

# Improving Color Constancy by Photometric Edge Weighting

Arjan Gijsenij, *Member, IEEE*, Theo Gevers, *Member, IEEE*, Joost van de Weijer, *Member, IEEE*

**Abstract**—Edge-based color constancy methods make use of image derivatives to estimate the illuminant. However, different edge types exist in real-world images such as material, shadow and highlight edges. These different edge types may have a distinctive influence on the performance of the illuminant estimation.

Therefore, in this paper, an extensive analysis is provided of different edge types on the performance of edge-based color constancy methods. First, an edge-based taxonomy is presented classifying edge types based on their photometric properties (e.g. material, shadow-geometry and highlights). Then, a performance evaluation of edge-based color constancy is provided using these different edge types. From this performance evaluation it is derived that specular and shadow edge types are more valuable than material edges for the estimation of the illuminant. To this end, the (iterative) weighted Grey-Edge algorithm is proposed in which these edge types are more emphasized for the estimation of the illuminant.

Images that are recorded under controlled circumstances demonstrate that the proposed iterative weighted Grey-Edge algorithm based on highlights reduces the median angular error with approximately 25%. In an uncontrolled environment, improvements in angular error up to 11% are obtained with respect to regular edge-based color constancy.

**Index Terms**—Color Constancy, Illuminant estimation, Grey-Edge, Edge classification



## 1 INTRODUCTION

CHANGES in illumination cause the measurements of object colors to be biased towards the color of the light source. The goal of color constancy is to provide invariance with respect to these changes. Color constancy facilitates many computer vision related tasks like color feature extraction [1] and color appearance models [2].

Many computational color constancy algorithms have been proposed, see e.g. [3, 4] for an overview. Traditionally, computational color constancy methods use pixel values of an image to estimate the illuminant. Examples of such methods include approaches based on low-level features, e.g. [5, 6, 7] and gamut-based algorithms [8, 9, 10]. Recently, methods that use derivatives (i.e. edges) and even higher-order statistics are proposed [11, 12, 13, 14].

The underlying assumption of Grey-World and Grey-Edge-based algorithms is that the distribution of the colors and edges is directed towards the illuminant direction. Hence, in order to accurately recover the color of the light source, ideally, only those pixels and edges should be used that coincide with the illuminant direction (highlights are one example of such pixels, perfect reflectances are another). Under the

assumption of neutral interface reflection, it is known that highlights roughly correspond to the color of the light source, making highlights particularly suited for estimating the color of the light source [15, 16, 17]. However, detecting highlight *pixels* has proven to be very challenging without prior knowledge of the scene [15, 18, 16, 19]. *Edges*, on the other hand, can automatically be classified into different types without much prior knowledge by using physics principles [20, 21, 22, 23]. For example, edges can be classified into material edges (e.g. edges between objects and object-background edges), shadow/shading edges (e.g. edges caused by the shape or position of an object with respect to the light source) and specular edges (i.e. highlights). These edges may have a distinctive influence on the performance of illuminant estimation.

In this paper, the use of distinct edge types is exploited to improve the performance of edge-based color constancy by computing a weighted average of the edges. The weights are computed using a photometric edge classification scheme. Since such methods often assume the scene is illuminated by a white light source, the automatic detection of such edges can become erroneous when the color of the light source is *not* white. To this end, an iterative weighting scheme is proposed that sequentially estimates the color of the light source and updates the computed edge weights. The rationale behind this approach is to fully exploit the information that is enclosed in the image, and simultaneously increase the accuracy of the illuminant estimation and (specular) edge detection.

This paper is organized as follows. In section 2,

- Arjan Gijsenij is with Alten PTS, the Netherlands.
- Theo Gevers is with the Intelligent Systems Lab Amsterdam (ISLA), University of Amsterdam, The Netherlands and with the Computer Vision Center (CVC), Campus Universitat Autònoma de Barcelona, Spain.
- Joost van de Weijer is with the Computer Vision Center (CVC), Campus Universitat Autònoma de Barcelona, Spain.
- e-mail: arjan.gijsenij@gmail.com; th.gevers@uva.nl; joost@cvc.uab.es

color constancy is discussed, followed by the introduction of color constancy by edge weighting in section 3. Then, in section 4, the performance of edge-based color constancy is analyzed with respect to different edge types. Finally, in sections 5 and 6, the method is evaluated and the obtained results are discussed.

