A Performance Evaluation of Color Constancy Methods based on Illumination Estimation

Chunxiao Li, Jiangming Kan*

Abstract—There are many color constancy methods based on illumination estimation and their error metrics, but it is unclear how they perform or which method is more appropriate in a particular scene. We chose Reproduction Angular Error (RAE) as the error measure to compare the performance of twenty-two illumination estimation methods in four different scenes (Open Country, City, Indoor, Lab), and present the best method choice to be used with each scene type.

Keywords—Color constancy, Illumination estimation，different scenes, performance evaluation.

I. INTRODUCTION

WHEN the spectral component of a light source changes, the color of an object's surface also changes. Under different lighting, the human visual system can identify the color of object surfaces, which is called color constancy [1]. Color constancy has been widely used in many fields. In the 20th century, Land and McCann proposed the Retinex system [2] [3] which performs well in image enhancement [4] [5]. In addition, color constancy can also be applied to video surveillance systems [6], color object recognition and detection [7], image indexing [8] [9], illumination estimation [10] [11] and other fields.

Color constancy methods are divided into two categories [12] [13]: methods based on color invariant descriptors and methods based on light estimation. The former extracts the color feature information which is independent of the illumination. The latter is a two-stage procedure; the illumination of a scene is estimated first, then the image color is adjusted according to the difference between estimated and canonical light [14], which is usually addressed by the Von Kries model [15]. For the color recovery of an image, if there is a transformation, illumination estimation is superior to color invariant descriptors [16]. The focus of this paper is to discuss different illumination estimation methods.

Illumination estimation methods include three groups: (1) unsupervised methods; (2) supervised methods; (3)

Jiangming Kan, with Beijing Forestry University, Beijing, 100083 China (corresponding author to provide phone: 861062337736; fax: 861062336137-706; e-mail: kanjm@bjfu.edu.cn).
Chunxiao Li, with Beijing Forestry University, Beijing, 100083 China (e-mail:jane950225@163.com).

combinational methods. The first group is based on the White Patch hypothesis [17] and Gray World hypothesis [18]. According to these hypothesis, Gijsenij et al. proposed an unsupervised method framework based on image edge [19]. The second group has two parts. One is statistic-based methods, and the other is based on machine learning. Many methods using statistics have been proposed, such as Gamut Mapping [19], derivative-constrained Gamut Mapping [20] and Color-by-Correlation [21]. And there are some methods that use learning phase including Bayesian method [22] [23], Color Cat [24], Smart Color Cat [25], and a convolutional neural network-based color constancy method [26]. Combination methods can be classified into two categories, which are direct combination and guided combination [14]. The direct combination method obtains an estimate by calculating the weighted combination of given estimates. The guided combination method obtains results from the attributes of image content, including the committee-based color constancy method [27], combination methods using natural image statistics, texture or other features [28]-[31].

There are four metrics to measure the performance of illumination estimation methods: angular error [32], chromaticity error [33], perceptual Euclidean distance (PED) [34], and reproduction angular error (RAE) [35]. Angular error calculates the angle between the RGBs of the measured illuminant and estimated illuminant colors; chromaticity error calculates the Euclidean distance between the estimated light and the true light in the chromaticity space; the weighted Euclidean distance between the estimated light and the true light is calculated by PED in the RGB space; the RAE is defined as the angle between the RGB of a white surface when the actual and estimated illuminations are 'divided out.'
There are many methods for illumination estimation, and they have different adaptabilities to different scenes [14]. In different scenes, the selection of the most optimal method is a critical dilemma. We chose RAE as the error measure to compare the performance of illumination estimation methods in different scenes, and we present the choice of methods for different scenes.

II. ILLUMINANT ESTIMATION METHODS

The illuminant estimation methods we discuss are the three types shown in section I : （1）unsupervised methods；（2）supervised methods；（3）combinational methods

A. Unsupervised methods

Unsupervised methods do not rely on prior knowledge, which can be estimated by low-level image feature statistics [36].

Among these of methods, two well-established methods, Gray-World and Max-RGB, were proposed early on. Gray-World is based on the Gray-World assumption that the average reflectance in a scene under a neutral light source is achromatic [18]; the Max-RGB method is based on the Max-RGB assumption that the maximum response in the RGB-channels is caused by a white patch [16].

To make Max-RGB and Gray-World more general, Finlayson et al. proposed the Shades of Gray method by introducing the Minkowski norm (p) into these two methods:

$$\left(\frac{\int f(x)^p \, dx}{\int dx} \right)^{\frac{1}{p}} = ke \qquad (1)$$

When $p = 1$, equation (1) is equivalent to Gray-World. In contrast, equation (1) is equivalent to the max-RGB when $p = \infty$. And Shades of Gray is equivalent to (1) when $1 < p < \infty$.

