Puma: A High-Quality Retinex-Based Tone Mapping Operator

Nikola Banić and Sven Lončarić
Image Processing Group
Faculty of Electrical Engineering and Computing
University of Zagreb, 10000 Zagreb, Croatia
E-mail: {nikola.banic, sven.loncaric}@fer.hr

Abstract—Tone mapping is the process of compressing high dynamic range (HDR) images to obtain low dynamic range (LDR) images in order to display them on standard display devices. The methods that perform tone mapping also known as tone mapping operators (TMOs) can be global and process all luminances in the same way, or they can be local and process the luminances with respect to their closer neighborhood. While the former tend to be faster, the latter are known to produce results of significantly higher quality. In this paper perceptually-based tone mapping is combined with one of the latest Retinex-based methods to create a high-quality TMO. The new TMO requires only a constant number of steps per pixel and experimental results show that it outperforms all but one state-of-the-art TMOs in terms of tone mapped LDR image quality. The source code is available at http://www.fer.unizg.hr/ipp/resources/color_constancy/.

Index Terms—HDR, image enhancement, LDR, Naka-Rushton equation, Retinex, tone mapping operator.

I. INTRODUCTION

Most contemporary image display devices support only low dynamic range (LDR) images, i.e. images with a low ratio between their largest and smallest non-zero intensity. Since high dynamic range (HDR) images are used ever more frequently [1], there is a need to first convert them to LDR images before displaying them on LDR display devices. This conversion is done by compressing the image luminance channel in a process known as tone mapping. The methods that perform tone mapping are called tone mapping operators (TMOs). A common choice for the luminance channel is the Y channel of the YUV colorspace, whose value for a pixel described in the RGB colorspace \( p = (R, G, B)^T \) is defined as [2]:

\[
Y = 0.299R + 0.587G + 0.114B. \tag{1}
\]

If after performing tone mapping the luminance value \( L \) of pixel \( p \) is changed to \( L' \), then \( p \) is by rescaling changed to

\[
p' = \frac{L'}{L}p = \left( \frac{L'}{L}R, \frac{L'}{L}G, \frac{L'}{L}B \right)^T. \tag{2}
\]

If a value of \( L \) for a pixel \( p \) is always mapped to the same \( L' \) regardless of \( p \)’s position in the image, then tone mapping is performed by a global TMO, otherwise it is performed by a local TMO. In the literature there are numerous examples of both global [3]–[10] and local TMOs [11]–[17]. While global TMOs tend to be faster, local TMOs are known to produce results of significantly higher quality as pointed out in several studies [18] [19]. In this paper a local TMO composed of two main steps is proposed. It first applies a simple and fast perceptual-based tone mapping that gives crude resulting images. Despite being of low quality, they also have a low dynamic range, which makes it possible to apply regular brightness adjustment methods without repercussions that occur when they are applied directly to high dynamic range images. It is shown that by applying a recently developed Retinex-based method the quality of these crude images is significantly improved up to the level obtained by state-of-the-art TMOs and higher. Since the used Retinex-based method performs only a constant number of steps per pixel, the whole combination results in a scalable high-quality local TMO.

The paper is structured as follows: in Section II a crude, but fast technique to reduce the initial dynamic range of a HDR image is described and the reasons for its introduction are explained, in Section III it is shown how the results of this technique can be improved with a Retinex-based brightness adjustment method, in Section IV experimental results are presented and discussed, and Section V concludes the paper.

II. CRUDE DYNAMIC RANGE COMPRESSION

A. Brightness adjustment and tone mapping

Tone mapping operators are very similar to brightness adjustment methods since both of them process the luminance channel in order to make it more suitable for display purposes. The main difference is that tone mapping operators are primarily designed to perform compression and they can be applied to images of various dynamic ranges. On the other hand brightness adjustment methods mostly work only on low dynamic range images and since they tend to increase the entropy of the image, they can actually increase the image size after lossless compression depending on the image content. At the end, however, both tone mapping and brightness adjustment are supposed to produce images that are visually pleasant. The problem is that in most cases applying a brightness adjustment method directly to high dynamic range images fails.

The reason for this failure are usually wide gaps between intensities of bordering image regions, which in case of local brightness adjustment methods tend to become the location of undesirable halo effects [17]. While there are methods that try to detect such regions in advance [16], the techniques used to do that may significantly increase the overall TMO complexity.
A simple conclusion from previous experience is that for a brightness adjustment method to become a practical TMO, usually some kind of specific modifications are required.

