

# A Retinal Adaptation Model for HDR Image Compression

Xuan Pu, Kaifu Yang, and Yongjie Li<sup>(✉)</sup>

Key Laboratory for Neuroinformation of Ministry of Education,  
University of Electronic Science and Technology of China,  
Chengdu 610054, China  
{yangkf,liyj}@uestc.edu.cn

**Abstract.** High dynamic range (HDR) images are usually used to capture more information of natural scenes, because the light intensity of real world scenes commonly varies in a very large range. Humans visual system is able to perceive this huge range of intensity benefiting from the visual adaptation mechanisms. In this paper, we propose a new visual adaptation model based on the cone- and rod-adaptation mechanisms in the retina. The input HDR scene is first processed in two separated channels (i.e., cone and rod channels) with different adaptation parameters. Then, a simple receptive field model is followed to enhance the local contrast of the visual scene and improve the visibility of details. Finally, the compressed HDR image is obtained by recovering the fused luminance distribution to the RGB color space. Experimental results suggest that the proposed retinal adaptation model can effectively compress the dynamic range of HDR images and preserve local details well.

**Keywords:** Visual adaptation · High dynamic range  
Image enhancement · Human visual system

## 1 Introduction

The range of the light intensity in natural scenes of the real world is so vast that the illumination intensity of sunlight can be 100 million times higher than that of starlight. Therefore, high dynamic range (HDR) images are usually used to capture more information of natural scenes. However, most display devices available to us have a limited dynamic range [21]. HDR tone mapping or dynamic range compression methods are usually required to reproduce the HDR images matching the dynamic range of the standard display devices, so that the details in both the dark and bright areas are as faithfully visible as possible.

In recent years, many HDR image compression methods have been proposed. Roughly, the existing methods can be classified into two categories: global operators and local operators. Global operators apply the same transformation to each pixel, without considering the spatial position. For example, Tumbling and

Rushmerier proposed a tone mapping operator in 1993 [25], which uses a single spatially invariant level for the scene and another adaptation level for the display. Although it can preserve the brightness value, but lost the visibility of high dynamic range scenes [12]. Instead of brightness, Ward [26] and Ferwerda et al. [8] aimed at preserving the contrast. They used a scaling factor to transform real-world luminance values to displayable range. They believe that a Just Noticeable Difference (JND) in the real world can be mapped as the JND on the display device. They built their global models based on the threshold versus intensity (TVI) function. Different from Ward’s contrast-based scale factor that only considers the photopic lighting conditions, Ferwerda et al. added a scotopic component. These models can preserve the contrast well but may lose visibility of the very high and very low intensity regions [12]. In the year of 2000, Pattanaik et al. [20] proposed a new time-dependent tone reproduction operator following the framework proposed by Tumblin and Rushmeier. In short, although these global models are low computational cost, they are difficult to preserve the details with high contrast scenes.

In contrast, local operators adapt the mapping functions to the statistics and contexts of local pixels. In comparison to the global operators, local operators perform better on preserving details. However, a fundamental problem with local operators is the halos artifacts appearing around the high-contrast edges. For example, Pattanaik et al. [19] described a comprehensive computational model of human visual system adaptation and spatial vision for realistic tone reproduction. This model is able to display HDR scenes on conventional display devices, but the dynamic range compression is performed by applying different gain-control factors to each band-pass, which causes strong halo effects. To solve this problem, Durand et al. [6] presented a method with edge-preserving filter (bilateral filter) to integrate local intensities for avoiding halo artifacts.

