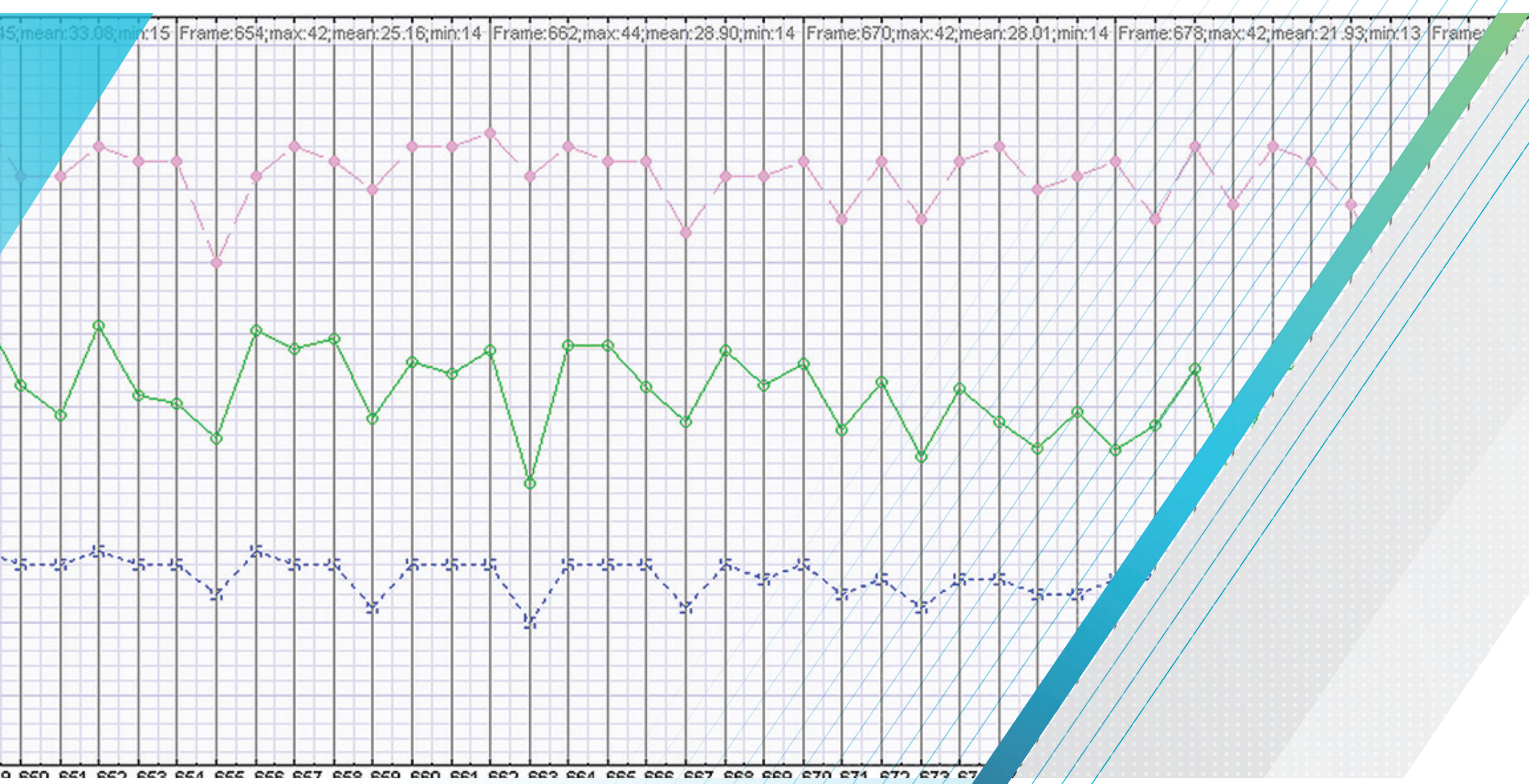
Ultra HDTV Testing Through MTS4EAV7 and PQA600B

APPLICATION NOTE



Tektronix®

With an increasing diversity of services, the growing popularity of HD video, and the emergence of beyond-HD formats (e.g. 4k×2k or 8k×4k resolution), there now exists a stronger need for coding efficiency superior to H.264/MPEG-4 AVC's capabilities. There is an increased desire for even higher quality and resolutions in mobile applications also.

High Efficiency Video Coding (HEVC) or H.265 is expected to deliver the same visual quality as Advanced Video Coding (AVC) or H.264/MPEG-4 while using about half the bandwidth. It can also support ultra-high definition resolutions such as 4K and 8K. This will allow content distributors to deliver HD video to mobile devices over mobile data networks at reduced bandwidth costs, as well as provide stunning 4K and beyond video to the home. HEVC has been designed to address essentially all existing applications of H.264/MPEG-4 AVC and to particularly focus on two key issues: increased video resolution and increased use of parallel processing architectures.

Efficient encoders are critical for achieving the bitrate reductions of 50% as compared with AVC and 70% as compared to MPEG-2. During the encoder design phase, designers need a tool which can help them make decisions in choosing the right encoder algorithm, or which can compare two different encoder algorithms and compares results for the given quality level of video. Also, when choosing the right video encoder and comparing many encoders you'd need statistics from many encoders for the same quality of video, whether each of them is producing the compliant bit-stream, if there are buffer violations while decoding that can affect the decoder and many more. The MTS4EAV7 AVC/HEVC analyzer verifies the compliance of the encoded video and provides statistics of the coding efficiency and helps encoder manufactures to debug and validate encoder designs.

Once we have determined that the compressed video stream is syntactically and semantically correct (legal according to the HEVC codec standard), we now need to look at the visual differences between the original video and the decoded HEVC video. To do so, we have many different choices and algorithms. The most basic is to simply play the source and decoded video on two different video monitors and ask human viewers to rate any visually differences. This method is highly subjective as each viewer will use his or her own personal preferences to determine video quality. The bad thing about using real humans is that we can occasionally say that two video clips are slightly different, when in fact they are identical. Also, we occasionally say that two videos look identical when in fact those are slightly different. The best method when using human viewers requires a rigidly controlled viewing room with a large population of human viewers sampled in order to gain valid statistical results. Because of the high cost of such a test, it is not preferred over the use of mathematical models used to determine video quality. The most common models used today are Peak Signal-to-Noise Ratio (PSNR), Differential Mean Opinion Score (DMOS), and Picture Quality Rating (PQR).

The PQA600B and PQASW support not only PSNR, DMOS, and PQR, but also a large offering of many other mathematical models, as well as the ability to customize each model based upon different viewing conditions (e.g., lighting, distance from TV, etc.). In the case of testing HEVC video, we will use the MTS4EAV7 to decode the compressed video file into one-byte or two-byte per sample YUV files (8-bit Vs. 10-bit HEVC). The PQA600B and PQASW can then objectively compare the original or reference YUV video file against the YUV file from an HEVC encoder.

This application note will first review the MTS4EAV7 product and how it can be used to determine interoperability as well as the HEVC schemes used during the encode process. We will also use the MTS4EAV7 to export the HEVC file to a YUV file. Finally, we will review the PQA600B and PQASW application to compare the reference and test video files. Normally, the reference video file would be from an uncompressed YUV source, but it is also possible to compare a lightly compressed file against a more heavily compressed video file.

Figure 1. MTS4EAV7 HEVC tool tip icons.

MTS4EAV7 Testing

The MTS4EAV7 application can test every frame of video against the compliance requirements defined in the HEVC standard, as well as many other compression standards. Any incompatibilities will be listed in the Alert Log. As the video frames are stepped though from 1 to N, the Alert log will keep track of any issues. The goal is to run or step through every frame without encountering any incompatibilities. Incompatibilities detected in the file denote that the file is not 100% compatible with a decoder standard. An Alert log without alerts means that the file is fully decodable by all compliant decoders. During the Step Frame or Play process, many additional display features are available to view each of the many choices that were made by the HEVC encoder. These additional graphical features will be discussed in the following pages.

HEVC Coding Efficiency

HEVC is a block based codec, consisting of bigger blocks of “Coding Tree Units (CTU)” (analogous to macroblock) which has the size selected by encoder. The encoder decides to use larger coding units where there is little or no encoding needed (more compression) and picks smaller coding units where the encoding is needed. The size of coding units can vary from 16, 32 or 64 samples, with large size enabling better compression. MTS4EAV7 offers many tool-tips to analyze the different coding parameters at Coding Tree Units, Coding Units and Luma and Chroma blocks. Figure 1 shows each of the tool tips.

These tool tips get enabled when the HEVC stream is opened with MTS4EAV7. The first of the many tool tips are show in Figure 2.

A video frame in HEVC consists of many coding tree units, which consists of coding units. Any coding tree unit or coding units will have different encoding parameters chosen by encoder, which can be observed with different tool tips

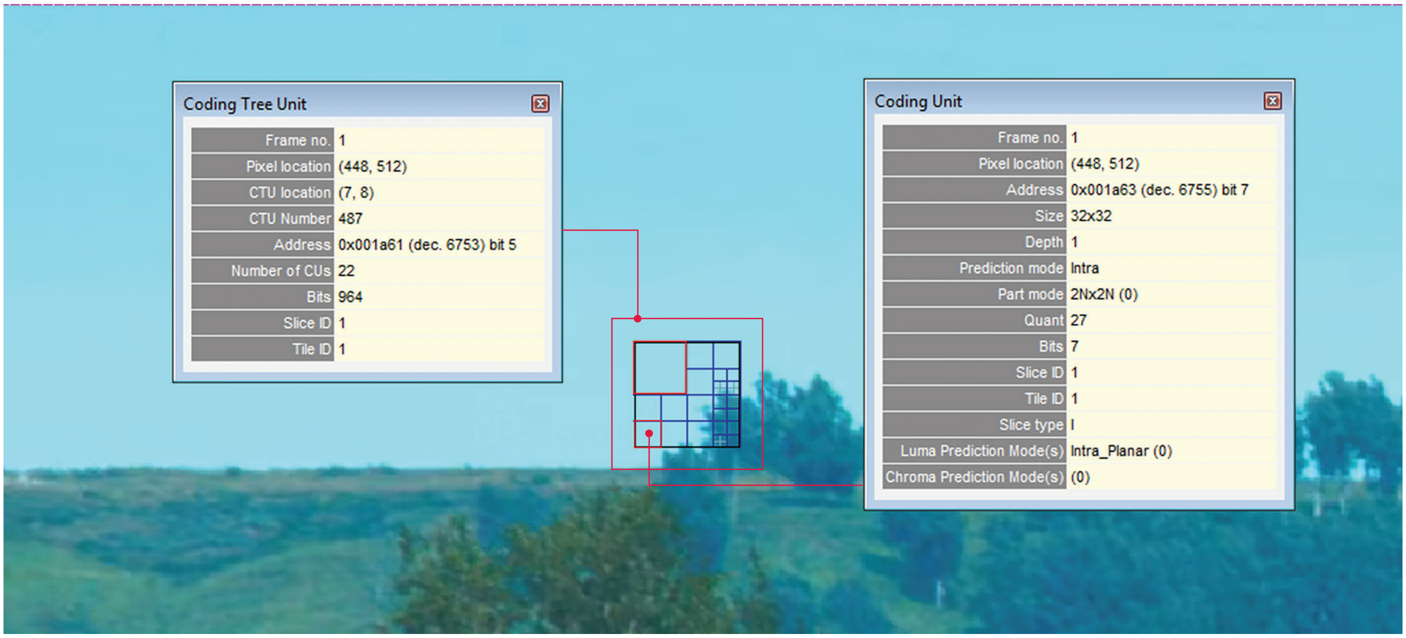


Figure 2. Coding Tree Unit and Coding Unit Tool tips. Video Scaled 400% (4k HEVC file from Elecard).

