An Analysis of MPEG Encoding Techniques on Picture Quality

A Video and Networking Division White Paper

By Roger Crooks

Product Marketing Manager

June 1998

Tektronix

Tektronix, Inc.
Video and Networking Division
Howard Vollum Park
PO Box 500, MS 58-965
Beaverton, OR 97077

US Sales	Tel: 800-547-8949
	Fax: (503) 627-7275
Europe	Fax: 33 (1) 69 07 0937
Asia	Fax: 85 (2) 2802 2996
InfoFAX	800-704-8843
URL	http://www.tek.com/VND

An Analysis of MPEG Encoding Techniques on Picture Quality

Abstract:

As a broadcaster, you now have a new option for video compression, the MPEG-2 4:2:2 profile@main level. It was adopted as a standard by SMPTE last year. Several new products from Tektronix and other manufacturers have recently been introduced using this format.

Motion JPEG has been the dominant compression format used in most video servers. Broadcasters have come to understand the quality they can expect at different bit rates. For instance, 24 Mb/s quality is equivalent to Beta SP, while 48 Mb/s is equal to Digital Betacam. Broadcasters simply choose a compression rate appropriate to their application.

MPEG introduces a new variable - temporal compression. Now you can improve quality by changing either the data rate or length of the group of pictures (GOP) over which redundant picture information is removed. A quick re-cap of compression variables will help us further understand their impact on picture quality. Some pundits have proposed that MPEG would provide the same quality at half the JPEG data rate, but the vote is still out on that one.

This paper will explore Tektronix PQA200 Picture Quality Analysis System and how it can be used to measure the video quality of material that has been digitized and compressed using the new PDR300 MPEG based Profile Video File server. We will compare the results of a series of experiments contrasting the quality achieved using JPEG and MPEG compression, 4:2:2 vs. 4:2:0 sampling, the impact of GOP structure, and the differences in using different types of video source material.

Picture Quality Analysis

Several years ago Tektronix, a leader in analog test and measurement equipment, began exploring the needs for test equipment in a digital television infrastructure. While there are direct corollaries to traditional baseband testing, new needs arise in a compressed digital operation. A method is needed that can help a facility identify the "operating margin" of a system. In other words, we want to know how close to the "cliff[1]" a system is operating.

When video compression is used, traditional measurement methods tend to break down. For instance, if a picture is too complex to be encoded at a given data rate, it typically "breaks" the compression algorithm and the compressed picture exhibits compression artifacts (as detailed later). These pictures have the ability to exhibit the same or even better signal to noise ratio as the input image. Yet, the picture is clearly impaired. With compressed digital signals a new method of testing is required. This method is based on what the eye sees and not what we have traditionally used to assess picture quality. The true test of picture quality is a measure of the viewer's satisfaction with the received image.

The industry has standards for subjective testing of a video signal, but these tests take a long time, require lots of people and are not at all practical for day-to-day monitoring within a facility.

[1] The cliff effect is described as the abrupt transition from good picture to no picture at all. Unlike analog systems, where picture quality degrades gradually, digital systems suffer from the cliff effect when the operating margin is inadequate.

Collaboration with Sarnoff Corporation

In 1996, Tektronix and Sarnoff Corp. started a collaboration to develop a product to measure digital video picture quality. Sarnoff is a leader in the creation and commercialization of electronic and information technologies, with a special focus in digital video. Researchers at Sarnoff have spent years studying the human visual system and applying that knowledge to television display and picture quality evaluation. Based on this work, the JNDmetrix (picture quality metric) automatically and accurately assesses the perceptual magnitude of differences between a test and reference video sequence.

Objective testing needs a valid algorithm, such as the JNDmetrix as its foundation. Naturally, implementation of a real-world measurement system must include a number of of the aspects. First, it must provide reference-scene motion sequences, a physical source for the reference scenes, and include input signal format conversion. It must also be able to test the impact of scene changes that can often overload a system operating at marginal data rates and include tests that can detect impairments due to processing in the non-compressed parts of the system. Finally, there must be accurate and repeatable alignment of the picture sequences that are used as inputs to the measurement algorithm.

