# PREDICTING THE COLORS OF REFERENCE SURFACES FOR COLOR CONSTANCY

*Dubuisson I., Muselet D., Basso-Bert Y., Trémeau A.*[*]

Univ Lyon, UJM, CNRS, IOGS
Laboratoire Hubert Curien UMR 5516
F-42023, SAINT-ETIENNE, France

*Laganière R.*

School of Electrical Engineering
and Computer Science
University of Ottawa, Canada

**ABSTRACT**

The classical color constancy algorithms concentrate only on the color of a grey surface to estimate the light color and to white balance the image. In this paper, we show that the quality of the whole process can be clearly improved by predicting and correcting the colors of a set of reference surfaces. The ground truth of the surface color under white light can be easily obtained with a set of images acquired by the considered camera under sun light. Thus, we design a deep network to predict the colors of the reference surfaces of a color checker as if it had been in the scene at acquisition time. We show that our solution improves the two steps of the color constancy process on 9 datasets and we claim that being able to synthetically insert a color chart in any image can help for many other tasks.

***Index Terms*—** Color constancy, light color estimation, white balance, color calibration, color checker.

## 1. INTRODUCTION

Color constancy is the ability of the human visual system to perceive constant colors despite illumination condition variations. Computational color constancy tries to mimic this behaviour with computer vision systems. Given a digital image, the idea is first to estimate the color of the illumination and then to remove the impact of this illumination from the pixel colors [1, 2, 3, 4]. In this paper, we propose to improve these two steps with an original and simple solution. First, the light color estimation in the classical methods consists in predicting the color of a grey surface as if it had been observed by the current camera under the same conditions. By looking at the color formation model (see next sections), it is clear that many different light spectral power distributions can provide identical colors for a single grey patch. So we claim that predicting the color of a single grey patch under one illumination is not enough for accurate characterization of this light. Second, in the second step of the color constancy process, the most widely used transform is based on a simple $3 \times 3$ diagonal matrix for white balancing the image. This transform exploits the Von Kries hypothesis [5] but is known to be inaccurate in many situations [6]. This solution is widely used because the first step provides only the color of a single grey patch and the only way to correct the color of a grey patch is to apply independent gains on each channel. However there are many image datasets that contain a color checker with a set of colored reference surfaces, in addition to a grey surface.

**Contributions -** In this paper, we propose to exploit such available data and design a regression convolutional deep network that predicts the colors of 19 reference surfaces including a grey one. We show that predicting the colors of a set of reference surfaces already improves the prediction of the grey surface color by leveraging color correlations between reference surfaces under a set of different illumination conditions. Second, we exploit these 19 colors to improve the color correction of the image by using a full $3 \times 3$ matrix transform.

## 2. GREY-PATCH-BASED COLOR CONSTANCY

In order to show the interest of relying on a set of patch colors instead of only on a grey patch, we propose to measure the sensitivity of such colors across illumination variations. In this aim, we have used the spectral sensitivities of the sensor of the camera Canon 20D, the spectral reflectances of the 24 patches of the MacBeth color checker and the spectral power distributions (SPD) of two illuminations, namely the illuminant D65 and the LED Cool White (see Fig. 1). With these spectral curves, we have evaluated the (r,g) chromaticity coordinates of the 24 patches under these two illuminations. Despite the large differences between the two light SPD, we can observe on Fig. 1 that the chromaticities of the grey patches (numbers 19 to 24) under these two lights are overlapping. This means that the color of a grey surface can remain stable even in case of large illumination variations. Fortunately, we also notice that the other patch colors show significant chromaticity differences across light variations, represented as black lines on the bottom plot in Fig. 1. This means that

[*]This work was supported by the French government administered by the French National Research Agency (ANR) under the "Investissement d'Avenir" (ANR-11-IDEX-0007) program with the reference LABEX MANUTECH-SISE (ANR-10-LABX-0075) of Université de Lyon and the reference EUR MANUTECH SLEIGHT (ANR-17-EURE-0026) and by the Région Auvergne-Rhône-Alpes.

looking at the colors of a set of different patches is much more accurate to characterize the light in a color image. This is a toy example illustrating the limits of using a single grey patch as groundtruth for color constancy, but other experimental results are provided in Sec. 5.

