Fast Segmentation-based Tone Mapping Operator

Jakkarin Singnoo

University of the Thai Chamber of Commerce, THAILAND

**jakkarin_sin@utcc.ac.th*

ABSTRACT

To realistically display HDR images on low-end devices such as mobile phone, a fast and efficiency TMO that free from halo artifacts is needed. In this work, a segmentation-based TMO that offers fast, efficiency, and free of halo artifacts has been developed. The idea is to divide the image into segments and only apply an unsharp masking to the segment that lack of local contrast. The results on several images show that the tone-mapped images have good dynamic range compression and contain no halos due to the benefit of segmentation-based approach. To evaluate the image quality of the proposed TMO, 2AFC psychophysics experiments are also conducted. The result shows that the proposed TMO in terms of both photorealistic and image preference is comparable to other high computational cost operators, indicating that the proposed TMO is suitable to be used in any devices since it demands much less computational costs.

KEYWORDS: *High Dynamic Range, Tone Mapping Operator, Psychophysics*

INTRODUCTION

The range of light in real scene can have up to ten order of magnitude, this number far more exceed the number of dynamic range that human and typical image capturing devices can handle. This discrepancy leads to the problem of how one can properly reproduce the scene luminances to display luminances to typically achieve photorealism or visual pleasing image, or both. This leads to the development of Tone-Mapping Operator (TMO). Over the years, many researchers have proposed numbers of TMO [10], [4], [2], [11], [6], [5], [8]. In general, these TMOs can be classified into four main categories based on the processing technique they use. These categories are: Global, Local, Frequency-based, and Segmentation-based operators.

Global operators apply the same operator to all pixels of the image. These operators typically result in tone-mapped images that are photorealistic, however since they do not preserve local contrast, the tone-mapped images lack local contrast and lost finer details of the original HDR, making the image look flat.

Local operators, on the other hand, do not suffer from local contrast lacking, since the operators take into account the spatial context. The tone-mapped images are locally boost and tend to be preferred by observers. However, this often leads to the unrealistic output, due to the spatial artifacts such as halos in the area around high contrast edges.

Frequency-based operators, as the name imply, apply the operator in the frequency domain (instead of spatial domain). The idea is to separate details (high frequency) from large features (low frequency). This results in the tone-mapped image that are similar to the local operators. In addition, both local and frequency-based operators typically consume much more computational costs compared to the global operators.

The newer approach to tone-map is the operator that is based on the concept of image segmentation. Segmentation-based operators divide the HDR image into regions and separately calculates an adaptation luminance for each region, and finally merge them to get the final output image.

In the next section, the proposed TMO that is motivated by the segmentation-based operator is discussed. The idea is to use the segmented high-passed component that is resulted from unsharp-masking algorithm, to enhance the flat look of the global tone-mapped image. The reason that we choose unsharp masking algorithm because it is fast and consumes less consume computational costs compared to other local operators, and have been used for many years in the publishing industry to sharpen images.

ALGORITHM

First, the operator globally compresses the luminance of the input HDR. The compressed luminance $L_{(x,y)}$ is computed by:

$$L(x,y) = \frac{a}{\bar{L}_{scene}} L_{scene}(x,y) \tag{1}$$

where a is the key value that control the key of the image and are bounded in the 0 to 1 interval. For normal key, $a = 0.18$. $L_s(x, y)$ is the scene luminance for pixel (x, y), $\bar{L}_s$ is the log average of the scene luminance that is computed by:

$$\bar{L}_{scene} = \frac{1}{N}\exp\left(\sum_{x,y}\log(L_{scene}(x, y) + \delta)\right) \tag{2}$$

where N is the total number of pixels of the image and δ is a small constant to avoid $\log(0)$. Next, the display luminance $L_d(x, y)$ is computed by applying a following equation:

$$\bar{L}_d(x, y) = \frac{L(x, y)\left(1 + \frac{L(x, y)}{L_{white}^2}\right)}{1 + L(x, y)} \tag{3}$$

where L_{white} is the luminance that intend to be mapped to the white in the display image. Also note that this equation guaranteed to spread all luminances within the 0 to 1 interval.