The PQA200 was developed around this very concept. Utilizing a variety of industry standard test sequences, the PQA200 can provide a realistic test. The compression rate selected for a required level of picture quality depends on the type of material being broadcast. The PQA200 stimulates the system under test. The server used in this example was a Profile PDR300 MPEG based server. The test sequence was recorded on the Profile at different bit rates and then played back and captured by the PQA200. Utilizing the Sarnoff human vision system model, the PQA200 contains the three necessary dimensions for evaluation of dynamic and complex motion test sequences; spatial analysis, temporal analysis and full color analysis. The result is a Picture Quality Rating (PQR) number. It provides a quantifiable number relating directly to perceived differences between the original and captured image.

Signal vs. Picture Quality

In the mixed environment of compressed and uncompressed signals, video quality measurements consist of two parts: signal quality and picture quality.

Signal quality measurements are made using a suite of test signals. They can be full field or as short as one line in the vertical interval. Such testing in uncompressed video systems is an indirect measurement since the test is outside the actual picture. Regardless, it provides a very good characterization of picture quality.

In digital compressed video systems picture quality changes based on data rate, picture complexity and the employed encoding algorithm. The static nature of traditional test signals don’t provide a true characterization of picture quality. Therefore, direct measurement of picture quality for these systems requires a suite of actual pictures, which are much more complex than traditional test signals. These complex sequences stress the capabilities of the encoder resulting in non-linear distortions that are a function of the picture content.

However, this doesn't eliminate the need for traditional signal measurements. For the most efficient compression, the input video should be as clean as possible. Problems in the amplitude or dc level may cause clipping of the picture or inefficient use of the 8-bit, 255 signal levels. If the resolution is low (soft picture), it is actually easier to encode. However, that may not represent the desired picture quality. Any defects that make the picture more complex will be reflected in the compressed image quality. These include ringing, jitter, noise, and composite-component processing artifacts. Therefore, these problems should be removed prior to compression.

The compression encoding process is sensitive to picture material. Setting a compression rate (usually expressed as Mb/s) tells the encoder how much data or how many bits can be used to compress the picture. A key advantage of using MPEG, with its temporal compression, is that the encoder can make decisions allocating the bits over a larger number of frames. This allows more complex pictures to be coded with more bits than less complex pictures in a given group of pictures (GOP) sequence.

Visual detail is directly related to the bit-rate target assigned to the encoder. If a picture requires a lot of bits because it contains high motion (temporal) or high spatial detail, the encoder has to discard some picture detail. This information is statistically discarded to try to minimize the perceived quality loss. Eventually, enough detail may be discarded that the loss becomes very visible. These artifacts or impairments can show up in many visible ways.

Compression Related Picture Quality Impairments

To the home viewer, the quality problems of digital signals are different than analog signals.The following is a list of the types of video impairments to look for in digital signals:

Blocking: This is obvious distortion of the received image characterized by the appearance of an underlying block encoding structure.

Blurring/Smearing: In a single frame (spatial example), reducing bandwidth in the number of pixels per horizontal line, causes a blurring or smearing effect.

Edge Busyness: Distortion concentrated at the edge of objects, characterized by temporally varying sharpness or spatially varying noise.

Error blocks: A form of block distortion where one or more blocks in the received image bear no resemblance to the current or previous scene and often contrast greatly with adjacent blocks.

Mosquito Noise: This is caused by quantizing errors between adjacent pixels, which is a result of compression. As the scene content varies, quantizing step sizes change, and the quantizing errors produced manifest themselves as shimmering black dots. This looks like "mosquitoes" and show up at random around objects within a scene.

Quantization Noise: Inaccurate digital representations of an analog signal that occurs during the analog-to-digital signal processing. Typically, the digital interpretation of video resolution is limited through the digital sampling of the analog video input signal.

How the PQA200 Works

The PQA200 cannot actually measure “picture quality.” It measures the changes between the original picture and the result at the output of a system, in this case the difference between the input and the output of the PDR300.

Subjective measurement methods, such as the ITU’s Rec. 500, use a viewer’s visual impression to establish the performance of a compressed television system. Many of the test sequences used in the PQA200 have established subjective metrics. We have seen an excellent correlation between the subjective measurements and the predicted ratings from the PQA200.