Table 1. unsupervised methods

Methods	Parameter	Equation	Hypothesis
Gray-World	$e^{0,1,0}$	$\left(\int \|f(x)\|dx\right)=ke$	The average reflectance in a scene under a neutral light source is achromatic
Max-RGB	$e^{0,\infty,0}$	$\left(\int\|f(x)\|^{\infty}dx\right)=ke$	The maximum reflectance in a scene under a neutral light source is achromatic
Shades of Gray	$e^{0,p,0}$	$\left(\int\|f(x)\|^{p}dx\right)^{\frac{1}{p}}=ke$	The p-order Minkowski paradigm in the scene is achromatic
General Gray-Edge	$e^{0,p,\sigma}$	$\left(\int\|f^{\sigma}(x)\|^{p}dx\right)^{\frac{1}{p}}=ke$	After smoothing, the p-order Minkowski paradigm in the scene is achromatic
1st Gray-Edge	$e^{1,p,\sigma}$	$\left(\int\|f_x^{\sigma}(x)\|^{p}dx\right)^{\frac{1}{p}}=ke$	The p-order Minkowski paradigm of the 1-order image in the scene is achromatic
2nd Gray-Edge	$e^{2,p,\sigma}$	$\left(\int\|f_{xx}^{\sigma}(x)\|^{p}dx\right)^{\frac{1}{p}}=ke$	The p-order Minkowski paradigm of the 2-order image in the scene is achromatic

J.V. Weijer et al. proposed the Gray-Edge hypothesis which is based on the observation of the color derivative distribution of the image color in the oppositional color space [19]; the difference between the average reflections of all physical surfaces in the scene is achromatic. The methods of General Gray-Edge, 1st Gray-Edge, and 2nd Gray-Edge filter the image by the order of 0, 1 and 2 respectively. J.V. Weijer et al. unified the Max-RGB, Gray-World, Gray-Edge and other methods into a computational framework, as seen in (2). The specific parameters are shown in Table 1.

$$ke^{n,p,\sigma} = \left(\int \left| \frac{\partial^n f^{\sigma}(x)}{\partial x^n} \right|^p dx \right)^{\frac{1}{p}} \qquad (2)$$

p is the Minkowski norm, σ is the scale parameter of the Gaussian filter, n is the derivative order, k is the multiplication scalar constraint. And $f^{\sigma}(x)$ is defined as the convolution of the image f and the n-order Gaussian filter.

Because different edges may contain different illumination information, Gijsenij et al. weighted the Gray-Edge [18]:

$$\left(\int \left| w(f)^k \bullet f_{c,x}(x) \right|^p dx \right)^{\frac{1}{p}} = ke_c \qquad (3)$$

$w(f)$ is a weighted function, and k can be added to the weight. This method uses mirror edge detection to calculate the weight.

B. Supervised methods

The supervised methods study the information of the known illumination color information, then estimate the illumination information of the input image by means of an obtained rule or model [38]. These methods are divided into two categories: statistics-based and machine learning-based. Statistics-based methods use a statistical study of the possible light colors in a scene and predict the illumination color of the input image with the statistical rule obtained. Machine learning-based methods, whose expected model is obtained through learning from a training set, predict the light color of the input image by the model.

1) Gamut Mapping

Gamut Mapping is a supervised method based on statistics, which was proposed by Forsyth et al. in 1990 [19]. It is based on the assumption that the image in the real world, for a given illumination, can only accept limited color [39]. If two colors are observed under a light, all light colors would also be observed between them, and in a particular light, a collection of all possible colors forming a convex hull could be seen. Figure 1 is the basic flow of Gamut Mapping. This method first learns the reflective surface of a standard gamut and image gamut as much as possible, and then maps the two, obtaining feasibility mapping set N, and finally choosing the optimal mapping to calculate the unknown illumination color.
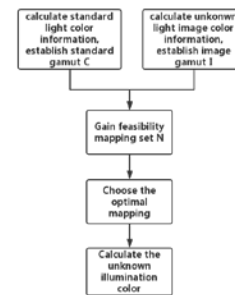

Fig1. the basic flow of Gamut Mapping

Finlayson et al. used convex optimization to improve color domain mapping, and improved it to a higher order image [40]. We can obtain eight methods by filtering the 0-order, 1-order, and 2-order image: Gamut pixel, Gamut fx, Gamut fy, Gamut 1grad, Gamut 2yy, Gamut 2xy, Gamut 2xx, Gamut 2grad.

2) Color Cat and Smart Color Cat

Color Cat (CC) is a machine learning method, which was proposed in 2012 by Nikola Banić and Sven Lončarić [24]. It relies on a normalized image chromaticity histogram with n3 bins in the RGB space. Firstly, it uses principal component analysis (PCA) to obtain a transformation matrix M and coefficients **c**; M can be used to extract the principal

components h'. Then, h' are used as features for linear regression with **c** to get $x \in [0,1]$, which is the normalized value between r_1 and r_2, values that were previously learned as likely extreme values. In the end, because of the high linear correlation between r and b in previously learned standard illumination, r and b can be calculated, and the color of estimated illumination is obtained. The details of Color Cat are shown in Table 2.