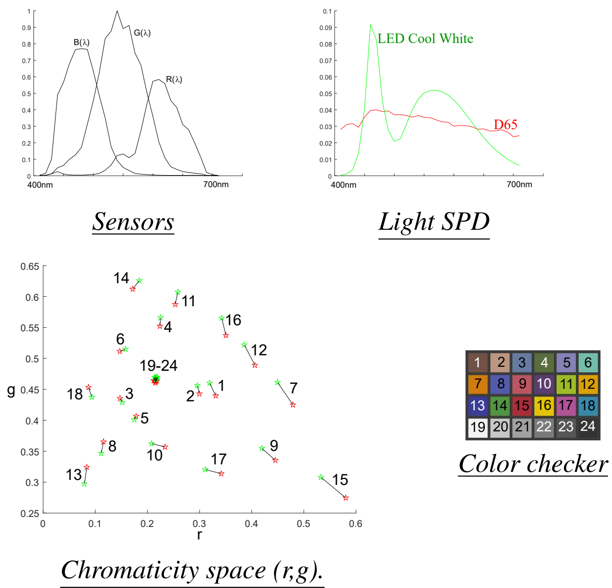


**Fig. 1**. Chromaticity variations across illumination conditions. The red (green, resp.) points correspond to the chromaticities of the color checker patches acquired by the displayed sensors under D65 (LED Cool White, resp.).

## 3. RELATED WORKS

**Light color estimation -** There exist three main categories of algorithms devoted to the color constancy problem. The solutions of the first category are exploiting the statistics of real color images and have all been unified in a general framework proposed by van de Weijer et al. [4]. They have been later extended with success in [7, 8, 9]. All these methods are based on empirical assumptions. The second category regroups the physics-based algorithms and mainly exploits the dichromatic reflection model from Shafer [10, 11, 12]. These methods require to detect grey surfaces [10], specularities [11] or to segment the images [12]. The third category contains all the learning-based approaches based on Gamut Mapping [13, 14], patch-based approaches [15] or the deep networks solutions [1, 2, 3, 16]. In this paper, we propose to take advantage and extend these last solutions so that they are able to predict the colors of a set of reference patches, and not only the color of a grey surface. From our knowledge, this is the first work that proposes to predict a set of colors to address the problem of color constancy.

**White balance -** After evaluating the light color, the next step consists in removing its impact on the colors of the pixels. In most of the color constancy approaches, this color correction relies on a simple diagonal $3 \times 3$ matrix, exploiting the von Kries hypothesis [5]. Some works have tried to improve this inaccurate transform [17, 5] but require to have access to the spectral sensitivities of the sensors [6]. Cheng et al. demonstrate that a full $3 \times 3$ matrix is much more accurate in many situations for color correction [6]. Nevertheless, they show that a diagonal transform provides good results in some cases and exploit this property to estimate the colors of reference patches under white light for a given camera. In our work, we exploit this groundtruth to estimate the quality of the whole color constancy process. Indeed, instead of only evaluating the rendering of a grey surface [18] as usually done by the color constancy approaches, we propose to analyze the rendering of a set of reference surfaces, which is much more complete.

## 4. OUR APPROACH

Considering a Lambertian surface characterized by a spectral reflectance $\beta(\lambda)$, illuminated by a light source with a SPD $E(\lambda)$ and observed by a camera whose sensor sensitivities are denoted $R(\lambda)$, $G(\lambda)$ and $B(\lambda)$, the output $C^k$ of the sensor $k \in \{R, G, B\}$ can be expressed as:

$$C^k = \alpha \int_\lambda E(\lambda)\beta(\lambda)k(\lambda)d\lambda, \tag{1}$$

where $\alpha$ is a constant across $k$ depending on the light irradiance and the orientation of the surface.

For the classical color constancy algorithms, the aim is to estimate the light color[1] whose component $E^k$ is expressed as: $E^k = \int_\lambda E(\lambda)k(\lambda)d\lambda$. This boils down to estimate the color component of a grey patch (with constant reflectance $\beta_{grey}$ over the whole spectrum): $E^k_{grey} = \int_\lambda E(\lambda)\beta_{grey}k(\lambda)d\lambda \propto E^k$.

