VISUALIZATION OF HIGH DYNAMIC RANGE IMAGES

*Alvaro Pardo**

IMERL & IIE
Facultad de Ingeniería
Universidad de la República
C.C. 30, Montevideo, Uruguay

Guillermo Sapiro[†]

Electrical and Computer Engineering
University of Minnesota
Minneapolis, MN 55455

ABSTRACT

A novel paradigm for information visualization in high dynamic range images is presented in this paper. These images, real or synthetic, have luminance with typical ranges many orders of magnitude higher than that of standard output devices, thereby requiring some processing for their visualization. In contrast with existent approaches, which compute a single image with reduced range, close in a given sense to the original data, we propose to look for a representative set of images. The goal is then to produce a minimal set of images capturing the information all over the high dynamic range data, while at the same time preserving a natural appearance for each one of the images in the set. A specific algorithm that achieves this goal is presented and tested on natural and synthetic data.

1. INTRODUCTION

High dynamic range (HDR) images contain a wide range in luminance, many times in the order of tens of thousands different values. These images could be natural, obtained for instance from multi-exposure photographs [1] or with a multiple exposure sensor [2, 3], or synthetic, in the case of computer graphics applications. These images have ranges that greatly exceed that of the output device. The question is then how can we *reproduce* and *visualize* such images in a standard output device.

Let us introduce some basic terminology. The *scene* is the real or synthetic picture we perceive without involving any output device between it and our eyes. An *image*, on the other hand, is what we can see using the output device, or its internal computer representation as an array of digital values. The key problem is how to translate from scenes

to images while preserving the relevant scene information, producing a natural looking image, and avoiding common artifacts such as *halos* (which are due to gradient reversals [4]).

HDR images can be applied in computer graphics for the production of synthetic images with realistic or hyper-realistic appearance. Another application is HDR photo graphs, which are able to capture much more detailed scene information than standard photographs. These systems make it possible to capture a highly detailed range representation of the scene and later process the data in order to select the image(s) that better fulfils the given requirements. These images could also improve computer vision and image analysis algorithms that usually rely on limited range data. This is particularly relevant in scenarios where we do not have complete control over the illumination. Therefore, there is a need to develop algorithms to perform the translation from scenes to images, algorithms such as the discussed and presented in this paper.

We can classify the existing approaches for the translation from scene to image in two main groups. The first group consists of algorithms that map the original range to the output range while attempting to preserve the subjective perception of the scene. Among the works which propose a tone mapping we have [7] and [5]. Larson *et. al.* [5] proposed a global tone mapping operator which adjusts the histogram of the scene based on psychophysical models for color, glare, and acuity perception. The results of this simple and elegant method, which reduces to just a global map, are very good and with high fidelity to the subjective perception of the scene. Although this idea of "tone mapping" works quite well, it has some caveats. First, it is not able to reproduce all the details present in the scene. Second, the method breaks down when the input range is too wide compared with the available output range. Note that since these methods try to preserve the original perception of the scene, details that are hard to see in the original scene will be difficult to see in the output image as well.

In the second group, we have algorithms that favor the visualization of details instead of the subjective perception

of the scene [6, 9]. They apply a multiscale decomposition to discriminate between illumination and details. The main problem with this idea is that, although correct in theory, it usually introduces halos in the output image. The local mappings produced by these techniques violate the basic monotonicity principle, i.e. the pixel value order is not necessarily preserved and darker (brighter) regions in the scene might become brighter (darker) in the image. Moreover, these approaches tend to have a large number of parameters, which are generally hard to control in an automatic fashion. In [8] a method for displaying synthetic images is presented, based on the segregation of lightning and surface layers.

The two groups of works just described are not completely equivalent, as they address slightly different problems. Furthermore, although similar, their solutions cannot be easily compared; one representation tries to capture the subjective appearance of the scene under the limitation of the output/display device, while the second group attempts to preserve the scene details. While enhancing details, we may be adding information not perceptually present in the original scene. On the other hand, while preserving visual appearance, details might be omitted.

