

Region-Based Illuminant Estimation for Effective Color Correction

Simone Bianco, Francesca Gasparini, and Raimondo Schettini

DISCo (Dipartimento di Informatica, Sistemistica e Comunicazione),
Università degli Studi di Milano-Bicocca, Viale Sarca 336, 20126 Milano, Italy
{bianco, gasparini, schettini}@disco.unimib.it

Abstract. Several algorithms were proposed in the literature to recover the illuminant chromaticity of the original scene. These algorithms work well only when prior assumptions are satisfied, and the best and the worst algorithms may be different for different scenes. In particular for certain images a do nothing strategy can be preferred. Starting from these considerations, we have developed a region-based color constancy algorithm able to automatically select (and/or blend) among different color corrections, including a conservative do nothing strategy. The strategy to be applied is selected without any a priori knowledge of the image content and only performing image low level analysis.

1 Introduction

The colors in a digital image depend not only on the surface properties of the objects present in the scene depicted, but also on the illuminant conditions (lighting geometry and illuminant color) and the characteristics of the capturing device. Unlike human vision, imaging devices (such as digital cameras) can not adapt their spectral responses to cope with different lighting conditions; as a result, the acquired image may have a cast, i.e. an undesirable shift in the entire color range. Color constancy aims to estimate the actual color in an acquired scene disregarding its illuminant. The different approaches can be broadly classified into color invariant and illuminant estimation [1]. The former approaches derive from the image data invariant color descriptors without estimating explicitly the scene illuminant. The latter is actually a two stage procedure: the scene illuminant is estimated from the image data, and the image colors are then corrected on the basis of this estimate to generate a new image of the scene as if it were taken under a known, canonical illuminant. Many illuminant estimation solutions have been proposed in the last few years although it is known that the problem addressed is actually ill-posed as its solution lack uniqueness or stability. To cope with this problem, different solutions usually exploit some assumptions about the statistical properties of the expected illuminants and/or of the object reflectances in the scene. We have recently considered some well known and widely used color constancy algorithms that are based on color image statistics, and we have shown that the best color correction algorithm with respect to all the possible image typologies does not exist [2]. In particular color

correction algorithms work well only when their prior assumptions are satisfied. Moreover in certain circumstances a do nothing strategy can be preferred with respect to a method that introduces a severe color distortion.

Most of the color correction algorithms based on a two step process, adopt for the color correction step a diagonal model of illumination change. This model is derived from the Von Kries hypothesis that color constancy is an independent gain regulation of the three cone signals, through three different gain coefficients [20]. The Gray World and the White Patch, which correspond respectively to average-to-gray and normalize-to-white, are perhaps the most commonly used. The Gray world algorithm assumes that, given an image with a sufficient amount of color variations, the mean value of the R, G, B components of the image will average out to a common gray value. Once we have chosen the common gray value, each image color component is scaled by applying a Von Kries transformation with coefficients:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_{out} = \begin{bmatrix} Gray_R/R_{avg} & 0 & 0 \\ 0 & Gray_G/G_{avg} & 0 \\ 0 & 0 & Gray_B/B_{avg} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{in} \quad (1)$$

where $R_{avg}, G_{avg}, B_{avg}$ are the averages of the three RGB channels, and $Gray_R, Gray_G, Gray_B$ represent the gray value chosen. There are several versions of the white patch algorithm: the basic concept is to set at white a point or a region that appears reasonably white in the real scene. The skill lies in identifying this point or region in the acquired image. The Von Kries coefficients are evaluated



Fig. 1. Examples of different color corrections. First column: original images; second column: corresponding images processed by a white patch algorithm; third column corresponding images processed by a Gray World algorithm. Note that for certain images a do nothing strategy is preferred (for example for the image of the last row.)

setting respectively the maximum value of each channel or a potential white region at a reference white.

Example of images processed with these simple color correction algorithms are reported in Figure 1. In the first column original images are shown. In the second column the corresponding images processed by a White Patch algorithm are reported. While the third column refers to images processed by a Gray World algorithm. Note that for the image of the first row the White Patch algorithm could be considered a good processing, while for the second row image the Gray World performs better. For the image of the last row instead a do nothing strategy is preferred.

Starting from these considerations, we have developed a region-based color constancy algorithm able to automatically select (and/or blend) “do nothing strategy”, Gray World and White Patch approach [10]. The selection/blending strategy to be applied is selected without any a priori knowledge of the image content and only performing image low level analysis.