As discussed and illustrated in the previous sections, we claim that both the estimate and the color correction are not accurate when using a single grey surface. Thus, we propose to characterize the light color of a set of surfaces characterized by diverse reflectance spectra, and not only a flat spectrum. This is a way to sample the light SPD at different wavelengths to enrich the extracted light characterization. Thus, in addition to $E^k_{grey}$, we propose to use the color components of a set of reference color surfaces $S_i$ illuminated by the current light:

$$E^k_{S_i} = \alpha \int_\lambda E(\lambda)\beta_{S_i}(\lambda)k(\lambda)d\lambda. \tag{2}$$

In total, the 19 reference patches (18 colored and 1 grey) from the MacBeth Color checker (see Fig. 1) are exploited. Our idea is to train a deep network to predict the colors of these 19 patches as if the color checker would have been in

[1]Since the light irradiance can't be estimated, these components are estimated up to a multiplicative constant.

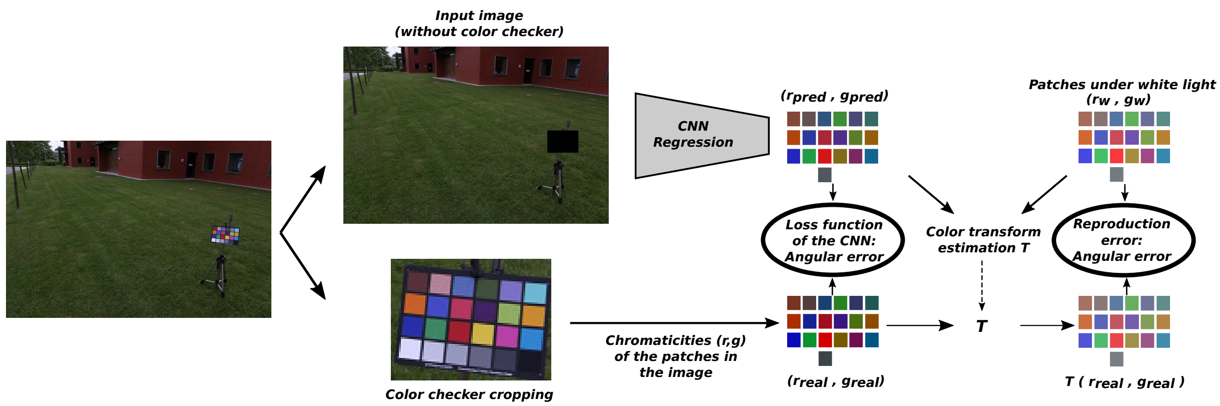


**Fig. 2**. The workflow of our method.

the scene while acquiring the image. We show in this paper that using these 19 predictions clearly improves the color constancy results. Obviously, having a model able to insert a synthetic color checker in any image acquired under uncontrolled conditions can help for many other tasks such as color calibration, mixed reality, light spectral estimation, ... out of the scope of this paper.

The workflow of our approach is illustrated in Fig. 2. Several datasets provide images with a color checker and we take advantage of this data to train our model. Given one image containing a color checker, we remove the color checker and extract its colors as groundtruth to evaluate the loss function used to train the network. Then, the color image without color checker is fed to the network. The model is trained to regress the chromaticity components $r = \frac{R}{R+G+B}$ and $g = \frac{G}{R+G+B}$ of the 19 patches of the color checker. Since the patches of the bottom row of the color checker are all grey, they have the same chromaticity coordinates, so we consider only 19 out of the 24 patches of the checker. Furthermore, the color coordinates $(R, G, B)$ of these patches depend on the orientation and location of the color checker as well as the light direction, position and irradiance, so we concentrate on the estimation of their chromaticity coordinates that are not sensitive to these parameters. When the regression network has been trained under the supervision of the real chromaticities $(r_{real}, g_{real})$, it is able to predict the chromaticities $(r_{pred}, g_{pred})$ of the 19 reference patches from any color image than does not contain any color checker.

