Random Spray Retinex Extensions Considering Region of Interest and Eye Movements

Midori Tanaka

College of Liberal Arts and Sciences, Chiba University, Chiba, Japan E-mail: midori@chiba-u.jp

Matteo Paolo Lanaro

Department of Computer Science, University of Milano, Milano, Italy

Takahiko Horiuchi

Graduate School of Engineering, Chiba University, Chiba, Japan

Alessandro Rizzi

Department of Computer Science, University of Milano, Milano, Italy

Abstract. The Random spray Retinex (RSR) algorithm was developed by taking into consideration the mathematical description of Milano-Retinex. The RSR substituted random paths with random sprays. Mimicking some characteristics of the human visual system (HVS), this article proposes two variants of RSR adding a mechanism of region of interest (ROI). In the first proposed model, a cone distribution based on anatomical data is considered as ROI. In the second model, the visual resolution depending on the visual field based on the knowledge of visual information processing is considered as ROI. We have measured actual eye movements using an eye-tracking system. By using the eye-tracking data, we have simulated the HVS using test images. Results show an interesting qualitative computation of the appearance of the processed area around real gaze points. © 2019 Society for Imaging Science and Technology.

[DOI: 10.2352/J.ImagingSci.Technol.2019.63.6.060403]

1. INTRODUCTION

Scenes perceived by the human visual system (HVS) may differ from the corresponding ideal photometric image. The perceived color will depend on the chromatic content of the neighboring objects or background. This phenomenon is named as "the locality of color perception."

Retinex by Land and McCann [1, 2] was the first computational model that explained how the HVS perceives color. The model presented in this article is a particular variant of a Milano-Retinex algorithm [3], named Random spray Retinex (RSR) [4]. The idea is to test a hybrid approach starting from the Retinex computational model, with the implementation of a particular method of image exploration.

Retinex is a model that accounts for the spatial formation of the visual color sensation in our vision system

Received July 14, 2019; accepted for publication Mar. 10, 2019; published online Nov. 26, 2019. Associate Editor: Susan Farnand. 1062-3701/2019/63(6)/060403/6/\$25.00

[1, 2]. The idea has been successful and has generated a wide set of implementations for the original method [5–7] and for the many versions that arose from the later simplified version by Land [8, 9]. A detailed description of Retinex history and of the differences among the many variants can be found in [10, 11].

Stated the biological inspiration of all the many existing Retinex models, a very important detail is their quantitative versus qualitative implementation. In fact, for a quantitative implementation, two different phases have to be inserted in the pipeline, the first at the beginning for a careful calibration of the input "stimulus" and the second at the end for the calibration of the final computed appearance. A more detailed description of this important point can be found in [12]. A qualitative implementation presented here has many more degrees of freedom since it does not aim at modeling the exact visual response of HVS but rather computing an image "enhanced" in the direction of the spatial changes performed by HVS.

The main characteristic of Retinex models is the spatial recomputation of each pixel value performed with a series of ratio operation in luminous intensity difference across the image. This idea of exploring the image content through paths is a common feature for many Retinex implementations (not all) and has been modeling studies [7, 13–19].

Retinex image exploration evokes spontaneously the natural exploration of the scene that derives from the foveal distribution of the retinal maximal acuity in HVS. The presented approach is an attempt to connect these two different mechanisms in order to test if this joint computation leads to an improvement in the final overall enhancement. Ohtera et al. proposed a Retinex based on the region of interest (ROI) [20]. This method was an image processing model based on center/surround processing, which is similar to the simplified version of Retinex [8, 9].

[▲] IS&T Members.

Instead of starting from a path Retinex version, we have decided to start from RSR [4]. RSR, a Milano-Retinex implementation, substitutes random paths with random sprays. This simplifies the implementation of a further spatial "movement." In this case, we have added a deeper exploration starting from the lightest point in the spray as will be presented in detail in the next sections.

