Comparison of Compression Algorithms for High Definition and Super High Definition Video Signals

Hrvoje Balasko
Audio Video Consulting Ltd., Karlovačka 36b, 10020 Zagreb, Croatia
E-mail: hrvoje.balasko@avc.hr

Abstract – Choosing the optimal compression algorithms for high definition and super high definition video signals including 2K and 4K resolutions and 3D stereoscopic is a complex task because of the technological needs and capabilities of different parts of the production, transmission and archiving chains. Specifics of recording, postproduction, contribution, distribution, transmission, archive and picture display are analyzed because of fully different technological capabilities. Further complexity of the proper compression system selection can be influenced by the issues related to the multiple cross-conversions between different compression algorithms and the often-present need for a very high compression rate related to the limited transmission bandwidth.

Keywords – High Definition and Super High Definition Video Signals, 2K, 4K, 3D, Video Compression Algorithm, Bandwidth

1. INTRODUCTION

Compression techniques were developed based on the studies of human perception. Data reduction is obviously the only option if a raw bitrate of HDTV (High Definition Television) or SHDTV (Super High Definition Television) picture is 1 to 10 Gbps and some parts of the whole chain, like transmission channels are well beyond that, with a bandwidth in the range of 1 to 30 MHz.

An encoding and decoding cycle is called a compression generation. In each generation, unperceived information is discarded and cannot be recovered. At some point, multiple compression generations will lead to noticeable artifacts. This has an impact on content production and distribution.

Video compression engines are divided into two broad categories.

• Lossless compression reduces the volume of data and, when reconstructed, restores it to its original state, without the loss of any information.
• Lossy compression discards data based on visual sensory characteristics and limits. Because of the sensory dependency, lossy encoders and decoders are called perceptual. When reconstructed, the sensory information is virtually indistinguishable from the original source.

Many methods of compressing digitized moving images exist.

Several versions and levels of Motion-JPEG, MPEG-2, MPEG-4 and many others are in use and they are being continuously further developed, improved and implemented [3].

2. COMPRESSION

Compression is the science of reducing the amount of data used to convey information. Compression relies on the fact that information, by its nature, is not random but exhibits order and patterning. If that order and patterning can be extracted, the essence of the information can be represented and transmitted using less data than would be needed for the original. Then, at the receiving point the original can be fully or very closely reconstructed.

There are several families of compression techniques, fundamentally different in their approach to the problem. Sometimes, these techniques can even be used sequentially to good advantage. Today complex compression systems use one or more techniques from each family to achieve the greatest possible reduction of data [2].

Two major compressions methods are to be differentiated – data and image compressions.

Data compression reduces the number of bits required to store or convey any kind of data (numeric, text, binary, image, sound, etc.) by exploiting statistical properties of the data. The reduction comes at the expense of the computational effort to compress and decompress. Data compression is, by definition, lossless. Most data compression techniques, including run-length encoding (RLE) and Lempel-Ziv-Welch (LZW), accomplish compression by taking advantage of repeated byte substrings which are, besides byte strings, often present in general computer applications binary data. The data lossless coding is generally restricted to compression ratio of around 2:1.
Image compression is based on the fact that typically there is a strong vertical and horizontal correlation among pixels or some wider parts of the image. For example, transform techniques are effective for the compression of continuous tone (grayscale or truecolor) image data. The discrete cosine transform (DCT) has been developed and optimized over the years and now is the method of choice for continuous-tone compression [1].

Some drawbacks of compression also exist. By definition, compression removes redundancy from signals. Redundancy is, however, essential to making data resistant to errors. As a result, compressed data are more sensitive to errors than uncompressed data. Thus transmission systems using compressed data must incorporate more powerful error-correction strategies and avoid compression techniques which are notoriously sensitive. Perceptive coders introduce noise and in concatenated system the second codec could be confused by the noise due to the first. Also the concatenation of different compression systems should be avoided if possible because this causes generation loss.

3. MOTION VIDEO COMPRESSION

High definition (HD) and super high definition (SHD) motion video compression can be handled using several methods. All of them are being continuously developed and their performances are being improved.

3.1 Motion-JPEG

Concept of Motion-JPEG is simple – each field or frame of the image sequence is intra coded as a JPEG image (usually with adaptive encoding to provide a fixed number of bits per frame) and the fields or frames are processed further or stored sequentially. No advantage is taken of temporal coherence but this compression technique is very convenient for editing, post-production and digital cinema storage (using both, 2K and 4K resolutions).

