Published in: Proceedings of the 47th IEEE Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, Nov. 3-6, 2013. ©IEEE

On the Use of SSIM in HEVC

Tiesong Zhao, Kai Zeng, Abdul Rehman and Zhou Wang
Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, N2L3G1, Canada
Email: {ztiesong, kzeng, abdul.rehman, zhou.wang}@uwaterloo.ca

***Abstract*—The Structural SIMilarity (SSIM) index has been attracting an increasing amount of attention recently in the video coding community as a perceptual criterion for testing and optimizing video codecs. Meanwhile, the arrival of the new MPEG-H/H.265 High Efficiency Video Coding (HEVC) standard creates new opportunities and challenges in perceptual video coding. In this paper, we first elaborate what are the attributes that make SSIM a good candidate for perception-based development of HEVC and future video coding standards for both testing and optimization purposes. We then address the computational issues in practical applications of SSIM in HEVC, in particular the trade-off between efficient computation and accurate estimation of SSIM when working with video codecs that have sophisticated block partitioning structures and aim for encoding videos with a wide range of spatial resolutions.**

***Keywords*—video coding, high efficiency video coding (HEVC), H.265, MPEG-H, video quality assessment, structure similarity (SSIM), perceptual optimization, perceptual video coding**

I. Introduction

The exponential growth in video traffic over various networks in the past years has created tremendous opportunities and challenges for video technology developers to produce high performance video codecs that can largely reduce the bandwidth of video streams while maintaining their integrity and quality. The most recent milestone in video codec development is the new MPEG-H/H.265 High Efficient Video Coding (HEVC) technique [2], which has just passed the stage of standard approval in 2013, but has already been quickly disseminated in the industry. Essentially all state-of-the-art video coding standards, including HEVC and the currently dominating MPEG-4/H.264 AVC standard [1], can be summarized by a hybrid framework of motion-handling and picture-coding, with a Rate-Distortion Optimization (RDO) technique to minimize the distortion subject to a constraint on Bit Rate (BR) [3]. A frequent criticism on the current approach is on the definition of distortion, for which traditional distortion/quality measures such as Sum of Absolute Difference (SAD), Mean Squared Error (MSE), and Peak-Signal-to-Noise-Ratio (PSNR), have been found to be poorly correlated with perceived video quality by the Human Visual System (HVS), which is the ultimate receivers of most visual communication applications.

To overcome the limitations of MSE/PSNR, another area that has undergone significant development in the past decade is perceptual objective video quality assessment (VQA). There have been several objective VQA methods that are widely noted to provide effective predictions of perceived video quality. These include Video Quality Metric (VQM) [7], the Structural SIMilarity (SSIM) index family [4], [5], [6], [23], and the MOtion-based Video Integrity Evaluation (MOVIE) index [8]. Among them, SSIM has appeared to be the most widely spread method in recent years as an alternative quality measure of PSNR in the video coding community. Indeed, a significant amount of effort has been made to improve video compression efficiency whilst maintaining SSIM-based video quality. For example, SSIM has been incorporated into H.264/AVC coding mode selection [12], Motion Estimation (ME) [13], and RDO algorithms [11], [14], [15], [16], [17], [18], [19], [20], [21], [22]. Most recently, SSIM optimization has also been embedded into HEVC, leading to notable coding gain [29]. It is also worth noting that in the use of SSIM in video coding, a number of variations have been employed in the computation of SSIM to better fit the computational structures of video coding algorithms. For example, as opposed to the 2D Gaussian sliding window employed in the original SSIM computation [5], different sizes of square windows have been used to calculate block SSIM, including 16×16 sliding windows in [11], [12], [13], [16], 8×8 sliding windows in [14], [15] and 8×8 non-overlapping windows in [17], [20], [21], [22]. In [11], [12], [13], [16], [17], [20], SSIM of a block is obtained by averaging the SSIM values of its 4×4 non-overlapping child blocks, in [14], [15], block SSIM is calculated in the full current block directly, while in [21], [22], 4×4 sliding windows are used for SSIM calculation in an extended block. Nevertheless, in spite of the fast-increasing usage of SSIM, discussions regarding why and how to use SSIM in video coding are lacking.