## 2 COLOR CONSTANCY

The image values  $\mathbf{f} = (f_R, f_G, f_B)^T$  for a Lambertian surface depend on the color of the light source  $I(\lambda)$ , the surface reflectance  $S(\mathbf{x}, \lambda)$  and the camera sensitivity function  $\boldsymbol{\rho}(\lambda) = (\rho_R(\lambda), \rho_G(\lambda), \rho_B(\lambda))^T$ , where  $\lambda$  is the wavelength of the light and  $\mathbf{x}$  is the spatial coordinate (e.g. [3, 10, 8]):

$$f_c(\mathbf{x}) = m(\mathbf{x}) \int_{\omega} I(\lambda) \rho_c(\lambda) S(\mathbf{x}, \lambda) d\lambda, \quad (1)$$

where  $\omega$  is the visible spectrum,  $m(\mathbf{x})$  is Lambertian shading and  $c = \{R, G, B\}$ . It is assumed that the scene is illuminated by uniform illumination and that the observed color of the light source  $\mathbf{e}$  depends on the color of the light source  $I(\lambda)$  as well as the camera sensitivity function  $\boldsymbol{\rho}(\lambda)$ :

$$\mathbf{e} = \begin{pmatrix} e_R \\ e_G \\ e_B \end{pmatrix} = \int_{\omega} I(\lambda) \boldsymbol{\rho}(\lambda) d\lambda. \quad (2)$$

Color constancy can be achieved by estimating the color of the light source  $\mathbf{e}$ , given the image values of  $\mathbf{f}$ , followed by a transformation of the original image values using this illuminant estimate [24]:

$$\mathbf{f}_t = \mathcal{D}_{u,t} \mathbf{f}_u, \quad (3)$$

where  $\mathbf{f}_u$  is the image taken under an unknown light source,  $\mathbf{f}_t$  is the same image transformed, so it appears if it was taken under the canonical illuminant, and  $\mathcal{D}_{u,t}$  is a diagonal matrix which maps colors that are taken under an unknown light source  $u$  to their corresponding colors under the canonical illuminant  $c$ . The aim of this transformation is not to scale the brightness level of the image, since the color constancy methods proposed and compared in this paper only correct for the *chromaticity* of the light source. Since both  $I(\lambda)$  and  $\boldsymbol{\rho}(\lambda)$  are, in general, unknown, the estimation of  $\mathbf{e}$  is an under-constrained problem that cannot be solved without further assumptions.

### 2.1 Pixel-based Color Constancy

Two well-known algorithms that are often used are based on the Retinex Theory proposed by Land [7]: the White-Patch and the Grey-World algorithm. The White-Patch algorithm is based on the White-Patch assumption, i.e. the assumption that *the maximum response in the RGB-channels is caused by a white patch*. In practice, this assumption is alleviated by considering the color channels separately, resulting in the

*max-RGB* algorithm. The Grey-World algorithm [5] is based on the Grey-World assumption, i.e. *the average reflectance in a scene is achromatic*. Another type of algorithms is gamut-based methods, originally proposed by Forsyth [10]. Gamut-based algorithms use more advanced statistical information about the image, and are based on the assumption, that *in real-world images, one observes, under a given illuminant, only a limited number of different colors*.

Another pixel-based algorithm, that is related to the current paper, is proposed by Tan et al. [17]. This approach is based on the dichromatic reflection model, and uses specularities or highlights for the estimation of the illuminant. By transforming the image to *inverse intensity chromaticity space*, pixels are identified that have a low body reflectance factor (effectively identifying pixels with a color that is close or identical to the color of the light source). However, the identification of such specular pixels remains a problem.

### 2.2 Edge-based Color Constancy

Extending pixel-based methods to incorporate derivative information, i.e. edges and higher-order statistics, resulted in the Grey-Edge [14] and the derivative-based Gamut mapping algorithms[13].

