



end, 3D scene geometry models are used to divide images into different geometrical regions. These models will be used to select appropriate color constancy algorithms per depth layer or geometrical section of the model.

This paper is organized as follows. First, in section 2, we briefly outline the color constancy framework, and discuss how to infer 3D scene geometry estimation from a single image. Then, the proposed method is described in section 3. After that, the experimental setup and results are presented in section 4. Finally, section 5 concludes this paper.

## 2. Preliminaries

In this section, color constancy and 3D scene geometry estimation is discussed in more detail.

### 2.1. Color constancy

Assuming Lambertian reflection, the image color  $\mathbf{f} = (R, G, B)^T$  recorded by electronic devices is dependent on the color of the light source  $e(\lambda)$ , the surface reflection  $s(\mathbf{x}, \lambda)$  and the camera sensitivity function  $\mathbf{c}(\lambda)$ :

$$\mathbf{f}(\mathbf{x}) = \int_{\omega} e(\lambda) \mathbf{c}(\lambda) s(\mathbf{x}, \lambda) d\lambda, \quad (1)$$

where  $\omega$  is the visible spectrum,  $\lambda$  is the wavelength of the light and  $\mathbf{x}$  is the spatial coordinate. Under the assumption that the observed color of the light source  $\mathbf{e}$  depends on the color of the light source  $e(\lambda)$  and the camera sensitivity function  $\mathbf{c}(\lambda)$ , the color of the light source is estimated by

$$\mathbf{e} = \int_{\omega} e(\lambda) \mathbf{c}(\lambda) d\lambda. \quad (2)$$

Color constancy is an under-constrained problem, since both  $e(\lambda)$  and  $\mathbf{c}(\lambda)$  are generally unknown. Therefore, in order to solve the color constancy problem, a number of assumptions are proposed such as the Grey-World assumption (*i.e.* the average pixel value is grey) [7] and the White-Patch assumption (*i.e.* the maximum pixel value is white) [7].

Recently, color constancy is addressed by exploiting higher order derivative information [19],

$$\left( \int \left| \frac{\partial^n \mathbf{f}^\sigma(\mathbf{x})}{\partial \mathbf{x}^n} \right|^p d\mathbf{x} \right)^{\frac{1}{p}} = k \mathbf{e}^{n,p,\sigma}, \quad (3)$$

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  $\sigma$ . Using Equation 3, different color constancy algorithms can be derived by varying one or more of parameter values. For example,

1. when  $n=0$ , pixel-based color constancy algorithms are obtained, such as the Grey-World algorithm ( $\mathbf{e}^{0,1,0}$ ), the White-Patch algorithm ( $\mathbf{e}^{0,-1,0}$ ) and the general Grey-World ( $\mathbf{e}^{0,13,2}$ );

2. when  $n=1$ , color constancy algorithms are obtained using first-order derivative information, *i.e.* image edges information. The Minkowski-norm  $p$  and smoothing parameter  $\sigma$  depend on specific dataset. The instantiation  $\mathbf{e}^{1,1,6}$  is applied in the paper;
3. when  $n=2$ , the framework provides color constancy methods based on second-order statistics. Similarly, the other two parameters  $p$  and  $\sigma$  vary with the dataset. We use  $\mathbf{e}^{2,1,5}$  in our experiments.

Although other color constancy methods can be used, we focus on the above instantiations which include pixel and derivative-based methods.

### 2.2. Stage models and segmentation

A number of methods have been proposed to estimate the geometry from single images [6, 16], where the work of Hoiem *et al.* [11] is the most influential. However, these methods are restricted to a number of classes restricting the applicability of the method. More general classes might be more useful instead of particular scene instantiations. To this end, we introduce an image classification scheme which derives scene information for a wider range of scene categories by using scene geometries.

Typical 3D scene geometries, called stages, are proposed in [15]. The authors infer 3D geometry for a wide range of stages. Image stages are considered as 3D geometrical models of the scene. A number of stage models, together with corresponding examples are shown in Figure 1. Each stage has a certain depth layout. These models are dependent on the inherent structure of natural images, resulting from image statistics and viewpoint constraints. In this paper, 13 different stages proposed by [15] are studied excluding *noDepth* or *tab+pers+bkg* as these stages are specific characteristics for the dataset used in [15], but not available in the dataset under consideration.

