SELECTIVE LOCAL TONE MAPPING

A. Artusi[1], *A. O. Akyüz*[2], *B. Roch*[3], *D. Michael*[4], *Y. Chrysanthou*[4], *A. Chalmers*[5]

[1]IMA, GiLab, University of Girona, Girona, Spain
[2]Middle East Technical University, Ankara, Turkey
[3]Vienna University of Technology, Vienna, Austria
[4]University of Cyprus, Nicosia, Cyprus
[5]University of Warwick, Coventry, UK

ABSTRACT

When preparing high dynamic range images (HDR) for display on standard monitors, it is often necessary to make a choice between *global* and *local* tone mapping. While the former is simple and efficient, it may fail to reproduce details in high contrast image regions. Although, the latter can better reproduce details in such regions, it often comes at the cost of increased complexity and computational time. In this paper, we present an algorithm that combines the best of both approaches. We perform local tone mapping only in high frequency image regions where the visibility of details can be an issue. In low frequency regions, we employ global tone mapping to save computational resources without degrading quality. Our algorithm is most suitable for tone mapping operators (TMOs) that utilize the concept of local adaptation luminances.

***Index Terms*—** Tone mapping, high dynamic range imaging, GPU, real-time.

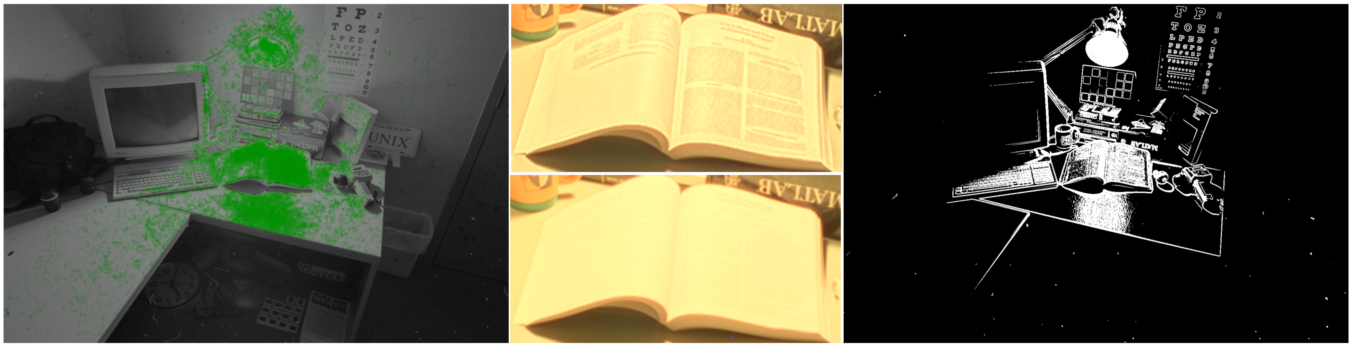
1. INTRODUCTION AND MOTIVATION

With increased use of HDR images in various application areas and relatively slow progress in HDR display technology, tone mapping remains to be a very important problem. During tone mapping, the user is generally presented with two choices: (1) use a fast and simple global operator which, in overall, conveys the general appearance of the scene but might lose small scale details; (2) use a more complex and computationally slower local operator, which preserves small scale details but may fail to convey the overall appearance as good as the global operator. We argue that this choice can be circumvented if one combines the strengths of both types of operators while avoiding their weaknesses. To understand how these two types of operators can best be combined we analyzed various HDR images tone mapped with global and local operators. We focused on which parts of the images exhibit the greatest difference when tone mapped with global and local operators. To this end, we used the dynamic range independent image quality assessment metric (DRIM) on pairs of globally and locally tone mapped images [1]. Figure 1 (a) depicts a representative result of our analysis. Here, the green pixels indicate the regions where contrast is *visibly* lost by global tone mapping but would be preserved by local tone mapping. This is confirmed by the close-up view in (b) where the top image shows the local result and the bottom image the global result. The gray pixels in (a) indicate that a human observer would not be able to perceive contrast changes between these two types of operators. Using the local operator in these regions would not bring additional benefits over using the global one. One can use the DRIM to determine the image regions which would benefit from local tone mapping. However, this would be counter-productive since the run time cost of this metric generally outweighs that of a local operator. As such, one needs to find a simpler function that approximates the output of the DRIM, but has a much lower computational cost. After experimenting with several such functions, we found that the Sobel edge detector [2] applied on the original HDR image yields a surprisingly similar output to that of the DRIM (Figure 1 (c)). Two further results supporting this finding are shown in Figure 2. The similarity may be explained by the fact that the sobel filter is detecting gradients which is a measure of contrast and DRIM itself has a contrast detection process. Based on these observations, we decided to identify the image regions that require local treatment using the Sobel edge detector.