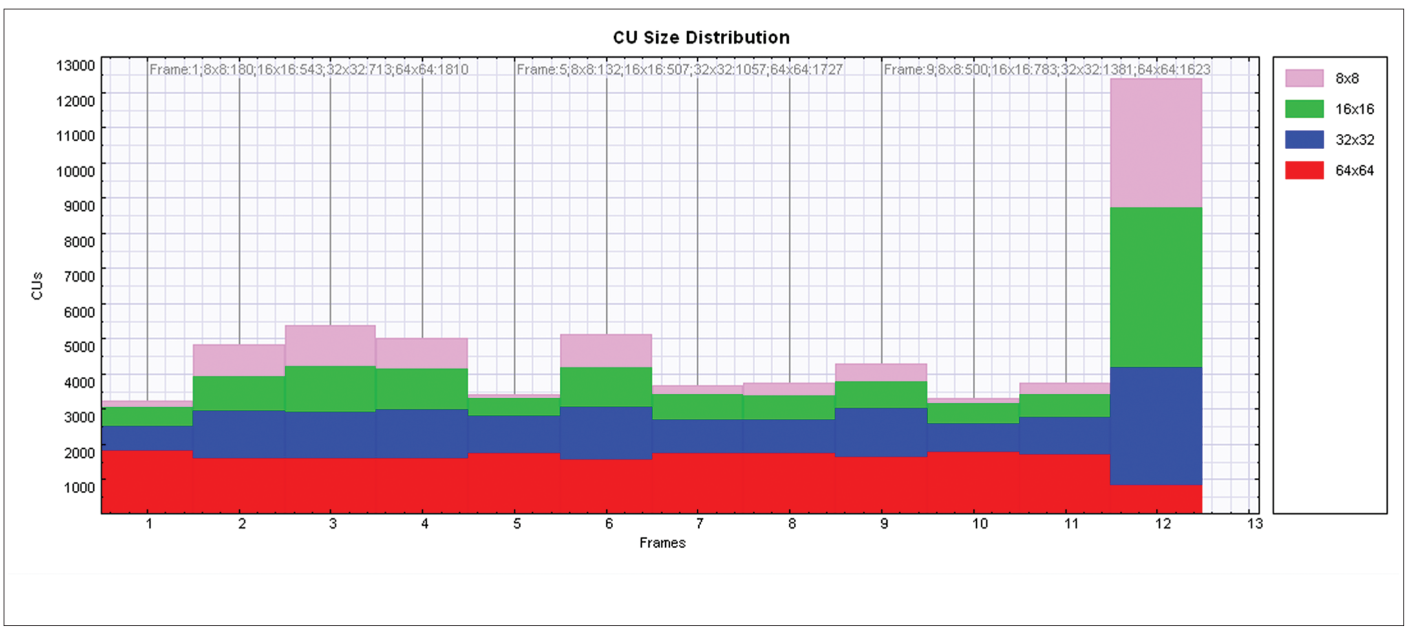


Figure 3. Varying Coding Unit sizes for each video frame.

provided in MTS4EAV7. The above picture shows different parameters such as Bits, Partition Modes, Luma Prediction modes, etc. The CU tooltip for Intra and Inter Prediction modes carry different information relevant to different prediction modes.

The CU Size distribution graph, shown above in Figure 3, indicates the different sizes of CU chosen by encoder. In the case of a uniform area of the picture, the higher the size of coding units, the better is the compression. It is expected that more complex scenes will use smaller CU sizes because they can more efficiently encode fine details compared to the large CU blocks.

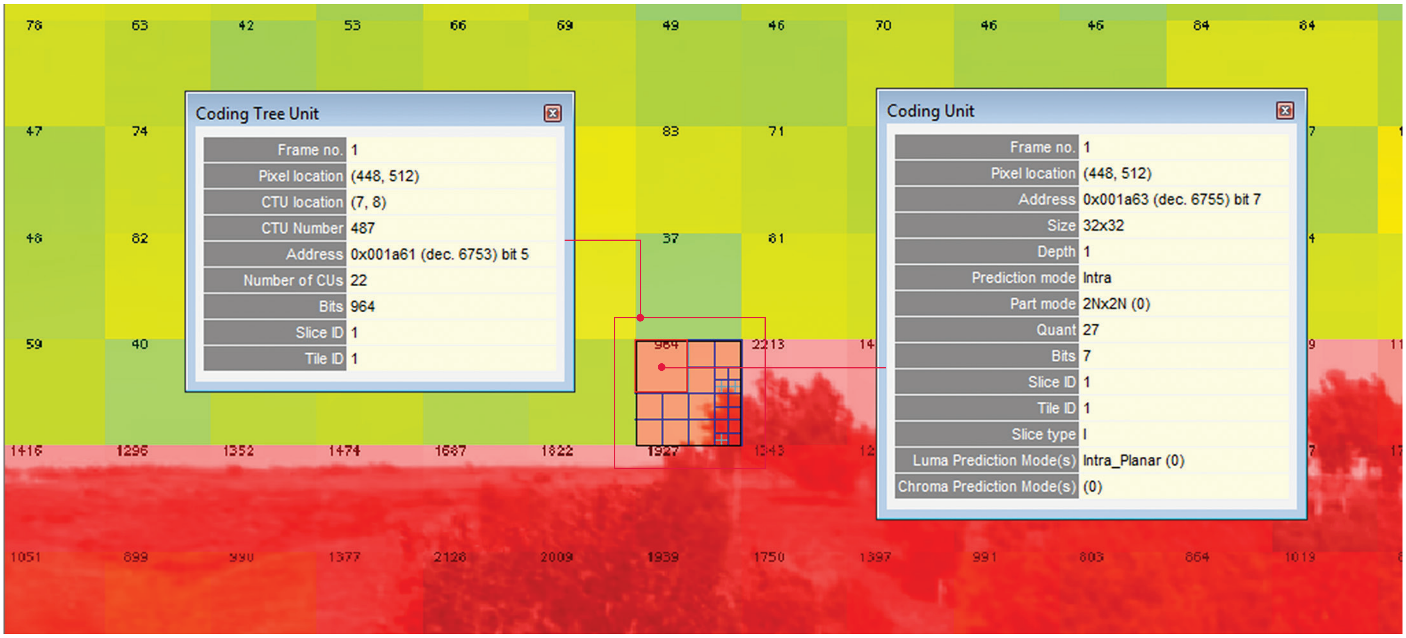


Figure 4. Picture Showing Bits overlay over CTU (same scene as in Figure 2).

Overlays

MTS4EAV7 offers different overlays at the frame layer and also at the bit-stream layer over the entire elementary stream which would help collect different statistics such as Bits, Quants, Slice and Tiles. Choose Overlay -> CTU/MB Statistics -> Bits to choose Bits per CTU. The example above in Figure 4 shows the Bits per CTU overlay. An additional color scheme is overlaid to more easily see which CTU's use the least amount of bits (green) while other colors show which CTU's use the most amount of bits (red).

The overlays for bits have an option to differentiate Luma and Chroma bits for each CTU. This would act as an indication of the bits being consumed for each CTU. Users can navigate to an interested CTU using zoom-in and out functionality provided with mouse controls. The encoder intelligently picks and chooses the type of coding units, allocates required number of bits for a given CTU. Each coding units can be Intra coded, Inter coded or Skipped for encoding. Intra coded blocks would consume more bits while Inter and skip CU would have less number of bits consumed. The encoder decision to choose different types and sizes can be observed, reviewed with CU type tooltip and Bits Overlay in MTS4EAV7.