As shown in Figure 1, the 3D depth structures of the stage models are demonstrated in different colors. Each stage has a unique depth pattern. The stage models, *i.e.*, 3D geometry structures, are used to determine how the image should be divided. For instance, images of the stage *sky+background+ground* should be divided into three parts: sky (in blue), background (in yellow) and ground (in brown). By contrast, if the image structure should be treated as a whole, no segmentation is applied. For instance, there is no need to segment images of stage *ground*, *diagonal-BackgroundLR*, *diagonalBackgroundRL*, because they are geometrically composed of only one plane.

## 3. Color constancy using 3D scene geometry

The proposed method using 3D scene geometry for color constancy consists of the following steps: first, images are classified into stages; then, according to the stage models,

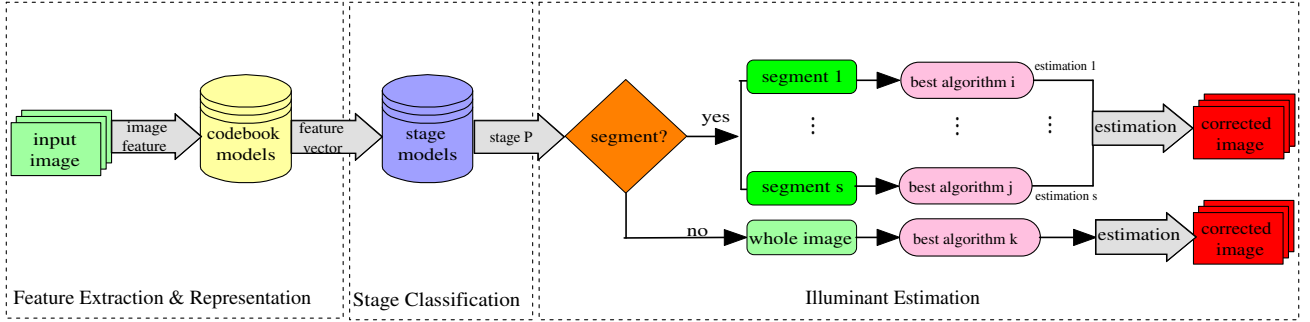


Figure 2. Outline of color constancy using 3D scene geometry. Note that the codebook models and the stage models are obtained off-line. Stage segmentation and best color constancy algorithm selection are trained on the dataset beforehand.

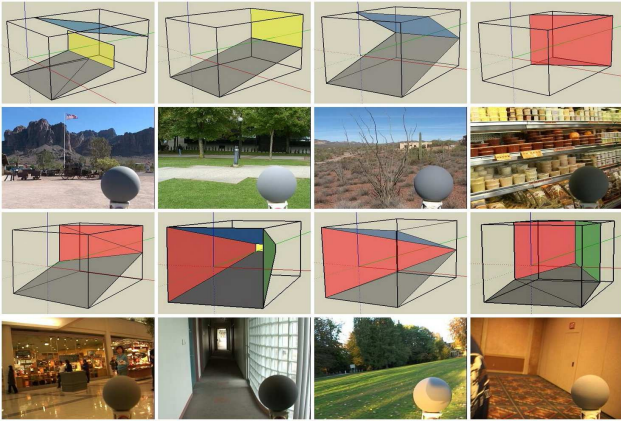


Figure 1. Stage models and their corresponding examples: top two rows, from left to right: sky+background+ground, background+ground, sky+ground, ground+diagonalBackgroundLR; bottom two rows: diagonalBackgroundLR, box, 1side-wallLR, corner.

images are partitioned into segments. Each segment corresponds to a specific geometrical structure in the image. Consecutively, for each segment, the most suitable color constancy algorithm is used. The final result for the whole image is the combination (*e.g.* weighted average) of the estimations for each segmentation. In the case of a single segment (*i.e.* whole image), illumination estimation will be computed using the best color constancy algorithm for that stage type. The whole process is demonstrated by the block diagram shown in Figure 2.

