

# Color Difference Matrix Index for Tone-mapped Images Quality Assessment

Cheng Guo and Xiuhua Jiang

Communication University of China, Beijing, China

**Abstract**—To visualize high dynamic range (HDR) images on common low dynamic range (LDR) displays, some tone-mapping operators (TMOs) are proposed in previous studies. TMOs convert specific HDR image to various LDR images, thus the quality assessment of these tone-mapped images or TMOs become a natural question. In this work, an objective quality index named color difference matrix index (CDMI) is proposed. It compares the CIEDE2000 color difference matrix between the HDR image and LDR images, thus the color information, which is ignored by many previous methods, is involved. CDMI is then compared and combined with other methods: TMQI and FSITM later in experiments. It is shown that CDMI have edges over other index in many aspects and have a satisfying result when combined with these two indices.

**Keywords**—high dynamic range images; tone-mapping operator; objective quality assessment; color difference

## I. INTRODUCTION

The dynamic range of an image is defined as the ratio of its maximum luminance to minimum. HDR images has been gaining increasing interest since they present well through human visual system (HVS) and produce better perceptual experience than LDR images. But current displays on the market are almost all in LDR, so tone-mapping operator are in need to convert dynamic range from high to low for the visualization of HDR images on LDR displays.

Unsurprisingly but unfortunately, each TMO performs differently on different HDR images. To be specific, such attributes like brightness, contrast, color, detail, and artifacts will influence perceptual quality of tone-mapped images [1]. That means we have to evaluate which TMO is better for each individual HDR image case. Obviously, subjective assessment for TMOs is the way that correspond HVS most, but the process recruiting participants and analyzing data is too time consuming to carry out.

### A. Related Works

The first study on objective quality assessment of tone-mapped images is the tone mapping quality index (TMQI) proposed in 2013 [2], TMQI use structural similarity index (SSIM) [3] to measure the structural fidelity of tone-mapped images, combined with a statistical naturalness measurement. But one shortcoming of TMQI is that only the luminance information is considered, and thus the color distortion is ignored.

The feature similarity index for tone-mapped images [4] (FSITM) was then proposed in 2015. It utilizes the images' local phase information and compares the origin HDR image's

locally-weighted mean phase angle map to that of tone-mapped image. But FSITM is calculated in separate RGB channel, and the author didn't provide a specific and reasonable way to fuse them.

TMQI was then combined with the visual saliency [5] in SHDR-TMQI in [6]. The author team also tried another two combinations: TMQI-NSS- $\sigma$  fuse TMQI with a natural scene statistics (NSS) model [7] based on mean subtracted contrast normalized (MSCN) pixels [8], and TMQI-NSS-Entropy use local entropy to improve the pooling process of structural fidelity score in TMQI. The latest revision of TMQI is TMQI-II [9] which improved both the structural fidelity and statistical naturalness parts.

In 2017, a novel method named TIQ [10] use a bag of features (BOF) which include up to eight features for assessment, due to the diversity of artifacts and distortions. These features range from structure fidelity, naturalness, to overall brightness. It's worth noting that some of these features are first extracted for tone-mapped image quality assessment.

Like common image quality assessment, these methods can be divided to full-reference (FR) and non-reference (NR) according to whether origin HDR image is required. An initial NR attempt called blind tone-mapped quality index (BTMQI) [11] was carried out in 2016, this method considers the information, naturalness, and structure of tone-mapped images.

The latest NR index for tone-mapped images is BLIQUE-TMI [12] proposed in 2017. This method utilizes sparse representation to calculate new features: sparse activity and primitive activity, to represent image's local features; and statistic modeling to calculate moment statistics and chromatic statistics to grasp global features.

### B. Purpose of CDMI

Almost none of the above method fully utilize color information of images or have a comprehensive evaluation of chromatic distortion. Since some TMOs consider chromatic information while tone-mapping HDR images, and color difference plays an important role in tone-mapped images' perception quality through HVS. It's necessary paying attention to color difference between HDR images and corresponding tone-mapped LDR images. That's the original intention of proposing CDMI.

