October 05, 2018

A MONOTONICITY MEASURE WITH A FAST ALGORITHM FOR OBJECTIVE EVALUATION OF TONE MAPPING METHODS

HP INC

A Monotonicity Measure with a Fast Algorithm for Objective Evaluation of Tone Mapping Methods

***Abstract*— The range of light intensity in the real world greatly exceeds what most existing devices can display. Various tone mapping methods have been developed to render HDR (high dynamic range) images or to increase local contrast of conventionally captured images. While local (or spatially varying) tone mapping methods are generally more effective they are also prone to artifacts such as halos. Most existing methods for evaluating tone-mapped images focus on preservation of informative details and may not identify artifacts effectively. This paper proposes an objective metric based on a monotonicity measure that may serve as a baseline measure for artifacts due to intensity reversal. A naïve method to compute the metric has a high computational complexity of $O(N^2)$, where N is the total number of pixels. To make the metric acceptable for interactive applications, a fast algorithm with the complexity of $O(N)$ is presented. Experimental results using real-world images are included to demonstrate the efficacy of both the metric and the fast algorithm.**

Keywords—tone mapping, high dynamic range (HDR), monotonicity, image quality assessment

I. Introduction

It is well known that the range of light intensity perceivable by human eyes greatly exceeds that which modern digital cameras and most display devices are capable of recording and reproducing [1][2][3][4]. To overcome these limitations, high dynamic range (HDR) imaging techniques such as multiple-exposure photography and tone mapping are used to capture and reproduce such scenes. As commercial-grade HDR software such as HDR Efex Pro (Nik Collection) became freely available to lay users, image manipulation using tone mapping is increasingly popular [5].

Tone mapping methods may be classified into two types: global (or spatially uniform) and local (or spatially varying) [1][2][3]. Global mapping applies the same transformation to all pixels in the image, while local mapping may apply different transformations to different pixels based on their neighborhood properties. Most early research relied on subjective psychophysical experiments [6][7]. Although such evaluations were directly based on human perception, they required elaborate experimental setups and considerable time investment, making them ill-suited for real-time and automated applications. In addition, since the performance of local tone mapping methods is generally image-dependent, a method that is deemed to be good based on certain subjective experiments may not necessarily be good for other images. Recent research has begun to focus on objective evaluations. Yeganeh and Wang proposed tone-mapped image quality index (TMQI), a metric based on a multi-scale signal fidelity measure and a naturalness measure [8]. The naturalness measure is based on a statistical model built on approximately 3,000 8-bits/pixel grayscale images, taking into account the global intensity and contrast. The two measures are then combined nonlinearly with three parameters determined by learning. Nafchi et al. proposed FSITM (Feature Similarity Index for Tone-Mapped Images) based on a phase-derived feature map derived from multi-scale log-Gabor wavelets [9]. Hadizadeh and Bajić proposed a "bag of features" approach that includes eight features [10]. On the other hand, several no-reference methods were also published recently. Gu et al proposed a blind tone-mapped quality index (BTMQI) based on global and local entropies, statistical naturalness and structural preservation [11]. Another method by Kundu et al. in is based on standard measurements of the bandpass and differential natural scene statistics (NSS) [12]. We noticed that the aforementioned methods mainly focused on preserving informative details and may neglect artifacts such as halos that may arise from localized processing [11].

Monotonicity is a fundamental requirement for histogram-modification techniques for regular grayscale images [13]. It is also a basic requirement for global tone mapping transforms (also known as tone reproduction curves) [2]. Violation of monotonicity may produce significant artifacts as a result of intensity reversal. However, because local tone mapping methods cannot be characterized by a common curve, this concept is not readily applicable.

In this paper, we extend the monotonicity concept to an image-reference pair that is applicable to wide-ranging situations including local tone mappings. We then propose a metric based on the counting of intensity reversed pairs among all possible pixel pairs. However, a naïve counting method has the complexity $O(N^2)$, making it too slow to be useful for most practical applications. To alleviate this problem, we present a fast $O(N)$ method for the most commonly encountered 8-bit images.

II. Proposed Metric and Fast Computation

First, we will extend the concept of monotonicity to an image-reference pair, and introduce a quantitative measure. We will then prove two important properties associated with the ordering problem and present a fast counting algorithm for 8-bit grayscale images.

