High Dynamic Range Tone Mapping Algorithm Based on Image Feature Maps

Matheus V. dos Santos and Mylène C.Q. Farias
Universidade de Brasília, Brasília, Brazil
matheusveleci@gmail.com and mylene@ene.unb.br

***Abstract*—Photographs produced by single-lens reflex (SLR) cameras are photos with Low Dynamic Range (LDR). To obtain photographs with a high dynamic range (HDR), i.e. images with radiance values that are closer to actual real life scenes, we need to use a combination of various LDR images captured at different exposure times. In this paper, we propose a new tone mapping algorithm that combines LDR images, taking into consideration a set of feature maps, which include a saliency map, a saturation map, and an error map. The proposed tone mapping algorithm shows good results when compared to state-of-the-art methods.**

***Keywords*-High Dynamic Range; image processing; tone mapping; photograph;**

I. Introduction

The first stage in a digital image processing system is the acquisition process. During acquisition, the scene luminance is captured and non-linear transformations are used to produce a series of picture elements (pixels) that represent the intensity of colors in the scene. One of the several parameters that directly influences the quality of the acquired image is the dynamic range, which is defined as the ratio between the highest and lowest value of a variable. In digital images, the dynamic range (or contrast ratio) denotes the ratio between black and white image pixels. Photographs captured by regular cameras have a low dynamic range (LDR), that is, they have a dynamic range order of only $10^2 : 1$. On the other hand, the dynamic range of an actual scene can be as high as $10^6 : 1$ and the human visual system can adapt to dynamic ranges of up to $10^4 : 1$.

Given that images captured by regular cameras correspond to a limited representation of the scenes perceived by human viewers, in the last decades researchers have been developing techniques that are able to capture and create images with a high dynamic range (HDR) [1], [6]. One popular approach is to create an HDR image using a series of LDR images, which are captured using a LDR camera with different expousure times (from 1/4000 to 30 seconds). As shown by Paul Debevec [1], LDR images with different exposure times can be combined to generate a better representation of a natural scene. More especifically, the images can be combined using a tone mapping algorithm that maps their color values into a LDR, approximatig the appearance of high-dynamic-range images. This process makes it possible to display better ('HDR-like') images on LDR displays [3].

In this paper, we propose a new tone mapping algorithm. The proposed algorithm uses the response camera function proposed by Paul Debevec [1]. Then, it obtains a set of feature maps, which includes a saliency map, a saturation map, and an error map. Finally, the algorithm combines the images captured at different exposure times, taking into account the corresponding feature maps. The proposed tone mapping algorithm shows good results when compared to state-of-the-art methods.

The paper is divided as follows. In Section II, we describe the proposed tone mapping algorithm. In Section II-F, we present our results. Finally, in Section III we draw our conclusions.

II. Proposed Tone Mapping Algorithm

The proposed tone mapping algorithm can be divided into six stages: (A) image acquisition, (B) camera response function, (C) extraction of image feature maps, (D) calculation of radiance maps, (E) tone mapping and (F) results . In this section, we describe each of these stages.

A. Image Acquisition

In the acquisition stage, we take photographs of static scenes with different exposure times. The exposure time (or shutter speed) consists of the amount of time the digital sensor is exposed to light. Naturally, the amount of light that is able to reach the digital sensor will be proportional to the exposure time. A short exposure time generates darker photos than a longer exposure time. In our work, we captured images with different exposure times using an SLR Camera (NIKON D90[1]) in manual mode. This mode allows varying the aperture and shutter speed to produce digital photographs with exposure times raging from 1/4000 to 30 seconds. In our tests, we captured 3, 5, and 7 photos for each scene.

B. Camera Response Function

In digital image processing, we define photos as a group of pixels with different intensity values. These different intensity values (radiance) represent a different color level in the scene. Frequently, the radiance range is not enough to represent the actual radiance values of a real scene. Therefore, the camera uses a series of different nonlinear mappings and transformations, producing an approximated radiance value for each pixel. We use Paul Debevec's algorithm [1] to reconstruct the response function of our camera. A block-diagram of this

[1]Nikon D90 - Resolution: 12.3 pixels, Kit Lens: 5.80x zoom, 18-105mm, Shutter: 1/4000 - 30 seconds, Max Aperture: 3.5.