2 Region-Based Color Correction

The region-based color constancy algorithm we propose is based on the Von Kries hypothesis. We look for what we have called the “White Balance Region (WBR)” i.e. the reference region to be neutralized to obtain the three gain coefficients. The idea is to select this region with respect to the color characteristic of the image so that the color correction will introduce less damage as possible. A dominant color in a scene could be a superimposed cast due to the illuminant, but it could also be the intrinsic color of the scene such as in the case of close-up images or portraits, or in case of images with a dominant portion of sky or sea or vegetation. We try to distinguish among true casts and intrinsic dominant colors without an image content analysis, but only analyzing the color distribution of the image in a proper color space with simple statistical tools. In this way it is possible not only to evaluate whether or not a cast is present, but also to classify it. The following color correction, (i.e. the estimate of the sensor scaling coefficients) is assimilated within the problem of quantifying the cast. Four cast classes are considered: i) no cast images; ii) evident cast images; iii) low cast images; iv) critical images: images with a predominant color that must be preserved or unclassifiable images. When we identify a cast, our approach tends to be a white patch approach, in case of low or ambiguous cast, we adopt a gray world like approach, while in the case of critical images (dominant intrinsic colors) or images without cast we adopt the do nothing strategy.

2.1 IPT Color Space

As we are interested in analyzing the image color distribution to correlate statistical measurements with the four cast classes defined above, we are interested in a color space that is possibly simple and perceptually uniform. With this aim we have adopted the IPT color space [11], that is more uniform in perceived hue than other commonly used color spaces, such as CIELAB or CIECAM97s. In

the IPT color space lightness dimension is denoted as I, the red-green dimension is denoted as P, while the yellow-blue dimension is denoted as T. The model consists of a 3x3 matrix, followed by a nonlinearity, followed by another 3x3 matrix. The model assumes input data is in CIEXYZ for the 1931 2-deg. observer with an illuminant of D65. The color transformation is described by Equations 2:

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix} = \begin{bmatrix} 0.4002 & 0.7075 & -0.0807 \\ -0.2280 & 1.1500 & 0.0612 \\ 0.0000 & 0.0000 & 0.9184 \end{bmatrix} \begin{bmatrix} X_{D65} \\ Y_{D65} \\ Z_{D65} \end{bmatrix}$$

$$\begin{aligned} L' &= L^{0.43}; L \geq 0 \\ L' &= -(-L)^{0.43}; L < 0 \\ M' &= M^{0.43}; M \geq 0 \\ M' &= -(-M)^{0.43}; M < 0 \\ S' &= S^{0.43}; S \geq 0 \\ S' &= -(-S)^{0.43}; S < 0 \end{aligned}$$

$$\begin{bmatrix} I \\ P \\ T \end{bmatrix} = \begin{bmatrix} 0.4000 & 0.4000 & 0.2000 \\ 4.4550 & -4.8510 & 0.3960 \\ 0.8056 & 0.3572 & -1.1628 \end{bmatrix} \begin{bmatrix} L' \\ M' \\ S' \end{bmatrix} \quad (2)$$

2.2 Color Cast Image Classification

The basis of our procedure is that statistical analysis of image color distribution can provide information about the color image characteristics, permitting to automatically adapt the color correction process. We assume here that the images are coded in terms of RGB color coordinates. These RGB values are mapped first into the CIEXYZ color space and then into the IPT color space following the color transformation described by Equations 2. To classify the color cast of the image into the four classes defined above, we analyze the 2D histogram $F(P, T)$ of the chromatic components PT. If the histogram is concentrated and far from the neutral axis, the colors of the image are thus confined to a small region in the PT chromatic diagram. The images are likely to have either an evident cast (to be removed), or an intrinsic color (to be preserved: widespread areas of vegetation, skin, sky, or sea or close-up image). Color histograms uniformly distributed around the neutral axis ($P = 0, T = 0$), instead, correspond to images without cast. Examples of images respectively with a strong cast, with no cast and with an intrinsic color, together with their corresponding color histograms, are reported in Figure 2.