Given a reference image I_0, the output image I_1, and a positive threshold t, the intensity order of a pair of pixels at two locations (i, j) and (m,n) is considered reversed if $sign(\Delta I_0(i, j; m, n)) \neq sign(\Delta I_1(i, j; m, n))$ and $(|\Delta I_0(i, j; m, n)| + |\Delta I_1(i, j; m, n)|) > t$, where $\Delta I_0(i, j; m, n) = I_0(i, j) - I_0(m, n)$ and

$\Delta I_1(i,j;m,n) = I_1(i,j) - I_1(m,n)$. Notice that the comparison with the threshold t may be defined in terms of percentage if the dynamic ranges of the input and output are different.

Based on the above definition, a metric measuring the monotonicity of the transform T: $\{I_1 = T(I_0)\}$ may be defined as:

$$\mu = 1 - N_r / N \quad (1)$$

Where N_r is the total number of pixel pairs with reversed order and N is the total number of pixel pairs. Notice that the definition does not require T to be explicitly defined. It should also be noted that since all possible pixel pairs are considered, this metric should take all possible scales into account. The metric value is in the range of [0,1]. A transform is considered totally monotonic if $\mu = 1$.

Clearly, the computational complexity falls completely on the counting process. For an image of total N pixels there are $N(N-1)/2$ pixel pairs. Consequently, a naïve treatment of the general counting problem has complexity $O(N^2)$, which is too high for most interactive applications. Towards a fast counting algorithm, we explore some important properties.

The differential ΔI as defined above is based on the pixel values at *two locations of the same image*. Another differential can be defined as $\delta I(i,j) = I_0(i,j) - I_1(i,j)$, which is based on the pixel values at the *same location of the two images*. The two differentials are related by

$$\delta I(i,j) - \delta I(m,n) = \Delta I_0(i,j;m,n) - \Delta I_1(i,j;m,n) \quad (2)$$

Using this relationship, we can prove two properties. For simplicity, notations ΔI_0 and ΔI_1 are used in the following.

Prop. 1. No order reversal for any pair of pixels if they have the same δI value.

Proof. For any two pixel locations (i,j) and (m,n), $\delta I(i,j) = \delta I(m,n)$ leads to $\Delta I_0 = \Delta I_1$ and consequently $sign(\Delta I_0) = sign(\Delta I_1)$. Therefore the intensity order remains the same by definition.

Prop. 2. For any pixel pair at (i,j) and (m,n) that satisfies the conditions $sign(\Delta I_0) \neq sign(\Delta I_1)$ and $(|\Delta I_0| + |\Delta I_1|) > t$, the following is true: $|\delta I(i,j) - \delta I(m,n)| > t$

Proof: For all pixel pairs with $sign(\Delta I_0) \neq sign(\Delta I_1)$, there are only two possibilities:

1. $\Delta I_0 \geq 0$ and $\Delta I_1 < 0$ or $\Delta I_0 > 0$ and $\Delta I_1 \leq 0$, in both cases, $\delta I(i,j) - \delta I(m,n) = \Delta I_0 - \Delta I_1 = |\Delta I_0| + |\Delta I_1| > t$
2. $\Delta I_0 \leq 0$ and $\Delta I_1 > 0$ or $\Delta I_0 < 0$ and $\Delta I_1 \geq 0$, in both cases, $\delta I(i,j) - \delta I(m,n) = \Delta I_0 - \Delta I_1 = -(|\Delta I_0| + |\Delta I_1|) < -t$

Therefore $|\delta I(i,j) - \delta I(m,n)| > t$.

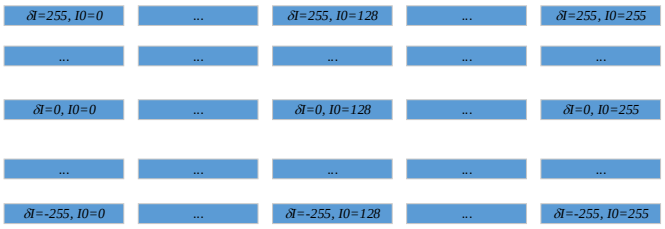

Figure 1. A matrix of counters for (δI, I_0) tuples.

However, the reverse may not be true. Although it can be proven easily that the condition $|\delta I(i,j) - \delta I(m,n)| > t$ leads to $(|\Delta I_0| + |\Delta I_1|) \geq |\Delta I_0 - \Delta I_1| = |\delta I(i,j) - \delta I(m,n)| > t$, it doesn't guarantee that $sign(\Delta I_0) \neq sign(\Delta I_1)$. This means the set of pixel pairs $\{(i,j),(m,n) | (sign(\Delta I_0) \neq sign(\Delta I_1)) \wedge (|\Delta I_0| + |\Delta I_1|) > t\}$ is a subset of $\{(i,j),(m,n) | |\delta I(i,j) - \delta I(m,n)| > t\}$.

