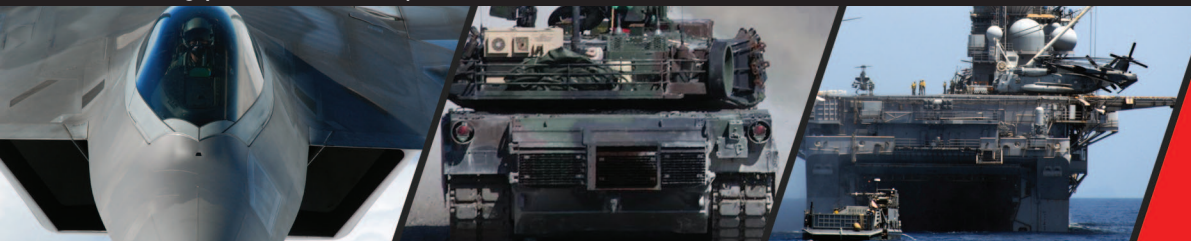


A Comparison of MPEG4 (H.264) and JPEG2000 Video Compression and Decompression Algorithms

cwcdefense.com



A Comparison of MPEG4 (H.264) and JPEG2000 Video Compression and Decompression Algorithms

Introduction

In the past decade digital video recording and distribution has been made much more viable by the introduction of new video compression and decompression techniques. Analog video applications have traditionally typically needed large investments in infrastructure to be able to provide any useful recording and distribution functionality (even at very low channel counts), even with the older generation of standard definition (SD) PAL/NTSC video.

With the introduction of High Definition (HD) video, much higher data rates need to be recorded and moved around in the digital domain meaning new technologies have had to be developed in order to accommodate these higher specification video streams.

In addition to hardware level innovations such as gigabit transceivers which allow much more data to be transmitted across single channels (rather than the blunt instrument of just adding more parallel infrastructure), hardware (or software) based coder/decoders (codecs) can help ease the issues of system development, not only by reducing the bandwidth required for each channel of video but also by reducing the storage requirements. This of course has benefits in terms of development time as well as overall cost for any given system.

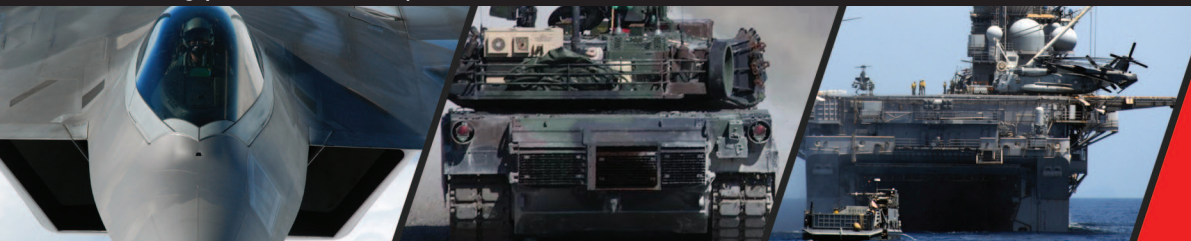
This white paper has been written to examine two of the more popular of these technologies, MPEG-4 Part 10 (also known as Advanced Video Coding (AVC) and ITU-T H.264) and JPEG2000. Both these codecs have distinct practical advantages and disadvantages in relation to each other, so to begin it will be helpful to look at how they both work at a high level before comparing the results of some tests so that we may draw conclusions on which applications each are best suited to. Note that audio and other associated metadata are not considered within the scope of this white paper.

Codec Technology

The core concept behind any video codec is that in any given picture there is redundancy in the image, i.e. information (regarded as 'invisible' to the human eye) can be removed reducing the amount of data required to represent the picture, at least up to a certain point. After this point a compressed then decompressed (reconstructed) image will contain certain artefacts – the nature of these artefacts differs depending upon the algorithm used to compress and decompress the image, as well as the image contents. Depending on the application, one codec's artefacts may be preferable over those of another.



Image courtesy of DefenseImagery.mil



Limitations on the bandwidth of any network being used are a very important consideration. If an existing network infrastructure is in place, it must be remembered that it will not take many channels of HD video to completely overwhelm even a Gigabit Ethernet network, and directly related to this are storage requirements. Some applications can require that several days worth of video is stored, in which case multiple terabytes of hard disk may be required, which has a related system cost. This can be not only in terms of the cost of the storage media itself, but the ability to remove the media, perhaps for debrief (for instance from a flight mission camera to a ground station playback PC). SSD storage is becoming considered more frequently for this sort of application and the cost/density/reliability aspects of that all have an influence as well.