### 3.1. Stage classification

Bag-of-words representations have become a well established approach for scene classification. The basic idea is to treat images as loose collections of independent patches. First, a representative set of patches is sampled from the image. Then, a visual descriptor for each patch is generated. Finally, the resulting distribution of samples in the descrip-

tor space is used for classification. For stage classification, we closely follow the scene classification system as proposed by Van de Sande et al. [18]. The method is based on a color modification [2] of the SIFT descriptor [14]. Specifically, for our experiments, we use a color constant RGB SIFT descriptor, evaluated in [18] as a good candidate for scene categorization. Furthermore, 1-vs-all SVM classifiers with precomputed  $\chi^2$  kernel are used, see [4, 18] for more details.

### 3.2. Segmentation

Image stages are dependent on the inherent structure of natural images, resulting from image statistics and view-point constraints [15]. Each stage has a unique partitioning, reflecting the 3D geometry of the underlying image. Each segment in the partitioning represents geometrical entities like walls, ground, and sky. We will use the segmentation provided by the stage model to learn which color constancy algorithm performs best for each segment. Both hard segments and soft segments is considered, the latter taking the uncertainty due to the rough outline of the stage geometry into account. Both of them are based on the occurrence probability in the training set. Ground truth is obtained by manually annotation, thereby dividing the training set according to the stage patterns, and fitting the parameters of each stage model (horizon, vanishing points) such as to visually best fit the underlying data.

Suppose that an image belongs to Stage  $S$ , which is composed of  $N$  parts, correspondingly there will be  $N$  mask maps. The mask map for the  $i^{th}$  partition  $MP_i$  is obtained by taking an average of the mask maps for each image:

$$MP_i(\mathbf{x}) = \frac{\sum_{j=1}^n M_{j,i}(\mathbf{x})}{n}, \quad (4)$$

where  $n$  is the total number of images in the training dataset, and  $M_{j,i}(\mathbf{x})$  is the mask map of the  $j^{th}$  image for the  $i^{th}$  partition. Note that  $M_{j,i}(\mathbf{x})$  is an indicator function:  $M_{j,i}(\mathbf{x}) = 1$ , if  $\mathbf{x}$  belongs to the  $i^{th}$  partition and 0 otherwise.

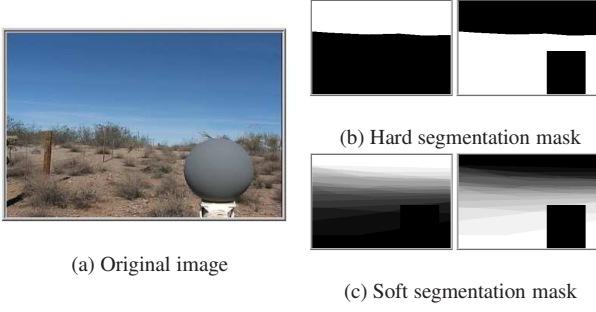


Figure 3. An example of hard and soft segmentation mask maps. The original image (a) belongs to stage *sky+ground*. The mask maps are of the same size as the original image. The difference between hard segmentation mask map, shown in figure (b), and the soft segmentation mask map, shown in figure (c), is that values in figure (b) are either 0 or 1 while values in figure (c) between 0 and 1.

### 3.2.1 Hard segmentation

Gijsenij *et al.* [9] show that images with different natural statistical distributions may prefer different color constancy algorithms. Therefore, mask maps are used to automatically divide the images. Assuming that the images of a stage can be partitioned into  $N$  parts, there exist  $N$  mask maps corresponding to the partitions in the training dataset. Then, the binary mask map is defined as follows:

$$MP'_i(\mathbf{x}) = \begin{cases} 1, & MP_i(\mathbf{x}) = \max_{j=1}^N MP_j(\mathbf{x}), \\ 0, & otherwise. \end{cases} \quad (5)$$

As a consequence, the values in the hard mask map are either 0 or 1, as shown in Figure 3(b).

After the maximum mask map is obtained, the color of the light source is estimated using pixels from one area while other pixels are ignored.

