Quality Improvement of Tone Mapped Images by TMQI-II based Optimization for the JPEG XT Standard

Artit Visavakitcharoen*, Yuma Kinoshita*, Hiroyuki Kobayashi† and Hitoshi Kiya*
*Department of Information and Communication Systems
Tokyo Metropolitan University, Hino, Tokyo 191-0065, Japan
†Tokyo Metropolitan College of Industrial Technology, Shinagawa,
Tokyo 140-0011, Japan
Email: visavakitcharoen-artit@ed.tmu.ac.jp, kinoshita-yuma@ed.tmu.ac.jp, hkob@metro-cit.ac.jp, kiya@tmu.ac.jp

***Abstract*—This paper proposes a quality improvement method of tone mapped images including compression distortion for the JPEG XT standard, which is a new compression standard for high dynamic range (HDR) images. HDR images are generally required to be mapped to low dynamic range (LDR) ones due to the limitation of display devices. Furthermore, the HDR ones include some compression distortion to be efficiently stored, in most cases. However, conventional tone mapping operations have not considered the effect of the distortion. We apply an iterative gradient ascend algorithm for improving the structural fidelity, which is based on the improved tone mapped image quality index (TMQI-II), under the use of the JPEG XT standard. Decompressed algorithm and the results are compared with those of original HDR ones. The experiment confirms that the images with better quality than images without optimization are provided by the proposed method and the quality is close to that of ones generated from the original HDR image without compression distortion.**

I. Introduction

High dynamic range (HDR) images are diffusing in many areas such as computer graphics, photography, medical imaging, and more. The dynamic range of pixel values in HDR images is wider than standard low dynamic range (LDR) images. On the other hand, HDR images supported display devices are not popular yet. Thus, a tone mapping operation (TMO) which is needed for compressing dynamic range of an HDR image to generate an LDR image. There are large number of research works about tone mapping which have been done. Many of these are focused on compression techniques [1]–[5], TMOs [6]–[10], and computational and memory cost during a TMO [8]–[11]. In this paper, we also focus on TMOs.

An LDR image, which expressed in integer value, can be generated by a TMO from an HDR image with floating-point format value. A TMO can be separated into two categories: global tone mapping and local tone mapping. A lot of TMOs have been been proposed so far. However, the conventional TMOs have not considered the effect of compression distortion, although images generally include the distortion due to the image compression. In this paper, we propose to apply an iterative gradient ascend algorithm to improve the quality of tone mapped images including some compression distortion under the use of the JPEG XT standard [1]–[5], which is the newest image compression standard for HDR images.

Recently, to evaluate the quality of tone mapping operation, the tone-mapped image quality index (TMQI) [12] was proposed, which is comprised of two measurements: the structural fidelity and the statistical naturalness. The structural fidelity part is calculated by the similarity between the luminances of an HDR image and tone mapped one. Then, the improved TMQI (TMQI-II) [13] was proposed in following years, in which an optimization process is applied to TMQI-II for generating more natural look images. However, the effect of compression distortion is not considered as well as other conventional TMOs. In this paper, an algorithm is proposed to improve the quality of tone mapped images by using a TMQI-II based optimization technique for JPEG XT standard.

The JPEG XT profile A, which is typically used for HDR image compression, and the TMQI-II index will be concisely reviewed in Section II. The proposed approach with considering the compression distortion will be explained in Section III. Then, the experiments and results of proposed method will be discussed in Section IV. Lastly, the summary of this paper is described in Section V.

II. PRELIMINATION

In this section, the JPEG XT standard and TMQI-II quality index are briefly reviewed.

A. JPEG XT Profile A

The encoding algorithm of JPEG XT profile A [1], [4] is shown as the block diagram in Fig. 1. To prepare base layer, a TMO can be applied for creating an LDR image LDR_0 from an HDR one HDR_0. Then, the LDR image is encoded by using the legacy JPEG encoder [14] to generate base layer.