This work was partially supported by the Spanish Ministry of Science and Innovation Sub-Program Ramón y Cajal RYC-2011-09372, the Catalan region CONES-2010-0027, EU Cost Action IC1005 and Cyprus Research Promotion Foundation grant STOQOS/0609/09.

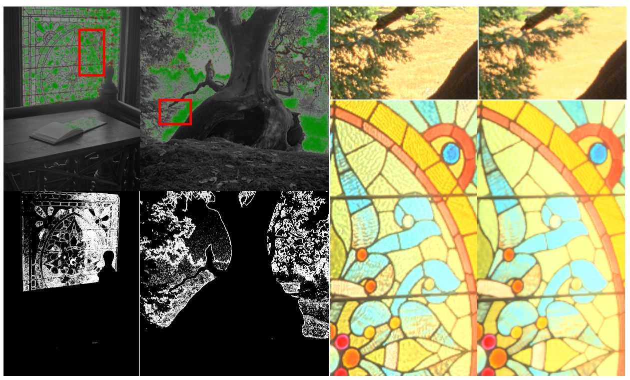
2. RELATED WORK

In the past several TMOs have been presented, and an extensive review can be found in [3, 4]. Hardware implementation of TMOs, tightly coupled with the current graphic hardware and FPGA's, have been presented in the past as well [5, 6, 7, 8]. In this section we will review only the works that are closely relevant to our study. The idea of using local adap-



(a) DRIM (b) Close-up view (c) Sobel Edge Map

Fig. 1. (a) DRIM [1] result where green regions indicate visible loss of contrast due to global tone mapping. (b) edge detection result using the Sobel operator. We used the 85^{th} percentile of gradient magnitudes for marking edges.



Local Global

Fig. 2. Comparison of the DRIM vs Sobel edge detection. Note that while the outputs are not identical there is a high degree of similarity. The close-ups show the difference between local and global tone mapping in marked regions.

tation luminance for tone mapping has been explored in several studies. These approaches either use difference of Gaussians as a measure of uniformity [9, 10, 11] or image segmentation [12, 13, 14, 15, 16] to localize the largest possible neighborhood around each input pixel. Then the local adaptation luminance is computed as the average luminance of this neighborhood. These techniques are not computationally efficient due to the use of complex filtering or segmentation approaches and in some cases they require to solve large linear systems [13]. Segmentation suffers from banding artifacts due to the need of precisely identifying the transition between bright an dark areas [14]. This can be eliminated, but further processing is required [14, 15, 12]. Efficient uses of bilateral filtering to the tone mapping problem have been presented [17, 18]. However, they may have memory management problems [17] and show artifacts when the bilateral filter is behaving as a pure Gaussian filter [18].

3. ALGORITHM

Our algorithm is comprised of three stages namely *identification*, *marking*, and *tone mapping*.

3.1. Identification

This step identifies the image regions that require local tone mapping. We simply define these regions to be the high frequency regions as detail loss due to global tone mapping is most severe in these regions. Based on our experiments, we find that a sophisticated algorithm for detecting such regions is not required; it is possible to use a simple edge detector for this purpose. Therefore, we use the Sobel edge detector applied on the luminance channel of the input image. Once the gradient magnitude map is generated, we threshold this map to only keep the strong gradients. The value of the threshold varies between 0 and the highest gradient magnitude computed by the Sobel filter; 0 means that the local operator is applied on all pixels of the input image. Increasing the threshold level reduces the number of pixels that are marked for local treatment (Figure 3). In practice, setting this threshold around the 85^{th} percentile of the gradient magnitudes yields good results.

