Combining Gray World and Retinex Theory for Automatic White Balance in Digital Photography

Edmund Y. Lam, *Member, IEEE*

Abstract — *White balancing is an important step in the digital camera processing pipeline to adjust the color of the pixels under different illuminations. Efficient automatic white balance is usually a required component of a consumer digital camera because many users would not prefer to handle this task manually. Gray world assumption and Retinex theory are two common methods used, but their aims vary and their applicability depends on the nature of the images. In this paper, we present an effective technique that combines the two together, while preserving the strength of the two methods. Experimental results confirm that our approach is a viable alternative to the two existing methods[1].*

Index Terms — **Digital photography, automatic white balance, camera design, efficient implementation.**

I. INTRODUCTION

White balancing is an important step in the digital camera processing pipeline to adjust the color of the pixels under different illuminations [1]. It can be performed either manually or automatically. For the former, usually the user would preset a certain illumination condition, and the camera would compute the color correction based on the preset values. Most amateur users however prefer automatic white balancing, for which there are mainly three approaches [2]. The first method incorporates the gray world assumption, which seeks to equalize the mean of the red (R), green (G), and blue (B) channels. This is quite effective in practice, except in situations where a certain color may dominate, such as a blue hue for the sky. In these situations, the second method is more effective. It is based on the Retinex theory of visual color constancy, which argues that perceived white is associated with the maximum cone signals [3]. The third method estimates the most likely illuminant from a predefined set using correlation. Variations of this method are also proposed recently with more sophisticated ways of finding the correlation [4], [5].

The correlation-based method and its variants are computationally intensive and may not be suitable for real-time applications. Both the gray world and Retinex theory lead to methods that are very efficient and produce a linear correction to the R and B channels. However, with a linear technique no method can guarantee to satisfy both assumptions. In this paper, we propose a quadratic correction to the two channels which can simultaneously satisfy the gray world and Retinex theory of white balancing. The paper is organized as follows: In Section II, we review the theory for color capture and explore the rationale behind the white balancing approaches. Then, in Section III, we describe our proposed method, including its derivation and implementation. In Section IV, simulation results are presented to verify the efficacy of our technique. We conclude with some practical suggestions on automatic white balance schemes in Section V.

II. COLOR CAPTURE IN DIGITAL PHOTOGRAPHY

Color is encoded as the intensity values of the red, green, and blue channels in a digital picture. Each of these values is affected by three physical quantities: the illumination, reflectance, and spectral sensitivity of the photodetectors.
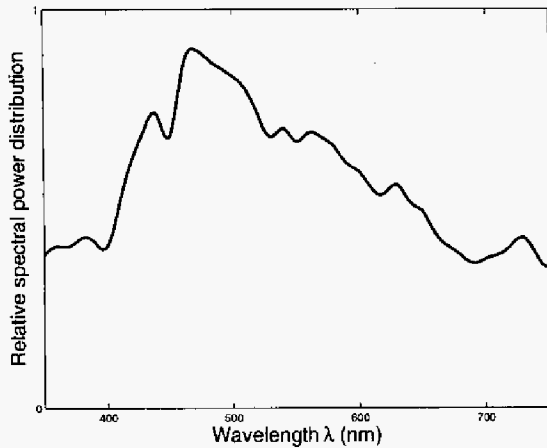
A. Illumination

Imaging begins with the source of light, also known as illumination. Virtually all illuminations consist of light with multiple wavelengths. Each one is then described with a curve showing the strength of the electromagnetic radiations at different values of wavelength λ. If we normalize the curves of various illuminations, the result is a very useful description of the spectral power distribution of the illumination as a function of wavelength. We can then compare the representative spectral power distributions of various common light sources. Note that we usually only describe the general characteristics of the illuminations; an actual measurement for sunlight, for example, would depend on the location, altitude, and atmospheric and weather conditions during the measurement [6].

Figure 1 shows the spectral power distribution of various common illumination sources. (a) is the curve for typical sunlight, which is continuous (although not uniform) over the visible spectrum. Tungsten light, as shown in (b), also appears to be rather smooth. In contrast, the fluorescent lamp consists of sharp spikes in the spectral power distribution, as shown in (c). We denote the spectral power distribution of an illumination as $I(\lambda)$. This will be important as we relate it with other color production factors.
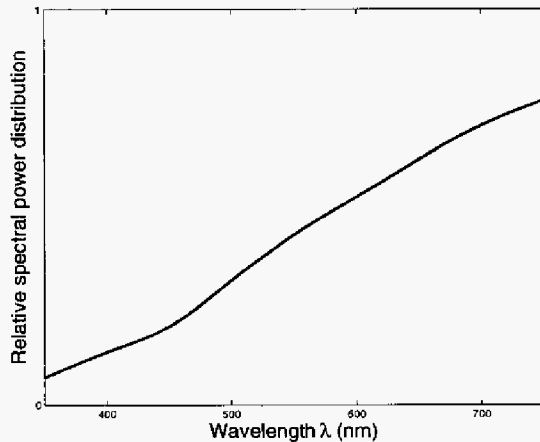
B. Object Reflectance

