IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received May 24, 2021, accepted June 2, 2021, date of publication June 7, 2021, date of current version June 15, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3087208

Correcting the Defects of the Default Tone Mapping Operators Implemented in the JPEGXT HDR Image Compression Standard

JIANYU LIN, (Senior Member, IEEE), AND WEI ZHANG
Department of Electronic Engineering, Shantou University, Shantou 515063, China
Guangdong Provincial Key Laboratory of Digital Signal and Image Processing, Shantou University, Shantou 515063, China
Key Laboratory of Intelligent Manufacturing Technology (Shantou University), Ministry of Education, Shantou 515063, China

Corresponding author: Jianyu Lin (jianyulin@stu.edu.cn)

This work was supported by Shantou University.

ABSTRACT This short article reports two imperfections or implementation bugs of the default tone mapping operators (TMOs) of the JPEGXT HDR image compression standard. In JPEGXT, profiles A and C use the same default TMO, while profile B uses a different default TMO. The first imperfection is with the default TMO of profile A and profile C. Specifically, the default TMO leads to a reduced dynamic range of the tone-mapped LDR images. Very often, the reduction can be severe. The mechanism of causing the dynamic range reduction stems from an implementation bug. To resolve the issue, the correction of adding the missed final normalization step to linearly stretch the dynamic range needs to be made. The second imperfection is the highlights clipping in the LDR images from the default TMO of profile B. The clipping could lead to unusable LDR images. Our suggestion is to unify the default TMO within JPEGXT, i.e., to use the corrected TMO of profiles A and C for profile B.

INDEX TERMS JPEGXT HDR image compression standard, default LDR tone mapping for the legacy JPEG compatibility.

I. INTRODUCTION

Digital images that are displayed on standard display devices are normally of the dynamic range of 8 bits per pixel (bpp) per color. With the advances of display devices and imaging technology, digital images of dynamic ranges higher than 8 bpp become very common. They are called high dynamic range (HDR) images.

HDR images are recorded using various formats. They are initially acquired by linearly recording the light intensities on each pixel of the image sensor. To increase accuracy, some HDR image formats use floating numbers. For fundamentals of HDR images, please refer to [1], [2]. In displaying HDR images, even if today's premium display devices have very high dynamic ranges, the dynamic ranges of the display devices are still normally much lower than the recorded dynamic ranges of the HDR images. Since the Human Visual System (HVS) does not perceive light intensities linearly, a simple linear scaling of the dynamic range does not provide pleasant views of the recorded HDR images.

For example, Fig. 1 (a) shows the popular HDR image "memorial" (downloaded from [3]) with the dynamic range linearly scaled to the dynamic range of the display device. Note, the response of a display device is not linear. Thus, to display light intensities linearly, gamma correction needs to be performed. In Fig. 1(a), we assumed $\gamma = 2.4$ for the display device. As can be seen, Fig. 1(a) completely differs from what is perceived by the HVS. Thus, to display HDR contents on a low (or standard) dynamic range (LDR) display device, suitable functions that mathematically map the dynamic range from HDR to LDR, which is referred to as the tone mapping operators (TMOs), are needed to make a conversion. Performing a simple tone-mapping provided in [4], Fig. 1(b) shows an 8-bit LDR image much closer to what is perceived by the HVS.

As mentioned, the display device with the highest dynamic range may still have a much lower dynamic range than that of the HDR image. Thus, a TMO is always needed. However, the TMO for LDR display devices and the TMO for HDR

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott.

(a) (b)

FIGURE 1. Two display methods for the HDR image "memorial": (a) The displayed light intensity of each pixel is proportional to the linearly recorded value ($\gamma = 2.4$ is assumed for the display device); (b) The TMO of [4] is applied.

display devices are different. Thus, in addition to linearly recorded HDR images, there is another category of HDR images, which uses bit-depth higher than 8 bpp and is already tone-mapped to be readily displayed on HDR display devices. The 10-bit and 12-bit videos belong to this category. To distinguish them from the linearly recorded HDR images, for such tone-mapped discrete value HDR images, we adopt the name from JPEGXT standard: the "intermediate dynamic range" (IDR) images. Fig. 4(a) shows such an IDR image. Using Adobe Photoshop, this 16-bit IDR image was generated from the camera raw image data downloaded from [5]. Because the image is for HDR display devices, it looks too dark on LDR display devices. Performing another TMO, Fig. 4(b) looks better on LDR display devices.