Another layer called residual layer is provided by the residual processing to generate the necessary information for HDR reconstruction. First, the luminance components of the HDR image and the LDR image are determined and represented as

978-1-5090-0629-8/16/$31.00

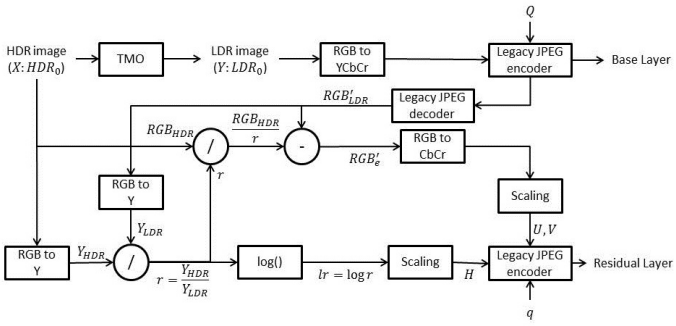

Fig. 1. JPEG XT profile A encoder [4]

Y_{HDR} and Y_{LDR}, respectively. The ratio of the luminances can be calculated as Eq. (1).

$$r = \frac{Y_{HDR}}{Y_{LDR}} \tag{1}$$

To obtain residual layer, the ratio r is passed to log operator for mapping the log-scale into linear-scale. Due to the legacy JPEG encoder, the luminance value for encoding has to limited between 0 and 255 (8-bit per channel). Thus the scaling is applied by following

$$H = \frac{lr - min(lr)}{max(lr) - min(lr)} \cdot 255, \tag{2}$$

where $lr = \log r$.

Besides, the chrominance components for residual encoding are represented as U and V. The HDR image is divided by the ratio r before determining the difference between the HDR image and the LDR_1. The results are converted to two chroma components i.e. Cb and Cr. The encoding chroma U and V can be written as Eq. (3) and (4).

$$U = \frac{Cb - min(Cb)}{max(Cb) - min(Cb)} \cdot 255 \tag{3}$$

$$V = \frac{Cr - min(Cr)}{max(Cr) - min(Cr)} \cdot 255. \tag{4}$$

Lastly, both layers are encoded by the legacy JPEG encoder independently to make the JPEG XT codestream that contains two layers, i.e. base layer and residual layer.

B. Tone Mapped Image Quality Index (TMQI-II)

To evaluate the quality of tone mapping results, the TMQI-II intends to measure both the structural similarity of both a tone mapped image and the original HDR image, and the naturalness of the images [13]. Let the luminances of the HDR image HDR_0 and the tone mapped one LDR_0 represent by X and Y, respectively. According to [13], the TMQI-II is defined by Eq. (5), which comprises two parts i.e. structural fidelity and statistical naturalness.

$$\text{TMQI-II}(X, Y) = a[S(X, Y)]^{\alpha} + (1 - a)[N(Y)]^{\beta}, \tag{5}$$

when a is the adjustable term for priority of both parts. Meanwhile, α and β control the sensitivity of each part.

The structural fidelity part $S(X, Y)$ is computed based on the similarity between luminances in the HDR and the LDR ones as

$$S(X, Y) = \frac{1}{N} \sum_{i=1}^{N} S_{local}(x_i, y_i), \tag{6}$$

where

$$S_{local}(x_i, y_i) = \frac{2\tilde{\sigma}_x \tilde{\sigma}_y + C_1}{\tilde{\sigma}_x^2 + \tilde{\sigma}_y^2 + C_1} \cdot \frac{\sigma_{xy} + C_2}{\sigma_x \sigma_y + C_2} \tag{7}$$

$$\tilde{\sigma} = \frac{1}{\sqrt{2\pi}\theta_\sigma} \int_{-\infty}^{\sigma} \exp\left[-\frac{(t - \tau_\sigma)^2}{2\theta_\sigma^2}\right] dt. \tag{8}$$

The sliding window is applied to both images for gathering two image patches; x_i and y_i, which correspond to HDR and tone mapped images, respectively. Then, Eq. (7) is expressed to measure the local structural fidelity, where σ_x, σ_y and σ_{xy} mean the standard deviations and covariance of two image patches, respectively. C_1 and C_2 are small positive values for preventing instability of S_{local}. This calculation is based on the contrast visibility and similarity of local structure. Besides, the $\tilde{\sigma}$ parameters for both images are provided by the Gaussian distribution function as Eq. (8), when their mean τ_σ and standard deviation θ_σ can be acquired by the contrast sensitivity function that $\theta_\sigma = \tau_\sigma / 3$.

Moreover, the statistical naturalness part $N(Y)$ is the objective assessment based on luminance values that correspond to the subjective evaluation, which reflects to the natural look of tone mapped images. In this paper, the structure fidelity part $S(X, Y)$ is used to improve the quality of tone mapped images.

