HIGH DYNAMIC RANGE FOR THE SECOND SCREEN

Sean T. McCarthy, Ph.D.

Independent Consultant, USA

ABSTRACT

High Dynamic Range (HDR) will soon be just as much about delivering great video experiences to tablets and smartphones as it is about delivering to the big UHD TV in the living room. For video service providers, however, the challenges are different. UHD HDR content is crafted with the big screen in mind. As a result, the small localized highlights and spatial details of HDR content can be obliterated when it is scaled, filtered, and encoded using adaptive streaming protocols to fit handheld screens.

This paper provides quantitative methods to optimize HDR for the second screen and adaptive bitrate services. We provide methods to measure and mitigate distortions of overall HDR luminance variations as well and the localized highlights and spatial details that are unique and significant in each video frame. A key take-away is that we provide methods to select the best combinations of bitrate and encoded resolution to use in multi-resolution adaptive streaming of HDR content.

INTRODUCTION

Adaptive streaming is rapidly becoming a dominant method for distribution of television to all screens from the big screens in living rooms to smaller-screen smartphones and tablets. At the same time, displays big and small are becoming much more capable of rendering the deep darks and bright highlights that make HDR (1,2) special. Together, adaptive streaming combined with advanced display technologies can enable HDR to become a high-value second screen experience.

Ideally, we would like to create a consistent HDR experience across all screens. A challenge is that small screens tend to have lower resolution than big screens. They also tend to be used more often at the edge of lower-bandwidth wireless networks.

Adaptive streaming protocols compensate for these issues by making several variants of media content available to adaptive streaming clients. The set of variants, often called an adaptation set, is made up of several representations of the original content at lower resolutions and at different bitrates. Adaptive streaming clients select a variant from those available in the adaptation set that makes best use of the client’s bandwidth availability and rendering capabilities. As bandwidth and other conditions fluctuate, the client can adapt by selecting different variants.

The main advantage of adaptive streaming is smooth uninterrupted playback. For HDR, however, there are potential negative side effects related to the loss of spatial detail and localized highlights when clients select lower-resolution lower-bitrate variants.

The key question for service providers is this: Can we design and construct adaptation sets in such a way as to minimize HDR distortions and promote consistency across screens? This paper provides methods to help achieve that goal. This paper provides methods to quantify HDR distortions, take steps to mitigate those distortions, and select the best combinations of bitrates and resolutions to include in HDR adaptation sets.

TEST SEQUENCES & PREPARATION

In this study, we used the HDR Wide Colour Gamut (WCG) test sequences shown in Figure 1. These sequences were created by the “HdM-HDR-2014 Project” (3,4) to provide professional quality cinematic wide gamut HDR video for the evaluation of tone mapping operators and HDR displays. All clips are 1920x1080p24 and colour graded for ITU-R BT.2020 (ref. 5) primaries & 0.005-4000 cd/m^2 luminance.

Figure 1 - HDR Test Sequences Used in this Study

We converted the original colour graded frames (RGB 48 bits per pixel TIFF files) to Perceptual Quantizer (PQ (ref. 6)) YCbCr v210 format (4:2:2 10 bit) using the equations defined in ITU-R BT.2020, ITU-R BT.1886 (ref. 7), and ITU-R BT.2100 (ref. 8).

For each test sequence, we created 50 variants having different encoded resolutions and bitrates. Of the 50 variants, 20 were raw uncompressed versions used to isolate the impact of different rescaling algorithms on HDR distortion. The remaining 30 variants for each test sequence were compressed using High-Efficiency Video Coding (HEVC) for each combination of encoded resolution (1920x1080, 1440x1080, 1280x720, 960x540, 720x540, and 640x360) and bitrate (10000, 3000, 1000, 300, and 100 kbps). All rescaling was performed in Matlab (9) using the *imresize* function. All compression was performed using command-line x265 (10) (main10 profile).

OPTIMIZING ADAPTIVE STREAMING RESOLUTIONS & BITRATES FOR HDR

HDR Distortion and Video Quality Metrics

For non-HDR adaptive streaming, our industry has experience using objective video quality metrics (11,12) to help create consistency. Commonly used objective video quality metrics such as Peak-Signal-to-Noise Ratio (PSNR), Mean-Squared Error (MSE), Structural Similarity Index (SSIM (ref. 13)), and Multiscale Structural Similarity Index (MS-SSIM (ref. 14)) are useful and practical even if it can be argued that none are yet a perfect substitute for human eyes (15).

For HDR, there is not yet an equivalent widely accepted and trusted objective video quality metric. MSE, PSNR, SSIM, MS-SSIM by themselves are not sufficient for the range of luminance found in HDR video, nor do they adequately report the heightened significance of localized highlights and details in HDR content. The development of HDR-aware metrics is progressing in academic and company labs (16-25), but they have not yet reached the level of maturity required for commercial television services.

In this paper, we instead leverage well-known distortion metrics such as MSE and linear-correlation coefficients but do so in a nuanced and focused manner that helps unveil HDR-specific differences between original and encoded content.

Decomposition of HDR

Our approach to measuring HDR distortion begins with decomposing each HDR frame into a Spatial Detail signal and a Basal Image, as illustrated in Figure 2.