1.1. Our contribution

Our proposed paradigm attempts to have the best of both groups mentioned above. We propose a method which captures the image details while preserving the natural appearance of the scene. As we will explain below, there is an intrinsic limitation in representing an HDR image with only *one standard* output image. Sometimes it is practically impossible to find an output image containing all the relevant information in the HDR image that represents the scene. For this reason, we argue for a method to obtain a *set of images* containing all the relevant information of the original scene. We could say that while the algorithms described above deal with the *reproduction* of the scene, the technique here proposed deals with its *visualization*. Moreover, we argue that not only the set of images has to accurately visualize the relevant information present in the scene, but also has to do it in a visually pleasant form. In the proposed algorithm, each image is produced by a different monotonic global map, thereby avoiding gradient reversals typical of the local schemes. The locality is achieved by letting this global map "stretch" different regions for each one of the images in the set.

2. THE METHOD

Our idea is to propose a simple and effective method to visualize *all* the information in the scene in a *pleasant* way. *All* means that we would like to capture as many details as possible, and *pleasant* means a procedure which appears

natural to the observer.

Let us assume we have a scene with dark and bright areas, and details all over it. If we want to visualize all the details we need first to have a minimum resolution available (number of output levels), and second to be able to "see" in every region. That is, in dark areas we must add light and reduce it on the bright ones. In both cases, the information is out there but it cannot be seen due to under-exposure or over-exposure. To solve this, we are going to modify the *luminance of the scene to capture all the details over it*. All the proposed operations are simply contrast changes. In this way artifacts such as halos are not introduced. Furthermore, this idea resembles the control of illumination during acquisition and is perceived as natural.

If we illuminate a given region, we will be stretching its output range, thereby using more output levels. On one hand we display this region with good light and resolution, while on the other hand we might be compressing and missing details in other regions. Hence, there is clearly a competition between the output fidelity of different regions, and unfortunately, it is difficult or impossible to find a satisfactory solution with a single output. To overcome this, we propose to generate a sequence of images with different resolutions in each scene region (space varying resolution). The basic idea is then to distribute the existing resources among different output images. This sequence could be either observed as a movie or as a set of still images could be extracted from it. The key point here is that for many applications more than one output image is a reasonable solution.

2.1. Outline of the algorithm

Before presenting our proposed algorithm, let us give some basic notation. (r, g, b): Input color primaries, L : Input luminance, $[L_w^m, L_w^M]$: Input luminance range, L': Modified luminance, $[L_d^m, L_d^M]$: Output luminance range, (R, G, B): Digital output values. We are now ready describe the different steps of the algorithm.

1-Compute image luminance: From the (r, g, b) primaries compute the luminance L (in cd/m^2) and the color information $(r/L, g/L, b/L)$. We process the luminance while preserving the color information.

2-Segment the image: Divide the image into two or more regions of interest. This is achieved splitting the histogram (see Section 3) into sub-intervals. From now on, we assume two regions, $[L_w^m, L_{w^*}]$ and $[L_{w^*}, L_w^M]$.

3-Modify the luminance: Apply Larson's histogram adjustment algorithm [5] to each interval. Map $[L_w^m, L_{w^*}]$ to $[L_d^m, L_{d^*}]$ and $[L_{w^*}, L_w^M]$ to $[L_{d^*}, L_d^M]$. The important point here is that when using the human contrast sensitivity function, the mapping does not produce a contrast greater than the one present in the original scene. Since we are working in regions, this means controlling the contrast over that given region. This step is not a traditional histogram

adjustment since it modifies the output range and the distribution within it.

If we select L_{d^*} close to L_d^M we will be visualizing the dark areas with a wider range than the bright ones. A first solution to the problem of visualization HDR data is to create a movie by increasing L_{d^*} from L_d^m to L_d^M. This gives us a sequence, which starting from the image with all resources allocated to the dark areas, slowly moves to an image with all resources allocated to the bright areas. This is just a nice way of visualizing all the information. A second possibility is to choose just a certain fixed number of images. We discuss below how to select these images. In the case of three intervals or more, the idea is the same. We start with most resources allocated to the first interval and we swap them to the next interval to the right. For details, see www.ece.umn.edu/users/guille/hdr.html.

