

使用 MTS4EAV7 和 PQA600B 进行超高清电视测试

应用指南

随着业务多元化程度不断提高，高清视频日益流行，超高清格式（如 4k × 2k 或 8k × 4k 分辨率）出现，现在人们需要的编码速率已经超出 H.264/MPEG-4 AVC 的能力，同时也对移动应用提出了更高的质量和分辨率要求。

高效视频编码 (HEVC) 或 H.265 预计将提供与高级视频编码 (AVC) 或 H.264/MPEG-4 同样的观感质量，而使用的带宽只有后者的一半。它还支持超高清分辨率，如 4K 和 8K。内容发行商可以以更低的带宽费用，通过移动数据网络把高清视频传送到移动设备，并为家庭提供精彩的 4K 及以上速率的视频。HEVC 是为满足 H.264/MPEG-4 AVC 的所有现有应用设计的，特别是重点解决两个关键问题：提高视频分辨率，提高并行处理结构的使用量。

高效编码器具有至关重要的意义，以便与 AVC 相比使位速率降低 50%，与 MPEG-2 相比使位速率降低 70%。在编码器设计阶段，设计人员需要一种工具，能够帮助他们做出决策，选择正确的编码器算法，或者能够比较两种不同的编码器算法，并在视频质量水平一定时比较结果。此外，在选择正确的视频编码器及比较多个编码器时，您需要同一视频质量的多个编码器的统计数据，每个编辑器是否生成满足标准的码流，在解码时是否有影响解码器的缓冲违规，等等。MTS4EAV7 AVC/HEVC 分析仪检验编码的视频是否满足标准，提供编码效率统计数据，帮助编码器制造商调试和验证编码器设计。

一旦我们确定压缩的视频流在句法和语义上是正确的(根据 HEVC 编码解码器标准是合法的),我们现在需要查看原始视频与解码后的 HEVC 视频之间的视觉差异。为此,我们有许多不同的选择和算法。最基本的方式是在两个不同的视频监视器上简单地播放源视频和解码后的视频,让观看者评估任何视觉差异。这种方式主观性特别强,因为每个观看者都会使用自己的个人偏好来确定视频质量。使用真人有一个坏处,我们偶尔会说两个视频短片略有差异,事实上它们完全相同。在使用观看者时,最好的方式是使用一个严格控制的观看室,并对大量的观看者样本抽样,以获得有效的统计结果。由于这种测试成本非常高,因此最好不使用这种方式,而是采用确定视频质量使用的数学模型。目前最常用的模型是峰值信噪比(PSNR)、差异平均主观评分(DMOS)和图像质量评价(PQR)。

PQA600B 和 PQASW 不仅支持 PSNR、DMOS 和 PQR,还支持许多其他数学模型,并能够根据不同的观看条件(如照明、距电视机的距离、等等)定制每个模型。在测试 HEVC 视频时,我们将使用 MTS4EAV7,把压缩的视频文件解码成每个样本一个字节或两个字节的 YUV 文件(8 位或 10 位 HEVC)。然后,PQA600B 和 PQASW 可以客观地比较原始或参考 YUV 视频文件及来自 HEVC 编码器的 YUV 文件。

本应用指南先考察 MTS4EAV7 产品、怎样使用它确定互操作能力以及在编码过程中使用 HEVC 方案。我们还将使用 MTS4EAV7 把 HEVC 文件导出到 YUV 文件。最后,我们将考察 PQA600B 和 PQASW 应用,比较参考视频文件和测试视频文件。在正常情况下,参考视频文件应来自没有压缩的 YUV 源,但也可以略微压缩的文件与压缩程度很高的视频文件。

MTS4EAV7 测试

MTS4EAV7 应用可以针对 HEVC 标准及许多其他压缩标准中规定的一致性要求测试每帧视频。将在 Alert Log (告



图 1. MTS4EAV7 HEVC 提示框图标。

警日志)中列出任何不兼容的地方。在从 1 到 N 逐步通过视频帧时,Alert 日志将追踪出现的任何问题。其目标是运行或逐步通过每个帧,而不会遇到任何不兼容问题。文件中检测到的不兼容问题表明文件并不是 100% 兼容某项解码器标准。如果 Alert 日志中没有告警,那么表示文件可以被所有满足标准的解码器全面解码。在 Step Frame 或 Play 过程中,有许多其他显示功能,查看 HEVC 编码器做出的多种选项中的每一项。我们将在下面讨论这些额外的图形功能。

HEVC 编码效率

HEVC 是一种基于数据块的编码译码器,由更大的“编码树单元(CTU)”数据块(类似于宏数据块)组成,其长度由编码器选择。在很少需要或不需要编码时(压缩更多),编码器决定使用较大的编码单元;在需要编码时,编码器选择较小的编码单元。编码单元的长度可以是 16 个样点、32 个样点或 64 个样点,长度越大,压缩程度越好。MTS4EAV7 提供了许多提示框,分析编码树单元、编码单元和 Luma 和 Chroma 数据块中不同的编码参数。图 1 显示了各个提示框。

这些提示框在使用 MTS4EAV7 打开 HEVC 流时启用。图 2 显示了多个提示框中的第一个提示框。

HEVC 中的视频帧由多个编码树单元组成,每个编码树单元又由多个编码单元组成。任何编码树单元或编码单元都有不同的编码参数,这些参数由编码器选择,可以使用 MTS4EAV7 中提供的不同的提示框观察。上图显示了不同的参数,如位、分区模式、Luma 预测模式、等等。预测内部模式和预测之间模式的编码单元提示框承载着与不同预测模式有关的不同信息。