This article proposes two models of the HVS by performing the RSR along ROI. In the first proposed model, a cone distribution based on anatomical knowledge is considered as ROI. In the second model, the visual resolution depending on the visual field based on the knowledge of the visual information processing is considered as ROI. The feasibility of the proposed models is verified by experiments.

2. MODELS CONSIDERING ROI

In this section, we propose two simulation models. In the first model, a cone distribution based on anatomical knowledge is considered as ROI. In the second model, the visual resolution depending on the visual field based on the knowledge of visual information processing is considered as ROI. Both models are built based on the principle of the RSR. Therefore, in Section 2.1, we will start with a description of the RSR.

2.1 Traditional RSR Model for Fixation Point

Consider an RGB digital image and a collection of N paths $\gamma_1, \ldots, \gamma_N$ composed of ordered chains of pixels starting with j_k and ending in i. Let n_k be the number of pixels traveled by the kth path, γ_k . The RSR [4] is a new implementation of the original Retinex model [1, 2, 8, 9] motivated by the results of the mathematical analysis of Retinex performed in [3]. In the RSR, the role of the path, γ_k , traveling n_k pixels and ending in the target, i, is performed by $Spray_k(i)$, a spray composed of n_k pixels and centered at i. In particular, the RSR explores the locality of a digital image by selecting pixels with dense samples called random sprays, located in a circular area around the pixel, i, with the density decreasing as a function of the distance from i. The function, $Spray_k$ (i), performs a pixel selection composed of n_k pixels centered at i to reveal different local filtering properties. Note that in the RSR, varying the number of points per spray with kis not required; hence, n and not n_k is used to denote the number of pixels per spray. To generate a spray, we first obtain a uniform random distribution of values within [0, 1]. Then it is multiplied to extend to the intervals of $[0, 2\pi]$ and $[0, 2\pi]$ R], where R is a real positive number corresponding to the radius of the spray, yielding the distributions, RAND_n[0, R] and RAND_n [0, 2π]. (i_x, i_y) being the coordinates of a pixel, *i*, we can obtain the polar coordinates, (j_x, j_y) , of a pixel *j* in $Spray_k(i)$ with

$$\begin{cases} j_x = i_x + \rho \cos(\theta) \\ j_y = i_y + \rho \sin(\theta) \end{cases} , \tag{1}$$

where $\rho \in \text{RAND}_n[0, R]$ and $\theta \in \text{RAND}_n[0, \pi]$. Sprays are then selected from N random precomputed sprays and used to search for the pixel with the highest intensity in all the

sprays. For each spray, the RSR performs N comparisons to determine the x_{Hk} pixel with the highest intensity traveled by the path, γ_k . By calculating a ratio between intensities I(i) for a pixel i and $I(x_{Hk})$ for the pixel x_{Hk} , the RSR recomputes the lightness, L(i), by averaging the contributions as follows:

$$L(i) = \frac{1}{N} \sum_{k=1}^{N} \frac{I(i)}{I(x_{H_k})}.$$
 (2)

2.2 Model Based on Cone Distribution

The RSR described in Section 2.1 is an algorithm for fixed viewpoints. In this section, we extend RSR to the function of ROI based on anatomical information. Anatomically, the cone cells present in the human retina are more distributed near the center of the retina (the fovea centralis), and they decrease rapidly as they move away from the fovea. We simulate the function for ROI by applying the RSR based on the density distributions of the cone cells.

Although individual differences are observed in the distributions of the cone cells, we assumed that the spatial density distributions of the cones follow a normal distribution in this study. In addition, there is evidence to suggest that the cone photoreceptor array plays a critical role in spatial sampling in the HVS in the region from approximately 1° to 10° [21–24]. Therefore, a visual field of 2° from the fixed point of view is set as ROI, and random numbers are generated in ROI to achieve a two-dimensional (2D) normal distribution as the processing pixels.

Let p be a pixel in ROI set for a fixed pixel, i, and the RSR algorithm calculates the lightness, L(i; p), as

$$L(i;p) = \frac{1}{N} \sum_{k=1}^{N} \frac{I(p)}{I(x_{H_k}(p))}, \exists p \in ROI(i).$$
 (3)

The lightness is calculated at all pixels, *p*, in ROI.

