Multi-Scale histogram tone mapping algorithm enables better object detection in wide dynamic range images

Conference Paper · August 2017

DOI: 10.1109/AVSS.2017.8078533

Jie Yang
The University of Calgary

Alain Horé
The University of Calgary

Multi-scale Histogram Tone Mapping Algorithm Enables Better Object Detection in Wide Dynamic Range Images

Jie Yang, Alain Horé, Ulian Shahnovich, Kenneth Lai,
S. N. Yanushkevich, *Senior Member, IEEE*, and Orly Yadid-Pecht, *Fellow IEEE*
University of Calgary,
2500 University Drive NW Calgary, AB T2N 1N4 CANADA
jie.yang2@ucalgary.ca

Abstract

In this paper, we present a novel tone mapping algorithm based on multi-scale histograms and fusion (MS-Hist), for displaying wide dynamic range (WDR) images and better detection of objects such as human faces. The proposed algorithm tone maps pixels based on multiple scale local histograms, where small scales are used to preserve local contrast and large scales allow to maintain the global brightness consistency. A database of WDR images of humans depicted in high-contrast light conditions was created to validate and compare the performance of various algorithms for face detection in tasks such as biometric based identification. Our experimental results show that the proposed MS-Hist algorithm preserves image detail, brightness and high local contrast, and can benefit tasks such as face detection in WDR images.

1. Introduction

In recent years, wide dynamic range (WDR) imaging is highly demanded in multiple applications including security monitoring, photography and consumer electronics. Dynamic range (DR) in imaging is defined as the ratio between the brightest pixel value and the darkest one in a scene. A higher dynamic range maintains more visible details and texture in the image. The dynamic range of a natural scene can reach up to 120 dB which exceeds the dynamic range of almost all modern image sensors. Conventionally, WDR images are made by combining multiple low dynamic range (LDR) images taken with different exposure levels [1,2]. However, the DR of traditional display devices such as LCD, CRT and LED are usually limited to 8 bits per pixel, thus making it difficult to reproduce the WDR image on these displays directly. In order to close the gap between the WDR images and the LDR display devices, tone mapping algorithms were developed. Their purpose is to transform the luminance of a WDR image to the luminance of the LDR range, while maintaining decent brightness and contrast. Image processing of tone mapped images produce better results than the processing of the original ones, thus making the WDR imaging important for security monitoring in safety critical places such as airports, tunnels and plazas where light intensity may change drastically and usually exhibits WDR. Fig.1 (a)-(e) show 5 images taken under

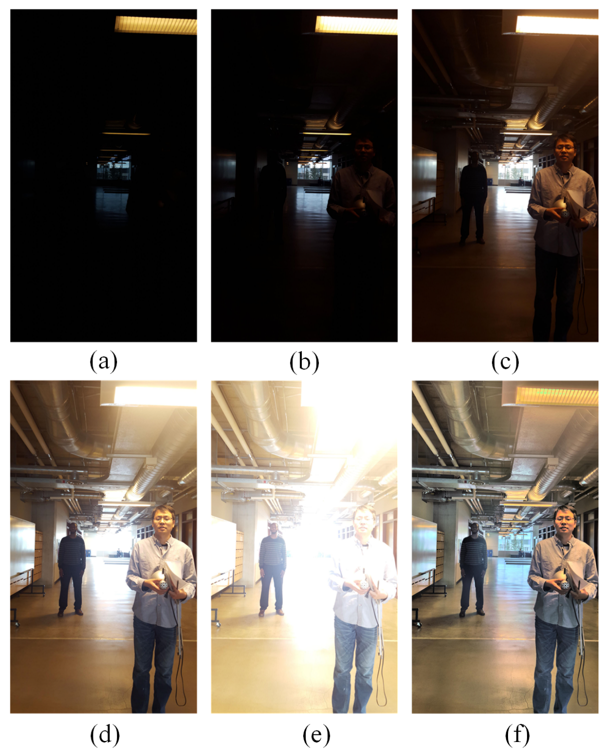

Figure 1. (a) - (e) Image taken under different exposure times. From (a) to (e), the exposure time is 1/1666 second, 1/310 second, 1/57 second, 1/10 second and 1/2 second, respectively. (f) Tone mapped image using the currently proposed MS-Hist algorithm.

978-1-5386-2939-0/17/$31.00 ©2017 IEEE

different exposure levels. However, none of the five images alone is good enough to view all places clearly. Fig.1 (f) shows a tone mapped image using the corresponding WDR file with all parts of the image seen clearly.