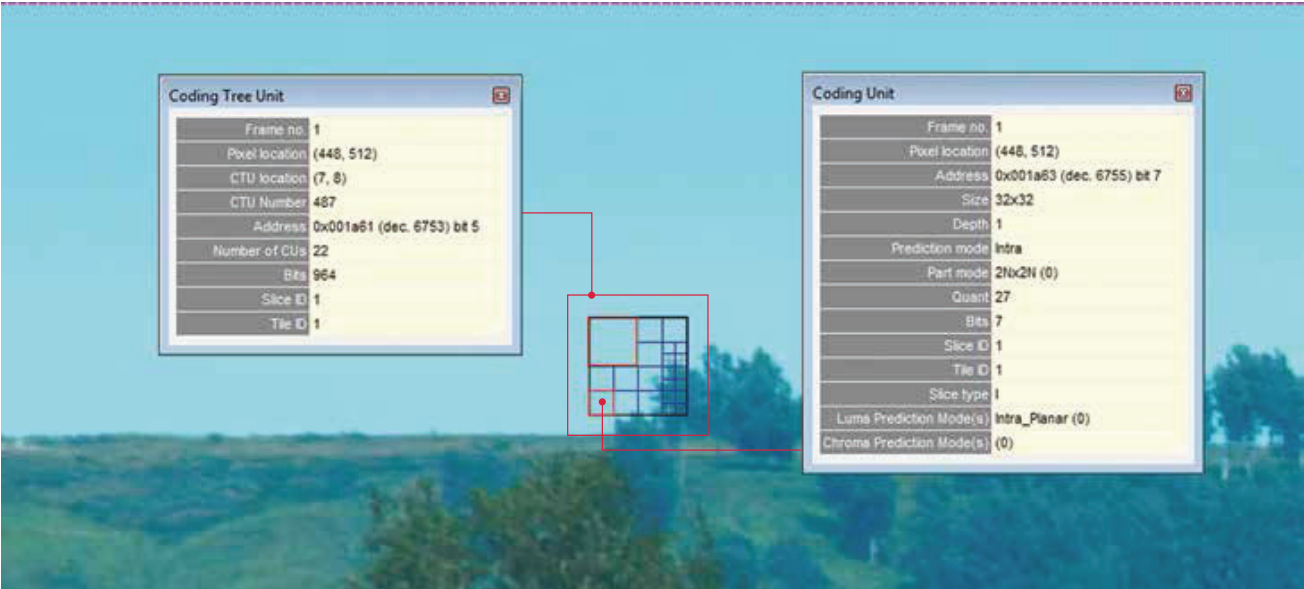


图 2. 编码树单元和编码单元提示框。视频放大 400% (4k HEVC 文件来自 Elecard)。

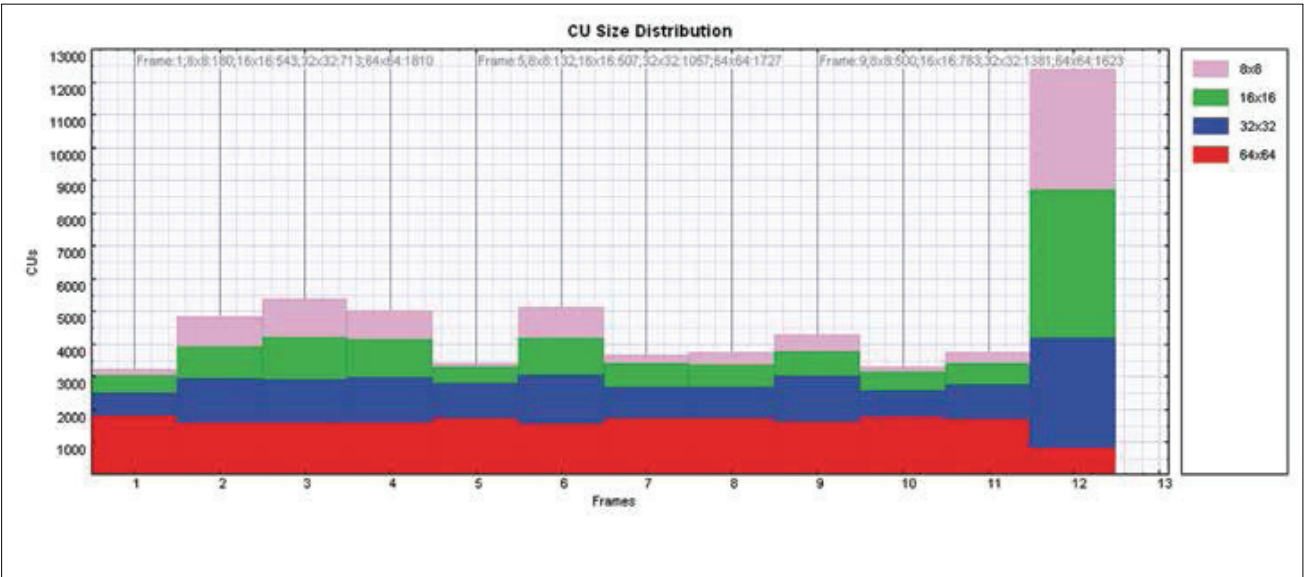


图 3. 每个视频帧变化的编码单元长度。

如上面图 3 所示，编码单元长度分布图指明了编码器选择的编码单元长度。在图片的正态区域，编码单元长度越高，压缩程度越好。预计方案越复杂，使用的编

码单元长度越小，因为与大的编码单元数据块相比，它们可以更有效地编码精密的细节。

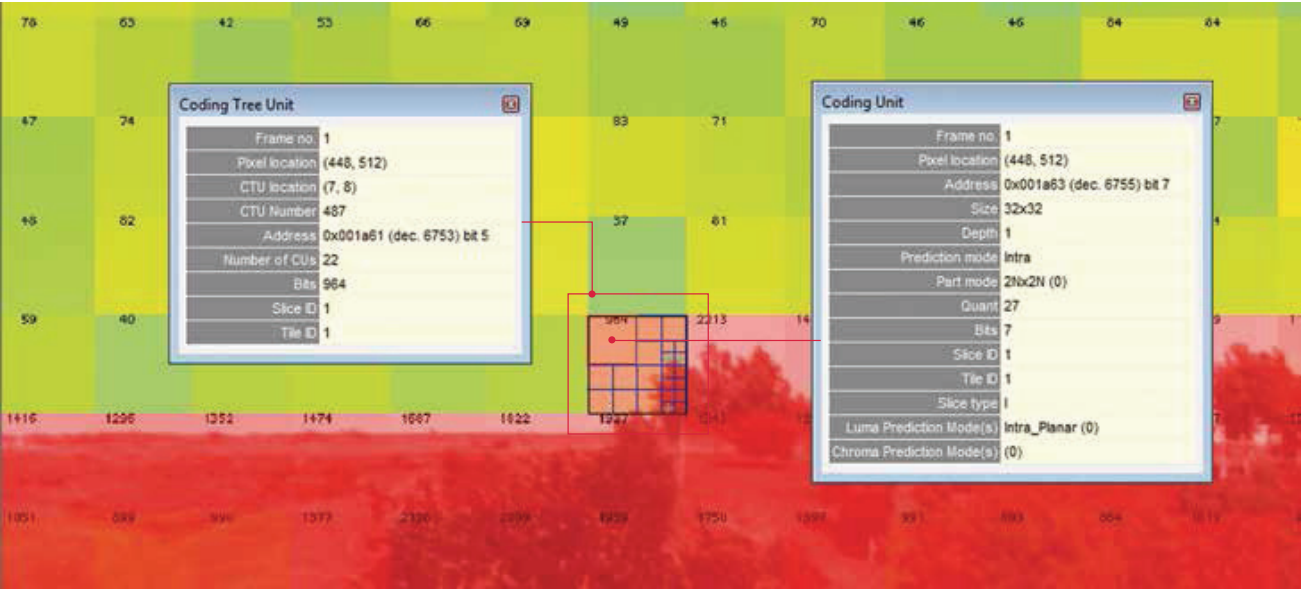


图 4. 显示 CTU 位叠加的图片 (场景同图 2)。

重叠

MTS4EAV7 在帧层提供了不同的叠加图，另外在整个基本流上码流层也提供了不同的叠加图，帮助收集不同的统计数据，如位数、量化、分片和磁贴。选择 Overlay -> CTU/MB Statistics -> Bits，选择 Bits per CTU。上面图 4 中的实例显示了 Bits per CTU 叠加图。它叠加了额外的色彩方案，以更简便地查看哪个 CTU 使用的位数量最少 (绿色)，其他颜色则显示哪个 CTU 使用的位数量最多 (红色)。

位叠加有一个选项，可以区分每个 CTU 的 Luma 位和 Chroma 位，这可以指明每个 CTU 占用多少位。用户可以使用缩放功能及鼠标控件转到感兴趣的 CTU。编码器智能选取和选择编码单元类型，为给定 CUT 分配要求的位数。每个编码单元可以采用 Intra 编码、Inter 编码或 Skipped for 编码。Intra 编码数据块占用的位数较多，Inter 和 skip 编码单元占用的位数较少。在 MTS4EAV7 中可以使用编码单元类型提示框和 Bits Overlay，观察和复核编码器选择不同类型和长度的决策。

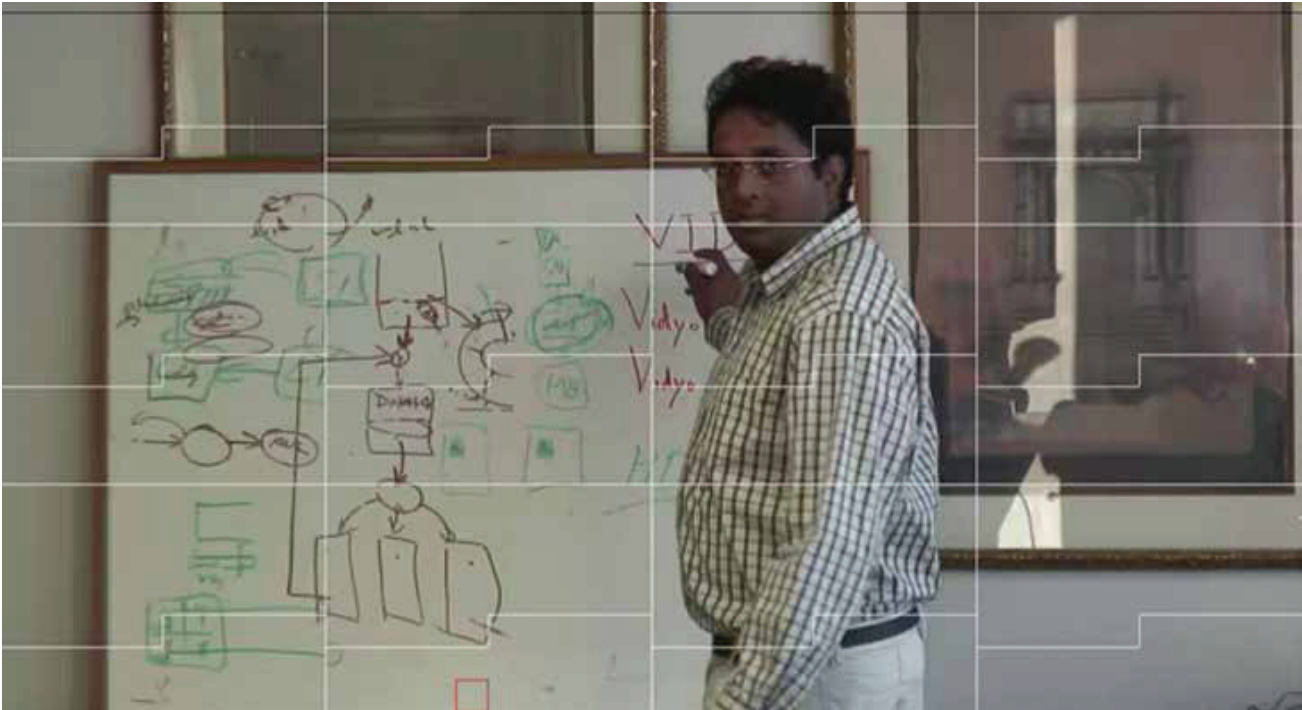


图 5. MTS4EAV7 显示 HEVC 分片分区。

Alert Log: merge_mondial_3840x2160p5094_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							

