

Bio-inspired color image enhancement

Laurence Meylan and Sabine Süsstrunk

LCAV, EPFL, Ecole Polytechnique Fédérale de Lausanne, CH-1015 Switzerland

ABSTRACT

Capturing and rendering an image that fulfills the observer's expectations is a difficult task. This is due to the fact that the signal reaching the eye is processed by a complex mechanism before forming a percept, whereas a capturing device only retains the physical value of light intensities. It is especially difficult to render complex scenes with highly varying luminances. For example, a picture taken inside a room where objects are visible through the windows will not be rendered correctly by a global technique. Either details in the dim room will be hidden in shadow or the objects viewed through the window will be too bright. The image has to be treated locally to resemble more closely to what the observer remembers. The purpose of this work is to develop a technique for rendering images based on human local adaptation.

We take inspiration from a model of color vision called Retinex. This model determines the perceived color given spatial relationships of the captured signals. Retinex has been used as a computational model for image rendering. In this article, we propose a new solution inspired by Retinex that is based on a single filter applied to the luminance channel. All parameters are image-dependent so that the process requires no parameter tuning. That makes the method more flexible than other existing ones. The presented results show that our method suitably enhances high dynamic range images.

Keywords: Retinex, High Dynamic Range Compression, Color Image Enhancement

1. INTRODUCTION

Digital cameras are becoming increasingly important and are now widely used by professional as well as non-professional photographers. Not only do they provide a convenient way to store images, but they also allow the users to manipulate their pictures easily using a personal computer. While digital cameras bring a simple way to manually treat images, automatically rendering an image that satisfy the customers' expectations remains a difficult task. In practice, the captured image often differs from what the observer remembers. This is due to the fact that the camera captures the physical values of light, while the observer perceives the result of the processing of his/her visual system. Indeed, the human visual system has a complex non-linear mechanism that determines the perceived color by spatial comparisons of color signals across a scene.

The human visual system easily handles scenes composed of a wide range of light intensities (eg. sunny outside taken through the window of a dark room). Such scenes are said to have a high dynamic range when the ratio between the highest luminance to the lowest luminance exceeds the one of the capture or output device. The human observer deals with high dynamic range scenes by adapting locally to each part of the scene and thus is able to retrieve details in low luminance as well as high luminance areas. Using a digital device is more problematic. The dynamic range of the scene has to be compressed, which often causes the captured image to lack detail in areas of low and high illumination.

Some recent developments made possible the capture of high dynamic range scenes.¹ The principle is to capture multiple pictures of the same scene with different exposure times. A so-called radiance map is built from the acquired pictures. This technique allows to obtain an accurate estimation of the scene despite the capture device limitation. Nevertheless, the problem of mapping the high dynamic range values expressed in floating point to the low dynamic range of the output device remains.