The second step of a color constancy algorithm is the color correction whose aim is to remove the impact of the estimated light on all the colors of the image. For this step, the chromaticities of the reference patches under the target light (typically white) with the considered camera are required. For this data we take advantage of the solution proposed by Cheng et al. [6]. Their idea is very simple and does not require any complex or expensive device. It just exploits few images of the color checker acquired under the sun light with the considered camera. Such images are available in many color constancy datasets. In our experiments, we exploit the groundtruth provided by Cheng et al., i.e. the $(r_w, g_w)$ of the 19 patches under white light for each considered camera.

From the predicted chromaticities $(r_{pred}, g_{pred})$ and the groundtruth ones $(r_w, g_w)$ under white light, we move to the 3D chromaticity coordinates $(r, g, b)$, where $b = 1 - r - g$ in order to evaluate a full $3 \times 3$ matrix for the color correction. This is possible thanks to the color homography proposed by Finlayson et al. in [19]. Indeed, this color homography allows to evaluate a $3 \times 3$ matrix between two sets of 3D colors without looking at their brightness $(R + G + B)$ levels. In other words, evaluating such a color homography between the 3D chromaticities provides the same $3 \times 3$ transform as if we had worked with the $(R, G, B)$ color coordinates. We refer the reader to [19] for further details. With this approach, we evaluate the $3 \times 3$ matrix $T$ that best maps the predicted chromaticities $(r_{pred}, g_{pred}, b_{pred})$ to the ground truth $(r_w, g_w, b_w)$ in the least square sense.

This transform $T$ is then applied to each pixel of the initial color image in order to remove the impact of the light on their colors. Correcting the image with this transform renders this image under white light but the quality of the result is not easy to assess. So, for quality assessment, we only consider the real colors $(r_{real}, g_{real})$ of the 19 patches as they actually appeared in the original image and compare their transforms $T(r_{real}, g_{real})$ with their groundtruth values $(r_w, g_w)$. This quality criteria is very similar to the reproduction error introduced in [18] for grey patches and extended in [6] to any colored patch. This measure assesses the quality of the whole color constancy process, i.e. both the light color estimation and the color correction of the image.

## 5. EXPERIMENTS

**Datasets -** For our experiments, we use the data from two different color constancy datasets, namely the ColorChecker RECommended (CCREC) Dataset [20] and the NUS-8 Camera (NUS) Dataset [21]. The CCREC dataset contains 568

| Dataset | | Prediction of 19 patches (Our approach) | Prediction of 1 grey patch (Classical appr.) |
|---|---|---|---|
| Colorchecker | | **3.62 / 2.84** | 4.08 / 3.11 |
| NUS | Canon1D | **3.20 / 2.84** | 3.32 / 2.86 |
| | Canon600D | 4.05 / 3.62 | **3.27 / 2.76** |
| | Fujifilm | **3.83 / 3.35** | 4.74 / 3.59 |
| | NikonD40 | **7.22 / 4.37** | 10.61 / 6.05 |
| | NikonD5200 | **4.15 / 3.12** | 5.32 / 4.74 |
| | Olympus | **3.60 / 3.26** | 4.87 / 3.40 |
| | Panasonic | **4.07** / 3.48 | 6.94 / **3.38** |
| | Samsung | **3.92 / 2.94** | 4.91 / 3.87 |
| | Sony | 4.46 / 3.72 | **4.00 / 3.08** |

*a) Error for the grey patch only*

| Dataset | | Prediction of 19 patches (Our approach) | Prediction of 1 grey patch (Classical app.) |
|---|---|---|---|
| Colorchecker | | **3.26 / 2.43** | 3.69 / 3.01 |
| NUS | Canon1D | **7.79 / 6.02** | 10.22 / 10.02 |
| | Canon600D | 12.97 / **10.97** | **11.72** / 11.34 |
| | Fujifilm | **8.29 / 6.39** | 10.07 / 9.21 |
| | NikonD40 | **5.85 / 3.67** | 9.05 / 5.69 |
| | NikonD5200 | **3.63 / 2.72** | 4.70 / 4.20 |
| | Olympus | **7.93 / 6.49** | 9.87 / 9.52 |
| | Panasonic | **6.40 / 5.23** | 9.10 / 7.60 |
| | Samsung | **3.50 / 2.59** | 4.15 / 3.53 |
| | Sony | **7.50 / 5.99** | 7.71 / 7.18 |