图 6. Alert 日志指明给定码流的 HEVC 标准违规情况。

分片和磁贴

Slice 叠加图允许目测观察帧中的分片分区。图 5 显示了 HEVC 编码器怎样把帧分成小的分片和磁贴。

码流一致性

MTS4EAV7 检查 HEVC 一致性，指明码流相对于标准的所有可能的违规。因此，预计从编码器中获得编码的任何码流都必须遵守 HEVC 标准，违规报告为许多不同码流层定义的告警。告警可以保存为 XML 文件格式，以进一步进行分析。

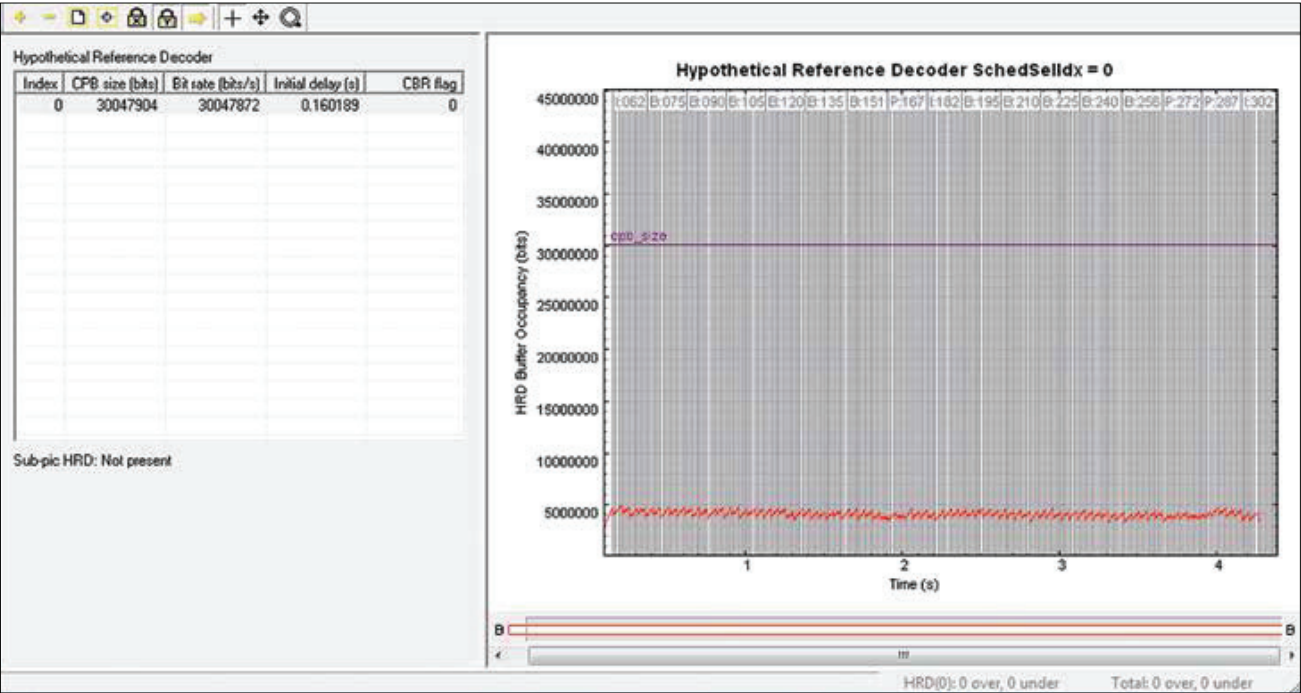


图 7. HRD 缓冲器分析。

缓冲器分析

如果码流没有溢出或下溢解码器的缓冲器内存要求，那么我们称这条码流满足标准。因此，在码流中指明的廓线和等级的任何 HEVC 码流都必须满足标准中提到的缓冲器要求。假想的参考解码器 (HRD) 为检查解码器缓冲器要求及解码器缓冲器占用情况提供了一条途径。

在 HEVC 流承载编码的 HRD 值时，MTS4EAV7 根据 HEVC 标准附录 C 检查码流是否溢出或下溢¹。上面的图 7 指明了视频流每个帧不同的缓冲器值和缓冲器占用程度。如果有任何缓冲器违规，都将作为一致性测试的一部分报告在 Alert Log 日志中。

¹(ITU-T H.265 04/2013 高效视频编码)。

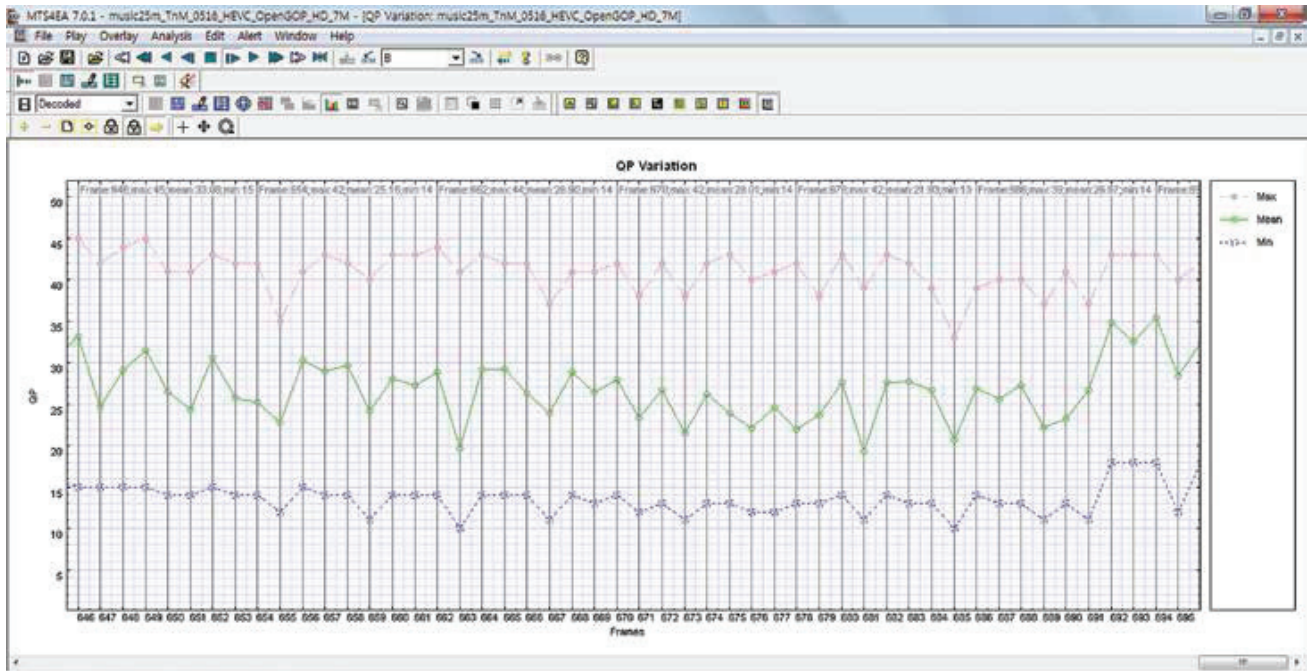


图 8. QP 变化图。

QP 变化

QP 指量化参数，再次指明了编码器的速率控制机制，以检查编码器在不同位速率时的性能。HEVC 编码器选择根据帧类型 (I、P 或 B) 分配不同的 QP 值，然后对 CU

级使用较少的位数。在位速率下降时，编码器控制着使用的量化等级和 CTU 长度。QP 变化图指明了编码器的速率控制机制，另外捕获整个视频持续时间内每个帧的最大、最小和平均 QP 变化。图 8 显示了最小、平均和最大 QP 变化。