4-Quantization: Quantize and gamma correct the recomputed primaries ($\frac{r}{L} * L'$, $\frac{g}{L} * L'$, $\frac{b}{L} * L'$) to obtain the digital output values (R, G, B).

2.2. Information assessment

If we consider each image as a message, its information can be measured with the entropy function. With it, we will be able to extract the best or set of best images in the sequence. From the histogram of the output image we find the probabilities of each output level, and with them we compute the entropy. Note that maximum entropy implies a flat histogram. Additionally, entropy maximization captures our subjective preference towards well-contrasted images.

3. ESTIMATION OF THE NUMBER OF IMAGES

The minimal number of images needed for a satisfactory visualization of all the details in the HDR data depends on the particular image being processed. To obtain an estimation of the number of images we use the number of clusters in the luminance histogram. To display the image without loss of information we would need, roughly, as many output levels as clusters present in the image. However, not all output levels can be used if we want to distinguish between them. We use an empirical rule of thumb that says that we can allocate around 200 clusters per image in the set. This is just a coarse estimation that works fairly well in practice.

Having an estimation of the number of output images required to satisfactory visualize the scene information, a segmentation is needed in order to find the location of the luminance intervals. To segregate dark and bright areas, a threshold t that maximizes the sum of the entropy of both distributions is computed. Using the histogram of the luminance and computing the above probabilities we find the threshold t that optimally segments dark and bright areas. If we use only two images, we only need to apply this pro-

cedure once; otherwise, we iterate this scheme. This procedure can be seen as a binary tree construction, where in each step an interval is divided in two. At the end, the user can select the important leaves from it if user interaction is constructed in the process.

4. RESULTS

We now present the results of our proposed algorithm. No user intervention was needed and the number of images and segmentation was performed as explained above. Movies showing the full set for this example and others can be found at www.ece.umn.edu/users/guille/hdr.html.

The first step is the segmentation of the original image. We found 419 clusters in the image luminance and therefore we use three output images (since we consider no more than 200 clusters per image). The intervals where found applying twice the algorithm described in Section 3. First, we obtained a segmentation of bright and dark areas and then we sub-segmented again the dark interval. Second, we assign the given resources to each region. In this case we assumed $L_d^M = 100 \ cd/m^2$ and $L_d^m = 2 \ cd/m^2$, and selected initial ranges of 86 cd/m^2 for the first interval and 6 cd/m^2 for the rest. For the next image in the set, the resources are re-allocated moving them from left to right in steps of 5 cd/m^2. After segmentation we obtained the following intervals: $[0.12, 2.27]$, $[2.27, 83.09]$ and $[83.09, 40416.77]$.

In Figure 1 we show the results obtained with our algorithm (each image represents the image with maximal entropy per region) together with the result of Larson's algorithm. With these images we can see more details than in the single image processed with Larson's algorithm in a fairly natural way. Hence, we managed to display the original scene in a set of images, making visible some information that was obscured in the original scene.

Since our method extends Larson's one, among all the images in the set, there should be one image close to the one obtained with Larson's algorithm. In other words, if we resign to use more than one image and visualize all the details in the scene, we should be able to extract a single satisfactory one from the computed set. To do that we take the image that maximizes the global entropy (see Figure 1).

5. CONCLUSIONS

In this paper, we have presented a new paradigm for the reproduction and visualization of information in HDR images. We argued for the use of a set of images instead of a single one as in traditional approaches.

More than being the last word about the problem of visualizing HDR data, with this work we attempted to illustrate the intrinsic limitation of working with only one image. We showed how going for more than one image we could

obtain a simple and nice solution to the problem of complete information visualization of high HDR images.

A number of questions remain open after this work. The specific algorithm here described for the computation of the set of images is just a particular example, and others should be developed. One of the crucial additional points is how to find the minimal number of images required to visualize all the relevant information. These images have also to be pleasant and hopefully with smooth transitions among them. We hope that the work here presented will open the door to works on these and other relevant questions.

6. REFERENCES

[1] P. Debevec and J. Malik, "Recovering High Dynamic Range Radiance Maps from Photographs," in *SIG-GRAPH 97*, 1997, pp. 369–378.