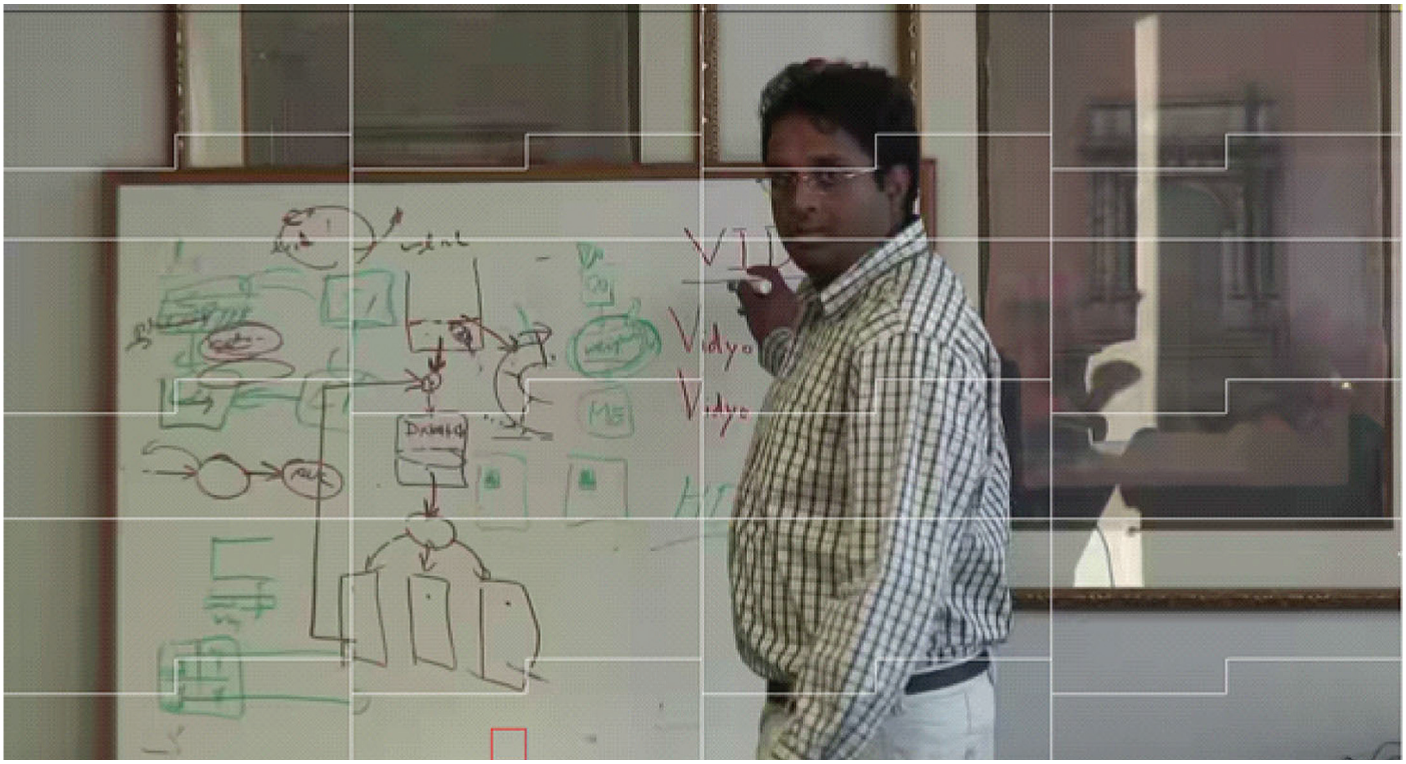


Figure 5. MTS4EAV7 showing HEVC slice partitions.

Alert Log: merge_mondial_3840x2160p5994_420_8Mbps(1).hevc

Level	ID	Source	Class	Title	Decode	Display	Details
Error	72557	Video	NAL unit	Zero Byte not present	1	1	The first NAL unit of an access unit in decoding order shall contain zero_byte.
Error	72554	Video	RBSP.VPS	Incorrect Profile Compatibility Flag	1	1	general_profile_compatibility_flag [1] is set to 0, for VPS with ID 0. The standard specifies it t...
Error	72585	Video	RBSP.SPS	Incorrect Maximum Decoder Picture Buffering	1	1	sps_max_dec_pic_buffering_minus1 [0] is 3, & vps_max_dec_pic_buffering_minus1 [0] is 0. T...
Error	72555	Video	RBSP.SPS	Incorrect Profile Compatibility Flag	1	1	general_profile_compatibility_flag [1] is set to 0, for SPS with ID 0. The standard specifies it t...
Error	72557	Video	NAL unit	Zero Byte not present	2	>1	The first NAL unit of an access unit in decoding order shall contain zero_byte.
Error	72557	Video	NAL unit	Zero Byte not present	3	>1	The first NAL unit of an access unit in decoding order shall contain zero_byte.
Error	72557	Video	NAL unit	Zero Byte not present	4	>1	The first NAL unit of an access unit in decoding order shall contain zero_byte.

Total: 7 alerts | All alerts shown

Figure 6. Alert log indicating HEVC standard violations of the given bit-stream.

Slices and Tiles

The Slice overlay allows you to visually see the slice partitions in the frame. Figure 5 shows how the HEVC encoder divided the frame into smaller Slices and Tiles.

Bit-stream Compliance

MTS4EAV7 checks for HEVC compliance to indicate all possible violations of the bit-stream against the standard. Hence it is expected that any bit-stream which gets encoded from an encoder must comply to the HEVC standard, the violations are reported as alerts which are defined at many different bit-stream layers. The alerts can be saved to XML file format for further analysis.

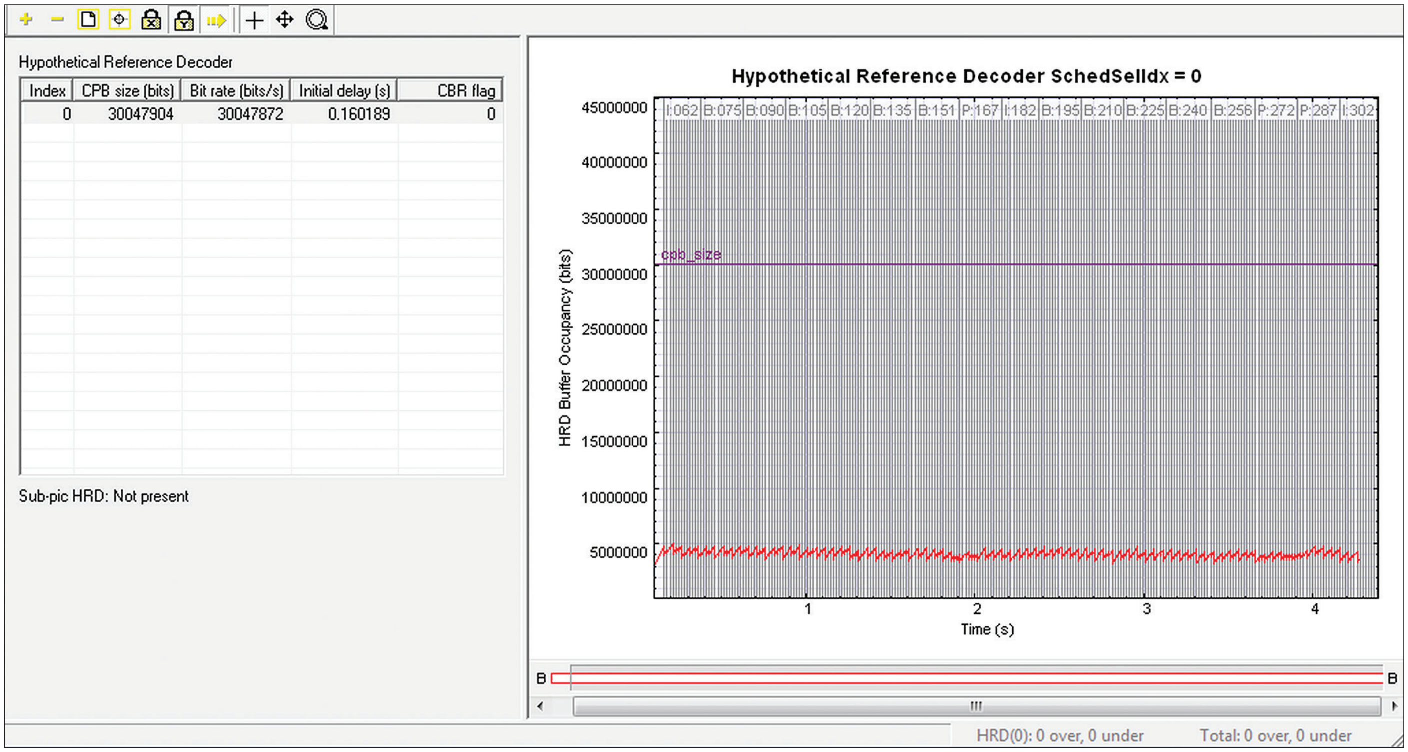


Figure 7. HRD buffer analysis.

Buffer Analysis

A bit-stream is said to be compliant when it doesn't overflow or underflow the decoder buffer memory requirements. Hence any HEVC bit-stream for the profile and level indicated in bit-stream, must adhere to the buffer requirements mentioned in standard. The Hypothetical Reference Decoder (HRD) provides a way to check for decoder buffer requirements and decoder buffer occupancy.

When the HEVC stream carries HRD values encoded, MTS4EAV7 checks if bit stream overflows or underflows according to Annex C of the HEVC standard.[1] The above graph in Figure 7 indicates different buffer values and buffer occupancy level for each frame of the video stream. The buffer violations, if any are reported in the Alert Log as part of the compliance.

[1] (ITU-T H.265 04/2013 High efficiency video coding).