III. PROPOSED APPROACH

In this paper, the structural fidelity part in TMQI-II is used for improving the quality of tone mapped images. The proposed method aims to improve the structural detail of tone mapped images under the use of the JPEG XT standard.

A. JPEG XT Decoder

To reconstruct original HDR images, the legacy JPEG decoder is applied to restore the base layer and residual layer images from JPEG XT codestream as the first step. The HDR reconstruction process for Profile A is illustrated in Fig. 2.

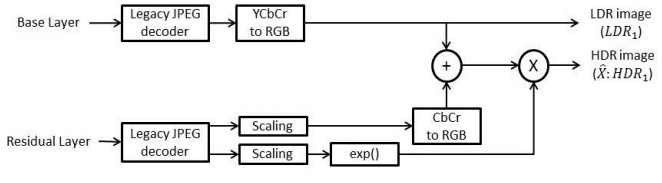

Fig. 2. JPEG XT profile A decoder [1]

We can obtain tone mapped images LDR_1 as LDR ones from base layer. Meanwhile, HDR images HDR_1 are reconstructed by combining base layer and residual layer. Therefore,

the TMO is applied to HDR_1 for generating the LDR_2. Next, we will consider a way to improve the LDR ones.

B. Iteratively Gradient Ascend Update

Since the tone mapped image LDR_2 that is provided from the HDR_2 is represented in RGB format, we need to convert the image into YCbCr color space. Then, we use Y component, which means the luminance of the image, for the optimization. At the same time, we store the chrominance components, i.e. Cb and Cr, for succeeding image reversion.

To improve the structural detail of tone mapped images, the optimization technique called gradient ascend is applied to the structure fidelity by using HDR_1 and LDR_2 tone mapped from HDR_1 shown in Fig. 3. Let the tone mapped image at iteration k-th be $\hat{Y}_k$, where $\hat{Y}$ and $\hat{X}$ are the luminances of LDR_2 and HDR_1, respectively. The iterative update process can be written in Eq. (9).

$$\hat{Y}_{k+1} = \hat{Y}_k + \lambda \nabla_Y S(\hat{X}, \hat{Y})|_{\hat{Y}=\hat{Y}_k}, \tag{9}$$

where the step size is shown by λ and the gradient of the local structure is represented by $\nabla_Y S(\hat{X}, \hat{Y})$, which is computed by substituting the local structural fidelity in Eq. (7). Eq. (7) is rewritten as

$$S_{local}(\hat{x}_i, \hat{y}_i) = \frac{A_1 A_2}{B_1 B_2}, \tag{10}$$

where

$$A_1 = 2\tilde{\sigma}_x \tilde{\sigma}_y + C_1 \tag{11}$$

$$A_2 = \sigma_{xy} + C_2 \tag{12}$$

$$B_1 = \tilde{\sigma}_x^2 + \tilde{\sigma}_y^2 + C_1 \tag{13}$$

$$B_2 = \sigma_x \sigma_y + C_2. \tag{14}$$

By applying the gradient operator to Eq. (10), the expression can be given by

$$\nabla_y S_{local}(\hat{x}_i, \hat{y}_i) = \frac{A_1 A_2' + A_1' A_2}{B_1 B_2} - \frac{A_1 A_2 (B_1 B_2' + B_1' B_2)}{(B_1 B_2)^2}, \tag{15}$$

where

$$A_1' = \nabla_y A_1 = 2\tilde{\sigma}_x \nabla_y \tilde{\sigma}_y \tag{16}$$

$$A_2' = \nabla_y A_2 = \nabla_y \sigma_{xy} \tag{17}$$

$$B_1' = \nabla_y B_1 = 2\tilde{\sigma}_y \nabla_y \tilde{\sigma}_y \tag{18}$$

$$B_2' = \nabla_y B_2 = \sigma_x \nabla_y \sigma_y. \tag{19}$$

Then, both image patches are treating to be column vectors, the statistical parameters can be rewrite as shown in following expressions

$$\mu_y = \frac{1}{N_w} \mathbf{1}^T \hat{\mathbf{y}} \tag{20}$$

$$\sigma_y^2 = \frac{1}{N_w} (\hat{\mathbf{y}} - \mu_y)^T (\hat{\mathbf{y}} - \mu_y) \tag{21}$$

