# Fast Automatic Compensation of Under/Over-Exposed Image Regions

Vassilios Vonikakis and Ioannis Andreadis

Democritus University of Thrace, Department of Electrical and Computer Engineering,
Laboratory of Electronics, Section of Electronics and Information Systems Technology,
Vas. Sofias, GR-67100 Xanthi, Greece
{bbonik, iandread}@ee.duth.gr
http://electronics.ee.duth.gr

**Abstract.** This paper presents a new algorithm for spatially modulated tone mapping in Standard Dynamic Range (SDR) images. The method performs image enhancement by lightening the tones in the under-exposed regions while darkening the tones in the over-exposed, without affecting the correctly exposed ones. The tone mapping function is inspired by the shunting characteristics of the center-surround cells of the Human Visual System (HVS). This function is modulated differently for every pixel, according to its surround. The surround is calculated using a new approach, based on the oriented cells of the HVS, which allows it to adapt its shape to the local contents of the image and, thus, minimize the halo effects. The method has low complexity and can render 1MPixel images in approximately 1 second when executed by a conventional PC.

**Keywords:** Image Enhancement, Tone Mapping, Human Visual System.

## 1  Introduction

Important differences often exist between the direct observation of a scene and the captured digital image. This comes as a direct result of the low dynamic range of the capturing device, compared to the dynamic range of natural scenes. Conventional SDR images (8-bits/channel) cannot acceptably reproduce High Dynamic Range (HDR) scenes, which is usual in outdoor conditions. As a result, recorded images suffer from loss in clarity of visual information within shadows (under-exposed regions), or near strong light sources (over-exposed regions).

A straight-forward solution to this problem is the use of HDR capturing devices instead of the conventional SDR ones. Nevertheless, HDR cameras cannot always provide a practical solution. Their increased cost has limited their use, while the majority of the existing vision systems are already designed to use SDR cameras. Another possible solution is to acquire an HDR image by combining multiple SDR images with varying exposures [1]. However efficient, this solution is by its definition time consuming and thus, cannot be used in time-critical applications. Consequently, an unsupervised tone-enhancement algorithm for the under/over-exposed regions of SDR images could at least partially solve the problem, by enhancing the visual information of those regions, while minimally affecting the others.

Many algorithms have been proposed in this direction the past decades. The most important of all is the Retinex family of algorithms. Retinex was first presented by Edwin Land in 1964 [2] and was motivated by some attributes of the HVS, which also defined its name (Retina & Cortex). The initial algorithm inspired several others and until today the most widespread version of Retinex is the Multi Scale Retinex with Color Restoration (MSRCR) [3], which has been extensively used by NASA and has been established in the market as commercial enhancement software (PhotoFlair). In MSRCR, the new pixel values are given by the logarithmic ratio of the initial value and a weighted average of its surround. Gaussian surround functions of three different scales are used in order to simultaneously achieve local contrast enhancement and tonal rendition. Recently, a new algorithm that has some common attributes with Retinex, called ACE (Automatic Color Equalization), has been reported in [4]. It uses a form of lateral inhibition mechanism, which adjusts every pixel value according to the local and global contents of the image. The new pixel values are then scaled to the dynamic range of the channel (0-255). As a result, it enhances the under/over-exposed image regions while and achieving White-Patch and Gray-World color correction. The main weakness of both algorithms is their computational burden. This derives from the convolution of the image with Gaussians of radiuses up to 240 pixels for the MSRCR or the participation of a great number of pixels for the ACE, respectively. Additionally, halo effects tend to appear in regions were strong intensity transitions exist, degrading the final output of the algorithms.

The thrust of the proposed method is to enhance the tones in the under/over-exposed regions of SDR images, with minimal computational burden and without affecting the correctly exposed ones. This is achieved by a spatially modulated tone mapping in which, the main tone mapping function is inspired by the shunting characteristics of the center-surround cells of the HVS. The calculation of the surround is based on a new approach inspired by the oriented cells of the HVS. According to this, the shape of the surround function is not constant and it is adapted to the local intensity distribution of the image. As a result, it avoids the strong intensity transitions, which lead to the emergence of halo effects. The results of the proposed method outperform the existing ones in this category. The execution time of the algorithm is faster than any of the other algorithms and can render 1MPixel images in approximately 1 sec when executed by a conventional PC.

