PROBABILISTIC COLOR CONSTANCY

Firas Laakom[†], *Jenni Raitoharju*[†§], *Alexandros Iosifidis*[⋆], *Uygar Tuna*[‡], *Jarno Nikkanen*[‡], *Moncef Gabbouj*[†]

[†]Tampere University, Faculty of Information Technology and Communication Sciences, Finland
[§] Finnish Environment Institute, Programme for Environmental Information, Finland
[⋆]Aarhus University, Department of Engineering, Denmark
[‡] Xiaomi, Finland

ABSTRACT

In this paper, we propose a novel unsupervised color constancy method, called Probabilistic Color Constancy (PCC). We define a framework for estimating the illumination of a scene by weighting the contribution of different image regions using a graph-based representation of the image. To estimate the weight of each (super-)pixel, we rely on two assumptions: (Super-)pixels with similar colors contribute similarly and darker (super-)pixels contribute less. The resulting system has one global optimum solution. The proposed method achieves competitive performance, compared to the state-of-the-art, on INTEL-TAU dataset.

Index Terms— Color constancy, illumination estimation, graph-based learning

1. INTRODUCTION

In the image processing pipeline of every camera, the impact of illumination on the image colors is attenuated. This well-established computer vision problem is usually referred to as color constancy [1]. Discounting the effect of the illumination and restoring the 'true' colors of objects in the scene improves the subjective image quality and it is an essential building block in several higher-order image processing systems, such as object tracking and object classification [2]. It allows objects and scene content to have consistent color distribution under different lighting conditions. Thus, they are easier tracked, classified, and recognized. Endowing digital cameras with this ability is a challenging and ill-posed problem. Computational color constancy approaches usually rely on a two-step process. In the first step, the global illumination is estimated and, in the second step, all color pixels of the scene are normalized using the estimated illuminant color. Thus, the color constancy problem is equivalent to illumination estimation.

Recently, with the advancement of machine learning techniques in general and deep learning in particular, several neural network-based illumination estimation techniques have been proposed [3, 4, 5, 6]. These methods have usually led to state-of-the-art results over several datasets. However, they require calibrated images with known ground-truth illumination to fine-tune their parameters and creating such images for a given sensor is time-consuming. Besides, supervised approaches rely on the assumption that the inference of a given scene colors distribution and illumination conditions are the same in the training and testing phases. Thus, they tend to over-fit to training data and are sensitive to camera models and scenes encountered during the training phase [5]. We thus argue there is still interest in unsupervised methods as they usually do not require any training data and are independent of the acquisition device. Furthermore, unsupervised methods are computationally fast and have small memory and power footprint compared to deep learning approaches. Thus, they are more convenient especially for low-power devices, such as mobile phones.

Unsupervised methods usually make assumptions about the nature of the colors and statistical properties of the scene to estimate illumination [7]. The Gray world algorithm [8], for example, assumes that illumination color is the average of colors present in a scene, while the White-Patch [9] uses the maximum RGB responses in a scene as an estimate for illumination. In [10] and [11], the problem of color constancy is reformulated as the search of grey pixels, i.e., pixels representing an achromatic surface and using the cast of the grey pixels as an illumination estimate. Most of these methods assume that the illumination color is one of the scene pixel colors. In this paper, we relax this constraint and assume that the color of the illumination belongs to the convex hull of the pixel colors present in the scene. Thus, it can be obtained as a convex combination of the scene colors. We propose a probability-based formulation defined on the color similarity graph of an image to obtain such a solution.