Figure 2 – Decomposition of original HDR frame into Spatial Detail and Basal Image

The method of calculating the Spatial Detail signal is in the Appendix and even more detail is in previous publications (26-30). In brief, the Spatial Detail signal can be thought of as the condensed essence of the original image. It isolates the features and details that are unique to the original while minimizing statistically expectable characteristics that the original image shares with images as a statistical class.

Images of natural and other complex scenes have an interesting statistical property: They have spatial-frequency magnitude spectra that tend to fall off with increasing spatial frequency in proportion to the inverse of spatial frequency (30). The magnitude spectra of individual images can vary significantly; but, as an ensemble-average statistical expectation, it can be said that “the magnitude spectra of images of natural and other complex scenes fall off as one-over-spatial-frequency.”

The Spatial Detail signal is effectively the result of de-emphasizing the statistically expectable one-over-frequency characteristic. As such, the Spatial Detail signal emphasizes the unique unexpectable details in an image.

The Basal Image is obtained by simply subtracting the Spatial Detail image from the original HDR image. As such, the Basal Image may be thought of as a special kind of low-pass filtered version of the original HDR image in which the unique spatial details are selectively attenuated. Although perhaps difficult to appreciate on the printed page, the Basal Image gives the visual sensation of being out of focus.

One way to think of the Basal Image is that it represents the overall contrast and luminance range of the HDR image, whereas the Spatial Detail signal can be thought of as representing localized contrast and luminance variations.

Spatial Detail Distortion is Most of the Total HDR Distortion

An advantage of decomposing each HDR frame into a Spatial Detail signal and Basal Image can be appreciated by examining the relative contribution of each to total HDR distortion.

Mean-Squared Error and PSNR are common and equivalent metrics of distortion. (PSNR is proportional to the logarithm of MSE.) MSE is the average over all pixels of the squared difference between an original image and a corresponding encoded variant.

MSE values calculated for five encoded resolutions and four candidate rescaling algorithms are shown in Figure 3. The candidate rescaling algorithms are nearest-neighbour interpolation, bilinear interpolation, bicubic interpolation, and lanczos3 resampling. The MSE values shown in Figure 3*A* are the average values of 15-second segments and of all test sequences. Lower MSE values indicate that the rescaled variant is less distorted from the original in terms of squared-error (and thus PSNR). The data show that lanczos3 resampling provides the lowest MSE values for all resolutions and should thus be considered the best choice in constructing adaptive streaming variants. If other considerations such as processing demands are significant, bicubic interpolation can deliver nearly as good results.

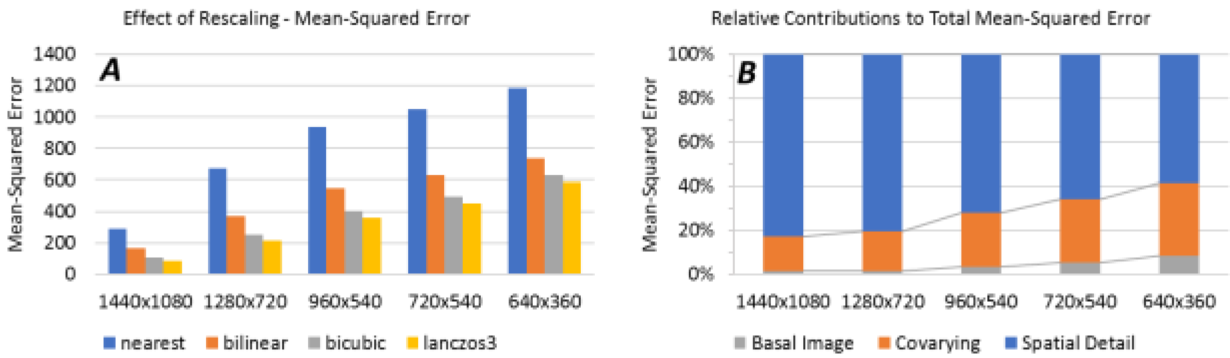


Figure 3 – Mean-Squared Error for Different Rescaling Algorithms (*A*) and the Relative Contribution of the Spatial Detail signal and Basal Image to Total MSE (*B*).

Total MSE can be thought of as the sum of the Spatial Detail MSE by itself, the Basal Image MSE by itself, and a contribution from the covariance of Spatial Detail signal and the Basal Image.

The data in Figure 3*B* show that the Basal Image MSE is a small fraction of the total MSE. The Basal Image contribution to total MSE increases with progressively more aggressive downscaling, which indicates that rescaling progressively distorts the underlying smoothly-varying luminance of the encoded video. Yet, even for the 3-fold downscaling from the original 1920x1080 to 640x360, the error associated with the Basal Image is only about 10% of the total error. Most of the total error is attributable to distortion of the Spatial Detail signal (approximately 80% for 1440x1080 and 60% for 640x360). (The data shown in Figure 3*B* are for lanczos3 resampling.)

Spatial Detail and the Basal Image also make different contributions to total MSE for HEVC-compressed variants, as illustrated in Figure 4. At high bitrates, the Spatial Detail signal (***A***) contributes ~60% and the Basal Image (***B***) contributes ~10% of the total MSE. At very low bitrates (100 kbps) the relative contributions become more equalized with Spatial Detail and the Basal Image both contributing ~35%.