It is important to remember that all codecs have their upsides and downsides. In considering any single algorithm for use the system architect must be aware of all factors and also that it is unlikely that any one codec will be completely ideal for its intended application.

MPEG-4 Part 10

MPEG-4 Part 10 was introduced in 2003, and as such is the slightly newer of the two codecs discussed here. MPEG stands for the Motion Picture Experts Group, and as the name suggests, the MPEG-4 codec is designed to encode and decode motion pictures, as opposed to single frames (such as a still picture). MPEG-4 is known as a 'lossy' codec as the reconstructed images will always be inferior to the original pictures.



Image courtesy of Defense.gov

The key principle behind MPEG4 encoding is the idea of the codec working on what is known as a 'Group of Pictures' (GOP). These consist of three types of image. The first type, the 'I-frame', is a stand-alone compressed frame, and there is one I-frame at the beginning of each GOP. Between this are two further types of image which predict the change (motion) between each I-Frame. The first level of predicted frame is the P-Frame, which contains the difference between the current and preceding frame. Secondly, the B-Frame contains the difference between the current frame and both the preceding and following frames. So, where before compression each frame is transmitted in full, after compression the bandwidth of the images is reduced as only a subset of information from each I-Frame is required. To reduce the bandwidth for any given moving image, the length of the GOP is increased. As the GOP is increased the bandwidth will fall, but so will the quality of the images as deeper prediction is required. Furthermore, as the number of I-frames in a video stream is reduced, the stream becomes less easily editable. Figure 1 shows a graphic representation of this.

This has an effect that is more pronounced depending upon the source material. For images where frame to frame changes are very small (such as a camera pointed at a sunset, or where there are large areas of the same color, or someone speaking to a camera where only his/her mouth is moving) then inter-frame changes are very small, leading to a good compression ratio and more accurate reconstruction at the decoder. For action movies, where there are many sharp cuts from scene to scene and fast moving objects, the inter-frame prediction starts to break down as the changes between successive frames are greater. This is one of the reasons why MPEG-4 Part 10 is popular in video conferencing applications as there is little frame to frame motion meaning very low streaming bit rates can be achieved.

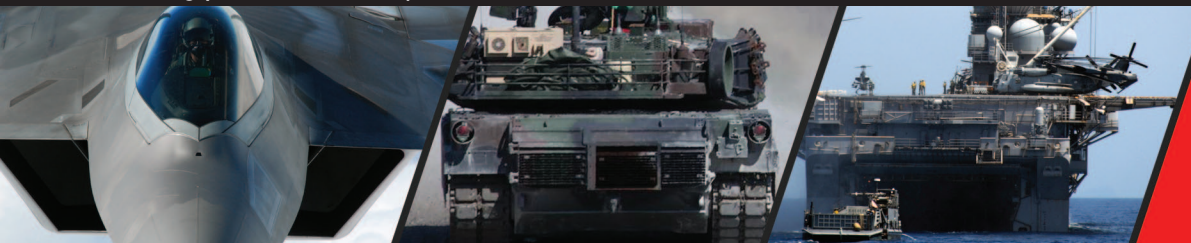
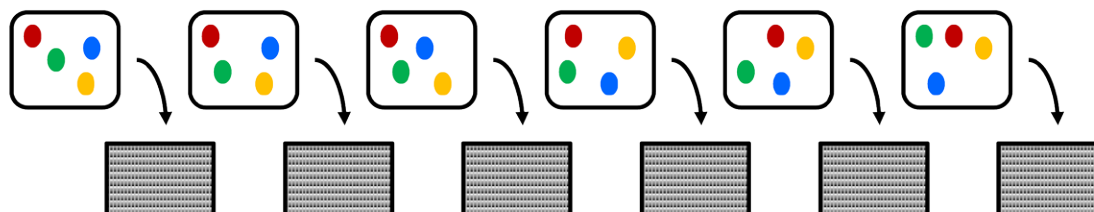
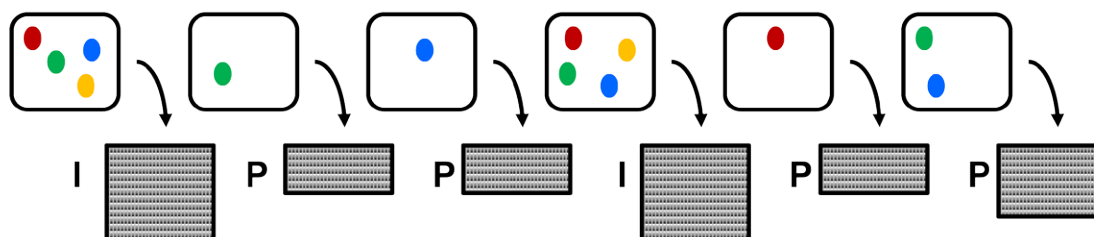


