

Optimum Parameter Estimation of Tone Mapping Operators by Natural Image Statistics

Okazaki, Daiki

Department of Communication Design Science, Kyushu University

Hara, Kenji

Department of Communication Design Science, Kyushu University

Inoue, Kohei

Department of Communication Design Science, Kyushu University

Urahama, Kiichi

Department of Communication Design Science, Kyushu University

<https://hdl.handle.net/2324/2552957>

出版情報 : Journal of the Institute of Industrial Applications Engineers. 7 (3), pp.106-109, 2019-07-26. 産業応用工学会

バージョン :

権利関係 : Creativecommons Attribution 4.0 International (CC BY 4.0)



Optimum Parameter Estimation of Tone Mapping Operators by Natural Image Statistics

DAIKI OKAZAKI* Non-member,
 KOHEI INOUE* Member,

KENJI HARA* Non-member
 KIICHI URAHAMA* Member

(Received December 4, 2018, revised July 6, 2019)

Abstract: In this paper, we propose a method that optimizes the parameters of tone mapping operators by compressing the dynamic range of HDR images using natural image statistics. First, a prior probability model of a natural image is constructed for color natural images based on a generalized Gaussian distribution. Then, an LDR image is generated by converting the HDR image using the tone mapping operator. Next, we generate a normalized histogram of the LDR image using a discrete wavelet transformation. Finally, the optimal parameters of the tone mapping operator are estimated by minimizing the Kullback–Leibler divergence of the probability density function and the normalized histogram. Using these parameters, it is possible to generate an LDR image that closely resembles the natural image.

Keywords: high dynamic range image, tone mapping operator, natural image statistics

1. Introduction

High dynamic range (HDR) images capture and represent real-world scenes more accurately than normal low dynamic range (LDR) images are able to do. HDR images maintain the same dynamic range as human visual characteristics (HVS), and have been applied in fields such as high-quality rendering, on-vehicle cameras, surveillance cameras, and medical imaging, owing to their high versatility [1]. An HDR image is usually generated by synthesizing multiple LDR images taken under different exposure conditions [1]. To display an HDR image on a general LDR display, a tone mapping operation is necessary to compress the dynamic range of the HDR image to RGB 8 bits or 256 gradations per pixel of the LDR image using a tone mapping operator (TMO).

Various tone mapping methods have been proposed in the literature [2]–[5]. The overall brightness and contrast of the LDR image obtained via these tone mapping processes depend largely on the values of the parameters of the respective methods. Therefore, it is necessary to set appropriate values for these parameters according to the HDR image. Currently, these values are often determined empirically or by trial and error.

In this paper, we propose a method that optimizes the parameters of tone mapping operators by compressing the dynamic range of HDR images using natural image statistics. First, a prior probability model of a natural image is constructed for color natural images based on a generalized Gaussian distribution. Then, an LDR image is generated by converting the HDR image using the tone mapping op-

erator. Next, we generate a normalized histogram of the LDR image using a discrete wavelet transformation. Finally, the optimal parameters of the tone mapping operator are estimated by minimizing the Kullback–Leibler divergence of the probability density function and the normalized histogram. Using these parameters, it is possible to generate an LDR image that closely resembles the natural image.

2. Tone Mapping Operation

Tone mapping operations need to compress the dynamic range and preserve an appropriate gradation of the HDR image. As such, several tone mapping methods have been proposed in the literature [2]–[5], and can be divided broadly into global and local tone mapping methods.

2.1 Global Tone Mapping Global tone mapping performs a luminance conversion of the entire image using a uniform conversion function [2] [3]. In order to handle a wide dynamic range, Drago et al. [2] proposed a tone mapping method using a logarithm function:

$$L_d = \frac{L_{d_{max}} \cdot 0.01}{\log_{10}(L_{max} + 1)} \cdot \frac{\log_{10}(L + 1)}{\log_{10} \left(2 + 8 \left(\left(\frac{L}{L_{max}} \right)^{\frac{\log_{10}(b)}{\log_{10}(0.5)}} \right) \right)}, \quad (1)$$

where L is the luminance value of the HDR image, L_{max} is the maximum luminance value of the HDR image, L_d is the luminance value of the LDR image after the tone mapping operation, $L_{d_{max}}$ is the maximum luminance value of the LDR image, and b is a parameter that controls the compression of high values and the visibility of details in dark areas. This method needs to adaptively change the parameter b of Eq.(1) according to the scene of the input image,

* Department of Communication Design Science, Kyushu University
 4–9–1, Shiobaru, Minami-ku, Fukuoka 815–8540, Japan

based on the view that the steep slope of the logarithmic function curve with a small base value is suitable for the luminance conversion of dark areas, and that the gentle slope of the curve with a large base value is suitable for compressing light areas. Global tone mapping operations tend to lose a small amount of contrast expressing the detailed texture of a local area because a uniform brightness is applied to the input image in order to preserve the receive magnitude of the brightness values among the pixels.