The Grey-Edge actually comprises a framework that incorporates zeroth-order methods (e.g. the Grey-World and the White-Patch algorithms), first-order methods (i.e. the Grey-Edge), as well as higher-order methods (e.g.  $2^{nd}$ -order Grey-Edge). Many different algorithms can be created by varying the three parameters in:

$$\mathbf{e}^{n,p,\sigma} = \left( \int \left| \frac{\partial^n f_{c,\sigma}(\mathbf{x})}{\partial \mathbf{x}^n} \right|^p d\mathbf{x} \right)^{\frac{1}{p}} = k e_c, \quad (4)$$

where  $|\cdot|$  indicates the absolute value,  $c = \{R, G, B\}$ ,  $n$  is the order of the derivative,  $p$  is the Minkowski-norm and  $k$  a multiplicative scalar constant chosen such that the illuminant vector  $\mathbf{e}$  has unit length.

Further, the derivative is defined as the convolution of the image with the derivative of a Gaussian filter with scale parameter  $\sigma$  [25]:

$$\frac{\partial^{s+t} f_{c,\sigma}}{\partial x^s \partial y^t} = f_c * \frac{\partial^{s+t} G_{\sigma}}{\partial x^s \partial y^t} \quad (5)$$

where  $*$  denotes the convolution and  $s + t = n$ . Good results are obtained by using instantiation  $\mathbf{e}^{1,1,\sigma}$ , i.e. a simple average of the edges at scale  $\sigma$  also called the Grey-Edge method [14].

Another pixel-based method which has been extended to incorporate derivative information is the Gamut mapping algorithm [13]. It can be proven that linear combinations of image values also form gamuts, thereby extending the Gamut mapping theory to incorporate image derivatives. In this paper, we assess the influence of various edge types on the

performance of both the Grey-Edge method and the derivative-based Gamut mapping method.

### 3 COLOR CONSTANCY BY EDGE WEIGHTING

Consider the following simplified version of eq. (4):

$$\left( \int |f_{c,x}(\mathbf{x})|^p d\mathbf{x} \right)^{\frac{1}{p}} = ke_c, \quad (6)$$

where  $f_{c,x}(\mathbf{x})$  is the derivative of color channel  $c \in \{R, G, B\}$  of image  $\mathbf{f}$  at a certain scale. Then, the weighted Grey-Edge algorithm is given by:

$$\left( \int |w(\mathbf{f})^\kappa f_{c,x}(\mathbf{x})|^p d\mathbf{x} \right)^{\frac{1}{p}} = ke_c, \quad (7)$$

where  $w(\mathbf{f})$  is a weighting function that assigns a weight to every value of  $\mathbf{f}$  and  $\kappa$  can be used to enforce the weights (a higher value of  $\kappa$  results in more emphasis on higher weights).

Numerous different weighting schemes can be incorporated in eq. (7), but in order to accurately estimate the color of the light source, a suitable weighting scheme should enforce relevant information about the color of the light source and disregard irrelevant information. Since it is known that highlights are a valuable cue for estimating the color of the light source [15, 16, 17], an obvious choice would be to compute weights using specular edge detection methods.

#### 3.1 Edge Types and Classification

Various edge types are considered, i.e. material edges, (colored) shadow or shading edges, specular edges and interreflection edges. Material edges are transitions between two different surfaces or objects. Shading edges are transitions that are caused by the geometry of an object, for instance by a change in surface orientation with respect to the illumination. Shadow edges are cast shadows, caused by an object that (partially) blocks the light source. Blocking of the light source often results in merely an intensity gradient, but sometimes a faint color gradient is introduced (provided two illuminants are present in a scene). When we refer to shadow edges in general, both intensity and colored shadow edges are implied. Finally, in real-world images, interreflection is an important aspect. Interreflection is the effect of light reflected from one surface onto a second surface. This effect changes the overall illumination that is received by the second surface, and hence the color of this surface. Finally, note that combinations of edge types (e.g. a shadow or shading edge that coincides with a material edge) can also occur, but are not handled explicitly here.