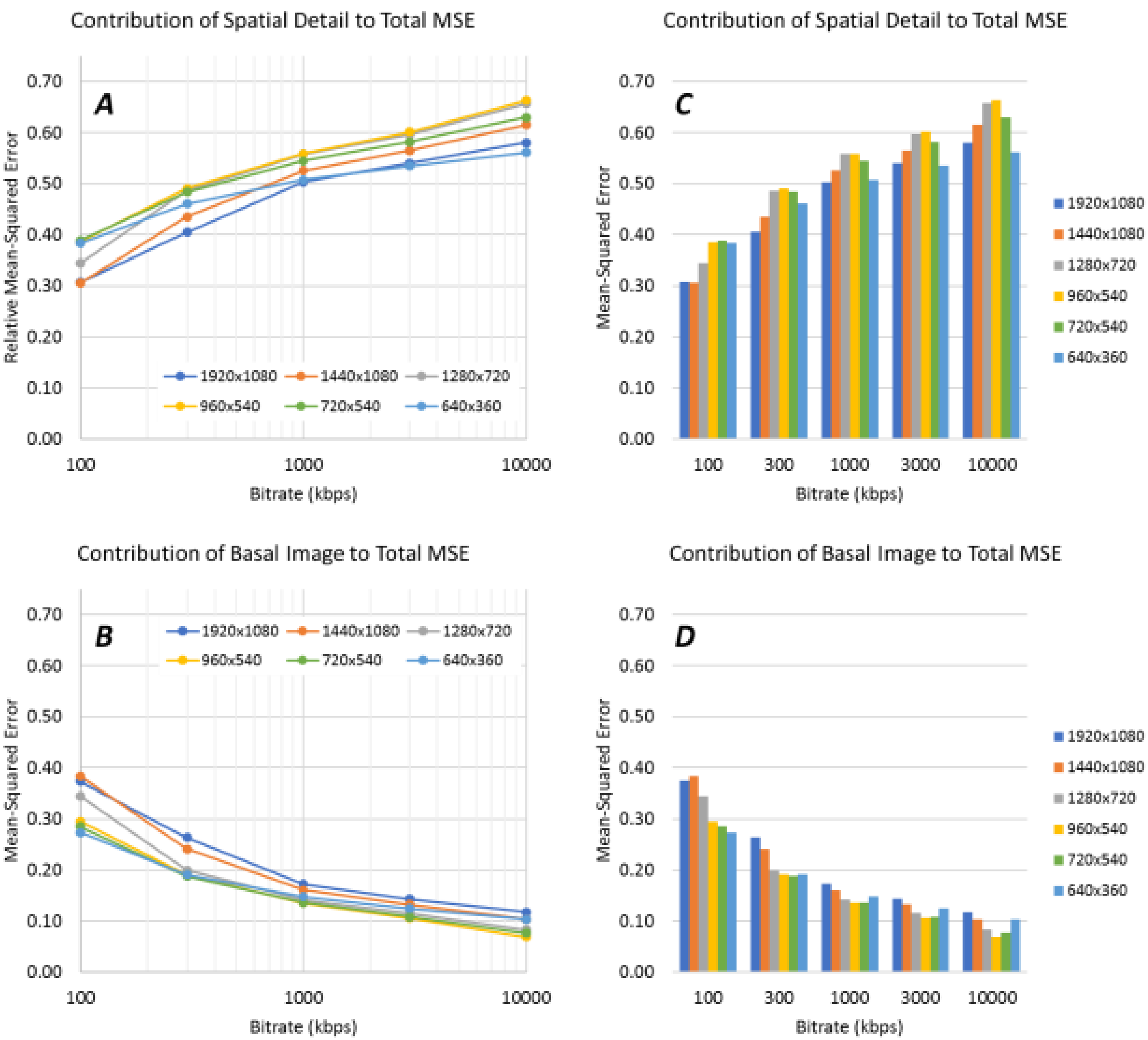


Figure 4 – Relative Mean-Squared Error for HEVC-Compressed Variants

Note that the relative contributions from Spatial Detail and the Basal Image tend to be complimentary. For example, variants with an encoded resolution of 960x540 maximize the Spatial Detail contribution (***C***) and minimize the Basal Image contribution (***D***) for almost all bitrates. Variants with higher and lower encoded resolutions tend to reduce the contribution from Spatial Detail and increase the contribution from the Basal Image.

Using Basal Image MSE to Exclude Low Video-Quality Variants

The size of the contribution of the Basal Image to total MSE can be used as an acceptance threshold to determine which combinations of bitrate and encoded resolutions should be considered for exclusion from an adaptation set. A large contribution by the Basal Image to total MSE, above 20% in our experience, is associated with very poor visual quality. A large contribution means that the major underlying contrast variations of the HDR content have been significantly and very noticeably distorted.

The data in Table 1 shows that the relative contribution of the Basal Image to total MSE can also be used to optimize selection of permissible variants in a content-adaptive manner. The red cells in the table denote combinations of bitrate and encoded resolution to be excluded when designing adaptation sets. The threshold used in this example is 0.20.