*b) Average error for the 19 patches*

**Table 1**. Reproduction angular error (mean/median) for a) the grey patch and for b) the 19 patches (1 grey and 18 colored) on the REcommended ColoChecker Dataset (Colorchecker) and the NUS-8 cameras dataset.

indoor and outdoor images acquired by two cameras Canon 1D and canon 5D. The NUS dataset contains images acquired by 8 different cameras (around 250 images for each camera). Our model is trained and tested on each camera independently, so that we have 8 different image sets in the NUS benchmark. For each image set and following the classical approaches, we have used 3-fold cross validation to evaluate our model.

**Settings -** For our tests, we exploit the recent and accurate approach from [3] that is based on deep bag-of-features to predict the light color in one image. Thus, we create a deep network with 3 convolutional layers (60, 30 and 30 kernels, respectively), followed by a Bag-of-Features (BoF) pooling with 150 codewords and finally 3 dense layers (128, 64 and $N_c$ neurons, respectively). The neurons of the last layer have a sigmoid activation while the other ones have ReLU activations. The number $N_c$ of neurons in the last layer is equal to 2 if we try to predict the chromaticity coordinates $(r, g)$ of a grey patch (like the classical approaches) or equal to 38 if we try to predict the chromaticity coordinates of the 19 reference patches (our approach). This is the only difference in the architectures of the two tested networks.

We use classical data augmentations (small rotations, shifts, zoom and flip) and the Adam optimizer with 500 epochs. The learning rate for predicting our 38 chromaticity values is $1.10^{-4}$, while it is $5.10^{-4}$ for predicting 2 chromaticity values. These values have been cross validated on the CCREC dataset, so that they are optimal for each approach for fair comparison.

Our aim is not to reach the state-of-the-art results for the grey patch prediction, but rather to have a raw baseline in order to measure the improvement provided by adding the new 18 predictions. Compared with the experiments in [3], we don't apply specific data augmentations, we use much less number of epochs, no attention and we predict 2 chromaticity values for each patch instead of 3 color values. This explains why our results are different from those presented in [3]. But our aim is just to show that moving from 2 predicted values to 38 helps for color constancy, whatever the used baseline.

**Results -** As we can see in Table 1, our approach has two main advantages. First, it decreases the reproduction angular error of the grey patch for almost all the datasets (see Table 1 a)). Our intuition is that our regression approach is leveraging the correlation between the chromaticities of all the 19 patches while predicting each patch, and thus improves each individual prediction. This additional information is not available when predicting only the grey patch chromaticities. For space constraints, we can not add results, but the value of reproduction error for the grey patch would be ranked $2^{nd}$ out of 23 color constancy methods in [20] on the CCREC dataset. Second, and this is the main contribution of our work, the reproduction quality over the 19 reference surfaces is clearly better with our approach than with the grey patch prediction (see Table 1 b)). By predicting a set of colors, we are able to define an accurate $3 \times 3$ matrix that corrects the colors of the image to a white light condition. The classical color constancy approaches just concentrate on a single grey patch and consequently only uses a diagonal transform to normalize the colors.

## 6. CONCLUSION

In this paper, we have proposed a generalization of the classical color constancy approaches by predicting the colors of a set of reference surfaces instead of just a grey patch color. We have shown that adding these predictions help in both light estimation step and color correction. Obviously, being able to predict the colors of a colorchecker in any image without colorchecker at acquisition time can help in many other computer vision tasks that will be studied in future works.

## 7. REFERENCES


[1] Yi-Chen Lo, Chia-Che Chang, Hsuan-Chao Chiu, Yu-Hao Huang, Chia-Ping Chen, Yu-Lin Chang, and Kevin Jou, "Clcc: Contrastive learning for color constancy," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2021, pp. 8053–8063.