Figure 1: Intra and Inter-Frame Coding Example



Intra-frame coding – uses only spatial redundancy



Inter-frame coding – uses spatial and temporal redundancy

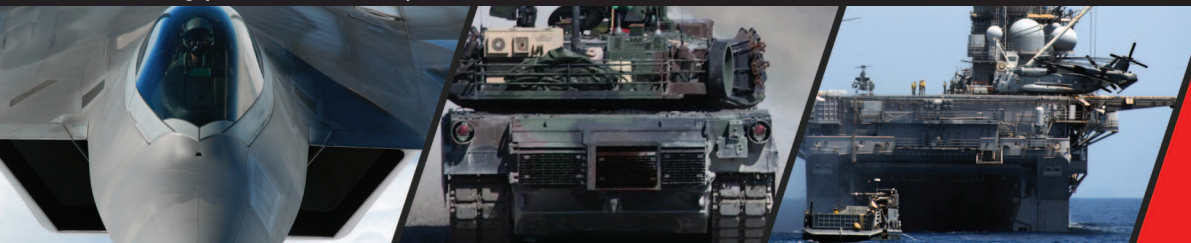
JPEG2000

JPEG stands for Joint Photographic Experts Group, and JPEG2000 is a successor to the earlier JPEG standard introduced in the early 1990's to encode and store still images, and, in its Motion JPEG2000 incarnation, moving images. Unlike MPEG4's use of both intra- and inter-frame encoding, JPEG employs only intra-frame encoding. This means that each frame is compressed individually, and no attention is paid by the encoder to either previous or following frames, and so it is not predictive in any way. In contrast to MPEG-4, JPEG2000 can be operated in both a lossless as well as a lossy mode. The Discrete Wavelet Transform at the heart of the algorithm can be based on either a reversible filter (lossless) or a non reversible filter which is lossy but provides a higher compression rate for any given material.

Because each frame is individually compressed, JPEG2000 offers a number of advantages over MPEG-4. For example, the latency to compress a frame is shorter (as the codec does not have to rely on generating forward and reverse differences between frames). In terms of video editing capabilities, as

no frames are dropped during encoding there is a direct correlation between each frame of encoded and decoded video, although with lossy compression the image quality of the reconstructed frame will be reduced. The disadvantage of this of course is that JPEG2000 has to encode and transmit each frame individually, unlike the MPEG-4 algorithm, which sends only the differences between a number of frames, ultimately meaning that the bandwidth requirements to transmit JPEG2000 can be higher, with an increase in storage needs. JPEG2000 is more resilient to errors in transmission than MPEG-4; a small loss of data in a JPEG2000 stream will be far less noticeable than the corresponding loss in an MPEG-4 stream.

Also, JPEG2000 is computationally much more complex to decode than MPEG-4. This means that though JPEG2000 can be decoded in software (for example on a generic PC CPU running a software application such as Kakadu), this can place quite a burden on the host CPU, leaving it less capable of running other tasks, or in extreme cases leading to frames being dropped. This means that most of the time it is desirable to accelerate JPEG2000 using a hardware-based decoder.



**CURTISS
WRIGHT** Controls
Defense Solutions

Test Setup

In order to get some real world metrics for compression quality versus file size for each of the codecs, a test setup was created to encode various types of material at different compression rates, such that the resultant files could be examined for both size and video quality after being decompressed. In both cases a hardware compression card was used based on a standard PC platform, and for decompression a software codec was used (though the frame rate suffers for JPEG2000 decompression, file size and image qualities are not affected). As the codecs do not have compression quality controls that directly correlate, the aim of the experiment is not to provide a direct comparison between the two, but rather to give a relative indication of storage requirements and reconstructed image quality for a given set of identical source material.

