Cluster Based Color Constancy

Weihua Xiong, Jiangtao Kuang, Xiaoyong Wang
Omnivision Technology Inc., Santa Clara, CA, USA

Abstract

Xiong et al. developed an improved color constancy solution, GSI, by identifying all of those potential gray surfaces and average them in the RGB space.[1] This strategy assumes that the gray surface, no matter what the lighting color is, have an S axis of nearly zero in the LIS coordinate system that roughly correspond to the variation of intensity, illumination and reflectance.[1,2] However, this assumption is often violated, where some non-gray surfaces under specific illumination would be mistaken as gray ones. Simply averaging all detected pixels would bias the illumination estimation. To overcome this problem, the GSI is extended by analyzing color distribution of all identified gray surfaces in perspective chromaticity space. We employ an unsupervised cluster technique to extract the group with maximal data distribution density among them. The center of the selected cluster is then used to approximate the illumination colors. The advantage of the cluster technique is that we don't need a large training data set to establish the relationship between the statistical properties of image data and the lighting color incident on it. The experiments based on two real image data set show that this method is comparative to other elaborate existing color constancy methods and has lower costs than most existing color constancy methods.

Key words

Cluster, Color Constancy, Auto White Balancing

Introduction

The color information from any imaging device depends on three factors: the spectral power distribution of the illumination incident on the objects, the underlying physical property of the objects, and the sensor sensitivity of the imaging system itself. Therefore, the same object will appear different in the image if the illumination color changes. Color constancy is a mechanism to remove the effect of light source and recover the original physical scene more accurately. The crucial point to solve this problem is recovering the illumination colors.

Color constancy is an ill-posed problem, so it cannot be solved without further constraints. During the past decades, the researchers have proposed several mathematical models to determine the properties of the illumination. Depending on the assumptions and techniques, all of these algorithms are roughly divided into two major categories. One category is analyzing and building statistical relationship between image distribution and light color, for example Color by Correlation, Neural Network, Support Vector Regression and Thin Plate Interpolation, Although they work well, they always require large image data for training.[3-6] The other one is based on the nature of the color components of image itself, so they are simple, straightforward and easy to implement. The methods belonging to this category include Gray World, Max-RGB, Shade of Gray and Gray Surface Identification [1,7-9].

Xiong et al. previously published an improved illumination chromaticity estimation solution, named Gray Surface Identification Color Constancy (GSI), by identifying those potential gray surfaces through a new coordinate system in Logarithm Space and averaged all of them in original RGB color space. [1] However, since some non-gray surfaces under a specific light may conflict with those of gray surfaces under another illumination, they will be identified as gray incorrectly so as to cause the estimation to be biased. To address this problem, we take the benefits of GSI but extend it by further analyzing the data distribution using an unsupervised cluster technique. The novel method proposed here is based on the observation that the color information from those truly gray surfaces is clustered compactly with maximal distribution density in a chromaticity. Therefore we can analyze these selected pixels' frequency of occurrence, group them into different clusters and pick up the cluster with maximal density inside a pre-defined region. The center of the cluster is used to approximate the illumination values.

Overview of Potential Gray Surface Identification

Given some assumptions and constrains that the illumination can be approximated as a blackbody radiator and the camera sensor sensitivity function is narrow enough, Xiong et al. prove that varying the illumination's color temperature or its intensity of a surface will cause (logR, logG, logB) to move within a plane and the planes from different surface reflectance are parallel to each other. [1,2] They further developed a new coordinate system, named LIS, that represent illumination change ('L'), reflectance ('S') and intensity ('I') as separately as possible. [2] The system is the extension of an illumination-invariant color chromaticity space developed by Finlayson. [14]

To determine the three axes of LIS coordinate system, PCA is applied to the logarithm of RGBs from distinct surfaces under different illuminations with various intensities. The vector corresponding to the maximal eigenvalue represents the intensity axis, the second vector is the illumination axis, and the vector corresponding to the least eigenvalue is the surface reflectance axis. Figure 1 shows a result for a calibrated SONY camera [9]. The RGB values are from three surface reflectances under 102 illuminations at 15 different intensities.