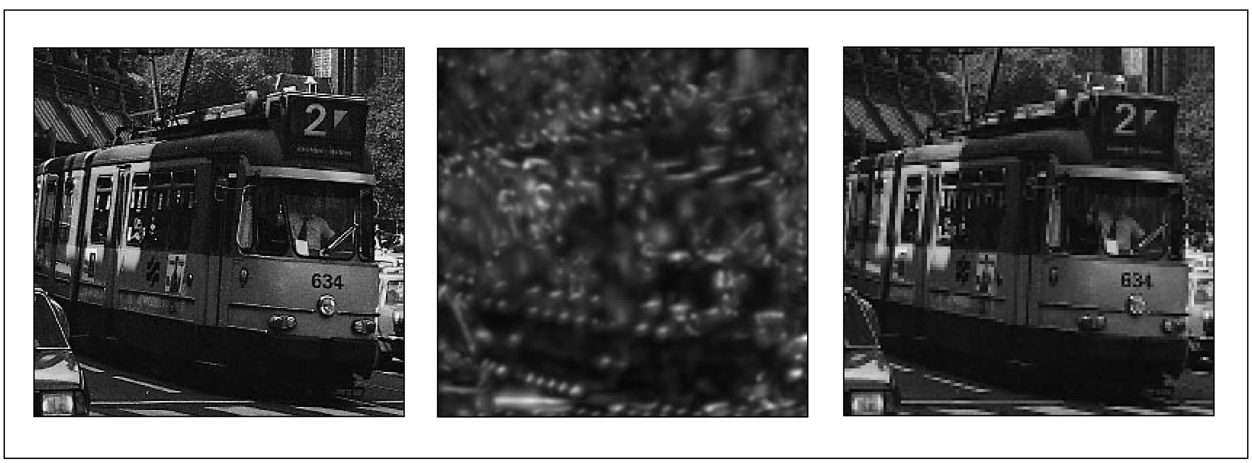
The PQA200 utilizes Sarnoff’s JNDmetrix (algorithm to evaluate the perceived differences between the original and processed picture. The pixel-by-pixel perceptual difference between the original and processed pictures is used to determine an overall objective Picture Quality Rating (PQR).

In addition to reporting PQR, the PQA200 provides animated maps. Their intensity is related to the perceived differences between the original and processed video sequences.

The PQR map indicates differences between the original and degraded pictures as seen in the example above. Brighter areas indicate greater perceptual differences. The bright area near the bottom left is due to the line on the street being distorted by the compression system. Just above that bright area is a series of dot-like bright areas due to another solid line on the street being broken into a series of dots.

The capability to view the relationship between the original material, captured material, and the PQR map provides invaluable information for evaluation and optimization of video compression systems.

The PQA200’s graphical presentation of PQR is also a very powerful analysis tool. The graphical presentation allows for quick visual identification of MPEG-2 compression GOP structure (I, B, and P frames). In addition, the graphical display mode’s cursor can be placed on any location within the graph to synchronize the views of the original video, captured video, and PQR/PSNR maps. This allows developers to explore the interrelationships between video complexity and compression artifacts both spatially (X & Y-axis) and temporally (field-by-field).



Original *PQR Map* *Processed*

PDR300 Video Server

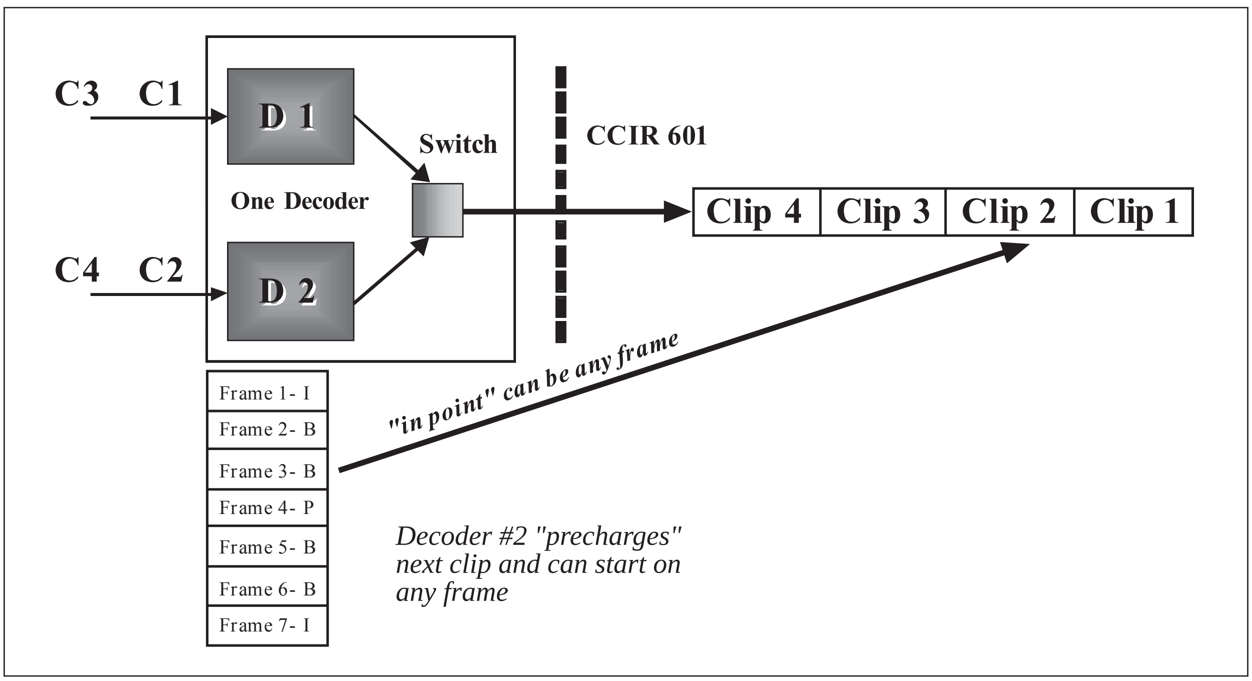
The PDR300 is a multi-channel video server using MPEG2 4:2:2 compression. It utilizes the IBM ME31 chip set. This three chip set encoder provides encoding bit rates from 4 Mb/s to 50 Mb/s, 4:2:2 and 4:2:0 encoding, and GOP lengths from I frame only to GOP of 16. The GOP sequence can be defined as I frame only, IB, or IBBP structure. The PDR300 also has two JPEG codecs to maintain backward compatibility with the over 3,000 JPEG based Profiles that are currently installed in facilities worldwide.