The statistical measures adopted to analyze the $F(P, T)$ color histogram are simply the mean and the variance of the 2D color histogram distribution. We define the mean values and the variances of the histogram projections along the two chromatic axes P and T as follows:

$$\mu_k = \int_k k F(P, T) dk \quad (3)$$

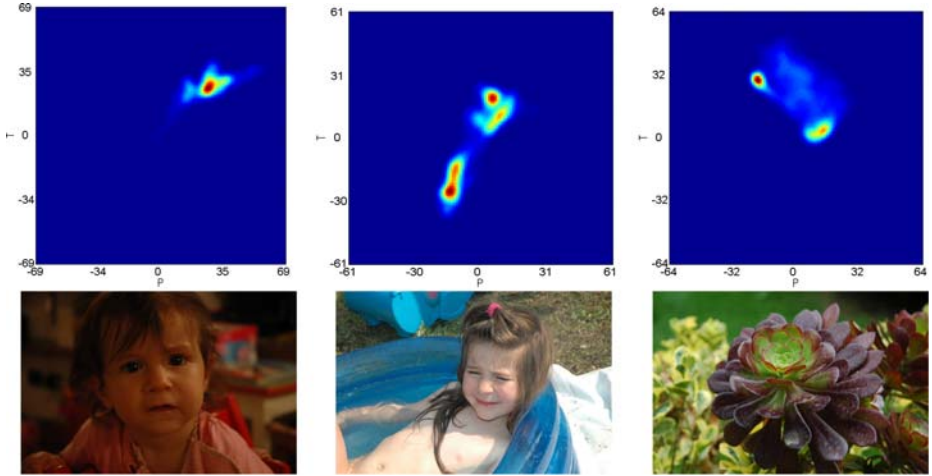


Fig. 2. Examples of 2D histogram of the chromatic components PT : evident cast (left), no cast (center), critical image (right)

and

$$\sigma_k^2 = \int_k (\mu_k - k)^2 F(P, T) dk \quad (4)$$

with $k = P, T$. Using these measures we can associate to each image histogram an Equivalent Circle (EC) with center $C = (\mu_P, \mu_T)$ and radius $\sigma = \sqrt{\sigma_P^2 + \sigma_T^2}$. To quantitatively characterize the EC we introduce a distance D defined as $D = \mu - \sigma$, where $\mu = \sqrt{\mu_P^2 + \mu_T^2}$, and the ratio $D_\sigma = D/\sigma$. D is thus a measure of how far the whole histogram (identified with its EC) lies from the neutral axis ($P = 0, T = 0$), while σ is a measure of how the histogram is spread. The idea is that analyzing these measurements makes it possible to quantify the strength of the cast. For instance, if the histogram (and thus its corresponding EC) is concentrated (small value of σ) and far from the neutral axis (high value of μ and D_σ), the colors of the image are confined to a small region. This means that there is either a strong cast (to be removed), or what we have called a dominant intrinsic color (to be preserved). On the other hand, color histograms corresponding to spread EC (high values of σ), centered near the neutral axis (low value of μ and D_σ) could correspond to no cast or feeble cast images. The boundaries that distinguish among these cast classes in terms of mean, variance, D and D_σ , were determined applying the Particle Swarm Optimization method described in the next section, on a suitable training set.

3 Particle Swarm Optimization Description

The parameters for our Region-based color constancy are found using Particle Swarm Optimization (PSO)[3] over the set of feasible solutions. PSO is a population based stochastic optimization technique which shares many similarities

with evolutionary computation techniques. A population of individuals is initialized as random guesses to the problem solutions; and a communication structure is also defined, assigning neighbours for each individual to interact with. These individuals are candidate solutions. An iterative process to improve these candidate solutions is set in motion. The particles iteratively evaluate the fitness of the candidate solutions and remember the location where they had their best success. The individual's best solution is called the particle best or the local best. Each particle makes this information available to its neighbours. They are also able to see where their neighbours have had success. Movements through the search space are guided by these successes. The swarm is typically modeled by particles in multidimensional space that have a position and a velocity. These particles fly through hyperspace and have two essential reasoning capabilities: their memory of their own best position and their knowledge of the global or their neighborhood's best position. Members of a swarm communicate good positions to each other and adjust their own position and velocity based on these good positions.

4 Experimental Results

4.1 Color Constancy Algorithms Used for Benchmarking

In this work, we chose for benchmark six widely used algorithms in the state of the art. These algorithms were chosen as exploiting only low-level image information.

These can be seen as different instantiations of a recently proposed equation [18] that unifies a variety of algorithms. The algorithms considered are:

1. Gray World (GW) algorithm [14], which is based on the assumption that the average reflectance in a scene is achromatic.
2. White Point (WP) algorithm [15], also known as Maximum RGB, which is based on the assumption that the maximum reflectance in a scene is achromatic.
3. Shades of Gray (SG) algorithm [16], which is based on the assumption that the p -th Minkowski norm of a scene is achromatic.
4. General Gray World (gGW) algorithm [17,18], which is based on the assumption that the p -th Minkowski norm of a scene after local smoothing is achromatic.
5. Gray Edge (GE1) algorithm [18], which is based on the assumption that the p -th Minkowski norm of the first order derivative in a scene is achromatic.
6. Second Order Gray Edge (GE2) algorithm [18], which is based on the assumption that the p -th Minkowski norm of the second order derivative in a scene is achromatic.