Utilizing the above two properties, a fast counting algorithm for *8-bit* images is presented in the following.

For 8-bit grayscale images, the range of δI is $[-255, 255]$. We designed a matrix $\sum$ of counters to count all pixels with all possible values of δI and I_0, as shown in Figure 1.

The matrix Σ has 511 rows corresponding to 511 possible δI values. Each row in turn has 256 columns, corresponding to possible values of I_0. Notice that cells of the same row have the same δI and therefore no order reversal among them. Only cells in different rows with distance $|\delta I_1 - \delta I_2| > t$ need to be checked. Additionally, since pixel pair (i,j) and (m,n) is considered the same as the pair (m,n) and (i,j), we only need to check $\delta I_1 - \delta I_2 > t$. Further computational savings may come from two aspects: 1) the matrix is sparse in many

List 1. Pseudocode for fast counting

```
int FastCount(I0, I1, t)  {
Initialize counters Σ to zeros.
// loop 1
for all locations (i,j){
  δI = I0(i,j) − I1(i,j);
  Σ(δI + 255, I0) + +;
}

count = 0;
for(i=510;i>0;i--) {
 δIi = i − 255;
 for(m=(i-t);m>0;m--) {
   δIm = m − 255;
   D = δIi − δIm;
    if D < t
       Continue;
    for( j=0;j<255;j++) {
       for( n=0;n<255;n++) {
         ΔI0 = j − n;
         ΔI1 = ΔI0 − D;
         if sign(ΔI0) ≠ sign(ΔI1)
           count+= Σ(i,j) · Σ(m,n);
       };
    };
 };
};
return count; }
```

cases, and 2) each cell generally represents many pixels such that one counting of cells may be equal to a substantial number of direct counting of pixels. Details of the algorithm are presented as a C-style pseudocode in List 1. It should be pointed out that the pseudocode is mainly for the purpose of complexity analysis and omits some minor programming tricks. For example, an additional counter can be added for each row such that if a row has no entry, further checking can be skipped.pseudocode in List 1. It should be pointed out that the pseudocode is mainly for the purpose of complexity analysis and omits some minor programming tricks. For example, an additional counter can be added for each .

Clearly, the complexity of the algorithm consists of two parts. The first part (the first loop) is $O(N)$, where N is the total number of pixels. The second part depends only on the fixed-size 511×256 matrix $\sum$ and is independent of N. Since the complexity for the second part is $O((511 \times 256)^2)$ the overall complexity is $O(N + (511 \times 256)^2)$. This means that for images larger than 511×256, the complexity of the proposed algorithm is guaranteed to be lower than that of naïve counting. In practical situations, there may be many pixels that do not change in intensity, resulting in a sparse matrix $\sum$. Therefore the proposed algorithm is faster even for images of smaller dimensions, which will be shown in the experimental section.

III. EXPERIMENTAL RESULTS

This section is intended to show the proposed metric computed on some representative cases, along with computation times in comparison with the naïve method. We also compared our metric with a recently published method using their publicly available implementation. For these experiments, we chose Adobe Photoshop CS6 and the Nik Collection by Google (Version 1.2.11) (Nik Collection has since been acquired by DXO [5]). In particular, Photoshop CS6 was used for interactive global tone adjustment, and HDR Efex Pro 2 (in Nik Collection) was used for local tone mapping of both multiple and single images. HDR Efex Pro 2 is a robust commercial-grade software. It can automatically produce visually appealing images in various preset styles for both multiple exposure and single exposure images without manual parameter tuning.

The first case is shown in Figure 2. Three images in (a), (b) and (c) were taken in raw 14-bit/pixel format and in three exposure compensations EV-2, EV0 and EV+1, with a Nikon D800E DSLR fixed on a tripod. By merging the three images in HDR Efex Pro 2, two 16-bit images of different styles were produced. Separately, the single raw image with EV0 (b) was adjusted in Adobe Photoshop Camera Raw 9.1.1.461 convertor. The parameters adjusted including "highlights", "shadows" and "whites". This is a global adjustment (tone mapping). For evaluation, all final images including the reference image (b) were converted to 8-bit. All images were further downsized to three sizes keeping the original aspect ratio. For each set of images in the three sizes, the proposed metric μ was computed (with $t=10$) and the elapsed times recorded. All computations were performed on a HP Z640 Workstation (Intel Xeon CPU @2.40GHz, 16GB RAM). For comparison, the same sets of images were evaluated using the no-reference methods [12]. The results are included in Table 1, where T_f and T_n are the elapsed time for the fast and naïve methods, respectively, H_1 and H_2 are computed scores using the publicly available codes of the HIGRADE-1 and HIGRADE-2 [12].