[2] S. Nayar and T. Mitsunaga, "High Dynamic Range Imaging: Spatially Varying Pixel Exposures," in *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, Jun. 2000, vol. 1, pp. 472–479.

[3] B. Wandell, P. Catrysse, J. DiCarlo, D. Yang, and A. Gamal, "Multiple Capture Single Image Architecture with a CMOS Sensor," in *International Symposium on Multispectral Imaging and Color Reproduction for Digital Archives. Japan.*, 1999, pp. 11–17.

[4] J. DiCarlo and B. Wandell, "Rendering High Dynamic Range Images," in *Proc. of the SPIE, Image Sensors.*, Jan. 2000, vol. 3965.

[5] G. Larson, H. Rushmeier, and C. Piatko, "A Visibility Matching Tone Reproduction Operator for High Dynamic Range Scenes," *IEEE Trans. Visualization and Computer Graphics*, vol. 3, no. 4, pp. 291–306, Oct.-Dec. 1997.

[6] J. Tumblin and G. Turk, "LCIS: A Boundary Hierarchy for Detail-Preserving Contrast Reduction," in *SIGGRAPH 99*, 1999.

[7] J. Tumblin and H. Rushmeier, "Tone Reproduction for Realistic Images," *IEEE Computer Graphics and Applications*, vol. 13, no. 6, pp. 42–48, Nov. 1993.

[8] J. Tumblin, J. Hodgins, and B. Guenter, "Two Methods for Display of High Contrasted Images," *ACM Trans. on Graphics*, vol. 18, no. 1, pp. 56–94, Jan. 1999.

[9] S. Pattanaik, J. Ferwerda, M. Fairchild, and D. Greenberg, "A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display," in *SIGGRAPH 98*, 1998, pp. 287–298.
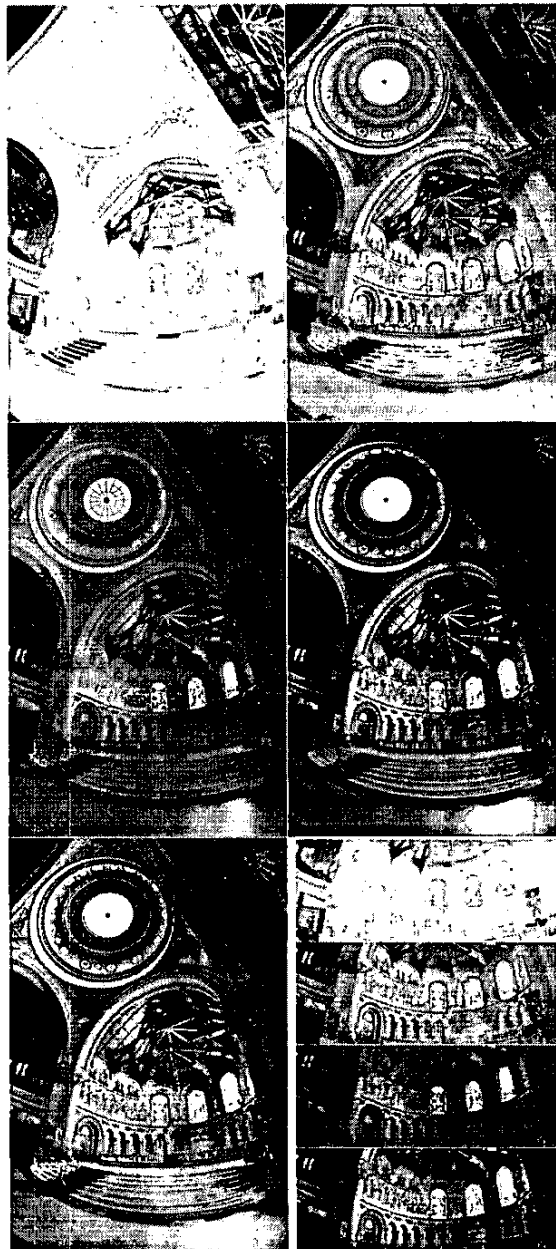
Fig. 1. Images displaying the first, second and third intervals (all with maximal entropy per region), Image obtained with Larson's algorithm, image with maximal global entropy, and zoomed versions of the first four.