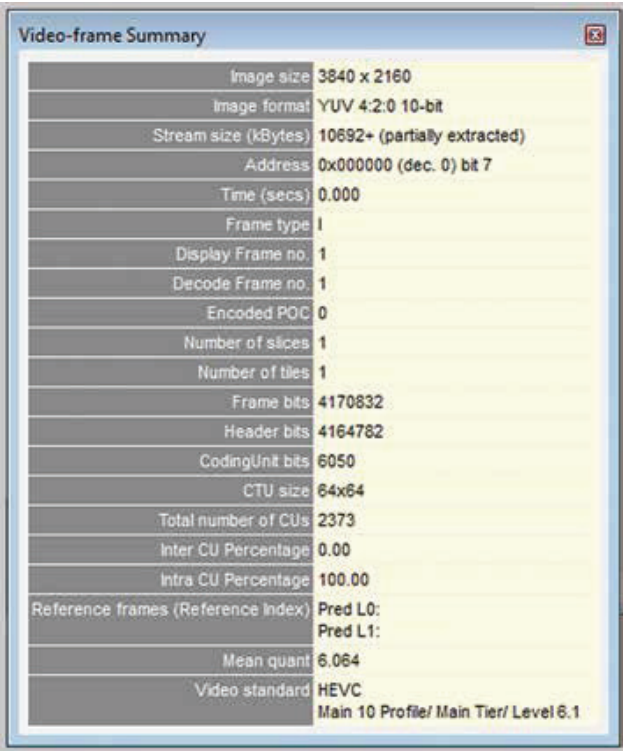


图 9. MTS4AV7 摘要，显示了 10 位模式下的 4k 格式。

MTS4EAV7 PSNR 测试

如果视频源是没有压缩的视频，用于参考，那么 MTS4EAV7 可以生成保真度分析指标，进一步量化原始视频的劣化情况。

把 8 位或 10 位 HEVC 视频导出到 YUV 文件中

MTS4EAV7 一直能够把压缩后的 4:2:2 或 4:2:0 视频导出为 YUV 文件。通过 HEVC，压缩后的视频通常采取 8 位形式或 10 位形式 (参见图 9)。

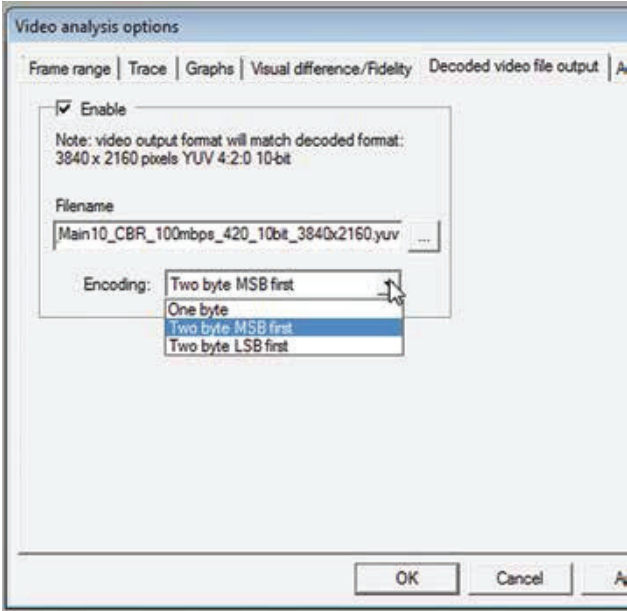


图 10. 把解码后的视频导出到 YUV 文件中。

在 HEVC 10 位视频中，MTS4EAV7 导出功能默认为两字节 MSB 文件格式。最新 PQA 中也为 YUV 导入提供了相同的文件格式 (参见图 10)。在我们的图像质量测试实例中，我们将使用两套不同的文件。带聚光灯的第一套文件都是 10 位 4k，第二套来自 Elecard，则是 8 位 4k 和 1080p HEVC 文件。

从 10 位 HEVC 源中导出视频的操作如下：

1. 打开视频文件。
2. 选择：Analysis–Output decoded video file…
3. 使用默认值，选择 OK
4. 逐帧移动视频，或 Play 连续步进
5. 按 Stop 关闭 YUV 视频文件，或等到一直到达文件最后

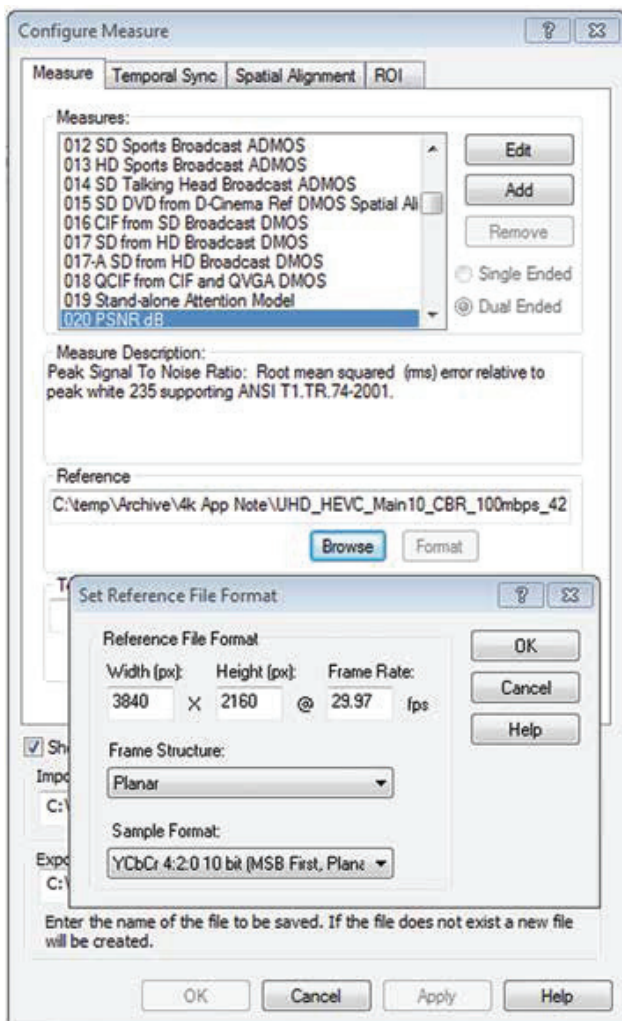


图 11. PQA Configure-Measure 菜单。

图像质量测试

PQA600B 和 PQASW 应用都支持许多不同的视频格式，包括最新的超高清电视格式 (4k 和 8k)。在 HEVC 中，我们将使用 MTS4EAV7 应用，把 8 位或 10 位 YUV 视频导出到磁盘上。在 PQA 应用内部，参考视频文件通常来自没有压缩的视频源，但参考文件也可以来自略微压缩的源文件中解码后的 YUV 文件。

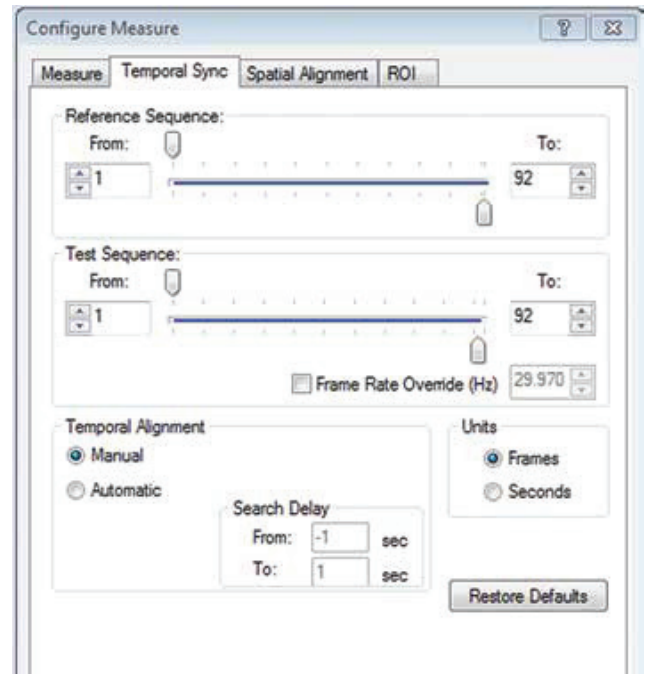


图 12. 时域帧同步。

4k 中的 PQA PSNR 测试

在这一测试中，第一步是选择 PQA Configure-Measures 菜单。图 11 显示了在多个可能的测试模型中只选择了 PSNR 模板。然后选择 HEVC YUV Reference 文件及 Test 文件。在这两种情况下，我们都将使用从 MTS4EAV7 应用中导出的一对 4k 文件。输入正确的横坐标、竖坐标、速率、结构和样本格式非常关键。YUV 文件格式不包括包头及这些细节，因此最好参考使用 MTS4EAV7 应用创建的文件名中的文本。

在选择 PSNR 和文件格式后，可能必需在时域上对准 Reference 帧和 Test 帧。图 12 显示有两种时域对准方式。一种是手动逐步通过 Reference 帧和 Test 帧，查找两个文件中唯一的场景。这在场景变化中最容易。第二种是使用自动模式，一帧一帧来回逐渐滑动文件，同时查看两个帧之间的差别。一旦在两个文件中找到最小点，那么将为每个文件分配一个偏置值。