The above operation is similar to the global part of the TMO proposed by [7].

As can be guessed, the above operation often generates a flat looking image due to nature of the global operation.

In this work, to handle with this problem, the modified version of equation (3) that is based on the segmentation approach is introduced.

Specifically, the unsharp masking with the concept of the segmentation is utilized here. The algorithm sharpens the compressed luminance $L_{(x,y)}$ by subtracting a blurred version of the compressed luminance. Here, the unsharp masking is defined as:

$$\hat{L}(x, y) = L(x, y) - \check{L}(x, y) \tag{4}$$

where $\hat{L}(x, y)$ denotes the high-passed component obtained by the algorithm, and $\check{L}(x, y)$ is a blurred version (low-passed component) of $L(x, y)$.

In addition, the amount of edge emphasis desired can be controlled with a concept of segmentation by introducing a threshold to the high-passed component $\hat{L}(x, y)$ as:

$$\hat{L}\begin{cases}\hat{L}(x, y) & \text{if } \hat{L} \geq T \\ 0 & \text{if } \hat{L} < T\end{cases} \tag{5}$$

where T is the threshold value.

This high-passed component later is added back to the compressed luminance $L_{(x,y)}$ to obtain the sharpen image $L_{sharp}(x, y)$.

$$L_{sharp}(x, y) = L(x, y) + \hat{L}(x, y) \tag{6}$$

Equation (3) are now becomes:

$$\bar{L}_d(x, y) = \frac{L_{sharp}(x, y)\left(1 + \frac{L_{sharp}(x, y)}{L_{white}^2}\right)}{1 + L(x, y)} \tag{7}$$

where $L_{sharp}(x, y)$ is the sharpen image resulted from the unsharp masking algorithm presented in the previous section.

In this way, the operator applies a local operator that is simple yet powerful enough to boost the details (edges) of the image only in the area that is specified by the threshold value T resulting in the tone-mapped image that is photorealistic, detail-preserved, and free from spatial artifacts. Figure 1 compares the outputs resulted from Equation (3) and Equation (7). The output images contains no halos artifact due to the benefit of unsharp-masking and segmentation-based approach. To demonstrate that the proposed TMO works well on a wide range of images, Figure 2 illustrates a selection of tone-mapped images using the proposed operator. We call this operator the Fast Segmentation-based TMO.

Figure 1: An example of the proposed operator. (left) Output of Equation (3). (right) Output of Equation (6). (Original image acquired from [1].)

Figure 2: Shows some of the output images.

EXPERIMENTAL

To evaluate the image quality of the Fast Segmentation-based TMO, two two-alternative forced choice (2AFC) psychophysics experiments (Engeldrum 2000) have been conducted. The goal is to compare the output of the proposed TMO with four well-known TMOs to generate preference scores for photorealism and image preference. Table 1 lists five TMOs that have been used in the experiment.

Table 1. TMOs used in the experiment

Name	Type	Label
Fast Segmentation-based (the proposed operator)	Segmentation-based	F
A Visibility Matching Tone Reproduction Operator [4]	Global Operator	V
Photographic Tone Reproduction [7]	Local Operator	P
Fast-Bilateral Filtering [2]	Frequency Domain	B
Segmentation and Adaptive Assimilation [11]	Segmentation-based	S

There are 60 tone-mapped images in total (five operators, twelve HDR scenes). These images were evaluated by 20 subjects with normal color vision, naïve to the goal of the experiment. The experiments were conducted in the controlled viewing environment. Also note that to avoid unfaithful implementation and bias from parameter selection, when possible the images used in the experiment were obtained from the author's online resources.

There was a total of 240 pairwise comparisons in the experiments (two experiments, eight scenes, 10 pairs of operators per scene). The subjects were asked to participate in two experiments. Participants were asked to make judgements based on photorealism and image preference, respectively. Specifically, for each pair, participants were instructed to observe the two tone-mapped images and select the one that correlates to the corresponding criteria the most. Images were shown randomly on the left or the right of the screen. All operators were shown roughly an equal number times left and right. The whole procedure took approximately 17-22 minutes. To evaluate the results, Thurstone's Law of Comparative Judgement Case V [9] is used to generated preference scores.