### 3.2.2 Soft segmentation

As the scene geometry per stage category is only roughly outlined, some pixels are more reliable than others to belong to a certain segment. Therefore, different confidence values are assigned to them. We set the confidence values of pixels to  $MP_i(\mathbf{x})$ , which indicates the occurrence frequency of pixels appearing in the training dataset. For example, given 100 training images, for a specific position  $\mathbf{x}$ , 80 training images indicate that this position belongs to the area  $A_i$ , then  $MP_i(\mathbf{x}) = 0.8$ . Hence, the value of  $MP_i(\mathbf{x})$  corresponds to a weight for pixels at this position in images of that stage. In particular, given a position, the larger the confidence value, the more probable it belongs to that partition. Note that the values of the soft segmentation mask map are between 0 and 1, as shown in Figure 3(c).

### 3.2.3 Illuminant estimation

**Without segmentation.** After stage classification, the most suitable color constancy algorithm is selected for each stage by considering the angular errors 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 on-line 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.

**With segmentation.** After the mask map has been obtained by hard or soft segmentation, images in the training dataset will be divided into several parts. The most suitable color constancy algorithm will be selected from the existing color constancy algorithms for each segment. This is achieved by analyzing the angular errors for all color constancy algorithms on images of this stage in the training dataset (see section 4.2 for details). The color constancy algorithm with the lowest angular error is assigned to this segment of the stage. In other words, the stage model is labeled with the segments and their corresponding color constancy algorithms which provide highest color constancy accuracy (*i.e.* lowest angular error). The weight of each segment is inverse to its angular error.

For each unseen image in the test dataset, the on-line process is as follows: first, it is classified into a specific stage  $S$ ; then it is divided according to the mask maps of stage  $S$ , which have been obtained in the training dataset beforehand. Further, the color of the light source is estimated in each segment using its selected color constancy algorithm. Finally, the illuminant estimation is obtained by combining the results of each segment by weighted average.

## 4. Experiments

In this section, the proposed method is evaluated on a real-world dataset and compared with the state-of-the-art color constancy algorithms on a large-scale dataset.

### 4.1. Data set

Two independent datasets are used in the experiments. In order to obtain a classifier with good generalization capabilities, a large dataset  $D_1$  with more than 3,500 images is used to train the stage classifiers [15]. The other dataset  $D_2$ , collected by Ciurea *et al.* [5] is used for testing. Note that  $D_2$  contains the ground truth for color constancy while no ground truth is available for  $D_1$ . Therefore, color constancy is evaluated on  $D_2$ .  $D_1$  is used to provide an inde-

pendent dataset for stage classification ensuring generalization of the proposed method.

Dataset  $D_2$  consists of more than 11,000 images, extracted from 2 hours video for a wide variety of settings such as indoor, outdoor, desert, cityscape, etc. There are totally 15 different video clips taken under different places and hence lighting conditions. As there exists high correlation among images of the same video clip, we test the color constancy algorithms on a subset of uncorrelated images composed of 711 images. These images are manually selected and annotated. A few examples of the dataset are shown in Figure 1. In each image, there is a grey ball at the right bottom, which is used to capture the ground truth of the light source. Note that the grey ball is masked when the illuminant is estimated.

## 4.2. Performance measure

Two performance measures are used in this paper: stage classification is evaluated using the average precision, while the angular error is used to validate the performance of the color constancy algorithms.

**Average precision.** The average precision is equivalent to the area under a precision-recall curve. It combines recall and precision in a single number. Mean average precision MAP is used to evaluate the performance of the features over all the stages, which is obtained by averaging the average precisions over all stages.

**Angular error.** In order to evaluate the performance of the color constancy algorithms, the angular error  $\varepsilon$  is used,

$$\varepsilon = \cos^{-1}(\hat{\mathbf{e}}_l \cdot \hat{\mathbf{e}}_e) , \quad (6)$$

where  $\hat{\mathbf{e}}_l$  is the normalized ground truth of the illuminant, while  $\hat{\mathbf{e}}_e$  is the normalized estimation. Both mean and median angular errors [13] are taken as performance indicator.

## 4.3. Stage classification

For the purpose of stage classification, we use generic 1-*vs-all*-based classifiers. Each classifier uses the RGB-SIFT feature vectors which are quantized by codebooks and outputs a single stage label, as this feature is shown to outperform other variants of the SIFT-feature [18]. There is a total of 13 classifiers corresponding to 13 stages concerned.