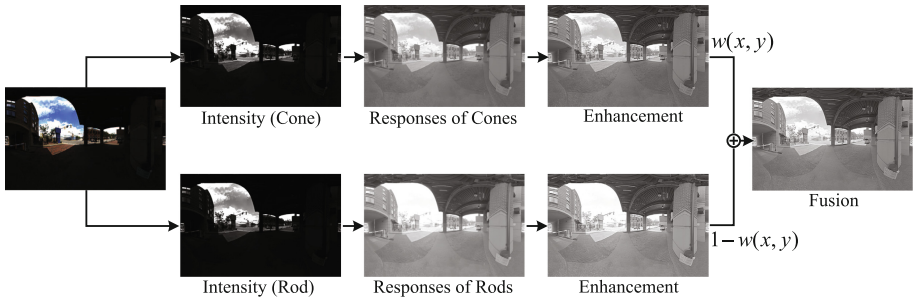
In 2004, Ledda et al. [13] proposed a local model based on the work of Pattanaik et al. [20]. The adaptation part of that method tries to model the mechanisms of human visual adaptation, with the help of a bilateral filter to avoid halo artifacts. Along this line, there are also some other works that attempt to build bio-inspired methods for HDR image rendering and dynamic range compression [9, 14, 28]. Biologically, human visual system can respond to huge luminance range, depending on the ability of visual adaptation, mainly the dark adaptation and light adaptation. Dark adaptation refers to how the visual system recovers its sensitivity when going from a bright environment to a dark one, while light adaptation indicates the processing that visual system recovers its sensitivity when we go from a dark environment to a bright one. In the retina, cone photoreceptors respond to higher light levels while rods are highly sensitive to the light of the dark and dim condition. Therefore, the light and dark adaptations is related to the switch of cones and rods.

In this paper, we propose a new visual adaptation model to compress the dynamic range of HDR images according to the mechanisms of cones and rods in the retina. We design two separated channels (cone and rod channels) with different adaptation properties to compress the dynamic range of light and dark information in the HDR images. The simple receptive field model is further

used to enhance the local contrast for improving the visibility of details. Finally, the compressed HDR image is obtained by recovering the fused luminance to the RGB color space. Experimental results suggest that the proposed retinal adaptation model can effectively compress the dynamic range of HDR image and preserve the scene details well.

## 2 The Proposed Model

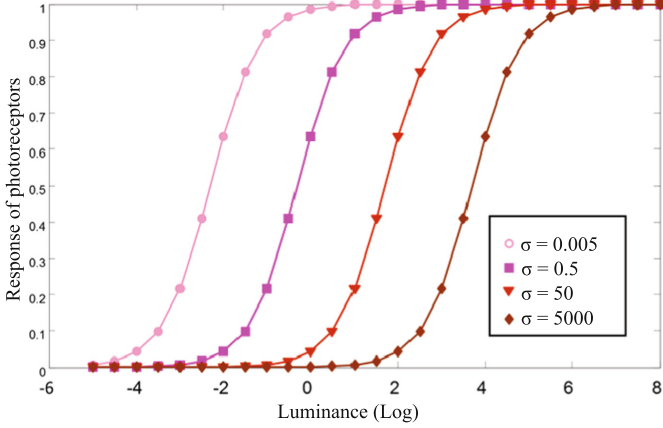
In this study, we proposed a new dynamic range compression method aiming at modeling the mechanisms of visual adaptation. The input image is first separated into two luminance distribution maps for the rod and cone channels. Then the responses of rods and cones are obtained with the Naka-Rushton adaptive model [16] according to the local luminance of the scene. A local enhancement operator based on the receptive field of ganglion cells is then used to enhance the local contrast in both the cone and rod channels. Finally, cone and rod channels are adaptively fused into a unified luminance map with compressed dynamic range and enhanced local contrast. The flowchart of dynamic range compression in the proposed model is shown in Fig. 1. Finally, the luminance information is recovered to the RGB space to obtain the color images.



**Fig. 1.** The flowchart of dynamic range compression in the proposed model. The spatially varying weighting  $w(x, y)$  for fusion is computed with Eq. (10).