Further author information: E-mail: laurence.meylan@epfl.ch

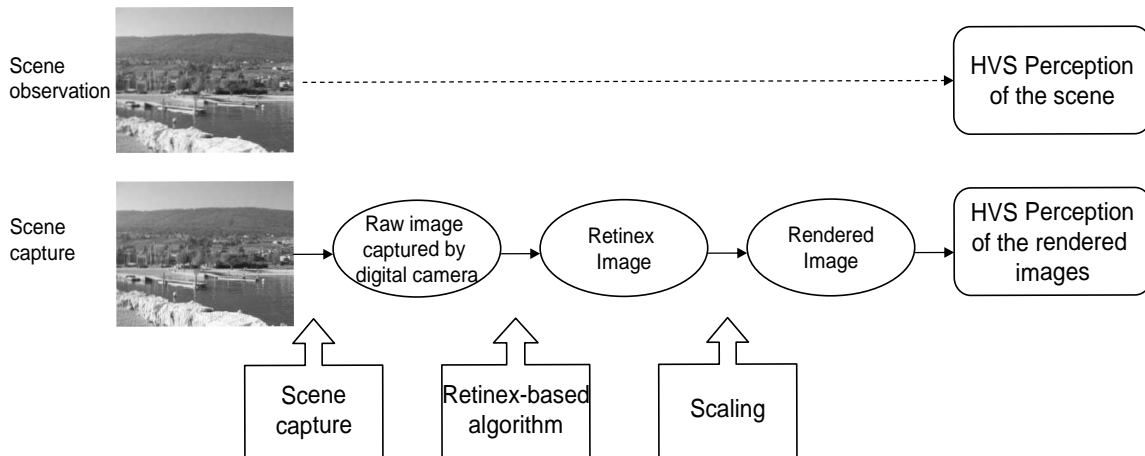


Figure 1. The purpose of the framework is to make the rendered image close to the observer percept of the scene. The framework has three main steps. i) The scene is captured ii) The Retinex-based algorithm is applied on the raw linear image iii) The Retinex image is scaled to the range of the output device

In this article, we address the problem of rendering images according to human preferences (illustrated in Figure 1). We believe that this problem can be solved by an algorithm that mimics the human visual system’s lateral interactions and local contrast adaptation. We take inspiration from an existing theory of color vision called Retinex, which has proved to compress dynamic range efficiently.^{2,3} Our aim is to develop a framework for rendering high dynamic range images so that the rendered image is close to the observers’ expectations. The treated images are either radiance maps⁴ or raw images shot with a “Canon Power Shot G2.” The core of the framework is a Retinex-based algorithm that compresses the dynamic range while maintaining local contrast in the low and high luminance parts of the image. We argue that it is possible to display images that accurately represent what the observer remembers by means of local adaptation. Indeed, since the dynamic range of a single neuron is usually lower than that of outdoor scenes,⁵ the human visual system faces the same problem. The way humans deal with high dynamic range scenes is by adapting locally to different parts of the scene in order to form a percept where all details are visible. The latter considerations show that there is no need to have high dynamic range output devices to render “nice” images. An appropriate local rendering operator is sufficient.

This article is structured as follows: Section 2 describes the main aspects of the Retinex theory and its evolution toward computational models. Previous computational models inspired by Retinex are presented. Section 3 introduces a new framework for rendering images and compares the technical differences with previously developed methods. Results are shown in section 4 and section 5 concludes the article.

2. RETINEX, THEORY AND APPLICATIONS

Retinex theory of color vision was developed by Land.⁶ It is based on a series of experiments carried out with a surface made of color patches* and three controllable independent light sources. During the experiments, it was possible to illuminate the surface and measure the signals coming from the different color patches. Using this setting, Land showed that two exact same stimuli (same color signal value) could lead to different sensations of color. Major conclusions about human color perception were drawn from these experiments. In particular, the perceived color of a unit area is determined by the relationship between this unit area and the rest of unit areas in the image independently in each waveband and does not depend on the absolute value of light intensities.⁷

Because Retinex theory is a simplified version of human color processing, it can conveniently be adapted for computational image rendering models. These Retinex-based algorithms can be classified in three classes.

*These surfaces were called “Mondrian” in reference to the dutch painter Piet Mondriaan

The first one is the class of paths-based algorithms where the new pixel value depends on the computation of ratios and products along paths in the image. The second class contains all algorithms where the new pixel value depends on the recursive comparison of surrounding pixels. The recursion depth determines the extent of interactions. The last class includes the center/surround versions of Retinex where the new pixel value depends on the ratios of the pixels included in a surrounding area. The ratios values are weighted by a surround function.

Class 1: Path version

The first Retinex computational model was developed by Land^{6,7} and is based on computing the product of ratios between pixels value along a set of paths in the image. The new pixel value is computed as the average over all paths. Further developments introduced randomly distributed paths using Brownian motion.⁸ The Retinex path version has also been extensively studied from a theoretical point of view. Brainard and Wandell described formally the path version using stochastic methods.⁹ They studied the convergence of the method for a large number of very long paths and found that for infinitely long paths, the Retinex image tends to the original image normalized by its maximum value. Horn extended the 2-dimensional path version of Land using a two-dimensional Laplacian.¹⁰ Hurlbert¹¹ formalized the lightness problem (retrieving sensation of color from absolute intensities) and showed that most of the algorithms (eg. Land, Horn) that aimed to solve it could be expressed with a single equation. This equation is a Poisson equation and can be solved using the Green function.