Table 2. the process of Color Cat method

1	Input image I
2	Gain chromaticity histogram h
3	principal component analysis, get transformation matrix M and coefficients **c**
4	h'=Mh
5	x=c^{T}h'
6	r=x(r$_1$-r$_0$)+r$_0$
7	b=a$_1$r+a$_0$
8	g=1-r-b
9	e=(r,g,b)^T

Color Cat has two problems using the chromaticity histogram. The first is that with the increase of n, the upper limit of n will be smaller if fewer training sets are used. Second, because a large number of bins are used to obtain the best value of n and the principal component k during the model selection process, the training time is long [25]. To improve the effect of Color Cat, Nikola Baniˊc and Sven Lonˇcariˊc proposed Smart Color Cat (SCC) [25], which saves training time. Because there are high correlations between individual histogram bins of image red chromaticity values and the red chromaticity component of the ground-truth, Smart Color Cat uses the red chromaticity histogram as the feature of the image instead of a chromaticity histogram; the rest of Smart Color Cat is the same as Color Cat, which can be seen in Table 3.

Table 3. the process of Smart Color Cat method

1	Input image I
2	Gain red chromaticity histogram h
3	principal component analysis, get transformation matrix M and coefficients **c**
4	h'=Mh
5	x=c^{T}h'
6	r=x(r$_1$-r$_0$)+r$_0$
7	b=a$_1$r+a$_0$
8	g=1-r-b
9	e=(r,g,b)^T

C. Combinational methods

Unsupervised methods have the advantages of little calculation and fast computation, and supervised methods can obtain a high accuracy of estimating illumination. However, both have an identical defect that the application scope is narrow. The combinational methods can improve the applicability of the method by integrating other methods.

Combinational methods can be divided into two groups: direct combination and guided combination [14]. Direct combination obtains an estimate by calculating given estimates, which can be further partitioned into a supervised combination and unsupervised combination. The unsupervised combination predefines the parameters, but the parameters of the supervised combination need to be trained. A guided combination uses the image semantic features to determine the choice of given estimates. In this paper, we focus on the unsupervised combination.

The performance of combinational methods is related to the selection of given estimates [41] [42]. We selected 15 methods to calculate given estimates: Gray-World, Max-RGB, Gray

Edge, 1st Gray Edge, 2nd Gray Edge, weighted Gray Edge, Gamut f$_x$, Gamut f$_y$, Gamut pixel, Gamut 2yy, Gamut 2xy, Gamut 2xx, Gamut 2grad, Gamut 1grad. Five unsupervised combinations were compared with other types of methods. The following is an introduction to the unsupervised combinations that we chose [14].

1) Simple Averaging (SA)

Average the given estimates and use this as the estimated illumination of the input image, as shown in (4).

$$c_e = \sum_{i=1}^{|E|} \frac{c_i}{|E|} \tag{4}$$

2) Nearest2 (N2)

Calculate the Euclidean distance between the given estimates in pairs, take the average of the pair of estimates, and use the smallest distance as the estimated illumination of the input image, the estimate is given by (5).

$$c_e = (c_n + c_m)/2 \quad d(c_n, c_m) = \min d(c_i, c_j) \tag{5}$$

3) Nearest-N% (N-N%)

Similar to N2, calculate the Euclidean distance two by two, take the average of (100 + N)% of the pair of estimates, and use the smallest distance as the estimated illumination of the input image. The formulation is (6).

$$c_e = \frac{\sum_{c_i \in K} c_i}{|K|} \tag{6}$$

$$K = \left\{ c_i \middle| \exists c_i \in K, s.t. d(c_i, c_j) \leq (100 + N)\% D_{\min} \right\}$$

4) No-N-Max (NNM)

Calculate the distances between each given estimate and other estimates, find the average, then discard the N distance values which are the maximum of the given estimates, and use the average of the rest of the given estimates c_i as the estimated illumination of the input image, the estimation is:

$$c_e = \sum_{i=1}^{|E|-N} \frac{c_i}{|E| - N} \tag{7}$$

5) Median (MD)

Similar to NNM, calculate the distances between each given estimate and other estimates and sum them, defined as $\sum D_{c_i}$. Next, take the minimum estimate as the estimated illumination of the input image.

$$c_e = \min(\sum D_{c_i}) \tag{8}$$

III. ERROR MEASUREMENT

There are four error metrics to measure how our chosen methods perform in different situations: angular error [32], chromaticity error [33], perceptual Euclidean distance (PED) [34], and reproduction angular error (RAE) [35].

A. Angular Error

This error metric represents the angle deviation of the estimated illumination $\vec{e}_e = (R_e, G_e, B_e)$ and the standard

illumination $\vec{e}_a = (R_a, G_a, B_a)$ in the RGB space, which is defined as:

$$E_a = \cos^{-1}\left(\frac{\vec{e}_a \bullet \vec{e}_e}{\|\vec{e}_a\| \bullet \|\vec{e}_e\|}\right) \qquad (9)$$

B. Chromaticity Error

The chromaticity error calculates the Euclidean distance of r and g components of the standard light and estimated illumination in RGB chromaticity space:

$$E_c = \sqrt{(r_a - r_e)^2 + (g_a - g_e)^2} \qquad (10)$$

r_a and g_a are the rg components of standard illumination, r_e and g_e are the rg components of the estimated illumination. The lower the value of E_c, and the smaller the distances of estimated and standard illumination, the better the method will perform.