### 2.1 Response of Photoreceptors

According to the visual adaptation mechanisms, cone and rod photoreceptors vary their response range dynamically to better adapt the available luminance of the environment. The model proposed by Naka and Rushton [16] has been widely used to describe the responses of cones [2, 11] and rods [4]. It indicates that the response of a photoreceptor at any adaptation level can be simply described as  $R = I^n / (I^n + \sigma^n)$ , where  $I$  is the light intensity and  $\sigma$  is the semi-saturation parameter determined by the adaptation level,  $n$  is a sensitivity-control exponent which is normally between the range of 0.7–1.0 [3, 17]. Thus, the dark and light adaptations can be explained by changing the value of  $\sigma$  at varying luminance levels. For example, on a sunny day, we cannot see well at the beginning when we



**Fig. 2.** Response to luminance with different adaptation levels.

enter a dark room. It is just because the value of  $\sigma$  is originally set at high adaptation level while  $I$  is of low intensity, which makes the response  $R$  almost zero. But after tens of minutes, visual sensitivity is restored after  $\sigma$  changes into a smaller adaptation level and the response  $R$  increases. Figure 2 shows the response curves changing with different adaptation levels by varying the value of  $\sigma$ .

HDR image compression mainly refers to compress the dynamic range of luminance information. In this work, we first convert the input HDR image (RGB) into two different luminance channels for further processing by rods and cones [19], i.e., the photopic luminance of cones ( $L_{cone}$ ) and scotopic luminance of rods ( $L_{rod}$ ):

$$L_{cone} = 0.25 \cdot I_{in}^R + 0.67 \cdot I_{in}^G + 0.065 \cdot I_{in}^B \quad (1)$$

$$L_{rod} = 0.702 \cdot X + 1.039 \cdot Y + 0.433 \cdot Z \quad (2)$$

where,  $I_{in}^R$ ,  $I_{in}^G$  and  $I_{in}^B$  are respectively the R, G and B components of the given HDR image, while  $X$ ,  $Y$  and  $Z$  are the three components when transforming the original RGB space to XYZ color space. Thus, we can obtain the cone and rod luminance responses referring to Naka-Rushton equation [16] according to

$$R_{cone}(x, y) = \frac{L_{cone}^n(x, y)}{L_{cone}^n(x, y) + \sigma_{cone}^n(x, y)} \quad (3)$$

$$R_{rod}(x, y) = \frac{L_{rod}^n(x, y)}{L_{rod}^n(x, y) + \sigma_{rod}^n(x, y)} \quad (4)$$

Following the previous work [27], there is an empirical relation between the adaptation parameter  $\sigma$  and visual intensity levels as

$$\sigma_{cone} = L_{cone}^\alpha \cdot \beta \quad (5)$$

$$\sigma_{rod} = L_{rod}^\alpha \cdot \beta \quad (6)$$

In this paper, we experimentally set the value of  $\alpha$  as 0.69 for both the cone and rod channels, but  $\beta$  is a constant that differs for the rods and cones, i.e.,  $\beta = 4$  for the cone channel and  $\beta = 2$  for the rod channel.

## 2.2 Local Enhancement and Fusion

The responses of cones and rods pass through the retina which contains the bipolar cells, horizontal cells, retinal ganglion cells, etc. In this paper, we just focus on the role of local enhancement by the retinal ganglion cells. We use a difference-of-Gaussians (DOG) model [23] to simulate the receptive field of ganglion cells for enhancing the local contrast of the visual scene. Thus the output of ganglion cells ( $G_{cone}$  and  $G_{rod}$ ) can be computed as:

$$G_{cone}(x, y) = R_{cone}(x, y) * (g(x, y; \sigma_{rf}^c) - k \cdot g(x, y; \sigma_{rf}^s)) \quad (7)$$

$$G_{rod}(x, y) = R_{rod}(x, y) * (g(x, y; \sigma_{rf}^c) - k \cdot g(x, y; \sigma_{rf}^s)) \quad (8)$$

$$g(x, y; \sigma_{rf}) = \frac{1}{2\pi\sigma_{rf}^2} \exp\left(-\frac{x^2 + y^2}{2\sigma_{rf}^2}\right) \quad (9)$$

where  $*$  denotes the convolution operator,  $\sigma_{rf}^c$  and  $\sigma_{rf}^s$  are respectively the standard deviations of Gaussian shaped receptive field center and its surround, which are experimentally set to be 0.5 and 2.0 in this work.  $k$  denotes the sensitivity of the inhibitory annular surround.