The rest of the paper is organized as follows: Section 2 presents a detailed description of the proposed algorithm. Section 3 demonstrates the experimental results. Finally, concluding remarks are made in section 4.

## 2   Description of the Method

Fig. 1 shows the block diagram of the proposed method. The input image is first converted to the YCbCr space, in order to decorrelate the chromatic and achromatic information. The algorithm enhances only the $Y$ component. The enhanced component $Y^{out}$ is used in the inverse color-space transformation to produce the final image.
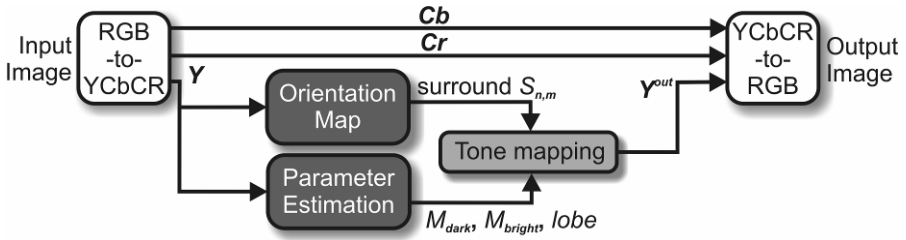
**Fig. 1.** The block diagram of the proposed method

## 2.1 Orientation Map

In order to design a surround function that is adaptable to the local distribution of the intensity, the orientation map *OM* of the *Y* component must be calculated. Instead of intensity values, *OM* contains the outputs of a set of orientation elements, similar to the orientation cells of the HVS. The orientation elements are a set of 60 binary kernels (Fig. 2a). Their size is 10×10 pixels and they are applied to non-overlapping regions of the *Y* component. Consequently, for an image of size *py*×*px*, the size of *OM* will be (*py*/10)×(*px*/10). Every kernel *K* has two distinct parts ($A_K$ and $B_K$). The two parts form a step edge in 12 different orientations (every 15°) and in all possible phases within the 10×10 kernel.



**Fig. 2.** a. The 60 kernels and the 44th kernel in magnification. b. The *Y* component of the original image. c. The orientation map *OM* of image b.

Let (*i,j*) be the coordinates of the upper left pixel in one of the non-overlapping regions of the *Y* component, in which the kernels are applied. Then *OM* is calculated as follows.

$$out_{u,v}^{K} = \left| M_{u,v}^{A,K} - M_{u,v}^{B,K} \right|, \; u = \frac{i}{10}, \; v = \frac{j}{10}, \; u,v \in Z \tag{1}$$

$$M_{u,v}^{A,K} = \frac{1}{N_{A,K}} \sum_{y=i}^{i+9} \sum_{x=j}^{j+9} Y_{y,x} \; \forall Y_{y,x} \in A_K \qquad M_{u,v}^{B,K} = \frac{1}{N_{B,K}} \sum_{y=i}^{i+9} \sum_{x=j}^{j+9} Y_{y,x} \; \forall Y_{y,x} \in B_K \tag{2}$$

$$OM_{u,v}^{A} = M_{u,v}^{A,K'}, \quad OM_{u,v}^{B} = M_{u,v}^{B,K'}, \quad K' : \max\left[ out_{u,v}^{K} \right]_{K=1}^{60} = out_{u,v}^{K'} \tag{3}$$

where $K$ is the number of the kernel, $M^{A,K}_{u,v}$, $M^{B,K}_{u,v}$ are the mean intensities of $A_K$ and $B_K$, respectively, $N_{A,K}$ and $N_{B,K}$ are the number of pixels of $A_K$ and $B_K$, respectively, and $u$ and $v$ are the coordinates of the $OM$. Equation (3) selects the kernel $K'$ with the highest output, whose phase and orientation match the local contents of region $(i,j)$. Every position $(u,v)$ of the $OM$ contains the number of the winner kernel $K'$ and the average intensity values of $M^{A,K}_{u,v}$ and $M^{B,K}_{u,v}$. Fig. 2b and Fig. 2c depict the original image and its $OM$, respectively.