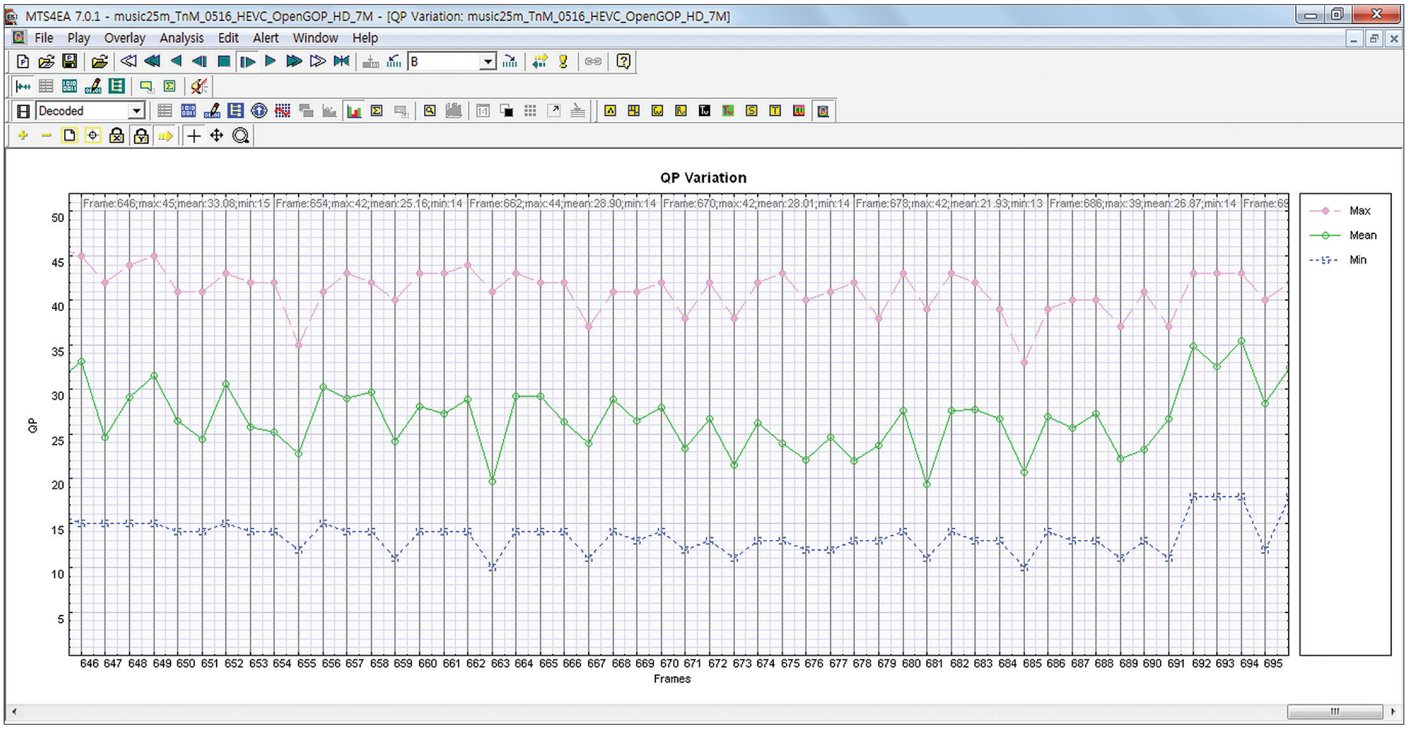


Figure 8. QP Variation graph.

QP Variation

QP indicates Quantization Parameter, which is again the indication of the rate control mechanism of the encoder to check how the encoder is performing to the different bitrates. HEVC encoders choose to allocate different QP values based on frame type (I, P or B) and then use the less number of bits for the CU level. As the bitrate is reduced, the encoders control the quantization levels applied and the size of the CTUs. QP variation graph indicates the rate control mechanism of the encoder and also captures max, min and mean QP variation per frame for the entire video duration. Figure 8 shows the Min, Mean, and Max of the QP variation.

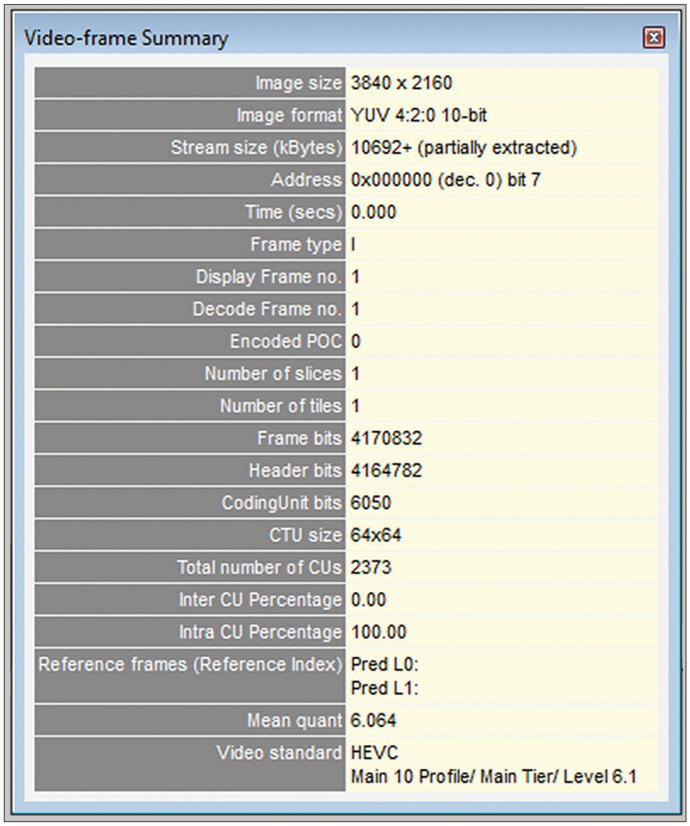


Figure 9. MTS4AV7 Summary showing 4k format in 10-bit mode.

MTS4EAV7 PSNR Testing

If the source is uncompressed video and is available for reference, then MTS4EAV7 is able to produce the fidelity analysis measurements to further quantify the degradation of the original video.

Exporting 8 or 10-bit HEVC Video to a YUV File

The MTS4EAV7 has always provided the ability to export the compressed 4:2:2 or 4:2:0 video to a YUV file. With HEVC, the compressed video will usually be in the 8-bit or 10-bit form (see Figure 9).

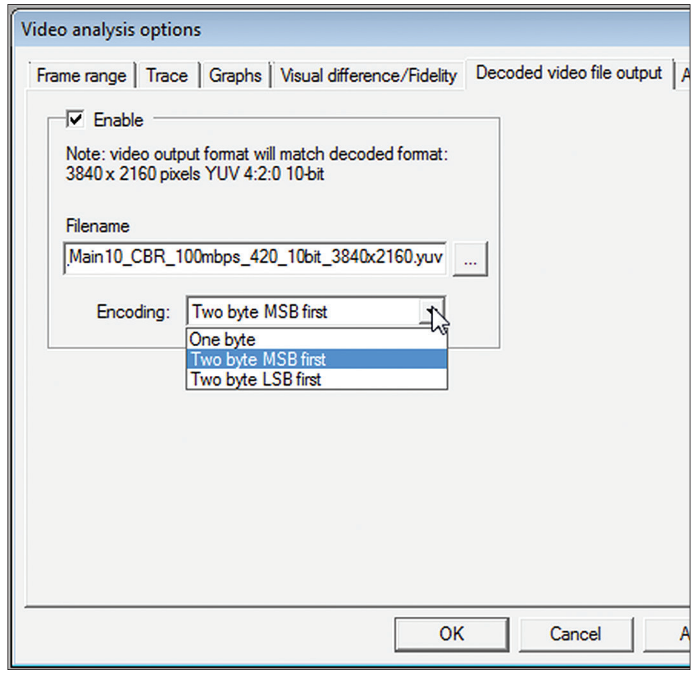


Figure 10. Exporting decoded video to a YUV file.

In the case of the HEVC 10-bit video, the MTS4EAV7 export function will default to two-byte MSB file format. The same file format is also available in the current PQA for YUV import (see Figure 10). For our Picture Quality testing examples, we will use two different sets of files. The first set with the spotlight are both 10-bit 4k, while the second set from Elecard are 8-bit 4k and 1080p HEVC files.

To export the video from a 10-bit HEVC source, do the following:

1. Open the video file
2. Select: Analysis-Output decoded video file…
3. Accept the defaults and select OK
4. Step the video frame by frame, or Play to step continuously
5. Press Stop to close the YUV video file, or wait to reach the end of the file

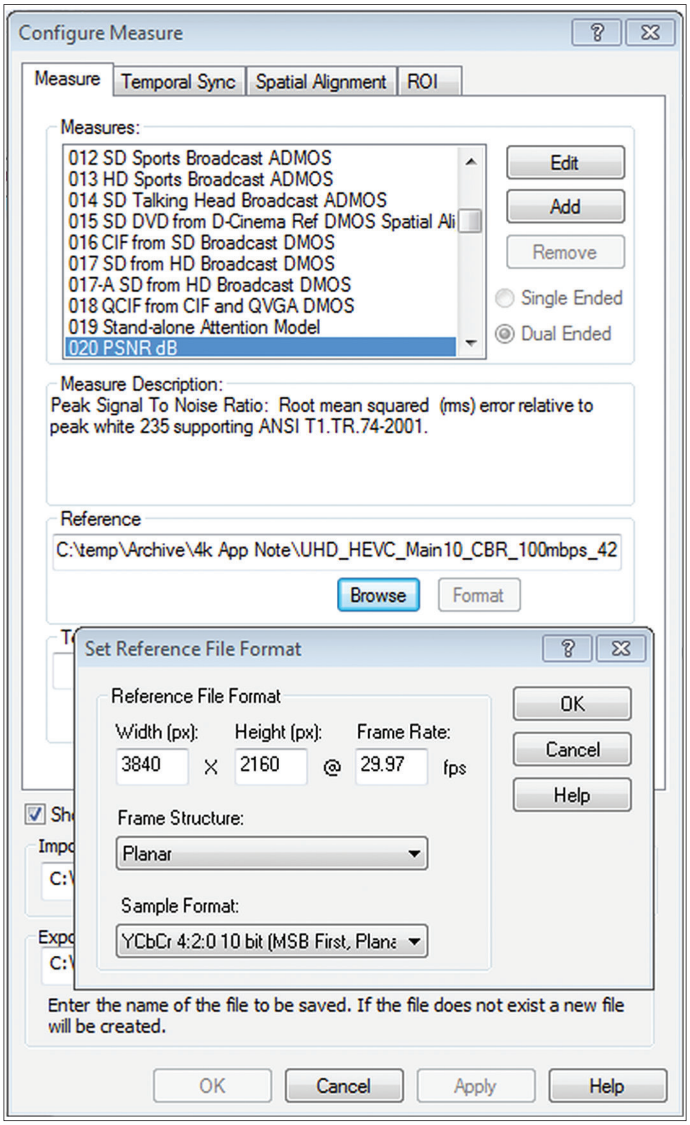


Figure 11. PQA Configure-Measure menu.

Picture Quality Testing

Both the PQA600B and PQASW applications support many different video formats including the newest Ultra HDTV formats (4k and 8k). In the case of HEVC, we will use the MTS4EAV7 application to export 8 or 10-bit YUV video to disk. Within the PQA application, the reference video file is usually from an uncompressed video source, but the reference can also be from a decoded YUV file from a lightly compressed source file.

