https://doi.org/10.2352/issn.2169-2629.2020.28.14

A Tone Mapping Model Based on Receptive Field for HDR Images

Imran Mehmood, Muhammad Usman Khan, Muhammad Farhan Mughal and Ming Ronnier Luo; State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou, China.

Abstract

High dynamic range (HDR) imaging has greater contrast reproduction capability than standard imaging techniques. It can achieve natural and pleasing appearance in terms of image quality. A tone mapping model (TMOz) is developed based on the center-surround properties of the mammalian ganglion cells of the human visual system for feature enhancement. The contrast of the HDR image is mapped adaptively to an SDR display range using a global method followed by contrast enhancement in local regions. A psychophysical experiment was conducted to refine the model for adaptivity of the contrast mapping function. Finally, the performance of the TMOz was evaluated using CIELAB (2:1) formula together with high quality reference images. The results showed that TMOz outperformed the other tone mapping operators (TMOs).

Introduction

The TMOs are methods that compress the luminance of HDR images to match the luminance of the display devices so that most of the information in the image is perceivable [1]. Human vision concentrates more on the image characteristics to be more appealing such as contrast and colorfulness. TMOs can be classified into two categories i.e., global and local operators. These algorithms are classified into major categories i.e., global and local operators. The global operators apply the same compression method to each of the image pixel without considering the neighborhood pixels. In contrast, the local operators consider the neighborhood pixels and other spatial features to compress the luminance of the image pixels.

The TMOs are based on photographic reproduction methods, human vision theory and other image processing methods. Schlick proposed a global method that was based on rational quantization technique [2]. The method preserves the contrast in shadows while compressing the luminance in highlights. It has three parameters that are to be adjusted manually to achieve better contrast in images. Reinhard *et al.* proposed a tone mapping method that was inspired by the photographic development methodology and printing process [3]. Its global version compress the luminance to the displayable range while the local version uses "dodging and burning" technique that improves the contrast in the local regions. The automatic version enables the method to work adaptively, freeing the user from setting parameters [4]. Drago *et al.* proposed an adaptive method based on logarithmic mapping for high contrast images [5]. It has one bias parameter, which can be varied depending upon the HDR image. Reinhard and Devlin proposed global and local versions of dynamic range reduction inspired by photoreceptor physiology [6]. The method works well, however it has three intuitive user parameters to adjust brightness, contrast and adaptation level. This study aims to develop an adaptive tone mapping model that frees the user from parameters and produces high quality images in terms contrast, details and colourfulness.

In the human visual system (HVS), retina behaves opposite to the artificial visual system i.e. cameras. Before passing the signal to the cortex, a series of actions is performed through the optic nerve. The optic nerve is composed of axons retinal ganglion cell (RGC) [7, 8]. The RGCs collectively transmit non-image forming and image forming visual signals from the retina in the form of an action potential or firing rate to several parts of the brain. Each ganglion cell bears a receptive field, which increases with the intensity of the light. For the largest receptive field, the light has to be more intense at the center.

Each receptive field is organized into a central disk "center" and the concentering ring called "surround". The center and surround behave oppositely to each other. For instance, the light in the surround might decrease the firing rate of RGCs and the light on the center would increase the firing rate. Therefore, there are two categories of the receptive field; on-center off-surround and off-center on-surround.

This work utilizes the properties of the on-center off-surround phenomenon of the RGCs for enhancing the details of the HDR image. As a feature enhancement algorithm, it is used to enhance fine details and visibility of the edges in the HDR images. The algorithm removes high frequency details that often include random noise in the images. Hence, it is capable of dealing with HDR images to have a high degree of noise. Furthermore, the HDR image contrast is mapped to standard display devices using logarithmic mapping and the bias parameter is automated for adaptivity of the model. Contrast is further improved in local regions using adaptive histogram equalization technique to match the probability density function (PDF) of uniform distribution.

The goal of this study is to develop a new tone mapping operator, TMOz, applying the testing method already established [9]. Liu *et al.,* [10] has explored colour difference evaluation using CMC [11], CIE94 [12], CIEDE2000 [13] and CIELAB [14-15], and determined that colour difference in images could be predicted with the Commission Internationale de I'Eclarage (CIE) colour difference formulae for image difference using CIELAB (K_L=2) and CIEDE2000. For simplicity, CIELAB (K_L=2) i.e., CIELAB (2:1) was adapted in this paper for evaluation of the TMO.