For every pixel $Y_{n,m}$ two surround values are calculated: one for the $Y$ component $(S_{1,n,m})$ and one for the $OM$ $(S_{2,n,m})$. The final surround value $S_{n,m}$ is calculated by an interpolation between the two values and will be discussed later. The surround, in the center of which the pixel $Y_{n,m}$ is located, was selected to be square, with a side of 51 pixels. The $S_{1,n,m}$ surround is the classic surround that has been also used in [3,4]. In these cases, a weighting function (Gaussian for MSRCR, Euclidian or Manhattan for ACE) is used to determine the weight of every pixel in connection to its distance from the central pixel. In the proposed method, for simplicity, no such function was used, and the $S_{1,n,m}$ surround is a simple average, as described by equation (4).

$$S_{1,n,m} = \frac{1}{51^2} \sum_{y=n-25}^{n+25} \sum_{x=m-25}^{m+25} Y_{y,x} \tag{4}$$

The $S_{1,n,m}$ surround tends to extract unwanted halo effects when calculated in a region with a sharp intensity transition (Fig. 3a). For this reason, the $S_{2,n,m}$ surround is also calculated, which adapts its size to the local intensity distribution of the image. The $S_{2,n,m}$ surround belongs to a region $H$ of 5×5 kernels in the $OM$, whose central kernel is the one in which, pixel $Y_{n,m}$ is located (Fig. 3b). Region $H$ is segmented into two distinct regions $E_1$ and $E_2$ (Fig. 3c) by a threshold $th_H$. These regions define the value of $S_{2,n,m}$, as the following equations indicate.

$$OM_H^{\max} = \max\left[OM_z^G\right], \ OM_H^{\min} = \min\left[OM_z^G\right], \forall G \in \{A,B\}, z \in H \tag{5}$$

$$th_H = \frac{OM_H^{\max} + OM_H^{\min}}{2} \tag{6}$$

$$ME_1 = \frac{1}{N_{E1}} \sum OM_z^G \ \text{if} \ OM_z^G < th_H \qquad ME_2 = \frac{1}{N_{E2}} \sum OM_z^G \ \text{if} \ OM_z^G \geq th_H \tag{7}$$

$$S_{2,n,m} = \begin{cases} ME_1 & \text{if} \ Y_{n,m} < th_H \\ ME_2 & \text{if} \ Y_{n,m} \geq th_H \end{cases} \tag{8}$$

where, $N_{E1}$ and $N_{E2}$ are the total number of kernel parts ($A$ or $B$) that constitute regions $E_1$ and $E_2$, respectively. Equation (5) defines the maximum and the minimum of all the kernel parts that belong to region $H$. These two extremes determine the threshold $th_H$, in equation (6), which is necessary to segment region $H$ into $E_1$ and $E_2$. Equation (7) shows that $ME_1$ and $ME_2$ are the mean intensities of $E_1$ and $E_2$, respectively. The final value of surround $S_{2,n,m}$ is determined by the value of the central pixel $Y_{n,m}$. If

$Y_{n,m}$ is lower than the threshold, it belongs to region $E_1$ (the darker region) and thus, $S_2$ acquires the value of $ME_1$. On the contrary, if $Y_{n,m}$ is higher than the threshold, it belongs to region $E_2$ (the brighter region) and thus, $S_{2,n,m}$ acquires the value of $ME_2$. Consequently, $S_{2,n,m}$ is allowed to obtain values only from one region but never from both, and thus, it does not contribute to the formation of halo effects. The difference $dif_H$ between the $OM_H^{max}$ and $OM_H^{min}$ is the factor that determines the final surround value of pixel $Y_{n,m}$ by interpolation.