The PDR300 can be configured as;

- One input, two outputs
- Two inputs, four outputs
- One input, six outputs
- Eight outputs (in this configuration, compressed data is loaded via a fibre channel network connection)

The PDR300 supports both 525 and 625 line standards and can be configured for Serial Digital or Analog Composite I/O. Audio in the system is uncompressed and can be configured for analog, AES/EBU digital or embedded digital audio.

The design of the PDR300 also provides for editing of the MPEG files regardless of bit rate or GOP structure and allowing cuts-only editing on any frame within a GOP structure be it an I frame, B frame, or P frame. Also clips of different bit-rates and GOP structures can play seamlessly back to back in any sequence. Frame by frame jogging is possible both forward and backward, again with any GOP structure. This is accomplished by using two decoders, each with their own buffer memory for each output channel. As seen in the diagram below switching between decoders make the above capabilities possible. Profile also provides scrub audio. You can cue material audibly in jog and shuttle modes. This works much like an analog VTR and is available on all audio tracks.



The diagram above shows how the PDR300 uses two decoders (D1 & D2) for each output channel. When playing a series of clips (C1, C2, C3, C4), as Clip 1 (C1) is being decoded by decoder #1 and played out, Clip 2 (C2) is being decoded by decoder #2 and is ready to play starting at any frame in the GOP. As Clip 2 is playing, Clip 3 is being decoded on decoder #1.

The Tests

Three series of tests were run;

Test # 1 — Relationship between the type of video, compression rate, and picture quality

This test records 11 different video clips at 8 Mb/s, 4:2:2 and compares at the PQR number for each clip.

Test #2 — Relationship between bit rate, 4:2:2 encoding, and MPEG vs. JPEG on picture quality

This test records the famous SMPTE Mobile/ Calendar sequence from 4 Mb/s to 40 Mb/s and examines the results for both MPEG and JPEG. An analysis of comparable video quality can be made between JPEG and MPEG. An analysis of 4:2:0 vs. 4:2:2 encoding can also be made.

Test #3 — Impact of different GOP structures on video quality

This test varies the IB, IBBP sequence from I frame only to a GOP of 16.

The Results

Test #1 — Relationship Between The Type Of Video, Compression Rate, And Picture Quality.

In this test we see the impact of program content on picture quality. Our tests involved a wide array of test picture sequences, each designed to stress encoders in different ways. All sequences were encoded at 8 Mb/s, 4:2:2 sampling and a GOP of 16 (IBBP). The results show a difference in picture quality from a PQR of less than 2.4 to a PQR of greater than 6. By examining the clips we can see a common trend over the range of tests.

The first group, Susie and Lily, contain little motion and is very easy for the compression engine. With little difference between frames, more bits can be used for detail. There is very little difference between the original and the compressed picture.