To combine the ganglion cell's outputs along the cone and rod channels, we use a sigmoid function to describe the weighting of cone and rod systems when fusing the two signals.

$$w(x, y) = \frac{1}{0.2 + L_{cone}(x, y)^{-0.1}} \quad (10)$$

where the fixed parameters (i.e., 0.2 and  $-0.1$ ) are experimentally set.

From Eq. (10), the value of  $w(x, y)$  is increased with the increasing of luminance, which basically matches the light and dark adaptation mechanisms. In the photopic condition, the cone system makes more contribution, while the rod system is more sensitive in the scotopic range. Thus, the fused luminance response is given by

$$L_{out}(x, y) = w(x, y) \cdot G_{cone}(x, y) + (1 - w(x, y)) \cdot G_{rod}(x, y) \quad (11)$$

## 2.3 Recovering of the Color Image

To recombine the luminance information into a color image and assure color remains stable, we keep the ratio between the color channels constant before and after compression [10]. Considering that for some input images with high saturation, the output images may appear over saturation when using this simple

rule, we add an exponent  $s$  to control the saturation of the output image [21], which is described as

$$I_{out}^c(x, y) = L_{out}(x, y) \left( \frac{I_{in}^c(x, y)}{L_{in}(x, y)} \right)^s, \quad c \in \{R, G, B\} \quad (12)$$

where  $I_{out}^c(x, y), c \in \{R, G, B\}$  are the RGB channels of the output image,  $L_{in}(x, y)$  is replaced using  $L_{cone}(x, y)$ . The exponent  $s$  is given as a parameter between 0 and 1, and we set  $s = 0.8$  for the most scenes.

### 3 Experimental Results

In this section, we evaluated the performance of the proposed method by comparing our method with some existing algorithms. We considered two representative local algorithms including the methods proposed by Durand and Dorsey [18] and Meylan et al. [15], and one typical global algorithm proposed by Pattanaik et al. [20]. The model of Durand et al. uses a bilateral filter to integrate the local intensity. The algorithm proposed by Meylan et al. is an adaptive Retinex model. The method of Pattanaik et al. is aimed at simulating certain mechanisms of human visual system, but with a global operation.

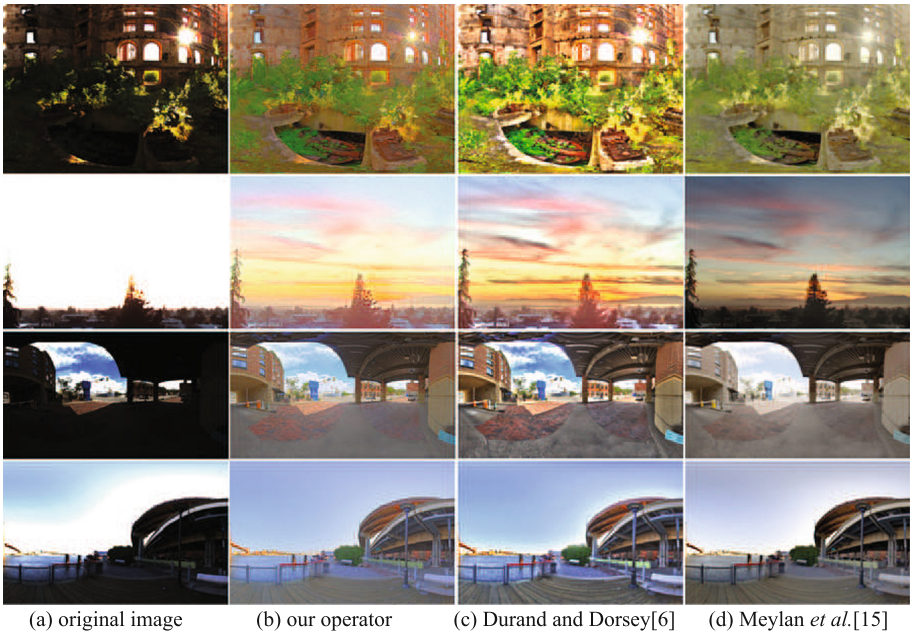
Experimental results are shown in Figs. 3, 4, 5, 6 and 7. In Fig. 3, we tested the methods with an indoor HDR scene named as “office”. Figure 3(a)–(d) lists



