FAST DYNAMIC RANGE COMPRESSION FOR GRAYSCALE IMAGES

Vasilios Vonikakis[1], Ioannis Andreadis[1] and Antonios Gasteratos[2]

[1] Laboratory of Electronics, Section of Electronics and Information Systems Technology,
Department of Electrical and Computer Engineering,
Democritus University of Thrace, GR-671 00 Xanthi, Greece
E-mail: {bbonik, iandread}@ee.duth.gr
[2] Laboratory of Robotics and Automation, Section of Production Systems,
Department of Production and Management Engineering
Democritus University of Thrace, GR-671 00 Xanthi, Greece
E-mail: agaster@pme.duth.gr

ABSTRACT

This paper presents a new center-surround network for the dynamic range compression of grayscale images. The proposed method exploits some of the shunting characteristics of biological center-surround networks, in order to reduce the effects of uneven illumination and improve the dynamic range of images. The main advantage of the proposed method is its low computational burden, which allows the rendition of high-resolution 5-MPixel images in approximately 1.3 seconds, when executed by a conventional personal computer. The method is compared to the latest commercial version of the Retinex algorithm, and exhibits promising results for a wide variety of real images and lighting conditions.

1. INTRODUCTION

The dynamic range of natural scenes, that is, the ratio between the maximum and minimum tonal values found in the scene, can reach up to great proportions [1]. Conventional 8-bit cameras have a dynamic range of just 256:1, while 14-bit high-end cameras can reach up to 16,384:1. This introduces an important problem to artificial vision systems, especially when the dynamic range of the scene exceeds the dynamic range of the camera. In these cases, certain parts of the image can become either underexposured or overexposured, reducing the compound visual information. Contrary to cameras, the human visual system (HVS) can accommodate a dynamic range of approximately 10,000:1 [1]. This usually results into significant differences between the image perceived by the HVS and the one captured by the camera.

Many algorithms have been presented in the past decade that attempt to solve this problem. The most important of all is the Retinex family of algorithms. Retinex was first presented by Edwin Land in 1972 [2] and was inspired by some attributes of the HVS, which also defined its name (Retina & Cortex). The initial algorithm inspired many others, the latest of which can be found in [3], while an extensive analysis of the algorithm can be found in [4,5]. The main idea of Retinex is the calculation of ratios between the original image and a gaussian-filtered version of the original image, by computing their differences in a logarithmic space. These ratios are used for the extraction of a lightness map which is independent from the scene illumination and depends only on the reflectances of objects in the scene. The algorithm is applied independently to each chromatic channel and in three different spatial scales. The final output is the weighted sum of the three scales. The main advantage of Retinex is the good dynamic range compression that achieves for a variety of different lighting conditions and the color constancy attributes that exhibits. It has been successfully used in many applications, such as shadow removal [6], color correction [7] and gamut mapping [8]. The main weakness of Retinex is its computational burden, which derives from the convolution of the image with Gaussian filters of radiuses up to 240 pixels. Additionally, halo effects tend to appear in regions where strong intensity transitions exist, degrading the final output of the algorithm. Other approaches to the dynamic range compression problem include the modeling of brightness perception [9], the manipulation of gradient in the luminance component [10] and the combination of images captured with different exposures [11, 12]. A detailed overview of dynamic range compression techniques can be found in [13].

The proposed method is partially inspired by the HVS. It particularly adopts some of the shunting characteristics of the on-center off-surround networks, in order to define the response function for a new artificial center-surround network. This network compares every pixel to its local average and assigns a new value in order to light the dark image regions, while minimally affecting the light ones. Histogram stretching is applied before the center-surround network, ensuring that the final output occupies the full dynamic range of the medium. The proposed algorithm allows the rendition of high resolution 5-MPixel images in approximately 1 second, even when executed by a conventional personal computer. Additionally, no multiple scales are needed and no halo effects are introduced in the strong transitions between light and dark regions. The algorithm exhibits good results for a wide variety of images and lighting conditions without requiring additional manual tuning. The evaluation of the proposed method is

carried out by comparing its outputs with the commercial software PhotoFlair, which utilizes the Multi-Scale Retinex (MSR) algorithm, as presented in [3]. The results of the comparison show that the proposed method exhibits comparable and many times superior results to the Retinex algorithm, in significantly lower execution times.