Graphs offer a rich and compact representation of the spatial and color-wise similarity relations between pixels/regions in an image. They allow to model constraints and properties of a particular problem in a single framework solving a joint optimisation problem. They have been used to model many tasks in computer vision, such as saliency estimation [12], object tracking [13], and image segmentation [14, 15, 16]. In this paper, we model the color constancy problem under a probabilistic framework encoding the illumination assumptions and smoothness constraints into an optimization problem. We rely on three key assumptions to model the connectivity between (super-)pixels [17] and the constraints used to define the optimization process. The first assumption is that (super-)pixels with similar colors encode the same information about the global illumination. The second assumption is that darker pixels encode less information about the scene illumination than the brighter ones. The third assumption is that the scene illumination can be obtained by a weighted sum of the RGB colors of the pixels in the scenes. The first two constraints enable us to define the connectivities of the graph. The adopted optimization scheme has a closed-form global optimum solution, which estimates the color of the light source in a scene.

The main contributions of the paper are as follows:

- We propose a novel probability-based unsupervised color constancy method, called Probabilistic Color Constancy (PCC). Following this approach, we formulate a novel optimization problem based on the color similarity graph of an image.
- The proposed approach was tested on the INTEL-TAU dataset, where it achieved comparable results with state-of-the art unsupervised color constancy methods.
- The proposed unsupervised approach can be used as a generic framework for color constancy approaches, where different prior information can be integrated in the future.

2. RELATED WORK

2.1. Color Constancy

The aim of color constancy methods is to estimate the global illumination of a scene in order to correct the colors of an image representing the scene. Several unsupervised methods were grouped into a single framework [18], where the illumination estimate $\mathbf{I}^{est}$ is expressed as follows:

$$\mathbf{I}^{est}(n,p,\sigma) = \frac{1}{k}\left(\int_x\int_y |\bigtriangledown^n \rho_\sigma(x,y)|^p dxdy\right)^{\frac{1}{p}}, \quad (1)$$

where n denotes the derivative order, p the order of the Minkowski norm, and k the normalization constant for $\mathbf{I}^{est}$. Also, $\rho_\sigma(x,y) = \rho(x,y) * g_\sigma(x,y)$ denotes the image convolution, .ie., $\rho(x,y)$, with a Gaussian filter $g_\sigma(x,y)$ having a scale parameter σ. This framework allows for deriving different algorithms simply by setting appropriate values for n, p and σ. The Gray-World method [8] corresponds to $(n=0, p=1, \sigma=0)$. It corresponds to the average of all the pixel values in the scene. The White-Patch [9] corresponds to $(n=0, p=\infty, \sigma=0)$. It considers the pixel with the maximum response on each channel as the main estimate of the illumination. The Gray-Edge methods [18] can be derived by using $(n=1, p=p, \sigma=\sigma)$. The input image can by subsequently corrected by dividing each pixel channel-wise with the estimated illumination $\mathbf{I}^{est}$.

2.2. Graph-based learning

Graph-based approaches have been successfully used to solve several computer vision tasks such as image segmentation based on semantic information [14, 15, 16] and weighting pixels of an image based on their saliency estimated from the color and proximity in the image [12]. In this paper, we argue that a graph-based learning approach can also be used to find image regions with similar color properties and to weight their contribution in estimating the global illumination.

In graph-based learning, an image is represented by a graph *G = (V, E)*, where each node $v_i \in V$ is represented by a vector $\mathbf{x}_i$, e.g., color of a pixel or a super-pixel in the image, and an edge $(v_i, v_j) \in E$ connects nodes v_i and v_j. Each edge has a weight s_{ij}, which typically represents the similarity of nodes v_i and v_j. Usually, graph-based approaches rely on two matrices, a similarity matrix $\mathbf{S} = \{s_{ij}\}$ and a diagonal matrix $\mathbf{Q}$, which can be used either for regularization or to encode prior knowledge and constraints depending on the problem. In a segmentation task, where the main goal is to divide an image into different regions possessing similar semantic properties, the problem becomes equivalent to dividing the graph based on paths with a minimal combined cost, where the cost for each edge is defined by the corresponding similarity s_{ij}. The similarity is defined based on the desired properties for the given task [14, 16]. For example, in a saliency estimation task, such as in [12], the main goal is to estimate the saliency of each pixel or region. In [12], the similarity matrix $\mathbf{S}$ is constructed based on spatial and color-wise distances and $\mathbf{Q}$ encodes prior knowledge that pixels at the borders of the image are not likely to be salient.