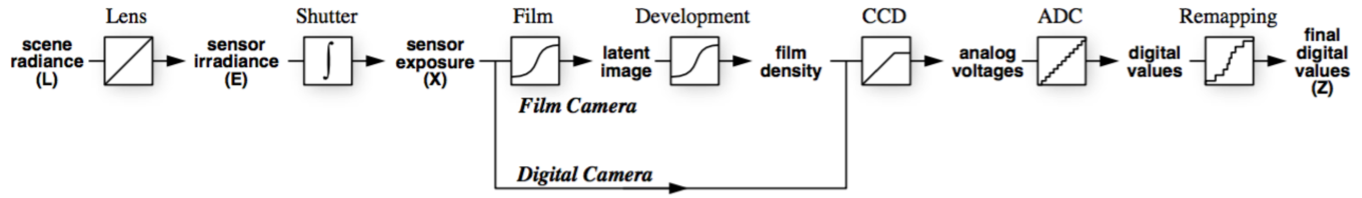

Fig. 1. Block diagram of the image acquisition pipeline, showing how scene radiance becomes pixel values for both film and digital cameras. (Original from Paul Debevec [1]).

process (original taken from Paul Debevec [1]) is shown in Fig. 1.

Notice from Fig. 1 that the relation between the 2-D signal captured in the sensor (sensor exposure), X, and the final pixel intensity digital values, Z, is given by a nonlinear function:

$$X = f^{-1}(Z), \tag{1}$$

where f is the response curve function for the camera. Knowing that X is the product between the specific exposure time, $\triangle t_j$, and radiance, E_j, we can recover the acquired radiance for each pixel Z_{ij}:

$$Z_{ij} = f(E_i \triangle t_j). \tag{2}$$

Considering the exposure time of the photograph and applying the monotonic property [1], we get:

$$f^{-1}(Z_{ij}) = E_i \triangle t_j. \tag{3}$$

Taking the natural logarithm of this equation, we have:

$$g(Z_{ij}) = \ln\left(f^{-1}(Z_{ij})\right) = \ln(E_i) + \ln(\triangle t_j). \tag{4}$$

At this point, the exposure time ($\triangle t_j$) and the pixel intensity values (Z_{ij}) are known, so we can find the actual radiance for each pixel with the following equation:

$$\ln(E_i) = g(Z_{ij}) - \ln(\triangle t_j). \tag{5}$$

To recover the closest response function corresponding to the camera, we introduce the weight function that models the basic shape of $g(Z)$ [2]. To reduce the systematic error (noise, ghosts and color errors), this function penalizes underexposed and overexposed pixels using the weight function proposed by Chaurasiya and Ramakrishnan [2]:

$$w(Z) = p^{\alpha}(1-p)^{\beta}, \tag{6}$$

where $0 \leq p \leq 1$ and $1 \leq \alpha \leq \beta$ [2]. Fig. 2(a) shows the shape of these weight functions [2], while Fig. 2(b) shows the camera response function for the NIKON D90.

C. Image Feature Maps

The third stage consists of acquiring the image feature maps. In this work we obtain the following feature maps for each image: saliency map, saturation map, and error maps. The combination of the input images with different exposure times, taking into consideration their feature maps, is used to generate the HDR image.