The rest of the paper is organized as follows: Section 2 presents some of the attributes of the ganglion cells of the HVS upon which the proposed method is based and a detailed description of the algorithm. Section 3 demonstrates the experimental results. Finally, concluding remarks are made in section 4.

2. DESCRIPTION OF THE METHOD

It is long known that a shunting on-center off-surround network of cells can reduce the effects of illumination (discount the illuminant) and extract the scene reflectances by adapting to a wide range of inputs [14, 15]. These attributes are the result of a biological adaptation mechanism known as 'shunting inhibition'. The steady state solution of the shunting differential equation, that shows the output of an on-center off-surround cell, when t $\rightarrow \infty$, is shown in equation 1.

$$\text{out}(C,S) = \frac{C-S}{g_{leak} + C + S} \qquad (1)$$

where C is the value of the center, S is the value of the surround and g_{leak} is the decay constant. When used in image processing applications, g_{leak} usually is the maximum value that a pixel may take, which in our case is 255. The 3-dimensional representation of equation (1) for an on-center off-surround cell with $0 \le C \le 255$ and $0 \le S \le 255$ is depicted in Fig. 1.
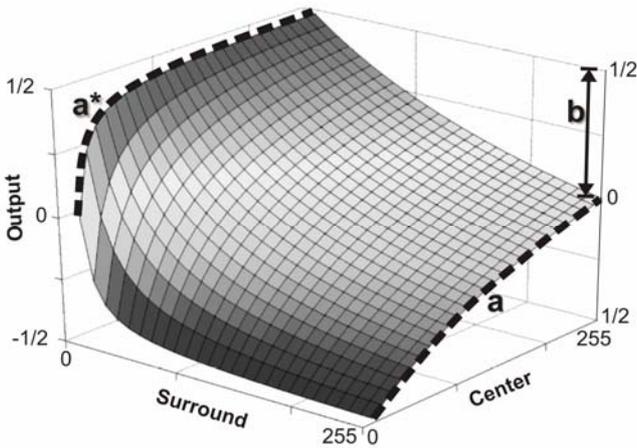

Figure 1: The 3-dimensional representation of equation (1).

The reason for which a center-surround cell with equation (1) as an activation function can discount the illuminant is evident in Fig. 1. In fact, equation (1) performs a comparison of the contrast between the center and the surround and adequately adjusts the output. For low surround values, which means that the cell is located in

dark image regions, the non-linearity 'a*' acts as a local nonlinear correction: it gives a high output even when the center has low values. This practically means that in dark image regions, where illumination is not adequate, the cell increases its response to the contrast between center and surround, compensating for the low illumination. On the contrary, for high surround values, which means that the cell is located in a light image region, the non-linearity 'a*' gradually vanishes to the almost-linearity 'a'. As a result, the output of the cell is linearly associated with the contrast between the center and the surround.

Equation (1) possesses several characteristics, which make it inappropriate for direct use in dynamic range compression. First, it produces a bipolar output, since it is essentially an edge detector. For this reason, it is usually followed by half-way rectification and a filling-in procedure [16] which is time consuming. Second, the output range is reduced to half 'b', as the surround values progress from low to high. In order to overcome these drawbacks, equation (2) is used as a basis, which describes the total activity of a shunting center-surround network [14]. It is then modified in order to eliminate the unwanted output reduction 'b'. The resulted function is equation (3). The graphical representations of both equations (2) and (3) are depicted in Fig. 2.