## II. PROPOSED CDMI INDEX

The color difference between the origin HDR image and tone-mapped LDR image is an essential feature which have a great impact on image's perceptual quality. The author of the

classic TMQI method claim himself in paper [2] that one of their shortcoming is the ignorance of chromatic information. This time, inspired by color difference which is frequently used in printing industry, and paper [13] by author's senior schoolfellow, we introduce the concept of color difference to the field of image quality assessment, to form the color difference matrix for assessing the chromatic difference in tone-mapped images.

The color difference matrix is consisted of the CIEDE2000 color difference [14] value of each corresponding pixel of HDR and LDR image. Since this value is calculated in CIELab color space (A color space which is perceptually uniform in HVS color vision.), and both HDR and LDR images are in RGB color space, a transformation from RGB to CIELab is required first. Formulas for the transformation can be easily acquired via published materials and easily implemented in MATLAB.

The next step is extracting the  $L^*$ ,  $a^*$  and  $b^*$  (Luminance, green-red and blue-yellow color components, respectively.) of corresponding pixel  $(i, j)$  in both HDR and LDR image. Once we get  $Lab_{HDR}(i, j) = (L_{HDR}^*(i, j), a_{HDR}^*(i, j), b_{HDR}^*(i, j))$  and  $Lab_{LDR}(i, j) = (L_{LDR}^*(i, j), a_{LDR}^*(i, j), b_{LDR}^*(i, j))$ , the CIEDE2000 color difference of pixel  $(i, j)$ :  $\Delta E_{00}(i, j)$  can be calculated as follow:

$$\Delta E_{00} = \sqrt{\left(\frac{\Delta L'}{k_L S_L}\right)^2 + \left(\frac{\Delta C'}{k_C S_C}\right)^2 + \left(\frac{\Delta H'}{k_H S_H}\right)^2 + R_T \frac{\Delta C'}{k_C S_C} \frac{\Delta H'}{k_H S_H}} \quad (1)$$

where  $\Delta L'$ ,  $\Delta C'$ , and  $\Delta H'$  denote the luminance difference, chroma difference, and hue difference of two pixels in CIELab color space, respectively. Meanwhile,  $S_L$ ,  $S_C$ ,  $S_H$ , and  $R_T$  denote luminance compensation, chroma compensation, hue compensation, and hue rotation, respectively. Coefficients  $k_L$ ,  $k_C$ , and  $k_H$  are usually all set to default value 1. Details about the calculation of seven parameters mentioned above can be found in [14], and the MATLAB complementation of CIEDE2000 color difference formula is available in [15].

Color difference matrix is then generated after traversing calculating the CIEDE2000 color difference value of each pixel. This matrix consists of each  $\Delta E_{00}(i, j)$  value of corresponding pixel. Since we are not considering the structural and positional weight of color difference. (Task like finding which position or part of color difference that counts most is not included in this study.), we choose the arithmetic mean value of  $\Delta E_{00}(i, j)$  to measure the overall color difference:

$$\Delta E_{00} = \frac{\sum_{i=1}^r \sum_{j=1}^c \Delta E_{00}(i, j)}{r \times c} \quad (2)$$

where  $r$  and  $c$  denote the pixel weight and height of image.

Visualizing the color difference matrix is an optional step of CDMI. It's implemented by regarding color difference matrix as a grayscale image. Obviously, "brighter" parts mean bigger color difference between the corresponding parts in HDR and tone-mapped images. The visualization of color difference matrix will be demonstrated later in this paper.

According to CIEDE2000's definition, pixels with greater color difference to origin pixel in HDR image will have bigger  $\Delta E_{00}(i, j)$  value. Hence, images with more significant overall

color difference (Worse perceptual quality.) will get a bigger  $\Delta E_{00}$  value. But CDMI is designed to be normalized to [0, 1] from bad to good, like other indices. So, a mapping from the origin  $\Delta E_{00}$  value's range to [0, 1] is required. Here we choose liner mapping for this process:

$$CDMI = 1 - \Delta E_{00} / \max(\Delta E_{00}) \quad (3)$$

CDMI upper bounded by 1 is finally calculated after those four steps: converting color space, traversal computation, averaging, and liner mapping.