It can be seen that for every HDR image, HDR display devices need an HDR digital version and LDR display devices need an LDR digital version. Keeping two versions increases the complexity for HDR image compression.

JPEGXT [6]–[8] is a new standard specifically developed for HDR image compression. It can compress the HDR version and the LDR version simultaneously. It accommodates both linearly recorded floating-number HDR images and discrete value IDR images. To achieve simultaneous compression on the HDR version and the LDR version, JPEGXT uses a 2-layer encoding structure. The LDR version is coded using the legacy JPEG, which is called the base layer encoding. Then, the difference between the HDR version and the decoded LDR image is coded further, which is referred to as the enhancement layer encoding. The main advantage of the 2-layer encoding structure is that JPEGXT provides compatibility to the legacy JPEG. At the same time, it can easily control the qualities of the compressed HDR and LDR versions, because the qualities are independently managed by the 2 separated layers. Currently, there are still ongoing researches on HDR image compression, for example [9]–[11], because HDR image compression and JPEGXT are relatively new research topics in image compression.

There are 3 profiles in JPEGXT, profile A, B, and C. All profiles accept floating number linear HDR images. Profile C can also compress IDR images of a bit-depth between 8 bits and 16 bits.

JPEGXT provides flexibility on the input. The input can be the HDR image only, or the HDR image together with an LDR version. This way, it provides the flexibility for users to use their own LDR versions. If the LDR version is not provided by the user, JPEGXT will use the default

(a) (b)

FIGURE 2. Tone-mapped LDR versions of the HDR image "dani_synagogue": (a) the LDR image from the default TMO of JPEGXT profiles A and C; (b) the dynamic range of the image in (a) is linearly stretched to the full 0 – 255 range.

tone mapping to generate an LDR version. Thus, using the legacy JPEG, an LDR version can always be decoded from the compressed output of JPEGXT. In other words, with the legacy JPEG, one can always display the JPEGXT output on an LDR display device, which is an important and convenient compatibility feature.

However, there are some imperfections in the default TMOs implemented in JPEGXT. These imperfections may lead to poor or even unusable output LDR images. Corrections are needed for these implementation oversights.

II. IMPERFECTIONS OF THE DEFAULT TMOS IMPLEMENTED IN JPEGXT AND THE CORRECTIONS

A. REDUCED DYNAMIC RANGE OF THE LDR IMAGES

In JPEGXT, profiles A and C use the same default TMO, while profile B uses a different default TMO. For the default TMO of Profiles A and C, the dynamic range of the tone-mapped LDR version is not fully used. First, for floating-number linear HDR input images, profiles A and C use the TMO of [4] as the default TMO. In the source code of the JPEGXT reference software [8], the exact TMO implementation is as follows. Find the luminance image Y first:

$$Y = 0.2126 I_r + 0.7152 I_g + 0.0722 I_b, \tag{1}$$

where I_r, I_g, and I_b are respectively the red, green, and blue component of the floating-number image. Then, parameters k and m are found from the logarithmic image $y = \log Y$:

$$k = \frac{\max(y) - \operatorname{mean}(y)}{\max(y) - \min(y)}, \quad (k \geq 0) \tag{2}$$

and

$$m = 0.3 + 0.7 k^{1.4}. \tag{3}$$

Finally, for all the 3 color channels I_r, I_g, and I_b, the HDR pixel value H is tone mapped to the LDR pixel value L by:

$$L = \operatorname{round}\left(255 \times \frac{H^{\frac{1}{\gamma}}}{A^m + H^{\frac{1}{\gamma}}}\right), \tag{4}$$

where $\gamma = 2.2$, $A = \operatorname{mean}(y)$. It can be easily seen that factor

$$f = \frac{H^{\frac{1}{\gamma}}}{A^m + H^{\frac{1}{\gamma}}} \tag{5}$$

in (4) is in the range $0 \leq f < 1$, provided the HDR pixel value $H \geq 0$ and $A > 0$. Since the derivative

$$\frac{df}{dH} = \frac{1}{\gamma} H^{\frac{\gamma-1}{\gamma}} \left(1 - \frac{H^{\frac{1}{\gamma}}}{A^m + H^{\frac{1}{\gamma}}}\right) = \frac{1}{\gamma} H^{\frac{\gamma-1}{\gamma}} (1 - f) > 0 \tag{6}$$

for $H > 0$, L is monotonically increasing on H.