$$dif_H = OM_H^{max} - OM_H^{min} \quad , dif_H \in [0, 255] \tag{9}$$

$$S_{n,m} = \frac{S_{1,n,m}(255 - dif_H) + S_{2,n,m} \cdot dif_H}{255} \tag{10}$$

When there is not a sharp intensity transition located within the surround window of pixel $Y_{n,m}$, $dif_H$ has a low value and as a result, $S_{1,n,m}$ contributes more to the final surround value $S_{n,m}$. Alternatively, if a sharp intensity transition is located within the surround window, $dif_H$ has a high value and $S_{2,n,m}$ is the main contributor to $S_{n,m}$.
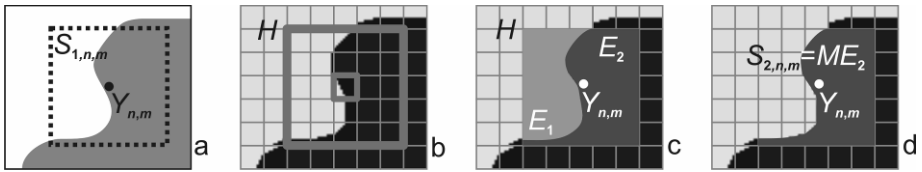


**Fig. 3.** a. A pixel $Y_{n,m}$ and its surround $S_{1,n,m}$ in the $Y$ component, located near an edge where halo effects may appear. b. The $H$ region of pixel $Y_{n,m}$ in the $OM$. c. Segmentation of region $H$ into regions $E_1$ and $E_2$. d. The position of $Y_{n,m}$ determines the value of $S_{2,n,m}$.

## 2.2 Tone Mapping

Center-surround cells of the HVS have to encode extremely high dynamic range visual signals into an output of variable, yet finite frequency. This is succeeded by modulating their response according to the following equation [5]:

$$F(x) = \frac{B \cdot x}{A + x} \quad \forall \, x \in [0, \infty), \, A, B \in \mathfrak{R}^+ \tag{11}$$

Equation (11) maps all inputs from [0,∞) to the interval [0,B] with a nonlinear degree, varying according to the A/B ratio. However, in the present application, all the inputs are bounded to the interval [0,255]. Fig. 4a depicts the graphical representation of equation (11) for inputs raging in the interval [0,B]. In this interval the maximum output of equation (11) is not constant (within the [0,B] interval) and depends on the A/B ratio. For this reason, equation (11) is transformed to equations (12) and (13), which retain a constant output, within the [0,B] interval, for all A/B ratios, as depicted in Fig. 4b and 4c.

$$G(x) = \frac{(B+A) \cdot x}{A+x} \quad \forall \, x \in [0, B], \; A, B \in \Re^{+} \tag{12}$$

$$H(x) = \frac{A \cdot x}{A+B-x} \quad \forall \, x \in [0, B], \; A, B \in \Re^{+} \tag{13}$$

Equations (12) and (13) can be used as adjustable mapping functions. Once B is defined according to the range of input data, varying A can result to different nonlinear curves, controlling the mapping between input and output. In the proposed method, the input range is [0,255] and thus B=255.
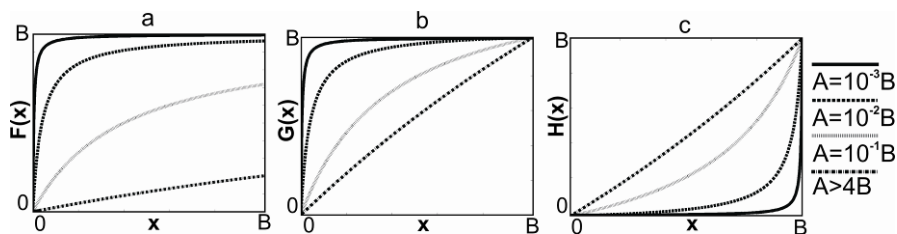


**Fig. 4.** a. Graphical representation of equation (11). b. Graphical representation of equation (12). c. Graphical representation of equation (13).