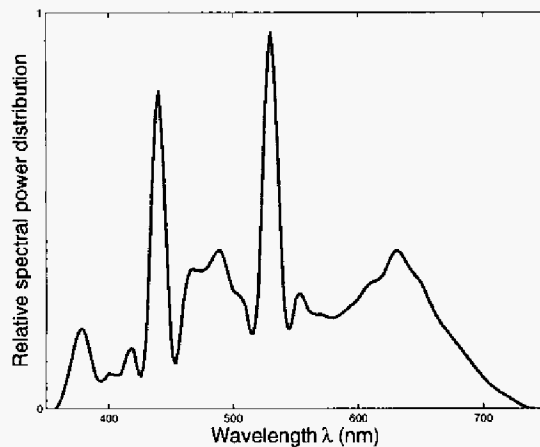When the electromagnetic radiations from the illuminat-

[1] Edmund Y. Lam is with the Department of Electrical and Electronic Engineering, University of Hong Kong, Pokfulam Road, Hong Kong SAR. (e-mail: elam@eee.hku.hk).

(a) sunlight

(b) tungsten

(c) fluorescent

Fig. 1. Typical spectral power distribution of various illuminations.

ion reach an object, they are partially absorbed, and partially reflected or transmitted (for transparent objects). For

different items, the proportion of the reflection or transmission varies with wavelengths, but this is an inherent property of the object irrespective of the illumination that has taken place. We can therefore characterize an object's spectral reflectance or spectral transmittance as a function of wavelength for comparison.

The spectral reflectance corresponding to several typical object colors is shown in Figure 2. For each plot, the y-axis denotes the fraction of light that is being reflected from the object. Figure 2(a), (b), and (c) show the spectral reflectance of a red, yellow, and gray patch of color respectively. These patches are taken from the Macbeth chart, which is often used to test digital camera performances [7]. As expected, a red patch absorbs most of the greenish-blue frequencies and reflects most of the higher-wavelength light that gives the sensation of red. However, it should be noted that some residual lower-wavelength frequencies are also reflected, only that the amount is much smaller. We observe a similar behavior for the yellow patch as well. For the gray patch, the spectral reflectance is roughly constant for the different wavelengths, causing the resulting gray sensation to be neutral in color. We denote the spectral reflectance of an object with $R(\lambda)$. Similarly, we can also define the spectral transmittance of an object with $T(\lambda)$, although we will not make much use of it in this paper.

Illumination and object reflectance (or transmittance) together determine what is known as color stimulus. The spectral power distribution of the illumination governs how much energy is incident on the object at every wavelength. For a reflective object, the spectral reflectance dictates what fraction of that radiation is reflected and will arrive at the eye or the sensor, again at every wavelength. Similarly, for a transmissive object, the spectral transmittance determines the fraction of the radiation being transmitted through the object. Therefore, the spectral power distribution of an object is the product of the spectral power of the illumination and the spectral reflectance of the object. This is also known as the color stimulus.
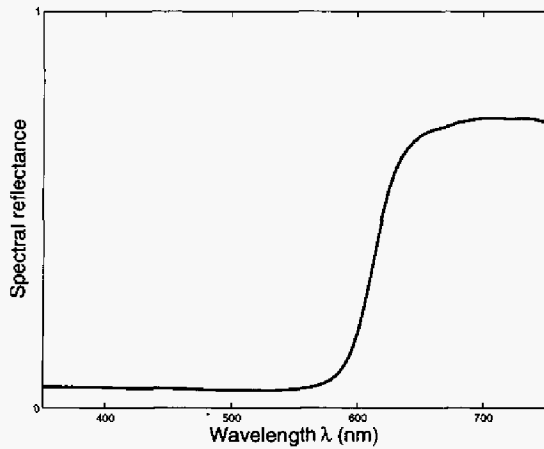
Mathematically, we denote the color stimulus with $S(\lambda)$. It relates to the illumination $I(\lambda)$ and object reflectance $R(\lambda)$ by