As can be noticed, four of the algorithms considered have one or two parameters that can be opportunely tuned for a particular image: SG has the p -th Minkowski norm, while gGW, GE1 and GE2 have both the p -th Minkowski norm and the smoothing parameter.

4.2 Dataset Selection

In our experiments we used a subset of the dataset presented by Ciurea and Funt [4]. The original dataset is commonly used in the evaluation of color constancy algorithms as it is labeled with the ground truth illuminants. In the dataset, 15 digital video clips were recorded (at 15 frames per second) in different settings such as indoor, outdoor, desert, markets, cityscape, etc. . . for a total of two hours of videos. From each clip, a set of images was extracted resulting in a dataset of more than 11,000 images. A gray sphere appears in the bottom right corner of the images and was used to estimate the true color of the scene illuminant. Since the dataset sources were video clips, the images extracted show high correlation. To remove this correlation, as already been done in [7,8], only a subset of images should be used from each set. Taking into account that the image sets came from video clips, we applied a two stage video-based analysis to select the image to be included in the final illuminant dataset.

In the first stage, a video clip is reconstructed from each set of images removing the right part of the images containing the gray sphere. The video clip is fed to a key frame extraction algorithm [5] which dynamically selects a set of candidate images by analyzing the visual contents of consecutive frames. Clips showing high variability in their pictorial contents will have a high number of images extracted while clips showing little or no variability will have only a single image extracted.

As a trade-off between the number of images to be included in the dataset and the correlation problem, we set the parameters of the key frame extraction algorithm so that the images extracted correspond to at least 10% of the clip size.

In the second stage, we further processed the extracted images with a visual summary post-processing algorithm [6]. For this work, we exploited only the key frame grouping processing step that eliminates pictorially similar images, using a hierarchical clustering algorithm.

The clustering algorithm further removes redundancies within the set of images. At each step one image is removed from the set and the clustering process stops when the number of remaining images is exactly 10% of the clip size. The final subset consists of 1,135 images.

4.3 Error Measure and Statistical Significance of the Results

To evaluate the performance of the algorithms on the dataset, we have used an intensity independent error measure. As suggested by Hordley and Finlayson [9], we use the angle between the RGB triplets of the illuminant color (ρ_w) and the algorithm's estimate of it ($\hat{\rho}_w$) as error measure:

$$e_{ANG} = \arccos \left(\frac{\rho_w^T \hat{\rho}_w}{\|\rho_w\| \|\hat{\rho}_w\|} \right). \quad (5)$$

It has been shown also [9] that the median error is a good descriptor of the angular error distribution.

Our color constancy algorithms and four of the color constancy algorithms considered for comparison (SG, gGW, GE1, GE2), needed a training phase to

opportunistically tune their parameters. The best parameters chosen for the SG, gGW, GE1, GE2 are the same used in [7,8], where the median error has been used as a cost function and it was optimized using a Pattern Search Method (PSM)[12,13]. The training set consisted of 300 randomly extracted images from the 1,135 images of the illuminant dataset. The remaining 835 images were used as test set, on which the algorithm performance are assessed. In order to compare the whole error distribution between different algorithms, together with the median angular error, we used the Wilcoxon Sign Test (WST) [21]. A score is generated by counting the number of times that the performance of a given algorithm has been considered to be better than the others.

The results in terms of both the angular error and the Wilcoxon score for the training and test sets are reported in Table 1. For what concerns the training set, the performance of almost all the benchmark algorithms considered were statistically equivalent. Our region-based algorithm, instead, performed statistically better than all the benchmark algorithms considered. On the test set instead, the benchmark algorithms tend to form two different statistically indistinguishable groups. Our region-based algorithm was still statistically the best one.

Table 1. Median angular error obtained by the color constancy algorithms on the training and test set. The best results for each column are reported in bold.

Method	Training set		Test set	
	Median	WSTs	Median	WSTs
GW	5.62	1	5.95	0
WP	7.76	0	5.48	3
SG	5.56	1	5.80	0
gGW	5.57	1	5.80	0
GE1	5.45	1	4.47	4
GE2	5.47	1	4.65	4
Region-based	3.89	6	3.74	6

5 Conclusions

In this work we have proposed a new region-based illuminant estimation algorithm for an effective color correction. The algorithm proposed is able to automatically select among different color corrections, including also a conservative do nothing strategy. The strategy to be applied is selected without any a priori knowledge of the image content and only performing image low level analysis.