Equations (12) and (13) are the basis of the proposed tone mapping function. In order to have spatially modulated tone mapping, factor A is substituted by a function of the surround $S_{n,m}$ and x by $Y_{n,m}$. The equations of the method are:

$$Y_{n,m}^{out}(Y,S) = \begin{cases} \dfrac{\left(B+A(S_{n,m})\right) \cdot Y_{n,m}}{A(S_{n,m})+Y_{n,m}} & \forall S_{n,m} < 128 \\[4mm] \dfrac{A(S_{n,m}) \cdot Y_{n,m}}{A(S_{n,m})+B-Y_{n,m}} & \forall S_{n,m} \geq 128 \end{cases} \tag{14}$$

$$A(S_{n,m}) = \begin{cases} \left[M_{dark}+q(S_{n,m})\right] \cdot d(S_{n,m}) & \forall S_{n,m} < 128 \\[2mm] \left[M_{bright}+q(255-S_{n,m})\right] \cdot d(255-S_{n,m}) & \forall S_{n,m} \geq 128 \end{cases} \tag{15}$$

$$q(x) = \frac{x^{2}}{lobe} \tag{16}$$

$$d(x) = \frac{128}{128-x} \quad \forall x \in [0, 128) \tag{17}$$

Equation (14) is the basic equation of the method. It combines equation (12) for surround values lower than 128, and equation (13) for surround values greater than 128. This is selected because under-exposed regions are darker than the middle of the range (i.e. 128) and need to increase their intensity. Equation (12) increases non-linearly the lower intensity values, while barely affecting the higher ones (Fig. 5a). On the contrary, over-exposed regions are brighter than the middle of the range and need to decrease their intensity. Equation (13) decreases non-linearly the higher intensity values, while barely affecting the lower ones. Equation (15) is the modulation function that controls the non-linearity degrees of equation (14). It comprises the two adaptation characteristics of the HVS: the global and local adaptation. Global adaptation is expressed by the constants $M_{dark}$ and $M_{bright}$ which depend on the global statistics of the image. They determine the initial non-linearity degrees a* (for $S_{n,m}$=0) and b* (for $S_{n,m}$=255), respectively (Fig. 5a). Local adaptation is expressed by the use of equation (16) in equation (15), which determines the transition of the two non-linearities a* and b* to the linearity a, in the middle of the surround axis ($S_{n,m}$=128). This transition is regulated by the local surround value $S_{n,m}$, since it is a local variable parameter, and the constant *lobe*, which depends on the global statistics of the image (Fig. 5b). Equation (17) is a necessary correction factor which ensures the smooth continuity of equation (14) in the middle of the surround axis ($S_{n,m}$=128), in the transition point between equations (12) and (13). If equation (17) is omitted from equation (15), equation (14) will not be smooth in the middle of the surround axis, and artifacts will be created in the mapping of input values with $S_{n,m} \approx 128$.
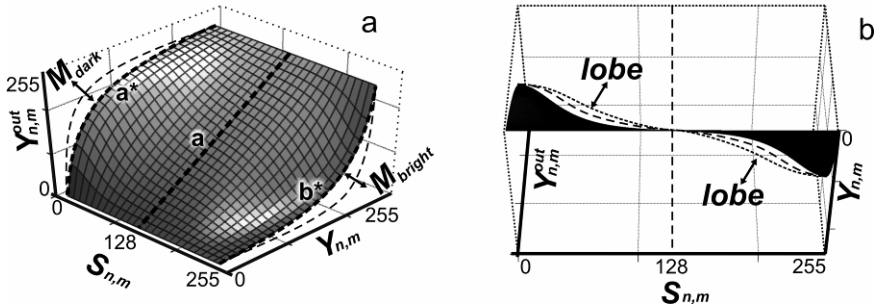