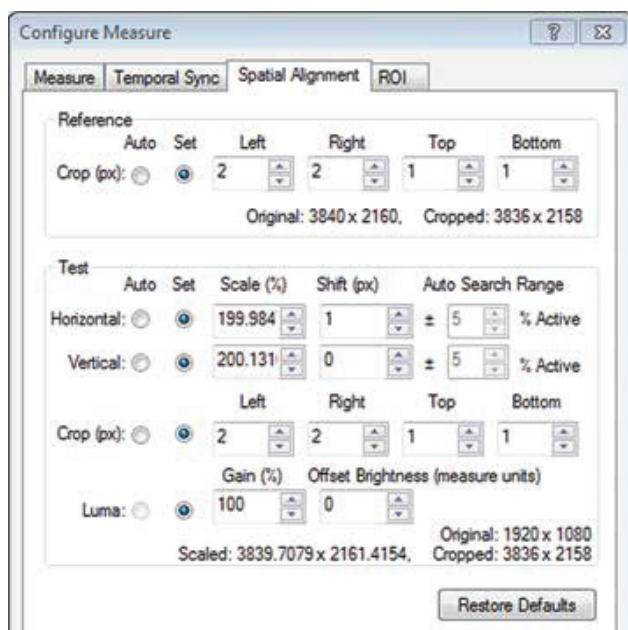


图 13. Spatial Alignment 空间对准。

在测量流程开始前，最后要检查的是可选的 Spatial Alignment 菜单，如图 13 所示。通常不需要使用这个菜单，但偶尔编码器可能会把编码后的视频帧向上、向下、向左、向右移动几个像素。此外，如果 Reference 文件和 Test 文件长度不同（比如 3840x2160 与 1920x1080），那么将要求使用 Spatial Alignment 菜单。在这些情况下，在已经设置 Temporal Alignment 后，在所有参数上使用自动对准模式会导致 Test 文件成比例放大和位移，以最好地匹配参考文件。在 4k Vs. HD 中，Test 文件将放大 200%，如图 13 和图 15 所示。

一旦配置 PQA 测量，我们可以退出菜单，选择图 14 左上角显示的 Measure 图标。Reference 视频将显示在左上象限。Test 视频将显示在右上象限。每个帧的像素差将显示在右下象限。最后，整体 PSNR 结果将显示在左下象限。

PSNR 测量提供了每个帧以及整个文件上加总的像素差的很好的数学结果。结果显示了 Test 文件与 Reference 文件的接近程度。得分在 80 dB 或以上意味着两个文件几乎完全相同。得分为零意味着两个文件根本不相关。得分大于 40 dB（如图 14 中的实例）表明两个文件相关程度非常紧密。我们不能比较 PSNR 结果与观看者试验，因为 PSNR 结果没有考虑人脑和人眼怎样协同工作。因此，良好的 PSNR 分值告诉我们两个文件接近，但我们需要考虑人类视觉模型，然后才能客观地评估这些文件的质量。

在前面的实例中，我们考察了比较两个 4k 超高清电视视频短片。在下面的 PSNR 例子中，我们将比较 4k 超高清电视 Reference 文件与面向广播就绪的 1080p 文件。您将注意到 PSNR 值比前面的 4k Vs 4k 测试高一位（图 15）。这符合预期，因为我们的 1080p Test 测试只是 4k Reference 文件长度的四分之一。

PSNR 结果在大约 40 dB 时仍非常好，但在从每个帧中去掉四分之三的像素时，其不可能达到 80 dB。在与 4k Reference 文件比较前，相应的 Test 图像必须扩大到 200%。