Experimental results, performed on a suitable subset of the widely used Funt and Ciurea dataset, demonstrate that our algorithm is able to improve the results with respect to widely used benchmark algorithms. From our experiments the region-based algorithm proposed reduced the median angular error by 19.6% with respect to the best benchmark algorithm considered.

As future work we plan to investigate if the results of the combination technique proposed can be improved using more performing illuminant estimation algorithms.

References

1. Hordley, S.D.: Scene illuminant estimation: Past, present, and future. *Color Res. Appl.* 31(4), 303–314 (2006)
2. Bianco, S., Gasparini, F., Schettini, R.: A consensus based framework for illuminant chromaticity estimation. *Journal of Electronic Imaging* 17, 023013-1–023013-9 (2008)
3. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: *Proc. IEEE International Conference on Neural Networks*, vol. 4, pp. 1942–1948 (1995)
4. Ciurea, F., Funt, B.: A Large Image Database for Color Constancy Research. In: *Proc. IS&T/SID 11th Color Imaging Conference*, pp. 160–164 (2003)
5. Ciocca, G., Schettini, R.: An Innovative Algorithm for Key Frame Extraction in Video Summarization. *Journal of Real-Time Image Processing* 1(1), 69–88 (2006)
6. Ciocca, G., Schettini, R.: Supervised And Unsupervised Classification Post-Processing for Visual Video Summaries. *IEEE Transactions on Consumer Electronics* 2(52), 630–638 (2006)
7. Bianco, S., Ciocca, G., Cusano, C., Schettini, R.: Classification-based Color Constancy. In: Sebillo, M., Vitiello, G., Schaefer, G. (eds.) *VISUAL 2008*. LNCS, vol. 5188, pp. 104–113. Springer, Heidelberg (2008)
8. Bianco, S., Ciocca, G., Cusano, C., Schettini, R.: Improving Color Constancy Using Indoor–Outdoor Image Classification. *IEEE Transactions on Image Processing* 17(12), 2381–2392 (2008)
9. Hordley, S.D., Finlayson, G.D.: Re-evaluating Color Constancy Algorithms. In: *Proc. 17th International Conference on Pattern Recognition*, pp. 76–79 (2004)
10. Land, E.: The retinex theory of color vision. *Scientific American* 237(6), 108–128 (1977)
11. Ebner, F., Fairchild, M.D.: IDevelopment and Testing of a Color Space (IPT) with Improved Hue Uniformity. In: *IS&T/SID Sixth Color Imaging Conference: Color Science, Systems and Applications*, vol. 6, pp. 8–13 (1998)
12. Lewis, R.M., Torczon, V.: Pattern search algorithms for bound constrained minimization. *SIAM Journal on Optimization* 9, 1082–1099 (1999)
13. Lewis, R.M., Torczon, V.: Pattern search methods for linearly constrained minimization. *SIAM Journal on Optimization* 10, 917–941 (2000)
14. Buchsbaum, G.: A spatial processor model for object color perception. *Journal of Franklin Institute* 310, 1–26 (1980)
15. Cardei, V., Funt, B., Barndard, K.: White point estimation for uncalibrated images. In: *Proc. IS&T/SID 7th Color Imaging Conference*, pp. 97–100 (1999)
16. Finlayson, G., Trezzi, E.: Shades of gray and colour constancy. In: *Proc. IS&T/SID 12th Color Imaging Conference*, pp. 37–41 (2004)
17. Barnard, K., Cardei, V., Funt, B.: A comparison of computational color constancy algorithms; part two: Experiments with image data. *IEEE Transactions on Image Processing* 11(9), 985–996 (2002)

18. van de Weijer, J., Gevers, T., Gijzenij, A.: Edge-based Color Constancy. *IEEE Transactions on Image Processing* 16(9), 2207–2214 (2007)
19. Funt, B., Barnard, K., Martin, L.: Is machine colour constancy good enough? In: *Proc. 5th European Conference on Computer Vision*, pp. 445–459 (1998)
20. Fairchild, M.D.: *Color Appearance Models*. Addison Wesley, Reading (1997)
21. Wilcoxon, F.: Individual comparisons by ranking methods. *Biometrics* 1, 80–83 (1945)