Previous Study

In our previous study [9], a methodology for the evaluation of the quality of TMOs was proposed. A set of reference images was obtained via a psychophysical experiment. In the experiment, a set of HDR images from the RIT image database [16] was used to render 3 attributes, contrast, sharpness and colourfulness. Each attribute was rendered with five levels. The observers rated each image either high or low quality. The raw data were analyzed and the images with the highest Z-scores were considered as reference images. The reference images were then used to evaluate five commonly used TMOs in a psychophysical experiment. Two colour difference metrics, CIELAB(2:1) [17] and S-CIELAB [18] were used. It was found that both metrics agreed well with the psychophysical data with a coefficient of determination (R^2)

0.84 and 0.85, respectively. However, it was found that the commonly used TMOs did not give accurate prediction according to the testing method described above.

Model Development

Inspired by the many existing TMOs, a new one called TMOz is developed. It is composed of various functions to achieve the desired level of details in the highlight and shadow areas with high contrast and colourfulness in the images. Figure 1 shows the workflow of the model including 6 stages. The input to the model is an HDR radiance map which was obtained from the multiple exposures of a scene [19]. The RGB radiance map was first transformed to the input CIE-$x_i y_i Y_i$ via forward GOG model [20], as the characterization model. Subsequently, only the input luminance channel Y_i was processed in Stages 1 and 2 while the input chromaticity coordinates $x_i y_i$ was preserved. It was further transformed to CIE-L_i*C_i*h space and contrast was then enhanced in the lightness channel L_i. At Stage 4, the Chroma channel CIELAB $C *_{ab,i}$ was enhanced by a scaling factor to make colours to appear more colourful. It is transformed back to new $x_n y_n Y_n$ coordinates keeping the range of Y_n to [0,1]. Finally, reverse GOG model was applied at stage 5 to transform $x_n y_n Y_n$ to display RGB values. Each stage in the workflow is described in the following sections, except stages 0 and 5, the forward and reverse GOG model.

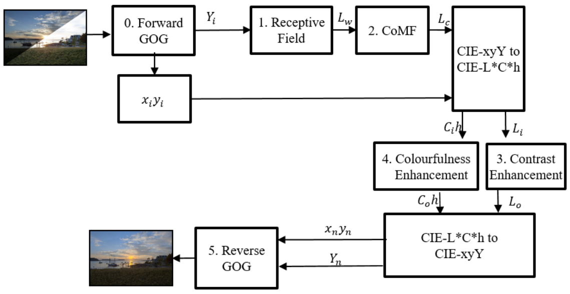

Figure 1. *Block diagram of the tone mapping model*

Stage1: The Baseline Receptive Field

The classical mathematical model for mimicking cat retinal ganglion cell response was analyzed by Rodieck [21]. He modeled the centre-surround mechanism in the form of the difference of Gaussians (DOG). The advantage for modeling centre-surround paradigm using DOG is to diminish the high frequency details that include random noise. A two dimensional DOG operator is given in equation (1).

$$\mathrm{DOG} = \frac{1}{\sqrt{2\pi}}\left(\frac{1}{\sigma_1}\mathrm{e}^{\frac{-(x^2+y^2)}{2\sigma_1^2}} - \frac{1}{\sigma_2}\mathrm{e}^{\frac{-(x^2+y^2)}{2\sigma_2^2}}\right) \quad (1)$$

where σ_1 and σ_2 represent the standard deviations of the excitory and inhibotry fields, respectively. The x and y are the Gaussins distribution variables. The DOG of an image I is calculated by subtracting a convolved version of the luminance Y_i of the input image I of Gaussians kernel with a standard deviation σ_2^2 from that of the Gaussian kernel with standard deviation of σ_1^2 as given in equation (2).

$$I_{\mathrm{DOG}} = Y_i * \left(\frac{A}{\sqrt{2\pi}}\left(\frac{1}{\sigma_1}\mathrm{e}^{\frac{-(x^2+y^2)}{2\sigma_1^2}} - \frac{1}{\sigma_2}\mathrm{e}^{\frac{-(x^2+y^2)}{2\sigma_2^2}}\right)\right) \quad (2)$$

where A is a constant used to scale the features.