2.2 Local Tone Mapping Local tone mapping performs a different luminance conversion for each local portion of an image in order to improve the reproducibility of the scene [4] [5]. Although the calculation cost is higher than that of a global tone mapping operation, it is possible to set a dynamic range suitable for each local area. In addition, a brightness conversion can be performed while maintaining information such as the texture and contrast of the HDR image. Reinhard et al. [4] proposed a tone mapping method based on a technique used to develop analog photos. First, the reference brightness in the scene of the input image is determined as:

$$\bar{L}_\omega = \frac{1}{N} \exp \left(\sum_{x,y} \log (\sigma + L_\omega(x, y)) \right), \quad (2)$$

where \bar{L}_ω is the average luminance value of the input image, N is the total number of pixels, $L_\omega(x, y)$ denotes pixel (x, y) of the input image, and σ is a small value used to avoid the singularity that occurs if back pixels are present in the image. Using the average luminance value \bar{L}_ω , each pixel value is scaled according to:

$$L(x, y) = \frac{a}{\bar{L}_\omega} L_\omega(x, y), \quad (3)$$

where $L(x, y)$ is the scaled luminance and a is a parameter that controls the brightness of the entire scene of the scaled luminance $L(x, y)$. Second, dodging-and-burning is performed to emphasize the contrast of pixels that are significantly different from the luminance value of the peripheral region. To begin with, we convolve the image with a circular convolution kernel defined using multiple spatial scales s_i ($i = 1, 2, \dots$). Let the obtained response be $\tilde{V}(x, y, s_i)$, and the local contrast according to the spatial scale s_i be given by:

$$V(x, y, s_i) = \frac{\tilde{V}(x, y, s_i) - \tilde{V}(x, y, s_{i+1})}{2^\phi a / s_i^2 + \tilde{V}(x, y, s_i)}, \quad (4)$$

where ϕ is a sharpness parameter that controls the edge emphasis. Reinhard et al. [4] consider eight scale levels for s_i : the minimum s_1 is 0.35, and is increased by 1.6 times thereafter. Third, s_m is found to first satisfy the following for each pixel:

$$|V(x, y, s_m)| < \epsilon, \quad (5)$$

where ϵ is the threshold. Finally, from the $V(x, y, s_m)$ generated in this way, the local tone mapping operation is performed using the following:

$$L_d(x, y) = \frac{L(x, y)}{1 + V(x, y, s_m)}, \quad (6)$$

where $L_d(x, y)$ is the LDR image after the local tone mapping operation. A natural LDR image can be generated by appropriately determining the parameter a and the spatial scale s_i . In local tone mapping, shadow-like noise that does not exist in the original image, called a halo artifact, may occur because the method tries to maintain the consistency of the luminance.

In the following, we describe a method for estimating the optimal values of the parameters used in these tone mapping operators.

3. Natural Image Statistics[6]

Here, we estimate the optimal values of the parameters of the tone mapping operators using natural image statistics. First, we construct a prior probability model for natural images based on a generalized Gaussian distribution (GGD). The probability density function of the GGD is expressed as follows [7]:

$$p(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-|x/\alpha|^\beta}, \quad (7)$$

where α is a scale parameter, β is a shape parameter, and $\Gamma(z) = \int_0^\infty e^{-t} t^{z-1} dt$ ($z > 0$) is the gamma function. From Eq.(7), the GGD is equal to the Gaussian distribution when $\beta = 2$ and to the Laplace distribution when $\beta = 1$. In addition, the GGD converges to the Dirac delta function distribution when $\beta \rightarrow 0+$ and to the uniform distribution when $\beta = 2$. Next, a large number of color natural images are prepared, and a discrete wavelet transform is applied to the CIELUV channel of each natural image. Then, a high-pass subband (SB) composed of three different frequency levels and directions (i.e., nine horizontal directions $\text{SB}\{cH_j^{(\lambda)}\}_{j=1,2,3,\lambda=L^*,u^*,v^*}$, nine vertical directions $\text{SB}\{cV_j^{(\lambda)}\}_{j=1,2,3,\lambda=L^*,u^*,v^*}$, and nine oblique directions $\text{SB}\{cD_j^{(\lambda)}\}_{j=1,2,3,\lambda=L^*,u^*,v^*}$) and low-pass $\text{SB}\{cA_j^{(\lambda)}\}_{j=1,2,3,\lambda=L^*,u^*,v^*}$ are extracted. Here, $j = 1, 2, 3$ is the resolution level of the discrete wavelet transform and $\lambda = L^*, u^*, v^*$ is the CIELUV color channel.