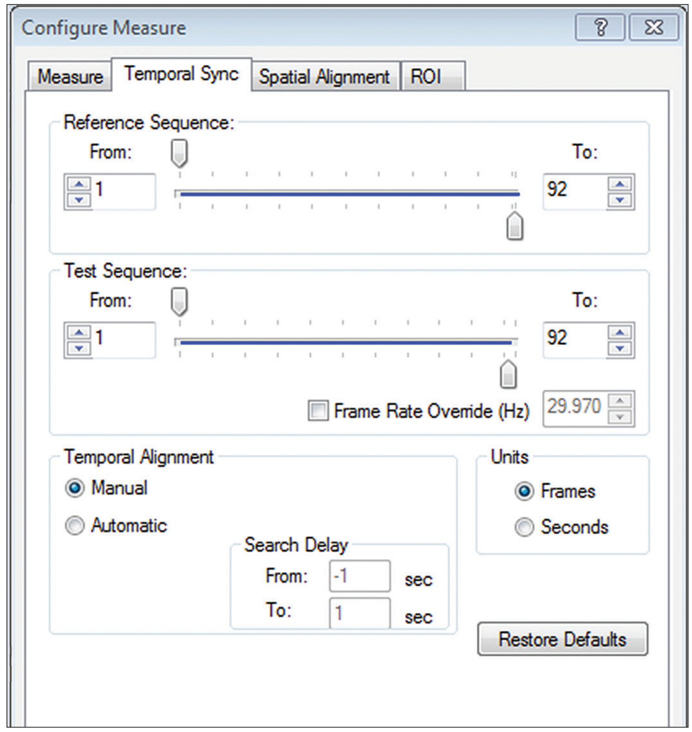


Figure 12. Temporal Frame Sync.

PQA PSNR Testing in 4k

The first step in this test is to select the PQA Configure-Measures menu. Figure 11 shows that only the PSNR template is selected out of the many possible test models. The next step is to select the HEVC YUV Reference file, as well as the Test file. In both of these cases, we will use a pair of 4k files exported from the MTS4EAV7 application. It is very critical to enter the correct horizontal, vertical, rate, structure, and sample format. The YUV file format does not include a header with these details, so it is best to refer to the text in the file name created by the MTS4EAV7 application.

After selecting PSNR and the file formats, is may be necessary to temporally align the Reference and Test frames. Figure 12 shows that there are two methods of temporal alignment. One is manually stepping through the Reference and Test frames looking for a unique scene in both files. This is easiest around a scene-change. The second method is to use the Automatic mode which incrementally slips each file one frame forward or backward while looking at the differences between the two frames. Once a minimal point is found in both files, the offsets in each will be assigned.

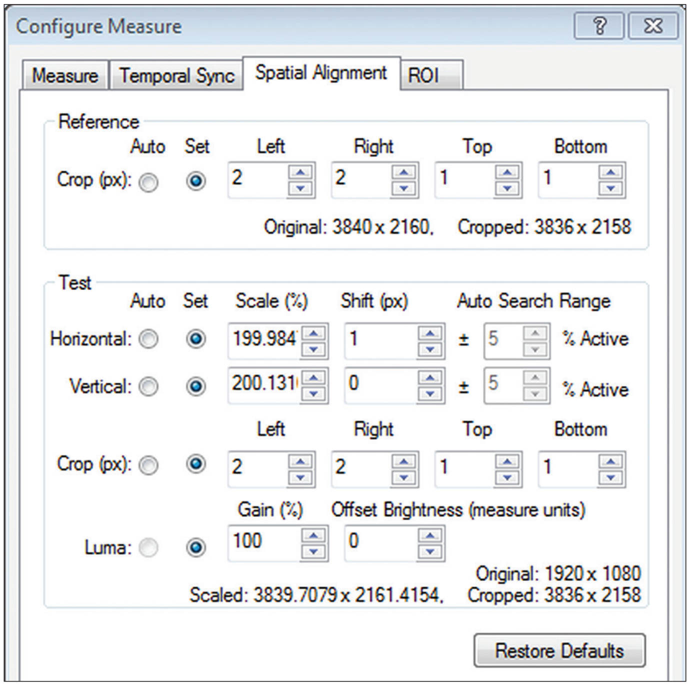


Figure 13. Spatial Alignment.

The last thing to check before starting the measurement process is the optional Spatial Alignment menu as shown in Figure 13. This is usually not needed, but occasionally, an encoder may shift the encoded video frame up, down, left, or right by a few pixels. Also, the Spatial Alignment menu will be required if the Reference and Test files are of different sizes (e.g., 3840x2160 Vs. 1920x1080). In these cases, and after Temporal Alignment has been set, applying the auto alignment mode on all parameters will cause the Test file to be scaled and shifted to best match the reference file. In the 4k Vs. HD case, the Test file would be scaled by 200% and can be seen in Figure 13 and Figure 15.

Once the PQA measurement has been configured, we can exit the menu and select the Measure icon shown in the upper left corner of Figure 14. The Reference video will be shown in the upper left quadrant. The Test video will be shown in the upper right quadrant. The pixel differences for each frame will be shown in the lower right quadrant. Finally, the overall PSNR results will be shown in the lower left quadrant.

The PSNR measurement gives a good mathematical result of the pixel differences summed over each frame, and also over entire file. The result shows how close the Test file is to the Reference file. A score of 80 dB or more means that the two files are virtually identical. A score of zero means that the two files are not related at all. A score greater than 40 dB, as in our example in Figure 14, shows that the two files are very closely related. We cannot compare a PSNR result to human viewer trials as a PSNR result does not take into account how the human brain and eyes work together. Therefore, a good PSNR score tells us that the two files are close, but we need to take into account the human vision model before we objectively rate the quality of these two files.

In the previous case, we looked at comparing two 4k Ultra HDTV video clips. In the next PSNR case, we will compare a 4k Ultra HDTV Reference file against a broadcast-ready 1080p file. You will notice that the PSNR values will be a bit higher than the previous 4k Vs 4k test (Figure 15). This is to be expected as our 1080p Test file is only one quarter the size of its 4k Reference file.

The PSNR results are still very good at about 40 dB, but it would be impossible to make it to 80 dB with three quarters of the pixels removed from each frame. Again, the right Test image had to be scaled up to 200% before comparing against its 4k Reference file.

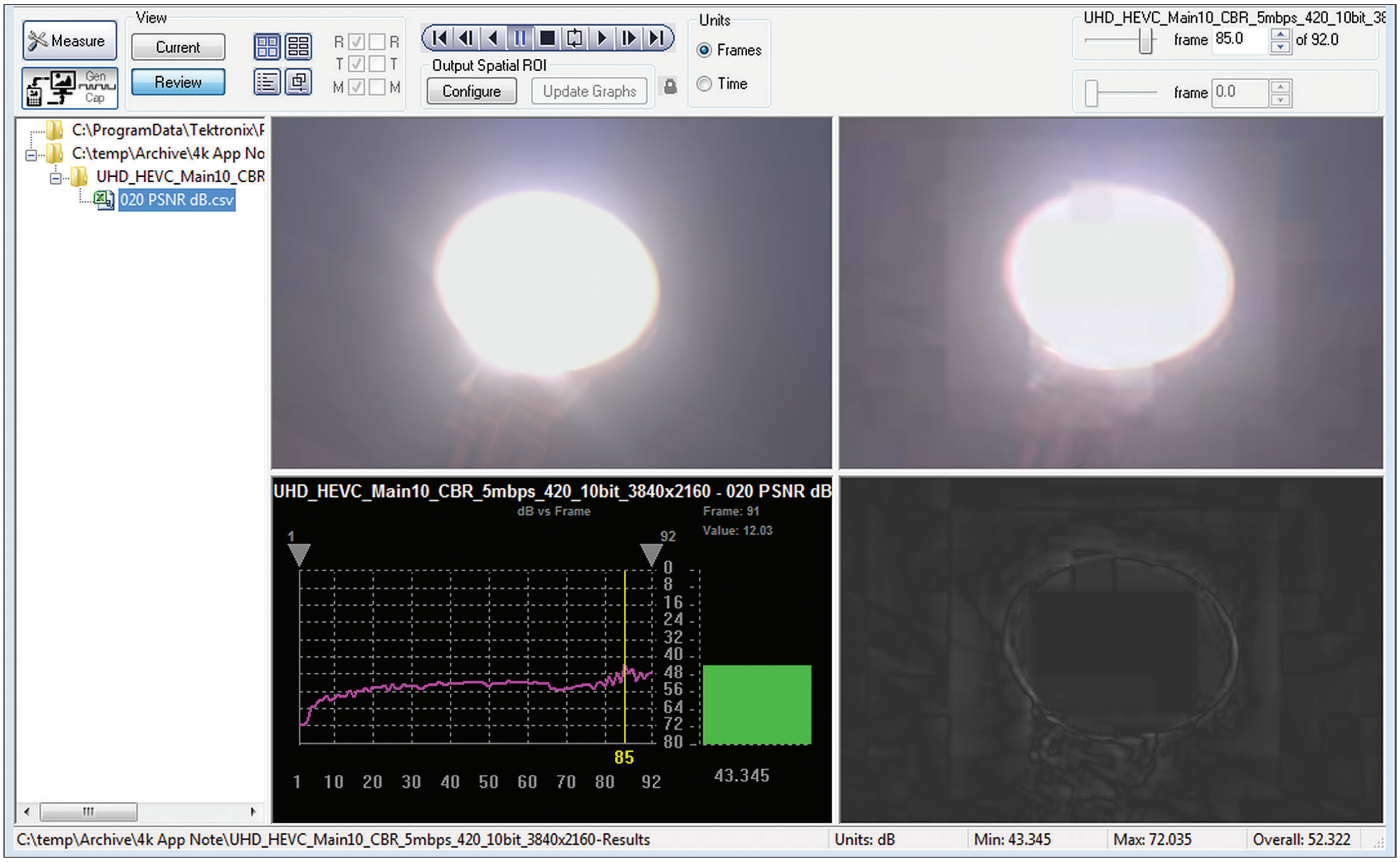


Figure 14. PQA PSNR results.