Practically, the main problem of Retinex path-based methods is its high computational complexity and the free parameters, such as the number of paths, their trajectories, and their lengths.

Class 2: Recursive version

An evolution of the path version was provided later by Frankle and McCann.¹² Its matlab code was published by Funt et al.¹³ In this version, the paths calculation are replaced by matrix computation where ratios and products are processed in parallel instead of being added and compared sequentially. The distance function previously determined by the paths length is defined by the number of iterations. This method works well for rendering images, provided that the number of iterations is well chosen. When this number increases, the image tends to its original version normalized by the maximum. There is a number of iterations for which the resulting image is visually the best but this number is hard to find automatically. Although some methods to automatically tune the parameters have been developed,¹⁴ finding the optimal number of iterations remains an issue.

Despite the undefined number of iterations problem, a number of authors have worked on developing further the “Frankle-McCann Retinex”. Sobol included a ratio modification operator to better compress large ratios while enhancing small ratios.¹⁵ Drago et al. aimed to reduce black halos around light sources.³ They modified the Retinex version of Sobol using a smooth function for the ratio modification operator instead of a fixed clipping value. In addition, they introduced a new reset value, more efficient for high dynamic range scenes. All these variants of the “Frankle-McCann Retinex” share the problem of undefined parameters. Another drawback of the recursive method is that the final result is hard to predict.

Class 3: Center/Surround version

A non iterative version of Retinex was proposed by Rizzi et al.^{16,17} Their algorithm computes the ratios of each pixel with all other pixels in the image or in a surrounding area. The influence of one pixel on the value of the treated pixel is controlled by a distance function. The increase in local contrast is controlled by a so-called relative lightness function. Their algorithm reproduces well the principles of Retinex (ratio, product, average) without the drawbacks of the recursive version. However, some parameters remain to be defined, such as the size of the surround area. Unfortunately, due to the high computational complexity of this method, the surround size is more often defined by the restriction in computational time than by some image-dependent features.

Another center/surround method is provided by Rahman et al.¹⁸ In their method, the ratio is not computed pixel after pixel. Instead, the new pixel value is given by the ratio of the treated pixel to the weighted average of the surrounding pixels. The weighted average of surrounding pixels is computed with a filter, whose weights

are defined by a surround function. They found that the best results were obtained by averaging the three images resulting from three different surround sizes. Later, they added a color restoration to overcome a graying out effect caused by the method. Their algorithm works correctly for limited dynamic range, but introduces important artifacts when applied on high dynamic range images.

The most important differences between the Retinex-based computational models are the order in which the pixels are explored and the influence of surrounding pixels, weighted by the distance. A common issue is the optimal choice of parameters.

In this article, we present a center/surround Retinex-based algorithm. The new pixel value is computed with a single filter. The filter coefficients are computed by a surround function, whose shape is defined by the image. This algorithm is the core of a global framework for image rendering, presented in section 3. Results in section 4 show that our framework efficiently compresses the dynamic range of images while improving detail visibility.

3. A FRAMEWORK FOR ENHANCING COLOR IMAGES OF NATURAL SCENES

In this section, we propose a framework for enhancing color images. Each step is described, from the image captured to the display on the output device. This framework can be used to enhance traditional 24-bit images as well as to compress high dynamic range images acquired in raw format or derived from a multiple exposures technique.

3.1. Capturing the images

Evaluation and testing were carried out on a set of high dynamic range images. The data set is composed of a variety of outdoor and indoor scenes. Among the data set are images of two different origins. Some are images taken with a “Canon Power Shot G2” camera. The others come from an open database of calibrated high dynamic range color images⁴ that were built using a multiple exposures technique. They are uploaded into the framework with the help of the tools provided on the database web page [†].