Many algorithms for tone mapping of WDR images have been developed so far. A simple way to achieve tone mapping is to apply a single global function to all pixels in the image such that identical pixels are given an identical value that does not change within the range of the display device. Tumblin and Rushmeier [3] developed the global operators for tone mapping that achieve dynamic range compression through matching the perceived brightness of the displayed image with that of the scene. Ward [4] developed a tone mapping method that aims to match the perceived contrast between the displayed image and the scene. Drago *et al*. [5] proposed an adaptive logarithmic mapping method that changes the base of the logarithmic function based on the brightness. Recently, Horé *et al*. [6] proposed a mapping algorithm that takes into account the images local statistics. Other similar algorithms are described in [7–10]. However, most of these algorithms suffer from low brightness of the tone mapped images, low contrast and loss of details due to the global compression of the dynamic range. Inspired by some features of the human visual system, some tone mapping algorithms try to mimic the dynamic range compression process of our photoreceptors [11–13]. Although these approaches may be effective in reducing the dynamic range, the toned images depict the internal representation rather than the luminance which is normally expected by human eyes. Some researchers consider the WDR compression to be a constrained optimization problem. Mantiuk *et al*. [14] considered the tone mapping as a minimum visible distortion problem. Ma *et al*. [15] proposed a tone mapping method by optimizing the tone mapped image quality index. However, solving these optimization problems may be computationally expensive and difficult to implement in real-time. In recent years, various algorithms emerged based on the Retinex theory [16–18]. Edge preserving filters [19–23] were used to separate the WDR image into the illuminance and reflectance images. The luminance image is regarded as a base layer whose information is believed to be less important for our visual system, hence, its dynamic range can be greatly compressed. Meanwhile, the reflectance image is treated as a detail layer which information is mostly preserved during tone mapping. The tone mapped images using the edge-preserving filters are visually appealing, and achieves state-of-the-art performance [21].

In this paper, we present a novel multi-scale tone mapping algorithm for displaying WDR images. We also show that it enables better detection of objects such as human faces. A dedicated WDR face database was created for testing the feasibility of the processed images for a biometric application, specifically, for face detection. This paper is organized as follows. In Section 2, we introduce the motivation and details of the proposed MS-Hist algorithm. The database, experimental results and comparison with other algorithms are reported in Section 3. Section 4 concludes the paper.

2. The proposed MS-Hist algorithm

The human visual system uses local adaptation to cope with large dynamic range in real world scenarios. Local adaptation is an ability to accommodate to the level of a certain visual field around the current fixation point. Experiments carried out by Stevens and Stevens [24] proved that the human visual system had different responses to the same luminance levels shown using different backgrounds. For example, higher luminance points can be perceived darker than lower ones when they are located in different backgrounds with uniform luminance. Their results also showing that overlapping reactions in the human visual system can extend our visual response range to cope with high contrast scenes in the real world. Web and Fechner Law revealed another important feature of the human visual system that the perception had a logarithmic relationship with the light intensity.

Motivated by the mentioned theories, we carry out the tone mapping in a local adaptation fashion and the calculation is done in the logarithmic domain. A direct approach to adapt to local areas is histogram equalization. It takes pixel distribution into account and was used for tone mapping in [6, 25, 26]. In our approach, we use histograms in the logarithmic domain. As shown in Fig. 2(a), any pixel is regarded as the center of a local adaptation window w. The histogram of w is shown in Fig. 2(b) where l is the logarithmic luminance of the pixels within w, u is the luminance of the displayed value, and $level$ is a user defined parameter which defines the number of bins in the histogram. If l_{max} and l_{min} are the maximum and minimum value within window w, then between any two adjacent bins b_k and b_{k+1}, there is a constant width in log-luminance value equal to $(l_{max} - l_{min})/level$. For the k-th bin, we define a linear function using the following equation:

$$u = a_k l + b_k \tag{1}$$

a_k and b_k are found as follows:

$$\begin{aligned} a_k &= \frac{u_k - u_{k-1}}{l^k_{max} - l^k_{min}} \\ b_k &= u_k - a_k \times l^k_{max} \end{aligned} \tag{2}$$

where

$$u_k = \frac{\sum_{i=1}^{k} p_i}{\sum_{i=1}^{n} p_i} \times 255 \tag{3}$$

l^k_{max} and l^k_{min} is the maximum and minimum value in the k-th bin and p_i is the number of pixels that fall into the i-th bin