3. PROBLEM FORMULATION

We assume that the illumination color is in the convex hull of the pixel colors present in the scene. As a result, the color constancy problem becomes equivalent to estimating the probability $p(\mathbf{x}_i)$ of each (super-)pixel $\mathbf{x}_i$, i.e., each node v_i of the graph *G*, to contain information on the global illumination. The probability $p(\mathbf{x}_i)$ is higher when the region $\mathbf{x}_i$ encodes more information about the illumination. Within this framework, we wish $\mathbf{p} = \{p(\mathbf{x}_1), ..., p(\mathbf{x}_n)\}$ to satisfy properties associated with the illumination assumptions. We make the assumption that pixels with similar color encode similar information about the illumination and thus they are expected to have similar probabilities. In addition, we assume that dark pixels, i.e., pixels with low luminance, encode less information about the light illumination and thus they should have lower probabilities. Therefore, the estimation of $\mathbf{p}$ can be formulated as an optimization problem with two terms. The first term enforces similar pixels to have similar probabilities and the second one encodes the prior knowledge and reflects the assumption that darker spots are usually less affected by the illumination. The joint optimization problem can be expressed as follows:

$$\begin{cases} \underset{\mathbf{p}}{\operatorname{argmin}} \sum_i p(\mathbf{x}_i)^2 q_i + \frac{1}{2}\sum_{i,j}(p(\mathbf{x}_i) - p(\mathbf{x}_j))^2 s_{ij}, \\ \text{subject to} \quad \sum_i p(\mathbf{x}_i) = 1. \end{cases} \tag{2}$$

The first term lessens the probability of a region $\mathbf{x}_i$ if there is prior information that this region does not encode information about the illumination, e.g., is dark. If q_i corresponding to $\mathbf{x}_i$ is high the system will tend to give a lower weight to $\mathbf{x}_i$ to ensure the minimization of the total sum. The weight s_{ij} encodes the similarity between two (super-)pixels $\mathbf{x}_i$ and $\mathbf{x}_j$ and thus if s_{ij} is high, the resulting probabilities $p(\mathbf{x}_i)$ and $p(\mathbf{x}_j)$ will be similar. $\sum_i p(\mathbf{x}_i) = 1$ is a normalization constraint ensuring that the result of the optimization problem (2) is a probability density function.

The above optimization problem can be framed as a graph-based learning problem with a similarity matrix formed by $\mathbf{S}_{ij} = s_{ij}$ and $\mathbf{Q} = diag(q_1, ..., q_n)$ and is similar to the problem defined in [12]. Thus, the graph-based formulation of the problem (2) can be expressed as follows:

$$\begin{cases} \underset{\mathbf{p}}{\operatorname{argmin}} \, \mathbf{p}^t\mathbf{H}\mathbf{p}, \\ \text{subject to} \quad \mathbf{p}^t\mathbf{1} = 1, \end{cases} \tag{3}$$

where $\mathbf{H} = \mathbf{D} - \mathbf{S} + \mathbf{Q}$ and $\mathbf{D}$ the graph degree matrix defined as $\mathbf{D}_{ii} = \sum_j \mathbf{S}_{ij}$. The optimization problem in (3) has one global optimum solution which can be obtained as

$$\mathbf{p}^* = \mathbf{H}^{-1}\mathbf{1}. \tag{4}$$

Having computed $\mathbf{p}$*, the illumination estimation can be computed as a weighted sum of the pixels color in the input image:

$$\mathbf{I}^{est} = \sum_i \mathbf{x}_i p^*(\mathbf{x}_i). \tag{5}$$

For the construction of $\mathbf{Q}$ and $\mathbf{S}$, we rely on the *Lab* color space. The *Lab* space is intuitive and appropriate in a color constancy context as it was designed to be perceptually uniform with respect to human color vision, meaning that the same amount of numerical change in these values corresponds to about the same amount of visually perceived change, and with respect to a given white point, it is device-independent. That is, it defines colors independent of how they are created or displayed. We propose two variants of $\mathbf{Q}$ in Eq. (6-7). The first, defined in Eq. (6), is a fixed threshold, while the second one, defined in Eq. (7), has a linear form which gives larger values to pixels with higher lightening:

$$\mathbf{Q}_1(\mathbf{x}_i) = \begin{cases} 0.1 \quad if \ L(\mathbf{x}_i) < q_0, \\ 0 \quad else \end{cases} \tag{6}$$

$$\mathbf{Q}_2(\mathbf{x}_i) = 0.1\frac{L_{max} - L(\mathbf{x}_i)}{L_{max} - L_{min}}. \tag{7}$$

$L(x)$ is the Lightness channel of *lab* color space. L_{max} and L_{min} are the maximum and the minimum values of $L(x)$ in the input image, respectively, and q_0 is a threshold value for $\mathbf{Q}_1$. The similarity function, used to calculate the elements of $\mathbf{S}$, reflects the assumption that pixels with similar colors are similarly affected by the illumination. It can be defined as follow:

$$S(\mathbf{x}_i, \mathbf{x}_j) = 1/(\epsilon + ||lab(\mathbf{x}_i) - lab(\mathbf{x}_j)||_2/\gamma), \tag{8}$$

where $lab(\mathbf{x}_i)$ is the color of the pixel $\mathbf{x}_i$ in the lab space, ϵ is a small constant, and γ is a scaling parameter. It should be noted that different forms of $\mathbf{Q}$ and $\mathbf{S}$ can be used and more sophisticated color constancy prior knowledge can be integrated in the future.

4. EXPERIMENTS AND DISCUSSION

We evaluated the proposed method on the recently released INTEL-TAU dataset [19]. The dataset contains 7022 images in total captured using three different camera models, Canon 5DSR, Nikon D810, and Sony IMX135. The dataset is the largest available high-resolution dataset for illumination estimation research. Due to the high dimension of the images (1080p), we abstract the image as (super)-pixels [17]. Each image is represented by 800 super-pixels. The parameter γ defined in Eq. (8) is set to 2. Another hyperparameter is the percentage of pixels used in the final estimate Eq. (5). We rely on 1% of the pixels with the highest probability values $\mathbf{p}^*(\mathbf{x}_i)$ computed using Eq. (4) to obtain the final estimate of the illumination. The threshold q_0 in Eq. (6) is set to 10% of L_{max}.

We report the mean of the top 25%, the mean, the median, Tukey's trimean, and the mean of the worst 25% of the Recovery angular error $e_{recovery}$ [20] and the 'Reproduction angular error' $e_{reproduction}$ [21] between the ground truth illuminant and the estimated illuminant:

$$e_{recovery}(\mathbf{I}^{gt}, \mathbf{I}^{est}) = \cos^{-1}(\frac{\mathbf{I}^{gt}\mathbf{I}^{est}}{\|\mathbf{I}^{gt}\|\|\mathbf{I}^{est}\|}) \tag{9}$$

$$e_{reproduction}(\mathbf{I}^{gt}, \mathbf{I}^{est}) = \cos^{-1}(\frac{\mathbf{I}^{gt}/\mathbf{I}^{est} \quad \mathbf{w}}{\|\mathbf{I}^{gt}/\mathbf{I}^{est}\|\sqrt{3}}), \tag{10}$$

where $\mathbf{I}^{gt}$ is the ground truth illumination for an image, $\mathbf{I}^{est}$ is the estimated illumination, / is the element-wise division operator, and $evaluated$ is defined as the unit vector, i.e., $\mathbf{w} = [1, 1, 1]^t$.