**Fig. 3.** Comparison of the results with local algorithms on the indoor “office” image.

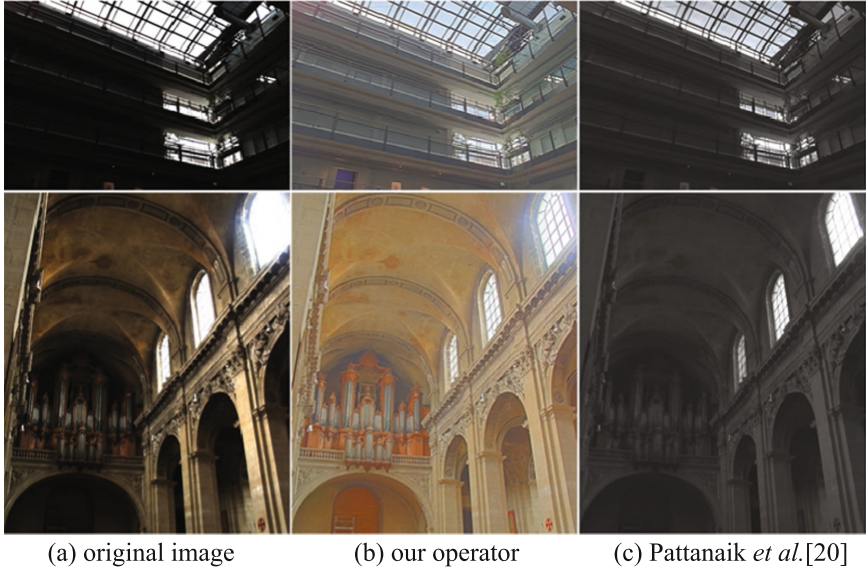


**Fig. 4.** Comparing the results with local algorithms (inside scenes).



**Fig. 5.** Comparison of the results with local algorithms on four outdoor scenes.

the results of the whole image (the first row), the local details in the dark area (the second row), and the local details in the bright area (the third row). For the dark regions, our and Meylan's results are better than that of the Durand's