使用 MTS4EAV7 和 PQA600B 进行超高清电视测试

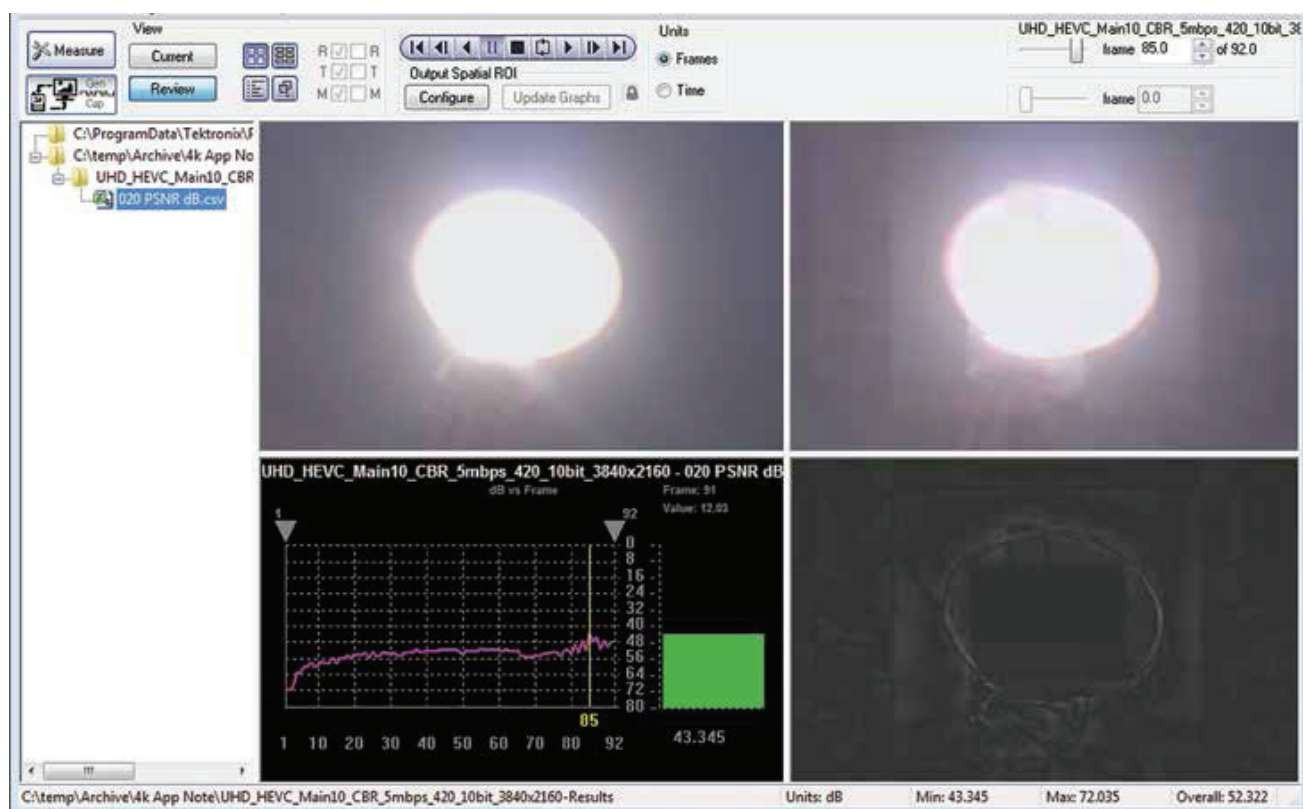


图 14. PQA PSNR 结果。

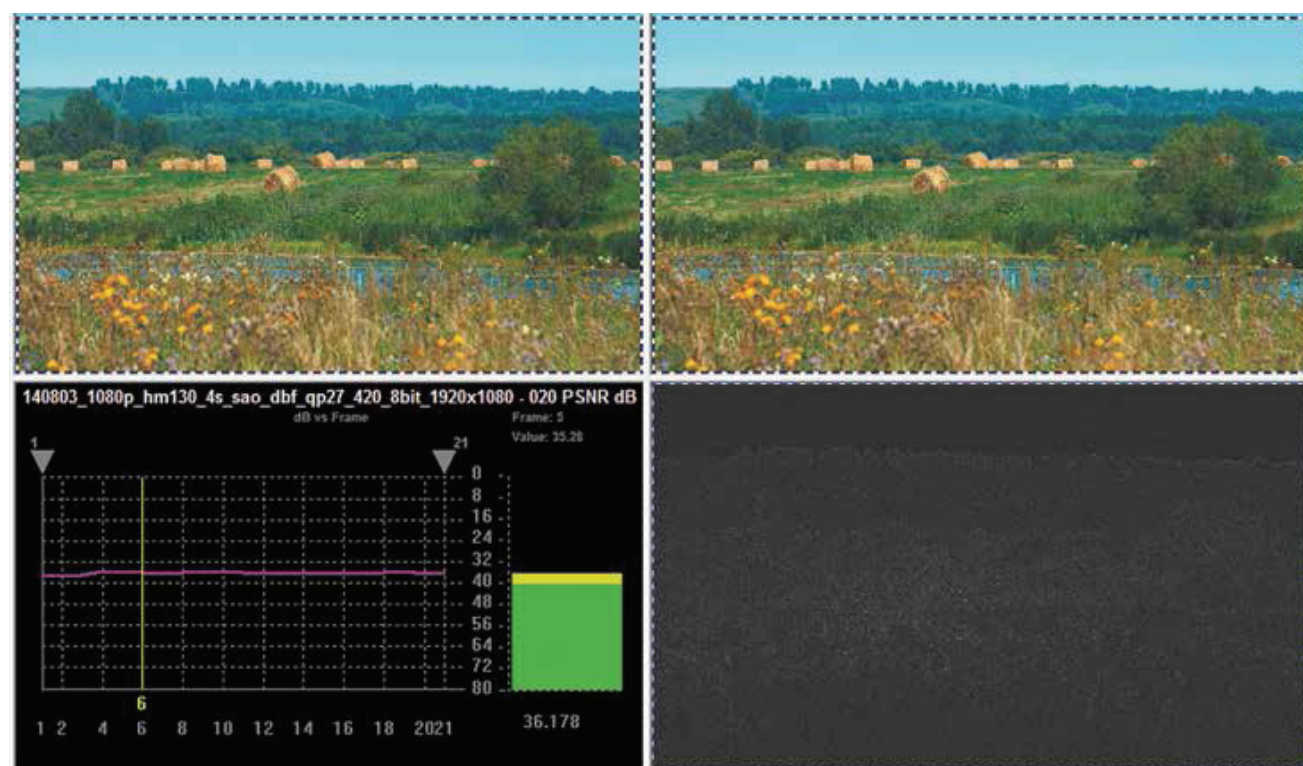


图 15. PQA PSNR 结果 (样本来自 Elecard)。

DMOS 测试

差异平均主观评分是 ITU Rec.500 规定的观看者比较两种几乎完全相同的视频的一种方法。受控条件包括电视监视器类型、房间光线、观看者距电视机的距离（距电视机的屏幕高度数）。对观看者来说，正常 ITU REC.500 试验中正常的观看距离是坐在距标准高清 CRT 电视机五个屏幕高度远的地方。这种方法对当今超高清电视显示器明显已经过时。目前的假设是 HDTV 为三个屏幕高度。对超高清电视显示器，推荐的观看者距离是大约为 1.5 个屏幕高度。对 90” 超高清电视，观看距离应为大约 5.6 英尺。

由于泰克在 Sarnoff 实验室创建的最小可觉差 (JND) 基础上增强了人眼视觉模型，因此 PQA DMOS 方式一直是自动实现的。这种 HVM 算法是泰克 30 多年来一直进行观看者视频试验的结果。DMOS 测试和 PQA 测试的基础都是 HVM 算法。在 DMOS 测试中，要求针对相同的参考文件运行多个不同的视频，以确定最坏情况场景。PSNR 测试中每次运行的测试文件都会生成一个 Minkowski 值。在开始测试前，应先把最坏情况值导入到 DMOS 配置中。这个 DMOS 测试的详细设置方式如下（假设已经完成时域对准和空间对准）：

1. 首先，选择 Configure Measure 菜单，然后只选择 HDTV DMOS 测试。选择 Edit（如图 16 所示），根据电视机和照明条件定制测试。选择 Edit，将创建一个 HDTV Broadcast DMOS 模板，然后可以根据具体测试重新命名。在这种情况下，我们将把测试重新命名为 Ultra HDTV Broadcast DMOS。

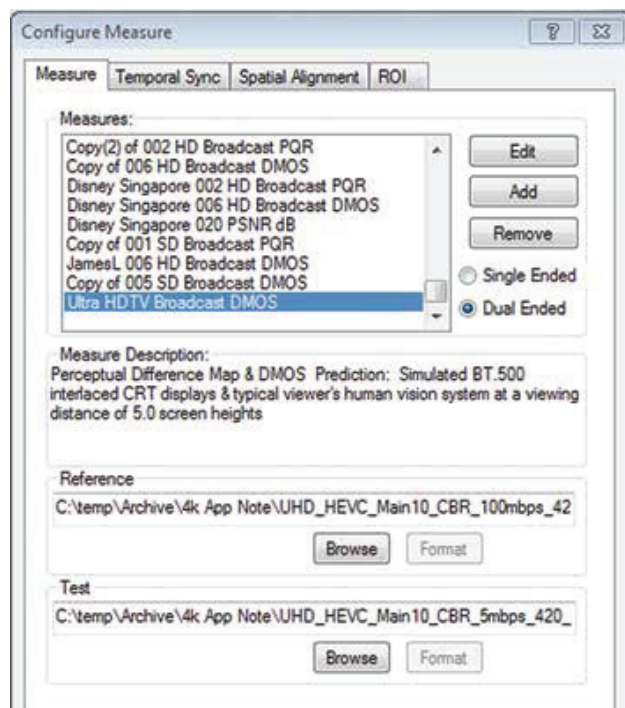


图 16. 定制 DMOS 模板。

2. 在选择 Edit 菜单后，将打开一个新的 Edit 菜单，如图 17 所示。

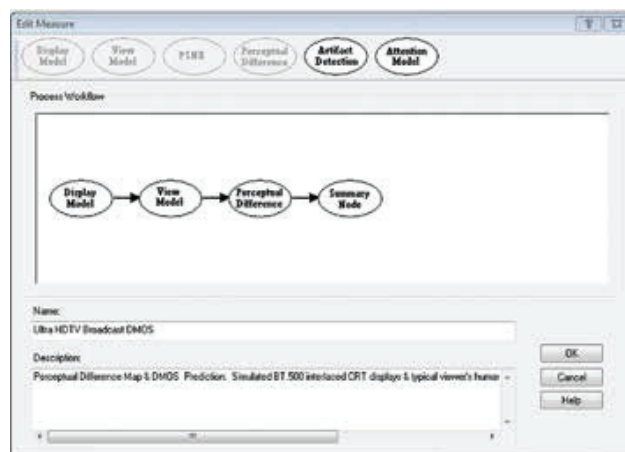


图 17. 编辑 DMOS 模板。

3. 在 Display Model 显示模型图标上按鼠标右键，打开菜单，如下面图 18 所示。把设置从 CRT 变成 LCD。

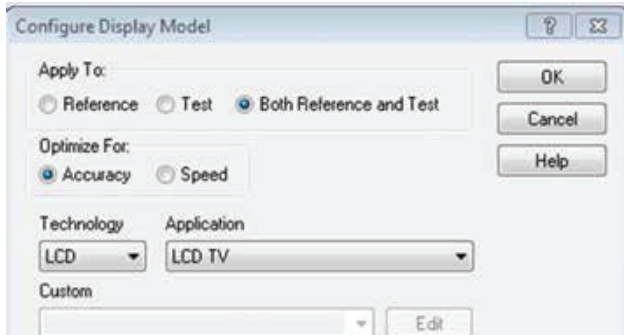


图 18. 显示模型。

4. 然后在 View Model 图标上按鼠标右键。把观看距离变成 1.5 个单位，如图 19 所示。

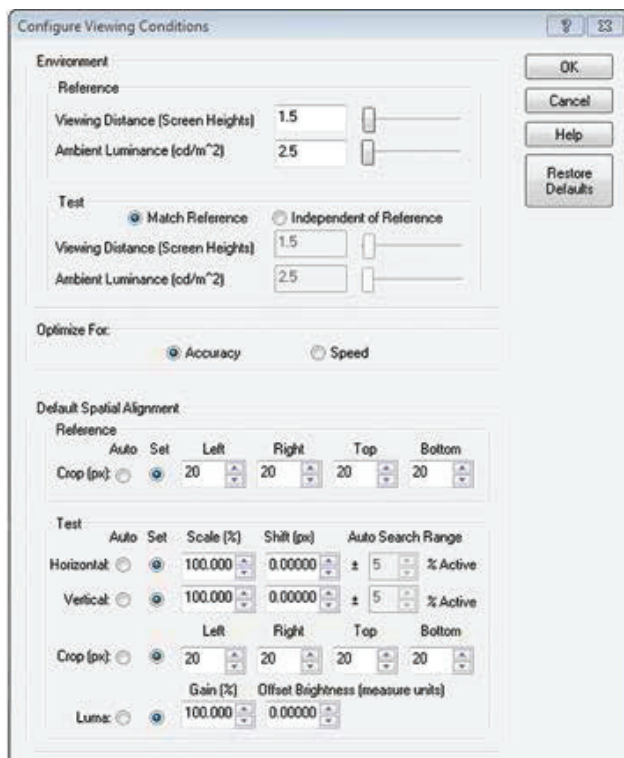


图 19. PQA 观看条件。

5. 最后，在 Summary 图标上按鼠标右键，从以前的 PSNR 试验中导入最坏情况 Minkowski 值。图 20 所示的 Import 图标将把您引导到存储 PSNR .CSV 结果文件的文件夹中。Import 功能将从 PSNR 文件中提取 Minkowski 值。