We compare the proposed method with the state-of-the-art unsupervised methods Grey-World [8], White-Patch [9], Grey-Edge variants [18], Shades-of-Grey [22], Weighted

Table 1. Results of benchmark methods on INTEL-TAU Dataset.

	Recovery angular error					Reproduction angular error				
Method	Best25%	Mean	Med.	Tri.	W.25%	Best25%	Mean	Med.	Tri.	W.25%
Grey-World [8]	1.0	4.9	3.9	4.1	10.5	1.2	6.1	4.9	5.2	13.0
White-Patch [9]	1.4	9.4	9.1	9.2	17.6	1.8	10.0	9.5	9.8	19.2
Grey-Edge [18]	1.0	5.9	4.0	4.6	13.8	1.2	6.8	4.9	5.5	13.5
2nd order Grey-Edge [18]	1.0	6.0	3.9	4.8	14.0	1.2	6.9	4.9	5.6	15.7
Shades-of-Grey [22]	0.9	5.2	3.8	4.3	11.9	1.1	6.3	4.7	5.1	13.9
Cheng et al. 2014 [23]	0.7	4.5	3.2	3.5	10.6	0.9	5.5	4.0	4.4	12.7
Weighted Grey-Edge [18]	0.8	6.1	3.7	4.6	15.1	1.1	6.9	4.5	5.4	16.5
Yang et al. 2015 [10]	0.6	3.2	2.2	2.4	7.6	0.7	4.1	2.7	3.1	9.6
Greyness Index 2019 [11]	0.5	3.9	2.3	2.7	9.8	0.6	4.9	3.0	3.5	12.1
Color Tiger [24]	1.0	4.2	2.6	3.2	9.9	1.1	5.3	3.3	4.1	12.7
PCC_Q1	0.6	4.1	2.6	3.0	10.1	0.7	5.3	3.7	4.2	12.3
PCC_Q2	0.6	3.9	2.4	2.8	9.6	0.7	5.1	3.5	4.0	11.9

Grey-Edge [18], Greyness Index [11], Color Tiger [24], and methods proposed in [23] and [10]. The performances of these methods are reported in Table 1. We also provide results for both variants of our approach, namely PCC_Q1 and PCC_Q2, obtained using the prior knowledge matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$, respectively.

In terms of recovery angular error, our method achieves better overall performance than Grey-World, White-Patch, Grey-Edge, second order Grey-Edge, Weighted Grey-Edge, method in [23], and Color Tiger. The method in [10] achieves the best overall results in terms of Recovery angular error statistics. In terms of Reproduction angular error, our method outperforms Color Tiger in the mean of Best 25% and worst 25% while Color Tiger yields a better median performance. Both Greyness Index and method in [10] outperform our approach mainly in terms of median and mean errors. We note that the linear second variant of $\mathbf{Q}$ outperforms the static first variant.

Figure 1 illustrates some visual results using the proposed approach. The middle column shows the weight masks obtained using the proposed approach. We see that depending on the scene characteristics, our approach generates different weight masks. It should be noted that using our approach one can generate local estimates per pixel by simply multiplying each pixel by its corresponding weight.

Fig. 1. Visual results on INTEL-TAU using our approach. From left to right, our input images from INTEL-TAU, the probability distribution mask obtained by PCC, our corrected image.

5. CONCLUSION

In this paper, a novel unsupervised color constancy approach is proposed. In this approach, we defined a probability-based method for estimating illumination of a scene using a graph-based representation of the image. We showed that the illumination estimation problem can be solved by estimating a probability density function defined over the image. To this end, we proposed two intuitive assumptions that enable us to define the graph connectivities. The proposed algorithm was tested on INTEL-TAU dataset, where we achieved comparable results to state-of-the-art methods. Future work includes proposing different variants of the unsupervised graph-based approach and incorporating other color constancy assumptions into the framework.