Stage 2: Contrast Mapping Function (CoMF)

The overall contrast should preserve both the details and contrast of HDR scene with luminance compression because contrast enhancement will be done later. The CoMF was inspired by Perlin and Hoofer's [22] bias power function as given in equation (3).

$$\mathrm{bias}_b(t) = t^{\log(b)/\log(0.5)} \quad (3)$$

where bias_b is a power function which maps the HDR luminance to the range of the display device when integrated with logarithmic tone mapping function as given in equations (4) and (5).

$$L_c = \frac{\log_{10}(1+L_\mathrm{w})}{\log_{10}(1+L_{\mathrm{w,max}})\log_{10}(L_r)} \quad (4)$$

$$L_r = 2 + 8\left(\frac{L_\mathrm{w}}{L_{\mathrm{w,max}}}\right)^{\log(b)/\log(0.5)} \quad (5)$$

where L_w is the logarithmic of the I_{DOG}, $L_{\mathrm{w,max}}$ is the maximum luminance of L_w, and L_c is the contrast mapping function which maps the HDR luminance in the range of [0,1]. The parameter b is the bias parameter over the range [0,1]. It changes the shape of the CoMF, thus, the contrast of the image. For brighter images, the value of b is larger and for dark images, the value is smaller. To make the L_c adaptive, a model is developed by minimizing CIELAB (2:1) colour difference [17], proposed by CIE for evaluating colour difference for images. The model was verified later using a psychophysical experiment, described later. The ten images from RIT database used in our earlier work [9, 10] were used, which were used to develop the reference images in our earlier study. The parameter b was varied manually to achieve minimum CIELAB (2:1) difference with the reference images. Afterward, the regression method was applied to develop a model for b in terms of a key of the image k, as given in equation (6).

$$b = 0.33k - 0.32 \quad (6)$$

where k is given by

$$k = 0.5 \times 4.5^{\frac{2\log_2 G_L - C_L}{C_L}} \quad (7)$$

The coefficients 0.5 and 4.5 were optimized empirically to achieve the best performance of the model [4]. G_L and C_L are the image statistics referred as geometric mean and contrast ratio, respectively [23], as given in equations (8) and (9).

$$C_L = log_2\, L_{w,max} - log_2\, L_{w,min} \quad (8)$$

$$G_L = exp\left(\frac{1}{N}\sum ln(\delta + L_w)\right) \quad (9)$$

where, $L_{w,\mathrm{min}}$ is the minimum luminance, N is the total number of pixels in the image and δ is very small number e.g., $\delta = 10^{-6}$, to avoid undefined values for the zero luminance levels.

Stage 3: Contrast Enhancement

Adaptive histogram equalization based on the clip limit was applied to enhance the contrast of the tone mapped images. It was used to increase the local contrast by dividing the image into various tiles. In this model, 8 × 8 tiles were used. The image was first transformed to CIE- L_i*C_i*h space. Pizer's model [24] was adopted, to match the PDF of uniform distribution given by equation (10).

$$L_o = [L_{\mathrm{imax}} - L_{\mathrm{imin}}] * P(f) + L_{\mathrm{imin}} \quad (10)$$

where L_{imax} and L_{imin} are maximum and minimum values of the transformed lightness image L_i, respectively. $P(f)$ is the cumulative probability distribution function of the histogram of the L_i. L_o is the output contrast enhanced lightness.

Stage 4: Colourfulness Enhancement

Due to the compression of the luminance to low dynamic range, the colourfulnes degrades. It was described earlier [9] that the reference images had higher colourfulness compared to the original images. Moreover, a suitable increase on colourfulness could result in pleasing effect to the images. Therefore, by minimize CIELAB (2:1) with the reference images, the colourfuness of the tone mapped images was enhanced. This was achieved by preserving the hue channel and increasing the Chroma channel by 15% i.e. $C_o = 1.15C_i$. The enhancement of the colourfulness is dependent on the specific display colour coordinates and the maximum increase in colorfulness is kept within the maximum allowed values of that display.

Subsequently, back transformation is applied from CIE-L_i*C_i*h to the XYZ coordinates. Finally, the GOG [20] display model is applied to transform the SDR image to the specific display image, transforming the SDR to the display dynamic range.