C. Perceptual Euclidean Distance

The human visual system is more sensitive to green than red or blue. Because of this feature, the PED evaluation criteria, proposed by Gijsenij [34], is a weighted Euclidean distance between the standard illumination $\vec{e}_a = (R_a, G_a, B_a)$ and the estimated illumination $\vec{e}_e = (R_e, G_e, B_e)$ in RGB space, defined as:

$$E_p = \sqrt{w_r(r_a - r_e)^2 + w_g(g_a - g_e)^2 + w_b(b_a - b_e)^2} \qquad (11)$$

A lower value of E_p will result in a more effective method.

D. Reproduction Angular Error

The RAE measures the angle between the reproduction of a true achromatic surface under a white light and the actual reproduction of an achromatic surface when an estimated illumination color is divided out:

$$E_{re} = \cos^{-1}\left(\frac{\vec{e}_a / \vec{e}_e \bullet U}{|\vec{e}_a / \vec{e}_e| \bullet \sqrt{3}}\right) \qquad (12)$$

$$\frac{\vec{e}_a}{\vec{e}_e} \approx U = \frac{\vec{e}_a}{\vec{e}_a} \qquad (13)$$

We compared each method's performance using RAE, and the median value (Med) and maximum value (Max) were chosen to evaluate the statistical performance.

IV. EXPERIMENTS

In this section, all the methods chosen were tested on two real-world image sets, which are the SFU image set [42] and the Gray Ball image set [43] established by Barnard. The performance is evaluated by the RAE error measures.

The methods we compared were: (1) unsupervised methods: General Gray-Edge, 1st Gray-Edge, 2nd Gray-Edge, Gray-World, Max-RGB, Shades of Gray, and Weighted Gray-Edge; (2) supervised methods: Gamut fx, Gamut fy,

Gamut pixel, Gamut 2yy, Gamut 2xy, Gamut 2xx, Gamut 2grad, Gamut 1grad, Smart Color Cat, and Color Cat; (3) combinational methods: SA, NN%, N2, NNM, and MD.

A. Results on the Gray Ball Image Set

The Gray Ball image set has 11345 images and 15 different scenes. We made a total of 4051 images taken from 6 scenes as Open Country scene, 3462 images taken from 4 scenes Indoor scene, and 708 images taken from 1 scene as City scene.

For the unsupervised methods, we set the parameters of General Gray-Edge, 1st Gray-Edge, and 2nd Gray-Edge based on reference [27]. Smart Color Cat and Color Cat were two supervised methods that used a 15-fold cross test [24]. The N% of N-N% was set to 30% and N in the No-N-Max was 3. The overall performance is shown in Table 4 and Figure 2.

Table 4. performance comparison of all methods in Open Country scene, City scene, and Indoor scene (Med: median RAE error Max: maximum RAE error)

Methods		Open Country		City		Indoor	
		Med(°)	Max(°)	Med(°)	Max(°)	Med(°)	Max(°)
Unsupervised Methods	General Gray-Edge $\left(e^{0,9,0}\right)$	5.436	21.386	5.761	16.541	5.931	22.506
	1st Gray-Edge $\left(e^{1,1,1}\right)$	3.834	20.995	2.005	15.832	8.24	26.03
	2nd Gray-Edge $\left(e^{2,1,2}\right)$	4.117	22.471	2.365	15.966	8.582	26.095
	Gray-World	7.911	27.498	5.718	15.304	6.258	37.162
	Max-RGB	3.033	19.644	1.572	17.078	10.811	27.354
	Shades of Gray	5.622	21.621	6.041	16.884	5.468	24.495
	Weighted Gray-Edge	3.834	24.595	2.361	17.355	10.278	34.272
Supervised Methods	Gamut fx	4.387	31.19	4.02	19.867	8.887	27.386
	Gamut fy	4.577	24.64	4.108	18.45	9.112	26.407
	Gamut pixel	2.754	19.61	1.854	17.078	10.187	26.549
	Gamut 2yy	5.167	27.489	5.156	27.06	9.096	30.927
	Gamut 2xy	4.353	27.185	3.804	20.859	10.21	27.22
	Gamut 2xx	4.442	25.016	3.744	23.725	10.032	30.739
	Gamut 2grad	4.95	23.386	4.337	20.186	9.524	27.922
	Gamut 1grad	4.047	22.87	3.292	18.269	9.162	25.281
	SCC	2.899	29.6	2.525	15.671	5.465	26.887
	CC	3.2	22.515	2.777	17.11	6.345	33.449
Combinational Methods	SA	3.934	20.722	3.042	17.316	7.761	23.327
	N-N%	4.747	21.203	4.306	16.286	6.71	24.673
	N2	3.943	21.738	2.169	15.899	8.447	26.063
	NNM	3.842	20.381	2.949	17.162	8.231	23.745
	MD	3.834	20.995	2.005	15.832	8.24	26.03