$$F(x) = \frac{B \cdot x}{A + x} \qquad (2)$$
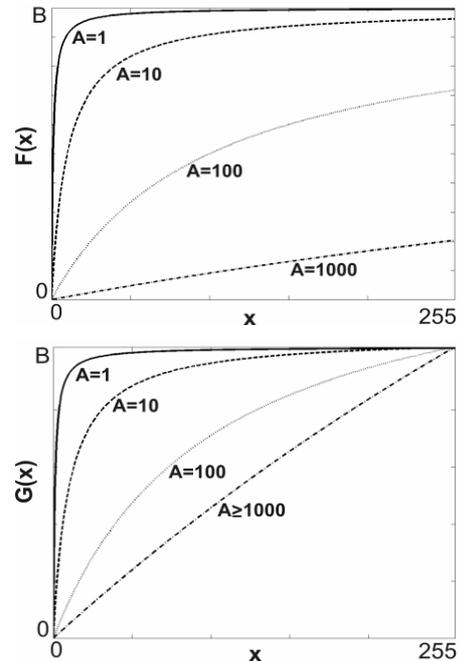
$$G(x) = \frac{(B + A) \cdot x}{A + x} \qquad (3)$$

Figure 2: Graphical representation of equations (2) and (3).

In equations (2) and (3), B is the maximum value the function can take and A is a constant that determines the degree of non-linearity. Equation (3) transitions from a sharp non-linearity (A=1) to a linearity (A≥1000), while maintaining the same output range $0 \le G(x) \le B$ for all the

possible values of A. If x is substituted by the center C of a center-surround cell and the surround S is correlated with the non-linearity factor A, a new, improved response function is formed. Since the maximum response of the cell must be equal to the maximum value of a pixel, B=255. The following equations describe the activation function of the center-surround cell of the proposed method.

$$C_{i,j,t+1}(C,S) = \frac{\left(255 + A\left(S_{i,j}\right)\right) \cdot C_{i,j,t}}{A\left(S_{i,j}\right) + C_{i,j,t}} \quad (4)$$

$$A(S_{i,j}) = S_{i,j} + m + q\left(S_{i,j}\right) \quad (5)$$

$$q(S_{i,j}) = \frac{255 \cdot S_{i,j}}{255 \text{-} S_{i,j}} \quad (6)$$

$$S_{i,j} = \frac{1}{9}\sum_{y=i-1}^{i+1}\sum_{x=j-1}^{j+1} p_{y,x} \quad (7)$$

Where, (i, j) denote the coordinates of a pixel in the image and p is its value. In the proposed method, the surround $S_{i,j}$ is the average of a 3×3 pixel region while the center $C_{i,j}$ is the central pixel of this region. Equation (5) describes the non-linearity factor A, as a function of the surround. The minimum value obtained by A(S), when S=0, is denoted by 'm'. This non-linearity affects the overall result of the algorithm and its value is determined by the statistics of the image. Equation (6) describes the transition between the non-linearity 'a*' and the linearity 'a', as the surround values increase. Fig. 3 shows the 3-dimensional representation of equation (4).
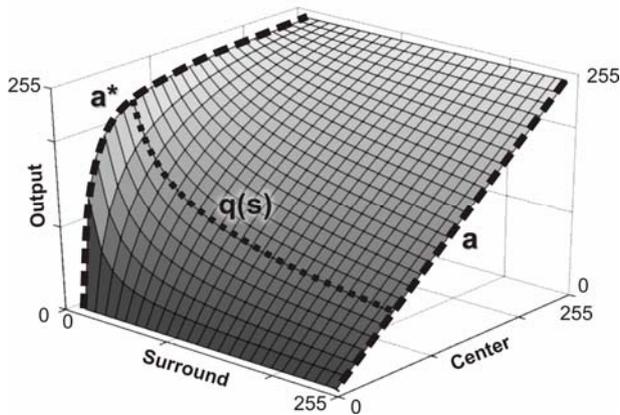

Figure 3: The 3-dimensional representation of equation (4).

The new activation function maintains the transition from the non-linearity 'a*' to the linearity 'a' correlated with the value of the surround. This means that when the surround has low values, something occurring in dark image regions, the non-linearity 'a*' increases the value of the center in order to increase the local contrast. On the contrary, when the surround comprises high values, something that happens in light image regions, the linearity 'a' does not alter the value of the center. For all the intermediate surround values, equation (6) determines the degree of non-linearity.

