COLOR CONSTANCY USING STAGE CLASSIFICATION

Rui Lu[1,2*], *Arjan Gijsenij*[2], *Theo Gevers*[2],
Koen van de Sande[2], *Jan-Mark Geusebroek*[2] *and De Xu*[1]

[1] School of Computer and Information Technology
Beijing Jiaotong University
{06112063, dxu}@bjtu.edu.cn

[2] Intelligent Systems Lab Amsterdam
University of Amsterdam
{a.gijsenij, th.gevers, ksande, geusebroek }@uva.nl

ABSTRACT

The aim of color constancy is to remove the effect of the color of the light source. Since color constancy is inherently an ill-posed problem, different assumptions have been proposed. Because existing color constancy algorithms are based on specific assumptions, none of them can be considered as universal. Therefore, how to select a proper algorithm for a given imaging configuration is an important question.

In this paper, image stage models are used to aid the selection of a specific color constancy algorithm. Image stages are $3D$ models of a scene. Based on stage classification, the most suitable color constancy algorithms is selected.

Experiments on large scale image datasets show that the proposed algorithm using stage classification outperforms state-of-the-art single color constancy algorithms with an improvement of almost 8%.

Index Terms— Color constancy, illuminant estimation, stage classification, SVM

1. INTRODUCTION

The color of objects is largely influenced by the color of the light source [1]. Therefore, the same object, taken by the same camera but under different illumination, may vary in its measured color appearance. This color variation may negatively affect the result of image and video processing methods for different applications such as image segmentation, object recognition and video retrieval. The aim of color constancy is to remove the effect of the color of the light source.

To this end, a considerable number of color constancy algorithms has been proposed, see [1, 2] for a review. Since color constancy is inherently an ill-posed problem, different assumptions have been proposed, such as the White-Patch assumption and the Grey-World assumption [2]. Because existing color constancy algorithms are based on specific assumptions, none of them can be considered as universal. Therefore, how to combine or select a proper color constancy for a given imaging configuration is an important research direction [2].

Recently, a number of color constancy methods have been proposed which take higher level visual information into account. For example, high-level visual information is used for color constancy in [3]. The image is modeled as a mixture of semantic classes, such as sky, grass, road and buildings. Illuminant estimation is steered by different classes by evaluating the likelihood of the semantic content. Similar to this, indoor-outdoor image information is used in [4]. Furthermore, image statistics are used in [5] to improve color constancy. It is shown that images with similar image statistics should be corrected by the same color constancy algorithms. Also, similar image statistics indicate a specific image category (i.e. scenes). For instance, the White-Patch color constancy algorithm is suitable for the forest category while the 1^{st}-order Grey-Edge is profitable for the street category. Hence, this indicates that there is a correlation between image statistics, scene types and color constancy algorithms.

In this paper, we take one step further in using high level information by considering depth information. It is shown that image statistics are influenced by depth patterns [6]. Furthermore, edge-based color constancy depends on the image statistics [5] caused by depth patterns. This relationship will be investigated in this paper. Moreover, the aim is to use image characteristics affected by depth information to achieve the selection of color constancy methods for a specific image.

This paper is organized as follows. First, in section 2, we briefly outline the color constancy framework. Then, the proposed method is described in section 3. The experimental setup and results are presented in section 4. Finally, section 5 concludes this paper.

2. COLOR CONSTANCY

Recently, a framework is proposed in [7]. It is based on the assumption that the average reflectance (and difference) in a image is achromatic. The framework is as follows:

$$\left(\int \left| \frac{\partial^n \mathbf{f}^\sigma(\mathbf{x})}{\partial \mathbf{x}^n} \right|^p d\mathbf{x} \right)^{\frac{1}{p}} = k\, \mathbf{e}^{n,p,\sigma}, \tag{1}$$

*The first author is funded in part by National Nature Science Foundation of China (60803072) and National Nature Science Foundation of China (90820013).

where n is the order of the derivative, p is the Minkowski-norm and $\mathbf{f}^\sigma(\mathbf{x}) = \mathbf{f} \otimes G^\sigma$ is the convolution of the image with a Gaussian filter with scale parameter σ. From (1), different color constancy algorithms can be derived by varying the parameter values. For the sake of simplicity, five most common algorithms, using different image (derivative) characteristics, are considered in this paper. This includes methods using pixels information, i.e., the Grey-World algorithm ($\mathbf{e}^{0,1,0}$), the Whilte-Patch algorithm ($\mathbf{e}^{0,\infty,0}$) and the general Grey-World algorithm ($\mathbf{e}^{0,13,2}$), using edges information ($\mathbf{e}^{1,1,6}$), and using higher-order statistics ($\mathbf{e}^{2,1,5}$).