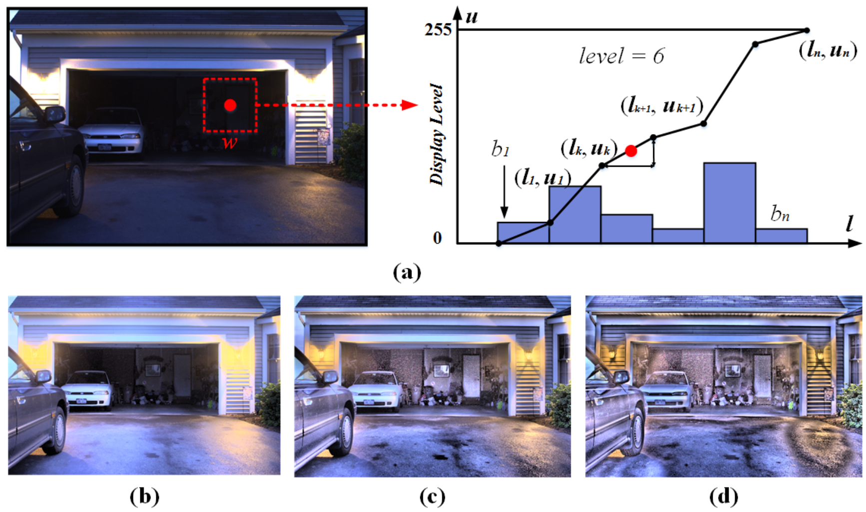

Figure 2. (a). Tone mapping function based on histogram in logarithmic domain. (b) Toned image using block size equal to image height and width. (c) Toned image using block size equal to half image height and width. (d) Toned image using block size equal to quarter image height and width.

of the histogram. The above equations ensure that the tone mapping in window w is monotonically increasing. For every pixel in a WDR image, we can always find a window w and use Eq. 1 to find a corresponding value between 0 and 255. The local histogram and the mapping function change with the scale of the window size w, hence, the toned value for the centre pixel changes accordingly. Fig. 2 (b)-(d) show the toned images for window of different sizes. It can be seen that smaller window scale reveals more detail of the WDR image and makes the image brighter, but there are some artifacts in uniform areas, such as the ground in Fig. 2(d). This is because in uniform areas, pixels are more likely falling into the same bin of the histogram, thus, a_k in Eq. 1 becomes a very large number, making small fluctuations in uniform area greatly exaggerated. Meanwhile, as shown in Fig. 2 (b), the large scale maintains more global brightness consistency and has no visible artifacts.

To have a good tone mapped image, it is important to maintain the brightness consistency while preserving the details. To achieve this goal, we intend to tone pixels in texture areas with smaller scales, while tone mapping pixels in uniform areas with larger scales. In order to achieve this goal, we need to detect uniform and texture areas in the WDR images first. There are a number of statistical metrics such as entropy or measures of dispersion that can be used for detecting uniform areas. In our case, we use the variance σ_w computed for block w, and calculate a_w value to measure the texture level of window w:

$$a_w = \frac{\sigma_w}{\sigma_w + \epsilon} \tag{4}$$

If the variance is much larger than ϵ, then a_w value is close to 1, which represents that the centre pixel is in a texture area. If the variance is much less than ϵ, the centre pixel is regarded as being in a uniform area. The block size w can be adjusted in order to detect uniform and texture areas at different scales. The variance value can be computed by:

$$\sigma_w = \frac{\sum_w I^2 - (\sum_w I)^2}{|w|^2} \tag{5}$$

$|w|^2$ is the number of pixels in window w. We use Eq.1 and Eq.5 to get the pixel value u_{w_i} and the corresponding a_{w_i} under different scales w_i. At each pixel location, we fuse the values using the following equation:

$$u = \frac{\sum_{i=1}^{n-1} a_{w_i}^i u_{w_i}}{\sum_{i=1}^{n-1} a_{w_i}^i} \tag{6}$$

In Eq. 6, n is the number of scales that are used for fusion. A larger i indicates smaller scale, therefore the weighting for small scales attenuate exponentially. In a texture area, a_{w_i} is close to 1, $a_{w_i}^i$ is also close to 1, which preserves more detail from small scales. In a uniform area, a_{w_i} is close to 0, $a_{w_i}^i$ will attenuate to 0 rapidly, which preserves more values from large scales. Our fusion function tends to give more weight to the large scales to make sure that there are no artifacts in the tone mapped images.