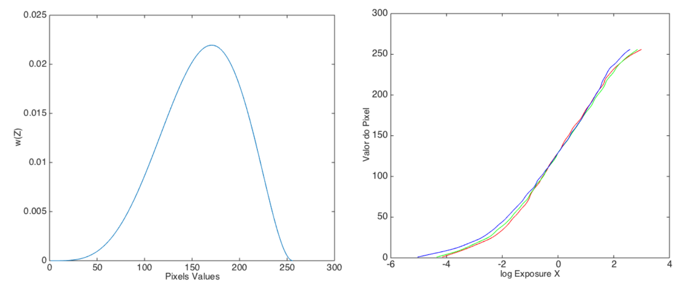

Fig. 2. Weight function (eq.6) used for the luminance values in the camera response function definition, considering $\alpha = 2$ and $\beta = 4$ [2] and the camera response function for the camera NIKON D90.

It is worth pointing out that unknown nonlinear mappings can occur during exposure, development, scanning, digitization, and remapping [1]. Underexposed photos might have dark areas at the image corners and, at the same time, not so many overall details. Overexposed photos, on the other hand, may be rich in details but have saturated areas (completely white areas). At the combination stage, saturated pixels, whose values do not correspond to *true* colors or to actual scene information, may negatively affect the final image. Therefore, to penalize or to eliminate these saturated pixels, in this work we propose to used a combination of three different features: (1) Error, (2) Saturation, and (3) Saliency. Next, we describe the procedure used to extract these features.

1) Error: In this map, we try to predict the pixel intensity of the next exposure, taking into account the linearity function $g(Z_{ij})$ [4]. Suppose that we have a pixel $Z_{ij}(x,y)$ in an image captured with exposure time $\triangle t_j$. In a second image, captured with a different exposure time $\triangle t_k$, the same pixel becomes $Z_{ik}(x,y)$. Using eq. 3, we obtain:

$$f^{-1}(Z_{ij}) = X_{ij} = E_i \triangle t_j \tag{7}$$

and

$$f^{-1}(Z_{ik}) = X_{ik} = E_i \triangle t_k. \tag{8}$$

These equations can be used to generate an estimate for all pixels in the second image $\tilde{Z}_{ik}$:

$$\tilde{Z}_{ik} = f\left(\frac{\triangle t_k}{\triangle t_j}\right) f^{-1}(Z_{ij}). \tag{9}$$

Then, we compare the differences between the pixels in the predicted and acquired images:

$$|\tilde{Z}_{ik} - Z_{ik}| < \epsilon. \quad (10)$$

where ϵ is a parameter defined by the the image variance. These differences determine which pixels are aligned with the correct exposure time. If a given predicted pixel has a big difference when compared to the original, this pixel gets a high value (closer to white) in the *error map*, while if the predicted pixel has a small difference it gets a small value (closer to black). Fig. 4 shows an example of an pixel error map corresponding to the images in Fig. 3.

Fig. 3. Photoghaphies captured with 1/3200 (left) and 1/400 (right) time exposures.

Fig. 4. Pixel Error Map for images in Fig. 3.

2) Saturation: Saturation is an important feature to take into account when combining images with different exposures. As mentioned earlier, these photos generally have overexposed or underexposed pixel areas that may be saturated towards '0' (dark areas) or towards '255' (bright areas). In order to obtain areas with a higher degree of details, these kinds of pixels must be penalized in the combination phase . This is achieved by creating a saturation map using the following equation:

$$\text{Sat}\left(Z_{ij}\right) = \begin{cases} 1, & \text{if } Z_{ij} = 255 \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

Fig. 5 shows an overexposed image (left), captured with 1/800 time exposure, and its corresponding saturation map (right).

3) Saliency: Salient areas are the regions of a scene that capture viewer's attention. The Saliency map can act as a weight map that penalizes only the pixels in areas with less visual importance. Therefore, using a saliency map in the combination stage can improve the quality of salient areas. To create the saliency map we use the Graph-Based Visual Saliency algorithm [5], which uses bottom-up characteristic of the image to obtain a map in which salient pixels have higher values (whiter). Fig. 6 shows an image (left) captured with 1/800 time exposure and its corresponding saliency map

Fig. 5. Overexposed image, captured with 1/800 time exposure (left), and its corresponding saturation map (right).

Fig. 6. Overexposed image, captured with 1/800 time exposure (left), and its corresponding saliency map (right).