**Fig. 6.** Comparing the results with a global algorithm.

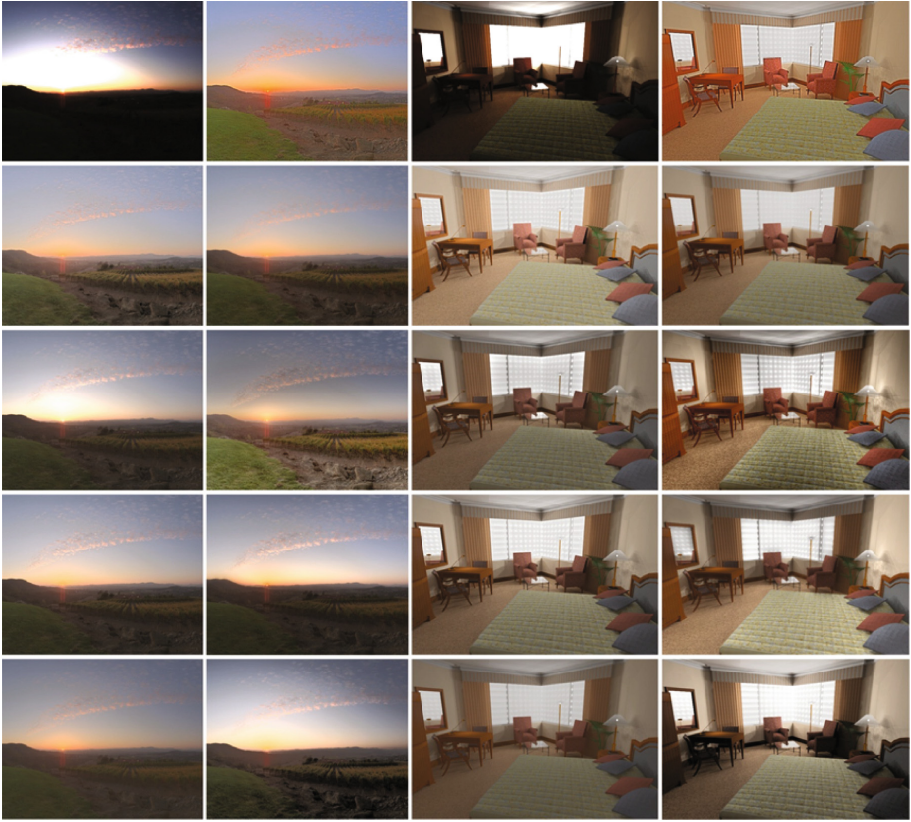
method. For the bright region, our method obtains good luminance compression and preserves more details. The last row (Fig. 3(e)) shows different exposure levels of the HDR image, it clearly shows that our model and Durand’s method still keep the wall a little blue which is similar to the original HDR image, but the Meylan’s result has serious color cast, e.g., the wall appears even a little yellow. Other two inside HDR scenes are also shown in Fig. 4. Durand’s model has serious contrast reversals artifacts around high contrast edges. For instance, unlocked lamps on the ceiling in top image of Fig. 4, their edges are out of shape.

We also selected four outdoor HDR scenes to give more comparisons. The results are shown in Fig. 5. In some situations, Durand’s algorithm could lose some information, while our method preserves more clear details in both the dark and light situations with less color cast.

We also compared our model with the method of Pattanaik et al. [20], a global tone reproduction operator (shown in Fig. 6). The results of Pattanaik et al. are cited from pfstools ([http://pfstools.sourceforge.net/tmo\\_gallery/](http://pfstools.sourceforge.net/tmo_gallery/)). We can see that Pattanaik’s results lose much color information and have lower contrast.

We further compared our method with Ashikhmin’s tone mapping algorithm [1], Drago’s adaptive logarithmic mapping [5], Durand’s bilateral filtering [6], Fattal’s gradient domain compression [7], Reinhard’s photographic tone reproduction [22], Tumblin’s fovea interactive method [24] and Ward’s contrast-based operator [26]. The results of these methods are downloaded from the Max Planck Institut Informatik (MPII) (<http://resources.mpi-inf.mpg.de/tmo/NewExperiment/TmoOverview.html>). Two examples of Napa Valley and Hotel Room scenes in MPII dataset are shown in Fig. 7. We calculated the entropy





**Fig. 7.** comparison on Napa Valley and Hotel Room from MPII. In every scene from top to bottom and left to right there are the original HDR image, our results, Ashikhmin [1], Drago [5], Durand [6], Fattal [7], Reinhard [22], Drago's Retinex, Tumblin [24] and Ward [26] respectively.