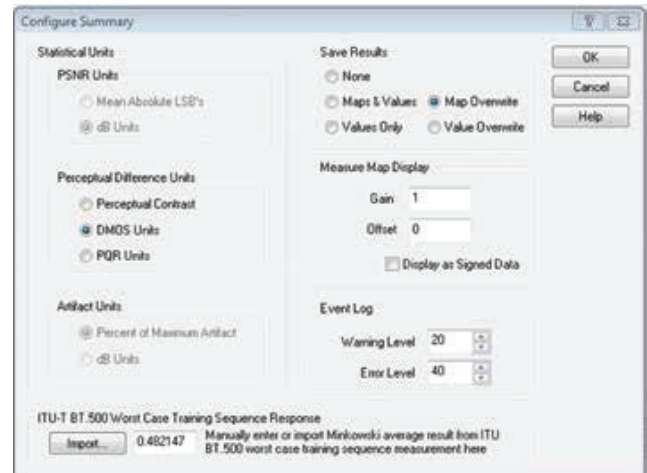


图 20. PQA Configure Summary 菜单。

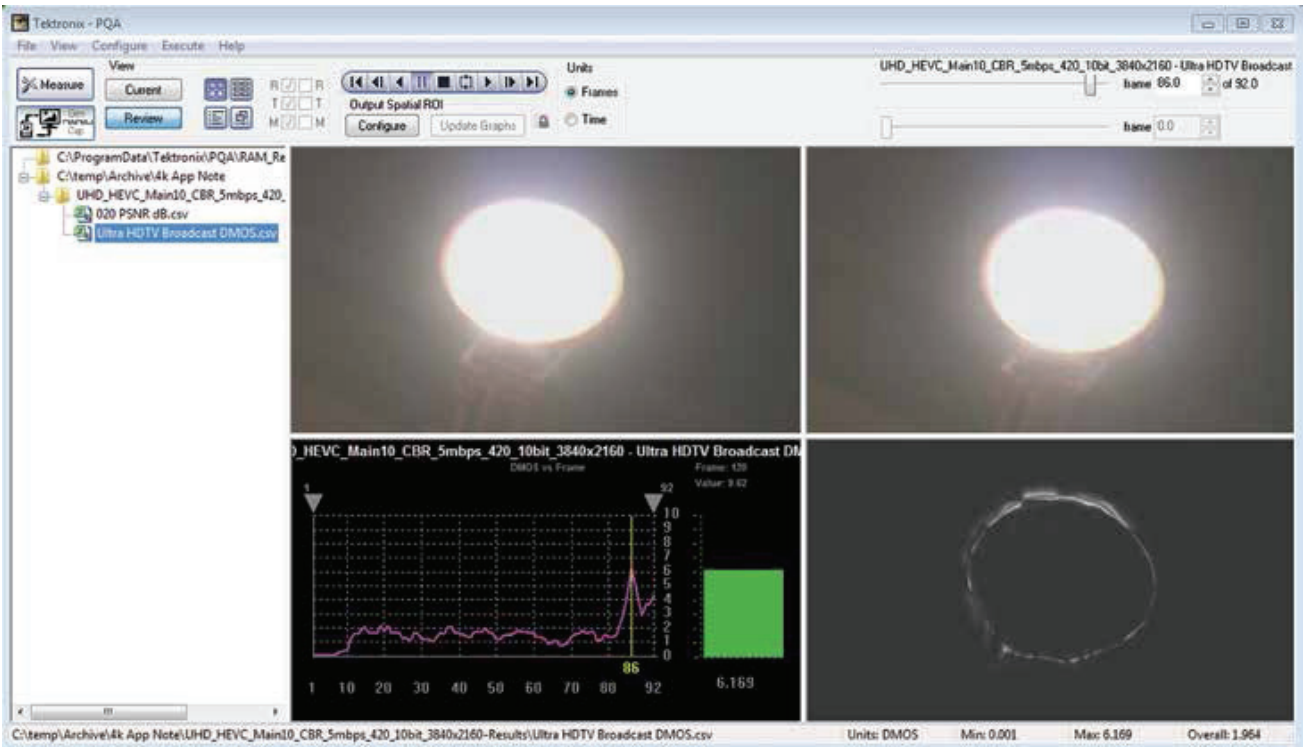


图 21. PQA DMOS 结果。

在对超高清电视 (4k) 测试完成 DMOS 配置变化后，我们现在可以在主要 PQA 画面中选择 Measure 图标。如图 21 所示，其结果与 PSNR 测量类似，但标度现在采用的是基于人类心理视觉模型的 DMOS 单位。为准确起见，

应对多个不同 Test 文件的每个文件应用 DMOS 模板。每个 Test 文件都来自略微不同的编码器设置。最接近零的测试结果表明观看者对文件的异议最少。

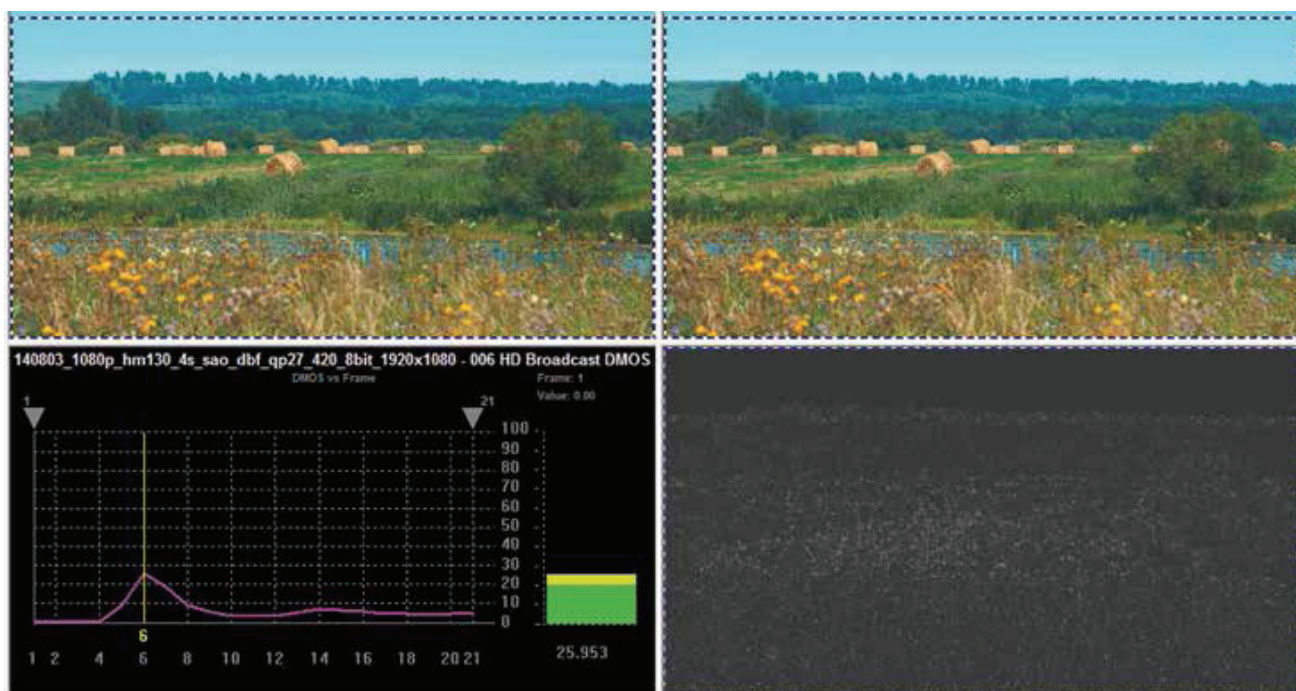


图 22. PQA DMOS 结果 (样本来自 Elecard)。

图 21 显示了 4k Test 文件与 4k Reference 文件比较结果。也可以针对 1080p Test 文件测试 4k Reference 文件的 DMOS，如图 22 所示。