### III. EXPERIMENT RESULTS

To have a simple evaluation of the proposed method, we choose a tone-mapped LDR images database provided by Yeganeh in his study [2]. This database contains fifteen sets of images, there are eight tone-mapped images produced by different TMOs along with its origin HDR image in each set. TMOs used in this database are namely Reinhard02 [16], Drago03 [17], Durand02 [18], Mantiuk05 [19], Pattanaik00 [20], and another three built-in TMOs in Adobe Photoshop, in turn from No.1 to No.8. All the tone-mapped images were ranked from 1(best perceptual quality) to 8(worst perceptual quality) by a subjective assessment conducted by twenty participants. This experiment used the mean rank of each tone-mapped to represent its mean opinion score (MOS).

#### A. Visualized Demonstration

As you can see in FIGURE Ibelow, tone-mapped images No.2, 3, 5, 7 from image set 7 in the database mentioned above are selected for visualized demonstration. The first row of FIGURE Iare different tone-mapped images, while the second row of FIGURE Idemonstrates color difference matrices of corresponding tone-mapped images. Images mentioned above are averagely ranked 2.35, 8, 5.7, 2.75 from subjective assessment, and scored 0.9007, 0.6946, 0.8494, 0.8539 (rank 1, 8, 6, 5) by CDMI. Images in smallest (best) and biggest (worst) subjective rank are also scored highest and lowest by CDMI. Besides, correlation between CMDI and subjective score will be discussed later in experiment.

As you can see from the color difference matrix, though image 5 and 7 have similar CDMI, the "brighter" part of image 7's color difference matrix is positioned near the window, while that of image 5 is around other area. That means TMO producing image 7 gains bigger chromatic distortion while tone-mapping dark area, while bigger color loss for bright area for image 5's TMO, on the contrary. That's an extra information which can be read by visualizing color difference matrix. But that's also one of the shortcoming of CDMI, since chromatic distortion in bright and dark area may affect the perceptual quality of HVS to varying degree.



FIGURE I. Image 2 (ranked 2.35)



Image 3 (ranked 8)

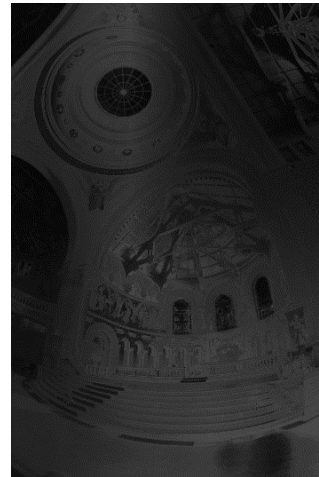
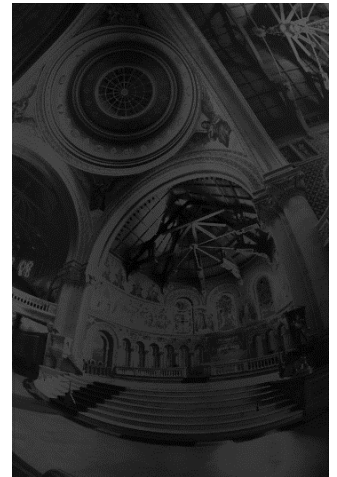


FIGURE I. CDMI = 0.8494



CDMI = 0.8539

FIGURE I. TONE-MAPPED IMAGES AND ITS VISUALIZED COLOR DIFFERENCE MATRIX WITH CDMI.

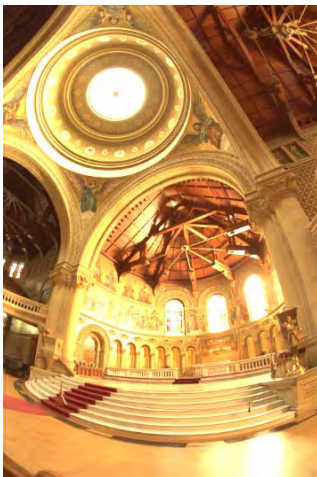


FIGURE I. Image 5 (ranked 5.7)



Image 7 (ranked 2.75)