B. The proposed preprocessing

Instead of directly upgrading a brightness adjustment method to make it able to carry out tone mapping, an alternative approach is to prepare a HDR image through an appropriate preprocessing to make it suitable for an unaltered brightness adjustment method to be applied. One such preprocessing may consist in applying a TMO in order to obtain a LDR image. This would bridge any potentially high intensity gaps between bordering image regions and remove the main obstacle to successful application of brightness adjustment methods. In this way after performing the initial tone mapping, any theoretically well defined brightness adjustment method could be used to do exactly what it was designed for. Some similar approaches have already been proposed, e.g. in [20], but the main problem is that such and similar solutions use a rather complex preprocessing method followed by an even more complex additional brightness adjustment method.

One way to find a simple enough preprocessing TMO could be to mimic the behaviour of human visual system (HVS). The light intensity response curve for rods and cones in retina can be fitted with the Michaelis-Menten equation [21]

$$r(I) = \frac{I^n}{I^n + I_s^n}$$

(3)

where $I$ is the light intensity, $I_s$ is the level at which the response is half maximum, $n$ is a sensitivity control exponent, and $r(I)$ is the response assumed to be correlated with the perceived brightness. Eq. (3) has a history of successful applications of its modified forms [6] [8] [20] despite the fact that there are more accurate models of HVS’s perception. For $n = 1$ Eq. (3) is simplified to the Naka-Rushton equation [22]. In [13] its extended and more practical form is given as

$$L' = \frac{L}{L_w} + \alpha = \frac{L}{L + \alpha L_w}$$

(4)

where $L$ is the initial luminance, $L_w$ is the image key, and $\alpha$ is a scaling parameter. Dividing the initial luminance by the image key is a widely used tone mapping step. The image key is approximated by geometric luminance mean [23] [5] [13]

$$L_w = \exp\left(\frac{1}{N} \sum_i \ln (L(i) + \epsilon)\right)$$

(5)

where $N$ is the number of pixels, $L(i)$ is the $i$-th pixel luminance, and $\epsilon$ is a small value to avoid logarithm of zero. The results of S-shaped functions similar to the one described by Eq. (4) are below the ones of state-of-the-art TMOs in terms of resulting LDR image quality [24] and usually some more sophisticated forms are used [13]. However, due to its simplicity, it is a good candidate for the mentioned initial tone mapping before handing the preprocessed image over to a brightness adjustment method for further processing.

In [25] it was shown how the quality of results of global TMOs can be significantly improved by applying them to the $V$ channel of the $HSV$ colorspace instead of applying them to the $Y$ channel of the $YUV$ colorspace. For this reason in the rest of the paper Eq. (4) will be applied to the $V$ channel and an example of such an application is given in Fig 1.

![Fig. 1: The same scene tone mapped by applying Eq. (4) to (a) the grayscale and (b) $V$ image channel.](image)

The parameter $\alpha$ in Eq. (4) controls how the luminance values are handled. Lower values of $\alpha$ give brighter images and vice versa as shown in Fig. 2, which seemingly makes them more useful. However, higher values introduce more contrast, which can be successfully used to obtain higher quality during brightness adjustment as will be demonstrated in the following sections. Since in a free interpretation the parameter $\alpha$ can be seen as a power regulator of an imaginary photographic flash, the global TMO obtained by applying Eq. (4) to the $V$ channel is named Flash for easier notation in the rest of the paper.

III. Retinex-based tone mapping

Now that Flash, a fast and simple preprocessing tone mapping has been proposed, the next step is to select a brightness adjustment method to improve the obtained LDR images. Examples of brightness adjustment methods include modifications of histogram equalization [26]–[28], adaptive histogram equalization [29] [30], Retinex-based methods [31]–[36], naturalness preserving methods [37] [38], unsharp masking [39] [40], automatic color equalization [41] [42]. One that was recently shown to be superior to most of them in terms of low complexity and resulting image quality is the Smart Light Random Sprays Retinex (SLRMSR) [43]. Therefore, SLRMSR is a suitable brightness adjustment method needed after the initial and crude tone mapping performed by Flash.

A. Smart Light Random Memory Sprays Retinex

SLRMSR relies on the Retinex theory and it processes each pixel individually. Its main processing part inherited from its predecessors can be summarised by using the equation

$$L^*(i) = \frac{L(i)}{L(x_{Hi})}$$

(6)

where $x_{Hi}$ is the index of the pixel with the highest luminance in a spray of $n$ pixels randomly chosen around the $i$-th pixel. The result obtained after performing a specific process of denoising described in more detail in [43] can be described as

$$L'(i) = Denoise(L^*(i)).$$

(7)
One of the final steps of SLRMSR is to prevent a possible over-enhancement of already bright pixels by introducing an adjustable intensity remapping mechanism, which is given as

$$L''(i) = \lambda(i)L(i) + (1 - \lambda(i))Y'(i),$$  \hspace{1cm} (8)

$$\lambda(i) = \left(\frac{Y(i)}{D}ight)^r$$  \hspace{1cm} (9)

where \(r\) is the intensity remapping adjustment parameter and \(L''(i)\) is the final result for \(i\)-th pixel, which can later be optionally further processed to improve the image sharpness. After several experiments the default values of SLRMSR parameters were in [43] determined to be \(n = 4\) and \(r = 0.55\). However, in order to simplify Eq. (9) and get better performance, in this paper \(r = 1\) will be used instead of \(r = 0.55\). Finally, it should be noted that in its denoising system SLRMSR uses guided image filtering [44] and kernels of default size \(25 \times 25\). Therefore, when larger images are processed, kernels of larger size may be more appropriate.

