related to the bitrate if CABAC is used. As an example, ATEME decoding engine is able to decode in real-time more than 150Mbps. Therefore the decoding time of fixed-size pictures encoded at this bitrate is equal to the picture period, 17ms in 30Hz systems.

End-to-end delay

Adding up all the individual delays in the encoding-decoding chain, we can conclude that the overall AVC-I system latency can be as low as 100ms in 30Hz systems. And it should be noted that this figure has been obtained without doing any compromise on conformance, video quality or assumption on the video source. Consequently, we have all the reasons to believe that dedicated AVC-I systems will reach the low latency levels that are currently obtained by other intra-only compression codecs at the expense of interoperability.

**AVC-I Video Quality**

Many academic papers have compared the coding efficiency of H.264/AVC in intra-only mode versus other intra codecs. For instance [5], [6] and [7] evaluate in great detail the relative
performance of AVC-I compared to JPEG-2000. While all these studies use the H.264/AVC reference software at various stages of the standard development, their findings are very similar: JPEG-2000 and AVC-I perform roughly equally well on most content, with a slight advantage toward AVC-I on interlaced material.

Those performance comparisons are carried out using objective metrics like PSNR or SSIM which may not reflect the visual perception. Furthermore, the simulation software configurations are far from broadcast common practices. Consequently, published studies do not necessarily reflect the visual experience of a given codec in the context of broadcast contribution.

For this reason, we have performed a visual evaluation of various intra codecs intended for broadcast contribution applications. It involved ATEME products like the H.264/AVC pair CM4101/DR8400 than can encode and decode AVC-I and MPEG-2 Intra up to 150Mbps, products from other vendors, but also reference software. This investigation was done by expert viewers on a large set of test sequences representative of high-definition broadcast contribution content, mostly interlaced.

The outcome of this evaluation is that two codecs are the most suitable for high bitrate intra uses: AVC-I and JPEG-2000\(^3\). The detail level appears to be about the same with both codecs on bitrates ranging from 50Mbps to 150Mbps. This confirmed that the coding efficiency of AVC-I and JPEG-2000 is close. However, since underlying algorithms are very different, coding artifacts are also different.

**AVC-I and JPEG-2000 artifacts**

At very-low bitrate, which is not the intended application but magnifies compression imperfections, both codecs produce unacceptable artifacts:

- JPEG-2000 blurs large areas of the pictures and the detail loss can be very significant. In addition, ringing and mosquito noise is experienced along sharp edges.
- AVC-I blurs also parts of the picture but on block basis. This effect is caused by the loop-filter that acts as a post-processing filter. If it is removed, then blocking artifacts appear.

At moderate bitrates, below 100Mbps, a problematic defect was observed similarly on both codecs: Pictures can exhibit an annoying flicker. This issue is caused by a temporal instability in the coding decisions, amplified by noise. It seems to appear below around 85Mbps with JPEG-2000 and below 75Mbps with AVC-I, and it worsens as the bitrate decreases. At 50Mbps and below, the flicker is extremely problematic and we felt that the video quality was too low for high quality broadcast contribution applications, even when the source is downscaled to 1440x1080 or 960x720. Furthermore, experiments with different screen technologies showed that some displays like high-end plasmas may attenuate the flicker while others could amplify it. Finally, this defect can be greatly reduced by de-noising the source video. Unfortunately, this is something that is to be avoided in broadcast contribution applications where the signal has to be transported as transparently as possible.

Around 100Mbps, both codecs perform well, even on challenging content. Pictures are flicker-free and coding artifacts are difficult to notice. However, noise or film-grain looks low-pass filtered and its structure seems sometimes slightly modified, but we did not feel this as an important issue.