(right). The salient areas (closer to white) should not be penalized as harshly as the non-salient areas. To create the final saliency map, we apply a threshold to the pixel values in this map, i.e. pixel values with saliency values greater than this threshold will be classified as pixels-of-interest. This final saliency map will work as a filter in the combination process.

D. HDR Radiance Maps

The next step of the algorithm is to recover the radiance map by combining the images with different exposure times, taking into account the weight function and the featured maps. Our combination is given by the following equation:

$$\ln(E_i) = \frac{\sum_{j=1}^{P} w(Z_{ij}) \left(g(Z_{ij}) - \ln(\triangle t_j)\right) \text{FM}(Z_{ij})}{\sum_{j=1}^{P} w(Z_{ij})} \quad (12)$$

where $w(Z_{ij})$ is the weight function in eq. 6 and $g(Z_{ij})$ is given by eq. 3. Also, $\text{FM}(Z_{ij})$ is the combined feature map, which is obtained by merging the saliency ($\text{Saliency}(Z_{ij})$), saturation ($\text{Sat}(Z_{ij})$), and error ($\text{Error}(Z_{ij})$) maps using the following equation:

$$\text{FM}(Z_{ij}) = (\text{Sat}(Z_{ij}) \vee \text{Error}(Z_{ij})) \wedge (\text{Saliency}(Z_{ij}))^c. \quad (13)$$

where $\wedge$ is the AND logical operation, $\vee$ is the OR logical operation and $(x)^c$ is the complement logical operation of x.

E. Tone Mapping

As mentioned earlier, HDR images have high levels of radiance. As a result, the map of radiance defined in eq. 12 is a better approximation of the actual radiance of the scene. Nevertheless, these radiance intensities are in a magnitude range that cannot be displayed in LDR monitors. Therefore, we must design a tone mapping algorithm, which can map the amplitudes of the radiance to the LDR domain. In other

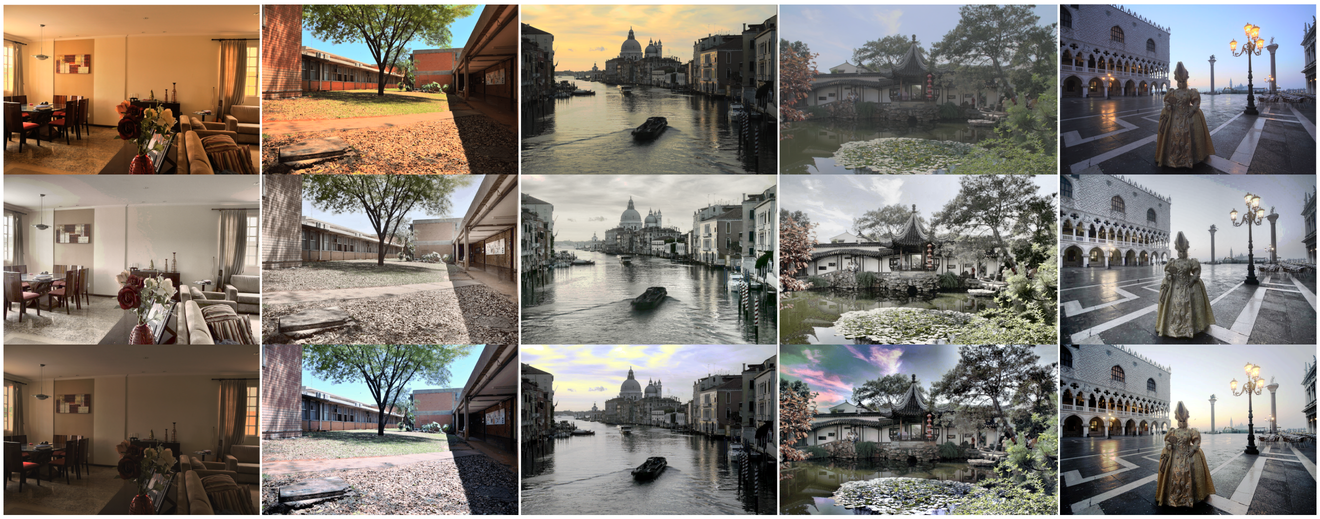

Fig. 7. Images showing the results obtained using: Paul Debevec (first row), Matlab (second row), proposed algorithm (thrid row).

words, we have to transform the HDR map of radiance to a default domain that can be understood by LDR displays.