In this paper, we attempt to address two issues regarding the use of SSIM in HEVC and possibly other existing and future video coding standards. First, given multiple options of VQA algorithms, why choose SSIM for video coding? In particular, what are the attributes that make SSIM stands out from other VQA models? If in the future we could find other alternative models to replace SSIM, what would be the must-have properties of these models? Second, given the wide range of target spatial resolutions and sophisticated block partitioning and coding structures of modern video codecs such as HEVC, how to compute or estimate SSIM efficiently without losing its quality prediction accuracy?

II. Why SSIM?

There are generally two ways of using a VQA measure in video coding. The first is to use it as a *testing tool*, so as to choose the best quality video from multiple coded videos, or the best video coding algorithm/configuration from multiple choices of coding methods or configurations. The second is to use it as an *optimization tool*, where the VQA measure is embedded into the design and optimization of video coding algorithms for best coding efficiency. There are a number of useful attributes of VQA measures, and depending on the way they are used, the desirability of these attributes could vary

TABLE I. DESIRABLE ATTRIBUTES OF VQA MEASURES FOR VIDEO CODING

Attribute	Desirability as Testing Tool	Desirability as Optimization Tool	PSNR	SSIM
Correlated with perceptual quality	High	High	Fair	Very Good
Low complexity	Moderate	High	Excellent	Very Good
Good mathematical properties	Low	High	Excellent	Good
Localized quality prediction	Moderate	High	Excellent	Excellent
Saturated at high rate	Moderate	High	Poor	Excellent
Popularity	High	High	Excellent	Excellent

Desirability scale: {Low, Moderate, High}; Attribute quality scale: {Poor, Fair, Good, Very Good, Excellent}

significantly. These are summarized in Table I and elaborated as follows.

TABLE II. QUALITY PREDICTION PERFORMANCE COMPARISON OF VQA ALGORITHMS

VQA Model	PLCC	MAE	RMS	SRCC	KRCC
PSNR	0.5408	1.1318	1.4768	0.5828	0.3987
VQM [7]	0.8302	0.7771	0.9768	0.8360	0.6243
MOVIE [8]	0.7164	0.9711	1.2249	0.6897	0.4720
SSIM [5]	0.8422	0.8102	0.9467	0.8344	0.6279
MS-SSIM [23]	0.8526	0.7802	0.9174	0.8409	0.6350

TABLE III. COMPUTATIONAL COMPLEXITY COMPARISON OF VQA ALGORITHMS (NORMALIZED BASED ON PSNR)

VQA Model	PSNR	VQM [7]	MOVIE [8]	SSIM [5]	MS-SSIM [23]
Complexity	1	1083	7229	5.874	11.36

1) Correlated with perceptual quality. There is no doubt that this is the most desirable attribute of VQA models for both testing and optimization purposes. In [25], a subjective test was conducted to evaluate the perceptual quality of HEVC-HM5.0 and H.264/AVC JM18.3 coded video sequences. The resulting Mean Opinion Scores (MOSs) given by human subjects are compared with popular VQA measures in [26]. The results are summarized in Table II, where five VQA models (PSNR, VQM [7], MOVIE [8], SSIM [5] and MS-SSIM [23]) are evaluated using five criteria, including Pearson Linear Correlation Coefficient (PLCC), Mean Absolute Error (MAE), Root Mean Squared Error (RMS), Spearman Rank-order Correlation Coefficient (SRCC), and Kendall Rank-order Correlation Coefficient (KRCC). Not surprisingly, PSNR performs poorly in the test. MS-SSIM delivers the best performance, but its advantage over SSIM and VQM may not be statistically significant.