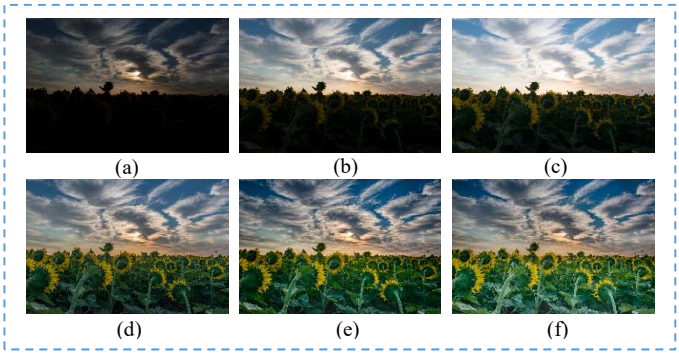

Figure 2. A multiple exposure case. (a)(b)(c) Three exposures with EV-2.0, EV0.0 and EV+1.0, respectively. (d) Global tone adjusted using only single image (b). (e) HDR merged and tone mapped (preset 26, landscape, vignette), (f) HDR merged and tone mapped (preset 23, outdoor1).

Table 1. Evaluation results of the Case 1.

		 Fig. 2(b)	 Fig. 2(d)	 Fig. 2(e)	 Fig. 2(f)
360 x 240	μ	0.95	0.89	0.81
	T_f (s)	0.13	0.33	0.38
	T_n(s)	7.28	8.55	9.22
	H_1	0.74	1.03	0.89
	H_2	0.56	0.56	0.56
512 x 342	μ	0.95	0.88	0.81
	T_f (s)	0.17	0.44	0.49
	T_n(s)	28.43	33.71	36.97
	H_1	0.92	1.02	1.17
	H_2	0.71	0.67	0.74
1024 x 683	μ	0.95	0.89	0.81
	T_f (s)	0.26	0.50	0.61
	T_n(s)	426.56	483.99	531.04
	H_1	1.06	1.14	1.21
	H_2	0.76	0.67	0.72

Some observations can be made from the results. The proposed metric μ achieved the highest score for the image of Fig. 2d obtained using global mapping. This is consistent with the results of subjective experiments [7]. The image (Fig. 2f) with the most significant local luminosity change scored the lowest. The metric is also consistent across three scales. In terms of computation time, the fast method is about 30 times faster than the naïve one even for the smallest image size of 360x240. With an approximately 8-second runtime even for the smallest image tested, the naïve method would be unacceptably slow for most applications. Regarding the methods HIGRADE-1 and HIGRADE-2 [12], we observed that the scores for individual images, and the ranking of those scores, were both inconsistent across scales. Moreover, H_1 scores exceed 1.0 without a clear range.

The second case is based on a single image, shown in Figure 3. The image's exposure was set based on the moon, causing the

forest to be too dark. It is desirable to keep the details of the moon while bringing back some details of the forest. Tone mappings similar to those in case 1 were applied. Application of global tone adjustments (including "highlights", "shadows", "blacks", and "whites") resulted in the image in Fig. 3b. The original image was also processed in HDR Efex Pro 2 in "Tone Mapping (Single Image)" mode. In two independent runs, images with the same presets as in case 1 were selected, as shown in Figure 3 (c) and (d). It can be observed that although the images adjusted by local tone mapping contain more details in the forest region than the globally adjusted images, they also contain substantially more non-uniform luminance artifacts. Compared to the original image, the images resulting from local tone mapping would appear unnatural. It should also be pointed out that the exact same local tone mapping method with default parameters produced different levels of artifacts in the two cases.

The results for case 2 are presented in Table 2. The behavior of the proposed metric is consistent across scale and with case 1. Runtime was faster than in case 1, probably due to the large region occupied by sky in the image. Again, HIGRADE-1 and HIGRADE-2 scores are inconsistent across scales. In this case, HIGRADE-2 even returned some negative scores. In both cases, images with more details were scored higher despite their significant artifacts.

IV. CONCLUSION

In this paper, we extend the concept of monotonicity to a general image-reference pair and introduce a quantitative metric for tone mapping methods. We also presented a fast counting method that makes the metric computation acceptable for interactive applications. Our results demonstrated that the metric is consistent across scales with scores reflect perceivable artifacts. The experiments also validated the dramatic speedup afforded by the fast algorithm in comparison with the naïve method. Our experiments also uncovered some limitations of a recently published method [12]. Future studies may expand on the proposed metric by adding an exponential element to control the rate of decay, and by taking the amount of intensity reversal into account, in addition to the number of occurrences.