The performance of stage classification on each stage is shown in Table 1. From this table, it can be derived that for some stages, such as *sky+bkg+gnd*, and *gnd*, the results are satisfying. For other stages, like *diagBkgLR* and *diagBkgRL*, the results still leave room for improvement. This is due to occlusions appearing in these categories, making it hard to classify them correctly. A few examples of misclassified images are given in Figure 4.

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
<b>MAP</b>	<b>0.320</b>	

Table 1. Stage classification results for each stage using the RGB-SIFT feature, as well as relative occurrence within  $D_2$ . Note that the last row gives the mean average precision over all stages.



Figure 4. Examples of misclassified images.

## 4.4. Single color constancy methods

The algorithms that are evaluated here are the five instantiations discussed in Section 2. The results for single algorithms are shown in Table 2. These methods are applied to each image in  $D_2$ . Table 2 shows that the edge-based methods (*i.e.* 1<sup>st</sup>-order Grey-Edge and 2<sup>nd</sup>-order Grey-Edge) outperform the pixel-based methods (*i.e.* Grey-World, White-Patch, and general Grey-World). Differences between 1<sup>st</sup>-order Grey-Edge and 2<sup>nd</sup>-order Grey-Edge are small, but the median angular error of the 1<sup>st</sup>-order Grey-Edge is slightly lower, so the 1<sup>st</sup>-order Grey-Edge is considered as our *baseline* in the remainder of this section.

## 4.5. Combination algorithms

In addition to the single algorithms, two combination algorithms are evaluated that using combination strategies to obtain a more accurate illuminant estimate. The first method is proposed by Bianco et al. [1] and distinguishes between indoor and outdoor images. An indoor-outdoor

Method	Mean	Median
Grey-World	7.4°	7.0°
White-Patch	7.3°	6.1°
general Grey-World	6.4°	5.8°
<b>1<sup>st</sup>-order Grey-Edge</b>	<b>6.0°</b>	<b>5.2°</b>
2 <sup>nd</sup> -order Grey-Edge	6.0°	5.4°
Combination using indoor-outdoor classification	7.0° (+17%)	6.5° (+25%)
Combination using natural image statistics	5.7° (-5%)	4.7° (-10%)
Proposed (auto): without segmentation	5.7° (-5%)	4.8° (-8%)
<b>Proposed (auto): hard segmentation</b>	<b>5.4° (-10%)</b>	<b>4.5° (-14%)</b>
Proposed (auto): soft segmentation	5.4° (-10%)	4.6° (-12%)
Proposed (manual): without segmentation	5.5° (-8%)	4.6° (-12%)
Proposed (manual): hard segmentation	4.7° (-22%)	3.7° (-29%)
<b>Proposed (manual): soft segmentation</b>	<b>4.7° (-22%)</b>	<b>3.6° (-31%)</b>

Table 2. Performance of color constancy algorithms over  $D_2$ . Proposed (auto) means that the proposed methods are applied to automatically classified images while proposed (manual) indicates that our methods are evaluated on manually classified images.