$$\sigma_{xy} = \frac{1}{N_w} (\hat{\mathbf{y}} - \mu_y)^T (\hat{\mathbf{x}} - \mu_x), \tag{22}$$

where $\mathbf{1}$ is the vector of ones. Therefore, the gradient of these parameters is determined as

$$\nabla_y \sigma_y = \frac{1}{N_w \sigma_y} (\hat{\mathbf{y}} - \mu_y) \tag{23}$$

$$\nabla_y \sigma_{xy} = \frac{1}{N_w} (\hat{\mathbf{x}} - \mu_x). \tag{24}$$

From the equations above, the gradient of the local structure can be completed by solving the Eq. (15). Finally, the overall gradient of structure fidelity is computed by placing each local patches back to the corresponding location in image as

$$\nabla_Y S(\hat{X}, \hat{Y}) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{R}_i^T \nabla_{\hat{y}} S_{local}(\hat{x}_i, \hat{y}_i), \tag{25}$$

where i is the number of image patches, while $\mathbf{R}_i^T$ is the operator to place the i-th image patch back to the corresponding location.

Lastly, we revert the YCbCr color components back to RGB image one by using the optimized luminance and the stored chrominances. Now, we can evaluate the image quality with some objective measurements and subjective evaluation further on section.

IV. EXPERIMENTS AND RESULTS

A. Experimental Setup

To testify our proposed approach, we set up the experiment by selecting "Memorial.hdr" image (512×768 pixels in size), which is widely used as an HDR test image. The experiment method is illustrated as shown in Fig. 3.

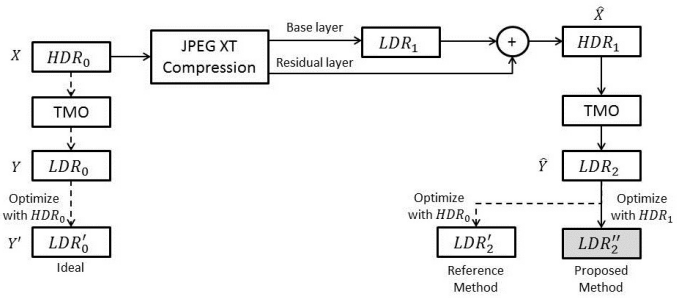

Fig. 3. Experimental approach

The original HDR image HDR_0 is mapped to be LDR_0 by using the Reinhard's global TMO [7], as a typical TMO. Then, the iterative optimization with HDR_0 shown in Eq. (9) is applied to LDR_0 to generate the optimized image one LDR_0'. This image LDR_0' is the ideal one since there is no compression distortion. Therefore, we set LDR_0' as the reference image to compare with other images including some distortion.

Next, we compress the HDR_0 with JPEG XT profile A [4], and LDR_1 and HDR_1 are reconstructed by the standard JPEG XT decoder respectively. In addition, the tone mapped version of HDR_1, defined as LDR_2, is also generated by using the same TMO. To improve the quality of LDR_2, we apply

the iterative optimization with HDR_1 for preparing LDR_2'', respectively. LDR_2 is also applied to the iterative optimization with HDR_0 to generate a reference LDR image LDR_2'. Note that HDR_0 that does not have any compression distortion can not be generally used in most actual cases. Therefore, we consider to use LDR_1 for the optimization instead of LDR_0.

The necessary parameters in the experiment are set as described below. The delta (for avoiding singularity) and key value (for luminance scaling) parameters of Reinhard's global TMO are set to 0.01 and 0.36, respectively. We assign the window size to be 11×11 to measure the local structural fidelity. Then, the step size of iteratively update process (λ) is fixed to be 0.5, and the optimization process is repeated for 100 iterations.

B. Experimental Results

The effectiveness of proposed method is evaluated by using three objective quality assessment indices; peak signal to noise ratio (PSNR), structure similarity (SSIM) [15], and feature similarity (FSIM) [16]. The aim of this experiment is to demonstrate that LDR_2'' generated by the proposed method has the almost same quality as that of LDR_2' optimized with the original image HDR_0. The measurement of the image quality can be defined by four subscriptions on each index as shown in Table I.