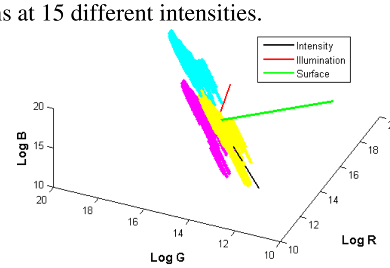

Figure 1 LIS coordinate system for a calibrated SONY camera. (Cited from figure 1 in [2])

Therefore, for each pixel from an image, the logarithm of color values, [$logR_i$, $logG_i$, $logB_i$], is projected onto the S axis of the LIS coordinate. If the resulting value is fallen into a specified threshold value, the pixel is considered as potential gray and will be retained for further processing; while those pixels falling outside the threshold will be withdrawn from future calculations.

Cluster Based Color Constancy Implementation

As noted in the previous section, not all potential gray pixels detected from the S axis in the LIS coordinate system are from a truly one. Therefore, instead of simply averaging them to estimate the illumination colors [8], we will project these identified pixels' R/G/B values into perspective chromaticity space[10]: (p_r, p_b) = (G/R, G/B).

We cluster these pixels into different groups in this space. For each group, the total pixels' number will be calculated inside a pre-fixed Density Region (DR). The cluster with the largest count will be considered. If the count is large enough, the center of the cluster is considered to be the estimated chromaticity values of incident illumination; otherwise, our algorithm will reduce to "Grey World" solution and the average of all pixels in the image will be considered as the light color.

Clustering is an unsupervised classification of data without knowing object labeling. It divides the data into different meaningful or useful groups for further analysis. The final goal is that the data in a group will be similar to each other while unrelated to data in other group. [11] There are a large variety of cluster techniques. In this paper, we use k-mean cluster technique since it can process the image pixel by pixel. So it is simple and easy to implement in real time system. This technique includes four steps: (1) Select k initial start points as cluster center; (2) Calculate each pixel's distance to the cluster center; (3) recalculate each cluster's center; (4) Repeat until converged to a stable status. Although the cluster number can be any number, three candidate illuminants equivalent to a blackbody radiator with a color temperature of 2856K, 4100K, and 6500K are mostly encounted, so we empirically set the k to be 3 in the following experiment, and the coordinates of three pre-selected illuminants serve as initial locations..

The image shown in figure 2 (a) consists of three Munsell papers: red, green and blue. It is lit under a reddish light and the ground truth of illumination color in perspective chromaticity is [0.6445, 0.7414]. The potential gray pixels identified from LIS are shown in white in Figure 2(b), it obviously indicates that some color pixels are detected as gray incorrectly. So if we simply average these potential gray pixels, some non-white pixels' color values will also be taken into account, and the illumination estimation will be [0.8709,0.7941]. However, if we group all of potential gray pixels into three clusters and analyze their distributions, we can find that the cluster with the highest density corresponds to those truly white surfaces, and its center is estimated to be the illumination value as [0.7227,0.7851].