The calculation of 'm' in equation (5) is as follows:

$$r = \frac{\sum_{i=1}^{py}\sum_{j=1}^{px} u\left(85\text{-}p_{i,j}\right)}{px \times py} \times 100 \quad (8)$$

$$m = \frac{190}{100}(100 - r) + 10 \quad (9)$$

Where u(x) is the unit step function (is 1 if x≥0 and 0 if x<0), px, py are the dimensions of image and $p_{i,j}$ the pixel value at position (i, j). The main idea is to calculate the percentage of pixels in the image that have values below 85. This is a rough estimation of the darkness in the image and therefore it is used to adjust the minimum value 'm' that the non-linear factor A can obtain. Equation (9) linearly regulates the value of 'm', between 10, when 100% of the image pixels are below 85 and 200, when 0% of the image pixels are below 85. Constant 85 was selected as it is the 1/3 of 255 and can be thought of as the first bin of a 3-bin histogram that divides the 255 intensities into 3 sets: dark, medium and light.

In order to achieve better results, a histogram clipping and stretching technique is required, prior to processing with the center-surround network. The technique is the one that has been extensively discussed in [4] and for this reason it is not described here.

3. EXPERIMENTAL RESULTS

The proposed method is compared to the MSR algorithm, extensively described in [3]. The MSR algorithm is generally used for color constancy applications. However, in the present study only its dynamic range compression characteristics are compared, by applying it only on grayscale images. The implementation of the MSR algorithm that was used for the tests was the commercial software PhotoFlair that features the MSR algorithm. The parameters of the MSR were the default that the authors have mentioned (3 scales with radiuses 5, 20 and 240 pixels and equal weights for every scale). The proposed method was implemented in C code. Both algorithms were executed on an Intel Pentium 4 processor, running at 3GHz, with 512 MB RAM and Windows XP.

3.1 Results with Real Images

This subsection presents the results of the comparison between the MSR and the proposed algorithm, in a set of real high-resolution grayscale images. Most of the images represent scenes that were captured with different digital cameras under different lighting conditions where dynamic range correction is required. Table 1 exhibits some of the results that were obtained by the proposed method and MSR. For every image, its size and the execution times of the two methods are included. It is important to mention that the PhotoFlair software that was used to obtain the MSR outputs has 3 different versions of the Retinex algorithm: Scenic Retinex, Portrait Retinex and Portrait Retinex followed by 'auto levels', which is a form of

histogram equalization. The Retinex results that are shown in Table 1 are always the best of the 3 versions, as selected by the typical human observer.

It is evident from the results that the proposed method can correct most of the problems caused by the limited dynamic range of the camera. Additionally, all images demonstrate the main advantage of the proposed method over MSR; the lack of halo effects. Most of the regions that MSR fails to restore are the result of significant halo effects which are caused in the boundaries of a dark region with a light one. The execution times of the proposed method are approximately 23 times lower than the ones required for the MSR, allowing the rendition of 5-MPixel images in 1.3 sec. It should not be neglected though, that MSR preserves better the light regions of the image, while the proposed method tends to light them more than necessary.

Original	MSR	Proposed
1. 1920×2560	30 sec	1.3 sec
2. 2560×1920	29 sec	1.3 sec
3. 2560×1920	30.5 sec	1.4 sec
4. 2304×1728	25.7 sec	1.1 sec
5. 1728×2304	27 sec	1.1 sec
6. 2560×1920	31.7 sec	1.4 sec

Table 1: Comparison of results of the proposed method with MSR in real images.