Table 1 – Setting a Threshold for Video Quality Using Basal Image MSE

	Relative Contribution of Basal Image to Total MSE									
Resolution	bistro					carousel fireworks				
	Bitrate (kbps)									
	100	300	1000	3000	10000	100	300	1000	3000	10000
1920x1080	0.38	0.21	0.16	0.13	0.11	0.31	0.25	0.17	0.14	0.12
1440x1080	0.30	0.17	0.12	0.10	0.08	0.33	0.23	0.16	0.13	0.10
1280x720	0.20	0.12	0.08	0.07	0.05	0.35	0.20	0.15	0.12	0.09
960x540	0.19	0.11	0.08	0.06	0.04	0.28	0.19	0.14	0.11	0.07
720x540	0.19	0.13	0.09	0.08	0.07	0.27	0.19	0.14	0.11	0.08
640x360	0.20	0.16	0.14	0.13	0.13	0.27	0.19	0.14	0.12	0.09

Using Total MSE to Choose Resolution & Bitrate Combinations for Adaptation Sets

The data in Figure 5 illustrate an existing method of selecting which combinations of bitrate and encoder resolution should be included when designing adaptation sets. The methodology is based on the ATIS-0800061(31) standard, which was developed as a joint effort by the Video Services Forum and the IPTV Interoperability Forum. The core concept is illustrated in Figure 5*A*. High-resolution representations of original content have better video quality (lower MSE) at high bitrate than do low-resolution representations. On the other hand, video quality deteriorates faster with bitrate for high-resolutions representations than for low-resolution representations. As a result, low-resolution representations have better video quality at low bit rates than do high-resolutions representations. Finding the points at which the high-resolution curves cross the low-resolution curves (upward arrows) provides a systematic way to determine which encoded resolution provides the best video quality at each bitrate. The selection process is illustrated in a different way in Figure 5*B*. For any set of possible encoded resolutions, there is one that results in the minimum total MSE at each bitrate (downward arrows).

Table 2 summarizes the results for the example shown in Figure 5. The green cells indicate the bitrate and encoded resolution combinations that minimize total MSE and might thus be considered for inclusion in the adaptation set.

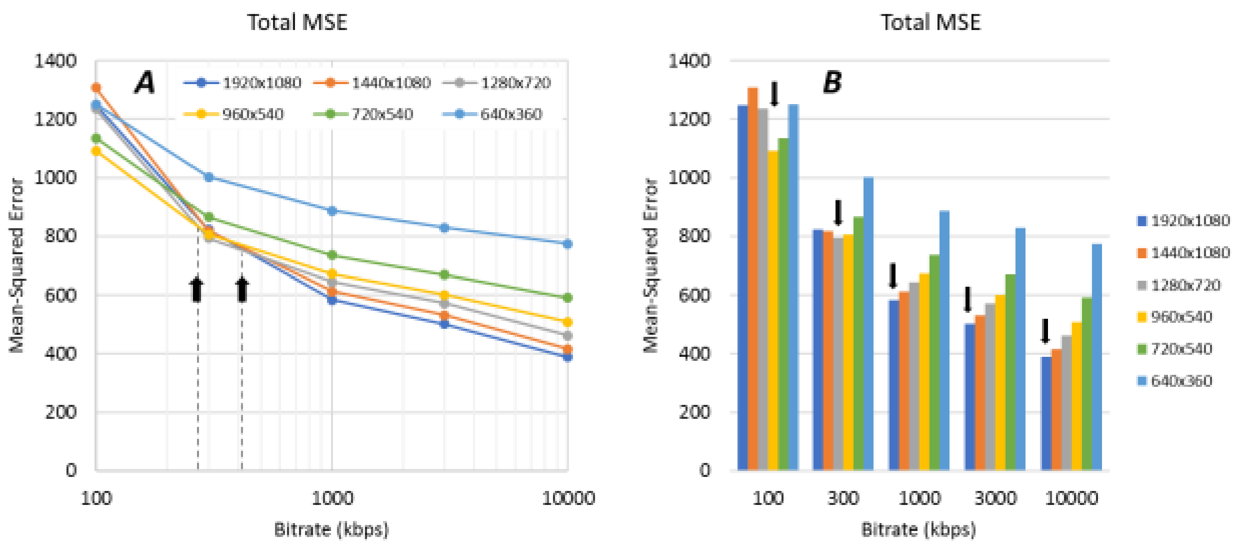


Figure 5 – Designing an Adaptiation Set Based on Mimimum Total Mean-Squared Error

Table 2 – Minima of Total Mean-Squared Error

Total Mean-Squared Error					
Resolution	**Bitrate (kbps)**				
	100	**300**	**1000**	**3000**	**10000**
1920x1080	1248	823	584	502	388
1440x1080	1308	818	613	532	416
1280x720	1235	794	645	572	463
960x540	1093	806	672	602	509
720x540	1135	865	736	669	591
640x360	1251	1003	888	829	775

Using Spatial Detail Correlation to Choose Resolutions & Bitrates for Adaptation Sets

Mean-squared error by itself is known to be a convenient but not very accurate predictor of human opinions of video quality, particularly regarding HDR content. Fortunately, the methodology illustrated in Figure 5 is not restricted to MSE. It can be applied to other objective measures of HDR distortion or video quality.

We applied the methodology to a metric that describes how well correlated the encoded Spatial Detail is with the original Spatial Detail. The metric we use is the coefficient of determination, R^2, the square of the Pearson correlation coefficient, R, (see ref. 33).

In this study, we calculated R^2 values using the average code values in the encoded frame compared to the corresponding code values in the original frame. The expectation is that the average code value of the encoded frame would be the same as the corresponding code value of the original frame. Mismatches are a manifestation of a lack of correlation and result in a lower value of R^2, which has a range of 0 to 1.