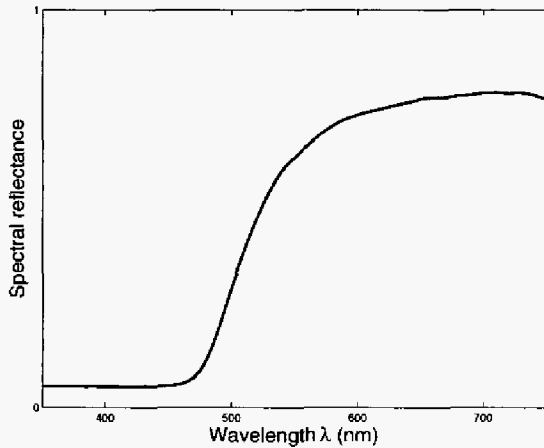$$S(\lambda) = I(\lambda)R(\lambda). \tag{1}$$

When we take a picture of the same object first under sunlight and then under fluorescent light, for example, the color stimuli vary significantly due to differences in illumination. We must "adapt" our camera to interpret the color stimuli differently, or otherwise the color of the photographs would look very different. This is the basis for the need for automatic white balance in digital photography.
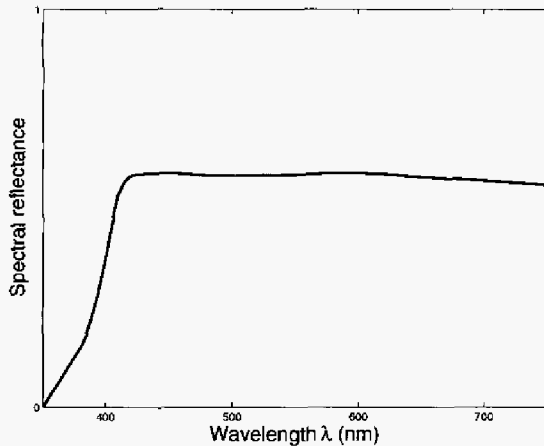
C. Spectral Sensitivity of Photodetectors

The sensor in the camera consists of three filters, typically red, green, and blue filters. The spectral sensitivities of a typical camera are shown in Figure 3. We use $x(\lambda), y(\lambda)$,
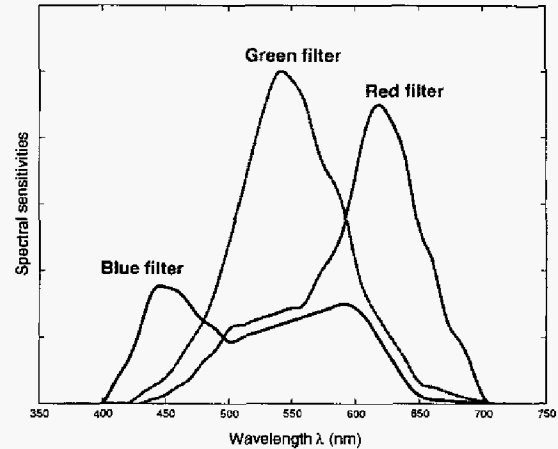
135

(a) red patch

Fig. 3. Spectral sensitivities of a typical digital camera.

ized to equal area. It is interesting to observe that they do not cover disjoint sections of the visible spectrum, nor do they cover it entirely.

When an object with stimulus $S(\lambda) = I(\lambda)R(\lambda)$ is observed, each of the three filters responds to the stimulus by summing up the reaction at all wavelengths. Therefore, three values are produced from the three filters, in accordance with the equations

$$X = \int_{400}^{700} x(\lambda)I(\lambda)R(\lambda)d\lambda$$

$$Y = \int_{400}^{700} y(\lambda)I(\lambda)R(\lambda)d\lambda \ . \tag{2}$$

$$Z = \int_{400}^{700} z(\lambda)I(\lambda)R(\lambda)d\lambda$$

The triplet (X,Y,Z) is called trichromatic responses. An important consequence of trichromatic response is that in the digital camera, we only require three numbers at each pixel to capture the color information. We do not need to record the color stimulus at all wavelengths. In fact, this gives rise to a useful phenomenon: Even if we consider the spectral sensitivities to be known, for any given triplet (X,Y,Z) there could be an infinite number of possibilities for the color stimulus according to (2).

(b) yellow patch

(c) gray patch

Fig. 2. Spectral reflectance of typical objects.

and $z(\lambda)$, to denote the responses of the three filters respectively. For comparison, the curves have been normal-

III. AUTOMATIC WHITE BALANCE

In the previous Section we explained the need and challenge of automatic white balance. Here we outline the common algorithms that are used today, and explain our technique that seeks to combine the contribution of two dominant techniques.