Achievable compression ratios are not very high and they are ranging from about 2:1 to about 20:1. Because of that it can not be used for HD transmission but there are certain applications where Motion-JPEG and Motion-JPEG2000 as the most recent version of that compression algorithm can be successfully used. In fact, measured performances are very closed to H.264/AVC, I frame only performances.

Motion JPEG is not fully standardized (it is not covered by ISO/IEC JPEG standard) so different manufacturers are varying this compression algorithm very often in their proprietary ways.

3.2 MPEG-2

MPEG is the acronym for the Moving Pictures Experts Group which was formed by the ISO (International Standards Organization) to set standards for audio and video compression and transmission.

First generated standard within MPEG was MPEG-1 which was related to the compression of low resolution video and very low bitrates. It is important that MPEG-1 introduced the great majority of the coding tools which would continue to be used in MPEG-2 and MPEG-4. These included an elementary stream syntax, bidirectional motion-compensated coding, buffering and rate control. Many of spatial coding principles of MPEG-1 were taken from JPEG.

MPEG-2 builds upon MPEG-1 by adding interlace capability as well as a greatly expanded range of picture sizes and bitrates. The use of scalable systems is also addressed, along with definitions of how multiple MPEG bitstreams can be multiplexed.

MPEG-2 (or called H.262 by ITU-T) has to many applications to solve with a single standard and so it is subdivided into Profiles and Levels. Profile describes the degree of complexity whereas a Level describes the picture size or resolution which goes with that Profile.

![Profiles and Levels in MPEG-2](image)

Group of Pictures (GOP) is a sequence of compressed video frames of a defined pattern and length. The pattern and length of a GOP impact amount of compression and video quality that is attained during compression process.

Frames are compressed as either:

- I frames: Intraframe, a frame that is encoded independent of any other frames
- P frames: Predictive, encoding is dependent on a previous I frame
- B frames: Bidirectional, encoding is dependent on previous or subsequent I or P frames
Compression of a video frame to produce an I frame is the simplest comparing it to P and B frames because it is compressed without reference to any other frames. The raster is subtracted from the anchor (I) frame and then divided into blocks of 8x8 pixels. Y, U and V are processed separately. Further steps reduce the amount of data:

- Discrete Cosine Transform
- Weighting and Requantization
- Variable Length Coding (Entropy/Huffman coding)
- Run Length Coding

Compression of P and B frames is processing only the difference between related previous or subsequent frames [3].

MPEG-2 supports a variety of bitstream types for various purposes as shown on Fig. 3. Elementary stream is the output of a single video (or audio) compressor. In transmission, many elementary streams can be combined to make a transport stream. Multiplexing requires blocks or packets of constant size. Single program transport stream carries only the elementary streams of one TV program. For certain purposes, such as recording a single elementary stream instead of an inappropriate transport stream a program stream can be used.

### 3.3 MPEG-4

MPEG-4 family of algorithms uses further coding tools with additional complexity to achieve higher compression ratios than MPEG-2. Besides more efficient coding of video, in the more complex Profiles, MPEG-4 compressed bitstream describes three-dimensional shapes and surface texture.

While earlier methods of MPEG compression divided the video pixel grid into fixed sized block of pixels, MPEG-4 has added an object-oriented paradigm to the compression tool kit. These items are video object, still object, mesh object and face and body animation object. The texture coding is also introduced and it breaks a scene down into background and foreground objects and each can be coded separately. Animation is enabled by applying vector information to faces and bodies.

All of these is significantly improving the compression rate compared to MPEG-2 which is leading to lower bitrate for the same picture quality or higher picture quality using identical bitrate.

### 3.4 H.264/AVC - MPEG-4 Part 10

Advanced Video Coding (AVC), developed by Joint Video Group (JVT) and approved as Recommendation H.264 (by ITU-T) and as MPEG-4 Part 10 (by ISO/IEC) is originally intended to compress moving images that take the of eight-bit 4:2:0 coded pixel arrays. Although complex, AVC offers between two and two and a half times the compression factor of MPEG-2 for the same picture quality [24].