3.2. Marking

Marking step is only necessary in the GPU implementation of our algorithm in order to discontinue the processing of the non-identified pixels in the rendering pipeline as early as possible. In theory, this step requires a simple conditional constructor. We have experimented with several approaches such as *branching* and *texkill* instructions, a tiling approach, and using the early Z test. We found that using the early Z buffer gives the best solution as it needs no additional logic and prevents processing of all fragments that fail the early Z test. To accomplish this, the output of the identification step is written to the depth buffer and an early Z test is performed to mask

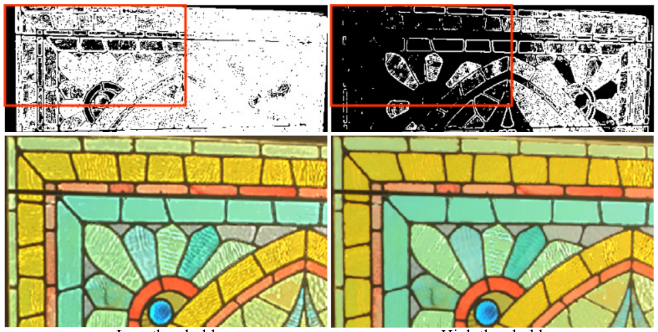


Fig. 3. Effect of the identification threshold: (left) 0.1 - (right) 1.0. Bottom row, the more pixels identified results in better reproduction of details.

out fragments which are not identified as high contrast[1]. To make this step more efficient, we make use of hi-Z, where the depth test is conducted on a tile basis.

3.3. Tone Mapping

The shader that is responsible for computing the local adaptation luminance makes use of the early Z buffer, and it is applied only on the pixels that pass the early Z test. The output of the complex shader is the computed local adaptation luminance. Afterwards, the global tone mapping curve is applied on all pixels by using the adaptation luminance computed for the identified pixels and the original luminance for the others.

4. RESULTS

We have performed a series of quality experiments between the ground truth and the output of our approach using different types of quality metrics. The ground truths are obtained by applying local luminance adaptation on the full image.

4.1. VDP and SSIM

The visual difference predictor (VDP) metric attempts to develop a general visual model for complex images [19]. As output it gives the percentage of pixels that are perceivably different according to the human visual system (HVS). Based on the fact that the HVS has the property to highly adapt to extract structural information from the viewing field, we have also used the structural similarity index (SSIM) as another metric in our comparison [20]. SSIM reports the degree of similarities between the two images in the range $[0, 1]$ with 1 meaning perfect similarity and 0 complete difference. Table 1 presents the results of applying these two metrics to compare our results with the ground truths. In the case of the VDP metric, the numerical values are percentage values. As it shown

[1]It is a hardware optimization that is enabled through the use of the depth buffer.

HDR Image	VDP	SSIM
16RPP	0.04	0.9968
Memorial	2.41	0.9876
Nave	1.47	0.9761
Rosette	0.38	0.9982
Belgium	0.21	0.9853
Desk	2.72	0.9925
FogMap	0.07	0.9925

Table 1. Results of the quality comparison using the VDP and SSIM. The HDR images are taken from http://www.anyhere.com/gward/hdrenc/pages/originals.html.

Fig. 5. Comparisons: (left) ground truth, (center) our and (right) DRIM outputs.

in the Table 1, this value is between 0.04 and 2.72 percent. This represents a very limited number of pixels that are perceived as different by the HVS with respect to the ground truth. On the other hand, the SSIM shows that the measure of the structural information change is minimal in all cases. This confirms the high degree of similarity between our results and the ground truth.

4.2. Dynamic Range Independent Quality Assessment

Recently Aydin et al. [1] introduced a novel metric, called DRIM which is capable of operating on image pairs of images with different dynamic ranges. This metric uses a model of the HVS and introduces a new definition of the visible distortion based on the detection and classification of visible changes in the image structure. Different types of changes measured by this metric are visualized using different colors: *green* for loss of visible contrast, *blue* for amplification of invisible contrast, and *red* for reversal of visible contrast. The results obtained by using the publicly available implementation of this metric are shown in Figures 5. As this figure shows the results of our algorithm and the ground truth are very similar except a few small isolated regions. Finally, in Figure 4 we