TABLE I
DEFINITION OF SUBSCRIPTIONS FOR EVALUATION INDICES

Subscription	Test image	Reference image
1 (Proposed)	LDR_2''	LDR_0'
2 (Reference)	LDR_2'	LDR_0'
3 (Without optimization)	LDR_2	LDR_0'
4	LDR_1	LDR_0'

Fig. 4 shows the comparison of PSNR values by varying the compression ratios of both base layer and residual layer, where two compression ratios for two JPEG encoders in the JPEG XT encoder are fixed at common values. The "Reference" method is used to evaluate the image generated by the optimization using HDR_0 as shown in Fig. 3, although the receiver could not get the original HDR image HDR_0 in general case. Therefore, we compare the proposed method with "Reference" instead. The "Without optimization" case is provided by the quality measurement of tone mapped images before applying the optimization to show the quality improvement through the optimization process. It is comfirmed from the Fig. 4 that the proposed method can provide LDR images with the almost same quality as that of LDR_2', nevertheless HDR ones used for the proposed method include some compression distortion.

Similarly, the measurements of SSIM and FSIM values are shown in Fig. 5 and 6, respectively. It is shown that the optimization can improve the structural detail of LDR_2 to be close to the ideal case LDR_0'.

To more clearly demonstrate the effectiveness, the examples of tone mapped LDR images are shown in Fig. 7. In Fig. 7, LDR_1 and LDR_2 have some block artifacts that are occurred

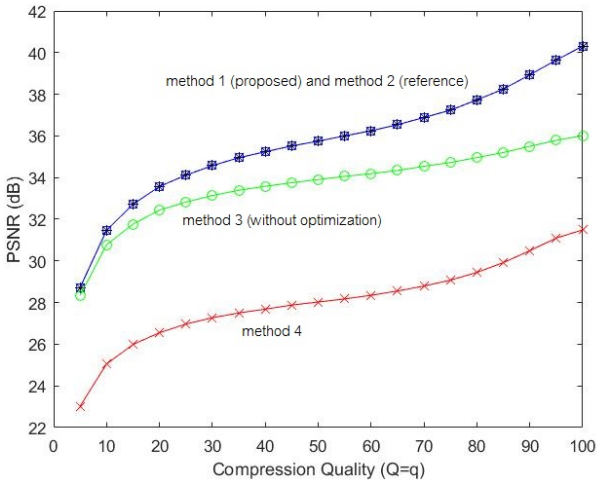

Fig. 4. PSNR comparison

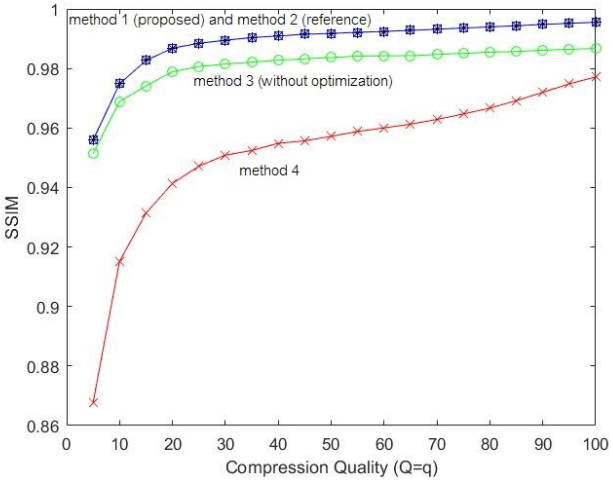

Fig. 5. SSIM comparison

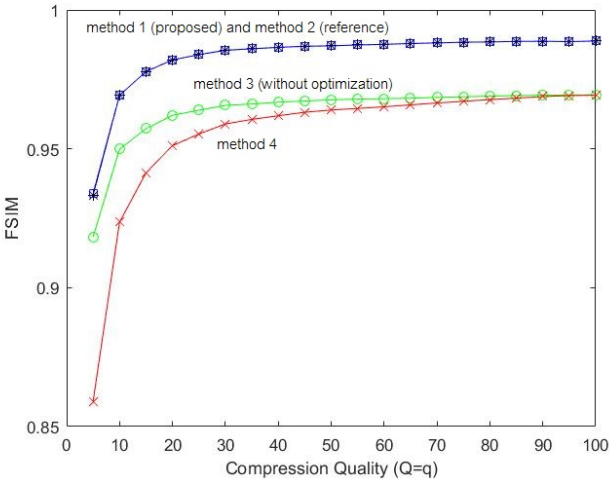

Fig. 6. FSIM comparison

(a) LDR'_0 (b) LDR_1 (25.05 dB)

(c) LDR_2 (30.77 dB) (d) LDR''_2 (31.46 dB)

Fig. 7. Examples of LDR images for $Q = q = 10$, (PSNR value).

through the legacy JPEG encoder, but the block artifacts in LDR''_2 faded away through the optimization because the effect of contrast sensitivity in the structural fidelity part.