### B. Combination

When Flash and SLRMSR are combined, there are three parameter values to be set: \(a\), \(n\), and \(r\). Since the values of \(n\) and \(r\) have already been discussed, what remains is to determine the default value for \(a\). As mentioned before and as shown in Fig. 3, for higher values of \(a\) Flash produces darker images, but after their brightness is adjusted by SLRMSR, their contrast is significantly better than in case of lower values of \(a\). With higher values of \(a\) the gain in contrast is only small and for this reason the default value of \(a\) is set to \(a = 10\). Since the described combination uses Flash to reduce the difference between intensities and then applies SLRMSR to perform the final processing, it was named Puma because pumas are known to be able to both jump very high and strike hard. Due to the fact that the initial Flash results are supposed to be dark, it makes sense to omit intensity remapping in SLRMSR. If this is done, then the obtained combination is named Light Puma.

### IV. EXPERIMENTAL RESULTS

#### A. Test images and metrics

To evaluate the performance of the proposed TMO and compare it to the performance of other TMOs, 33 images from the publicly available HDR dataset given at [46] were used. The resulting LDR image was assessed by means of objective quality measures Tone Mapped image Quality Index (TMQI) [47] and Feature Similarity Index For Tone-Mapped Images (FSITM). Both of them compare a given LDR image to its original HDR image and produce a number in interval \([0, 1]\) with a higher number meaning higher quality. TMQI evaluates structural fidelity and statistical naturalness of a tone mapped image, while FSITM relies on local phase information of an image. Their combination \(FSITM^{C\_TMQI}\), where \(C\) stands for a color channel, gives results that were shown to approximate the subjective image quality assessment even more closely. In this paper the green (G) channel is used in the mentioned combination. The TMQI-II [48] measure was

![Fig. 2: The same scene tone mapped by applying Eq. (4) for (a) \(a = 0.25\), (b) \(a = 0.5\), (c) \(a = 1\), (d) \(a = 2\), and (e) \(a = 4\).](image)

![Fig. 3: Combining Flash and SLRMSR\(_{100}\) for different values of parameter \(a\): (a) \(a = 1\) and (b) \(a = 10\). First row: the initial Flash results; second row: SLRMSR processing of Flash results.](image)
not used since FSITM$^G$-TMQI was shown to be better [49].

### B. Obtained results

To produce the results of other well-known TMOs, the Luminance HDR open-source software was used with all parameters for all nine provided individual TMOs set to default values. This was done due to the fact that a similar procedure was used in [48] and because the results can be recreated by using the same software. For Color Badger the values of $p_1$ and $p_2$ were set to 10 and 7, respectively. Table I shows the numerical results for chosen metrics obtained on the test images. In terms of TMQI Puma and Light Puma outperform all other TMOs, while for FSITM$^G$-TMQI they are outperformed only by Reinhard’s TMO [13]. These results demonstrate the quality of Puma’s and Light Puma’s Retinex-based tone mapping performance and they experimentally justify the preprocessing based on the Naka-Rushton equation. Fig. 4 shows examples of application of different TMOs.

The proposed Puma TMO and previously proposed Color Badger TMO [17] both inherently rely on the Light Random Sprays Retinex (LRSR) algorithm [35], but there are several crucial differences. Color Badger uses an ad hoc power approach to bridge the high dynamic range, while Puma uses the Naka-Rushton equation, which has solid foundations in mimicking the behaviour of human visual system. Additionally, Puma has a theoretically lower complexity because of SLRMSR. Another possible confusion when comparing Puma to Color Badger is that in the original Color Badger paper the experimental results were performed so that for each test image and tested method the best combination of parameter values was searched, while here only the default parameter values were used, which is closer to real-world application.

### V. Conclusions and future research

A local TMO based on the Naka-Rushton equation and the Retinex theory has been proposed. It has been shown to outperform most of the tested TMOs in terms of resulting LDR image quality. Since the proposed TMO relies both on the Naka-Rushton equation and the Retinex theory, there are at least two possible directions for future research. One of them includes using other brightness adjustment methods to improve the initial results of Flash, while the other is to examine which other preprocessing methods to use before applying SLRMSR.

### REFERENCES