**Fig. 5.** a. 3D representation of equation (14). b. Different view-angle of a.

The coefficients $M_{dark}$, $M_{bright}$ and *lobe*, are statistically determined by the $Y$ component and adapt globally the mapping function to the input image, according to its dominant tones. The dominant tones of the image are defined by calculating the % percentage of pixels that belong to the interval [0, 85) for the dark tones, [85, 170) for the middle tones and [170, 255] for the light tones. These percentages indicate roughly whether the image is subjectively light, normal or dark. $M_{dark}$ and $M_{bright}$ determine the higher non-linearity degrees a* and b*, which can be applied in the under-exposed and over-exposed regions, respectively (Fig. 5a). Their values are in the interval [10,255] and are inversely proportional to the % percentage of the dark tones (*bin_low*) for $M_{dark}$ and light tones (*bin_high*) for $M_{bright}$.

$$M_{dark} = \frac{245}{100}(100 - bin\_low) + 10 \qquad (18)$$

$$M_{bright} = \frac{245}{100}(100 - bin\_high) + 10 \qquad (19)$$

A high percentage of dark tones, indicates a globally dark image, resulting to a low $M_{dark}$ value and a higher non-linear correction for the dark regions. A high percentage of light tones, indicates a globally bright image, resulting to a low $M_{bright}$ value and a higher non-linear correction for the bright regions.

Coefficient *lobe* determines the shape of the transition of the two non-linearities (Fig. 5b). Its value is in the interval [1,30] and is inversely proportional to the % percentage of the middle tones (*bin_middle*).

$$lobe = \frac{29}{100}(100 - bin\_middle) + 1 \qquad (20)$$

Low *lobe* values, limit the non-linearities to low surround values, leaving middle tones intact. High *lobe* values allow the non-linearities to affect the middle tones also.

## 3   Experimental Results

The proposed method was compared to MSRCR and ACE. These algorithms were selected because they belong to the same category with the proposed method and they are also inspired by the HVS. The proposed method was implemented in C. All algorithms were executed by an Intel Pentium 4 at 3 GHz, under Windows XP. The implementation of the MSRCR that was used in the evaluation is the commercial software Photoflair [6]. The parameters of the MSRCR are the default ones as reported in [3] (3 scales of 5, 20 and 240 pixels with equal weights, a color restoration process and no auto-level post-processing). The implementation of ACE can be found in [7]. The parameters that are used in the testing are the defaults: sub-sampling factor=8, slope=5, distance alpha=0.01, and a dynamic tone reproduction scaling of WP+GW. The implementation of the proposed algorithms is available in [8]. All the results depicted from the proposed method were derived by the parameters described in section 2.2, without any manual tuning.

Fig. 6 depicts the results of the algorithms for three images with strong under-exposed or over-exposed regions. In the first image, the proposed method extracts more details in the under-exposed region. MSRCR and ACE are affected by the strong intensity transition and fail to extract details in some of the under-exposed regions. The second image has both an under-exposed and an over-exposed region. MSRCR is heavily affected by the sharp intensity transition and extracts a strong halo effect. ACE tends to lighten the correctly exposed regions. The proposed method does not extract a hallo effect and is the only one that compensates for the over-exposed region of image 2. The proposed method has a faster execution time compared to the other two methods and needs approximately 1 sec for the

**Fig. 6.** Line 1: Results for image 1. Line 2: Magnified portion of image 1. Line 3: Results for image 2. Halo effects are marked with "H". Line 4: Magnified portion of image 2. Compensated over-exposed area is marked with a circle. Line 5: Results for image 2. Line 6: Magnified portion of image 3.