In summary, the optimization allows to improve the local structural fidelity of tone mapped images, so that the block artifacts can be reduced.

V. Conclusion

This paper has proposed a method to improve the quality of tone mapped images for JPEG XT standard based on TMQI-II, in which a TMQI-II based optimization technique is used. To obtain the better quality of the tone mapped image, the structural fidelity is optimized with an iterative gradient ascend algorithm. Furthermore, the reconstructed HDR image, which includes the compression distortion, is used in the optimization instead of the original HDR image in the proposed method. The experimental results have demonstrated that the proposed method can provide images with the almost same quality as that of tone mapped images generated by using the original HDR image. It was also shown that block artifacts due to the use of the JPEG compression, are reduced by proposed method.

Acknowledgment

The financial support provided by the Tokyo Metropolitan Government (TMG) through the Tokyo Human Resources Fund for City Diplomacy Scholarship for Artit Visavakitcharoen is gratefully acknowledged.

References

[1] *Information technology: Scalable compression and coding of continuous-tone still images extensions for high-dynamic range images*, International Standard ISO/IEC 18477 Std., 2015.

[2] S. Choi, O. Kwon, D. Jang, and S. Choi, "Performance evaluation of JPEG XT standard for high dynamic range image coding," in *Proc. of the World Congress on Engineering 2015 (WCE 2015)*, vol. I, London, U.K., July 2015, pp. 552–556.

[3] A. Artusi, R. K. Mantiuk, T. Richter, P. Korshunov, P. Hanhart, T. Ebrahimi, and M. Agostinelli, "JPEG XT: A compression standard for hdr and wcg images," *IEEE Signal Processing Mag.*, vol. 33, pp. 118–124, Mar 2016.

[4] O. Watanabe, R. Suzuki, and H. Kiya, "A structure of JPEG XT encoder considering effect of quantization error," in *IEEE International Symposium on Circuits and Systems*, Motreal, Canada, May 2016, pp. 810–813.

[5] M. Iwahashi and H. Kiya, "Two layer lossless coding of hdr images," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 1340–1344.

[6] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," *Comput. Graph. Forum*, vol. 22, no. 3, pp. 419–426, 2003.

[7] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tone reproduction for digital images," *ACM Trans. Graph.*, vol. 21, no. 3, pp. 267–276, July 2002.

[8] T. Dobashi, T. Murofushi, M. Iwahashi, and H. Kiya, "A fixed-point global tone mapping operation for hdr images in the rgbe format," in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2013 Asia-Pacific*, 2013, pp. 1–4.

[9] T. Dobashi, A. Tashiro, M. Iwahashi, and H. Kiya, "A fixed-point implementation of tone mapping operation for hdr images expressed in floating-point format," *APSIPA Trans. Signal and Information Processing*, vol. 3, no. e11, pp. 1–11, 2014.

[10] T. Dobashi, T. Murofushi, M. Iwahashi, and H. Kiya, "A fixed-point global tone mapping operation for hdr images in the rgbe format," *IEICE Trans. on Fundamentals of Electronics, Communications and Computer Sciences*, vol. 97, no. 11, pp. 2147–2153, 2014.

[11] T. Murofushi, M. Iwahashi, and H. Kiya, "An integer tone mapping operation for hdr images expressed in floating point data," in *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP)*, May 2013, pp. 2479–2483.

[12] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 657–667, Feb 2013.

[13] K. Ma, H. Yeganeh, K. Zeng, and Z. Wang, "High dynamic range image compression by optimizing tone mapped image quality index," *IEEE Trans. Image Process.*, vol. 24, no. 10, pp. 3086–3097, Oct 2015.

[14] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. Consumer Electronics*, vol. 38, no. 1, pp. xviii–xxxiv, Feb 1992.

[15] 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.

[16] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Trans. Image Process.*, vol. 20, no. 8, pp. 2378–2386, Aug 2011.