Table 5 shows the performance of Gamut Mapping methods in the Open Country scene. From both Table 4 and Figure 2(a), the combinational methods do not show any improvement over the unsupervised and supervised methods. For unsupervised methods, the median RAE error of Max-RGB is 3.033, which is in third place. However, the Gray-World performs the worst of all methods with a median RAE error of 7.911. For Gray-Edge methods, a 1-order method (1st Gray-Edge) shows the best result. For supervised methods, a 0-order method (Gamut pixel) performs best of all the methods, which is much better than 1-order methods (Gamut fx, Gamut fy, Gamut 1grad) for the Gamut Mapping methods (Table 5). SCC is in second place overall, which is slightly better than CC.
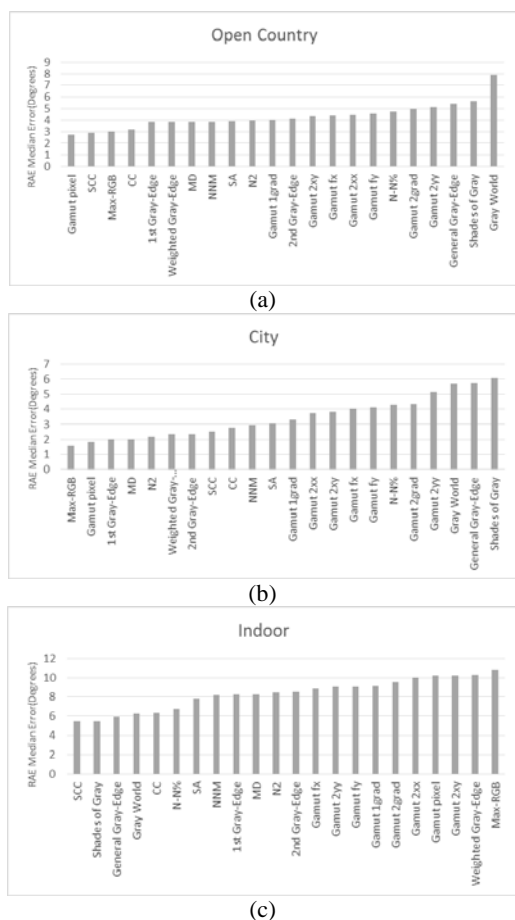
(a)

(b)

(c)

Fig 2. performance ranking of the methods based on the median of RAE error in 3 scenes (a) Open Country (b) City (c) Indoor

Table 5. performance comparison of Gamut Mapping methods in the Open Country scene(Med: median RAE error   M: the mean of median RAE error within each category)

	Methods	Med(°)	M(°)
0-order	Gamut pixel	2.754	2.754
1-order	Gamut fx	4.387	4.337
	Gamut fy	4.577	
	Gamut 1grad	4.047	
2-order	Gamut 2yy	5.167	4.728
	Gamut 2xy	4.353	
	Gamut 2xx	4.442	
	Gamut 2grad	4.950	

Table 6. performance comparison of Gamut Mapping methods in the City scene(Med: median RAE error   M: the mean of median RAE error within each category)

	Methods	Med(°)	M(°)
0-order	Gamut pixel	1.854	1.854
1-order	Gamut fx	4.020	3.807
	Gamut fy	4.108	
	Gamut 1grad	3.292	
2-order	Gamut 2yy	5.156	4.260
	Gamut 2xy	3.804	
	Gamut 2xx	3.744	
	Gamut 2grad	4.337	

Table 7. performance comparison of Gamut Mapping methods in the Indoor scene (Med: median RAE error   M: the mean of median RAE error within each category)

	Methods	Med(°)	M(°)
0-order	Gamut pixel	10.187	10.187
1-order	Gamut fx	8.887	9.054
	Gamut fy	9.112	
	Gamut 1grad	9.162	
2-order	Gamut 2yy	9.096	9.716
	Gamut 2xy	10.210	
	Gamut 2xx	10.032	
	Gamut 2grad	9.524	

Table 6 shows the performance of Gamut Mapping methods in the City scene. From Figure 2(b), in addition to the General Gray-Edge and Shades of Gray, the remaining methods in the City scene have a lower RAE error median. Similar to the Open Country scene, the combinational methods fail to improve the accuracy of unsupervised methods and supervised methods for the City scene. In the unsupervised methods, Max-RGB is the best method, with the median RAE error of 1.572; the median RAE error of Shades of Gray is 6.041, which is the worst of all methods. For Gray-Edge methods, a 1-order method (1st Gray-Edge) still outperforms other Gray-Edge methods. Among the supervised methods, SCC and CC do not show superiority; For Gamut Mapping methods, shown in Table 6, a 0-order method (Gamut pixel) is still the best and is in second place of all methods, followed by 1-order methods (Gamut fx, Gamut fy, Gamut 1grad).