As a result, there are two properties of the H to L mapping expressed by (4). Firstly, if A^m is not small compared with $\max(H^{1/\gamma})$, then, $\max(f)$ becomes much less than 1, or equivalently, $\max(L)$ is much less than 255. Secondly, if $\min(H) > 0$, then, $\min(f) > 0$. This could lead to a $\min(L)$ much greater than 0. In both cases, the dynamic range of L is dramatically reduced.

For example, when the HDR image "dani_synagogue" (downloaded from [3]) is tone mapped to the LDR image by the JPEGXT default TMO, the maximum L value is 167, and the minimum L value is 24, leading to a dynamic range of $167 - 24 = 143$, which is only about 56% of the 8-bit 0 – 255 dynamic range, i.e., about half the dynamic range is unused! The LDR image is shown in Figure 2(a).

Besides the maximum value cannot reach 255, the minimum value could also be much greater than 0. When the HDR

(a) (b)

FIGURE 3. Tone-mapped LDR versions of the HDR image "rend03": (a) the LDR image from the default TMO of JPEGXT profiles A and C; (b) the dynamic range of the image in (a) is linearly stretched to the full 0 – 255 range.

image "rend03" (downloaded from [3]) is tone mapped to the LDR image by the JPEGXT default TMO, the maximum L value is 245, and the minimum L value is 79. In addition to about a 35% reduction in the dynamic range, the darkest value in the image is 79. The image looks unpleasantly bright with low contrast as shown in Fig. 3(a).

Next, for the discrete-value IDR image compression in profile C, the implemented TMO is very similar to the HDR TMO of (4). The only difference is, there is no gamma correction, i.e.,

$$L = \text{round}\left(255 \times \frac{H}{A^m + H}\right). \quad (7)$$

Parameters A and m are found using the same method as described earlier for the HDR TMO of (4), except that the IDR image values H in (7) are normalized to the range of 0 to 1 before performing the TMO. Thus, there is the same reduced dynamic range issue as the HDR TMO (4). Fig. 4(c) shows the tone-mapped LDR version of the IDR image in Fig. 4(a) using the default TMO of JPEGXT profile C. The maximum L value of the LDR image shown in Fig 4(c) is only 148, and the minimum L value is 0, which leads to a dynamic range reduction of about 42%.

The above-described dynamic-range reduction problem of the default TMO in profiles A and C happens very often. Users do not have to search many images to find a case with a significant dynamic-range reduction similar to the above examples. As already mentioned, JPEGXT has the convenient compatibility feature that any JPEGXT compressed HDR image can be viewed on an LDR display device with the legacy JPEG. With this imperfection, users must provide their own LDR versions in order to avoid unpleasant dynamic range reduced LDR versions, and the above JPEGXT compatibility feature is degraded greatly. However, the problem can be easily resolved by stretching the dynamic range linearly to the 0 – 255 8-bit LDR dynamic range, i.e.,

$$L = \text{round}\left(255\frac{f - \min(f)}{\max(f) - \min(f)}\right), \quad (8)$$