Experiment to Verify Bias Parameter

The goal of the experiment was to verify the bias parameter model given in equation (6), which was developed by minimizing CIELAB (2:1) with the references images, had best visual preference as well. A psychophysical experiment was conducted by assessing the image quality using the pair comparison method. Fifteen images from RIT database [16] were used, including the ten images previously used to develop the reference images [9]. Each image was rendered at five levels of the bias parameter. One image was generated with the actual value calculated using equation (6) named as original image. Two images were generated manually by varying the bias parameter prior to and following the actual value and two levels in the neighborhood of the latter two types. Each value of the bias parameter was selected such that, just noticeable difference (JND) was seen between the two rendered images. Each image was rendered with five levels, therefore the total number of the rendered images were 75. All the images were selected by the authors in the experimental conditions. So, there are a total of 150 pairs, 15(5x4)/2.

Figure 2 shows the setup for the experiment. A pair of rendered images were displayed side by side. The observers were asked to select the preferred image based on overall image quality, in contrast to the difference of the images from the reference images, to make sure that, the model had a minimum difference from the images as well as the best visual preference.

The images were displayed on a NEC PA302W AH-IPS LCD display located in a dark room. The wall reflectance of the darkroom was approximately 4%. The peak luminance of the peak white of the display was set at CIE D65 and 1931 standard colorimetric observer at a luminance of 287 cd/m2. For evaluation of spatial uniformity, the display was divided into 3 by 3 segments and the mean colour difference calculated between the centre and each segment was 1.21 ΔE*ab. As mentioned earlier, the GOG display model was used for display characterization. The 24 colours on the ColorChecker chart was used to check the performance of the model. The performance was 0.64 ΔE*ab with a range of 0.37 to 1.66. The processed images were transformed to the display RGB using this model.

The display was located at 45 cm away from the observer. Observers were tested for colour vision using Ishihara colour vision test. A greyscale image was displayed for 30 seconds before the actual experiment to adapt to the environment. Twenty subjects from two ethnic group i.e. Pakistani (10 males and 3 females) and Chinese (3 males and 4 females), participated in the experiment. In total, 3,300 judgements were made, i.e. 20 observers x 15 images × 10 pairs + 300 repeated images.

Results and Discussion

This section presents results of bias parameter model and performance of TMOz.

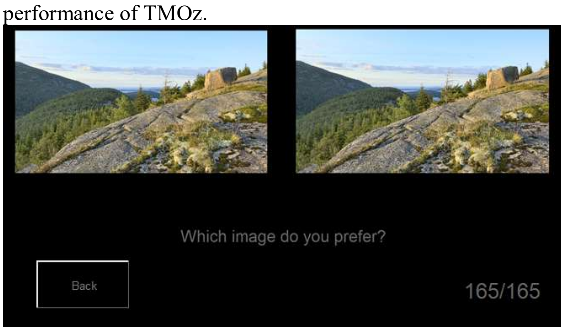

Figure 2. Setup for the verification of the model

To analyze the validity of the bias parameter, inter- and intra-observer variations were calculated using the coefficient of variation (CV) method [25]. For inter-observer variability, the standard deviation was divided by the mean of the two values for the repeated images. The same pair of images shown to different observers was used to calculated intra-observer variability. A lower value of CV implied better observer performance and lower value of intra-observer variation inferred better reliability of the experiment. The mean values inter-observer variability was 32 whereas the values for intra-observer variability were 23.

The experimental data was converted to Z-scores using Thurstone's law of comparative judgment [26]. Figure 3 shows a box and whisker plot for the Z-scores. The red circles in the figure showed the Z-scores of the original images i.e., the images generated using the bias parameter model given in equation (6). Higher values of Z-scores correspond to the more preferred images. The plot showed that in most cases, the original images had the highest value of Z-scores. This indicates that the original images were preferred over the other rendered images. However, in a few cases, instead of the original images, other rendered images were preferred. Figure 4 shows the correlation between experimental values and model values. The coefficient of determination R^2 was 0.86, which implied the psychophysical data agreed well with the bias parameter model performance.