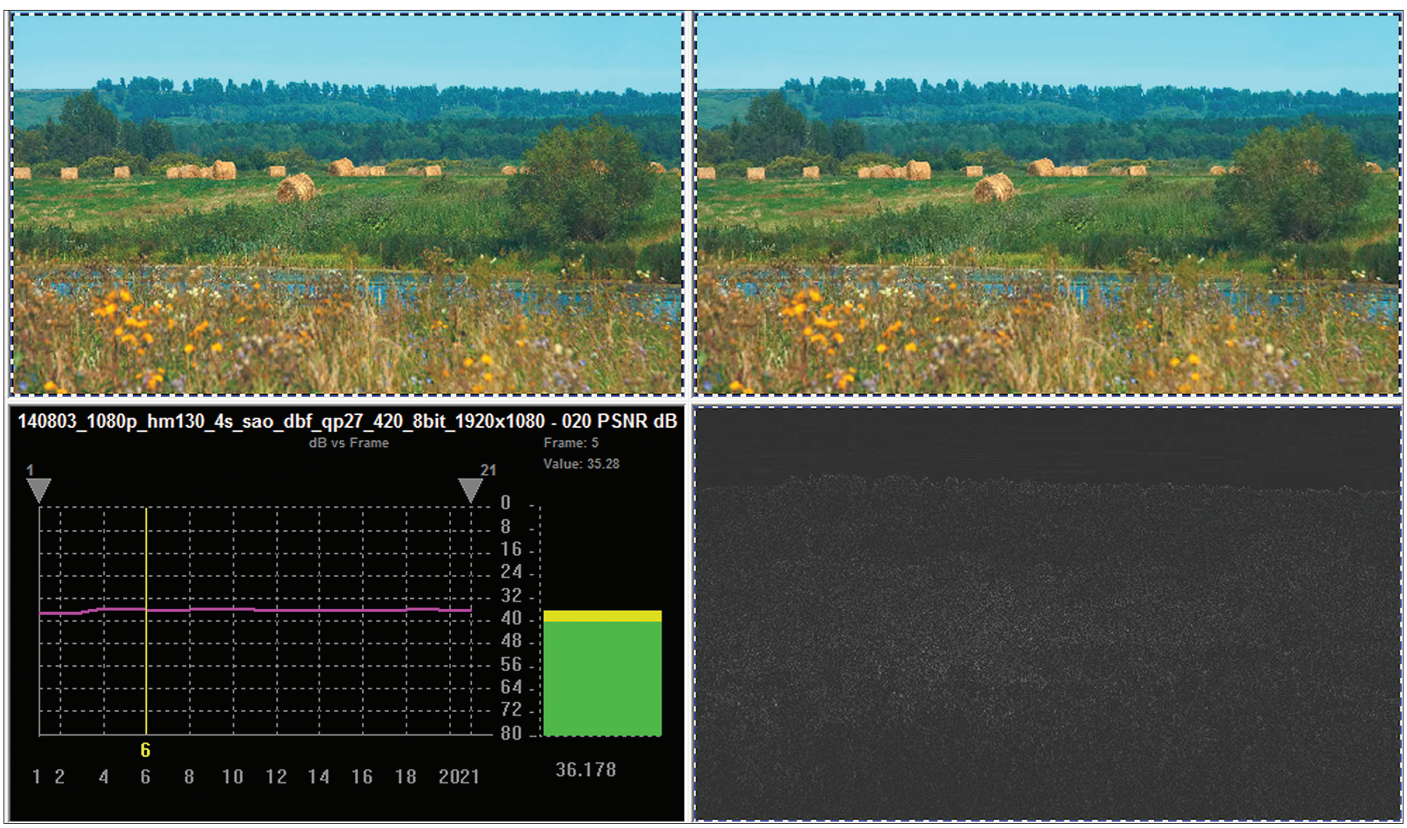


Figure 15. PQA PSNR results (samples from Elecard).

DMOS Testing

Differential Mean Opinion Score is a method defined by ITU Rec.500 for human viewers comparing two nearly identical videos. The controlled conditions include the TV monitor type, room lighting, distance from the viewer to the TV (in screen heights away from the TV). The normal viewing distance in the normal ITU Rec.500 trial is for a viewer to sit five screen heights away form a standard definition CRT TV. This is obviously out of date for today's Ultra HDTV displays. The current assumption is for three screen heights for an HDTV. For Ultra HDTV displays, the recommended distance for the viewer is about 1.5 screen heights. For a 90" Ultra HDTV, the viewing distance should be about 5.6 feet.

The PQA DMOS method has been automated thanks to the Human Vision Model enhanced by Tektronix based on the Just Noticeable Difference (JND) algorithm created by Sarnoff Laboratories. This HVM algorithm is the result of over 30 years of conducting human viewer video trials. The HVM algorithm is the basis for both the DMOS and PQA tests. In the case of the DMOS test, there is a requirement to run several different videos against the same reference file to determine a worst case scenario. Each test file run through the PSNR test will generate a Minkowski value. The worst case value should be imported into the DMOS configuration before starting the test. The details of setting up this DMOS test are as follows (assuming that temporal and spatial alignments are complete):

1. First, select the Configure Measure menu and then select only the HDTV DMOS test. Select Edit (as seen in Figure 16) to customize the test to your TV and lighting conditions. By selecting Edit, a copy of HDTV Broadcast DMOS template will be created and can be renamed according to the specific test. In this case, we will rename the test to Ultra HDTV Broadcast DMOS.

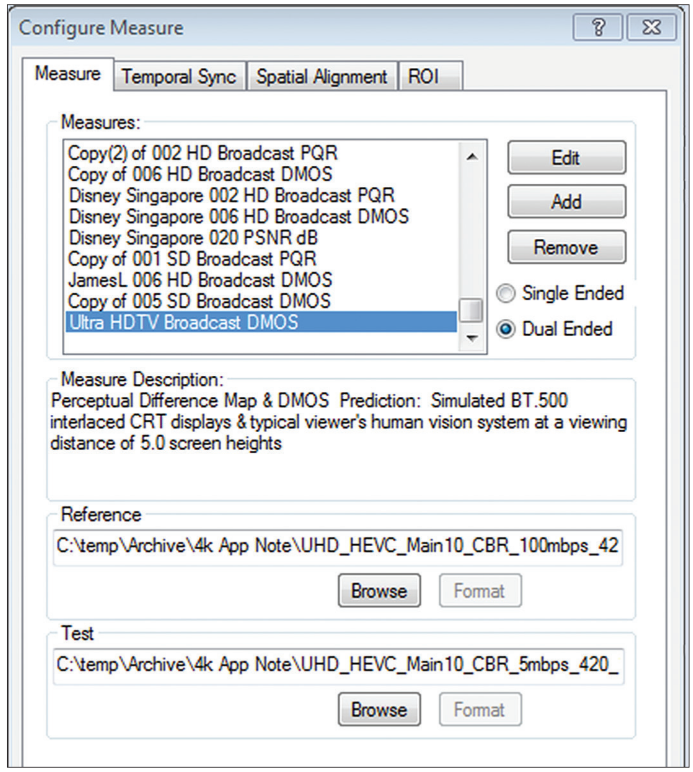


Figure 16. Customizing the DMOS template.

2. After selecting the Edit menu, a new Edit menu will open as shown in Figure 17.

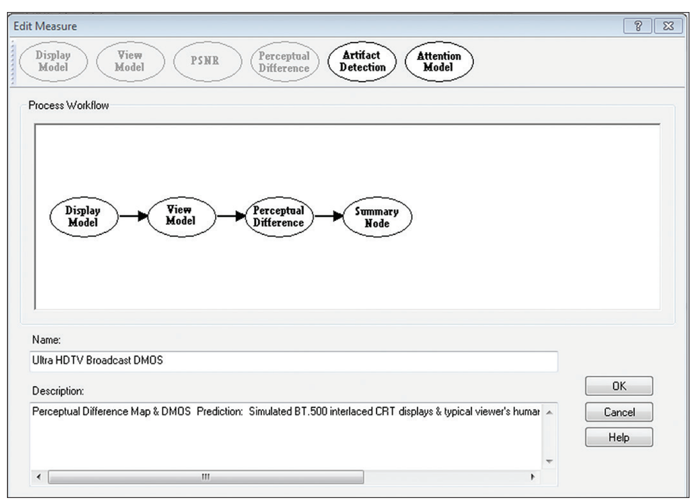


Figure 17. Editing the DMOS template.

3. Right-Click on the Display Model icon to open the menu show below in Figure 18. Change the setup from CRT to LCD.

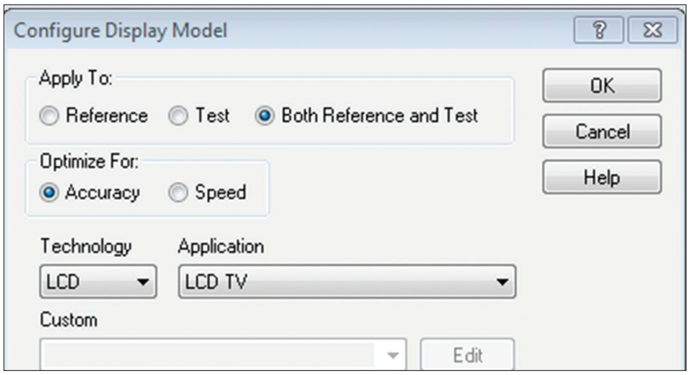


Figure 18. Display model.

4. Next, right-click the View Model icon. Change your viewing distance to 1.5 units as shown in Figure 19.

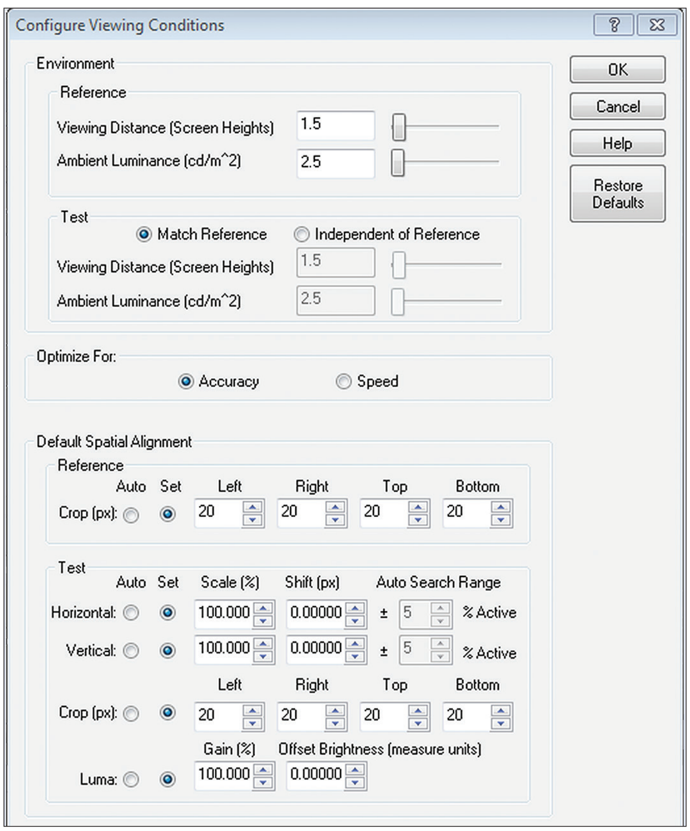


Figure 19. PQA Viewing Conditions.