3. COLOR CONSTANCY USING STAGE CLASSIFICATION

In this section, scene/stage models are introduced and the stage classification is described. Further, our color constancy method using stage classification is proposed.

3.1. Stage models

Spatial image structures are valuable visual cues for scene classification. Typical 3D scene geometries, called stages, are proposed in [6]. A few examples are shown in Figure 1. Each stage has a certain depth layout. These models depend on the inherent structure of natural images. In this paper, 13 different stages proposed by [6] are studied excluding $noDepth$ and $tab + pers + bkg$ as these stages are not available in the dataset under consideration.

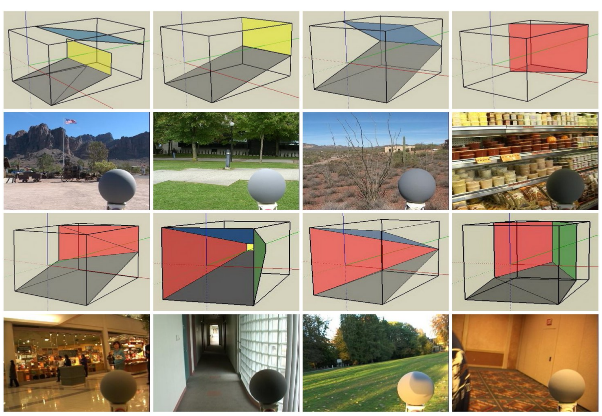


Fig. 1. Stage models and their corresponding examples: top two rows, from left to right: sky+bkg+gnd, bkg+gnd, sky+gnd, gnd+diagBkgLR; bottom two rows: diagBkgLR, box, 1side-wallLR, corner.

3.2. Stage classification

Bag-of-words representations have become a well established approach for scene classification. The method represents images by loose collections of independent patches [8]. First, a representative set of patches from the image is computed. Then, visual descriptors are generated for each patch. Finally, the resulting distribution of samples in descriptor space is used as the features for an image. SIFT features proposed by Lowe [9] are commonly adopted in bag-of-words representations. It describes the local shape of a region using edge orientation histograms. To increase the color stability and descriptive power among objects, color SIFT rather than intensity-based SIFT is preferred [10]. Therefore, in this paper, the RGB SIFT is used for stage classification. RGB SIFT is robust against changes in the intensity and color of the light source [10].

To this end, images are divided into 1×2, 2×2 and 1×3 regions by a spatial pyramid. For each region, dense sampling is used with a distance of 6 pixels. The RGB SIFT descriptors are computed for sample points [10]. Then, the descriptors are vector-quantized using a codebook. The quantized weights are 1 for the whole image, $1/4$ for 2×2 partitions and $1/3$ for 1×3 partitions. The feature vectors are the input of the stage classifier. A 1-*vs-all* SVM classifier with precomputed χ^2 kernel is used here, see [10] for details.

3.3. Stage based color constancy

After stage classification, the most suitable color constancy algorithm is selected for each stage by considering the angular error (see eq. 2) of the five different color constancy algorithms. Algorithms are applied on the training images of a specific stage. The algorithm with the lowest angular error is assigned to the stage under consideration. In this way, for each stage, the most proper color constancy algorithm is assigned. Note that this step is processed off-line. Then, the online processing is to predict which stage an (unknown) image belongs to by using the trained classifier. Finally, the color constancy algorithm that has been assigned to that stage will be used to estimate the light source to correct the image. This process is outlined in Figure 2.

4. EXPERIMENTS

4.1. Data set

Two independent datasets are used in the experiments. A large dataset D_1 with more than 3,500 images is used to train the stage classifier [6]. The other dataset D_2, collected by F. Ciurea et.al [11], is used for testing. Note that D_2 contains the ground truth for color constancy as a grey ball is present in each image. No ground truth is available for D_1. Therefore, color constancy is evaluated on D_2. D_1 is used to provide an independent dataset for training ensuring generalization of the proposed method. Dataset D_2 consists of 11,346 images, extracted from 15 different video clips taken at different places and hence lighting conditions. Note that the images in D_2 are not linear, as they were gamma-corrected (with unknown value for gamma) after capturing. Since the value of gamma is unknown, images are not corrected before estimating the

