HUMAN VISUAL SYSTEM INSPIRED SALIENCY GUIDED EDGE PRESERVING TONE-MAPPING FOR HIGH DYNAMIC RANGE IMAGING

Nipu R. Barai[1], Matthew Kyan[2], Dimitrios Androutsos[1]

[1] Ryerson University, Dept. of Electrical & Computer Engineering, Toronto, Canada.
[2] York University, Dept. of Electrical Engineering & Computer Science, Toronto, Canada

ABSTRACT

With the growing popularity of High Dynamic Range Imaging (HDR), the necessity for advanced tone–mapping techniques has greatly increased. We propose a Human Visual System (HVS)-inspired Saliency Guided Edge-preserving Tone Mapping method (HSGETM) that uses saliency region information of an HDR image as input to the guided filter for a base and detail image layer separation. After detail layer enhancement and base layer compression with constant weights, a new edge preserved tone mapped image is composed by adding the layers back together with saturation and exposure adjustments. The filter operation is faster due to the use of the guided filter which has *O(N)* time operation with *N* number of pixels. Experimental results demonstrate that the HSGETM has higher edge, and naturalness preserving capability, which is homologous to the HVS, as compared to other state-of-the-art tone mapping approaches.

Index Terms - Human Visual System, Saliency map, High Dynamic Range Imaging, Tone mapping, Guided filter.

1. INTRODUCTION

There is a growing trend for HDR to replace Low Dynamic Range (LDR) imaging because of its extremely versatile applications in photography, architecture, surveillance, astronomy, remote sensing, human vision studies, medical imaging, computer graphics and in many other related fields [1]. The HDR files cannot be directly rendered to most displays (which render only LDR) and need to be re-mapped such that image fidelity is maintained. The process is known as tone-reproduction (TR) or tone mapping (TM). Though there exist various TM methods, Human Visual System (HVS) inspired tone mapping methods are highly incremental. Thus, we have focused on developing a HVS inspired TM approach.

Salient areas are a key factor in attracting the human eyes within a scene. A saliency map is an arranged topographical map that represents the visually dominant regions of an image. Such maps are often designed as an input to convert selective attention [2]. Saliency detection could be bottom up or top down. Bottom up approaches detect areas that grab human attention naturally. Top down approaches detect areas, which are intentionally looked for by an observer. Very few methods have been proposed for saliency guided tone mapping [3-4].

As salient regions more readily draw a viewer's attention, it would seem obvious that tone mapping should prioritize a higher fidelity reconstruction of image content in these regions. Motivated by this, we propose an HVS inspired saliency guided edge preserving tone mapping method in which we present a way of using saliency map information directly to guide the tone mapping. In our method, the High Definition Human Visual Saliency (HDHVS) method [5] is applied to obtain the saliency map of HDR images. The resulting saliency map is then used to modulate and guide filters that preserve edge information during the tone mapping operation.

2. SUPPORTING METHODS

2.1. HDHVS method

The HDHVS detection model is a wavelet inspired implementation of Itti's model [6] that facilitates the extraction of a high resolution saliency map. The popular center-surround (CS) method by Itti is a biologically inspired saliency detection method, where the saliency map is obtained by comparing a small local region to its surrounding region, responding to the presence of opposing intensity/spatial features between them. Itti's method is limited to low resolution saliency map and high computational time. To modify it for high resolution map, Saber *et al.* [5] proposed HDHVS method that use wavelets to perform the center surround operation. It eliminates the limitation of lower resolution and makes the process faster.

In the HDHVS method the gray scale for intensity feature; red, green, blue and yellow channels for color opponency features are extracted using linear filters. An intensity feature map is obtained by center-surround operation in the gray image. The center is assumed to be the approximate image after wavelet transformation and surround is the main gray image. Red-Green and blue-yellow feature maps are obtained by performing the same operation as the intensity map. The only difference is with red approximate as center with green channel is used as surround and vice versa. Similarly, for the blue-yellow channel. This is performed to mimic HVS color opponency as per Itti's operation. After that, All individual feature maps are added to form final saliency map. The down sampling of the surround scale to center scale is done by bi-cubic interpolation and point by point subtraction.

Due to use of wavelets, the resultant maps from each feature type can be composed back to high resolution image using inverse wavelet transform with the vertical, horizontal and detailed components of the image.