5. Finally, right-click the Summary icon to import the worst case Minkowski value from previous PSNR trials. The Import icon as shown in Figure 20 will guide you to the folder were the PSNR .CSV result files are stored. The Import function will automatically extract the Minkowski value from the PSNR file.

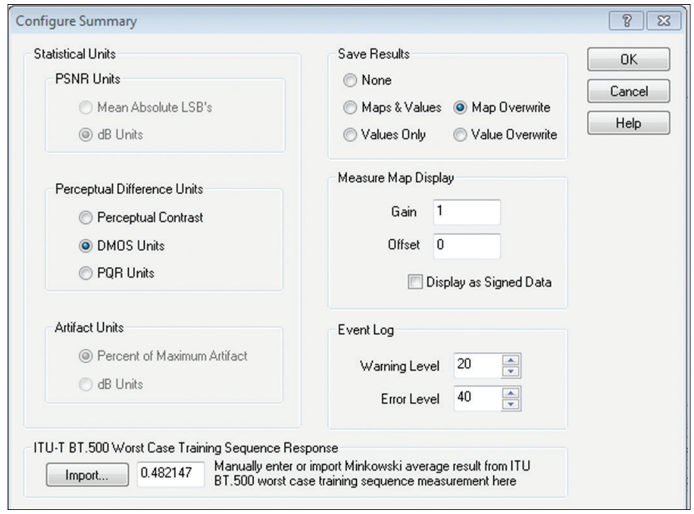


Figure 20. PQA Configure Summary menu.

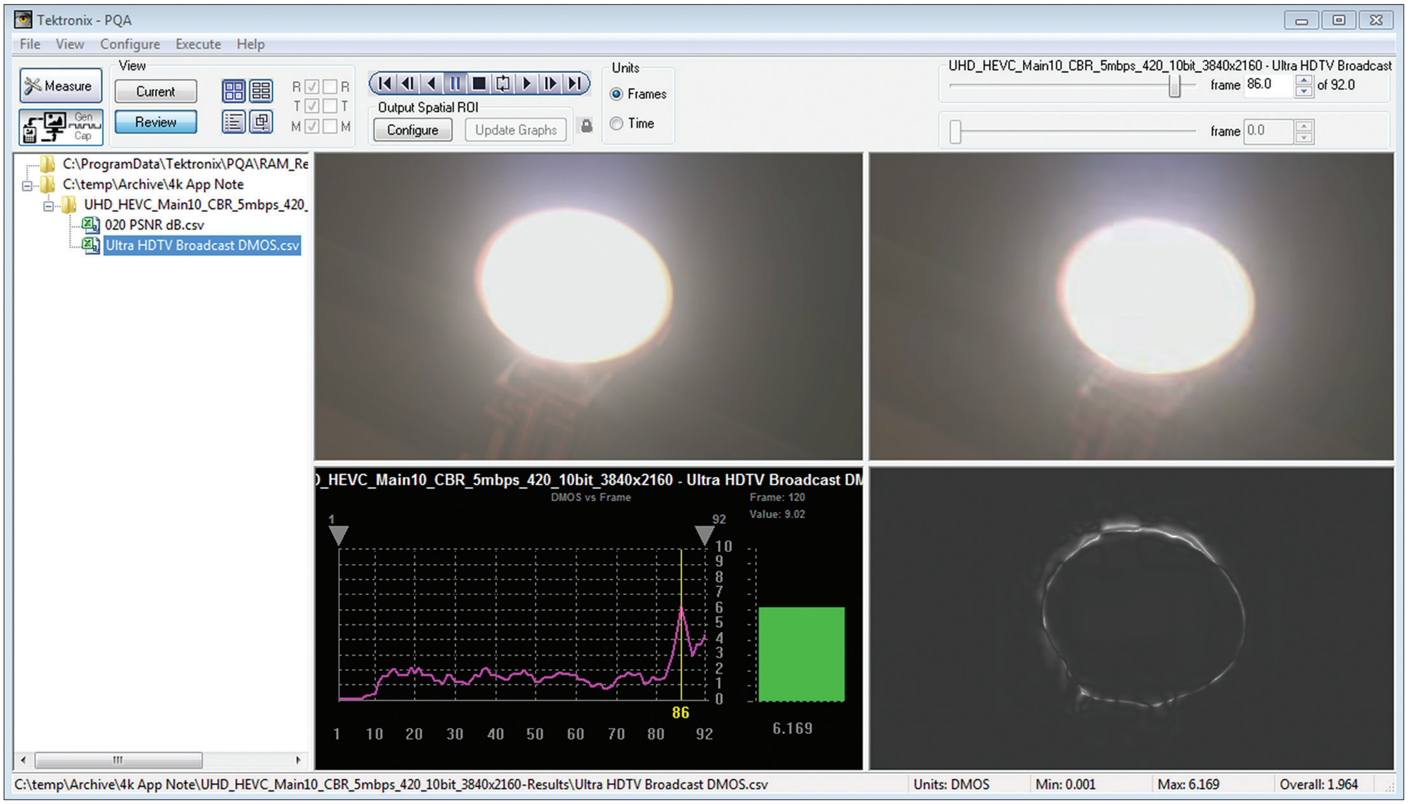


Figure 21. PQA DMOS results.

After completing the DMOS configuration changes for Ultra HDTV (4k) testing, we can now select the Measure icon in the main PQA display. The results, as shown in Figure 21, will be similar to the PSNR measurements, but the scale will now be in DMOS units based upon human psycho-visual models.

To be accurate, the DMOS template should be applied to each of the many different Test files. Each Test file is from a slightly different encoder setting. The test results closest to zero will denote the file with the least objections from human viewers.

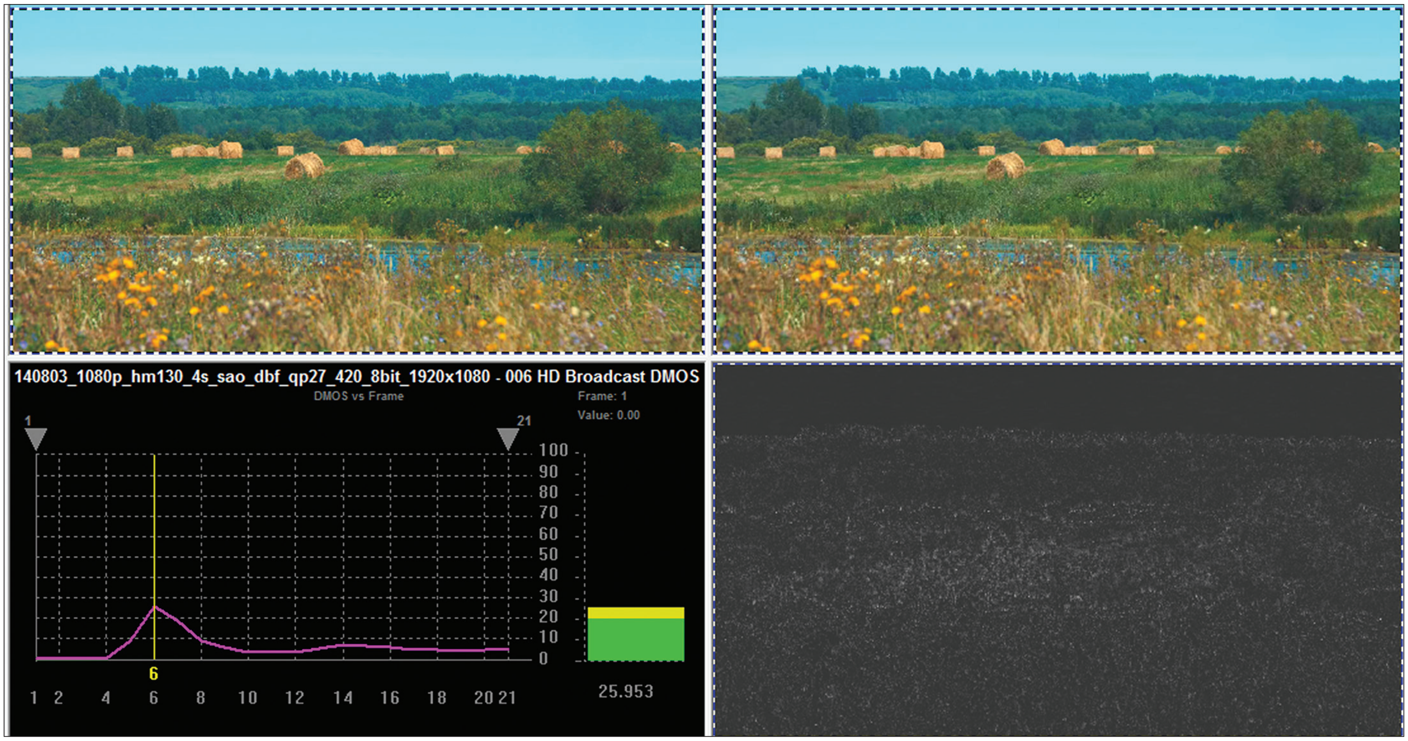


Figure 22. PQA DMOS results (samples from Elecard).

Figure 21 showed a 4k Reference file compared against a 4k Test file. It is also possible to test the DMOS of a 4k Reference file against a 1080p Test file as seen in Figure 22.

The PQA DMOS results are better in the last part of the short clip when the scene was more stable and the foreground grass was not moving as much.