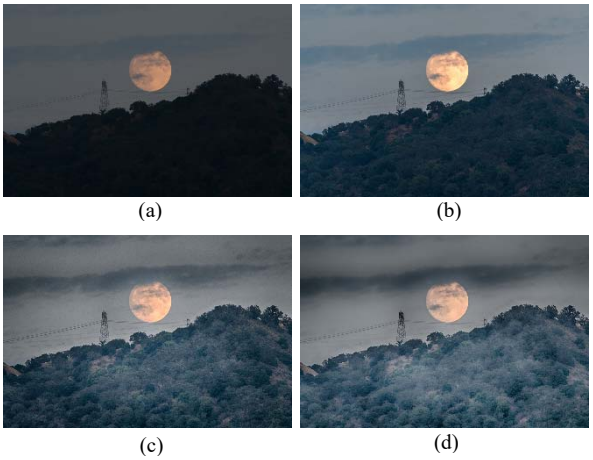

Figure 3. A single exposure case. (a) Captured image. (b) Global tone adjusted. (c) Single-image tone mapped (preset 26, landscape, vignette), (d) Single-image tone mapped (preset 23, outdoor1).

Table 2. Evaluation results of the Case 2.

		 Fig. 3(a)	 Fig. 3(b)	 Fig. 3(c)	 Fig. 3(d)
360x240	μ		0.99	0.87	0.76
	T_f (s)		0.013	0.041	0.049
	T_n(s)		9.79	11.13	12.23
	H_1		0.36	0.70	0.74
	H_2		0.036	0.17	0.39
512x341	μ		0.99	0.87	0.76
	T_f (s)		0.021	0.057	0.076
	T_n(s)		38.48	43.97	46.70
	H_1		0.36	0.90	0.94
	H_2		0.28	0.29	0.71
1024x683	μ		0.99	0.87	0.76
	T_f (s)		0.045	0.11	0.14
	T_n(s)		668.51	726.07	814.12
	H_1		0.04	0.32	0.85
	H_2		-0.26	-0.46	0.30

REFERENCES

[1] Kate Devlin, "A review of tone reproduction techniques", Computer Science, University of Bristol, Tech. Rep. CSTR-02-005, 2002

[2] Jeffery M. DiCarlo and Brian A. Wandell, "Rendering high dynamic range images", Proc. SPIE: Image Sensors 3965, pp. 392–401, 2000.

[3] G. Eilertsen, R. K. Mantiuk and J. Unger, "A comparative review of tone-mapping algorithms or high dynamic range video", Computer Graphics Forum archive, Vol. 36, Issue 2, pp. 565-592, May 2017

[4] Frédo Durand and Julie Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images", SIGGRAPH '02 Proceedings of the 29th annual conference on Computer graphics and interactive techniques, pp. 257-266

[5] https://www.google.com/nikcollection/; https://nikcollection.dxo.com/

[6] Akiko Yoshida, Volker Blanz, *et al*, "Perceptual evaluation of tone mapping operators with real-world scenes", Proc. SPIE, Human Vision and Electronic Imaging X, Vol. 5666, pp. 192–203, March 2005

[7] Martin Cadik, et al., "Evaluation of HDR tone mapping methods using essential perceptual attributes", Computers & Graphics 32 (2008) 330–349

[8] H. Yeganeh and Z. Wang, "Objective quality assessment of tone-mapped images", IEEE Trans. Image Proc. Vol. 22, No. 2, pp. 657-667, Feb. 2013

[9] H. Z. Nafchi, et al., "FSITM: a feature similarity index for tone-mapped images", IEEE Sig. Proc. Lett., Vol. 22, No. 8, pp. 1026-1029, Aug. 2015

[10] Hadizadeh and Bajic, "Full-reference objective quality assessment of tone-mapped images", IEEE Trans. Multimedia, Vol. 20, No. 2, pp. 392-404, Feb. 2018.

[11] Ke Gu, et al., "Blind quality assessment of tone-mapped images via analysis of information, naturalness, and structure", IEEE Trans. Multimedia, Vol. 18, No. 3, pp. 432-443, Mar. 2016

[12] Debarati Kundu, et al., "No-Reference quality assessment of tone-mapped HDR pictures", IEEE Trans. Imag. Proc., Vol. 26, No. 6, pp. 2957-2971, Jun. 2017; http://live.ece.utexas.edu/research/Quality/higradeRelease.zip

[13] R. C. Gonzalez and P. Wintz, Digital Image Processing, 2E, Chapter 4, Addison-Wesley, 1987

Disclosed by Jian Fan, HP Inc.