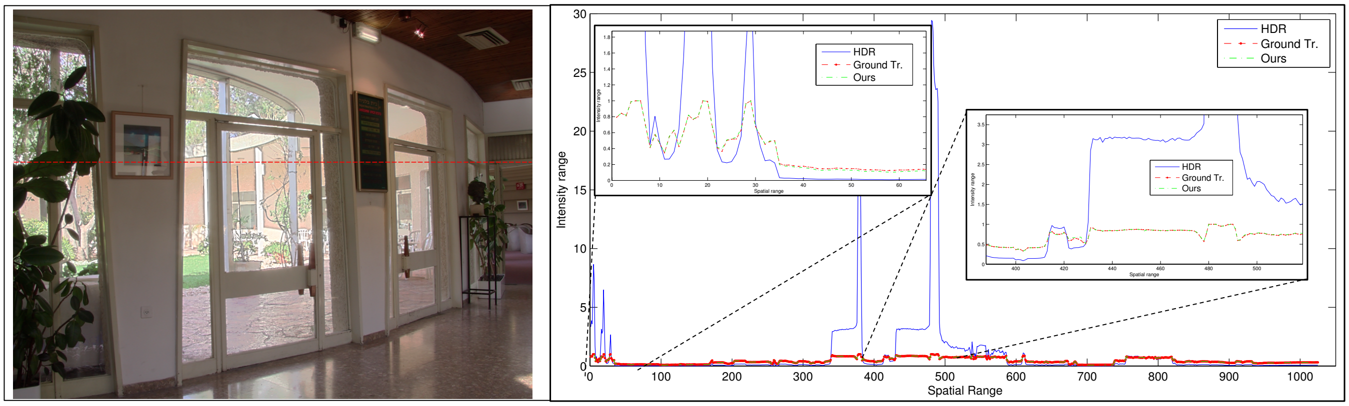


Fig. 4. Intensity profile comparison of the ground truth (red), our approach (green) and of the HDR input image (blue) for scanline 300. The close-ups show columns 1-80 and 380-520 respectively.

compare the intensity profiles of the proposed method (green) with the ground truth (red). The close-ups show the similarity of the scanline profiles between the ground truth and the proposed method.

4.3. Timing Performances

Next, we show the computational performance results. First, we show the run time in frame per second (FPS) of our results varying the edge threshold value and the input frame resolution in Table 2. Second, we show the run time of the proposed method compared with the ground truth (full local operator), and the global operator varying the input frame resolution in Table 3. We used Ashikhmin's TMO to generate both tables [9]. The HDR image used is shown in Figure 6. Both experiments were run on an Nvidia GTX560M. As can be seen from these tables, by using our method, one can reach interactive frame rates even for very high resolution HDR images. As expected the performance of our algorithm lies between that of the global and local methods.

Frame	97^{th}	87^{th}	77^{th}
5097x2889	6	5	5
1/2	23	21	18
1/4	72	68	55
1/8	207	194	167

Table 2. Run time of the proposed method (in FPS) for different image resolutions and edge thresholds (in percentile).

Fig. 6. Image used for the performance experiments.

Frame	Global	Ground truth	Ours
5097x2889	12	1	5
1/2	37	7	18
1/4	136	25	55
1/8	420	87	167

Table 3. Run time varying the resolution of the input frame for the global, local and our operators. The edge threshold value used for our operator was 77^{th} percentile.

5. CONCLUSIONS

We developed a novel technique to address the preservation of local contrast and details while keeping lower computational costs and complexity. Local tone mapping operators tend to be applied on the whole image regardless the information contained in the HDR input image. As a result the computational costs are prohibitive for very large HDR images and textures. Our algorithm is based on the simple concept of detecting high frequency image regions and applying local tone mapping only in these regions. This allows us to save computational resources without degrading image quality. In our experiments, we have not seen any need for blending between locally and globally tone mapped regions. We attribute this to the fact that we use the same TMOs local and global components. If one mixes and matches different operators, blending could be needed. In this case, smoothing of the edge map by a low-pass filter could be satisfactory.

6. REFERENCES

[1] Tunç Ozan Aydin, Rafał Mantiuk, Karol Myszkowski, and Hans-Peter Seidel, "Dynamic range independent image quality assessment," *ACM Trans. Graph.*, vol. 27, no. 3, pp. 69:1–69:10, Aug. 2008.

[2] M Sonka, V Hlavac, and R Boyle, *Image Processing, Analysis, and Machine Vision*, Brooks Cole, 1998.