2) Low complexity. Given the large volume of video data being created and transmitted every day, speed becomes a key issue in the practical use of VQA techniques. High-accuracy high-complexity VQA models are still useful but their applications are very limited. In time-critical applications such as real-time network quality monitoring, low complexity VQA systems are highly desirable. The meaning of complexity here is not only restricted to computational complexity. The complexity in the structures and implementations of the VQA models should also be taken into consideration. This is particularly important when the models are used as optimization tools in video coding systems, which may require iterative applications of the VQA measures. In [26], the computational complexity of popular VQA algorithms are compared, and the results are given in Table III, where all complexity values are normalized based on PSNR. It can be observed that different VQA models could have drastically different complexity. Combining the results in Tables II and III, SSIM and MS-SSIM achieve excellent trade-off between accuracy and complexity, relative to VQM and MOVIE. Note that the results here are based on unoptimized Matlab implementations. Further speedup can be achieved by adapting the computation to video coding structures (as will be discussed in Section III) or through algorithm, software and hardware optimizations [27].

3) Good mathematical properties. This might not be a very useful attribute when the VQA models are employed solely as testing tools, but becomes critically important when they are deeply embedded into the design of video codec systems as optimization tools, e.g., in the RDO process. In this aspect, PSNR (and associated MSE) is excellent because it bears almost all the desirable mathematical properties, which are often overlooked in the traditional development of VQA models [10]. Fortunately, SSIM (and its direction variations) holds several mathematical properties, making it a much better option than other VQA models in video codec optimizations. Specifically, $\sqrt{1-SSIM}$ can be a valid distance metric (that satisfies identity and symmetry properties and triangular inequality). It is differentiable, locally convex and quasi-convex, and is distance preserving under orthogonal or unitary transformations. More detailed discussions can be found in [28]. These advantages have already been partially demonstrated in recent SSIM-optimal H.264 and HEVC video codec development [20], [21], [22], [29].

4) Localized quality prediction. This attribute allows us to create a "quality map" of the video being analyzed, which indicates how the quality varies as a function of space and time. The value of the quality map is not only in the diagnosis of quality degradations (such that severe artifacts can be detected), but also in the optimal bit allocation during the video coding process. In such a process, a quality map serves as a guide that directs the available bits to be distributed to the spatiotemporal locations that most require quality improvement. Both PSNR and SSIM are excellent in this attribute, which may not be true for other VQA models. In particular, SSIM has pixel precision in quality evaluation and local window precision in the computation being performed.

5) Saturated at high rate. With the increase of bit rate, the quality of the reconstructed video created by most video codecs improves. However, when the rate is higher than certain level, the reconstructed video becomes perceptually identical to the original video, and further increase of bit rate has no impact on visual quality. An ideal VQA model should well reflect such saturation effect of the HVS, especially when it is used as an optimization tool to guide the design of video codecs. Unfortunately, PSNR is not saturated at high rate, as exemplified in Fig. 1, where a sharp increase of PSNR is observed at high rate. Theoretically, the value of PSNR can be infinity. By contrast, SSIM is bounded by 1 and is saturated at high rate, as shown in Fig. 1. As a result, when SSIM

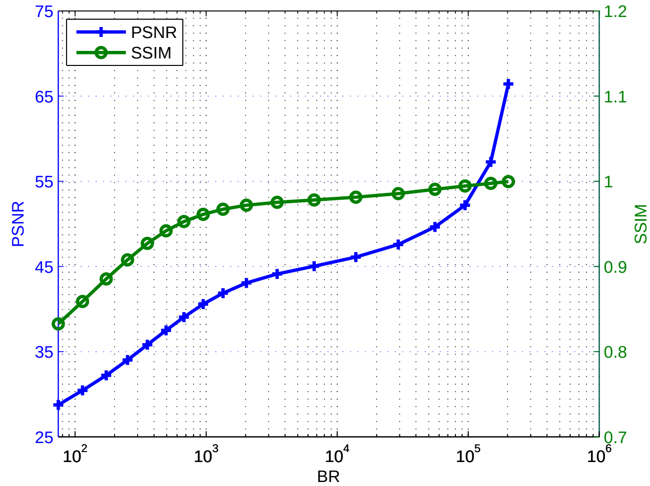

Fig. 1. A typical example of PSNR and SSIM versus bit rate (BR).

is employed in the optimization process of video coding, it will not guide the encoder to assign bits to the spatiotemporal locations where perceptual quality has already been saturated (but PSNR values may not). In general, the behavior of SSIM is very close to that of MOS at high rate, making it an excellent choice for perceptual optimization.