Furthermore, a normalized histogram (hereafter, an SB histogram) is generated for the 27 (= 9×3) SB components $\{cG_j^{(\lambda)}\}_{G=H,V,D,j=1,2,3,\lambda=L^*,u^*,v^*}$ of each natural image. Here, the bin width of the SB histogram is set to $H = 3.5S/N^{1/3}$ (where N is the number of pixels and S is the standard deviation of the pixel values), as per Scott's formula [8]. For each of these SB histograms, the parameters of the GGD are obtained by using an MLE. Then, for each of the 27 parameters of the GGD for each SB component α and β , the average value $\{\{\hat{\alpha}_{G_j^{(\lambda)}}, \hat{\beta}_{G_j^{(\lambda)}}\}_{G=H,V,D,j=1,2,3,\lambda=L^*,u^*,v^*}\}$ per natural image is calculated and used as a preliminary model of the natural image. In the experiment, 6000 images were used for the above learning.

4. Parameter Optimization

Using a tone mapping operator with parameters chosen such that the difference between the SB histogram of the output image and the probability density function under the GGD parameter A is small, we expect to obtain images

with reduced texture information and reduced halo artifacts. Kullback–Leibler divergence (KLD) is used as a measure of the distance between probability distributions. Here, we optimize the parameter b in Eq.(1) of the global tone mapping method of Drago et al. [2] and the parameter a and the spatial scale s_i in Eq.(3)–(6) of the local tone mapping method of Reinhard et al. [4]. We use KLD to solve each of the following:

$$\bar{b} = \underset{b}{\operatorname{argmin}} \sum_{\lambda \in \{L^*, u^*, v^*\}} \sum_{j=1}^3 \sum_{G \in \{H, V, D\}} D_{KL}(P_{G_j^\lambda} \| Q_{G_j^\lambda}(b)) \quad (8)$$

$$(\bar{a}, \bar{s}_m) = \underset{a, s_m}{\operatorname{argmin}} \sum_{\lambda \in \{L^*, u^*, v^*\}} \sum_{j=1}^3 \sum_{G \in \{H, V, D\}} D_{KL}(P_{G_j^\lambda} \| Q_{G_j^\lambda}(a, s_m)), \quad (9)$$

where $D_{KL}(P_{G_j^\lambda} \| Q_{G_j^\lambda}(b))$ is the KLD of the probability distributions $P_{G_j^\lambda}$ and $Q_{G_j^\lambda}(b)$, and $D_{KL}(P_{G_j^\lambda} \| Q_{G_j^\lambda}(a, s_m))$ is the KLD of the probability distributions $P_{G_j^\lambda}$ and $Q_{G_j^\lambda}(a, s_m)$. We solve for \bar{b} in Eq.(8) as the optimum parameter of the tone mapping operator of Drago et al. [2], and for (\bar{a}, \bar{s}_m) in Eq.(9) as the optimum parameters and spatial scale of the tone mapping operator of Reinhard et al. [4]. Here, Eq.(8) and Eq.(9) are rewritten as follows,

$$D_{KL}(P_{G_j^\lambda} \| Q_{G_j^\lambda}(b)) = \sum_i p(x_i; \hat{\alpha}_{G_j^\lambda}, \hat{\beta}_{G_j^\lambda}) \log \frac{p(x_i; \hat{\alpha}_{G_j^\lambda}, \hat{\beta}_{G_j^\lambda})}{Q_{G_j^\lambda}(i|b)}, \quad (10)$$

$$D_{KL}(P_{G_j^\lambda} \| Q_{G_j^\lambda}(a, s_m)) = \sum_i p(x_i; \hat{\alpha}_{G_j^\lambda}, \hat{\beta}_{G_j^\lambda}) \log \frac{p(x_i; \hat{\alpha}_{G_j^\lambda}, \hat{\beta}_{G_j^\lambda})}{Q_{G_j^\lambda}(i|a, s_m)}. \quad (11)$$