(A note on calculating average code values – Each average code values in the encoded frame was calculated by finding all the pixel locations in the original frame that have a specific code value, for example, 312 out the possible range of 64 to 940 for 10-bit encoding. The encoded code values for the corresponding pixel locations tend to have a distribution of values because of compression and scaling. We use the average over the distribution. The resulting average encoded code values were compared to the original code values to calculate R^2 values. The R^2 values in this study are thus a measure of the correlation between actual mean values and expected values.)

The data in Figure 6*A* illustrate the correlation between original luma code values and encoded luma code values for an encoded resolution of 1920x1080 and bitrates from 100 kbps to 10 Mbps. The encoded values display an almost perfect correlation with the original values: The encoded values fall along a straight line with a slope of 1. The exception is for dark regions having code values below ~200. The corresponding dark regions in the encoded frames tend to brighter. In other words, HEVC compression is causing an elevation of the average black levels of HDR content. The amount of brightening increases with decreasing bitrate. (We have been investigating the elevation of blacks in HEVC-compressed video. It appears to be a result of low-pass filtering of textures and film grain by the internal HEVC image processing operations.)

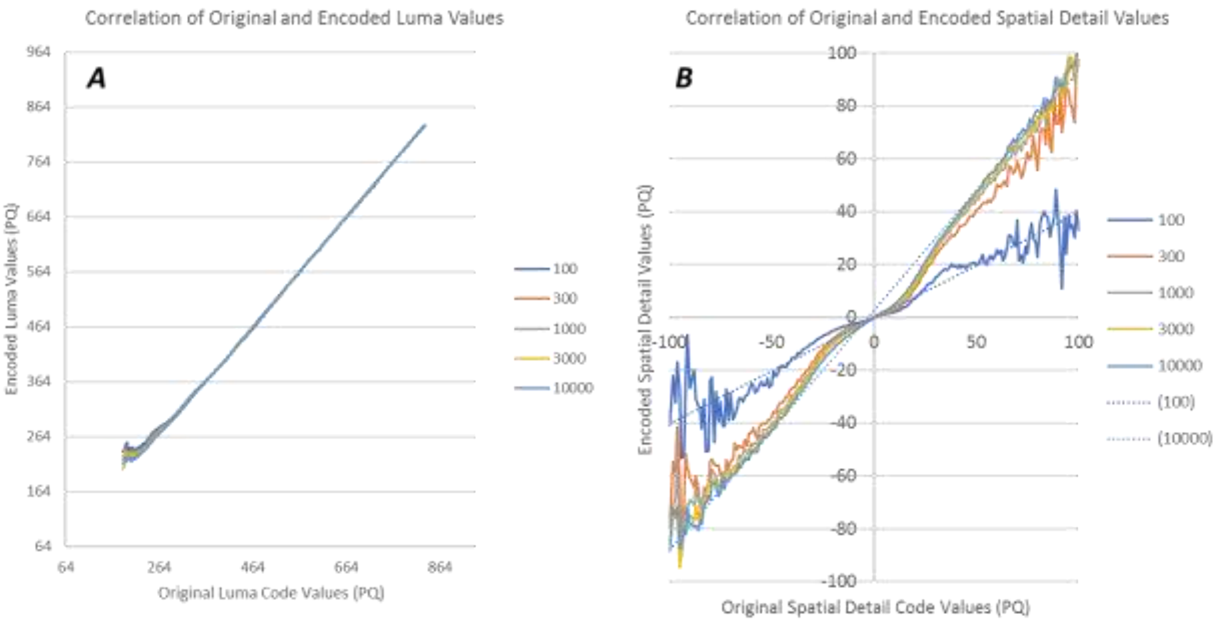


Figure 6 – Correlations of Luma Code Values and Spatial Detail Signals

The data in Figure 6*B* show the correlations of the original and encoded Spatial Detail signals for the corresponding luma correlations shown in Figure 6*A*. The encoded Spatial Detail signals display weaker correlations with the original Spatial Detail signals compared to the corresponding luma signals. There is a general overall linear relationship between encoded and original Spatial Detail, but there are four significant differences worth noting. First, there is much greater variation around the straight-line fitted using least-mean-squared regression (shown as dashed lines). Second, the magnitude of the variation around the fitted straight-line increases with decreasing bitrate. This indicates that Spatial

Detail correlation decreases with decreasing bitrate. Third, the slope of the fitted straight-line decreases with decreasing bitrate. This indicates that localized contrasts of textures and HDR highlights are diminished with decreasing bitrate. Fourth, the slope of the correlation is flatter neat the origin (small Spatial Detail values) than for large Spatial Detail values. This indicates that low contrast textures and details (such as those typically associated with faces and background textures) are systematically impacted more severely by HEVC compression than are high contrast HDR textures.

The variation of the data around the best-fitting least-mean-squared regression line (dashed lines) is quantified by the value of R^2, which measures the "goodness of fit" between the fitted straight line and the actual data.

Table 3 summarizes the goodness-of-fit, R^2 values, for the bitrate-resolution combinations used in this study. The cells highlighted in green indicate which encoded resolution maximizes the goodness-of-fit for each bitrate. In other words, the green-highlighted cells correspond to the bitrate and resolution combinations that maximize the correlation between encoded Spatial Detail and original Spatial Detail. These are the bitrate-resolution combinations that best preserve the textures, highlights, and local contrast variations in HDR video.