Figure 1 shows an example of a set of pixels, *p*, in ROI, indicated in blue color. The gray point at the center shows a fixed pixel, *i*. The pixels, *p*, are randomly generated in ROI to achieve a 2D normal distribution as the processing pixels.

2.3 Model Based on Visual Resolution

As a different approach from Section 2.2, we extend RSR to the function of ROI based on visual information processing. In particular, we focused on the resolution of the eye.

In the HSV, the first 2° of the visual angle has the best resolution performance, following which the performance gradually declines to 10° of the visual angle. To apply this property to the RSR described in Section 2.1, the visual acuity is considered to set a new spatially variable parameter depending on the visual angle from a fixed point. We newly introduce a weight, w(R), where R represents the Euclidean distance between the fixed point, i, and a pixel p in ROI as $R = \|i - p\|^2$. Here, we set ROI as the 10° visual degree

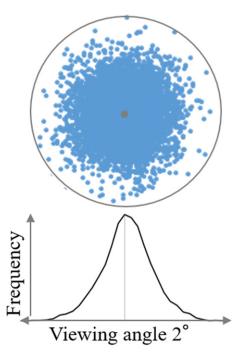


Figure 1. Set of points *p* in the ROI. (Top) Top view of the two-dimensional distribution. (Bottom) Distribution frequency histogram.

region. Then the weight, w(R), is defined as

$$w(R) = \begin{cases} 1.0, & (R \le r_2) \\ \frac{I(x_{H_k}(p)) - 1}{r_{10} - r_2} (R - r_2) + 1, & (r_2 < R \le r_{10}) \\ I(x_{H_k}(p)), & (r_{10} < R) \end{cases} , (4)$$

where r_2 is the radius of the 2° visual degree region and r_{10} is the outer radius of the 10° visual degree region. Then, the RSR algorithm calculates the lightness, L(i; p) as

$$L(i; p) = \frac{1}{N} \sum_{k=1}^{N} \frac{I(p)}{I(x_{H_k}(p))} w(\|i - p\|^2), \forall p \in ROI(i).$$
(5)

Note that unlike the model described in Section 2.2, this is defined for all the pixels, p, in ROI with a viewing angle of 10° . Figure 2 illustrates the definition of the weight, w(R). As shown in the figure, Eq. (4) represents that the weight, w(R), within a viewing angle of 2° is 1.0, and the RSR works completely. However, when the viewing angle is larger than 10° , the weight is constant $I(x_{H_k}(p))$, and the output of L(i; p) matches the original intensity, I(p), of a pixel p.

2.4 Eye Movement Processing

In Sections 2.2 and 2.3, we proposed a model for a fixation point and its ROI. In this section, a processing model for the case of the viewpoint moving with time will be described.

Basically, even when the viewpoint moves, RSR processing may be performed on each sampling point of the sequential viewpoint. However, note that we may look at the same point multiple times along the time sequence. Therefore, we can form two hypotheses. One hypothesis is that the effect of the RSR is repetitively amplified by viewing

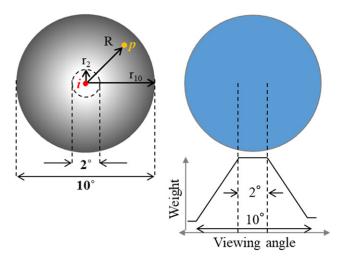


Figure 2. Definition of the weight, w(R). (Left) Illustration of the symbols. (Right) Illustration of the weight for the viewing angle.

multiple times, and the other one is that the effect of the RSR is constant regardless of the number of times. Because the former hypothesis needs a complex visual processing model considering latency memory and afterimage, this study adopted the latter hypothesis. However, in the processing for a pixel p in ROI, the model in Section 2.3 may have different weights depending on the distance to the fixation point, i, and may have different outputs of lightness. Therefore, we adopt a larger weight, w.