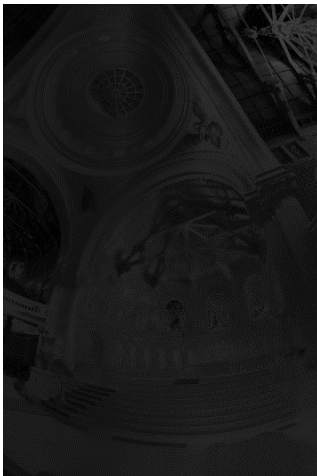
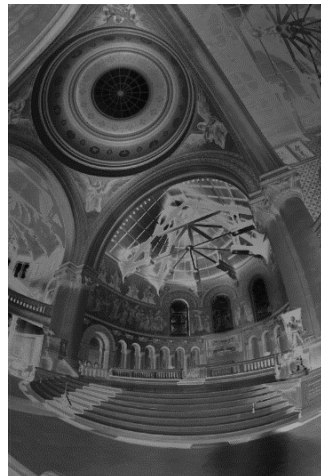


FIGURE I. CDMI = 0.9007



CDMI = 0.6946

### B. Correlation Validation

To validate the accuracy of proposed CDMI, Spearman rank-order correlation coefficient (SRCC), Kendall rank-order correlation coefficient (KRCC), Pearson linear correlation coefficient (PLCC) and root-mean-square error (RMSE) are selected to describe the correlation between subjective perception and CDMI. CDMI is then compared to two classic method: TMQI and FSTIM. Considering that CDMI is an index that only describe chromatic distortion, and maybe unqualified to become an independent quality assessment method, it's combined with them for better performance.

We used the dataset and its subjective experiment result mentioned above to validate six methods below:

- 1) *CDMI*: The proposed index as an independent method, without combining any other methods.
- 2) *TMQI*: The originator of qulity assessment of tone-mapped images, introduced in [2].
- 3) *FSTIM(RGB)*: FSTIM is another widely recognized method, introduced in [4]. FSTIM(RGB) is calculated as the arithmetic mean of FSTIM of each R, G, and B channnel.
- 4) *CDMI\_FSTIM(RGB)*: The arithmetic mean of CDMI and FSTIM(RGB).
- 5) *CDMI\_TMQI*: The arithmetic mean of CDMI and TMQI.
- 6) *CMDI\_TMQI\_FSTIM(RGB)*: The arithmetic mean of CDMI, TMQI and FSTIM(RGB).

We compute SRCC, KRCC, PLCC, and RMSE of scores given by six different methods on each image set. For each method, there are thirteen groups of SRCC, KRCC, PLCC, and RMSE. Their mean value and variance are calculated, as you can see in TABLE I.

The best performance of each eight aspects are highlighted in bold. We are going to discuss CDMI and its fusion method from the following angles:



TABLE I. PERFORMANCE COMPARISON OF SIX METHODS

Corr Indices	Methods Tested					
	1)	2)	3)	4)	5)	6)
<i>Mean Value</i>						
SRCC	0.5983	0.7925	0.8300	0.8044	0.8309	<b>0.8630</b>
KRCC	0.4816	0.6684	0.7069	0.6850	0.7179	<b>0.7565</b>
PLCC	0.6169	0.7948	0.8461	0.8186	0.8282	<b>0.8658</b>
RMSE	1.8036	1.3601	1.1435	1.2573	1.2476	<b>1.0919</b>
<i>Variance</i>						
SRCC	0.2749	<b>0.0899</b>	0.1396	0.1673	0.0963	0.0950
KRCC	0.2385	<b>0.0999</b>	0.1790	0.1621	0.1249	0.1395
PLCC	0.2568	0.0928	0.1179	0.1586	0.0918	<b>0.0911</b>
RMSE	0.5462	<b>0.3062</b>	0.4074	0.4071	0.3202	0.3715

- CDMI (1) alone gains the worst performance. We are not surprised about this result, since CDMI is an index that only measure chromatic information, rather than a complete method. But its performance boosts when combined with other method(s), see below:
- CDMI\_TMQUI (5) outperforms TMQUI (2) itself and CDMI\_FSTIM(RGB) (4) in all correlation indices. It's also better than FSTIM(RGB) (3) in terms of SRCC, KRCC, and RMSE. The ultimate fusion of three methods (6) scores the best even if they're combined with arithmetic mean, the simplest form.
- Even (6) has the best correlation with subjective results, the classic TMQUI still has the smallest variance of SRCC, KRCC, and RMSE, that means TMQUI is still the most robust one. But CDMI\_TMQUI (5) and (6) have an approximate robustness.