Although we use only two formats, any images that reflect the scene radiance linearly can be used as input. With linear images, we ensure that the algorithm input corresponds to the color signal that reached the observer’s eye.

3.2. The framework

The whole framework is described in Figure 2. As specified above, the input image is linear with respect to focal plane irradiance. Since human sensitivity to light is approximately logarithmic, the first step is to apply a logarithm on the linear image. The second step is a global tone mapping that was introduced for extremely high dynamic range images. Step 3 separates the image in three components. At step 4, the “Retinex with 1 filter” algorithm is applied on the luminance channel. The treated components are transformed back into an RGB image at step 5. Step 6 scales the image to a displayable range.

3.2.1. Step 2: Introduction of a global tone-mapping

To understand the occasional need of additional global tone mapping, the reader should first understand the trade-off between dynamic range compression and correct rendering of images involved with the “Retinex with 1 filter” method. The smaller the filter is, the better is the compression; but small filters introduce more artifacts than large filters. The compression artifacts are graying out effects in smooth areas and apparition of halos around sharp edges. In most cases, the input image can efficiently be rendered by step 3 without additional global tone mapping. Only extremely high dynamic range images require this additional treatment to avoid compression artifacts. Although these cases are rare, we introduced a global treatment to handle images exceeding a given dynamic range.

[†]<http://white.stanford.edu/hdri/>

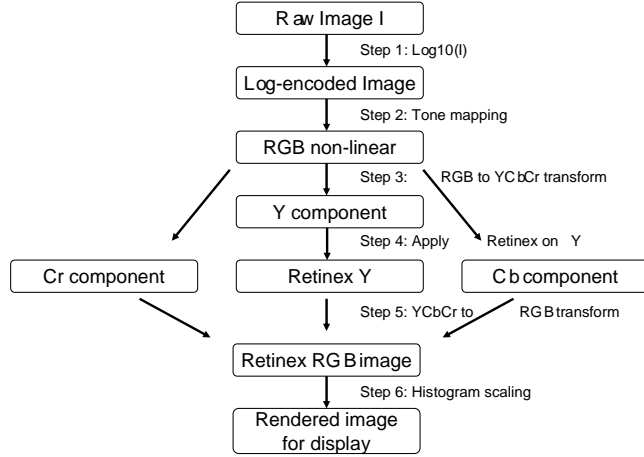


Figure 2. The input image is first log-encoded. Then, a global tone mapping is applied when needed. The RGB non-linear image is transformed into a YCbCr image. A Retinex-based algorithm is applied on the luminance channel. The resulting image is transformed back to RGB. It is then scaled and displayed on the output device.

$$\log_{10}(I)' = \log_{10}(I)^{\frac{1}{x}} \quad (1)$$

where the value of x is determined by global statistics of the image such as mean and variance. This step can be equated to the first adaptation state in the human visual system where an adaptation to global illumination takes place.

3.2.2. Step 3: Separating the color non-linear image in three component

The non-linear image obtained after the previous steps is transformed into the Luminance Chrominance Chrominance YCbCr color space so that the luminance channel can be treated independently. This choice was motivated by the fact that applying “Retinex with 1 filter” on R,G and B components led to color artifacts. We only treat the luminance component Y and do not change the chrominance components Cb and Cr.

3.2.3. Step 4: “Retinex with 1 filter”

Our “Retinex with 1 filter” takes inspiration from the Retinex theory. It determines the new pixel value by computing the ratio of the treated pixel to a weighting average of other pixels in the image.

Let the Retinex image for the luminance channel R_Y be defined as:

$$R_Y = \log_{10}(I)'_Y - \log_{10}(mask) \quad (2)$$

where $\log_{10}(I)'_Y$ is the Y component of the non-linear image computed at step 2 and transformed into YCbCr color space. $mask$ is computed by convolving the I_Y component of the original image I with a surround function F .

$$mask = I_Y * F \quad (3)$$

Based on the previous work of Rahman and Jobson,^{18,19} we chose to use a surround function whose weights are defined by an addition of Gaussian functions. Our method differs from their in the way all spatial constants are included in a single filter, whereas Rahman and Jobson compute a mask for each size of Gaussian and take

the average for the final mask. In addition, the way we define the size of Gaussians is more flexible. The Gaussian spatial constants are image-dependent, while Rahman and Jobson use fixed scales.