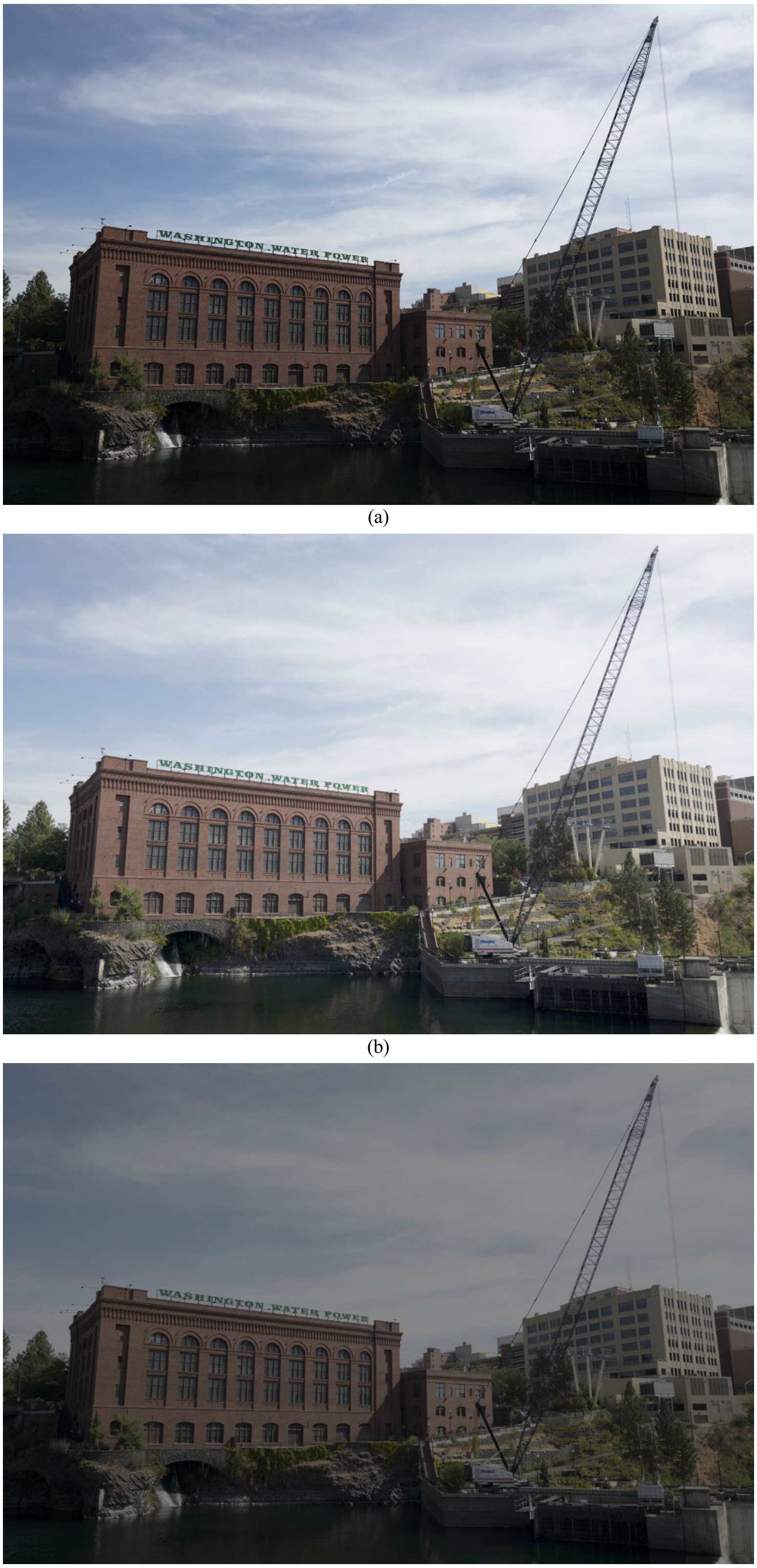
where f is given by (5), with $\gamma = 1$ for the discrete IDR TMO in profile C. With this minor correction, the LDR 8-bit dynamic range is fully exploited. Fig. 2(b), 3(b), and 4(b) show the dynamic range linearly stretched versions using (8).

Note, in [4], when the TMOs are plotted, the mapping is from interval [0, 1] to interval [0, 1] on the real axis. This suggests a normalization equivalent to equation (8) above. In other words, the imperfection of the default TMO of profiles A and C is just an oversight during the implementation of [4]. However, the consequence of the implementation bug is not trivial.

B. CLIPPING ON THE HIGHLIGHTS

The default TMO of profile B seems to go to the other extreme. It does provide higher contrast in many cases. However, it can easily clip the highlights in the LDR images. For example, Fig. 5(b) shows the LDR version of the HDR image "SpheronNapaValley" (downloaded from [3]). Fig. 5(a) shows the LDR version from the default TMO of profile B. Because of the highlights clipping, the Fig. 5(a) image is simply unusable. Thus, we suggest that JPEGXT use the same default TMO for all 3 profiles. This TMO can be the corrected default TMO for profiles A and C shown in Part A above. In fact, the Fig. 5(b) LDR version is obtained from the corrected default TMO for profiles A and C.