**Table 1.** Quantitative comparison on images from MPII with entropy as metric.

	Napa Valley	Hotel Room	Atrium Morning	Atrium Night	Memorial
Ashikmin [1]	7.416	6.966	7.778	7.972	7.769
Drago [5]	7.381	6.756	7.421	7.907	7.753
Durand [6]	7.182	6.695	7.046	7.424	7.291
Fattal [7]	7.600	7.445	8.031	7.455	8.504
Reinhard [22]	7.232	6.817	7.375	7.860	7.837
Tumblin [24]	7.016	6.391	6.995	7.491	7.668
Ward [26]	7.416	6.694	7.251	7.915	7.783
<b>Proposed</b>	<b>7.941</b>	<b>7.267</b>	<b>7.830</b>	<b>8.309</b>	<b>8.030</b>

values as quantitative metric. The entropy of our method and the seven compared methods on five scenes are listed in Table 1. A higher entropy score means the richer information contained in an image, indicating the better performance of the method.

## 4 Conclusion

In this paper, we proposed a visual adaptation model based on the retinal adaptation mechanisms to compress the dynamic range of HDR images. The model has considered the sensitivity changes based on the light and dark adaptation mechanisms and the receptive field property of retinal ganglion cells. We have compared our results with several typical compression operators, on both indoor and outdoor scenes. In general, the proposed model can efficiently compress the dynamic range and enhance the local contrast. Considering that some high level information in the visual system can usually serve as an important guidance to improve the low level visual processing, in the future work we will consider to add the global visual effect when compressing the dynamic range.

**Acknowledgments.** This work was supported by the Major State Basic Research Program under Grant 2013CB329401, and the Natural Science Foundations of China under Grant 61375115, 91420105. The work was also supported by the Fundamental Research Funds for the Central Universities under Grant ZYGX2016KYQD139.

## References

1. Ashikhmin, M.: A tone mapping algorithm for high contrast images. In: 13th Eurographics Workshop on Rendering, Pisa, Italy, 26–28 June 2002, p. 145. Association for Computing Machinery (2002)
2. Baylor, D., Fuortes, M.: Electrical responses of single cones in the retina of the turtle. *J. Physiol.* **207**(1), 77–92 (1970)
3. Boynton, R.M., Whitten, D.N.: Visual adaptation in monkey cones: recordings of late receptor potentials. *Science* **170**(3965), 1423–1426 (1970)
4. Dowling, J.E., Ripps, H.: Adaptation in skate photoreceptors. *J. Gen. Physiol.* **60**(6), 698–719 (1972)
5. Drago, F., Myszkowski, K., Annen, T., Chiba, N.: Adaptive logarithmic mapping for displaying high contrast scenes. *Comput. Graph. Forum* **22**, 419–426 (2003). Wiley Online Library
6. Durand, F., Dorsey, J.: Fast bilateral filtering for the display of high-dynamic-range images. *ACM Trans. Graph. (TOG)* **21**, 257–266 (2002). ACM
7. Fattal, R., Lischinski, D., Werman, M.: Gradient domain high dynamic range compression. *ACM Trans. Graph. (TOG)* **21**, 249–256 (2002). ACM
8. Ferwerda, J.A., Pattanaik, S.N., Shirley, P., Greenberg, D.P.: A model of visual adaptation for realistic image synthesis. In: Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques, pp. 249–258. ACM (1996)
9. Gao, S., Han, W., Ren, Y., Li, Y.: High dynamic range image rendering with a luminance-chromaticity independent model. In: He, X., Gao, X., Zhang, Y., Zhou, Z.-H., Liu, Z.-Y., Fu, B., Hu, F., Zhang, Z. (eds.) *ISCIde 2015*. LNCS, vol. 9242, pp. 220–230. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-23989-7\\_23](https://doi.org/10.1007/978-3-319-23989-7_23)