The surround function is defined as follow:

$$F_{(x,y)} = \sum_{i=1}^{\lfloor \log_2(K) \rfloor} e^{-\frac{(x^2+y^2)}{2^{2i}}} \quad (4)$$

$$K = \frac{\max(\text{size}(I))}{8} \quad (5)$$

where (x, y) is the coordinate of the filter entry and $\max(\text{size}(I))$ is the size of larger pixel dimension of I_c . x , respectively y go from 1 to K . Since F is four-fold symmetric, eq. 4 gives the weights of only one quadrant. The other quadrants are just a flipped copy of it.

Eq. 4 is a sum of Gaussians, whose spatial constants are powers of 2. K is the size of the filter basis. This parameter will decide the dynamic range compression. If K is large, the compression is low. If K is too small, artifacts such as black halos appear. The global tone mapping applied before the Retinex local operator takes care of images that need a greater compression. The selected K (eq. 5) shows good performance on all images and appears to be an appropriate value for the surround size of local adaptation.

Figure 3 shows the filter F . The pointed shape allows good compression and artifacts are reduced by the large base of the filter.

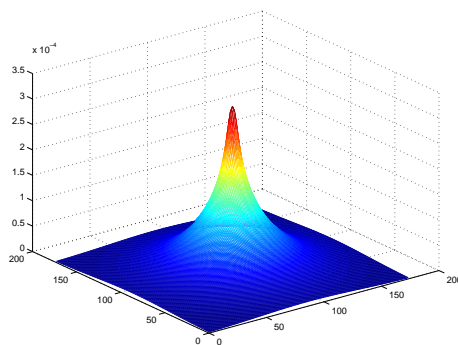


Figure 3. Filter F

The filter F thus obtained is convolved with the image to form the mask. To avoid convolution with too large filters, we use a multiplication in the Fourier domain.

$$\text{mask} = DFT^{-1}(DFT(F^{pad}) * DFT(I_Y^{pad})) \quad (6)$$

where F^{pad} and I_Y^{pad} are F respectively I_Y padded with zeros so that the number of pixels in each dimension are doubled. Without the zero-padding, the result of the multiplication in the Fourier domain is equivalent to a circular convolution. Using a padding makes eq. 3 and eq. 6 equivalent.



Figure 4. Left: Log-encoded original image. **Middle:** Image rendered with “Retinex with 1 filter” method. **Right:** Image treated with the multiscale with color restoration (Demonstration version).

3.2.4. Step 5: Transforming the result back to RGB color space

After having computed the three components, the YCbCr image is transformed back to RGB color space using a simple matrix transformation.

3.2.5. Step 6: Histogram Scaling

The image is then scaled to a displayable range using a linear histogram scaling. The blackest and the whitest digital values are found in the histogram. The non significant values at the tails are cut down. All values are scaled linearly so that they fit into the dynamic range of the display device.

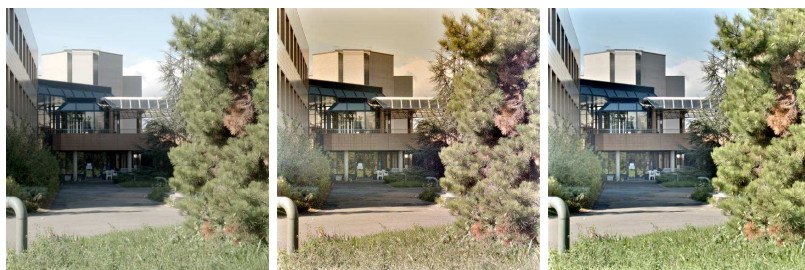


Figure 5. Left: Log-encoded original image. **Middle:** “Retinex with 1 filter” applied on R,G and B components independently. **Right:** “Retinex with 1 filter” applied on the luminance channel of a YCbCr-encoded image.