Here, x_i is the median of the i -th bin in the SB histogram generated for the output image by the tone mapping operator, $p(x_i; \hat{\alpha}_{G_j^\lambda}, \hat{\beta}_{G_j^\lambda})$ ($G = H, V, D; j = 1, 2, 3; \lambda = L^*, u^*, v^*$) is the value in $x = x_i$ of the probability density function of the GGD when the parameter value is $(\alpha, \beta) = (\hat{\alpha}_{G_j^\lambda}, \hat{\beta}_{G_j^\lambda})$ ($G = H, V, D; j = 1, 2, 3; \lambda = L^*, u^*, v^*$), $Q_{G_j^\lambda}(i|b)$ ($G = H, V, D; j = 1, 2, 3; \lambda = L^*, u^*, v^*$) is the frequency of the i -th bin of each SB histogram generated for the output image by the tone mapping operator of Drago et al. [2], and $Q_{G_j^\lambda}(i|a, s_m)$ ($G = H, V, D; j = 1, 2, 3; \lambda = L^*, u^*, v^*$) is the frequency of the i -th bin of each SB histogram generated for the output image by Reinhard et al. [4] tone mapping operator.

5. Experimental Results

In this experiment, we use HDR images collected from the Internet and HDR images obtained by converting multiple LDR images generated by shooting an actual scene, while varying the exposure time using Debevec's method [1]. We shot 40 multiple exposure images to create one HDR image. The experiment uses Drago's method [2] as a global

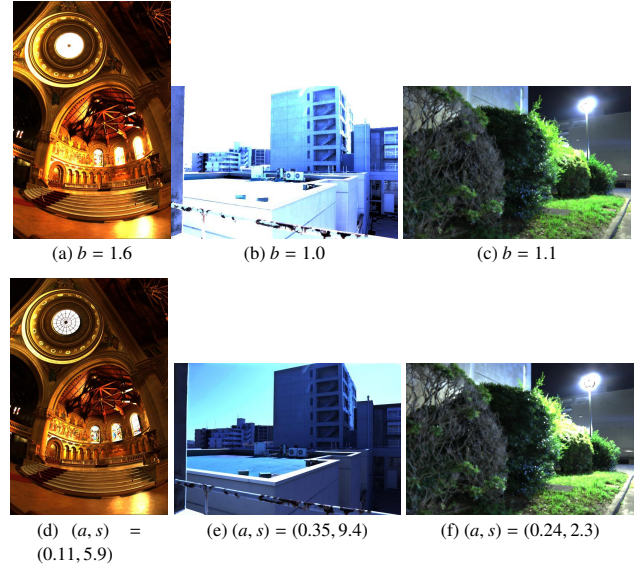


Figure 1: LDR images generated by performing the tone mapping operation using the proposed method. (a)–(c) ours–Drago, (d)–(f) ours–Reinhard.

tone mapping operator and Reinhard's method [4] as a local tone mapping operator. The parameter a and the spatial scale s are defined according to Reinhard et al. [4] and the tone mapping operation is applied as a conventional method (hereafter, referred to as Reinhard). The tone mapping operation performed using the parameter of the proposed method is applied to (1) Drago et al. (hereinafter referred to as “ours–Drago”) and (2) Reinhard et al. (hereinafter referred to as “ours–Reinhard”). Then, we compare the performance of the methods.

Fig.1 shows the results of the LDR image generated by processing the HDR image with the tone mapping operation corresponding to the minimum KLD solution.

Fig.2 compares each LDR image generated by applying the local tone mapping operation by the proposed method and the conventional method. We compare the results generated by Reinhard as a conventional method and ours–Reinhard as the proposed method. The results show that the texture is clearer when using the proposed method than when using the conventional method.

Tables 1 and 2 summarize the quantitative evaluation of the proposed method. In the experiment, we evaluated four sheets, shown in Fig.1(b),(c),(e),(f). First of all, we chose two images, one with off-white skipping and one with black crushed, from the LDR image generated by shooting an actual scene. Next, an area which is not white skipping is arbitrarily selected from the image which is white skipping, and an area which is not black crushed is similarly selected from the image having black crush. For the global tone mapping operation, we calculate the correlation coefficient between the two areas and the same two areas of each LDR image subjected to the tone mapping process with the value of ± 0.1 of the parameter b of the proposed method and calculate the average value of the two correlation coefficients (Table 1). To evaluate the local tone mapping op-

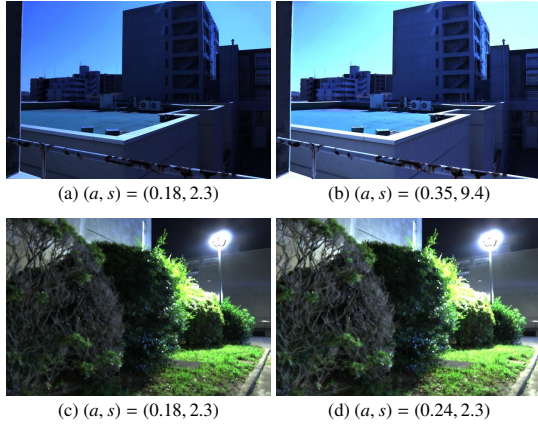


Figure 2: Comparison between proposed method and conventional method. (a) Reinhard, (b) ours-Reinhard, (c) Reinhard, (d) ours-Reinhard.