3. Experimental results and discussion

3.1. Customized App and WDR faces database

We developed an Android and iOS application called CaptureWDR to capture bracketed LDR images with different exposures. The user can define various exposure time

Figure 3. Original LDR image and results of applying different algorithms on the database. (a) Original image. (b) MS-Hist (currently proposed). (c) Result of Durand *et al*. [19]. (d) Result of Gu *et al*. [21]. (e) Result of applying a Gammar function. (f) Original image. (g) Result of applying Drago *et al*. [5] (h) Result of Fattal *et al*. [27]. (i) Result of Reinhard *et al*. [11]. (j) Result of applying a logarithmic function.

Algorithm used	Algorithm used
MS-Hist (Currently proposed)	Fattal *et al*. [27]
Durand *et al*. [19]	Reinhard *et al*. [11]
Gu *et al*. [21]	Logarithmic function
Gamma function	Wards *et al*. [4]
Drago *et al*. [5]	Exponential function

Table 1. Algorithms used on the database

and the number of bracketed images. The application automatically takes images with different exposure time and save the resulting images on the phone memory. A new database of WDR images containing faces was created by the captured LDR images. In total, 28 images of 2988 $\times$ 5312 pixels were collected under different lighting conditions including sunny outdoor, dark indoor and shaded areas. WDR images cannot be displayed on the monitor or used for face detection, therefore, we apply different tone mapping algorithms to convert the WDR images to LDR images. We have chosen 10 tone mapping methods ranging from basic to state-of-the-art ones (see Table 1). Using these tone mapping methods on the 28 WDR images leads to 10 sets of 28 tone mapped images. Some examples are shown in Fig. 3. The leftmost column shows the original WDR images. The tone mapped one follow in the remaining columns.

3.2. Evaluation of MS-Hist for face detection

To evaluate the impact of different tone mapping algorithms on an application-specific process such as face detection, we use the well-known Viola-Jones face detection algorithm [28]. We used its implementation available in OpenCV library. The 280 images of our unique database were used as input. If a face is correctly detected in any image, the corresponding tone mapping method gets one score. The higher the score, the better rating of the tone mapping algorithm. Fig. 4 shows an example of detection. The brightness and contrast of the face image varies greatly among different algorithms. In Fig. 4 (a), (c) and (d), the face area is too dark and blurry, therefore, OpenCV fails to detect the face. In Fig. 4 (b), (e) and (f), where faces are brighter and with more contrast, they are correctly detected. Performance of various algorithms is depicted in Fig. 5. Given the 28 WDR images, 20 faces were detected in tone mapped images using Durand [19] and the proposed MS-Hist algorithm. Tone mapping using Gu's approach [21] detected 19 faces. Other methods fail to detect more than 15 faces. We use overall brightness, sharpness, local stan-

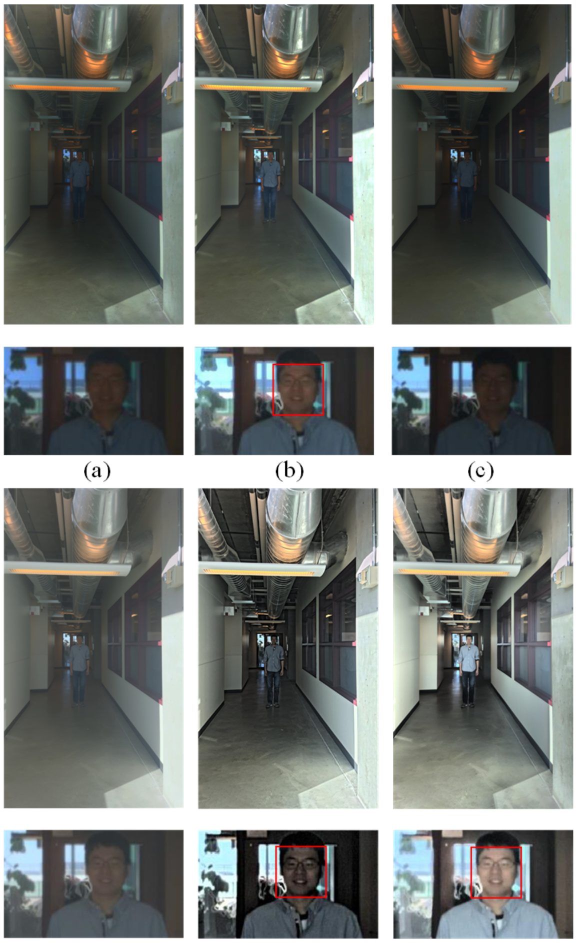
(a) (b) (c)
(d) (e) (f)