3. EXPERIMENT

To verify the feasibility of our proposed models, we conducted simulation by measuring ROI based on actual eye tracking.

3.1 Measurement of ROI

To simulate the ROI from the actual eye movement, we conducted an experiment to measure visual point movement when an observer looks at a displayed image.

In our HVS, the diameter of one cone is about 2 μm (30" of visual angle), and the distribution interval between the cones in the central fovea is 35". When we assume that the above-mentioned 30" of the visual angle corresponds to one pixel on a displayed image in real world, we can realize the simulation of a more realistic retinal image. In this experiment, we used a laptop display (VAIO Corp., VAIO Z VJZ13A1). The laptop has a display size of 20.6 cm \times 16.5 cm. Therefore, to realize one pixel on a digital image as one cone, we set our observation condition (viewing distance: 50 cm, viewing angle: $23.28^{\circ} \times 18.74^{\circ}$ and resolution: 2288×1830 pixels) because one pixel corresponds to 36.63" (it is similar to 30" of the viewing angle by one cone).

The gaze points were measured by an eye tracker (NAC Image Technology Inc., EMR-9) using the positions of both eyes. Figure 3(a) shows a snapshot of the experimental environment. The target image displayed on a laptop monitor is used as an experimental stimulus. The participant observes the experimental stimulus by changing only the visual

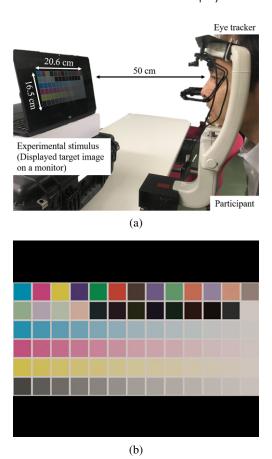


Figure 3. Experimental environment. (a) System setup. (b) A stimulus.

point under a fixed chin support. Fig. 3(b) shows that the experimental stimulus was a simple color chart. We also prepared test stimuli even for images that were more complex than the color chart and performed simulations using the proposed models. The test images were natural scenes under low illumination conditions (710×613 pixels), and a cube (1000 × 800 pixels) was from the YACCD (yet another color constancy database) dataset [25]. These images are approximately contained in the 10° field of view. A participant freely observed three target images for 30 s. The results of this step yield a very low visual impact because each fixation lasts on an average 200-300 ms, and in 30 s of gazing at the image, only \sim 120 pixels will be affected. Therefore, 1800 detected gaze points (60 Hz \times 30 s) were reduced to 200 points after excluding the noise data by using the area threshold (outside the target image area) and the frequently detected data with the same coordinate. All the experiments were conducted in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki). Written informed consent was obtained from all the participants.

3.2 Results for Single Gaze Point

In this section, we consider the processing results for a single gaze point. In the experiment, the diameters of the circles with viewing angles of 2° and 10° corresponded to 1.75 (194 pixels) and 8.75 cm (972 pixels) on the display.

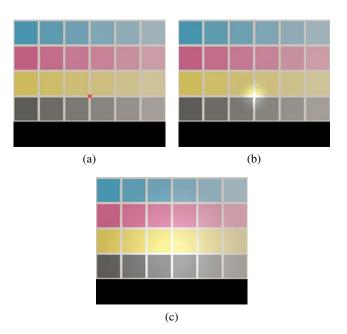


Figure 4. Results for a fixation point. (a) Original image. (b) Model based on the cone distribution (Section 2.2) (number of pixels p=16,325). (c) Model based on the visual resolution (Section 2.3) (number of pixels p=741,977).

Figure 4 shows the simulation results of close-up images for the color chart image. In Fig. 4(a), the red "x" represents a fixation point. Figs. 4(b) and 4(c) show the simulation results for the fixation point by the proposed models in Sections 2.2 and 2.3, respectively. Although the number of processed pixels, p, in ROI varies significantly between the two models, in both the simulation results, it is observed that the brightness is improved in the area around the fixation point, and the color of the original image is maintained as it approaches the border of ROI.