Our algorithm uses the global tone mapping technique developed by Reinhard [3], which first sets the global luminance (L_w) and defines the tonal range using a key luminance value, which is calculated using the following equation:

$$L_w = \frac{1}{N} \exp \left(\sum_{w,y} \log(\delta + L_w(x, y)) \right), \tag{14}$$

where $L_w(x, y)$ is the *scene* luminance for each pixel [3], N is the total number of pixels, and δ is a small constant value used to avoid having a zero. We can use this key luminance (L_w) value to scale the image values between 0 and 1, using the following equation:

$$L(x, y) = \frac{a \cdot L(x, y)}{L(x, y)}, \tag{15}$$

where a is equal to 0.18, as suggested by Reinhard [3]. To make sure that the scale is from '0' to '1' and, at the same time, penalize high luminances, we define:

$$L_d(x, y) = \frac{L(x, y)}{1 + L(x, y)}. \tag{16}$$

Finally, to generate the final image, we apply the classics algorithms of histogram equalization, which has the goal of improving the quality of the result of the tone mapping step [7].

F. Results

We compare the proposed algorithm with the algorithm proposed by Paul Debevec [1] and the tone mapping implementation available in the Matlab platform. Fig. 7 shows sample results obtained for four images. In this figure, each column corresponds to a different image, while each line corresponds to a different tone mapping algorithm (Paul Debevec [1]), Matlab implementation available on Image Processing Toolbox and the proposed algorithm). As it can be seen from the images in Fig. 7, when compared to the other two algorithms, the proposed tone mapping algorithm has a good performance. The images generated with the proposed algorithm have a high quality images as we can perceive, showing the details of the scene without introducing visible artifacts.

III. Conclusions

In this paper, we propose a new tone mapping algorithm that, besides using the response camera function proposed by Paul Debevec [1], also takes into account image features. In particular, it considers a saliency map, a saturation map, and an error map to combine images captured at different exposure times. The proposed tone mapping algorithm shows good results when compared to state-of-the-art methods, reducing color distortions and false-contours artifacts.

IV. ACKNOWLEDGEMENTS

This work was supported by CAPES, CNPQ, and by the University of Brasília.

References

[1] P. Debevec and J. Malik. Recovering High Dynamic Range Radiance Maps from Photographs. In Proc. of 24th conference on Compute Graphics and Interactive Techniques, (ACM), New York, USA, pp. 369378. August 1997.

[2] P. K. Chaurasiya and K.R. Ramakrishnan. "High Dynamic Range Imaging". 2013 International Conference on Communication Systems and Network Technologies.

[3] E. Reinhard, M. Stark, P. Shirley, J. Ferwerda, Photographic Tone Reproduction for Digital Images. ACM Transections on Graphics, 21 (3). New York, USA. July 2002.

[4] G. Thorsten, "Fast and Robust High Dynamic Range Image Generation with Camera and Object Movement", December 2005.

[5] H. Jonathan, K. Christof, and P. Pietro. "Graph-Based Visual Saliency", 2013.

[6] E. Reinhard, S. Pattanaik, G. Ward and P. Debevec. "High Dynamic Range Imaging: acquisition, display, and image-based lighting", 2006.

[7] A. Boschetti, N. Adami, R Leonardi and M. Okuda ."High Dynamic Range image tone mapping based on local Histogram Equalization", Multimedia and Expo (ICME), 2010 IEEE International Conference on July 2010.

[8] G Ward. "High Dynamic Range Encondings". Color and Imaging Conference, 2001.