A standard laptop PC with HDMI output was set up with the open source video playback application VLC. Three types of source material were selected, and each played 3 times for a 5 minute period with a different compression quality selected each time. The playback resolution out of the laptop in all cases was set to 720p at 60 frames per second.

The source video material was selected to test each codec for compression rate and subsequent reconstructed quality. The first file used was stock library footage (NASA) where there was a high amount of motion in the picture, and the video on the whole had high contrast between the colors in any given frame. The second stock film (NATURE) had a relatively small amount of motion between frames (it being slow motion to begin with), with the range of colors, being 'natural', that did not cover a very wide dynamic range. Finally, a synthesized source of a radar video display was used which for both codecs should stress the ability to handle text and graphics.

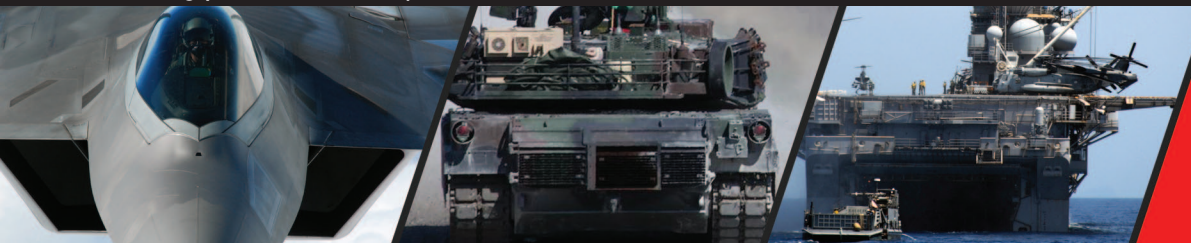


For the MPEG-4 codec test the output of the laptop was fed into the DVI input (after HDMI-DVI conversion) of a Curtiss-Wright Controls Defense Solutions MPEG-4 H.264 SPMC-281 codec card hosted on a standard PC. Codec quality is controlled on a sliding scale from 1 – 51, 1 representing the least amount of compression and 51 representing the heaviest available. A nominal setting for the codec is given as 30, therefore the values of 15, 30 and 45 were chosen for the experiments.

Similarly, for JPEG2000 a Curtiss-Wright XMC-280 card with DVI input was used. For JPEG2000 the main quality control is the number of Megabits per second (Mbps), so for the JPEG2000 tests 30, 20 and 10 Mbps were used on the same material as the MPEG-4 test.

Both cards had the DVI input capture controlled by Curtiss-Wright's Sentric2 software, version 1.8.1 running on Ubuntu 10.4 LTS operating system.

In the interests of a fair comparison, both codecs were used in lossy mode.



Results

The results for both experiments are shown in the table below.

Table 1: Intra and Inter Frame coding example

File Size (MB)	MPEG-4			JPEG2000		
	Comp. Factor 15	Comp. Factor 30	Comp. Factor 45	30 Mbps	20 Mbps	10 Mbps
Stock 1 Video: High Entropy (NASA)	451	67	17	1100	730	378
Stock 2 Video: Medium Entropy (NATURE)	387	56	19	1000	714	375
Synthesized Graphics and Text (SOFT SCAN)	81	35	17	1100	702	349

MPEG-4

At a compression setting of 15, with MPEG-4 the reconstructed images are almost lossless to the naked eye – no ‘macro-blocking’ can be seen and motion is smooth. At a setting of 30, artefacts start to introduce themselves as a mosaic effect (restricted to certain parts of the image) and the motion between frames is just a little more juddery in comparison to the lowest setting, but the effects are subtle and the video is still highly watchable. Predictably, at the highest compression setting used in this experiment the replayed video is very heavily pixelated into larger macro-blocks, frame to frame motion is very poor and while it is still easy enough to make out what the images are it cannot claim to be high quality.

Where there have been encoding errors on playback this is largely noticeable in that parts of the picture are not reconstructed at all, leading to strips 15 to 20 pixels wide of a single color.