[2] Partha Das, Yang Liu, Sezer Karaoglu, and Theo Gevers, "Generative models for multi-illumination color constancy," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 1194–1203.

[3] Firas Laakom, Nikolaos Passalis, Jenni Raitoharju, Jarno Nikkanen, Anastasios Tefas, Alexandros Iosifidis, and Moncef Gabbouj, "Bag of color features for color constancy," *IEEE Transactions on Image Processing*, vol. 29, pp. 7722–7734, 2020.

[4] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE Transactions on Image Processing*, vol. 16, no. 9, pp. 2207–2214, Sept 2007.

[5] Hamilton Y. Chong, Steven J. Gortler, and Todd Zickler, "The von kries hypothesis and a basis for color constancy," in *2007 IEEE 11th International Conference on Computer Vision*, 2007, pp. 1–8.

[6] Dongliang Cheng, Brian Price, Scott Cohen, and Michael S. Brown, "Beyond white: Ground truth colors for color constancy correction," in *2015 IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 298–306.

[7] D. Cheng, D. K. Prasad, and M. S. Brown, "Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution," *J. Opt. Soc. Am. A*, vol. 31, no. 5, pp. 1049–1058, May 2014.

[8] S. Gao, W. Han, K. Yang, C. Li, and Y. Li, "Efficient color constancy with local surface reflectance statistics," in *European Conference on Computer Vision (ECCV)*, 2014, vol. 8690, pp. 158–173.

[9] K.-F. Yang, S.-B. Gao, and Y.-J. Li, "Efficient illuminant estimation for color constancy using grey pixels," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2015)*, 2015.

[10] Yanlin Qian, Jarno Nikkanen, J. Kämäräinen, and Jiri Matas, "On finding gray pixels," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 8054–8062, 2019.

[11] R. T. Tan, K. Nishino, and K. Ikeuchi, "Color constancy through inverse-intensity chromaticity space," *J. Opt. Soc. Am. A*, vol. 21, no. 3, pp. 321–334, Mar 2004.

[12] G. Finlayson and G. Schaefer, "Solving for colour constancy using a constrained dichromatic reflection model," *International Journal of Computer Vision*, vol. 42, no. 3, pp. 127–144, 2001.

[13] A. Gijsenij, T. Gevers, and J. van de Weijer, "Generalized gamut mapping using image derivative structures for color constancy," *International Journal of Computer Vision*, vol. 86, no. 2-3, pp. 127–139, 2010.

[14] G.D. Finlayson, "Corrected-moment illuminant estimation," in *IEEE International Conference on Computer Vision (ICCV)*, Dec 2013, pp. 1904–1911.

[15] H.R.V. Joze and M.S. Drew, "Exemplar-based color constancy and multiple illumination," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 5, pp. 860–873, May 2014.

[16] Damien Fourure, Rémi Emonet, Elisa Fromont, Damien Muselet, Alain Trémeau, and Christian Wolf, "Mixed pooling neural networks for color constancy," in *IEEE International Conference on Image Processing (ICIP 2016)*, 2016.

[17] G.D. Finlayson, M.S. Drew, and B.V. Funt, "Diagonal transforms suffice for color constancy," in *1993 (4th) International Conference on Computer Vision*, 1993, pp. 164–171.

[18] Graham D. Finlayson and Roshanak Zakizadeh, "Reproduction angular error: An improved performance metric for illuminant estimation," in *British Machine Vision Conference (BMVC)*, 2014.

[19] Graham D. Finlayson, Han Gong, and Robert B. Fisher, "Color homography: Theory and applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 1, pp. 20–33, 2019.

[20] Ghalia Hemrit, Graham D. Finlayson, Arjan Gijsenij, Peter V. Gehler, Simone Bianco, and Mark S. Drew, "Rehabilitating the color checker dataset for illuminant estimation," *ArXiv*, vol. abs/1805.12262, 2018.

[21] D. Cheng, D. K. Prasad, and M. S. Brown, "Illuminant estimation for color constancy: why spatial-domain methods work and the role of the color distribution," *J. Opt. Soc. Am. A*, vol. 31, no. 5, pp. 1049–1058, 2014.