6) Popularity. Both PSNR and SSIM are popular VQA measures that are widely known, well understood, and frequently used in the image processing and video coding community. Although popularity is not a technical issue, it is highly desirable such that researchers and engineers in the field have common languages, and the experimental results can be easily reproduced and directly compared.

In summary, by going through the above list of desirable attributes, we have a much better understanding about what constitutes a good VQA measure for the purposes of testing and optimizing video codecs. As shown in Table I, among existing VQA models, SSIM is perhaps the only one that deserves a score of 'Good" or higher in all attributes. This does not mean that further improvement is not necessary, but SSIM, its derivatives, or direct improvement upon them (e.g., by incorporating temporal assessment approaches) would be currently the top choice in the context of video coding, especially when perceptual optimization is the major target. In fact, SSIM-optimal video coding has achieved notable success and demonstrated great potentials in recent years in both H.264 and HEVC codec designs [20], [21], [22], [29].

III. Computational Issues

The original default computation of SSIM employs an 11×11 sliding Gaussian window [5], which does not match the typical block partitioning structures in video coding standards. Such a mismatch is cumbersome when one wishes to use SSIM to optimize these standard video codecs. A series of questions we would like to ask here are: Can we adjust the windowing approach in the original SSIM to match the block partitioning structures in standard video codecs? How do these adjustments affect the quality prediction performance of SSIM? Should we use the same window size for video sequences with different resolution? How does downsampling operation affect the quality prediction performance of SSIM?

We study these problems empirically by comparing the quality prediction performance of SSIM with different window size and window type (sliding or non-overlapping). Specifically, we compute the SSIM values of HEVC-HM5.0 compressed video and compare them with the MOS values given in [30]. Five Class B sequences (1920×1080, including *Kimono*, *ParkScene*, *Cactus*, *BasketballDrive* and *BQTerrace*) and four Class C sequences (832×480, including *BasketballDrill*, *BQMall*, *PartyScene* and *RaceHorses*) are tested with different Quantization Parameters (Qps) using HEVC Random Access (RA) coding structure. The evaluation results based on PLCC, KRCC and SRCC between objective measure and MOS are summarized in Table IV, where 32×32, 16×16, 8×8 and 4×4 are the sizes of the square windows used in SSIM computation, and "S" and "N" denote sliding and non-overlapping windows, respectively. The evaluation results for PSNR and MS-SSIM are also included for reference.

We have several observations in Table IV. Firstly, among all objective measures, MS-SSIM appears to be the best VQA measure that achieves the best and most consistent performances. However, due to its multi-scale computation, it may not be easily fitted to the codec design. Secondly, when the same window size is used in SSIM computation, the sliding window ("S") and non-overlapping window ("N") approaches have very similar performance across all video sequences. Considering that the computational and implementation complexity of non-overlapping window approach is much lower and it better fits the blocking structures in video coding standards, it would be the preferred method in video coding as both testing and optimizing tools. Thirdly, the best window size depends on the spatial resolution of the video sequence. From Table IV, the best window sizes for Class B (1920×1080) and Class C (832×480) are 32×32N and 8×8N, respectively, which are larger than the size of the smallest coding unit in HEVC. For HD videos such as those in Class B, using small window sizes should be avoided. Specifically, SSIM with 4×4 window size results in the lowest quality prediction performance.