![Fig. 4. AVC encoder with spatial (for I frames) and temporal (for P and B frames) predictors](image)

![Fig. 5. Comparison of Texture Coding Systems of MPEG-2 (a) / MPEG-4 (b) / AVC (c) [6]](image)
In the original standard following three profiles were defined:

- Baseline (BP)
- Extended (XP)
- Main (MP)

Later on, because of a market and technology requirements new high profiles were added in Fidelity Range Extensions (FRExt) Amendment [12]:

- High (HP)
- High 10 (Hi10P)
- High 4:2:2 (Hi422P)
- High 4:4:4 (Hi444P)

Sixteen different levels are defined within H.264/AVC. It is important to constrain the processing power and memory size needed for the implementation. Picture size and frame rate play the main role in influencing those parameters. Levels also provide constraints on the number of reference pictures and the maximum compressed bitrate can be used [8].

<table>
<thead>
<tr>
<th>Level Number</th>
<th>Typical Picture Size</th>
<th>Typical frame rate</th>
<th>Maximum compressed bitrate (for VC1.1 in Non-FRExt profiles)</th>
<th>Maximum number of reference frames for typical picture rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>QCIF</td>
<td>15</td>
<td>44 Mbps</td>
<td>4</td>
</tr>
<tr>
<td>26</td>
<td>QCIF</td>
<td>15</td>
<td>32 Mbps</td>
<td>4</td>
</tr>
<tr>
<td>3, 12</td>
<td>CIF on QCIF</td>
<td>15</td>
<td>30 (QCIF)</td>
<td>2 (QCIF)</td>
</tr>
<tr>
<td>2, 7</td>
<td>CIF</td>
<td>15</td>
<td>30 (QCIF)</td>
<td>2 (QCIF)</td>
</tr>
<tr>
<td>1</td>
<td>CIF</td>
<td>15</td>
<td>30 (QCIF)</td>
<td>2 (QCIF)</td>
</tr>
<tr>
<td>2, 7</td>
<td>CIF</td>
<td>15</td>
<td>30 (QCIF)</td>
<td>2 (QCIF)</td>
</tr>
<tr>
<td>2</td>
<td>1/3x (256x144)</td>
<td>30</td>
<td>5,048</td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>30</td>
<td>5,048</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>1280x720p</td>
<td>15</td>
<td>14 Mbps</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>1280x720p</td>
<td>15</td>
<td>14 Mbps</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>HD Format 1/3 (720p or 1080i)</td>
<td>59y</td>
<td>50 Mbps</td>
<td>4</td>
</tr>
<tr>
<td>4, 2</td>
<td>HD Format 1/3 (720p or 1080i)</td>
<td>59y</td>
<td>50 Mbps</td>
<td>4</td>
</tr>
<tr>
<td>5, 10</td>
<td>576x432</td>
<td>25</td>
<td>13 Mbps</td>
<td>3</td>
</tr>
<tr>
<td>5, 10</td>
<td>576x432</td>
<td>25</td>
<td>13 Mbps</td>
<td>3</td>
</tr>
</tbody>
</table>

Fig. 6. Levels in H.264/AVC [8]

As FRExt is still rather new, and as some of the benefit of FRExt is perceptual rather then objective, it is somewhat more difficult to measure its capability. The results of a subjective quality evaluation done by Blue-ray Disc Association are reproduced in Fig. 7 below. The test was conducted on 24 frame/sec film content with 1920x1080 progressive-scanning shows the following:

- The High profile of AVC FRExt produced better video quality than MPEG-2 using only one third as many bits (8 Mbps versus 24 Mbps)
- The High profile of AVC FRExt produced nominally transparent (difficult to distinguish from original uncompressed video) video quality at only 16 Mbps [8]

Fig. 7. Perceptual Test of H.264/AVC FRExt High Profile Capability by Blue-ray Disc Association [8]

Some objective tests like an example objective (PSNR) comparison performed by FastVDO confirm the strong performance of the High profile. 8 Mbps of AVC FRExt outperforms 20 Mbps of MPEG-2, as shown in Fig. 8.

Fig. 8. Objective performance of AVC FRExt High profile vs. MPEG-2 on a test 720p clip "BigShips" [8]

There are issues related to different compression algorithms efficient transcoding, like between widely used MPEG-2 and AVC/H.264 [16]. For certain applications like 3D TV new developments like H.264/SVC (Scalable Video Coding) algorithm are to be considered. It is also developed by Joint Video Team (JVT) and it supports spatial, temporal and quality scalability for video [28].

4. APPLICATIONS

4.1 Recording and Postproduction

Picture shooting, recording and postproduction is the part of a chain where minimum possible compression level is to be used. Picture shooting is the key of final product picture quality so any form of compression has to be avoided or kept on a minimum level. Recording devices internally often use mild compressions. High end HD postproduction is mostly using intra-frame compression if needed and general postproduction is also accepting inter-frame compression methods.