Figure 4. Results of various tone mapping using the newly collected WDR face database and the following algorithms: (a) Fattal *et al*. [27], (b) Durand *et al*. [19], (c) Drago *et al*. [5], (d) Reinhard *et al*. [11]. (e) Gu *et al*. [21]. (f) MS-Hist (currently proposed).

dard deviation value of the images to compare the images quality. The sharpness value s and the brightness value b are computed as follows:

$$s = \frac{1}{N}\sum |\nabla I|,\ b = \frac{1}{N}\sum I \tag{7}$$

where N is the number of pixels in image I. To further compare the three algorithms that show the best performance in face detection, the average s, b and local standard deviation value of the 28 LDR images are calculated and shown in Table 2. The higher the s and b values are, the brighter and sharper the image is. The standard deviation value measures the image texture level. The currently proposed MS-Hist tone mapping approach achieves the highest value in terms of overall brightness, sharpness and local standard deviation.

Algorithm	Brightness	Sharpness	Standard deviation
Durand *et al*. [19]	123.0681	10.8610	21.3761
Gu *et al*. [21]	127.2465	13.5740	31.1089
MS-Hist (Currently proposed)	**130.8593**	**14.8941**	**34.7047**

Table 2. Average value of brightness, sharpness and standard deviation of the best three tone mapping algorithms.

4. Conclusions

In this paper, we proposed a tone mapping algorithm based on the multiple scale histogram and fusion for display of WDR images. A database of WDR images was created to validate and compare the effect of various tone mapping algorithms on face detection. Our experimental results show that the proposed MS-Hist algorithm can produce good brightness and high local contrast images that can also benefit detection of objects such as faces in WDR images.

References

[1] P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” in *ACM SIGGRAPH 2008 classes*, p. 31, ACM, 2008.

[2] A. Spivak, A. Belenky, A. Fish, and O. Yadid-Pecht, “Wide-dynamic-range cmos image sensorscomparative performance analysis,” *IEEE transactions on electron devices*, vol. 56, no. 11, pp. 2446–2461, 2009.

[3] J. Tumblin and H. Rushmeier, “Tone reproduction for realistic images,” *IEEE Computer graphics and Applications*, vol. 13, no. 6, pp. 42–48, 1993.

[4] G. Ward, “A contrast-based scalefactor for luminance display,” *Graphics gems IV*, pp. 415–421, 1994.

[5] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, “Adaptive logarithmic mapping for displaying high contrast scenes,” in *Computer Graphics Forum*, no. 3, pp. 419–426, Wiley Online Library, 2003.

[6] A. Hore and O. Yadid-Pecht, “A statistical derivation of an automatic tone mapping algorithm for wide dynamic range display,” in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 2475–2479, IEEE, 2014.

[7] A. Horé and O. Yadid-Pecht, “A new filter for reducing halo artifacts in tone mapped images,” in *Pattern Recognition (ICPR), 2014 22nd International Conference on*, pp. 889–894, IEEE, 2014.

[8] G. W. Larson, H. Rushmeier, and C. Piatko, “A visibility matching tone reproduction operator for high dynamic range scenes,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 3, no. 4, pp. 291–306, 1997.

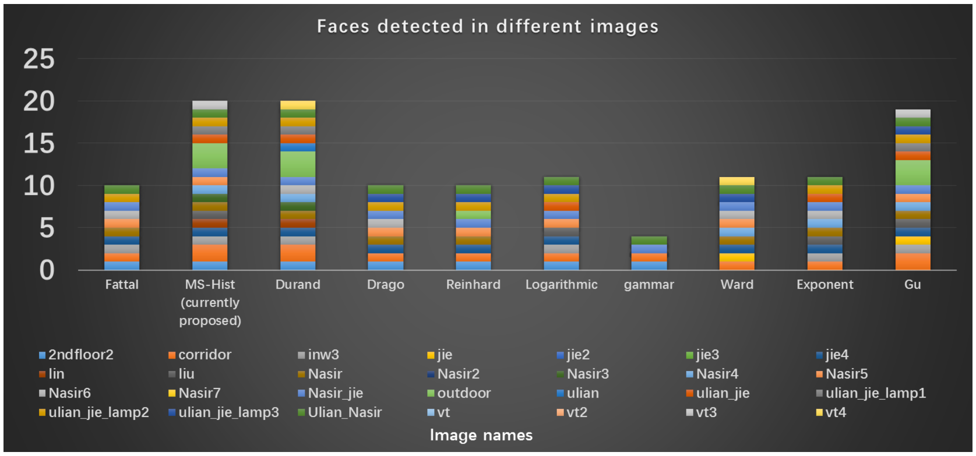

Figure 5. Number of faces detected on the images after aplying various tone mapping algorithms.