Table 3 – Maxima of Spatial Detail Correlation (R^2)

Resolution	Bitrate (kbps)				
	100	300	1000	3000	10000
1920x1080	0.900	0.933	0.946	0.966	0.983
1440x1080	0.893	0.939	0.949	0.968	0.982
1280x720	0.919	0.942	0.955	0.965	0.976
960x540	0.903	0.934	0.946	0.955	0.962
720x540	0.892	0.924	0.935	0.943	0.950
640x360	0.869	0.897	0.902	0.909	0.914

Table 4 illustrates the use of Spatial Detail correlation to choose the best variants to include in HDR adaptation in a content-aware manner. Note that the *carousel_fireworks* test sequence benefits from lower resolution variants than does the *bistro* test sequence. The *carousel_fireworks* test sequence is more challenging because it has a wider range of luminance and more motion.

Table 4 – Choosing Bitrate & Resolution Combinations based on Spatial Detail Correlation

Resolution	bistro					carousel_fireworks				
	100	300	1000	3000	10000	100	300	1000	3000	10000
1920x1080	0.958	0.987	0.995	0.997	0.999	0.841	0.848	0.878	0.942	0.985
1440x1080	0.956	0.990	0.995	0.997	0.998	0.848	0.879	0.898	0.956	0.989
1280x720	0.979	0.990	0.993	0.995	0.996	0.841	0.901	0.932	0.966	0.986
960x540	0.963	0.985	0.987	0.988	0.988	0.860	0.906	0.927	0.948	0.968
720x540	0.966	0.983	0.988	0.988	0.989	0.866	0.916	0.931	0.946	0.960
640x360	0.950	0.966	0.967	0.967	0.968	0.869	0.912	0.920	0.927	0.929

DISCUSSION

In this paper, we provide methods to quantify HDR distortions and take steps to mitigate those distortions by selecting the best combinations of bitrates and resolutions to include in HDR adaptation sets used in adaptive streaming services.

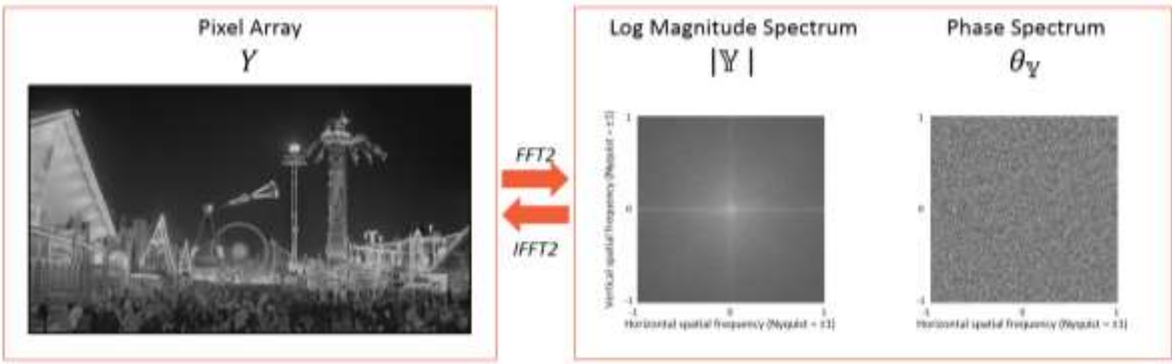
A key part of our approach is to decompose HDR video into two components. A Basal Image contains the overall luminance and contrast variations in HDR video. A Spatial Detail signal contains the localized luminance and contrast variations.

We showed that most of the distortion in encoded HDR content is a result of distortions of the Spatial Detail signal. On the other hand, the mean-squared error associated with distortions of the Basal Layer is a useful metric with which to set minimum-video-quality. We propose that a good video-quality acceptance-threshold is the point at which the Basal Image contributes 20% of the total mean-squared error between the original and encoded HDR videos.

We also propose that the correlation between the encoded and original Spatial Detail is a useful metric by which to select the best combinations of bitrate and resolution for inclusion when designing adaptation sets. This approach maximizes the similarity of the local contrasts and highlights between encoded and original HDR videos. These local contrasts and highlights add significant impact to HDR video and represent a large part of the content creator's original intent.

APPENDIX – SPATIAL DETAIL AND BASAL IMAGE

The method of creating the Spatial Detail signal can perhaps best be understood by thinking of an image in terms of spatial frequency spectra as illustrated in Figure A1 (only the luma channel is shown). Any 2-dimensional array of pixel values can also be represented without loss of information as the product of a magnitude spectrum and a phase spectrum in 2-dimensional spatial frequency space. Spatial-frequency spectra can be obtained from an image pixel array by performing a 2-dimensional Fast Fourier Transform (FFT2). The pixel array can be recovered by performing a 2-dimensional Inverse Fast Fourier Transform (IFFT2). FFT2 and IFFT2 are well known signal processing operations that can be calculated quickly in modern processors.



$$FFT2(Y(x,y)) = \mathbb{Y}(k_x,k_y) = |\mathbb{Y}(k_x,k_y)| * exp\left(i\theta_{\mathbb{Y}}(k_x,k_y)\right)$$