3.3 Results for Eye-Tracking Data

The proposed model was applied to the time series gaze tracking data described in Section 3.1. Figure 5 shows the simulated images of the color chart image obtained by each model. Fig. 5(a) shows an original test image and the sampled gaze points indicated as red "x"s. Figs. 5(b)–(d) show the simulated images using the traditional RSR, proposed model based on the cone distribution and another proposed model based on visual resolution, respectively. Compared to Fig. 5(b), the proposed model simulates the HVS in ROI. As shown in Fig. 5(a), the participant observes each color patch and grid line to compare the difference in the lightness and color. A participant observes the same gaze points several times. It is reported that the appearance is constant and not repetitively amplified by viewing multiple times in an introspection survey. Therefore, our hypothesis of the effect of the RSR being constant regardless of the number of times seems to be reasonable to simulate the appearance of human vision.

Other results are presented in Figures 6 and 7. As shown in Fig. 6, the appearance of the processed area around

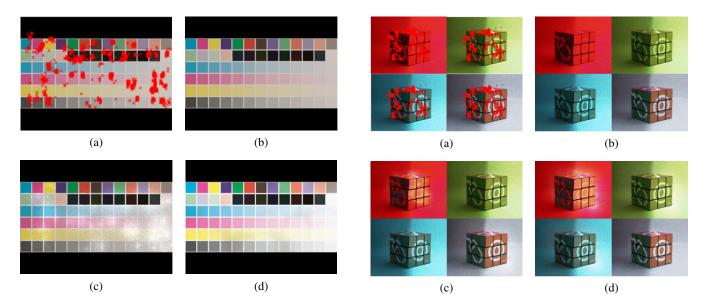


Figure 5. Results for the color chart image. (a) Original image. (b) Traditional RSR (Section 2.1). (c) Model based on the cone distribution (Section 2.2) (total number of pixels p = 101, 575). (d) Model based on the visual resolution (Section 2.3) (total number of pixels p = 364, 820).

Figure 7. Results for the cube image. (a) Original image ("x": 200 measured gaze points). (b) Traditional RSR (Section 2.1). (c) Model based on the cone distribution (Section 2.2) (total number of pixels p = 86, 735). (d) Model based on the visual resolution (Section 2.3) (total number of pixels p = 324, 494).

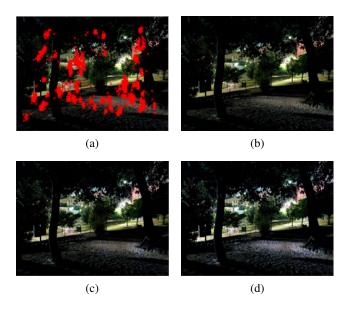


Figure 6. Results for the natural scene image. (a) Original image ("x": 200 measured gaze points). (b) Traditional RSR (Section 2.1). (c) Model based on the cone distribution (Section 2.2) (total number of pixels p = 94,704). (d) Model based on the visual resolution (Section 2.3) (total number of pixels p = 366,276).

the real gaze points results natural. For the cube images shown in Fig. 7, the same fixation points measured for the red-light condition are used as the eye-tracking data for the other cube images. Figs. 7(c) and (d) confirm the chromatic normalization effect typical of SCAs (spatial color algorithms) [12].

4. CONSIDERATION

We conducted an additional experiment to detect the characteristics of the eye movement during gazing at the target point on an object. Our eyes are typically not perfectly still during gazing but are perturbed by small undetectable ocular motions called involuntary eye movements. Their role is to counteract the effects of a neural adaptation during unvarying stimuli, preventing the effect of perceptual fading. This processing method considers the involuntary eye movements in the computation of the RSR. Three types of eye motions can be distinguished during a gazing: tremor, drift and microsaccades. A tremor is defined as an aperiodic, wave-like motion of the eyes with a frequency from 30 to 100 Hz and an angular extent of 10-20 s of the arc, about the diameter of one cone in the fovea [26]. A visual tremor is also reported to be generally independent in both eyes, generating a physical limit on the ability of matching the corresponding visual points at the retina during a stereoscopic vision. Drifts are movements that occur in conjunction with a tremor, resulting in a slow motion of the eyes, causing the image to move across an extent of 5 min of the arc. Micro saccades are extremely small and unnoticeable jerking movements of an eye occurring during a voluntary fixation, usually moving the retinal image across 2-25 min of the arc and occurring in about 25 ms [27]. These movements have been theorized to be functional during fixations, similar to the saccadic movements during the exploration of a scene.