[3] Erik Reinhard, Greg Ward, Sumanta Pattanaik, and Paul Debevec, *High Dynamic Range Imaging: Acquisition, Display and Image-Based Lighting*, Morgan Kaufmann, San Francisco, second edition edition, 2010.

[4] Francesco Banterle, Alessandro Artusi, Kurt Debattista, and Alan Chalmers, *Advanced High Dynamic Range Imaging: Theory and Practice*, CRC Press, (AK Peters Ltd), 2011.

[5] A.Artusi, J. Bittner, M. Wimmer, and A. Wilkie, "Delivering interactivity to complex tone mapping operators," in *Proceeings of Eurographics Symposium on Rendering 2003*, 2003, pp. 38–44.

[6] N. Goodnight, R. Wang, C. Woolley, and G. Humphreys, "Interactive time-dependent tone mapping using programmable graphics hardware," in *EGSR03, Eurographics Symposium on rendering 2003*, 2003, pp. 26–37.

[7] B. Roch, A. Artusi, D. Michael, Y. Chrysanthou, and A. Chalmers, "Interactive local tone mapping operator with teh support of graphics hardware," in *Proceedings of the Spring Conference in Computer Graphics, Proceedings SCCG 2007*, 2007.

[8] Chris Kiser, Erik Reinhard, Mike Tocci, and Nora Tocci, "Real-time automated tone mapping system for hdr video," in *Proceedings of IEEE International Conference on Image Processing*, 2012.

[9] Michael Ashikhmin, "A tone mapping algorithm for high contrast images," in *Rendering Techniques '02 (Proceedings of the 13th Eurographics Workshop on Rendering)*, 2002, pp. 145–156.

[10] Erik Reinhard, Michael Stark, Peter Shirley, and Jim Ferwerda, "Photographic tone reproduction for digital images," in *Proceedings of SIGGRAPH 2002*, 2002, pp. 267–276.

[11] Sumanta Pattanaik and Hector Yee, "Adaptive gain control for high dynamic range image display," in *SCCG '02: Proceedings of the 18th spring conference on Computer graphics*, New York, NY, USA, 2002, pp. 83–87, ACM.

[12] Hector Yee and Sumanta Pattanaik, "Segmentation and adaptive assimilation for detail-preserving display of high-dynamic range images," *The Visual Computer*, vol. 19, no. 7/8, pp. 457–466, December 2003.

[13] Dani Lischinski, Zeev Farbman, Matt Uyttendaele, and Richard Szeliski, "Interactive local adjustment of tonal values," *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH 2006)*, vol. 25, no. 3, aug 2006.

[14] Grzegorz Krawczyk, Karol Myszkowski, and Hans-Peter Seidel, "Lightness perception in tone reproduction for high dynamic range images," in *The European Association for Computer Graphics 26th Annual Conference EUROGRAPHICS 2005*, Dublin, Ireland, 2005, vol. 24 of *Computer Graphics Forum*, pp. 635–645, Blackwell.

[15] M. Cadik, "Perception motivated hybrid approach to tone mapping," in *Proceedings WSCG 2007*, 2007.

[16] Francesco Banterle, Alessandro Artusi, Elena Sikudova, Thomas Bashford-Rogers, Patrick Ledda, Marina Bloj, and Alan Chalmers, "Dynamic range compression by differential zone mapping based on psychophysical experiments," in *Proceedings ACM Symposium on Applied Perception (SAP)*, 2012.

[17] Jiawen Chen, Sylvain Paris, and Fredo Durand, "Real-time edge-aware image processing with the bilateral grid," in *Proceedings ACM Siggraph 2007*, 2007.

[18] Francesco Banterle, Massimiliano Corsini, Paolo Cignoni, and Roberto Scopigno, "A low-memory, straightforward and fast bilateral filter through subsampling in spatial domain," *Computer Graphics Forum*, vol. 31, no. 1, pp. 19–32, February 2012.

[19] S Daly, "The visible difference predictor: an algorithm for the assessment of image fidelity," in *Digital Images and Human Vision*, A B Watson, Ed. 1993, pp. 179–206, MIT Press.

[20] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assesment: From error visibility to structural similarity," *IEEE Transaction on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.