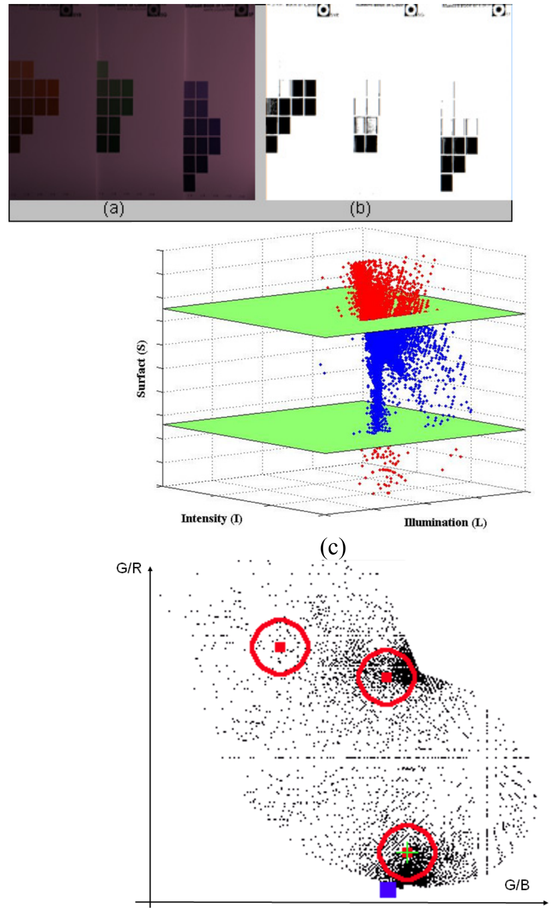

Figure 2 (a) Input Image (b) The Pixels identified as potential gray are indicated in white, otherwise, they are indicated in back (c) Image pixels in LIS Coordinate system. The pixels between thresholds (Represented by two Green Planes) identified as potential gray are indicated in Blue, other pixels are indicated in Red (d) Detected potential gray surfaces' color distribution in perspective chromaticity space. The image data of potential gray surfaces are represented by 'Black Dots'. Three clusters' centers are shown in 'Red Square' and true scene illumination is shown in 'Blue Square'. 'Red' line defines the boundary of 'DR'. The 'Green Cross' indicate the center of the cluster with the maximal density

Experiment

To evaluate the performance of cluster based color constancy, we will compare it to other existing methods in terms of the distance and angle errors in another commonly used chromaticity, [r,g]= [R/(R+G+B) G/(R+G+B)]. So we will change the output from perspective space into which one by
$r = (p_g)/(p_g+p_r*p_g+p_r)$, $g = p_r*p_g/(p_g+p_r*p_g+p_r)$ and $b=1-r-g$.

The distance error in two dimensions and angular error in three dimensions between estimated illumination chromaticity (r_{esti},g_{esti},b_{esti}) and ground truth (r_{real},g_{real},b_{real}) are calculated as:

$$Dist_i = \sqrt{(r_{real} - r_{esti})^2 + (g_{real} - g_{esti})^2}$$

$$Angular_i = COS^{-1}[\frac{(r_{real}, g_{real}, b_{real}) \circ (r_{esti}, g_{esti}, b_{esti})}{\sqrt{r_{real}^2 + g_{real}^2 + b_{real}^2} \times \sqrt{r_{esti}^2 + g_{esti}^2 + b_{esti}^2}}] \times \frac{2\pi}{360}$$

Then for any database with N images, statistical distance and angular errors, such as mean, median and RMS, are also given out. In particularly, the RMS is defined as:

$$RMS_{dist} = \frac{1}{N}\sqrt{\sum_{i=1}^{N} Dist_i^2}$$

$$RMS_{angular} = \frac{1}{N}\sqrt{\sum_{i=1}^{N} Angular_i^2}$$

To evaluate whether the difference between two methods is statistically trivial, we also apply the Wilcoxon signed-rank test based on the angular errors[15]. The error rate for accepting or rejecting null hypothesis is always set to 0.01.

In the first experiment, we use 321 images captured in the lab from SFU image database. [12] These images are captured by the calibrated SONY camera including 33 scenes under 11 different lights. The LIS system is acquired by the method proposed in [1]. Now there are two parameters to be set: the threshold in S axis for detecting potential gray surfaces (G_TH) and the size of density region (DR_S). We allow the G_TH parameter to be chosen from a range between 0.001 and 0.5 with the step size of 0.002, and the DR_S to be chosen between 0.03 and 0.25 with the step size of 0.015. The leave-one-out method on these 321 images is conducted. In each procedure, 320 images are used as training set, the pairs of these two parameters are tried to find the best choice that minimizes the mean angular error over this set, and then the selected parameters are applied on the remaining image. The whole procedure repeats 321 times so that each image is tested. The results and their comparison to other typical methods, GSI [1],Thin Plate Spline(TPS)[3], Support Vector Regression(SVR) [4], Shade of Gray(SoG) [7], MAX RGB [8] and Gray World(GW) [9], are listed in Table 1 and 2.

Our second experiment is based on another real world image database including 120 natural images of indoor and outdoor scenes. These images are captured by a 2MB CMOS sensor. For each image, the same scene with a Macbeth Color Checker is also taken, so the average R/G/B values of those saturated-exclusively gray patches in the Color Checker is used as the ground-truth of the illumination of the scene.

Since the camera is not calibrated, we will use the Color Checker again and capture its images under seven typical lights: A, D65,D75, indoor, CWF, fluorescent light and TL84. The color values of the gray patch in the Color Checker are applied to determine LIS coordinate system of the camera. The coordinates of three lights, A, CWF and D65, worked as initial locations of cluster. Same leave-one-out experiments are conducted again. The performance comparison between proposed method and others is given out in table 3 and table 4.