For readers' convenience, test images used in this paper, as well as the compiled JPEGXT, are available at [12]. As mentioned, the imperfections appear very often, readers can easily find some other examples that lead to the imperfect LDR images without a long search on other images.



(a)

(b)

(c)

FIGURE 4. Tone-mapped LDR versions of an IDR image: (a) the original IDR image; (b) the LDR image obtained from the default TMO of JPEGXT profile C corrected using linear normalization; (c) the LDR image obtained from the default TMO of JPEGXT profile C directly.

(a)

(b)

FIGURE 5. **Tone-mapped LDR versions of the HDR image "SpheronNapaValley": (a) the LDR image from the default TMO of JPEGXT profile B; (b) the correct LDR version of the HDR image "SpheronNapaValley" (generated using the default TMO of profile A corrected with linearly stretched dynamic range).**

III. CONCLUSION

There are imperfections in the default TMOs implemented in the JPEGXT HDR image compression standard, which often lead to substantially unused dynamic range for the tone-mapped LDR images (profile A and C), or lead to unusable highlights-clipped LDR images (profile B). The reduced-dynamic-range imperfection in profiles A and C stems from an implementation oversight and thus can be easily corrected by linearly normalizing the dynamic range to the full 8-bit range. For the other imperfection, we suggest unifying the default TMO within the standard. The corrected TMO eliminates the defects and produces much more pleasant LDR images as shown. Thus, the minor corrections secure the important JPEGXT compatibility feature, which is, the legacy JPEG can always display the JPEGXT compressed results no matter whether the LDRs are provided in the compression inputs.

REFERENCES

[1] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, W. Heidrich, and K. Myszkowski, *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*, 2nd ed. San Mateo, CA, USA: Morgan Kaufmann, 2010.

[2] F. Banterle, A. Artusi, K. Debattista, A. Chalmers, *Advanced High Dynamic Range Imaging*, 2nd ed. Boca Raton, FL, USA: CRC Press, 2018.

[3] Accessed: May 2021. [Online]. Available: http://www.anyhere.com/gward/hdrenc/pages/originals.html

[4] E. Reinhard and K. Devlin, ''Dynamic range reduction inspired by photoreceptor physiology,'' *IEEE Trans. Vis. Comput. Graphics*, vol. 11, no. 1, pp. 13–24, Jan. 2005.

[5] Accessed: May 2021. [Online]. Available: https://www.dpreview.com/sample-galleries/5682971467/nikon-d850-sample-gallery

[6] 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 Process. Mag.*, vol. 33, no. 2, pp. 118–124, Mar. 2016.

[7] A. Artusi, R. K. Mantiuk, T. Richter, P. Hanhart, P. Korshunov, M. Agostinelli, A. Ten, and T. Ebrahimi, ''Overview and evaluation of the JPEG XT HDR image compression standard,'' *J. Real-Time Image Process.*, vol. 16, pp. 413–428, Apr. 2019.

[8] *JPEGXT Source Codes*. Accessed: Mar. 2021. [Online]. Available: https://jpeg.org/jpegxt/software.html

[9] O.-J. Kwon, S. Choi, and D. Shin, ''Improvement of JPEG XT floating-point HDR image coding using region adaptive prediction,'' *IEEE Access*, vol. 6, pp. 3321–3335, 2018.

[10] D. Gommelet, A. Roumy, C. Guillemot, M. Ropert, and J. L. Tanou, ''Gradient-based tone mapping for rate-distortion optimized backward-compatible high dynamic range compression,'' *IEEE Trans. Image Process.*, vol. 26, no. 12, pp. 5936–5949, Dec. 2017.

[11] D. Gommelet, J. L. Tanou, A. Roumy, M. Ropert, and C. Guillemot, ''Optical-flow based nonlinear weighted prediction for SDR and backward compatible HDR video coding,'' *IEEE Trans. Image Process.*, vol. 29, pp. 2123–2138, 2020.

[12] Accessed: May 2021. [Online]. Available: https://stumail-my.sharepoint.cn/:u:/g/personal/jianyulin_stu_edu_cn/ER8nXEH9loBGnqU8AJ9xaAwBHiLf5-r8lzTRS00EdMSSFA?e=ZJSa51