In the scenarios where computational complexity is a critical issue, there may be several approaches to further reduce the complexity. The question is whether these approaches degrade SSIM's quality prediction performance. In the first approach, a large primary window (e.g., 32×32N) is divided into four sub-blocks with quarter sizes (16×16N). The SSIM values are then evaluated in each sub-block and then summed to produce an overall SSIM measure of the primary window. From Table IV, we can conclude that this approach should be used cautiously, because the quality prediction performance may be affected in either positive or negative ways, depending on the size of the primary window as well as the spatial resolution of the video sequence.

The second approach that may be used to reduce the computational complexity is to downsample the video spatially before the SSIM evaluation. This approach was actually employed in the original SSIM paper [5]. To test it, we downsample the video sequences to half, quarter and one-eighth sizes both horizontally and vertically, and test SSIM with different window sizes. The results are given in Table V.

TABLE IV. PLCC, KRCC AND SRCC PERFORMANCE EVALUATION OF PSNR, MS-SSIM AND SSIM WITH DIFFERENT CHOICES OF WINDOW TYPES AND WINDOW SIZES

Evaluation Method	Video Sequence	PSNR	SSIM								MS-SSIM
			32×32S	32×32N	16×16S	16×16N	8×8S	8×8N	4×4S	4×4N	
PLCC	Class B	0.566	0.840	0.843	0.784	0.788	0.675	0.681	0.421	0.423	0.945
	Class C	0.766	0.809	0.805	0.850	0.848	0.888	0.887	0.829	0.827	0.846
KRCC	Class B	0.430	0.653	0.684	0.621	0.621	0.483	0.483	0.302	0.302	0.801
	Class C	0.600	0.617	0.600	0.700	0.700	0.717	0.717	0.667	0.667	0.750
SRCC	Class B	0.602	0.818	0.833	0.767	0.767	0.646	0.646	0.401	0.401	0.932
	Class C	0.782	0.800	0.791	0.868	0.868	0.894	0.894	0.865	0.865	0.900

TABLE V. PLCC, KRCC AND SRCC PERFORMANCE EVALUATION OF SSIM WITH DIFFERENT CHOICES OF DOWNSAMPLING FACTORS

Evaluation Method	Downsampling Factor	Video Sequence	SSIM							
			32×32S	32×32N	16×16S	16×16N	8×8S	8×8N	4×4S	4×4N
PLCC	2	Class B	0.973	0.973	0.966	0.967	0.959	0.959	0.930	0.931
		Class C	0.786	0.783	0.787	0.782	0.822	0.820	0.875	0.874
	4	Class B	0.955	0.951	0.942	0.940	0.936	0.936	0.941	0.941
		Class C	0.800	0.822	0.803	0.804	0.803	0.800	0.833	0.832
	8	Class B	0.885	0.866	0.893	0.884	0.878	0.876	0.877	0.878
		Class C	0.723	0.792	0.770	0.807	0.793	0.802	0.803	0.803
KRCC	2	Class B	0.875	0.875	0.865	0.854	0.865	0.865	0.780	0.801
		Class C	0.600	0.600	0.633	0.633	0.683	0.683	0.750	0.750
	4	Class B	0.833	0.812	0.801	0.801	0.790	0.790	0.780	0.790
		Class C	0.617	0.650	0.650	0.667	0.667	0.667	0.717	0.717
	8	Class B	0.759	0.737	0.790	0.790	0.769	0.780	0.737	0.759
		Class C	0.600	0.633	0.583	0.650	0.600	0.633	0.683	0.700
SRCC	2	Class B	0.969	0.969	0.964	0.962	0.964	0.964	0.921	0.939
		Class C	0.785	0.785	0.809	0.809	0.868	0.868	0.900	0.900
	4	Class B	0.949	0.939	0.936	0.936	0.935	0.935	0.929	0.936
		Class C	0.794	0.809	0.818	0.826	0.838	0.838	0.882	0.882
	8	Class B	0.920	0.900	0.931	0.931	0.927	0.928	0.903	0.920
		Class C	0.776	0.788	0.765	0.809	0.785	0.812	0.865	0.879

It can be observed that this approach is quite effective at improving the quality prediction performance for HD video (Class B), but may result in negative impact on low-resolution videos (Class C), though the performance degradation is not significant. Considering the significant reduction in computational complexity, this is a valuable option in the practical use of SSIM in HEVC.