[9] G. Qiu and J. Duan, "An optimal tone reproduction curve operator for the display of high dynamic range images," in *2005 IEEE International Symposium on Circuits and Systems*, pp. 6276–6279, IEEE, 2005.

[10] C.-Y. Tsai, "A fast dynamic range compression with local contrast preservation algorithm and its application to real-time video enhancement," *IEEE transactions on multimedia*, vol. 14, no. 4, pp. 1140–1152, 2012.

[11] E. Reinhard and K. Devlin, "Dynamic range reduction inspired by photoreceptor physiology," *IEEE Transactions on Visualization and Computer Graphics*, vol. 11, no. 1, pp. 13–24, 2005.

[12] J. Van Hateren, "Encoding of high dynamic range video with a model of human cones," *ACM Transactions on Graphics (TOG)*, vol. 25, no. 4, pp. 1380–1399, 2006.

[13] H. Spitzer, Y. Karasik, and S. Einav, "Biological gain control for high dynamic range compression," in *Color and Imaging Conference*, vol. 2003, pp. 42–50, Society for Imaging Science and Technology, 2003.

[14] R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," in *ACM Transactions on Graphics (TOG)*, vol. 27, p. 68, ACM, 2008.

[15] K. Ma, H. Yeganeh, K. Zeng, and Z. Wang, "High dynamic range image tone mapping by optimizing tone mapped image quality index," in *2014 IEEE International Conference on Multimedia and Expo (ICME)*, pp. 1–6, IEEE, 2014.

[16] E. H. Land and J. J. McCann, "Lightness and retinex theory," *JOSA*, vol. 61, no. 1, pp. 1–11, 1971.

[17] L. Meylan and S. Susstrunk, "High dynamic range image rendering with a retinex-based adaptive filter," *IEEE Transactions on image processing*, vol. 15, no. 9, pp. 2820–2830, 2006.

[18] M. Herscovitz and O. Yadid-Pecht, "A modified multi scale retinex algorithm with an improved global impressionof brightness for wide dynamic range pictures," *Machine Vision and Applications*, vol. 15, no. 4, pp. 220–228, 2004.

[19] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *ACM transactions on graphics (TOG)*, vol. 21, pp. 257–266, ACM, 2002.

[20] K. He, J. Sun, and X. Tang, "Guided image filtering," in *European conference on computer vision*, pp. 1–14, Springer, 2010.

[21] B. Gu, W. Li, M. Zhu, and M. Wang, "Local edge-preserving multiscale decomposition for high dynamic range image tone mapping," *IEEE Transactions on image Processing*, vol. 22, no. 1, pp. 70–79, 2013.

[22] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," in *ACM Transactions on Graphics (TOG)*, vol. 27, p. 67, ACM, 2008.

[23] S. Paris, S. W. Hasinoff, and J. Kautz, "Local laplacian filters: edge-aware image processing with a laplacian pyramid," *Communications of the ACM*, vol. 58, no. 3, pp. 81–91, 2015.

[24] J. Stevens and S. S. Stevens, "Brightness function: Effects of adaptation," *JOSA*, vol. 53, no. 3, pp. 375–385, 1963.

[25] J. Duan, M. Bressan, C. Dance, and G. Qiu, "Tone-mapping high dynamic range images by novel histogram adjustment," *Pattern Recognition*, vol. 43, no. 5, pp. 1847–1862, 2010.

[26] G. Eilertsen, R. K. Mantiuk, and J. Unger, "Real-time noise-aware tone mapping.," *ACM Trans. Graph.*, vol. 34, no. 6, p. 198, 2015.

[27] R. Fattal, D. Lischinski, and M. Werman, "Gradient domain high dynamic range compression," in *ACM Transactions on Graphics (TOG)*, vol. 21, pp. 249–256, ACM, 2002.

[28] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004.