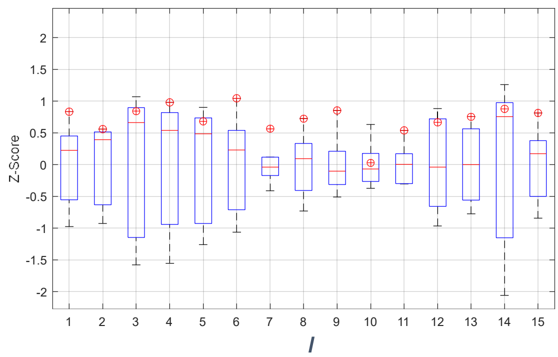

***Figure 3.** Box and whisker plot showing Z-scores for each Image I*

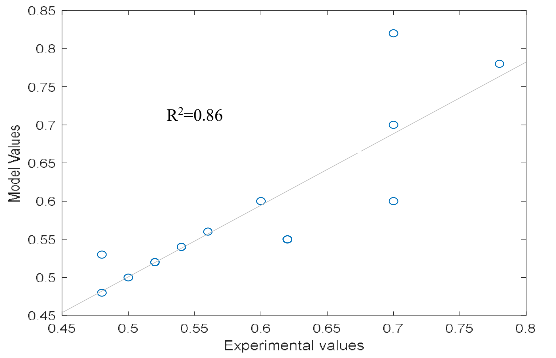

Figure 4. *The correlation between model values and experimental values*

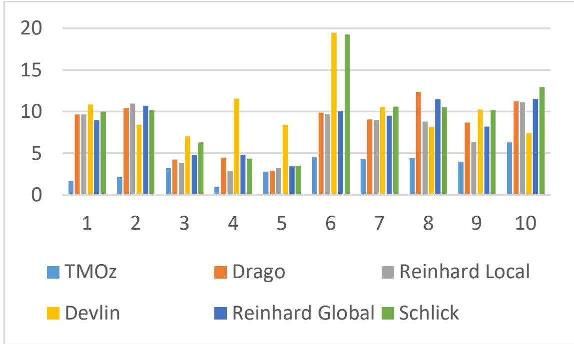

Figure 5. *Comparison of CIELAB(2:1) $\Delta E_{a,b}$ for 10 HDR Images*

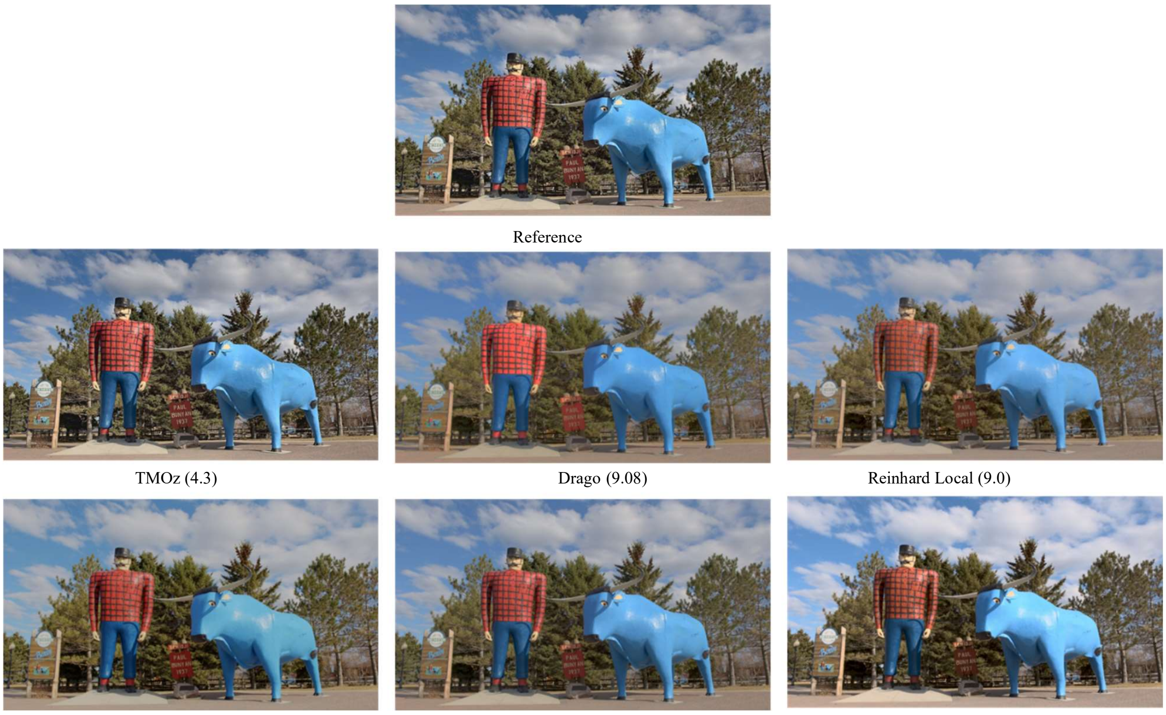

Figure 6. *Comparison of TMOz with reference image and other TMOs. The number in the parenthesis is the CIELAB(2:1) $\Delta E_{a,b}$*

The performance of the TMOz was analyzed using CIELAB (2:1) image quality metric and reference images. There are sufficient evidence to show that the reference images developed via psychophysical experiment were high quality images in terms of contrast, sharpness and colorfulness [9]. Moreover, the results showed that the colour difference metric CIELAB (2:1) $\Delta E_{a,b}$ agreed well with the psychophysical data.

To evaluate the performance of TMOz, the ten HDR images used in our previous study [9] were tone mapped and CIELAB (2:1) colour difference was calculated using the reference images. In CIELAB (2:1), the lightness channel was divided by 2 to see more effect of the a* and b* channels. Figure 5 compares TMOz with five TMOs i.e. Schlick [2], Reinhard's photographic tone reproduction global and local versions [3], Drago's adaptive logarithmic [5], Reinhard & Devlin's dynamic range reduction [6] and the reference image, in terms of CIELAB (2:1) colour differences. For all of the 10 images, TMOz had the lowest difference from the reference images (the most accurate model).

Figure 6 shows the comparison of the Image 7 tone mapped image using TMOz with five common TMOs and reference images. The CIELAB (2:1) $\Delta E_{a,b}$ is also given in parenthesis. The figure showed that the image tone mapped with TMOz is mapped with TMOz is high contrast and colourfulness compared to the other TMOs and it is close to the reference image.

Conclusions