In summary, the empirical results presented in this section suggest that the use of SSIM in HEVC (or other video coding standards) can be flexible based on the block partitioning structures in the video codec. In general, non-overlapping windows largely reduce complexity without sacrificing SSIM's quality prediction performance. The best SSIM window size depends on the spatial resolution of the video. Downsampling of the video, if used properly, may lead to better quality prediction performance with reduced computational complexity.

IV. CONCLUSION

In this paper, we focus on the integration of two state-of-the-art topics in video coding – HEVC, which is the newly arrived video coding standard; and SSIM, which has recently become a top candidate to replace the traditional PSNR measure as the perceptual criterion in the evaluation and optimization of video codecs. Our goal is to provide a thorough discussion regarding why SSIM (or its derivatives) is a good quality model for the current and future development of video coding techniques, where VQA models may be employed as both testing and optimization tools. We have also attempted to address the computational issues in the practical applications of SSIM in HEVC or possibly future video coding standards, which have sophisticated block partitioning structures (in prediction, transformation, and coding), and aim for encoding videos with a wide range of spatial resolutions. In the future, the discussion can be extended in many directions. These include how to best model the rate-SSIM performance of a given video codec, and how to compute the coding gain between two video codecs based on their rate-SSIM performance.

REFERENCES

[1] "Advanced video coding for generic audiovisual services", ISO/IEC 14496-10:2005(E) ITU-T Rec. H.264 (E), Mar. 2005.

[2] B. Bross, W.-J. Han, G. J. Sullivan, J.-R. Ohm, and T. Wiegand, "High efficiency video coding (HEVC) text specification draft 8," Doc. JCTVC-J1003, July 2012.

[3] G. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Process. Mag.*, vol. 15, no. 6, pp. 74-90, Nov.1998.

[4] Z. Wang and A. C. Bovik, "A universal image quality index" *IEEE Signal Process. Lett.*, vol. 9, pp. 81-84, Mar. 2002.

[5] Z. Wang, A.C. Bovik, H.R. Sheikh, and E.P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600-612, Apr. 2004.

[6] Z. Wang, L. Lu, and A.C. Bovik, "Video Quality Assessment Based on Structural Distortion Measurement," *Signal Processing: Image Communication. Special issue on Objective Video Quality Metrics*, vol. 19, no. 2, pp. 121-132, Feb. 2004.

[7] M.H. Pinson, and S. Wolf, "A new standardized method for objectively meausuring video quality," *IEEE Trans. Broadcast.*, vol. 50, no. 3, pp. 312-322, Sept. 2004.

[8] K. Seshadrinathan and A. C. Bovik, "Motion tuned spatio-temporal quality assessment of natural videos," *IEEE Trans. Image Process.*, vol. 19, no. 2, pp. 335-350, Feb. 2010.

[9] H.R. Sheikh, M.F. Sabir, and A.C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3449-3451, Nov. 2006.

[10] Z. Wang and A.C. Bovik, "Mean squared error: love it or leave it? A new look at signal fidelity measures," *IEEE Signal Process. Mag.*, vol. 26, no. 1, pp. 98-117, Jan. 2009.

[11] Z.-Y. Mai, C.-L. Yang, L.-M. Po, and S.-L. Xie, "A new rate-distortion optimization using structural information in H.264 I-frame encoder," in *Proc. ACIVS'05*, pp. 435-441, 2005.

[12] Z.-Y. Mai, C.-L. Yang, and S.-L. Xie, "Improved best prediction mode(s) selection methods based on structural similarity in H.264 I-frame encoder," in *Proc. SMC'05*, pp. 2673-2678, Oct. 2005.