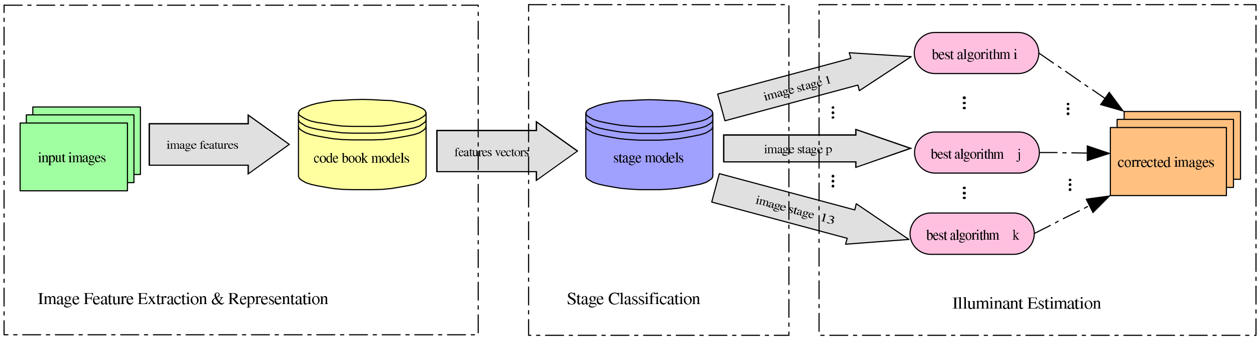


Fig. 2. Outline of the stage-based classification for color constancy.

illuminants. As there exists a high correlation among images of the same video clip, we test the color constancy algorithms on a subset of uncorrelated images (manually selected) composed of 711 images, see Figure 1 for example. Note that the grey ball is masked when the illuminant is estimated.

4.2. Performance measure

In order to evaluate the performance of the color constancy algorithms, the angular error ε is used:

$$\varepsilon = \cos^{-1}(\hat{\mathbf{e}}_\mathbf{l} \cdot \hat{\mathbf{e}}_\mathbf{e}), \qquad (2)$$

where $\hat{\mathbf{e}}_\mathbf{l}$ is the normalized ground truth of the illuminant while $\hat{\mathbf{e}}_\mathbf{e}$ is the normalized estimation. Since the distribution of errors, in general, is non-Gaussian (e.g. skewness), the mean and standard deviation values are not suitable for summarizing the errors. Consequently, the median angular error is reported [12].

4.3. Experimental results

Stage classification. All images used in the experiment are manually annotated. Because images in D_2 are not uniformly distributed in each stage, the average precision(AP) is taken to reduce the influence of data imbalance. The average precision is a single-valued measure that is proportional to the area under the precision-recall curve. The stage classifier results are shown in Table 1, together with the relative occurrence of each stage in the dataset.

From Table 1, it can be derived that for a number of stages, such as *sky+bkg+gnd*, and *gnd*, proper results are obtained. Others, like *diagBkgLR* and *diagBkgRL*, the stage classification performs less satisfying. This is due to the many occlusions appearing in these categories.

Illuminant estimation. Only images in dataset D_2 (with groundtruth) are used for illuminant estimation. A leave-one-out method is adopted. The results of the proposed method (*SBCC(auto)*), together with all of the five algorithms are shown in Figure 3 while the results on the entire data set are given in Table 2. To determine how the performance

Name	**% in dataset**	**AP**
skyBkgGnd	9.1%	0.65
bkgGnd	9.9%	0.34
skyGnd	2.7%	0.34
gnd	12.1%	0.67
gndDiagBkgLR	6.6%	0.16
gndDiagBkgRL	4.6%	0.16
diagBkgLR	4.6%	0.12
diagBkgRL	3.8%	0.15
box	8.0%	0.37
1side-wallLR	12.9%	0.46
1side-wallRL	15.6%	0.41
corner	6.5%	0.15
persBkg	3.5%	0.19
MAP	0.32	

Table 1. Relative occurrence and the mean average precisions for each stage under consideration within dataset D_2. The last row gives the mean average precision over all stages.

Method	**Median**
Grey-World	6.9°
White-Patch	7.1°
general Grey-World	5.8°
1^{st}-order Grey-Edge	5.2°
2^{nd}-order Grey-Edge	5.4°
SBCC(auto)	4.8° (-8%)
SBCC(manual)	4.6° (-12%)

Table 2. Median angulars on the entire dataset D_2.

of the classifier influences color constancy performance, we have also applied the proposed method on a "perfect" stage classifier by manually assigning the images to the right stage denoted by *SBCC(manual)*.