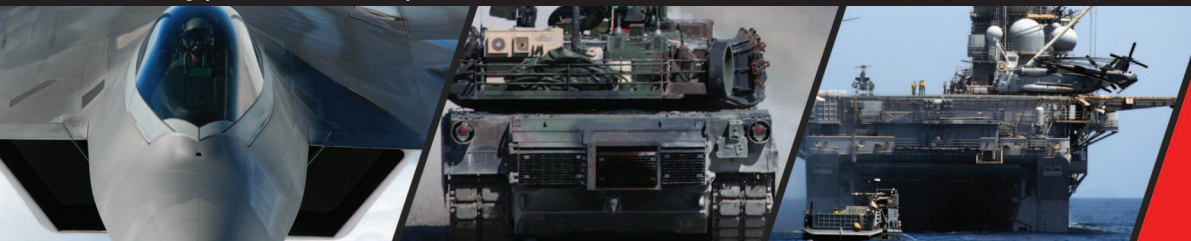
Interestingly, the file sizes for all three video sources seem to converge as the compression quality moves towards maximum compression.

JPEG2000

In terms of quality JPEG2000 fares extremely well at 30Mbps, with the reconstructed video being nearly identical to the source, with only the smallest amount of fidelity lost which is almost invisible to the naked eye. At 20Mbps quality was only reduced slightly but there is uniform noise over the entire image – however, the picture is still fairly high quality with most of clarity maintained, especially around the edges of objects. Down at 10Mbps the noise that appears at 20Mbps becomes much more pronounced, becoming a uniform ‘watery’ like effect across each frame. Unlike the MPEG-4 at its lowest setting this picture is still very watchable, and it is still possible to make out distinct objects and their edges.



Image courtesy of DefenseImagery.mil



**CURTISS
WRIGHT** Controls
Defense Solutions

Conclusion

From the results, it is quite clear that in terms of reduced requirements for both network streaming and storage, MPEG-4 is the winner. However, JPEG2000 is the winner in terms of picture quality. JPEG2000 is far more resilient in terms of transmission or encoding error than MPEG-4 and because of its frame-by-frame nature can be easily manipulated in its decompressed form – MPEG-4 is ‘digital VHS’ meaning little can be done with it in terms of editing. An interesting point here is also in terms of legal use – in some countries MPEG-4 video is inadmissible as evidence due to its inter-frame encoding whereas JPEG2000 has been successfully used as evidence in a number of legal cases worldwide.

Overall, although we cannot compare the two codecs directly due to their different compression methods (meaning that ‘compression quality’ settings between the two never directly correlate) it can be seen that JPEG2000 compression appears to degrade in terms of the reconstructed pictures much more gracefully between the lowest and highest compression settings, but the price that is paid in terms of storage and network bandwidth (for network transmission or streaming) purposes is relatively high. Conversely, though MPEG-4 is lower quality, it has many advantages in the network/storage domain, even though the reconstructed video may be of poorer quality and the end user has less ability to edit it.

In some applications both codecs have their place. As an example, in a recent application Curtiss-Wright was asked to supply a video recording and distribution system which required that remote CCTV sensors streamed live video to remote operators over a WAN. The WAN was highly bandwidth limited (it was already in place, and time shared between many users); however, the highest quality of video recording needed to be stored for later analysis. The solution to this was to provide both codecs in parallel, an MPEG-4 encoder streaming live video to the operators over the network while a local RAID array stored the high quality JPEG2000 video.

Further Information

Curtiss-Wright Controls Defense Solutions

(Video Distribution Systems) – Homepage at: http://www.cwcembedded.com/video_distribution_system.htm

MPEG

Motion Picture Experts Group. The committee is responsible for developing the MPEG standards. Homepage at: <http://mpeg.chiariglione.org/>

SPMC-281 Video Compression Card

http://www.cwcembedded.com/pmc-281_video_compression.htm

JPEG

Joint Photographic Experts Group. The committee is responsible for developing the JPEG and JPEG2000 standards. Homepage at: <http://www.jpeg.org>

XMC-280 Video Compression Card

http://www.cwcembedded.com/xmc-280_video_compression.htm

Contact Information

To find your appropriate sales representative:
Website: www.cwcdefense.com/sales
Email: defensesales@curtisswright.com

Technical Support

For technical support:
Website: www.cwcdefense.com/support
Email: support1@cwcembedded.com

The information in this document is subject to change without notice and should not be construed as a commitment by Curtiss-Wright Controls Defense Solutions. While reasonable precautions have been taken, Curtiss-Wright assumes no responsibility for any errors that may appear in this document. All products shown or mentioned are trademarks or registered trademarks of their respective owners.



defensesales@curtisswright.com

cwcdefense.com