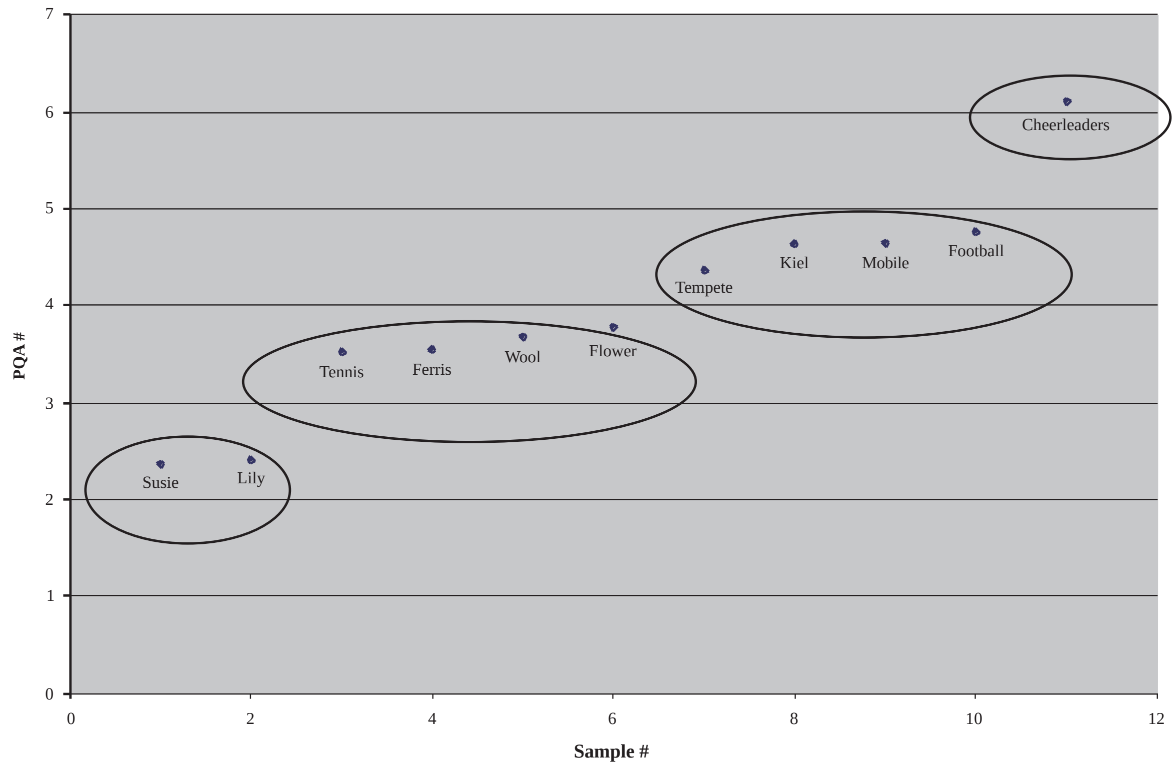
The second group contains moderate motion or a panning of the camera. The compression engine is stressed a bit more, using more bits to track the differences between frames, and thus leads to greater differences after compression.

The third group has more rapid motion and finer detail. The Kiel sequence is a pan and zoom with lots of vertical details while the football sequence contains periods of rapid motion.

The fourth group is the cheerleader sequence, which contain lots of motion and a very complex background.

The obvious results of this test shows that one compression rate does not solve all needs. In terms of costs, MPEG allows significant savings on storage costs by choosing compression settings based on the material you are encoding. Sports footage requires significantly higher bit rates than an interview show. Also clips with lots of chrominance, typical of many commercials, require a higher bit rate to maintain quality, something advertisers will be looking at with a keen eye.

Various Clips @ 8Mb/s 4:2:2



Video Test Sequences

Video Test Sequence	Motion	Characteristic
Susie	Slow	Skin Tone, talking head
Lily	Still	Luminance Resolution
Tennis	Pan	Multiple random motion, sports
Ferris	Fast Complex	Luminance and color details
Wool	Medium	Moving colors
Flower	Slow Pan	Color details, landscape
Tempete	Random Motion	Horizontal, vertical, luminance, color detail
Kiel	Zoom	Luminance detail, landscape
Mobile	Slow	Random motion of objects, color detail
Football	Random motion	Sports, busy, large objects
Cheerleaders	Rapid motion	Fast complex sports, rich background

Test #2 — Relationship Between Bit Rate, 4:2:2 Encoding, and MPEG vs. JPEG on Picture Quality