It can be derived from Figure 3, that the proposed method achieves improved results compared to all the other color constancy algorithms on almost all stages. Note that the proposed method maps 13 stages to just 5 color constancy algorithms. Hence, a misclassification can still result in applying the cor-

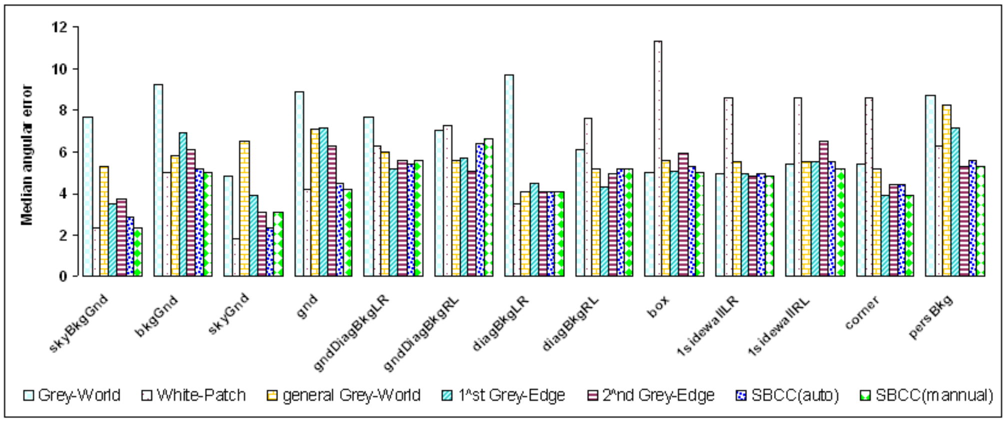
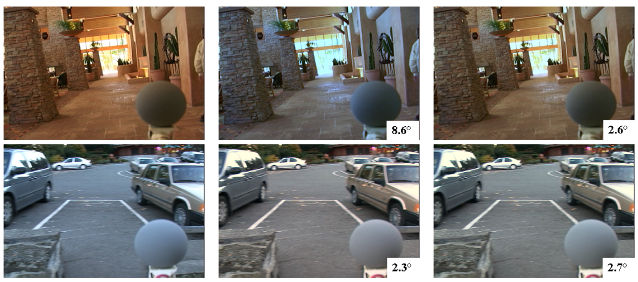


Fig. 3. Median angular error for the different color constancy algorithms for each stage.



(a) original image (b) GW Correction (c) SBCC Correction

Fig. 4. Results for the Grey-World algorithm (GW) and the proposed algorithm (SBCC (auto)). The image at the top row is correctly classified while the one at the bottom is misclassified due to the occlusion. The angular errors are shown in the right bottom corner.

rect algorithm, as the misclassified stage can correspond to the same method as the correctly classified stage. This accounts for the good performance of automatic classification, with respect to manual classification. In summary, Table 2 shows that the proposed method outperforms the five single color constancy methods. Using the proposed algorithm lead to an improvement of nearly 8% over the best-performing single algorithm, i.e., 1^{st}-order Grey-Edge. Two specific examples are shown in Figure 4.

5. CONCLUSIONS

Although there are many color constancy algorithms available, none of them is universal. Therefore, in this paper, stage models have been used to select a proper algorithm for a given image. Experiments demonstrates that the proposed color constancy algorithm outperforms state-of-the-art single color constancy algorithms with an improvement of nearly 8%. Ideally, the proposed algorithm can make an improvement of more than 12% over the best single algorithm.

6. REFERENCES

[1] Marc Ebner, *Color constancy*, Wiley, 2007.

[2] G. Hordley, "Scene illuminant estimation:past, present, and future," *CRA*, vol. 31, no. 4, pp. 303–314, 2006.

[3] J. van de Weijer, C. Schmid, , and J. Verbeek, "Using high-level visual information for color constancy," in *ICCV*, Oct. 2007, pp. 1–8.

[4] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, "Improving color constancy using indoor-outdoor image classification," *IEEE Trans. on Image Processing*, vol. 17, no. 12, pp. 2381–92, 2008.

[5] A. Gijsenij and Th. Gevers, "Color constancy using natural image statistics," in *CVPR*, June 2007, pp. 1–8.

[6] V. Nedovic, A. Smeulders, A. Redert, and J.-M. Geusebroek, "Depth information by stage classification," in *ICCV*, Oct. 2007, pp. 1–8.

[7] J. van de Weijer, Th. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE TIP*, vol. 16, no. 9, pp. 2207–2214, 2007.

[8] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-of-features image classification," in *ECCV*, 2006, pp. 490–503.

[9] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *IJCV*, vol. 60, pp. 91–110, 2004.

[10] K. van de Sande, Th. Gevers, and C. Snoek, "Evaluation of color descriptors for object and scene recognition," in *CVPR*, June 2008, pp. 1–8.

[11] F. Ciurea and B. Funt, "A large image database for color constancy research," in *CIC*, July 2003, pp. 160–164.

[12] S.D. Hordley and G.D. Finlayson, "Reevaluation of color constancy algorithm performance," *Optical Society of America*, vol. 23, no. 5, pp. 1008–1020, 2006.