10. Hall, R.: *Illumination and Color in Computer Generated Imagery*. Springer, New York (2012)
11. Hiroto, I., Hirano, M., Tomita, H.: Electromyographic investigation of human vocal cord paralysis. *Ann. Otol. Rhinol. Laryngol.* **77**(2), 296–304 (1968)
12. Larson, G.W., Rushmeier, H., Piatko, C.: A visibility matching tone reproduction operator for high dynamic range scenes. *IEEE Trans. Vis. Comput. Graph.* **3**(4), 291–306 (1997)
13. Ledda, P., Santos, L.P., Chalmers, A.: A local model of eye adaptation for high dynamic range images. In: *Proceedings of the 3rd International Conference on Computer Graphics, Virtual Reality, Visualisation and Interaction in Africa*, pp. 151–160. ACM (2004)
14. Li, Y., Pu, X., Li, H., Li, C.: A retinal adaptation model for HDR image compression. In: *Perception*, vol. 45, pp. 45–45. Sage Publications Ltd., London (2016)
15. Meylan, L., Susstrunk, S.: High dynamic range image rendering with a retinex-based adaptive filter. *IEEE Trans. Image Process.* **15**(9), 2820–2830 (2006)
16. Naka, K., Rushton, W.: S-potentials from colour units in the retina of fish (cyprinidae). *J. Physiol.* **185**(3), 536–555 (1966)
17. Normann, R.A., Werblin, F.S.: Control of retinal sensitivity. *J. Gen. Physiol.* **63**(1), 37–61 (1974)
18. Paris, S., Durand, F.: A fast approximation of the bilateral filter using a signal processing approach. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) *ECCV 2006*. LNCS, vol. 3954, pp. 568–580. Springer, Heidelberg (2006). [https://doi.org/10.1007/11744085\\_44](https://doi.org/10.1007/11744085_44)
19. Pattanaik, S.N., Ferwerda, J.A., Fairchild, M.D., Greenberg, D.P.: A multiscale model of adaptation and spatial vision for realistic image display. In: *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 287–298. ACM (1998)
20. Pattanaik, S.N., Tumblin, J., Yee, H., Greenberg, D.P.: Time-dependent visual adaptation for fast realistic image display. In: *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques*, pp. 47–54. ACM Press/Addison-Wesley Publishing Co. (2000)
21. Reinhard, E., Heidrich, W., Debevec, P., Pattanaik, S., Ward, G., Myszkowski, K.: *High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting*. Morgan Kaufmann, Burlington (2010)
22. Reinhard, E., Stark, M., Shirley, P., Ferwerda, J.: Photographic tone reproduction for digital images. *ACM Trans. Graph. (TOG)* **21**(3), 267–276 (2002)
23. Rodieck, R.W., Stone, J.: Response of cat retinal ganglion cells to moving visual patterns. *J. Neurophysiol.* **28**(5), 819–832 (1965)
24. Tumblin, J., Hodgins, J.K., Guenter, B.K.: Two methods for display of high contrast images. *ACM Trans. Graph. (TOG)* **18**(1), 56–94 (1999)
25. Tumblin, J., Rushmeier, H.: Tone reproduction for realistic images. *IEEE Comput. Graph. Appl.* **13**(6), 42–48 (1993)
26. Ward, G.: A contrast-based scalefactor for luminance display. In: *Graphics Gems IV*, pp. 415–421. Academic Press Professional Inc., San Diego (1994)
27. Xie, Z., Stockham, T.G.: Toward the unification of three visual laws and two visual models in brightness perception. *IEEE Trans. Syst. Man Cybern.* **19**(2), 379–387 (1989)
28. Zhang, X.-S., Li, Y.-J.: A retina inspired model for high dynamic range image rendering. In: Liu, C.-L., Hussain, A., Luo, B., Tan, K.C., Zeng, Y., Zhang, Z. (eds.) *BICS 2016*. LNCS, vol. 10023, pp. 68–79. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-49685-6\\_7](https://doi.org/10.1007/978-3-319-49685-6_7)