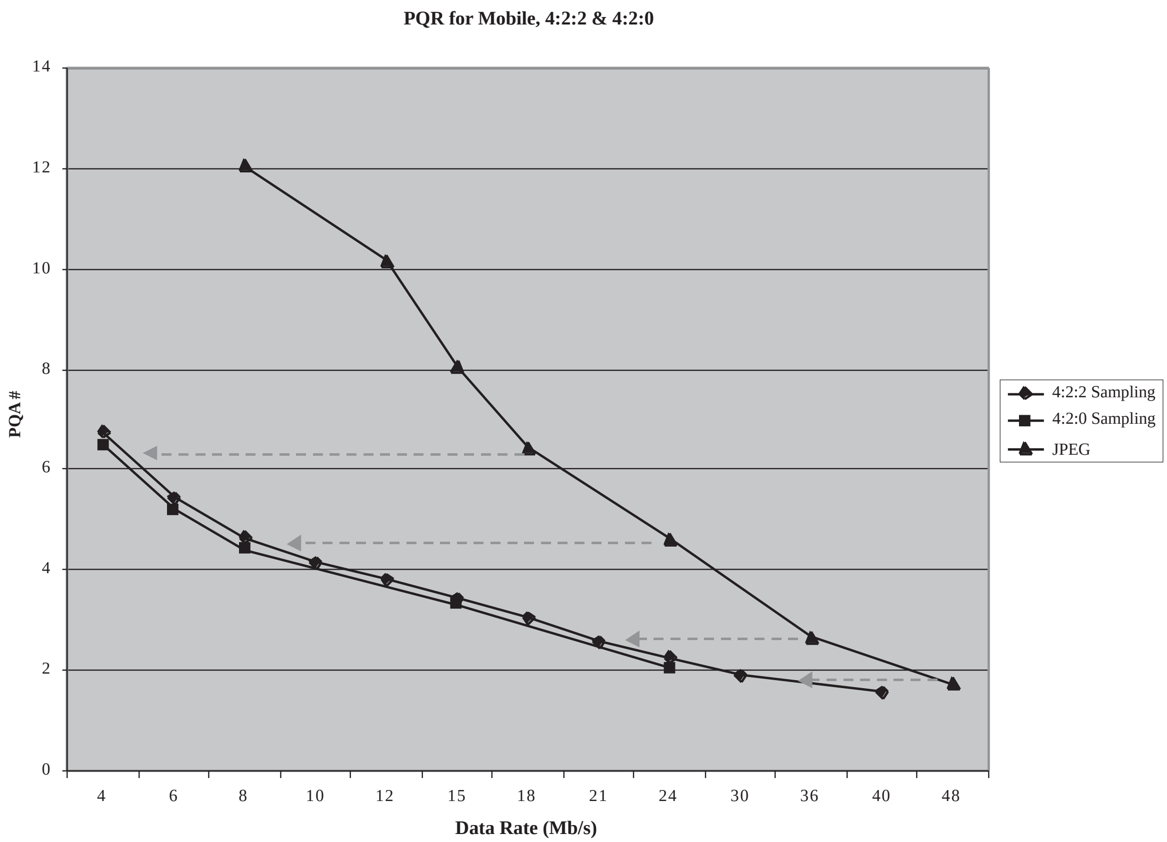
Keeping the test sequence constant, we can now see the impact of compression on picture quality.

First, all compression systems have a knee in their picture quality curve. Below the knee in bit rate; the picture quality begins to drop dramatically. In a JPEG system, the knee is around 18 Mb/s. For MPEG, it's around 8 Mb/s. And, as expected, at higher bit rates, you get diminishing returns. However, keep in mind that higher bit rates should be used for any production work that requires multiple generations or when the material will be subject to up-conversion to HDTV. Low bit rates extract too much information to allow multiple passes through a codec.

Second, we see that 4:2:0 sampling provides a slightly better picture below 10 Mb/s. This is to be expected as 4:2:0 provides less chroma information to encode. This means that there are more bits allocated to code the luminance information — this has a double impact since not only is the luminance represented more faithfully, the eye is less sensitive to the loss of chroma information than it is to loss of luminance. However, one should remember that 4:2:2 is already optimized to take advantage of the different characteristics of the eye's response to chroma and luma - reducing chroma in half does have an impact on image quality, which is especially evident in a production application. On a sequence with more chrominance variance, the difference will be greater.