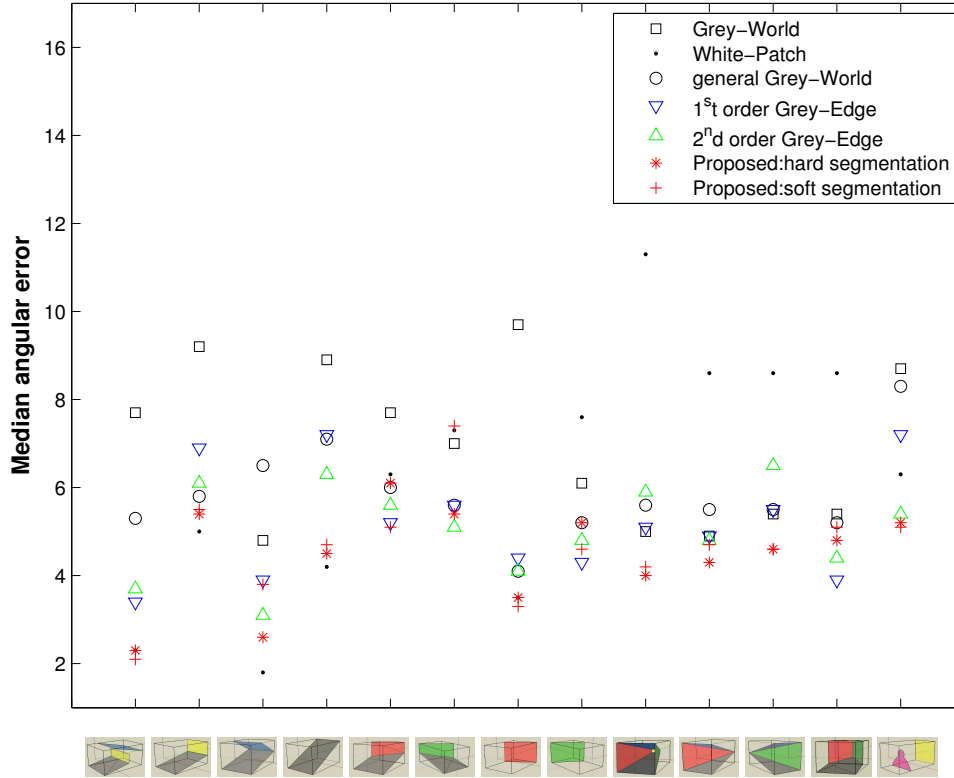


Figure 5. Median angular errors of color constancy algorithms for each stage in  $D_2$ . The stage models are shown on the horizontal axis: sky+background+ground, background+ground, sky+ground, ground, ground+diagonalBackgroundLR, ground+diagonalBackgroundRL, diagonalBackgroundLR, diagonalBackgroundRL, box, 1sidewallLR, 1sidewallRL, corner, person+Background.

classifier is proposed that is used to apply different color constancy algorithms to indoor and outdoor images. In spe-

cific, they propose to use the shades-of-grey method for indoor images and the 2<sup>nd</sup>-order Grey-Edge method for out-

door images. For convenience, we used manual annotation of indoor and outdoor images instead of the indoor-outdoor classifier proposed in [1]. As can be seen in table 2, the accuracy of the illuminant estimates does not improve with respect to the single color constancy algorithms.

Another combination method that is evaluated in this paper is proposed by Gijsenij and Gevers [9], and use global image statistics for selection of the most appropriate color constancy algorithm. Results indicate that the performance indeed improves with respect to the best-performing single algorithm, see table 2. However, as this method uses global selection of the most appropriate color constancy algorithm, the scene geometry is not taken into account, so there is little room for further improvement of the results using this approach.

#### 4.6. Color constancy using stage classification

Evaluation of the proposed methods is performed using the leave-one-out cross validation method. In order to obtain the mask map, all images in the training dataset are manually annotated and segmented.

**Without segmentation** Illuminant estimation is computed using the entire image. The median angular error of the proposed method without segmentation is  $4.8^\circ$  as shown in Table 2. Compared with the best-performing algorithm, *i.e.*, the 1<sup>st</sup>-order Grey-Edge, an increase of almost 8% on the median angular error is reached. Results on individual stages reveal that most color constancy algorithms have a preference for specific stages. For instance, 0<sup>th</sup>-order methods like the White-Patch and the general Grey-World prefer stages where the depth can become quite high, like the stage *sky+background+ground*. Such stages with a large depth can contain haze, which causes a relatively low signal-to-noise ratio, and it is known from [9] that methods that are based on higher-order statistics like the 2<sup>nd</sup>-order Grey-Edge do not perform well on such images. On the other hand, the 2<sup>nd</sup>-order Grey-Edge algorithm performs better on images with a high amount of information, *e.g.* many edges. This is reflected in a preference for stages like *diagonal BackgroundLR* and *diagonal BackgroundRL* that generally contain images with much contrast and many edges.