Figure A1 - Representation of a Video Frame in Terms of Spatial Frequency

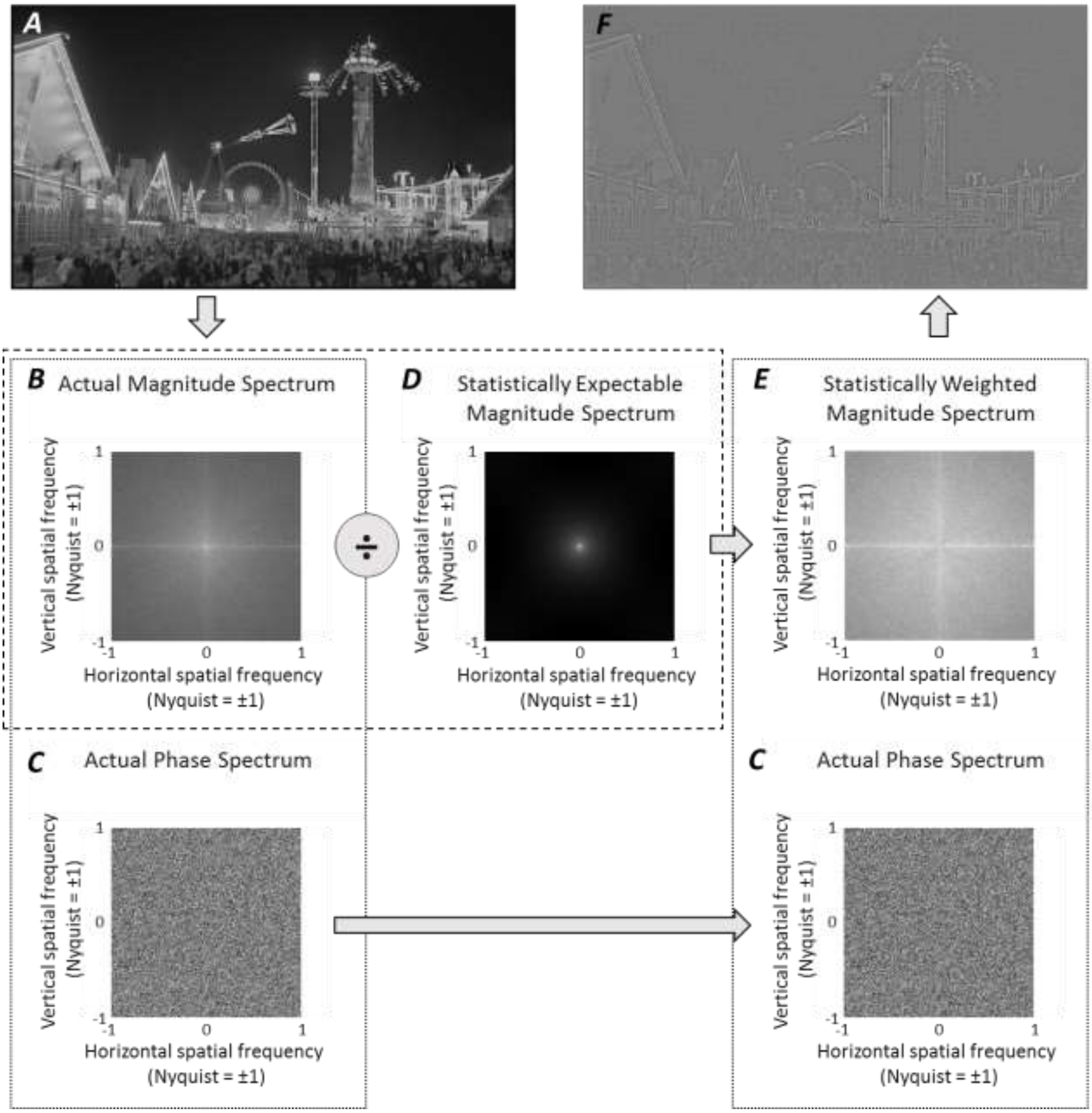


Figure A2 - Method of Calculating the Spatial Detail Signal

The Spatial Detail signal is calculated as illustrated in Figure A2. First, the magnitude (***B***) and phase spectra (***C***, shown twice) are calculated from the image pixel array (***A***). Next, a predetermined archetype of the statistically expectable one-over-frequency magnitude spectrum (***D***) is divided into the actual magnitude spectrum to produce a statistically weighted magnitude spectrum (***E***). Third, the statistically weighted magnitude spectrum is combined with the actual phase spectrum (***C***). Finally, a 2-dimensional Inverse Fast Fourier Transform is applied to produce a pixel array that we call the Spatial Detail signal (***F***).