Some new codec’s like Dirac Pro are also being standardized in SMPTE as VC-2 and used for HD and Film postproduction and lossless coding [11].

Comparison with JPEG2000 and AVC/H.264 shows that performances od Dirac Pro can be in the same range [28].
4.2 Contribution

Improved bandwidth efficiencies are continuously requested for all contribution applications. Beside the lower operational costs in many cases economic migration from SD to HD is enabled by higher compression efficiency.

Chosen compression algorithm is related to contribution application and picture quality request, delivery method, bitrate, picture resolution, sampling and latency requirement.

<table>
<thead>
<tr>
<th>Compression</th>
<th>Video</th>
<th>Audio</th>
<th>Aspect Ratio</th>
<th>Sampling Rate</th>
<th>Latency</th>
<th>Other Requirements</th>
</tr>
</thead>
<tbody>
<tr>
<td>NVGathering</td>
<td>SD/HD</td>
<td>SD</td>
<td>4:3</td>
<td>25/30 fps</td>
<td>Low</td>
<td>Multi-channel audio</td>
</tr>
<tr>
<td>Mid-range contribution</td>
<td>SD/HD</td>
<td>SD</td>
<td>4:3</td>
<td>25/30 fps</td>
<td>Low</td>
<td>Multi-channel audio</td>
</tr>
<tr>
<td>Premium contribution</td>
<td>SD/HD</td>
<td>SD</td>
<td>4:3</td>
<td>25/30 fps</td>
<td>Low</td>
<td>Multi-channel audio</td>
</tr>
</tbody>
</table>

Fig. 9. Segmenting the Contribution market to identify key encoding requirements [10]

Currently, MPEG-2 is mostly being used but for HD applications it is fast migrating to MPEG-4 and for some special applications and if high bitrates are available also JPEG-2000 can be considered.

![Fig. 10. Possible applications of MPEG-2, MPEG-4 and JPEG-2000 in HD contribution [10]](image)

4.3 Distribution and Transmission

Distribution and transmission are two parts of a chain which request maximum available compression ratios. Most of transmission principles are strongly limited by available bandwidths. RF terrestrial and satellite bands, cable and many other paths have hard limits which are making delivery of HD video content to the consumers impossible without powerful compression. Major compression algorithm in SD, MPEG-2 is being gradually replaced by, for HD delivery, much more efficient H.264/AVC.

Certain MPEG-2 efficiency improvements are performed on some new encoders but that will hardly slower the migration to H.264/AVC [13].

4.4 Storage

Optimal compression for storage applications is related to the specific applications. Deep archive might be related to high compression ratios but near archive which is being often used requires low level compression. If high resolution Digital Cinema is considered, intra-frame compressions like JPEG-2000 are used for storage/archiving solutions.

4.5 3D Stereoscopic Broadcasting

Idea of 3D stereoscopic broadcasting in HD resolution is currently booming. Transmitting of two full quality HD signals as independent streams is impractical as it uses up significant bandwidth and risks two signals picking up unwanted differential artifacts or getting out of sync.

There are several schemes that aim to remove this including SideBy Side, Checker Board and 3D specific compression.

In Side By Side, a single signal is created that "squashes" two pictures into a single one. The viewing device expands that out. Despite the related reduction in horizontal resolution final 3D picture is more than acceptable.

Checker Board use a boarding scheme of displaying single video stream to a viewer.

3D specific compression, besides standard single video stream compression principles, takes advantage of a redundant data between left and right views. Some of recent versions of MPEG-4 like SVC/H.264 are obviously the right tool for 3D compression [32].

Some new approaches like depth-image-based rendering (DIBR) are also being considered [9].

4.6 Digital Cinema

Digital cinema today is related to 2K and 4K resolutions. Preferred compression scheme is definitely (Motion)JPEG-2000 although some versions of intra-frame MPEG are also considered. Fast migration from 2K to 4K and implementation of 3D in digital cinemas is influencing development of related playout equipment, compression schemes and content delivery channels [21].

5. CONCLUSION

Compression should not be used for its own sake, but only where the genuine bandwidth or cost bottleneck exists. Even then the mildest compression possible should be used. Whilst high compression factors are permissible for final delivery of material to the consumer, they are not advisable prior to any post-production stages. For contribution material, lower compression factors are essential and this is sometimes referred to as mezzanine level compression.

REFERENCES