3.3. Discussion

The framework described above takes inspiration from Retinex and previous works on that topic. In particular the computation of the single channel Retinex share a few similarities with the multiscale Retinex of Rahman et al.¹⁸

A difference is that all our parameters are image dependent. In particular, the spatial constant are not fixed but gets adapted depending on the image. Another major difference is that we include all the parameters in the coefficients of a single filter, whereas Rahman et al. compute the single Retinex for three fixed scales and average the results to obtain the final image. In addition, Rahman et al. introduce a color restoration factor to overcome the graying out of smooth areas. A drawback is that this factor introduces unnatural color or artifacts in dark areas as illustrated in Figure 4. By applying “Retinex with 1 filter” on the luminance channel only, we avoid these problems of unnatural colors and graying out effects. The benefit of restricting the computation to the Y component of a YCbCr-encoded image is illustrated in Figure 5.



Figure 6. Left: Log-encoded image. Right: “Retinex with 1 filter” image

4. RESULTS

In this section, we present the result of our method applied on different images (Figure 6)[‡]. For each pair of images representing the same scene, the image on the left is the original high dynamic range image that was compressed using a global tone mapping technique. A logarithm function was applied followed by a linear scaling. The one on the right is the image obtained with the “Retinex with 1 filter” method.

The effect of the algorithm is clearly visible. The algorithm increases local contrast. The pointed shape of the single filter allows to retrieve small details in shadows while conserving a good rendition of the image. It also provides a sharpening effect.

A common problem to most Retinex-based algorithms is the apparition of halos around high contrasted areas. In the currently presented version, it is reduced by choosing an appropriate filter size and by introducing global tone mapping for extremely high dynamic range images. The next two images are examples where artifacts are still visible.

Figure 7 shows an extreme high dynamic range scene with important features in both lights and shadows. The image is quite difficult to enhance. The black halos problem of Retinex becomes visible around the hair of the person. The front-head is darker than the rest of the image.



Figure 7. Left: Log-encoded image.**Right:** “Retinex with 1 filter” image

In Figure 8, a white halo is visible in the middle. The door seems to be illuminated by an additional light source.

The halo problem is due to the circular shape of the filter. We are currently implementing a version that uses a filter whose base has an adaptive shape. The shape of the surround will not be circular as in the current version but will depend on the local contrast in the image. Large surrounds should be used for low contrast areas while small surrounds should treat high contrasted areas.

5. CONCLUSION

The problem of rendering an image has often been treated as a global tone mapping problem where a function determines the new value of each pixel without taking into account spatial relations in the image. Global tone mapping techniques are fast as they can be implemented with a look-up table but the produced results are not always satisfying. Global compression can lead to a loss of detail visibility in areas of low or high luminance. Local operators are required to achieve better performances. The idea of local operators is to compute the new pixel value not only depending on its actual value but also using surrounding information. A solution to design

[‡]All images are available at <http://ivrgwww/index.php?name=EI04>



Figure 8. Left: Log-encoded image.**Right:** “Retinex with 1 filter” image

good local operators is to take inspiration from features of the human visual system. Indeed, the human visual system efficiently adapts to different lighting conditions.

In this paper, we present a framework for enhancing natural color images that takes inspiration from a model of human color vision called Retinex. We collected a variety of high dynamic range images for evaluation purpose. The input to the framework are linear RGB images. The first two steps apply a global non-linear function on the input values. The non-linear image is then transformed into YCbCr color space and the “Retinex with 1 filter” method is applied on the luminance component. The “Retinex with 1 filter” method computes a new intensity value for each pixel, given by the ratio of the treated pixel to the weighted average of a surrounding area. The weighted average is computed by convolving the original image with a surround function defined as a sum of Gaussians. The computations are effected in the Fourier domain to reduce the high number of computations due to the large filters involved in the convolution. The new image thus obtained is transformed back to RGB color space and scaled to the output device dynamic range.