From Table 4, the results of all methods in the Indoor scene are worse than in the other scenes. As shown in Figure 2(c), similar to the result of combinational methods in the Open Country scene and City scene, the combinational methods still have poorer performance that the other methods. A very interesting phenomenon in this scene is that Shades of Gray is ranked second Max-RGB is ranked last, which is quite different from the results in the previous two scenes. A 0-order method (General Gray-Edge) has the best result of the Gray-Edge methods. For supervised methods, SCC is clearly the best with median RAE error of 5.465, outperforming CC. Table 7 shows the performance of Gamut Mapping methods for the Indoor scene. It can be seen that 1-order methods (Gamut fx, Gamut fy, Gamut 1grad) have the lowest error, which is different from the results of the other two scenes.

B. Results on the SFU Image Set

The SFU image set has 321 images with 22 kinds of objects, and 11 different scene light sources in the laboratory. Because of the characteristics of the images in the SFU image set, we categorized it as Lab scene.

For General Gray-Edge, 1st Gray-Edge, 2nd Gray-Edge, we set the parameters as summarized in Table 8 [35] [19]. Weighted Gray-Edge had the same parameters as in the Gray Ball image set. For Smart Color Cat and Color Cat, we used a 2-fold cross test. In the combinational methods, N% is 30% in the N-N% and N is 3 in the No-N-Max method.

Table 8. performance comparison of all methods in Lab scene (Med: median RAE error Max: maximum RAE error)

|	Methods	Med(°)	Max(°)
Unsupervised Methods	General Gray-Edge $(e^{0,9,0})$	3.369	30.077
	1st Gray-Edge $(e^{1,14,4})$	3.976	31.683
	2nd Gray-Edge $(e^{2,15,10})$	3.843	40.843
	Gray-World	7.492	36.975
	Max-RGB	7.444	36.317
	Shades of Gray	4.249	30.864
	Weighted Gray-Edge	3.862	44.879
	Gamut fx	4.068	37.705
	Gamut fy	4.403	35.559
	Gamut pixel	3.644	29.046
Supervised Methods	Gamut 2yy	4.708	36.194
	Gamut 2xy	4.052	42.118
	Gamut 2xx	4.759	28.449
	Gamut 2grad	4.709	30.445
	Gamut 1grad	3.989	28.375
	SCC	9.965	25.38
	CC	9.446	32.337
Combinational Methods	SA	3.611	28.705
	N-N%	3.467	29.046
	N2	3.467	29.046
	NNM	3.406	28.878
	MD	3.989	28.375
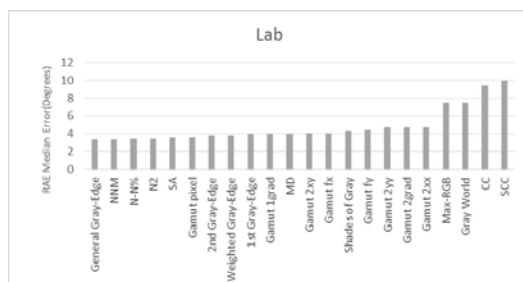
Fig 3. performance ranking of the methods based on the median of RAE error in the Lab scene

Table 8 and Figure 3 show the overall performance in the Lab scene, and Table 9 shows the performance of Gamut Mapping methods in the Lab scene. It can be seen in Figure 3 that, as a group, the ranking of combinational methods has notable improvement. Except for the methods in the last 4 bits, the median error of the other methods is not significant. Among the unsupervised methods, General Gray-Edge (0 order) is clearly the best of all methods with a median RAE error of 3.369 (Table 8) and a maximum RAE error of 30.077 (Table 8). In contrast, Gray-World and Max-RGB do not perform well. For the Gamut Mapping of supervised methods, as shown in Table 9, a 0-order method (Gamut pixel) is best and 2-order methods (Gamut 2yy, Gamut 2xy, Gamut 2xx, Gamut 2grad) are the worst. SCC and CC performance are significantly different than in the previous three scenes. This is probably because there are only 321 images in the set and the components of r and b in the rb chromaticity space are lower,

which can be seen from Figure 4: (a) 98.98%, (b) 95.66%, (c) 99.24%, (d) 95.25%.

Table 9. Performance comparison of Gamut Mapping methods in the Lab scene(Med: median RAE error   M: the mean of median RAE error within each category)

	Methods	Med(°)	M(°)
0-order	Gamut pixel	3.644	3.644
1-order	Gamut fx	4.068	4.153
	Gamut fy	4.403	
	Gamut 1grad	3.989	
2-order	Gamut 2yy	4.708	4.557
	Gamut 2xy	4.052	
	Gamut 2xx	4.759	
	Gamut 2grad	4.709	

(a)   (b)

(c)   (d)
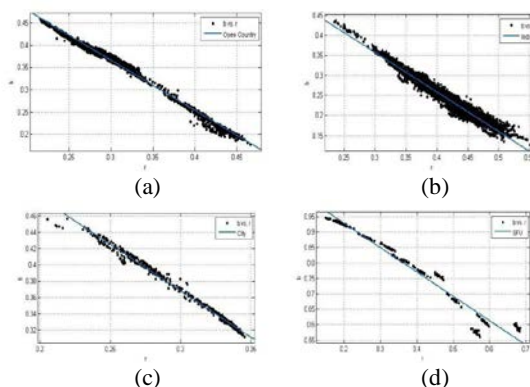
Fig 4. the correlation analysis of the ground-truth illuminations in the rb chromaticity space (a) Open Country (b) Indoor (c) City (d) Lab

V. CONCLUSION

In this paper, we have presented an experimental evaluation of several illumination estimation methods in various scenes. The goal was to compare the performance of those methods in specific scenes and obtain suggestions for selecting methods for different scenes.