RESULTS AND DISCUSSION

The average preference scores of the experiments are given in Figure 3. The error bars represent the 95% confidence interval (For 20 observers the error bar is 0.3). To interpret the result, the preference score (y-axis) can be interpreted as a z-score, thus, if one operator were strongly preferred, the score should be 2. On the other hand, if the operator were strongly dislike, the score should be close to -2. In cases where operators do not deliver strong preference, the scores should be cluster around 0. If the error bars do not overlap then one operator is said to have higher preference than another at the specified confidence interval (95% in this case).

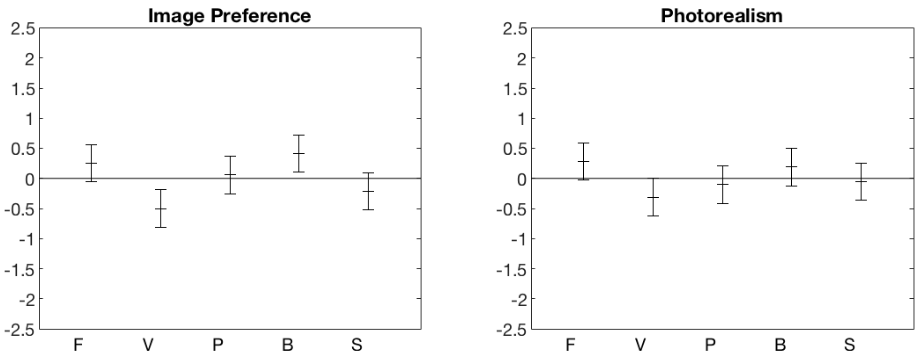

Figure 3: Preference scores of 5 different TMOs for image preference (left), for photorealism (Right).

For the image preference, the result indicates that B operator has the highest score, follow by F (the proposed TMO), P, S, and V. For the photorealism, F (the proposed operator) has highest preference scores among all the tested operators, follow by B, S, P, and V. The scores of the two experiments share a similar pattern, indicating a strong correlation between the two criteria the participants were made. Also note that, most of the operators (except V operator) have the scores cluster around 0, and their error bars are also overlap, indicating that none of them performs better than the others.

CONCLUSION

We have developed a fast segmentation-based TMO that incorporates the use of the well-known sharpening algorithm called unsharp-masking. The proposed TMO realistically reproduces tone-mapped images that are both photorealism and visually pleasing, making it well-suited for applications where a photorealistic, artifact-free images and limited computational resources available is the goal.

REFERENCES

[1] Banterle F., A. Artusi, K. Debattista & A. Chalmers 2011. *Advanced high dynamic range imaging: theory and practice*.

[2] Durand F. & Dorsey J. 2002. *Fast bilateral filtering for the display of high-dynamic-range images*. ACM transactions on graphics (TOG), 257-266.

[3] EngeldrumM P.G. 2000. *Psychometric scaling: a toolkit for imaging systems development*. Imcotek.

[4] Larson G.W. & Rushmeier H. 1997. *A visibility matching tone reproduction operator for high dynamic range scenes*. IEEE Transactions on image processing.

[5] Mantiuk R. & Seidel H.P. 2008. *Modeling a Generic Tone-mapping Operator*. Computer Graphics Forum.

[6] Meylan L. & Susstrunk S. 2006. *High dynamic range image rendering with a retinex-based adaptive filter*. IEEE Transactions on image processing.

[7] Reinhard E., Stark M., Shirley P. & Ferwerda J. 2002. *Photographic tone reproduction for digital images*. ACM transactions on image processing.

[8] Singnoo J. & Finlayson G.D. 2012. *Optimal global approximation to spatially varying tone mapping operators*. Conference on Colour in Graphics, Imaging, and Vision.

[9] Thurstone L.L. 1927. *A law of comparative judgment*. Psychological review.

[10] Tumblin J. & Rushmeier H. 1993. *Tone reproduction for realistic images*. IEEE Computer graphics and and Applications.

[11] Yee Y.H. & Pattanaik S. 2003. *Segmentation and adaptive assimilation for detail-preserving display of high-dynamic range images*. The Visual Computer.