rendition of a 1-MPixel image. Image 3 has both under/over-exposed regions. It is clear that the proposed method compensates better for both these regions, in comparison to the other methods.
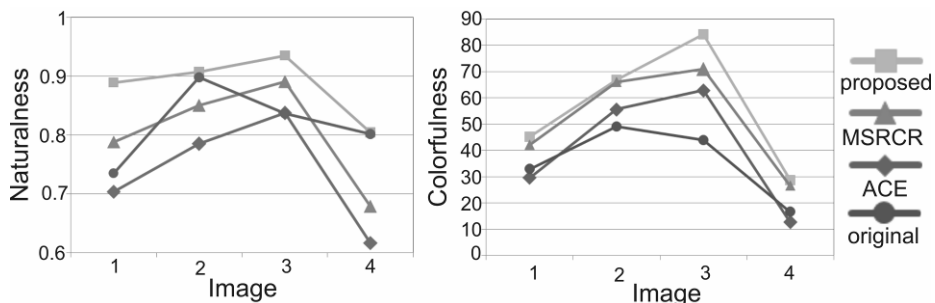
**Fig. 7.** Naturalness and colorfulness of the images of Fig. 7



**Fig. 8.** The four images that were used in the experiment of Fig. 8 and the results of the compared methods

The main objective of the proposed method is to enhance the original image and produce a better looking image to the human observer. For this reason, two image quality metrics, which were proposed after psychophysical experimentation with human observers, were used in the evaluation. The first metric is *naturalness*, which is the degree of correspondence between human perception and reality world and has a value in the interval [0,1], with 1 accounting for the most natural image. The second metric is *colorfulness*, which presents the color vividness degree of an image. High colorfulness values indicate high color vividness. These metrics were found to have strong correlation with the perception of the human observer [9] and were successfully used in the enhancement algorithm of [10]. The algorithms were applied to the four images of Fig. 8, and both metrics were calculated for their results and for the original image. The four images were selected because they have under-exposed and over-exposed regions. The calculated naturalness and colorfulness are depicted in Fig. 7.

The proposed method achieved higher degrees of naturalness and colorfulness for all the four images, indicating that its results are more probable to be ranked first by a human observer.

## 4   Conclusions

A new method for spatially modulated tone mapping has been presented in this paper. Its main objective is to enhance the under/over-exposed regions of images, while leaving intact the correctly exposed ones. The method utilizes orientation kernels, similar to the orientation cells of the HVS, in order to achieve varying surrounds that adapt their shape to the local intensity distribution of the image. Thus, no averaging is performed between two regions with strong intensity differences. This, results to the elimination of halo effects, which are a consequence of the wide surrounds. The surround of every pixel, regulates the tone mapping function that will be applied to the pixel. This function is inspired by the shunting characteristics of the center-surround cells of the HVS. The proposed method exhibits at least comparable and many times better results than other established methods in the same category. More importantly, the execution times of the proposed method are lower than those of the existing ones.

## References

1. Battiato, S., Castorina, A., Mancuso, M.: High dynamic range imaging for digital still camera: an overview. Journal of Electronic Imaging 12, 459–469 (2003)
2. Land, E.: The Retinex. American Scientist 52(2), 247–264 (1964)
3. Jobson, D.J., Rahman, Z., Woodell, G.A.: A multi-scale Retinex for bridging the gap between color images and the human observation of scenes. IEEE Trans. Image Processing 6, 965–976 (1997)
4. Rizzi, A., Gatta, C., Marini, D.: A new algorithm for unsupervised global and local color correction. Pattern Recognition Letters 24, 1663–1677 (2003)
5. Ellias, S., Grossberg, S.: Pattern formation, contrast control and oscillations in the short term memory of shunting on-center off-surround networks. Biological Cybernetics 20, 69–98 (1975)

6. Truview (2007), http://www.truview.com/
7. Eidomatica (2007), http://eidomatica.dico.unimi.it/ita/ricerca/ace.html
8. Electronics, http://electronics.ee.duth.gr/vonikakis.htm
9. Hasler, S., Susstrunk, S.: Measuring colorfulness in real images. In: Proc. SPIE Electron. Imag.: Hum. Vision Electron. Imag. VIII, SPIE 5007, pp. 87–95 (2003)
10. Huang, K.-Q., Wang, Q., Wu, Z.-Y.: Natural color image enhancement and evaluation algorithm based on human visual system. Computer Vision and Image Understanding 103, 52–63 (2006)