Firstly, the results show that Smart Color Cat and Color Cat have better performance in the Open Country and Indoor scenes. So, if the image belonged to the Open Country scene and City scene, it would be suitable to choose one of these methods under the premise that there are plenty of images for training. For the Open Country and City scenes, Max-RGB or Gamut pixel would be a fine choice. However, General Gray-Edge or Shades of Gray can be used for the Indoor scene. The combinational methods do not show any superiority in these scenes.

A second conclusion can be drawn that in the Lab scene, General Gray-Edge would be an appropriate choice because Smart Color Cat and Color Cat are hampered. Combinational methods work better than other methods in the Lab scene.

Finally, according to the results of the Gray-Edge methods and Gamut Mapping methods, for the methods that are high order, if the scene image is like the Open Country or City scenes, 1-order methods are the best choice; in contrast,

0-order methods are better when the scene image is similar to the Indoor or Lab scenes.

References

[1] Foster, D. H., "Color constancy," *VISION RESEARCH*, vol. 51, pp.674-700, 2011.

[2] Funt, B., F. Ciurea, and J. McCann, "Retinex in MATLAB (TM)," *J ELECTRON IMAGING*, vol.13, pp.48-57, 2004.

[3] Land, E. H., "Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image," *Proceedings of the National Academy of Sciences of the United States of America*, vol.80, pp.5163-5169, 1983.

[4] Li, D., Y. Zhang, P. Wen, and L. Bai, "A Retinex Method for Image Enhancement Based on Recursive Bilateral Filtering," In *11th International Conference on Computational Intelligence and Security*, pp 154-157. 2015.

[5] Petro, A., C. Sbert, and J. Morel, "Multiscale Retinex," *Image Processing On Line*, vol.4, pp.71-88, 2014.

[6] Mikic, I. P. C. Cosman, Name G T Kogut, "Moving shadow and object detection in traffic scenes," In *Proceedings of the International Conference on Pattern Recognition. IEEE*, pp 321-324, 2000.

[7] van de Sande, K., T. Gevers, C. Snoek, I. C. Society, "Evaluating Color Descriptors for Object and Scene Recognition," *IEEE T PATTERN ANAL*, vol.32, pp.1582-1596, 2010.

[8] S. Huang, T., S. Mehrotra, K. Ramchandran, "Multimedia Analysis and Retrieval System (MARS) Project," in *Proceedings of the 33rd Annual Clinic on Library Application of Data Processing - Digital Image Access and Retrieval*, University of Illinois at Urbana-Champaign, pp.553-554, 1996.

[9] Rui, Y., T. S. Huang, S. Mehrotra, "Content-based image retrieval with relevance feedback in MARS," in *Proceedings of IEEE Int. Conf. on Image Processing*, Santa Barbara, California, pp.26-29, 1997.

[10] van de Sande, K. E. A., T. Gevers, C. G. M., "Snoek Evaluating Color Descriptors for Object and Scene Recognition," *IEEE T PATTERN ANAL* , vol.32, pp.1582-1596, 2010.

[11] Hordley, S. D., "Scene illuminant estimation: Past, present, and future.," *COLOR RESEARCH AND APPLICATION*, vol. 31, pp.303-314, 2006.

[12] Gijsenij, A., T. Gevers, J. van de Weijer, "Computational Color Constancy: Survey and Experiments," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol. 20, pp.2475-2489, 2011.

[13] Xingsheng Yuan, "Research on Color Constancy Computation and Its Application on Machine Vision," Changsha: National University of Defense Technology, 2014.

[14] Li, B., W. Xiong, W. Hu, B. Funt, I. S. P. Society, "Evaluating Combinational Illumination Estimation Methods on Real-World Images," *IEEE Transactions on Image Processing*, vol.23, pp.1194-1209, 2014.

[15] J. von Kries, "Influence of adaptation on the effects produced by luminous stimuli," in *Sources of Color Vision, D. MacAdam, Ed. MIT Press*, pp.109–119, 1970 .

[16] Zheng Tang, Hongzhe Liu, Jiazheng Yuan, "Advances Research on Color Constancy Computation under Single Illuminant," *Computer science*, vol.43, pp.12-18, 2016.

[17] Land, E. H., "The retinex theory of color vision," *Scientific American*, vol.237, pp.108-128, 1977.

[18] Buchsbaum, G., "A spatial processor model for object colour perception," *Journal of the Franklin Institute*, vol.310, pp.1-26, 1980.