In our models, we desired to incorporate involuntary eye movements in the RSR computation. As shown in Figure 8, we record the data from microsaccades and drifts using an eye-tracker set to a detection rate of 60 Hz while fixating the target gaze point marked as a red symbol for 30 s.

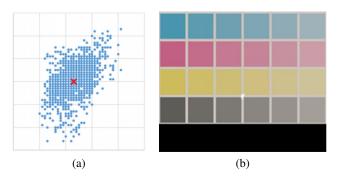


Figure 8. Consideration of the involuntary eye movements. (a) Measurement result. "x" is the target point. Blue dots indicate the detected points as involuntary eye movements. (b) Visualization result.

We considered the results for the dominant right eye of one of the authors with a viewing distance of 50 cm while observing a target in the color chart (2288×1830 pixels), with one pixel corresponding to 36.63" of the viewing angle. Fig. 8(a) shows the detected points, with a range of about 30×50 pixels (18.3 × 30.5'), distributed over an approximate oval shape. The visualization result using the traditional RSR is shown in Fig. 8(b). Because of the drifted points, the eye movements might have an anisotropy depending on the viewing condition, such as dominant eye, target point in the field of view and focus distance. Therefore, isotropic modeling like Brownian motion does not seem to be appropriate for the simulation. Moreover, an involuntary eye movement is affected by the color and luminance of the target object [28]. To incorporate involuntary eye movements into our model, considerations of an anisotropic modeling based on the eye characteristics and the color and luminance of the target object in the Retinex model are needed by our experiments.

5. CONCLUSIONS

This article proposed two models extending RSR along ROI. In the first proposed model, a cone distribution based on anatomical knowledge was considered as ROI. Based on the knowledge that the cone photoreceptor array plays a critical role in spatial sampling in the HVS in the region from approximately 1° to 10°, the visual field of 2° from the fixed point of view was set as ROI. Moreover, random numbers were generated in ROI to achieve a two-dimensional normal distribution as the processing pixels. In the second model, the visual resolution depending on the visual field based on the knowledge of visual information processing was considered as ROI. Based on the knowledge of the visual information processing, we developed a model in which the first 2° of the visual angle had the best resolution performance, which then gradually declined to 10° of the visual angle by using the weight for the RSR.

As a support for the proposed models, we measured actual eye movements using an eye-tracking system. By using the eye-tracking data, we acquired the HVS movements using test images. Results confirmed the interest of such an extension of the RSR model. A wide and more accurate

comparison of the results of the proposed extension will be the subject of further research.