[13] Z.-Y. Mai, C.-L. Yang, K.-Z. Kuang, and L.-M. Po, "A novel motion estimation method based on structural similarity for H.264 inter prediction," in *Proc. ICASSP'06*, pp. 913-916, May 2006.

[14] C.-L. Yang, H.-X. Wang, and L.-M. Po, "Improved inter prediction based on structural similarity in H.264," in *Proc. ICSPC'07*, pp. 340-343, Nov. 2007.

[15] C.-L. Yang, R.-K. Leung, L.-M. Po and Z.-Y. Mai, "An SSIM-optimal H.264/AVC inter frame encoder," in *Proc. ICIS'09*, pp. 291-295, Nov. 2009.

[16] B.H.K. Aswathappa and K.R.Rao, "Rate-distortion optimization using structural information in H.264 strictly intra-frame encoder," in *South Eastern Symposium on System Theory (SSST)*, pp. 367-370, Mar. 2010.

[17] Y.-H. Huang, T.-S. Ou, and H.H. Chen, "Perceptual-based coding mode decision," in *Proc. ISCAS'10*, pp. 393-396, May-Jun. 2010.

[18] T.-S. Ou, Y.-H. Huang, and H.H. Chen, "A perceptual-based approach to bit allocation for H.264 encoder," *SPIE Visual Commun. Image Process.*, Jul. 2010.

[19] H.H. Chen, Y.-H. Huang, P.-Y. Su and T.-S. Ou, "Improving video coding quality by perceptual rate-distortion optimization," in *Proc. ICME'10*, pp. 1287-1292, Jul. 2010.

[20] Y.-H. Huang, T.-S. Ou, P.-Y. Su and H.H. Chen, "Perceptual rate-distortion optimization using structural similarity index as quality metric," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 11, pp. 1614-1624, Nov. 2010.

[21] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "SSIM-motivated rate distortion optimization for video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 4, pp. 516-529, Apr. 2012.

[22] S. Wang, A. Rehman, Z. Wang, S. Ma, and W. Gao, "Perceptual video coding based on SSIM-inspired divisive normalization," *IEEE Trans. Image Process.*, vol. 22, no. 4, pp. 1418-1429, Apr. 2013.

[23] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multi-scale structural similarity for image quality assessment," *IEEE Asilomar Conf. Signals, Systems and Computers*, Nov. 2003.

[24] A. Rehman, "SSIM-Inspired Quality Assessment, Compression, and Processing for Visual Communications," Ph.D. Thesis. University of Waterloo: Canada, 2013.

[25] V. Baroncini, J.R. Ohm, G.J. Sullivan, "Report on preliminary subjective testing of HEVC compression capability," JCT-VC of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11, 8th Meeting: San Jose, CA, USA, 1-10 Feb. 2012.

[26] K. Zeng, A. Rehman, J. Wang and Z. Wang, "From H.264 to HEVC: Coding Gain Predicted by Objective Video Quality Assessment Models," VPQM 2013, Scottsdale, AZ, USA, Jan.-Feb. 2013.

[27] M.-J. Chen and A.C. Bovik,"Fast structural similarity index algorithm," *J. Real-Time Image Proc.*, vol. 6, no. 4, pp. 281-287, Dec. 2011.

[28] D.Brunet, E.R. Vrscay, and Z. Wang, "On the Mathematical Properties of the Structural Similarity Index," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1488-1499, Apr. 2012.

[29] A. Rehman and Z. Wang, "SSIM-Inspired Perceptual Video Coding for HEVC," in *IEEE International Conference on Multimedia and Expo (ICME)*, Jul. 2012, pp. 497-502.

[30] TK Tan, A. Fujibayashi, J. Takiue, "Objective and subjective evaluation of HM5.0," Doc. JVTVC-H0116, San Jose, CA, Feb. 2012.