#### IV. CONCLUSION

In this work, CDMI, an objective index for tone-mapped images quality assessment was proposed. CDMI is the first index to introduce the concept of color difference in printing industry into tone-mapped images quality assessment, thus chromatic information which is ignored by most of previous methods is considered this time. Fusion of CDMI with other methods is needed if a better performance is required, since CDMI only measures one of many aspects that affect HVS's perceptual quality of tone-mapped images. That's also why CDMI is called an "index" rather than being qualified to be called as a "method".

A more comprehensive tone-mapped images database with more HDR images and latest TMOs are required for further studies. A subjective experiment is also required to get the MOS of new database. We will try merging more indices in a more complicated form rather than arithmetic mean and using machine learning to determine each indices' coefficients that create the best performance.

#### REFERENCES

- [1] Reinhard, E., Ward, G., Pattanaik, S., and Debevec, P. E. High dynamic range imaging : acquisition, display, and image-based lighting. Political parties and the state :. Princeton University Press., 2005.
- [2] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images," IEEE Trans. Image Process., vol. 22, no. 2, pp. 657–667, 2013.
- [3] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Trans. Image Process., vol. 13, no. 4, pp. 600–612, 2004.
- [4] H. Z. Nafchi, A. Shahkolaei, R. Farrahi, and M. Cheriet, "FSITM: A feature similarity index for tone-mapped images," IEEE Signal Process. Lett., vol. 22, no. 8, pp. 1026–1029, 2015.
- [5] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 11, pp. 1254–1259, 1998.
- [6] D. Kundu and B. L. Evans, "Visual attention guided quality assessment of tone-mapped images using scene statistics," in Proc. IEEE Int. Conf. Image Process, pp. 25–28, 2016.
- [7] D. L. Ruderman, "The statistics of natural images," Netw. Comput. Neural Systems, vol. 5, no. 4, pp. 517–548, 1994.
- [8] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," IEEE Trans. Image Process., vol. 21, no. 12, pp. 4695–4708, 2012.
- [9] K. Ma, H. Yeganeh, K. Zeng, and Z. Wang, "High dynamic range image compression by optimizing tone mapped image quality index," IEEE Trans. Image Process., vol. 24, no. 10, pp. 3086–3097, 2015.
- [10] Hadizadeh, H., & Bajic, I. V. Full-reference objective quality assessment of tone-mapped images. IEEE Transactions on Multimedia, pp. 99, 1-1, 2017.
- [11] K. Gu, S. Wang, G. Zhai, S. Ma, X. Yang, W. Lin, W. Zhang, and W. Gao, "Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure," IEEE Transactions on Multimedia, vol. 18, no. 3, pp. 432–443, 2016.
- [12] Jiang, Q., Shao, F., Lin, W., and Jiang, G. Blique-tmi: blind quality evaluator for tone-mapped images based on local and global feature analyses. IEEE Transactions on Circuits & Systems for Video Technology, pp. 99, 1-1, 2017.
- [13] Xu, L., and Jiang, X. Visual perception features of color damage in High Definition video compression. International Congress on Image and Signal Processing, Biomedical Engineering and Informatics. IEEE, pp.129-134, 2017.
- [14] Improvement to industrial colour-difference evaluation. Publication CIE No. 142, 2001.
- [15] Sharma, G., Wu, W., and Dalal, E. N. The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations. Color Research & Application, 30(1), 21-30, 2010.
- [16] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic tonereproduction for digital images," in Proc. 29th Annu. Conf. Comput.Graph. Interact. Tech., vol. 21, pp. 267–276, 2002.
- [17] F. Drago, K. Myszkowski, T. Annen, and N. Chiba, "Adaptive logarithmic mapping for displaying high contrast scenes," Comput. Graph. Forum, vol. 22, no. 3, pp. 419–426, 2003.
- [18] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of highdynamic-range images," ACM Trans. Graph., vol. 21, no. 3, pp. 257–266, 2002.
- [19] R. Mantiuk, K. Myszkowski, and H. Seidel, "A perceptual framework for contrast processing of high dynamic range images," in Proc. 2<sup>nd</sup> Symp. Appl. Percept. Graph. Visual, pp. 87–94, 2005.
- [20] S. N. Pattanaik, J. Tumblin, H. Yee, and D. P. Greenberg, "Timedependent visual adaptation for fast realistic image display," in Proc. ACM SIGGRAPH Conf. Comput. Graph, pp. 47–54, 2000.