2.2. Guided filtering

A Guided Filter is a structure transferring smoothing filter [7]. It requires two input images: the guidance image (I) and the input image (p). The filtered output, q, is assumed to be a linear transform of the guidance image in a defined window. Equation (1) shows the linear relation.

$$q_i = a_k I_i + b_k, \forall i \in \omega_k \tag{1}$$

(a_k, b_k) are some linear coefficients assumed to be constant in window ω_k. The cost function developed to reduce difference between the output q and input p is:

$$E(a_k, b_k) = \sum_{i \in \omega_k} ((a_k I_i + b_k - p_i)^2 + \epsilon {a_k}^2) \tag{2}$$

The regularization parameter ϵ controls the coefficient a_k. Coefficient values are obtained by solving Equation (2) using linear regression. The main idea behind its edge preserving property is, a pixel value remains unchanged in a high variance region and averaged in a low variance region. The algorithm is explained in details in the next two sections.

3. HVS INSPIRED SALIENCY GUIDED FILTERING

The separation of images into detail and base layer is very common in image manipulation. There have been many TM algorithms that follow this process [3, 4, 8-10].

The proposed method is similar to Durand's [8] and Li's [3] TM approaches. However, Durand uses the bilateral filter for separating a base and detail layer. Li uses a guided filter; however, their method has a histogram based saliency weight to manipulate the detail layer rather than using it directly in the filter.

We have used the guided filter for our proposed method. The advantages of using a guided filter over the bilateral filter are many: the guided filter has better edge preserving performance and structure transferring capability than the bilateral filter. It also eliminates the halo and gradient reversal artifacts which are limitations of the bilateral filter. The guided filter is also faster due to $O(N)$ time operation, where, N is the number of pixels. In the case of bilateral filtering, it is $O(Nr^2)$, where r is the radius of the filter.

The guided filter can take two input images. Both the inputs could be same input images or different images. One is the input image and another is the guidance image. We use the HDHVS guided saliency map as the guidance image and the original HDR as the input image:

$$SM_{HDR} = HDHVS(I_{HDR}) \tag{3}$$

HDR saliency computational time is longer than the LDR saliency processing computational time. The saliency map has the same resolution as the input image thus; no rescaling is required before using it in the guided filter.

$$q_{GFHDR} = f(a) * (SM_{HDR}) + f(b) \tag{4}$$

The guided filter algorithm for both gray and color image is available online [13]. We have used the luminance image as the input of the guided filtering.

$$a = \frac{f(SM_{HDR} * LL) - [f(SM_{HDR}) * f(LL)]}{[f(SM_{HDR} * SM_{HDR}) - f(SM_{HDR}) * f(SM_{HDR})] + eps} \tag{5}$$

$$b = f\,(f(LL) - a * f\,(SM_{HDR})) \tag{6}$$

In Equation (2) to (4), LL is the input image, SM_{HDR} is the guidance image, r is the radius of the window for the mean filter and eps is the regularization parameter. q_{GFHDR} is the guided filtered HDR image which is the linear transform of the guidance saliency map *f()* denotes a box filtered operation where, each pixel is replaced by the average cumulative sum of its neighbouring pixels. The square window of the box filter is defined with window radius r.

The overall idea behind using saliency guided filtering is to transfer the structure information of the saliency map to the output image while smoothing the image. Highly variant areas mean these regions have high gradient variation e.g. corners or edges in an image. We want to preserve details in the tone mapped image and compress only the information, which will not make any difference visually.

4. HSGETM ALGORITHM

The proposed TM algorithm can be divided into three parts: selecting the input and guidance image; guided filtering to separate the base and detail layer; and composing back the new developed detail and base layer to obtain the tone mapped image. Figure 1 shows the work-flow of the saliency guided tone mapping method.