Lastly, the results of this test provide a good guideline for the comparison between JPEG and MPEG. The Picture Quality Rating summarizes it succinctly.

- 48 Mb/s JPEG = about 33 Mb/s MPEG
- 36 Mb/s JPEG = about 21 Mb/s MPEG
- 24 Mb/s JPEG = about 8 Mb/s MPEG
- 18 Mb/s JPEG = about 4.5 Mb/s MPEG



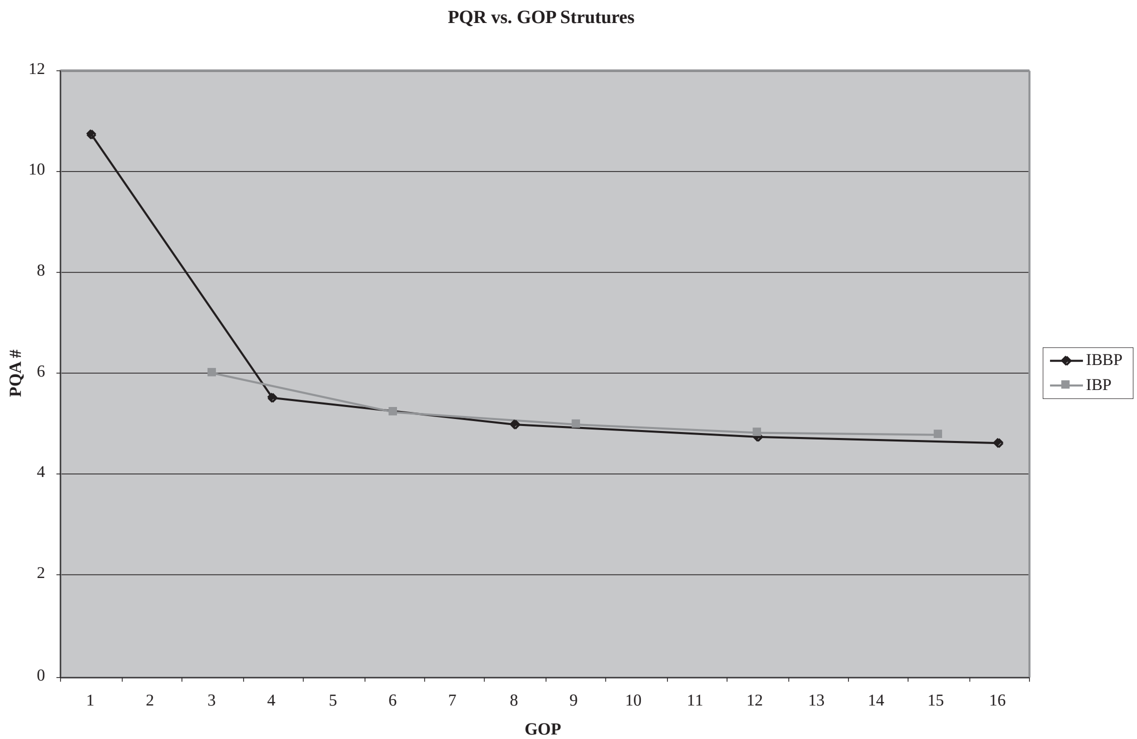
In the chart above, the dotted arrows indicate the equivalent PQR numbers for JPEG and MPEG

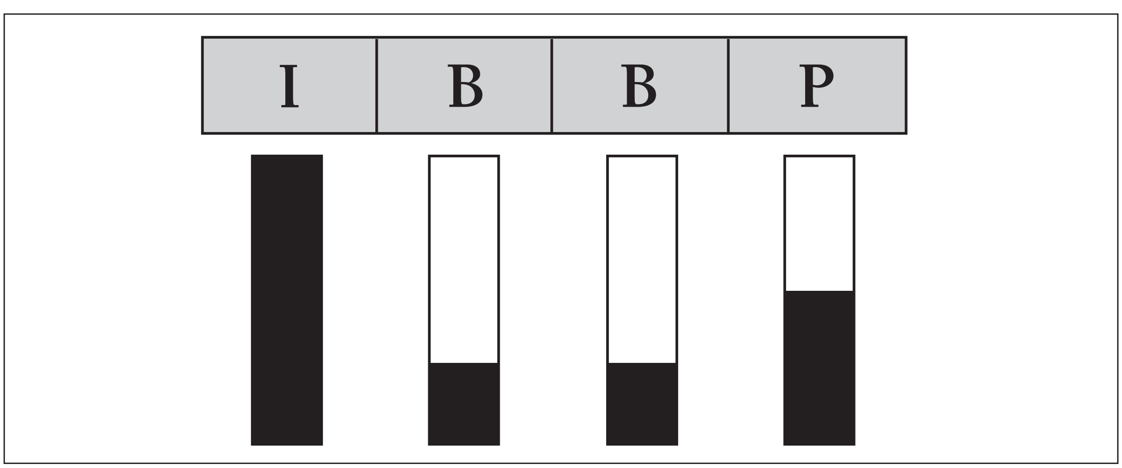
Test #3 — Impact of GOP on picture quality