在场景比较稳定，前景草地移动不多时，短片最后部分的 PQA DMOS 效果比较好。

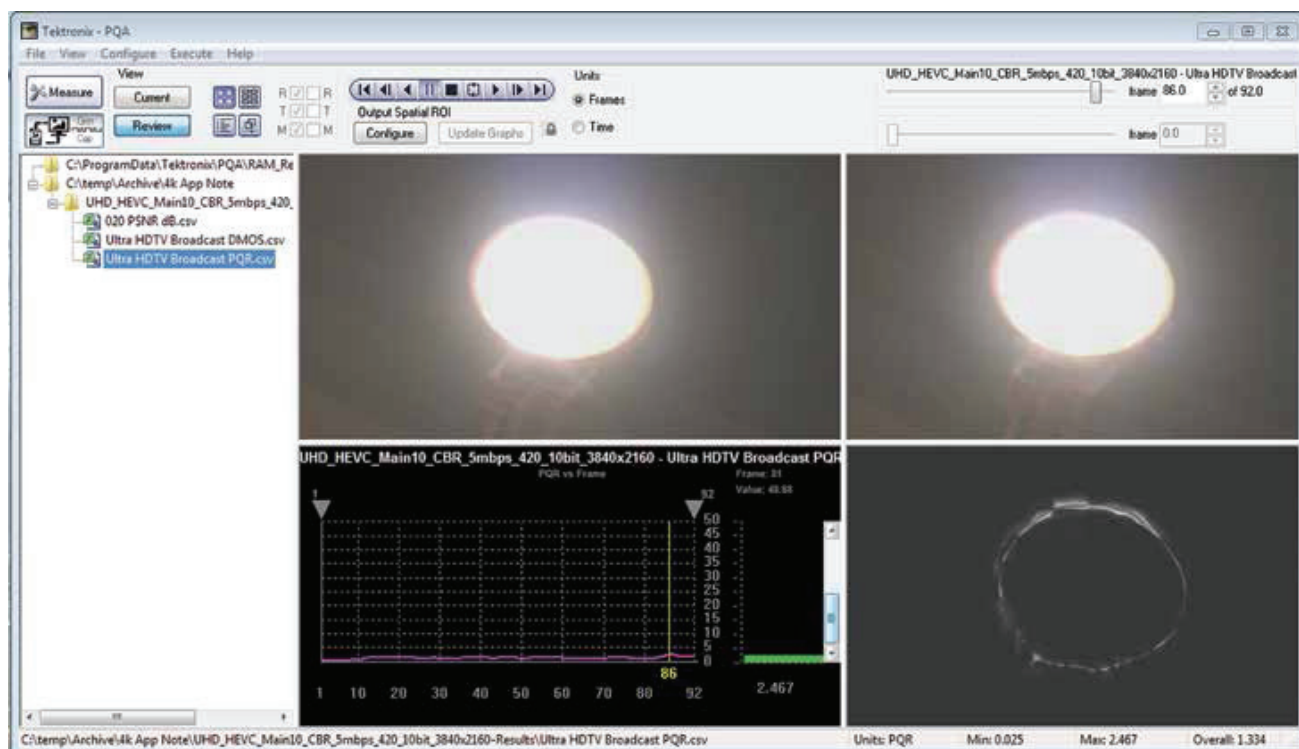


图 23. PQA PQR 结果。

PQR 测试

DMOS 方式和 PQR 方式都基于 HVM 算法，两者之间一个关键差别在于，PQR 方式在开始测试前不要求多个训练序列。此外，有一个 1.0 PQR 参考点，也就是说四分之三的观看者会注意到 Reference 文件与 Test 文件之间的某些差异。前面的 DMOS 模板配置变化还

可以用来配置 PQR 模板（如照明、距离、显示类型、等等）。因此在我们开始测试前，和前面一样，我们先选择 Configure-Measure 菜单，但这次要确保只选择 HD Broadcast PQR 方式，然后调节 Display Model 和 Viewing Conditions。一旦完成，选择 Measure 图标，在左下象限中查看结果，如图 23 所示。

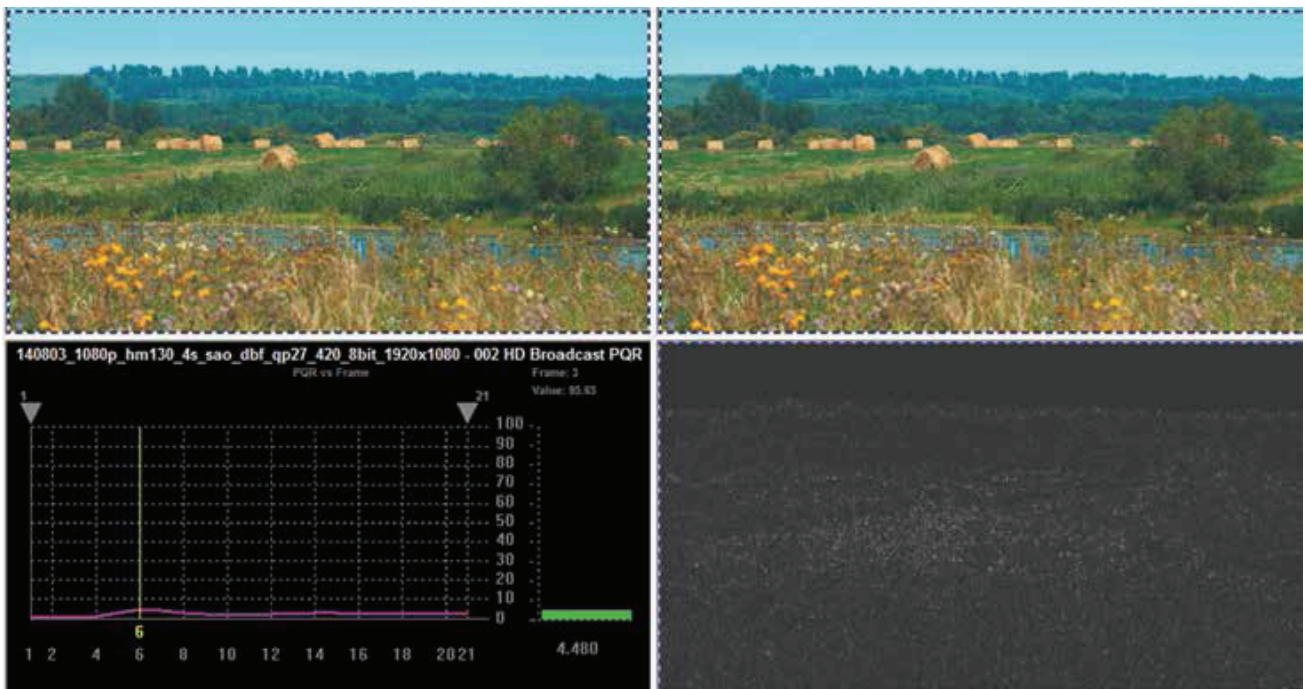


图 24. PQA PQR 结果 (样本来自 Elecard)。

在这个实例中，平均 PQR 值约为 1.0，高峰在第 86 帧上，峰值为 2.467。这意味着大约四分之三的观看者能够注意到 Reference 文件与 Test 文件之间的某些差异。这里很重要的一点是，Test 文件是 5 Mbps 的 4k，而 Reference 文件的运行速率是 100 Mbps。这对 5 Mbps 相当好，但采用相同编码器配置、但速率更高 (如 10, 15, 20 Mbps) 的其它 Test 文件将导致 PQR 结果明显低得多，也就是说，大多数人不能区分广播速率较低的数据流与贡献率高得多的数据流之间的差异。

图 24 显示了 4k Reference 文件与 1080p 文件相比的图像质量评分。结果非常好，但由于整体结果大于 1.0 的 PQR，因此我们知道至少四分之三的观看者能够发现 4k Reference 文件与 1080p Test 文件之间的差异。

小结

如前所述，新的 HEVC 编码译码器的运行带宽大约是上一代产品 (AVC 或 H.264) 的一半，最多是 MPEG-2 的四分之一。HEVC 支持 SD 和 HD 格式，支持超高清电视大得多的 4k 和 8k 格式，以及移动设备较小的格式。随着 4k 显示面板量产，同时随着世界各地许多广播公司已经使用 HEVC 进行广播，我们开始在世界各地的数字电视中采用 HEVC。MTS4EAV7 一直是 HEVC 互操作能力测试使用的参考一致性测试软件，PQA600B 和 PQASW 则能够在将来为 8k 及以上速率提供图像。

使用 MTS4EAV7 和 PQA600B 进行超高清电视测试

泰克科技(中国)有限公司
上海市浦东新区川桥路1227号
邮编：201206
电话：(86 21) 5031 2000
传真：(86 21) 5899 3156

泰克北京办事处
北京市海淀区花园路4号
通恒大厦1楼101室
邮编：100088
电话：(86 10) 5795 0700
传真：(86 10) 6235 1236

泰克上海办事处
上海市徐汇区宜山路900号
科技大楼C楼7楼
邮编：200233
电话：(86 21) 3397 0800
传真：(86 21) 6289 7267

泰克深圳办事处
深圳市福田区南园路68号
上步大厦21层G/H/I/J室
邮编：518031
电话：(86 755) 8246 0909
传真：(86 755) 8246 1539

泰克成都办事处
成都市锦江区三色路38号
博瑞创意成都B座1604
邮编：610063
电话：(86 28) 6530 4900
传真：(86 28) 8527 0053

泰克西安办事处
西安市二环南路西段88号
老三届世纪星大厦26层C座
邮编：710065
电话：(86 29) 8723 1794
传真：(86 29) 8721 8549

泰克武汉办事处
武汉市解放大道686号
世贸广场1806室
邮编：430022
电话：(86 27) 8781 2760/2831

泰克香港办事处
香港九龙尖沙咀弥敦道132号
美丽华大厦808-809室
电话：(852) 2585 6688
传真：(852) 2598 6260

有关信息

泰克公司备有内容丰富的各种应用文章、技术简介和其他资料，并不断予以充实，可为从事前沿技术研究的工程师提供帮助。请访问泰克公司网站 cn.tektronix.com

©2014 年泰克公司版权所有，侵权必究。泰克产品受到已经签发及正在申请的美国专利及外国专利的保护。本文中的信息代替以前出版的材料中的所有信息。本文中的技术数据和价格如有变更，恕不另行通告。TEKTRONIX 和 TEK 是泰克公司的注册商标。本文中提到的所有商号均为各自公司的服务标志、商标或注册商标。

11/14 DM/WWW

2AC-31033-0

Tektronix®