REFERENCES

1. ITU-R Report BT.2246-2 (2017) "The present state of ultra-high definition television."
2. ITU-R Report BT.2390-2 (2017) "High dynamic range television for production and international programme exchange."
3. J., Froehlich, et al., "HdM-HDR-2014 Project," http://www.hdm-stuttgart.de/~froehlichj/hdm-hdr-2014
4. Froehlich, J., Grandinetti, S., Eberhardt, B., Walter, S., Schillin, A., and Brendel, H. 2014. "Creating cinematic wide gamut HDR-video for the evaluation of tone mapping operators and HDR-displays," *Proc. SPIE 9023, Digital Photography X*
5. ITU-R Rec. BT.2020-2 (2015) "Parameter values for ultra-high definition television systems for production and international programme exchange."
6. Miller, S., Nezamabadi, M., and Daly, S. 2013. "Perceptual Signal Coding for More Efficient Usage of Bit Codes," *SMPTE Motion Imaging Journal*
7. ITU-R Rec. BT.1886 (2011) "Reference electro-optical transfer function for flat panel displays used in HDTV studio production."
8. ITU-R Rec. BT.2100-0 (2016) "Image parameter values for high dynamic range television for use in production and international programme exchange."
9. MathWorks, MATLAB. https://www.mathworks.com/
10. x265 (x265-64bit-10bit-2017-05-01.exe) https://builds.x265.eu/
11. Winkler, S. 2005 Digital Video Quality: Vision Models and Metrics, John Wiley & Sons
12. VQEG. 2000. "Final report from the video quality experts group on the validation of objective models of video quality assessment."
13. Wang, Z., Bovik, A.C., Sheikh, H.R., and Simoncelli, E.P. 2004. "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612*
14. Wang, Z., Simoncelli, E. P., and Bovik, A.C. 2003. "Multi-scale structural similarity for image quality assessment." *Proc. of the 37th IEEE Asilomar Conference on Signals, Systems, and Computers*
15. Wang, Z. and Bovik, A.C. 2009. "Mean squared error: love it or leave it? - A new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98-117
16. Mantiuk, R., Daly, S., Myszkowski, K., and Seidel, H.-P. 2005. "Predicting visible differences in high dynamic range images: model and its calibration," *SPIE Human Vision and Electronic Imaging X*, vol. 5666
17. Aydin, T.O., Mantiuk, R., and Seidel, H.-P. 2008. "Extending quality metrics to full luminance range images," Human Vision and Electronic Imaging XIII. Edited by Rogowitz, Bernice E.; Pappas, Thrasyvoulos N. Proceedings of the SPIE, Volume 6806
18. Mantiuk, R., Kim, K.J., Rempel, A.G., and Heidrich, W.. 2011. "HDR-VDP-2: A calibrated visual metric for visibility and quality predictions in all luminance conditions," *ACM Trans. Graph.* 30(4), 40:1–40:14
19. Hanhart, P., Kroshunov, P., and Ebrahimi, T. 2014. "Subjective evaluation of higher dynamic range video," *Proceedings of SPIE - The International Society for Optical Engineering*

20. Hanhart, P., Bernardo, M.V., Korshunov, P., and Pereira, M. 2014. “HDR image compression: A new challenge for objective quality metrics,” in *Sixth International Workshop on Quality of Multimedia Experience (QoMEX)*
21. G. Valenzise, F. De Simone, P. Lauga, and F. Dufaux, “Performance evaluation of objective quality metrics for HDR image compression,” *Proc. SPIE 9217, Applications of Digital Image Processing XXXVII,* 2014
22. Mantel, C., Ferchiu, S.C., and Forchhammer, S. 2014 “Comparing subjective and objective quality assessment of HDR images compressed with JPEG-XT,” *in 16th International Workshop on Multimedia Signal Processing (MMSP), IEEE*
23. Hanhart, P., Bernado, M.V., Pereira, M., Pinheiro, M.G., and Ebrahimi, T. 2015. “Benchmarking of objective quality metrics for HDR image quality assessment,” *EURASIP J. Image & Video Processing*
24. Narwaria, M., Mantiuk, R.K., Perreira Da Silva, M., and Le Callet, P. 2015. “HDR-VDP-2.2: a calibrated method for objective quality prediction of high-dynamic range and standard images,” *J. Electron. Imaging.* 24(1), 010501
25. Rerabek, M., Hanhart, P., Korshunov, P., and Ebrahimi, T. 2015. “Subjective and objective evaluation of HDR compression,” *International Workshop on Video Processing and Quality Metrics for Consumer Electronics - VPQM,* Chandler, Arizona, USA
26. McCarthy, S.T. 2017. “A Biologically-Inspired Approach to Making HDR Video Quality Assessment Easier” SMPTE Motion Imaging Journal, vol. 124, no. 4, pp 47-58
27. McCarthy, S.T. 2014. “Theory and practice of perceptual video processing in broadcast encoders for cable, IPTV, satellite, and internet distribution,” *Proc. SPIE 9014, Human Vision and Electronic Imaging XIX*
28. McCarthy, S. 2012. “A Biological Framework for Perceptual Video Processing and Compression,” *SMPTE Mot. Imag. J.*, 119(8):24-32, Nov/Dec.
29. McCarthy, S.T. and Owen, W.G. 2006. “Apparatus and Methods for Image and Signal Processing”. US Pat. 6014468 (2000). US Pat. 6360021 (2002), US Pat. 7046852 (2006)
30. Field, D.J. 1987. “Relationship between the statistics of natural images and the response properties of cortical cells,” *J. Opt. Soc. Am. A. Vol. 4, No. 12*
31. ATIS 0800061 (2103) “Methodology for subjective or objective video quality assessment in multiple bit rate adaptive streaming”
32. Bansal, G. 2017. “What is the difference between coefficient of determinations and coefficient of correlation?” http://blog.uwgb.edu/bansalg/statistics-data-analytics/linear-regression/what-is-the-difference-between-coefficient-of-determination-and-coefficient-of-correlation/