**Hard segmentation.** The performance of the proposed method using hard segmentation on each stage is shown in Figure 5. The performance of the proposed method on the entire dataset  $D_2$  is given in Table 2: the median angular error equals  $4.5^\circ$ . Compared with the *baseline*, the median angular error is reduced by almost 14%. Specific examples are shown in Figure 6. Note that, due to the stage classification, we not only improve upon overall illuminant estimation accuracy, but we are also in a position to assess the illuminant color of the various geometrical constellations in the scene. This is outlined in Figure 5, where each stage “wall” is assigned its best illuminant (algorithm). Ex-



	Segmentation	Estimation	Groundtruth
	sky	(0.60, 0.59, 0.54)	(0.55, 0.57, 0.60)
	background	(0.48, 0.56, 0.67)	
	ground	(0.58, 0.58, 0.58)	
	IsidewallLR	(0.61, 0.56, 0.57)	(0.59, 0.59, 0.54)
	ceil	(0.62, 0.45, 0.64)	
	floor	(0.68, 0.58, 0.45)	

Figure 6. Examples of the proposed method using hard segmentation. The red lines in images are the boundaries of rough segmentation. Illuminant estimation of each region is obtained by 1<sup>st</sup>-order Grey-Edge, which is the best among the considered single algorithms on  $D_2$ .

panding from this, a trivial extension is the estimation of the light source color at various depth layers as indicated by the 3D stage model. This allows the estimation of a distant light source, and to distinguish a nearby illuminant (indoor, shadow) from a far away illuminant (outdoor, sunlight). As such ground truth is not yet available in current color constancy datasets, we did not pursue evaluation of these extensions.

**Soft segmentation.** The performance of the proposed method using soft segmentation on each stage is demonstrated in Figure 5 while the result over the whole dataset is shown in Table 2: the median angular error is  $4.6^\circ$ , which is quite similar to the proposed method using hard segmentation. The proposed method using soft segmentation makes an improvement 12% in median angular error over the *baseline*.

**Discussion.** In addition to automatic classification, manual classification is used to determine how the stage classification performance influences the final results. Results are shown in Table 2. Using this perfect classifier (*i.e.* the mean average precision is 1), the median angular errors of the method without segmentation is reduced to  $4.6^\circ$ . Using hard segmentation, the best-performance that can be obtained is  $3.7^\circ$  for the median angular error, while the median angular error can be further reduced to  $3.6^\circ$  by using soft segmentation. In conclusion, improving stage classification will further improve the color constancy results significantly.

Figure 7 presents three images, two of which are correctly classified while the other is misclassified due to the occlusion. The proposed method using soft segmentation is more effective in the presence of shadow or shading edges.

## 5. Conclusions

In this paper, 3D geometry models have been used to determine which color constancy method to use for different scene geometry settings. To achieve this, first images have been classified into stages. Then images have been divided into different regions using hard or soft segmentation. Af-

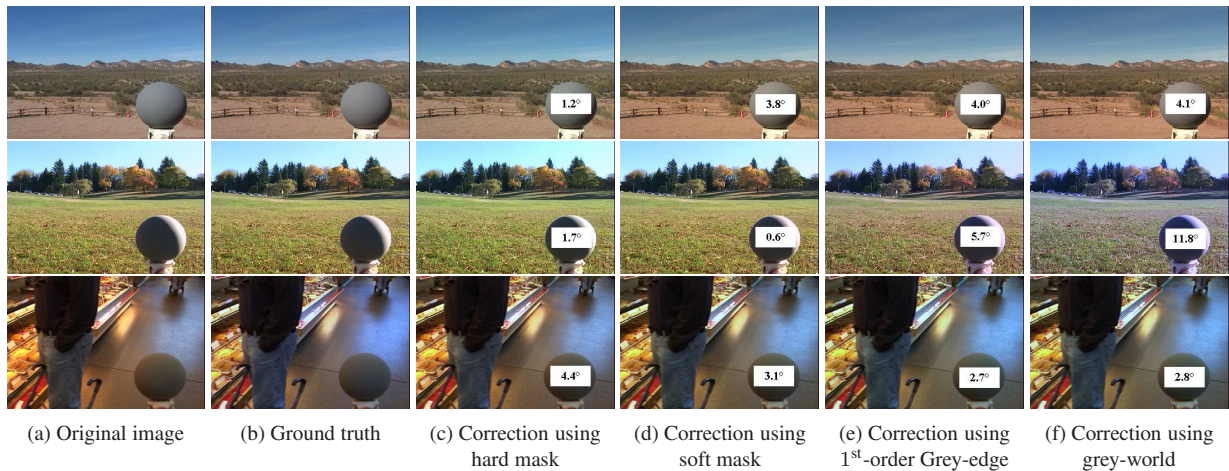


Figure 7. Results of color constancy. The angular error is given on the gray ball, which is masked during illuminant estimation.