REFERENCES

- ¹ E. H. Land and J. J. McCann, "Lightness and the Retinex theory," J. Opt. Soc. Am. **61**, 1–11 (1971).
- ² E. H. Land, "The Retinex theory of color vision," Sci. Am. **237**, 108-128 (1977).
- ³ E. Provenzi, L. D. Carli, A. Rizzi, and D. Marini, "Mathematical definition and analysis of the Retinex algorithm," J. Opt. Soc. Am. A 22, 2613–2621 (2005).
- ⁴ E. Provenzi, M. Fierro, A. Rizzi, L. D. Carli, D. Gadia, and D. Marini, "Random spray Retinex: a new Retinex implementation to investigate the local properties of the model," IEEE Trans. Image Process. **16**, 162–171 (2007)
- ⁵ J. J. McCann, "Retinex at 50: color theory and spatial algorithms, a review," J. Electron. Imaging 26, 031204 (2017).
- ⁶ A. Rizzi and C. Bonanomi, "Milano Retinex family," J. Electron. Imaging **26**, 1–7 (2017).
- ⁷ M. Lecca, G. Simone, C. Bonanomi, and A. Rizzi, "Point-based spatial color sampling in Milano-Retinex: a survey," IET Image Process. 12, 833–849 (2018).
- ⁸ E. H. Land, "Recent advances in Retinex theory and some implications for cortical computations: Color vision and the natural image," Proc. Natl Acad. Sci. USA 80, 5163–5169 (1983).
- ⁹ E. H. Land, "An alternative technique for the computation of the designator in the Retinex theory of color vision," Proc. Natl Acad. Sci. USA 83, 3078-3081 (1986).
- ¹⁰ J. J. McCann and A. Rizzi, *The Art and Science of HDR Imaging* (John Wiley, Hoboken, NJ, 2011).
- ¹¹ E. Provenzi, Computational Color Science: Variational Retinex-like Methods (John Wiley, Hoboken, NJ, 2017).
- ¹² A. Rizzi and J. J. McCann, "On the behavior of spatial models of color," Proc. SPIE **6493** (2007).
- ¹³ G. Gianini and A. Rizzi, "A fuzzy set approach to Retinex spray sampling," Multimedia Tools Appl. 76, 24723–24748 (2017).
- ¹⁴ M. Lecca, A. Rizzi, and R. P. Serapioni, "GRASS: A gradient-based random sampling scheme for Retinex," IEEE Trans. Image Process. 26, 2767–2780 (2017).
- ¹⁵ M. Lecca, A. Rizzi, and R. P. Serapioni, "GREAT: a gradient-based color sampling scheme for Retinex," J. Opt. Soc. Am. A 34, 513–522 (2017).
- 16 G. Gianini, M. Lecca, and A. Rizzi, "A population-based approach to point-sampling spatial color algorithms," J. Opt. Soc. Am. A 33, 2396–2413 (2016).
- ¹⁷ M. Lecca, A. Rizzi, and G. Gianini, "Energy-driven path search for termite Retinex," J. Opt. Soc. Am. A 33, 31–39 (2016).
- ¹⁸ G. Gianini, A. Rizzi, and E. Damiani, "A Retinex model based on absorbing Markov chains," Inf. Sci. 327, 149–174 (2016).
- ¹⁹ G. Gianini, A. Manenti, and A. Rizzi, "QBRIX: a quantile-based approach to Retinex," J. Opt. Soc. Am. A 31, 2663–2673 (2014).
- ²⁰ R. Ohtera, T. Horiuchi, and S. Tominaga, "Implementation of Retinex algorithm by eyegaze tracking interface," Proc. SPIE 7241, 724107 (2009).
- ²¹ D. R. Williams and R. Collier, "Consequences of spatial sampling by a human photoreceptor mosaic," Science 221, 385–387 (1983).
- D. R. Williams, "Seeing through the photoreceptor mosaic," Trends Neurosci. 9, 193–198 (1986).
- ²³ D. R. Williams and N. J. Coletta, "Cone spacing and the visual resolution limit," J. Opt. Soc. Am. A 4, 1514–1523 (1987).
- ²⁴ J. Hirsch and C. A. Curcio, "The spatial resolution capacity of human foveal retina," Vis. Res. 29, 1095–1101 (1989).
- ²⁵ A. Rizzi, C. Gatta, and D. Marini, "YACCD: yet another color constancy database," Proc. SPIE **5008**, 24–36 (2003).
- ²⁶ L. A. Riggs, F. Ratliff, J. C. Cornsweet, and T. N. Cornsweet, "The disappearance of steadily fixated visual test objects," J. Opt. Soc. Am. 43, 495–501 (1953).
- ²⁷ F. Ratliff and L. A. Riggs, "Involuntary motions of the eye during monocular fixation," J. Exp. Psychol. 40, 687–701 (1950).
- ²⁸ R. M. Steinman and R. J. Cunitz, "Fixation of targets near the absolute foveal threshold," Vis. Res. 8, 277–286 (1968).