This test examined the impact of GOP structure and length. Again using the Mobile/Calendar sequence at 8 Mb/s, we see little picture quality difference once you go past a GOP of 4-6, however there are significant storage efficiencies that may be recognized at the longer GOP. Also notice there is very little difference between IBP and IBBP encoding but again IBBP encoding will provide more efficient storage.

This test also shows the dramatic impact of GOP on picture quality. From I frame only to a GOP of three, there was dramatic improvement in the picture. But after that, picture improvements were minimal.

As a test sequence becomes more stressful, this curve will take longer to flatten out, as temporal compression becomes more difficult.





In a Group of Pictures (GOP), a B frame contains approximately 25% of the information (as indicated by the shaded area) of an I frame and a P frame contains about 50%.

Summary

This paper has quantified some basic attributes of MPEG compression. Many of the results were as expected, but without a quantifiable test capability, it would be difficult to prove the same results based solely on visual observation.

The results offer some basic guidelines when moving into an MPEG 4:2:2 compression format.

1. Initially, select a video file server or compression engine with the ability to vary compression rates and encoding capability.

2. It's important to understand the type of video you will be broadcasting and select the appropriate compression settings.

3. There is little benefit in choosing IBP encoding over IBBP encoding. IBBP provides more efficient storage.

4. There is also little benefit in choosing shorter GOP structures for program playout. Choose the longest GOP possible as the storage benefits far outweigh the picture quality improvements. I frame only and short GOP structures are useful for editing applications which need flexibility in performing insert edits.

5. Select 4:2:2 encoding whenever possible, especially when you know the clip will either be upconverted or will experience multiple generations. The 4:2:2 signal will maintain it's quality better through the production process.

   However if you do need to run at very low bit-rates, the ability to select 4:2:0 encoding is a big advantage.

6. Finally, having the ability to measure your digitally compressed signals and gather similar information that you were used to having with your analog signals is critical in the setup and running of a digital infrastructure. Products like the PQA200 take the guesswork out of video quality tradeoffs and enable you to maintain high quality images out of your facility

References:

1) A Guide to Picture Quality Measurements for Modern Televisions Systems — Tektronix #25W-11419-0

2) JNDmetrix™ is fully described in a paper at HYPERLINK
http://stdsbbs.ieee.org/groups/videocomp/jnd-1.zip
http://grouper.ieee.org/groups/videocomp
Filename: jnd-1.zip

3) PQA200 Data Sheet = HYPERLINK
http://www.tek.com/Measurement/Products/catalog/pqa200
http://www.tek.com/Measurement/Products/catalog/pqa200

4) PQA200 Application Note = HYPERLINK
http://www.tek.com/Measurement/App_Notes/PicQuality
http://www.tek.com/Measurement/App_Notes/PicQuality

5) IRT / Tektronix Comparison of Subjective and Objective Picture Quality Rating of MPEG-2 Video Report = HYPERLINK
http://stdsbbs.ieee.org/groups/videocomp/tek-irt.zip
http://grouper.ieee.org/groups/videocomp
Filename: tek-irt.zip

6) MPEG-2 Fundamentals for Broadcast and Post-Production Engineers — Tektronix # 2AW-1061 or HYPERLINK
http://www.tek.com/VND/PDFs/MPEG2.PDF
http://www.tek.com/VND/Support/White_Papers/Profile/mpeg2.html

7) Profile Video File Server - HYPERLINK
http://www.tek.com/VND/PDFs/pdr200ds.pdf
http://www.tek.com/VND/Products/Profile/pdr200.html

8) Profile MPEG Support - HYPERLINK
http://www.tek.com/VND/Products/Profile/pdr200-ug-ds.html
http://www.tek.com/VND/Products/Profile/pdr200-ug-ds.html

2AW-8033 FL5505 8/98