The aim of the framework is to render the original image so that it reproduces what the observer remembers of the captured scene. That involves compressing the dynamic range while conserving detail visibility in all parts of the image. The results presented in section 4 show that our method renders high dynamic range images efficiently while retrieving details in dark and bright areas.

ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation under grant number 21-101681.

REFERENCES

1. P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” in *SIG-GRAPH’97*, pp. 369–378, 1997.
2. J. McCann, “Lessons learned from mondrians applied to real images and color gamuts,” in *The 7th Color Imaging Conference: Color Science, Systems and Applications*, 1999.
3. F. Drago, W. L. Martens, K. Myszkowski, and N. Chiba, “Design of a tone mapping operator for high dynamic range images based upon psychophysical evaluation and preference mapping,” in *IS&T/SPIE Electronic Imaging 2003. The Human Vision and Electronic Imaging VII Conference.*, **5007**, April 2003.
4. F. Xiao, J. DiCarlo, P. Catrysse, and B. Wandell, “High dynamic range imaging of natural scenes,” in *The Tenth Color Imaging Conference*, (Scottsdale, USA), 2002.
5. S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, and D. P. Greenberg, “A multiscale model of adaptation and spatial vision for realistic image display,” *Computer Graphics* **32**(Annual Conference Series), pp. 287–298, 1998.

6. E. Land, "The Retinex," *American Scientist* **52**, pp. 247–264, 1964.
7. E. H. Land, "Recent advances in retinex theory," *Vision Research* **26**(1), pp. 7–21, 1986.
8. A. Rizzi, D. Marini, L. Rovati, and F. Docchio, "Unsupervised corrections of unknown chromatic dominants using a brownian-path-based retinex algorithm," *Journal of Electronic Imaging* **12**, pp. 431–440, July 2003.
9. D. Brainard and B. Wandell, "Analysis of the retinex theory of color vision," *Journal of the Optical Society of America* **3**, pp. 1651–1661, 1986.
10. B. K. P. Horn, "Determining lightness from an image," *Computer Graphics and Image Processing* **3**, pp. 277–299, 1974.
11. A. Hurlbert, "Formal connections between lightness algorithms," *Journal of the Optical Society of America* **3**, pp. 1684–1692, 1986.
12. J. Frankle and J. McCann, "Method and apparatus for lightness imaging." US Patent #4, 384,336, May 1983.
13. B. Funt, F. Ciuera, and J. McCann, "Retinex in matlab," in *IS&T/SID Eighth Color Imaging Conference*, pp. 112–121, (Scottsdale), 2000.
14. B. Funt, F. Ciurea, and J. McCann, "Tuning retinex parameters," in *IS&T/SPIE Electronic Imaging 2002. The Human Vision and Electronic Imaging VII Conference.*, **4662**, pp. 358–366, (San Jose), 2002.
15. R. Sobol, "Improving the retinex algorithm for rendering wide dynamic range photographs," in *IS&T/SPIE Electronic Imaging 2002. The Human Vision and Electronic Imaging VII Conference.*, **4662**, pp. 341–348, (San Jose), 2002.
16. A. Rizzi, C. Gatta, and D. Marini, "A new algorithm for unsupervised global and local color correction," *Pattern Recognition Letters* **24**, pp. 1663–1677, July 2003.
17. A. Rizzi, C. Gatta, and D. Marini, "Color correction between gray world and white patch," in *IS&T/SPIE Electronic Imaging 2002. The human Vision and Electronic Imaging VII Conference.*, **4662**, pp. 367–375, (San Jose), 2002.
18. Z. Rahman, D. Jobson, and G. Woodell, "Retinex processing for automatic image enhancement," in *IS&T/SPIE Electronic Imaging 2002. The Human Vision and Electronic Imaging VII Conference.*, **4662**, pp. 390–401, (San Jose), 2002.
19. D. Jobson and Z. Rahman, "Properties and performance of a center/surround retinex," *IEEE Transactions on image processing* **6**, pp. 451–461, March 1997.