4.1. Input image

The input HDR radiance map is converted into a luminance image. The logarithm of the luminance is used for further processing because of easy manipulation of image layers in the log scale without effecting the image.

$$LL = log\,(I_L) \tag{7}$$

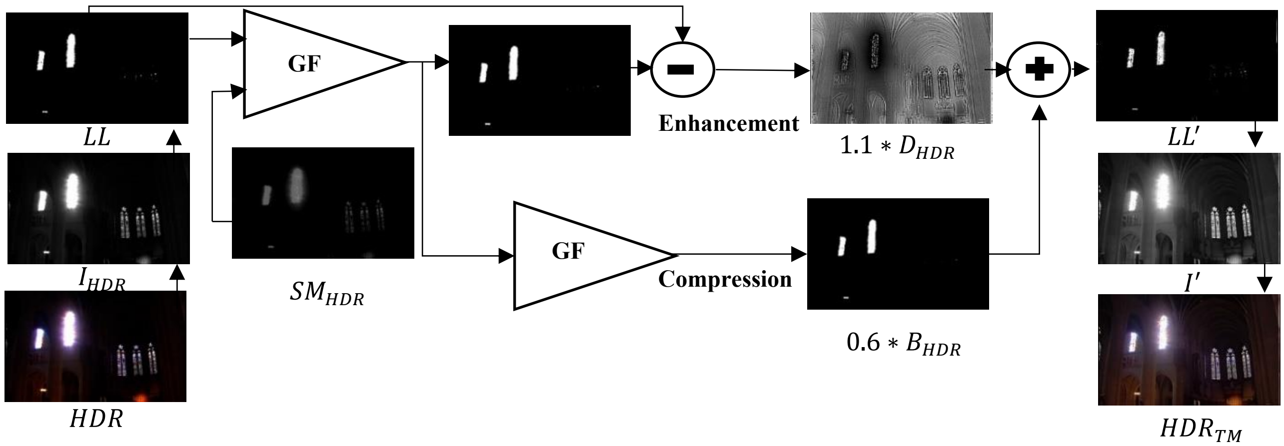


Fig 1. The flow chart showing the workflow of the proposed algorithm. GF is guided filter, I is the intensity image and LL is log luminance of the intensity image, $(0.6 * B_{HDR})$ is the compressed base layer and $(1.1 * D_{HDR})$ is the enhanced detail layer.

The HDHVS map of the input HDR image is used as the guidance image in the filter as Equation (1).

4.2. Guided filtering

Next step is the filtering process. Equations (2), (3) and (4) are the guided filtering equations. We have used the guided filter two times. This way the base layer can be compressed more without distorting the image in anyway. The detail layer is obtained by the difference between the original LL and the guided filtered output using radius 2. The first filtered output is further filtered using with a radius 3 to obtain a base layer. Both the inputs of the guided filter are the previous filtered output during base layer computation.

$$D_{HDR} = LL - P_{GFHDR} \tag{8}$$

4.3. Tone mapped output

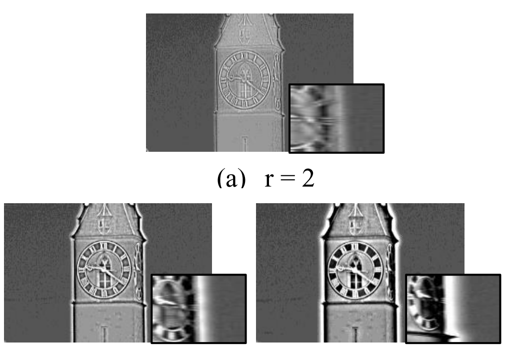
The detailed and base layer are enhanced and compressed respectively with a constant weight of 1.1 and 0.6 in the proposed algorithm. This combination of weights avoids introduction of artifacts in the final result. The detail layer is median filtered with a 2x2 neighbourhood for smoother and artifact free output.

The saturation parameter is 0.5 for the proposed algorithm. The exposure 0.6 was used by Li *et al.* [3] however, 0.6 eliminates the naturalness of some of our tested image which can be avoided using a value of 0.5. Exposure is changed according to the mean pixel value of the original HDR image using Equation (10).

$$HDR_{TM} = f_{new}[exp(LL'))] \tag{9}$$

f_{new} denotes the operation to get back to color tone mapped image from new luminance. The exposure equation applied for the experiment is developed by user defined exposure values of 32 images from EMPA dataset. It is observed that the mean pixel (M_{Pixel}) of the HDR image and exposure values are inversely related. Thus, the best fitted power equation of the exposure in terms of mean pixel value, is used for the proposed tone mapping method.

$$Exposure = (1.276 * M_{Pixel})^{-0.3535} - 0.1418) \tag{10}$$



(a) r = 2

(b) $r = 5$ (c) $r = 10$

Fig. 2. Effects of the luminance image layers for different radius values. Increasing radius introduces artifacts.

5. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, three publicly available HDR databases, EMPA [12], RIT MCSL [13], and Paul Debevec's database [14], are used to evaluate the performance of our approach. The first one contains 32 images; the second one contains 62 images. Thirteen common HDR images from the Paul Debevec's website are tested with the proposed method as well. The RIT MCSL database includes the images with various lighting conditions including both indoor and outdoor images. It has clear salient regions as well. The EMPA database contains all outdoor images with high color

variation. All the images used for the experiment are RGBE formatted radiance maps proposed by Greg Ward [1].
The radii for the guided filter are selected as 2 and 3 for the experiment. This parameter selection is made by observing the introduced artifacts in the resultant images. Radius value of the guided filter is varied from 1 to 10 to observe the artifacts in the final result. It should be pointed out that a radius greater than 3 introduces halo and gradient reversal artifacts in the resultant image. Thus, the value could be varied between 2 and 3 for a fixed regularization parameter. Figure 2 shows the resulting artifact with an increasing radius value.

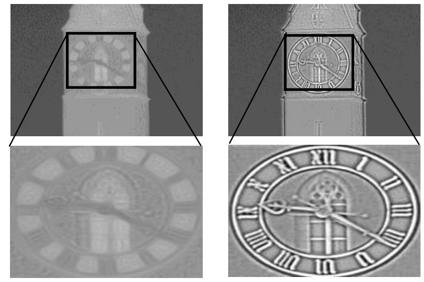
The exposure equation obtained by 32 tested images worked well for all the test images. The saturation parameter is selected to be 0.5 for the proposed method. It worked well for all other test images. If the images are too bright, the saturation parameter might vary between 0.65 – 0.75.

The resultant images are tested with two different edge strength based image quality assessment methods. One is the feature based image quality index metric (FSIM) [15] and the other one is the edge strength based image quality index metric (ESSIM) [16]. The features used in evaluating a score comparing the original and test image, are based on phase congruency and gradient strength. These two offer an objective edge-aware evaluation. ESSIM generates a score based on the edge similarity between the original and the test result. Other tone mapping methods selected for the comparison in our experiment are Durand's [10], Fattal's [11] and Mantiuk's [17] tone mapping approaches. They are widely used HVS inspired methods. We don't observe the FSIM scores of all the tone mapped versions of each images. Rather, we observe the percentage of high scores of FSIM for each method. This is because, for each individual image, the FSIM score range is different and the test images are completely different from each other. Table 1 shows the percentage of higher scores for HSGETM, Durand's, Fattal's and Mantiuk's methods using both FSIM and ESSIM quality assessment, respectively.

Table 1. Percentage of high FSIM (a) and ESSIM (b) scores for tested image

Database	Proposed		Durand		Fattal		Mantiuk	
	(a)	(b)	(a)	(b)	(a)	(b)	(a)	(b)
EMPA	35	97	10	1	50	1	5	1
RIT MCSL	70	48	5	8	20	26	5	18
Debevec	47	97	1	1	47	1	5	1
Overall	56	70	6	5	33	15	5	10

Overall, the proposed method clearly indicates better FSIM performance than the other three approaches. Fattal's method provided better result after the proposed method. This is because the method is gradient based and the feature used in scoring the FSIM is based on gradient magnitude. The datasets used for the testing, contain high variation in lighting condition and exposures. The proposed method demonstrates better performance for RIT MCSL images which has great variation in lighting and contains more salient regions. The EMPA dataset consists of four different exposed images for every HDR acquisition. The proposed method provides the higher percentage of ESSIM high scores for all the tested images. This proves the tone mapped images have better edge preserving capability compare to the other three methods.



(a) (b)

Fig. 3. Difference in the detail layer using saliency map. (a) detail layer obtained without using saliency map as a guidance and (b) detail layer obtained using the HDHVS saliency map as guidance.

A subjective quality assessment is also performed using 20 HDR images. For each image, five different tone mapped versions including the proposed method are generated. The other methods are Durand's [10], Fattal's [11], Mantiuk's [17] and Reinhard's [1]. A total of 51 observers have ranked the images on a scale of 1 to 5. 1 being the "best" and 5 being the "worst", based on the naturalness of the image and capability of preserving edge details. The names of the methods have not been mentioned to the observers, and the images have been mixed randomly. Thus, the resultant scores are unbiased. For half of the images (10 out of 20), our proposed method provided the lowest numerical ranking, which implies the "best" qualify. Our method was, overall, the best performing method, with Mantiuk's method second (which had the lowest numerical ranking in 9 cases) Therefore, our method achieves improved fidelity in terms of naturalness and edge preserving capability.