3.2 Quantitative Comparison

The following experiment was carried out in order to achieve a numerical comparison between the two algorithms. Different levels of a computer generated shadow were embedded in a real image that had otherwise no need for dynamic range correction. The shadow level was determined by the factor 'sh', ranging between 50%-95%. Half of the image pixels were multiplied by 1-(sh/100), thus producing a shadow in half of the image. Fig. 4 shows some of the computer generated shadows, as well as the outputs for these particular images.
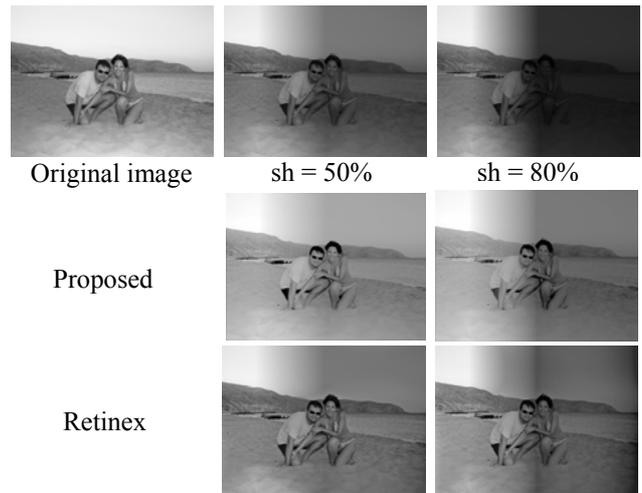

Figure 4: The original image and two examples of the computer generated shadows, as well as results for these examples.
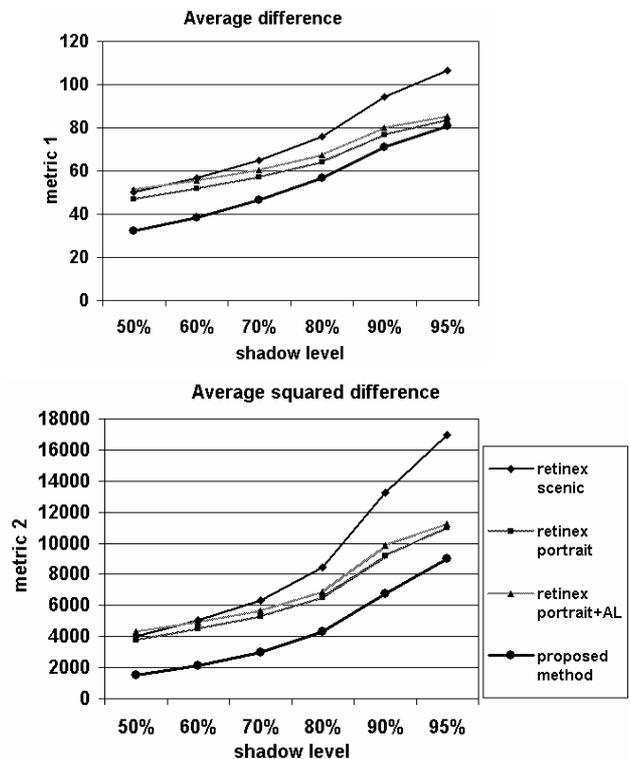

Figure 5: Results from the comparison for different shadow levels.

The results of the two algorithms were compared to the original image by the use of two metrics. The first was the average difference, which is the sum of the pixel differences divided by the total number of pixels. The second was the average squared difference, which is the sum of the squared pixel differences divided by the total number of pixels. The first metric gives an overall impression of the difference of the two images. The second metric assigns more weight to bigger differences than smaller ones, due to the presence of the square. It is thus more sensitive to big differences, such the ones derived by unwanted halo effects. Fig. 5 shows the results of the comparison. All the 3 versions of Retinex were tested. The Retinex images that are shown in Fig. 4 are the best of the 3 versions, as selected by the typical human observer. The results of the experiment show that the proposed method outperforms all the 3 versions of the MSR for both the metrics and for all the shadow levels.

4. CONCLUSIONS

This paper presents a new algorithm for fast dynamic range compression. The proposed algorithm exploits some of the characteristics of the shunting on-center off-surround networks of the HVS. A new activation function has been introduced that compensates for the low illumination in the dark image regions and enhances the contrast.