The first algorithm is based on the gray world assumption, which argues that for a typical scene, the average intensity of the red, green, and blue channels should

be equal. This method has its root in film photography, where for the negatives the average is biased towards dark regions of the scene which tends to be neutral [2]. In terms of implementation, let an image $I(x,y)$ have size $M \times N$, where x and y denote the indices of the pixel position. Furthermore, let $I_r(x,y)$, $I_g(x,y)$, and $I_b(x,y)$ denote the red, green, and blue channels of the image respectively. We compute

$$R_{avg} = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} I_r(x,y)$$

$$G_{avg} = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} I_g(x,y). \tag{3}$$

$$B_{avg} = \frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} I_b(x,y)$$

If the three values are identical, the image already satisfies the gray world assumption. In general, they may not be. Customarily, we keep the green channel unchanged, and define the gain for the red and blue channels as

$$\hat{\alpha} = \frac{G_{avg}}{R_{avg}} \quad \text{and} \quad \hat{\beta} = \frac{G_{avg}}{B_{avg}}. \tag{4}$$

We then adjust the red and blue pixels by

$$\hat{I}_r(x,y) = \hat{\alpha} I_r(x,y)$$
$$\hat{I}_b(x,y) = \hat{\beta} I_b(x,y). \tag{5}$$

Alternatively, under the Retinex theory, it is argued that the perceived white is associated with the maximum cone signals of the human visual system [3]. As such, the method for white balance should be to equalize the maximum values of the red, green, and blue channels. To avoid disturbances to the calculation caused by a few bright pixels, one can treat clusters of pixels or lowpass the image [2]. To implement this, we compute

$$R_{max} = \max_{x,y}\{I_r(x,y)\}$$
$$G_{max} = \max_{x,y}\{I_g(x,y)\} \tag{6}$$
$$B_{max} = \max_{x,y}\{I_b(x,y)\}$$

As with the gray world assumption, we keep the green channel unchanged. We define the gain for the red and blue channels as

$$\tilde{\alpha} = \frac{G_{avg}}{R_{avg}} \quad \text{and} \quad \tilde{\beta} = \frac{G_{avg}}{B_{avg}}. \tag{7}$$

We then adjust the red and blue pixels by

$$\tilde{I}_r(x,y) = \tilde{\alpha} I_r(x,y)$$
$$\tilde{I}_b(x,y) = \tilde{\beta} I_b(x,y). \tag{8}$$

The two methods have their respective strengths. Upon a closer investigation, we can make the following observations:

- For most images, the two methods produce different results. In other words, the corrected image can rarely satisfy both the gray world assumption and the Retinex theory.
- Equations (5) and (8) are both linear adjustments to the pixel intensities. Furthermore, there is also a fixed point in the mappings: for pixels with zero intensity, the two mappings would not affect their values.

To incorporate the rationales of both methods, we propose a simple adjustment here with a quadratic mapping of intensities. As with the methods above, we keep the green channel unchanged. Let the change to the red channel be represented as

$$\bar{I}_r(x,y) = \mu I_r^2(x,y) + v I_r(x,y), \tag{9}$$

where (μ, v) are the parameters for automatic white balance. To satisfy the gray world assumption, we require that

$$\sum_{x=1}^{M} \sum_{y=1}^{N} \bar{I}_r(x,y) = \sum_{x=1}^{M} \sum_{y=1}^{N} I_g(x,y) \tag{10}$$

i.e.

$$\mu \sum_{x=1}^{M} \sum_{y=1}^{N} I_r^2(x,y) + v \sum_{x=1}^{M} \sum_{y=1}^{N} I_r(x,y) = \sum_{x=1}^{M} \sum_{y=1}^{N} I_g(x,y) \tag{11}$$

Similarly, to satisfy the Retinex theory, we need

$$\mu \max_{x,y}\{I_r^2(x,y)\} + v \max_{x,y}\{I_r(x,y)\} = \max_{x,y}\{I_g(x,y)\} \tag{12}$$

Equations (11) and (12) together form two equations in two unknowns. We can represent them in a matrix form

$$\begin{bmatrix} \sum\sum I_r^2 & \sum\sum I_r \\ \max I_r^2 & \max I_r \end{bmatrix} \begin{bmatrix} \mu \\ v \end{bmatrix} = \begin{bmatrix} \sum\sum I_g \\ \max I_g \end{bmatrix} \tag{13}$$

where we have omitted the subscripts for simplicity. This can be solved analytically, either with Gaussian elimination or using Cramer's rule [8]. The white balance for the blue channels can be computed in an analogous manner.