ter that the best color constancy algorithm has been selected for each geometry segment. Besides improving color constancy algorithms by exploiting the 3D geometry structure of the scene, we have opened up the possibility to access illuminant color over the depth field in a single image. Hence, estimating which illuminant is directly in front of the camera, and the one that is illuminating the more distant scene.

Experiments on large scale image datasets show that the proposed algorithm outperforms state-of-the-art single color constancy algorithms with an improvement of almost 14% of median angular error. When using an ideal classifier (i.e., all of the test images are correctly (manually) classified into stages), the performance of the proposed method improves the median angular error as much as 31%. This gain in performance can largely be explained by the fact that most color constancy algorithms are specifically suited for images with certain image statistics, like a high (or low) signal-to-noise ratio. Further, it is shown that extracting local geometry features is more efficient than applying a global selection or combination algorithm.

## Acknowledgement

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

## References

- [1] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini. Improving color constancy using indoor-outdoor image classification. *TIP*, 17(12):2381–92, 2008.
- [2] G. J. Burghouts and J. Geusebroek. Performance evaluation of local colour invariants. *CVIU*, 13(113):48–62, 2009.
- [3] A. Chakrabarti, K. Hirakawa, and T. Zickler. Color constancy beyond bags of pixels. In *CVPR*, 2008.
- [4] C.-C. Chang and C.-J. Lin. *LIBSVM: a library for support vector machines*, 2001.
- [5] F. Ciurea and B. Funt. A large image database for color constancy research. In *CIC*, pages 160–164, 2003.
- [6] E. Delage, H. Lee, and A. Y. Ng. A dynamic bayesian network model for autonomous 3d reconstruction from a single indoor image. In *CVPR*, pages 2418–2428, 2006.
- [7] M. Ebner. *Color constancy*. Wiley, 2007.
- [8] G. Finlayson, S. Hordley, and C. Lu. On the removal of shadows from images. *TPAMI*, 28(1):59–68, 2006.
- [9] A. Gijsenij and T. Gevers. Color constancy using natural image statistics. In *CVPR*, 2007.
- [10] R. Henry, S. Mahadev, S. Urquijo, and D. Chitwood. Color perception through atmospheric haze. *J. Opt. Soc. Am. A*, 17(5):831–835, 2000.
- [11] D. Hoiem, A. A. Efros, and M. Hebert. Geometric context from a single image. In *ICCV*, pages 654–661, 2005.
- [12] G. Hordley. Scene illuminant estimation: past, present, and future. *Color Res. and App.*, 31(4):303–314, 2006.
- [13] S. Hordley and G. Finlayson. Reevaluation of color constancy algorithm performance. *J. Opt. Soc. Am. A*, 23(5):1008–1020, 2006.
- [14] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 60:91–110, 2004.
- [15] V. Nedović, A. Smeulders, A. Redert, and J.-M. Geusebroek. Depth information by stage classification. In *ICCV*, 2007.
- [16] E. Sudderth, A. Torralba, W. Freeman, and A. Willsky. Depth from familiar objects: A hierarchical model for 3d scenes. In *CVPR*, pages 2410–2417, 2006.
- [17] A. Torralba and A. Oliva. Depth estimation from image structure. *TPAMI*, 24(9):1226–1238, 2002.
- [18] K. van de Sande, T. Gevers, and C. Snoek. Evaluation of color descriptors for object and scene recognition. In *CVPR*, 2008.
- [19] J. van de Weijer, T. Gevers, and A. Gijsenij. Edge-based color constancy. *TIP*, 16(9):2207–2214, 2007.
- [20] J. van de Weijer, C. Schmid, and J. Verbeek. Using high-level visual information for color constancy. In *ICCV*, 2007.