ACKNOWLEDGEMENT

This work was supported by the NSF-Business Finland Center for Visual and Decision Informatics (CVDI) project AMALIA 2019.

6. REFERENCES

[1] L. T. Maloney and B. A. Wandell, “Color constancy: a method for recovering surface spectral reflectance,” *JOSA A*, 1986.

[2] A. Gijsenij, T. Gevers, and J. Van De Weijer, “Computational color constancy: Survey and experiments,” *IEEE Transactions on Image Processing*, 2011.

[3] S. Bianco, C. Cusano, and R. Schettini, “Color constancy using CNNs,” in *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2015.

[4] Y. Hu, B. Wang, and S. Lin, “FC4: Fully convolutional color constancy with confidence-weighted pooling,” in *IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[5] F. Laakom, J. Raitoharju, A. Iosifidis, J. Nikkanen, and M. Gabbouj, “Color constancy convolutional autoencoder,” in *Symposium Series on Computational Intelligence*, 2019.

[6] F. Laakom, N. Passalis, J. Raitoharju, J. Nikkanen, A. Tefas, A. Iosifidis, and M. Gabbouj, “Bag of color features for color constancy,” 2019.

[7] A. Gijsenij, T. Gevers, and J. van de Weijer, “Computational color constancy: Survey and experiments,” *IEEE Transactions on Image Processing*, 2011.

[8] J. Cepeda-Negrete and R.E. Sanchez-Yanez, “Gray-world assumption on perceptual color spaces,” in *Image and Video Technology*, 2014.

[9] A. Rizzi, C. Gatta, and D. Marini, “Color correction between gray world and white patch,” in *The International Society for Optical Engineering*, 2002.

[10] K. Yang, S. Gao, and Y.J. Li, “Efficient illuminant estimation for color constancy using grey pixels,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015.

[11] Y. Qian, J.K. Kamarainen, J. Nikkanen, and J. Matas, “On finding gray pixels,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019.

[12] C. Aytekin, A. Iosifidis, and M. Gabbouj, “Probabilistic saliency estimation,” *Pattern Recognition*, 2018.

[13] J. Malcolm, Y. Rathi, and A. Tannenbaum, “Multi-object tracking through clutter using graph cuts,” in *2007 IEEE 11th International Conference on Computer Vision*, 2007.

[14] S. Vicente, V. Kolmogorov, and C. Rother, “Graph cut based image segmentation with connectivity priors,” in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, 2008.

[15] C. Aytekin, A. Iosifidis, S. Kiranyaz, and M. Gabbouj, “Learning graph affinities for spectral graph-based salient object detection,” *Pattern Recognition*, 2017.

[16] J. Shi and J. Malik, “Normalized cuts and image segmentation,” *Departmental Papers (CIS)*, 2000.

[17] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, “Slic superpixels,” Tech. Rep., 2010.

[18] J. van de Weijer, T. Gevers, and A. Gijsenij, “Edge-based color constancy,” *IEEE Transactions on Image Processing*, 2007.

[19] F. Laakom, J. Raitoharju, A. Iosifidis, J. Nikkanen, and M. Gabbouj, “Intel-tau: A color constancy dataset,” *arXiv preprint arXiv:1910.10404*, 2019.

[20] S. Hordley and G. Finlayson, “Re-evaluating colour constancy algorithms,” in *International Conference on Pattern Recognition*, 2004.

[21] G. Finlayson and R. Zakizadeh, “Reproduction angular error: An improved performance metric for illuminant estimation,” *perception*, 2014.

[22] G. Finlayson and E. Trezzi, “Shades of gray and colour constancy,” in *Color Imaging Conference*, 2004.

[23] D. Cheng, D. Prasad, and M. Brown, “Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution,” *Journal of the Optical Society of America. A, Optics, image science, and vision*, 2014.

[24] N. Banić, K. Koščević, and S. Lončarić, “Unsupervised learning for color constancy,” *arXiv preprint arXiv:1712.00436*, 2017.