IV. SIMULATION

In this Section we present a simple simulation that demonstrates the efficacy of our approach. We apply our method on the image shown in Figure 4(a). We can see that the color tends to be too reddish. If we apply the gray world assumption in automatic white balance, the result is shown in (b). If we apply the Retinex theory, the result is shown in (c). Either case, the result is not too satisfactory. The former is too grayish while the latter has a purple hue. On the other hand, if we apply our white balancing technique with quadratic correction, we obtain the result in (d). The color balance is significantly improved.

To provide more insight into the algorithm in addition to

137

(a) raw image

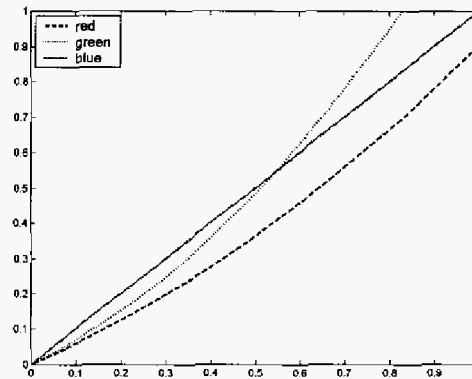(b) using gray world assumption

(c) using Retinex theory assumption

(d) using our quadratic correction

Fig. 4. An example of automatic white balance using the proposed scheme.

the visual appearance of test images, we plot the input-output mapping of the color channels in Figure 5. As expected, the mappings are quadratic in nature. Note that we

(c) input-output intensity mapping

Fig. 5. An example of the input-output intensity mapping with the proposed scheme.

did not make changes to the green channel in our algorithm. Moreover, Table I and II show the numerical values of the gray point and maximum cone signal using the three fast methods mentioned: the gray world, Retinex, and our proposed technique. Again, we confirm that our method can simultaneously satisfy the two goals stated above.

TABLE I
GRAY POINT OF IMAGES

Method	Red	Green	Blue
Gray world	0.37	0.37	0.37
Retinex theory	0.41	0.37	0.42
Proposed scheme	0.37	0.37	0.37

TABLE II
MAXIMAL CONE SIGNAL

Method	Red	Green	Blue
Gray world	0.65	0.71	0.62
Retinex theory	0.71	0.71	0.71
Proposed scheme	0.71	0.71	0.71

V. CONCLUSION

In this paper we have described the formation of color signal in digital photography, and explained the need for automatic white balance. We then reviewed two widely used assumptions for this task, namely gray world and Retinex theory, and then proposed a method that can simultaneously satisfy these two requirements. We also produced simulation results that confirm the efficacy of our proposed method.

REFERENCES

[1] E. Lam, "Image restoration in digital photography," *IEEE Transactions on Consumer Electronics*, vol. 49, no. 2, pp. 269-274, May 2003.

[2] R. Hunt, *The Reproduction of Colour.* 6th edition, Wiley, 2004.

[3] E. Land and J. McCann, "Lightness and Retinex theory," *Journal of the Optical Society of America,* vol. 61, no. 1, pp. 1-11, 1971.

[4] N. Kehtarnavaz, H. Oh, and Y. Yoo, "Development and real-time implementation of auto white balancing scoring algorithm," *Real-Time Imaging,* vol. 8, pp. 379-386, 2002.

[5] K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms – part I: Methodology and experiments with synthesized data," *IEEE Transactions on Image Processing,* vol. 11, no. 9, September 2002.

[6] E. Giorgianni and T. Madden, *Digital Color Management,* Addison Wesley, 1998.

[7] N. Sampat, S. Venkataraman, and R. Kremens, "System implications of implementing white balance on consumer digital cameras," *Proceedings of the SPIE,* vol. 3965, pp. 362-368, January 2000.

[8] G. Strang, *Linear Algebra and Its Applications.* HBJ, 1988.

Edmund Y. Lam (S'97-M'00) received the B.S. degree in 1995, the M.S. degree in 1996, and the Ph.D. degree in 2000, all in electrical engineering from Stanford University, U.S.A. At Stanford, he was a member of the Information Systems Laboratory, conducting research for the Stanford Programmable Digital Camera project. His focus was on developing image restoration algorithms for digital photography. Outside Stanford, he also consulted for industry in the areas of digital camera systems design and algorithms development. He is now an Assistant Professor of Electrical and Electronic Engineering at the University of Hong Kong.