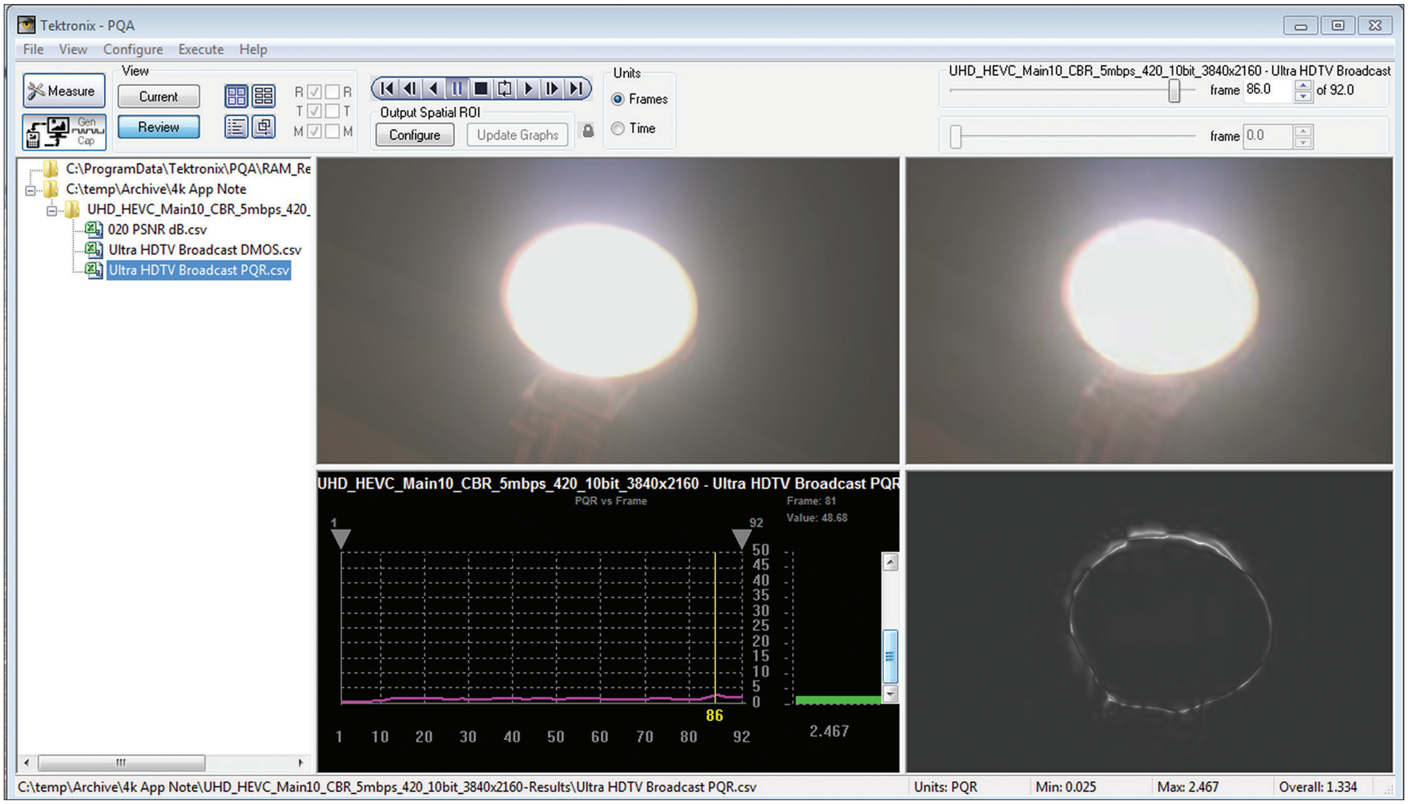


Figure 23. PQA PQR results.

PQR Testing

Both the DMOS and PQR methods are based upon the HVM algorithm. One key difference between them is that the PQR method does not require several training sequences before starting the test. Also, there is a reference point of 1.0 PQR meaning that three out of four human viewers would notice some difference between the Reference and Test files. The previous DMOS template-configure changes can also be used to configure PQR template too (i.e., lighting, distance, display type, etc.). So before we begin, let us select the Configure-Measure menu as we did earlier, but this time please ensure that only the HD Broadcast PQR method is selected, and then adjust the Display Model and Viewing Conditions. Once completed, select the Measure icon and view the results in the lower left quadrant as shown in Figure 23.

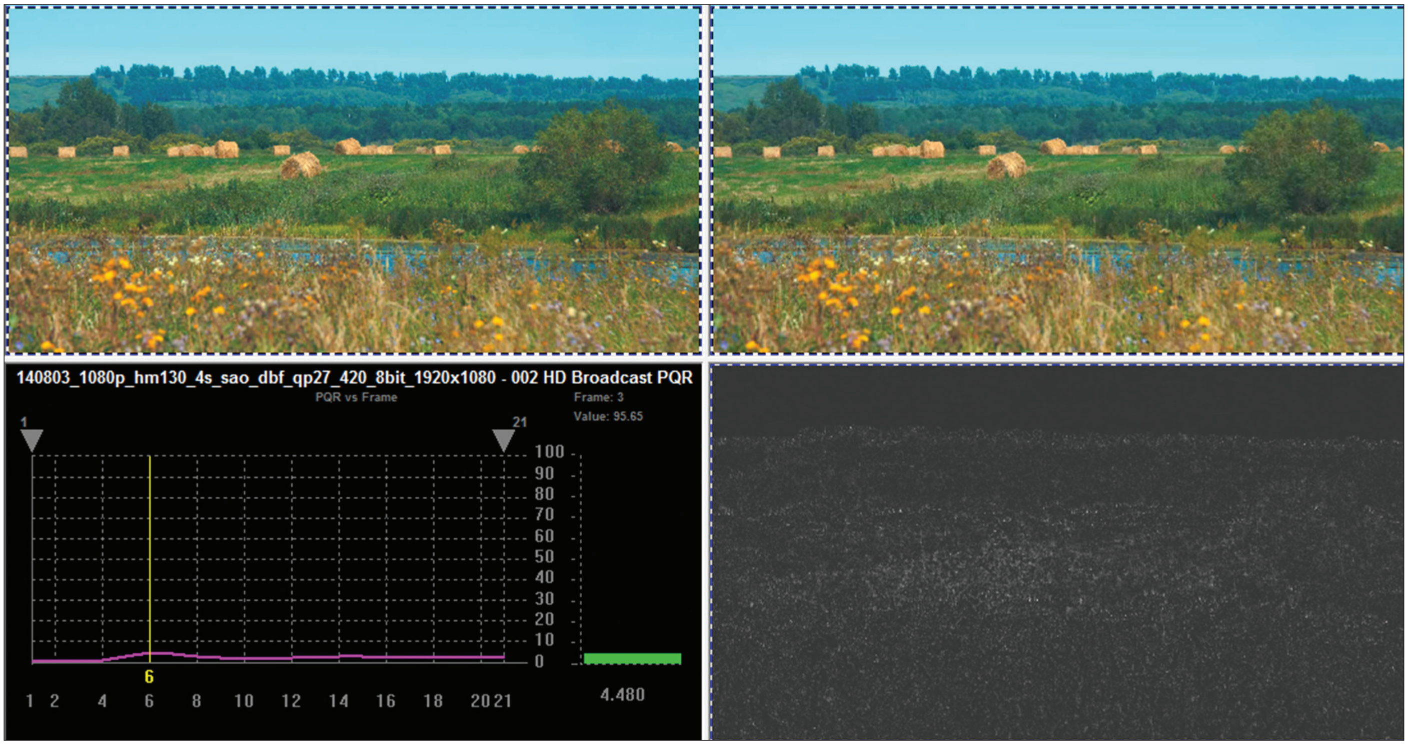


Figure 24. PQA PQR results (samples from Elecard).

The average PQR value in this example is about 1.0, with a high peak at 2.467 on frame 86. This means that about three out of four viewers would be able to notice some differences between the Reference and Test files. One important point here is that the Test file is 4k at 5 Mbps, while the Reference file was running at 100 Mbps. This is quite good for 5 Mbps, but other Test files running with the same encoder configurations, but at higher rates (e.g., 10, 15, 20 Mbps), would result in significantly lower PQR results meaning that most people could not tell the difference between a stream at a lower broadcast rate and a much higher contribution rate.

Figure 24 shows the Picture Quality Rating on a 4k Reference file compared against a 1080p file. The result are very good, but because the overall results are greater than a PQR of 1.0, then we know that at least three out of four human viewers would be able to detect a difference between the 4k Reference and the 1080p Test file.

Summary

As noted earlier, the new HEVC codec is capable of running at about half the bandwidth of its predecessor (AVC or H.264), and up to one quarter of MPEG-2. HEVC supports SD and HD formats, as well as the much larger 4k and 8k formats of Ultra HDTV, and smaller formats of mobile devices. With 4k display panels well into production, and many broadcasters around the world already broadcasting with HEVC, we are well on our way to adopting HEVC throughout the world for digital TV. The MTS4EAV7 continues to be a reference compliance application for HEVC interoperability testing, as well as the PQA600B and PQASW for providing picture quality testing across all video formats used today, as well as in the future with 8k and beyond.

Contact Information:

ASEAN / Australia (65) 6356 3900
Austria 00800 2255 4835
Balkans, Israel, South Africa and other ISE Countries +41 52 675 3777
Belgium 00800 2255 4835
Brazil +55 (11) 3759 7627
Canada 1 800 833 9200
Central East Europe / Baltics +41 52 675 3777
Central Europe / Greece +41 52 675 3777
Denmark +45 80 88 1401
Finland +41 52 675 3777
France 00800 2255 4835
Germany 00800 2255 4835
Hong Kong 400 820 5835
India 000 800 650 1835
Italy 00800 2255 4835
Japan 81 (3) 6714 3010
Luxembourg +41 52 675 3777
Mexico, Central/South America and Caribbean 52 (55) 56 04 50 90
Middle East, Asia, and North Africa +41 52 675 3777
The Netherlands 00800 2255 4835
Norway 800 16098
People's Republic of China 400 820 5835
Poland +41 52 675 3777
Portugal 80 08 12370
Republic of Korea 001 800 8255 2835
Russia / CIS +7 (495) 6647564
South Africa +41 52 675 3777
Spain 00800 2255 4835
Sweden 00800 2255 4835
Switzerland 00800 2255 4835
Taiwan 886 (2) 2656 6688
United Kingdom / Ireland 00800 2255 4835
USA 1 800 833 9200

Rev. 01/16

For Further Information
Tektronix maintains a comprehensive, constantly expanding collection of application notes, technical briefs and other resources to help engineers working on the cutting edge of technology. Please visit www.tek.com

01/16 EA 2AM-31033-1