\(^3\) While we found some disparities among JPEG-2000 products from various vendors, we considered only the best one, as this was only a codec evaluation and not a product or vendor assessment.
All those defects are less visible and annoying as the bitrate is increased. But while AVC-I picture quality raises uniformly, some JPEG-2000 products may still exhibit blurriness artifacts even at 180Mbps. Further experiments with a JPEG-2000 software [8] did not produce similar flaws. This leads us to believe that this issue is related to current implementations of JPEG-2000, but not to the codec itself. Nevertheless, using available JPEG-2000 contribution pairs, we were not able to find a bitrate at which compression is visually lossless on all high-definition broadcast content. On the other hand, ATEME AVC-I encoder could be considered visually lossless at 150Mbps, even when encoding grainy content like movies.

Finally, we were surprised to observe that the subjective impression at a given bit-rate is just slightly better in 720p than in 1080i. This showed that both technologies are able to manage properly interlaced material. However, we found it logical to notice that 1080p50/60 content needs about twice the bitrate of 1080i or 720p to achieve an equivalent perceived video quality.

**Bitrates to use in Broadcast Contribution**

The subjective analysis of an actual AVC-I implementation on various on broadcast contribution content permits us to categorize its usage according to the available transmission bandwidth. Table 2 present our findings on 1080i25 and 720p50 high definition formats.

<table>
<thead>
<tr>
<th>Bitrate in AVC-I</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤ 50Mbps</td>
<td>Video quality is too low for high quality broadcast contribution applications</td>
</tr>
<tr>
<td>50Mbps - 75Mbps</td>
<td>Acceptable on low-noise sources but poor on most sequences</td>
</tr>
<tr>
<td>75Mbps - 90Mbps</td>
<td>Acceptable</td>
</tr>
<tr>
<td>90Mbps – 110Mbps</td>
<td>Good</td>
</tr>
<tr>
<td>110Mbps - 150Mbps</td>
<td>Excellent</td>
</tr>
<tr>
<td>≥ 150Mbps</td>
<td>Visually lossless</td>
</tr>
</tbody>
</table>

Table 2. AVC-I bitrate versus video quality for 1080i25 - 720p50 content.

Since AVC-I does not make use of temporal redundancies, 30Hz content (1080i30 or 720p60) are more difficult to encode than 25Hz material. And bitrates to achieve the same perceived video quality level have to be raised by 20%.

**Conclusion**

The availability of high speed networks for contribution applications enables Broadcasters to use Intra-only video compression codecs instead of the more traditional long-GOP formats. This allows them to benefit from distinctive advantages like:

- Low encoding and decoding delays
- More constant video quality
- Easy editability when the content is stored
- Lower sensitivity to transmission errors

However, currently available intra-only video codecs require to choose between interoperability and coding efficiency.
AVC-I being just the restriction of standard H.264/AVC to intra-only coding, avoids making difficult compromises. It is more efficient than any other available intra-only codec but more importantly, it benefits from the strong standardization efforts that permitted H.264/AVC to replace of MPEG-2 in all broadcast applications.

Low end-to-end latencies is a major requirement of broadcast contribution applications. And it remains to be seen if AVC-I is able to achieve low delay constraints. ATEME CM4101/DR8400 encoder/decoder pair being the first contribution equipments to offer an AVC-I profile on top of standard H.264/AVC contribution profiles, it provides end-to-end delays similar to other intra-only technologies when they were introduced and which are adequate in most contribution applications. But we have shown that much lower values can be achieved while still being fully compliant with standards and operational practices. Therefore, AVC-I, can achieve the very low delays required by some applications.

The other major aspect of a video compression format is the video quality it may provide at a given bitrate. Academic research has shown that AVC-I is at least as efficient as the best intra-only codec currently available. A subjective study was performed using ATEME CM4101/DR8400 pair and products from various vendors to evaluate the video quality of AVC-I compared to other intra-only codecs. This analysis permitted to identify specific coding artifacts and confirmed the superiority of AVC-I at high bitrates.

Acknowledgements

Thanks to my colleagues Jérôme Vieron and Marc Baillaioine for their valuable help during the preparation of this paper.

References