A tone mapping model, TOMz, based on the receptive field of memory ganglion and logarithmic mapping was proposed. In order to enhance the details and contrast at the edges of HDR image, the center-surround operation was applied. The contrast was further compressed by a global adaptive logarithmic mapping function and adaptive histogram equalization method was used for local contrast adjustments. The bias parameter, which manipulates the shape of the contrast mapping function, was made adaptive using CEIALB (2:1) colour difference and was verified by the psychophysical experiment. The reference images were used to test the performance of the TMOz. The CEIALB (2:1) colour difference results showed that the proposed method had minimum difference with the reference images as compared to the other familiar TMOs. Moreover, the psychophysical experiment showed that, the images processed using bias model had best visual preference and quite good correlation with the model. It can be concluded that the proposed method produces high quality images in terms of contrast, details and colourfulness.

Acknowledgement

This work was supported by Zhejiang University, Two-High doctoral fellowship.

References

[1] Erik Reinhard, Wolfgang Heidrich, Paul Debevec, Sumanta Pattanaik, Greg Ward & Karol Myszkowski, High dynamic range imaging: acquisition, display, and image-based lighting (Morgan Kaufmann, 2010).

[2] Schlick, C., “Quantization Techniques for the Visualization of High Dynamic Range Pictures” P. Shirley, G. Sakas, and S. Müller (eds.), Photorealistic Rendering Techniques, New York: Springer-Verlag, 7 (1994).

[3] Reinhard, E., Stark, M., Shirley, P., & Ferwerda, J., “Photographic tone reproduction for digital images” ACM transactions on graphics (TOG), 21(3), 267 (2002).

[4] Reinhard, E., “Parameter estimation for photographic tone reproduction” Journal of graphics tools, 7(1), 45 (2002).

[5] Drago, F., Myszkowski, K., Annen, T., & Chiba, N., “Adaptive logarithmic mapping for displaying high contrast scenes” Computer graphics forum. 22(3), 419 (2003).

[6] Reinhard, E., & Devlin, K., “Dynamic range reduction inspired by photoreceptor physiology” IEEE Transactions on Visualization & Computer Graphics, 13. (2005).

[7] Grossberg Stephen & Simon Hong, A neural model of surface perception: Lightness, anchoring, and filling-in (Spatial Vision, 2006) pg. 263-322.