The main advantage of the proposed method is its fast execution. The major reason for this is the fact that the method avoids computationally intensive convolutions and computes the new pixel values by processing a 3×3 neighborhood. On the contrary, MSR, as presented in [3, 4], is based on the convolution of the original image with Gaussian filters of large radiuses, ranging up to 240 pixels. Consequently, it is much more computationally expensive than the proposed method. This is reflected also in the results, where the proposed method renders the test images approximately 20 times faster than MSR. The reader can assess the quality of the results of the proposed method, which is at least comparable with that of MSR. When strong transitions between dark and light image regions occur, MSR tends to exhibit halo effects, whereas the proposed method does not. A numerical assessment to quality was also carried out to confirm these conclusions.

However, it should be mentioned that the proposed method tends to light the correct image areas more than necessary, whereas the MSR does not. This can be justified due to the fact that all the results and tests were carried out without manual tuning. This is important because it shows that the proposed method can achieve satisfactory results for a wide variety of images, without any supervision. Taking also into account its simplicity and the short execution times, the algorithm it can be used in time-critical applications that require automatic dynamic range compression.

5. REFERENCES

[1] K. Devlin, "A Review of Tone Reproduction Techniques," Technical Report CSTR-02-005, University of Bristol, 2002.

[2] E. Land and J. McCann, "Lightness and Retinex Theory," Journal of the Optical Society of America, Vol. 61. No.1.pp. 1-11, 1971.

[3] D. Jobson, Z. Rahman, and G. Woodell, "A Multi-Scale Retinex For Bridging the Gap Between Color Images and the Human Observation of Scenes," IEEE Transactions on Image Processing: Vol. 6.pp. 965-976, 1997.

[4] D. Jobson, Z. Rahman, and G. Woodell, "Properties and Performance of a Center/Surround Retinex," IEEE Transactions on Image Processing, Vol. 6.pp. 451-462, 1997.

[5] K. Barnard and B. Funt, "Investigations into Multi-scale Retinex," Color Imaging: Vision and Technology, pp. 9–17, 1999.

[6] G. Finlayson, S. Hordley and M. Drew, "Removing Shadows from Images using Retinex," Color Imaging Conference, pp. 73-79, 2002.

[7] D. Marini and A. Rizzi, "A Computational Approach to Color Adaptation Effects," Image and Vision Computing Vol. 18. pp. 1005-1014, 2000.

[8] J. McCann, "Lessons Learned from Mondrians Applied to Real Images and Color Gemuts," The Seventh Color Imaging Conference, Color Science, Systems and Applications, pp. 1-8, 1999.

[9] V. Brajovic, "Brightness Perception, Dynamic Range and Noise: A unifying model for adaptive image sensors," CVPR'04 Vol 2. pp. 189-196, 2004.

[10] R. Fattal, D. Lischinski and M. Werman, "Gradient Domain High Dynamic Range Compression," 29th International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH), 2002.

[11] P. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," SIGGRAPH 97, pp. 369– 378, 1997.

[12] S. Nayar and T. Mitsunaga, "High Dynamic Range Imaging: Spatially Varying Pixel Exposures," IEEE Conference on Computer Vision and Pattern Recognition Vol. 1. pp. 472 – 479, 2000.

[13] S. Battiato, A. Castorina, and M. Mancuso, "High dynamic range imaging for digital still camera: an overview," Journal of Electronic Imaging Vol. 12. pp. 459-469, 2003.

[14] S. Ellias and S. Grossberg, "Pattern Formation, Contrast Control and Oscillations in the Short Term Memory of Shunting On-Center Off-Surround Networks," Biological Cybernetics Vol. 20. pp. 69-98, 1975.

[15] S. Grossberg, "Visual Boundaries and Surfaces," The Visual Neurosciences Vol. 2, MIT press, 2004.

[16] E. Mingolla, W. Ross and S. Grossberg, "A Neural Network for Enhancing Boundaries and Surfaces in Synthetic Aperture Radar Images," Neural Networks Vol. 12. pp. 499-511, 1999.