6. CONCLUSION

In this paper, we proposed an HVS inspired HDHVS map guided edge preserving tone mapping approach for HDR imaging. We use saliency information directly to guide the filter while separating an image layer with edge or detailed information of the image, and a base layer with less variant information of the image. The two layers are manipulated based on importance of the information. The detail layer is more important and thus, enhanced and filtered for preserving enhanced edge information in the output image. The base layer is compressed with a constant weight. We get a new luminance image by adding the two newly formed layers. The saturation parameter and exposure parameter are adjusted according to HVS. Both the subjective and objective survey show that the methods preserve edges and naturalness better than similar tone mapping methods. Our future work is focused on developing a saliency aware tone mapping quality metric for more efficient quality assessment.

7. REFERENCES

[1] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, "High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting", Morgan Kau_man, 2005.

[2] L. Zhang, W. Lin, "Selective visual attention: Computational models and applications", Wiley – IEEE Press, 2013.

[3] Z. Li, J. Zheng, "Visual-Salience-Based Tone Mapping for High Dynamic Range Images", IEEE Transactions on Industrial Electronics, vol. 61, pp. 7076-7082, 2014.

[4] Y. Mei, G. Qiu, K. Lam, "Saliency Modulated High Dynamic Range Image Tone Mapping", IEEE sixth International Conference on Image and Graphics, pp. 22-27, 2011.

[5] Y. Saber, M. Kyan, "High Resolution Biologically Inspired Salient Region Detection", IEEE International Conference on Image Processing, pp. 649 – 652, 2011.

[6] L. Itti, C. Koch, E. Niebur, "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis", Pattern Analysis and Machine Intelligence, Vol. 20, Issue. 11, PP 1254 – 1259, 1998.

[7] K. He, J. Sun, X. Tang, "Guided Image Filtering", IEEE Transactions on Pattern Analysis and Machine Intelligence , vol.35, Issue: 6, pp. 1397 – 1409, 2013.

[8] F. Durand, J. Dorsey, "Fast bilateral filtering for the display of high dynamic range images", ACM Transactions. on Graphics, SIGGRAPH'02, 2002.

[9] R. Fattal, D. Lischinski, M. Werman, "Gradient domain high dynamic range compression", 29th annual conference on Computer graphics and interactive techniques, pp. 249- 256, 2002.

[10] Z. Farbman, R. Fattal, D. Lischinski, R. Szeliski, "Edge-Preserving decompositions for multi-scale tone and detail manipulation", ACM Trans. On Graphics, vol. 27, Issue 3, 2008.

[11] K. He, "Guided Image Filtering", Microsoft Research. Website Visited on Nov 10th, 2016. Website: http://kaiminghe.com/eccv10/index.html.

[12] EMPA Media Technology, HDR database, 2011-2013. Website: http://www.empamedia.ethz.ch/hdrdatabase/.

[13] RIT MCSL High Dynamic Range Image database. Website:http://www.cis.rit.edu/research/mcsl2/icam/hdr/rit_hdr/.

[14] Paul E. Debevec website. Website visited on Oct 31st, 2016. Website: http://www.pauldebevec.com/

[15] L. Zhang, L. Zhang, X. Mou, D. Zhang, "FSIM: A Feature Similarity Index for Image Quality Assessment", IEEE Transactions on Image Processing, vol. 20, pp. 2378 – 2386, 2011.

[16] X. Zhang, X. Feng, W. Wang, W. Xue, "Edge Strength Similarity for Image Quality Assessment", IEEE Signal Processing Letters, vol. 20, pp. 319 – 322, 2013.

[17] R. Mantiuk, S. Daly, L. Kerofsky, "Display Adaptive Tone Mapping", SIGGRAPH, vol. 27, Issue 3, 2008.

[18] Eilertsen, Gabriel and Mantiuk, Rafal K. and Unger, Jonas, "A comparative Review of Tone mapping Algorithms for High Dynamic Range Video", Computer Graphics Forum, The Eurographics Association and John Wiley & Sons Ltd, 2017.