[8] Helga Kolb, How the retina works: Much of the construction of an image takes place in the retina itself through the use of specialized neural circuits (American scientist 2003) pg. 28-35.

[9] Mehmood, I., Khan, M. U., Luo, M. R., & Mughal, M. F. Tone Mapping Operators Evaluation Based on High Quality Reference Images, Color and Imaging Conference, Society for Imaging Science and Technology, vol. 2019, no. 1, pg. 268 (2019).

[10] Liu, H., M. Huang, G. Cui, M.R. Luo, and M. Melgosa, “Color-difference evaluation for digital images using a categorical judgment method” JOSA A, 30(4), 616-626 (2013).

[11] Clarke, F., R. McDonald, and B. Rigg, “Modification to the JPC79 colour–difference formula” Journal of the Society of Dyers and Colourists, 100(4), 128-132. (1984).

[12] McDonald, R. and K. Smith, “CIE94-a new colour-difference formula” Journal of the Society of Dyers and Colourists, 111(12), 376-379 (1995).

[13] Luo, M.R., G. Cui, and B. Rigg, The development of the CIE 2000 colour-difference formula: CIEDE2000. Color Research & Application: Endorsed by Inter-Society Color Council, The Colour Group (Great Britain), Canadian Society for Color, Color Science Association of Japan, Dutch Society for the Study of Color, The Swedish Colour Centre Foundation, Colour Society of Australia, Centre Français de la Couleur, 26(5), 340-350 (2001).

[14] CIE, I., Colorimetry—part 4: 1976 Lab colour space, Joint ISO/CIE Standard: ISO 11664-4:2008(E)/CIE S 014-4/ E:2007. Australia (2008).

[15] Luo, M.R. CIE Division 8: a servant for the imaging industry, Color Science and Imaging Technologies. International Society for Optics and Photonics (2002).

[16] Fairchild, M. D., The HDR photographic survey, Colour and imaging conference, Society for Imaging Science and Technology, vol. 2007, no. 1, pg. 233. (2007).

[17] CIE 199:2011 Methods for evaluating colour differences in images, Vienna, Austria.

[18] Zhang, X. and B.A. Wandell. A spatial extension of CIELAB for digital color image reproduction. SID international symposium digest of technical papers, Citeseer, (1996).

[19] Debevec, P. E., & Malik, J., Recovering high dynamic range radiance maps from photographs, Proceedings of the 24th annual conference on Computer graphics and interactive techniques, pg. 369. (2008).

[20] Berns, R. S., Methods for Characterizing CRT Displays, Displays 16(4), pg.173 (1996).

[21] Rodieck, R.W. and J. Stone, “Analysis of receptive fields of cat retinal ganglion cells” Journal of neurophysiology, 28(5), 833 (1965).

[22] Perlin, K. & E.M. Hoffert, Hypertexture, Proceedings of the 16th annual conference on Computer graphics and interactive techniques (1989).

[23] Francesco Banterle, Alessandro Artusi, Kurt Debattista & Alan Chalmers. Advanced high dynamic range imaging: theory and practice. (AK Peters 2011).

[24] Pizer, S. M., Amburn, E. P., Austin, J. D., Cromartie, R., Geselowitz, A., Greer, T. & Zuiderveld, K., “Adaptive histogram equalization and its variations”, Computer vision, graphics, and image processing, 39(3), 355 (1987).

[25] Reed, G.F., F. Lynn, and B.D. Meade, “Use of coefficient of variation in assessing variability of quantitative assays” Clin. Diagn. Lab. Immunol., 9(6), pg. 1235 (2002).

[26] Thurstone, L.L., “A law of comparative judgment” Psychological review, 34(4), pg. 273 (1927).

Author Biography

Imran Mehmood received his M.Sc. and M.Phil. degrees from Department of Electronics, Quaid-i-Azam University, Islamabad, Pakistan. He worked on wireless communication and computer vision. He is currently doing his Ph.D. in Optical Science and Engineering from Zhejiang University Hangzhou, China, under the supervision of Professor Ming Ronnier Luo. He is working on testing and developing generic tone mapping operators in HDR imaging for real time applications. His research interest includes HDR imaging, tone mapping, image quality and colour science etc.