(a)

(b)

Figure 3 (a) Input Image (2) The image of the same scene including a Color Checker, the illuminant is determined from R/G/B values of saturated-exclusively gray patches in it

The results from Table 1 and Table 2 show that the proposed method is better than MAX, GW and GSI and comparable to SoG. Although its performance is not as accurate as TPS and SVR 3D, it doesn't require large training data set, and thus it is easier for hardware implementation.

Conclusion

A new illumination estimation method based on clustering technique is presented in this paper. It integrates the effectiveness of GSI [1] and those statistical-based color constancy solutions. The method analyzes those potential gray surfaces' color distribution so as to find the cluster with highest density in perspective chromaticity space. The center of the cluster is considered to be the illumination value. Although the proposed method solves color constancy problem from making assumptions requiring relationship between statistics of the RGB colors arising in the image and the color of the light, it does not require training data set to extract the relationship between the RGB color statistics and the color of light, therefore it is easy to implement with low costs. Our experiments show that the proposed algorithm is comparable to other existing solutions.

Reference

[1] W.H. Xiong, B.V. Funt, and Lilong , Shi, "Auto White Balanacing via Gray Surface Identification", Proceedings of *15th Color Imaging Conference*, pp. 143-146, (2007)

[2] W.H. Xiong, and B.V. Funt, "Color Space for Stereo Retinex", *Third International Conference on Color in Graphics, Imaging, and Vision,* Leeds, pp. 24-28, (2006)

[3] W.H. Xiong, Lilong Shi, and B.V. Funt, "Illumination Estiamtion via Thin-plate Spline Interpolation", Proceedings of *15th Color Imaging Conference*, pp. 25-29, (2007)

[4] *W.H. Xiong, and B.V. Funt,Estimating Illumination Chromaticity via Support Vector Regression*, Journal of Imaging Science and Technology Vol. 50, No.4 pp. 341-348, July/August, (2006)

[5] V. Cardei, B. Funt. and K. Barnard, "Estimating the Scene Illumination Chromaticity Using a Neural Network," *Journal of the Optical Society of America A,* Vol. 19, No. 12, pp. 2374-2386, (2002)

[6] G. D. Finlayson, S. D. Hordley and P. M. Hubel, "Color by Correlation: A Simple, Unifying Framework for Color Constancy," *IEEE Transactions on Pattern Analysis and Machine Intelligence* Vol. 23, No. 11, pp. 1209-1221, (2001)

[7] G. D. Finlayson and E. Trezzi, "Shades of Gray and Colour Constancy," *Proc. IS&T/SID Twelfth Color Imaging Conference: Color Science, Systems and Applications*, *Society for Imaging Science and Technology*, Scottsdale, pp. 37-41, (2004)

[8] G.D. Finlayson. "Retinex viewed as a gamut mapping theory of color constancy," *Proc. AIC International Color Association Color 97,* Vol. 2, pp. 527-530, (1997)

[9] B. Buchsbaum. "A spatial processor model for object color perception," *Journal of the Franklin Institute*, Vol. 31, pp. 1-26, (1980)

[10] G.D. Finlayson, "Color in Perspective", IEEE transactions on Pattern Analysis and Machine Intelligent, Vol. 18, Issue 10, pp1034-1038, 1996

[11] Jia wei han, Micheline Kamber, " Data Mining: Concepts and techniques", Chapter 8, pp335-485, Morgan Kaufmann, 2001

[12] K. Barnard, L. Martin, L., B.V. Funt, and A. Coath, "A Data Set for Color Research," *Color Research and Application*, Vol. 27, No. 3, pp. 140-147, (2002). (Data from: www.cs.sfu.ca/~colour)

[13] W.H. Xiong, "Separating Illumination From Reflectance in Colour Imagery" , PH.D. THESIS, SFU LIBRARY, 2007

[14] Finlayson, G.D. and Hordley, S.D., "Color Constancy at a pixel", *Journal of the Optical Society of America A*, Vol. 18, issue 2, pp. 253-64, 2001

[15] S.D. Hordley, G.D. Finlayson, "Reevaluation of color constancy algorithm performance*", Journal of the Optical Society of America A*, vol. 23, Issue 5, 2006, pp.1008-1020

Author Biography

Dr. Weihua Xiong is now an imaging scientist in Omnivision Technologies, USA. He got his Ph.D. from Simon Fraser University in 2007. He received his Master's degree from Beijing University (1996) and Bachelor's degree from Beijing Information University (1993). His primary research interest is color science, automatic white balancing, stereo vision and image processing. One of his papers, "Stero Retinex", was awarded as "Best Vision Paper" in 3rd Canadian Conference on Computational and Robot Vision in 2006. He is a member of Digital BioColor Society and Society of Imaging Science and Technology.