[19] Van de Weijer, J., T. Gevers, A. Gijsenij, "Edge-based color constancy," *IEEE TRANSACTIONS ON IMAGE PROCESSING*, vol.16, pp.2207-2214, 2007.

[20] Gijsenij, A., T. Gevers, and J. van de Weijer, "Generalized Gamut Mapping using Image Derivative Structures for Color Constancy," *INTERNATIONAL JOURNAL OF COMPUTER VISION*, vol. 86, pp.127-139, 2010.

[21] Finlayson, G. D., S. D. Hordley, P. M. Hubel, "Color by Correlation: A Simple, Unifying Framework for Color Constancy," *IEEE Trans. Pattern Anal. Mach. Intell*, vol.23, pp.1209-1221, 2001.

[22] Brainard, D. H., W. T. Freeman, "Bayesian color constancy," *JOURNAL OF VISION*, vol.14, pp.1393-1411, 1997.

[23] R. Rosenberg, C., T. P. Minka, A. Ladsariya, "Bayesian Color Constancy with Non-Gaussian Models," in *Neural Information Processing Systems*, 2003.

[24] Banic, N., S. Loncaric, "Color Cat: Remembering Colors for Illumination Estimation," *IEEE Signal Processing Letters*, vol.22, pp.651-655, 2015.

[25] Banic, N. and S. Loncaric, "Using the red chromaticity for illumination estimation," in *2015 9th International Symposium*, pp.131-136, 2015.

[26] L.TÖRÖK, ÁKOS ZARÁNDY, "CNN based color constancy method," in *Proceedings of the 7th IEEE International Workshop Frankfurt*, pp.22-24, 2002.

[27] C. Cardei, V. and B. Funt, "Committee-Based Color Constancy," in *IS&T/SID's Color Imaging Conference*, pp.311-313, 1999.

[28] Bianco. S, Ciocca.G, Cusano.C, Schettini.R, "Automatic color constancy algorithm selection and combination," PATTERN RECOGNITION, vol.43, pp.695-705, 2010.

[29] Li, B., D. Xu, C. Lang, "Colour constancy based on texture similarity for natural images," *COLOR TECHNOL*, vol.125, pp.328-333, 2009.

[30] Bianco, S., G. Ciocca, C. Cusano, R. Schettini, "Automatic color constancy method selection and combination," *PATTERN RECOGN*, vol.43, pp.695-705, 2010.

[31] Wu, M., J. Sun, J. Zhou, G. Xue, "Color constancy based on texture pyramid matching and regularized local regression," *JOURNAL OF THE OPTICAL SOCIETY OF AMERICA A-OPTICS IMAGE SCIENCE AND VISION*, vol.27, pp.2097-2105, 2010.

[32] Barnard, K., V. Cardei, B. Funt, "A comparison of computational color constancy methods - Part I: Methodology and experiments with synthesized data," *IEEE TIP*, vol.11, pp.972-983, 2002.

[33] Gijsenij, A., T. Gevers, M. P. Lucassen, "Perceptual analysis of distance measures for color constancy methods," *Journal of the Optical Society of America. A, Optics, image science, and vision*, vol.26, pp.2243-2256, 2009.

[34] Gijsenij, A., T. Gevers, M. P. Lucassen, "A Perceptual Comparison of Distance Measures for Color Constancy Methods," in *Proceedings of the 10th European Conference on Computer Vision: Part I. Springer-Verlag*, Marseille, France, pp 208-221, 2008.

[35] Finlayson, G., R. Zakizadeh, "Reproduction Angular Error: An Improved Performance Metric for Illuminant Estimation," in *Proceedings of the British Machine Vision Conference*, 2014.

[36] Bianco, Simone Schettini, Raimondo, "Computational color constancy," *3rd European Workshop on Visual Information Processing, EUVIP 2011 - Final Program*, pp.1-7,2011.

[37] Bing Li, "Research on Color Constancy Computation," Beijing: Beijing Jiaotong University, 2009.

[38] Martinez-Verdu, F., M. J. Luque, J. Malo, A. Felipe, J. M. Artigas, "Implementations of a novel method for colour constancy," *Vision research*, vol.37, pp.1829-1844, 1997.

[39] Gijsenij, A., T. Gevers, J. van de Weijer, "Generalized Gamut Mapping using Image Derivative Structures for Color Constancy," *INTERNATIONAL JOURNAL OF COMPUTER VISION*, vol.86, pp.127-139, 2010.

[40] Bianco, S., F. Gasparini, R. Schettini, "Consensus-based framework for illuminant chromaticity estimation," *Journal of Electronic Imaging*, vol.17, 2008.

[41] Barnard, K., L. Martin, B. Funt, A. Coath, "A data set for color research," *Color Research and Application*, vol.27, pp.147-151, 2002.

[42] Ciurea, F., B. Funt, "A Large Image Database for Color Constancy Research," in *Proceedings of the Imaging Science and Technology Eleventh Color Imaging Conference*, pp.160-164, 2003.