Table 1: Quantitative comparison of global tone mapping operation.

| Parameter b | Image | Correlation coefficient |
|---------------|----------|-------------------------|
| 0.9 | Fig.1(b) | 0.7459 |
| 1.0 | | 0.7466 |
| 1.1 | | 0.7455 |
| 1.0 | Fig.1(c) | 0.8528 |
| 1.1 | | 0.8528 |
| 1.2 | | 0.8527 |

Table 2: Quantitative comparison of local tone mapping operation.

| Method | Image | Correlation coefficient |
|---------------|----------|-------------------------|
| Reinhard | Fig.2(a) | 0.8373 |
| ours-Reinhard | Fig.2(b) | 0.8658 |
| Reinhard | Fig.2(c) | 0.8357 |
| ours-Reinhard | Fig.2(d) | 0.8574 |

eration, we calculate the correlation coefficient between the two areas and the same two areas of the LDR image obtained using Reinhard and ours-Reinhard and calculate the average value of the two correlation coefficients (Table 2). Tables 1 and 2 show that the correlation coefficient is maximized for the parameter value corresponding to the minimum KLD solution.

6. Conclusion

We have proposed a parameter optimization method for a tone mapping operator using a prior probability model of a natural image based on the generalized Gaussian distribution. Then, we confirmed the effectiveness of the proposed method in an experiment using actual HDR images. As a future task, we will attempt to speed up the proposed method.

References

- [1] P. E. Debevec, and J. Malik, “Recovering high dynamic range radiance maps from photographs”, *In Proceedings of SIGGRAPH’97*, pp.369–378, 1997.

- [2] F. Darago, K. Myszkowski, T. Annen, and N. Chiba, “Adaptive logarithmic mapping for displaying high contrast scenes”, *Computer Graphics Forum*, vol.3, pp.419–426, 2003.
- [3] J. Tumblin and, G. Turk, “LCIS: A boundary hierarchy for detail-preserving contrast reduction”, *In Proceedings of SIGGRAPH’99*, pp.83–90, 1999.
- [4] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, “Photographic tone reproduction for digital images”, *In Proceedings of SIGGRAPH’02*, vol.21, no.3, pp.267–276, July 2002.
- [5] R. Fattal, D. Lischinski, and M. Werman, “Gradient domain high dynamic range compression”, *In Proceedings of SIGGRAPH’02*, vol.21, pp.249–256, 2002.
- [6] A. Morinaga, K. Hara, K. Inoue, and K. Urahama, “Classification between natural and graphics images based on generalized Gaussian distributions”, *Information Processing Letters (IPL)*, no.138, pp.31–34, Oct 2018.
- [7] G.E.P. Box, and G.C. Tiao, “Bayesian Inference in Statistical Analysis”, *John Wiley & Sons*, New York, 1992.
- [8] D.W. Scott, “On optimal and data-based histograms”, *Biometrika*, vol.66, no.3, pp.605–610, 1979.



Daiki Okazaki (Non-member) He is now studying for a master’s degree in the Kyushu University. His research interests include omnidirectional image and high dynamic range image.



Kenji Hara (Non-member) He received the BE and ME degrees from Kyoto University in 1987 and 1989, respectively, and the PhD degree from Kyushu University in 1999. He is currently an Associate Professor in Kyushu University. His research interests include physics-based vision and geometric modeling.



Kohei Inoue (Member) He received B.Des., M.Des. and D.Eng. degrees from Kyushu Institute of Design in 1996, 1998 and 2000, respectively. He is currently an Associate Professor in Kyushu University. His research interests include pattern recognition and image processing.



Kiichi Urahama (Member) He received M.Eng and D.Eng. degrees from Kyushu University in 1976 and 1980. From 1980 to 1995 he was an Associate Professor in Kyushu Institute of Technology. He is now a Professor in Kyushu University. His research interests include pattern recognition, image processing and computer graphics.