Mr. Xiaoyong Wang got his bachelor's degree and master's degree in Optical & Electronic Engineering of Zhejiang University in 2002 and 2006 respectively. Now he works in Omnivision Technologies as an imaging scientist. His research focuses on digital image stabilization and CMOS Sensor image processing.

Method	SVR Dimension/ Norm Power	Median Angle	RMS Angle	Max Angle	Median Dist ($\times 10^2$)	RMS Dist ($\times 10^2$)	Max Dist ($\times 10^2$)
Cluster		3.62	9.29	29.86	2.51	6.92	20.72
TPS *		0.64	2.10	14.43	0.53	1.55	10.42
GSI *		3.91	10.11	33.79	2.71	7.15	22.65
SVR *	2D	4.65	10.06	22.99	3.41	7.5	16.41
	3D	2.17	8.069	24.66	3.07	6.3	16.03
SoG *	6	3.97	9.027	28.70	2.83	6.21	19.77
Max RGB *		6.44	12.28	36.24	4.46	8.25	25.01
GW *		7.04	13.58	37.31	5.68	11.12	35.38
Color by Correlation+			10.9			8.1	
Neural Network+			9.5			7.0	

Table 1 Comparison of Cluster based color constancy performance to TPS, GSI, 2D/3D SVR, SoG, Max RGB, Grayworld , Color by Coorelation, Neural Network performance. The results involve real-data training and testing on the 321 SONY images captured in SFU lab. Errors are based on leave-one-out cross validation evaluation. The data of other methods marked by '*' are extracted from [13] (Table 4 page 76). The data of other methods marked by '+' are extracted from [4] (Table 3 page 345).

	Cluster	TPS	GSI	2D SVR	3D SVR	SoG (norm=6)	Max RGB	GW
Cluster		-	+	+	-	+	+	+
TPS *	+		+			+	+	+
GSI *	-	-		=	-	=	+	+
2D SVR*	-	-	=		-	=	+	+
3D SVR*	+	-	+	+		+	+	+
SoG* (norm =6)	-	-	=	=	-		+	+
Max RGB *	-	-	-	-	-	-		-
GW *	-	-	-	-	-	-	+	

Table 2 Comparison of the different algorithms via the Wilcoxon signed-rank test based on angular errors. A '+' means the algorithm listed in the corresponding row is better than the one in corresponding column; a '-' indicates the opposite; an '=' indicates that the performance difference between two algorithms is trivial. The data of other methods marked by '*' are extracted from [13] (Table 5 page 76).

Method	Norm Power	Median Angle	RMS Angle	Max Angle	Median Dist ($\times 10^2$)	RMS Dist ($\times 10^2$)	Max Dist ($\times 10^2$)
Cluster		1.15	3.78	16.30	1.02	2.96	12.44
GSI		1.99	4.19	16.59	1.50	3.22	12.59
SoG	6	2.34	5.17	24.83	1.60	3.97	23.42
Max RGB		7.87	9.41	22.62	5.74	7.26	17.24
GW		3.71	5.21	20.64	1.19	4.22	18.30

Table 3 Comparison of Cluster based color constancy performance to GSI, SoG, Max RGB, Grayworld performance. The results involve real-data training and testing on the 120 natural images. Errors are based on leave-one-out cross validation evaluation.

	Cluster	GSI	SoG(Norm=6)	Max RGB	GW
Cluster		+	+	+	+
GSI	-		+	+	+
SoG(norm=6)	-	-		+	+
Max RGB	-	-	-		-
GW	-	-	-	+	

Table 4 Comparison of the different algorithms via the Wilcoxon signed-rank test based on angular errors. Labeling '+', '-', '=' as for Table 2.