Generally, edge classification is based on photometric information. For instance, Finlayson et al. [20] propose to project a 2D log-chromaticity image onto

the direction orthogonal to the light source, resulting in a new image that is invariant to light intensity and color. Shadow edges are then detected by subtracting the derivatives of the invariant image and the original image. Alternatively, Geusebroek et al. [21] propose a rule-based approach in combination with a set of color invariants derived from the Kubelka-Munk theory for colorant layers. A slightly different rule-based approach is proposed by van de Weijer et al. [23], which is based on the same photometric invariant principles. Geometric information is mostly ignored for general edge classification methods, although some recent advancements show that shadow edge detection can benefit from geometric features [26, 27]. The weighted Grey-Edge method given by eq. (7) can incorporate classifications based on both photometric and geometric features, but in this paper the focus will be mainly on photometric features. More specifically, the quasi-invariants [23] are used to design several different soft weighting schemes, resulting in an elegant incorporation of weighting schemes into edge-based color constancy.

#### 3.2 Edge Weighting Schemes

Quasi-invariants [23] are computed using the derivative of an image,  $\mathbf{f}_x = (f_{R,x}, f_{G,x}, f_{B,x})^T$ , and a set of three photometric variants. By removing the variance from the derivative of the image, a complementary set of derivatives is constructed called *quasi-invariants*. The edge energy contained in the three variant directions is indicative for the type of edge, e.g. if most energy is contained in the specular direction, the edge is most likely to be a highlight. Using the quasi-invariants, three different weighting schemes can be derived (including all combinations of these schemes): the specular weighting scheme, the shadow weighting scheme and the material weighting scheme.

**Specular Edge Weighting Scheme.** The quasi-invariants decompose a derivative image into three directions. The projection of the derivative on the specular direction (i.e. the color of the light source) is called the specular variant and is defined as:

$$\mathbf{O}_x = (\mathbf{f}_x \cdot \hat{\mathbf{c}}^i) \hat{\mathbf{c}}^i, \quad (8)$$

where  $\hat{\mathbf{c}}^i = \frac{1}{\sqrt{3}}(1, 1, 1)^T$  is the color of the light source (assumed to be white here) and the dot indicates the vector inner product. The specular variant is that part of the derivative which could be caused by highlights. What remains after subtraction of the variant from the derivative is called the specular quasi-invariant:

$$\mathbf{O}_x^t = \mathbf{f}_x - \mathbf{O}_x, \quad (9)$$

The quasi-invariant  $\mathbf{O}_x^t$  only contains shadow-shading and material edges, and is insensitive to highlight edges. Since all derivative energy of an image  $\mathbf{f}$  is contained in either of the three variant directions, the ratio of the energy in the specular variant versus the

total amount of derivative energy is an indication if an edge is specular or not. This ratio translates to the following specular weighting scheme:

$$w_{s,\text{specular}}(\mathbf{f}_x) = \frac{|\mathbf{O}_x|}{\|\mathbf{f}_x\|}, \quad (10)$$

where  $|\mathbf{O}_x|$  is the absolute value of  $\mathbf{O}_x$  and  $\|\mathbf{f}_x\| = \sqrt{f_{R,x}^2 + f_{G,x}^2 + f_{B,x}^2}$ .

**Shadow Edge Weighting Scheme.** Using the same reasoning on the shadow-shading direction  $\hat{\mathbf{f}}$  (i.e. intensity), a shadow-shading invariant and quasi-invariant is obtained:

$$\mathbf{S}_x = (\mathbf{f}_x \cdot \hat{\mathbf{f}}) \hat{\mathbf{f}}, \quad (11)$$

$$\mathbf{S}_x^t = \mathbf{f}_x - \mathbf{S}_x, \quad (12)$$

where  $\hat{\mathbf{f}} = \frac{1}{\sqrt{R^2 + G^2 + B^2}}(R, G, B)^T$ . It can be derived that the shadow-shading quasi-invariant is insensitive to shadow edges. Translating this variant to a shadow weighting scheme yields the following result:

$$w_{s,\text{shadow}}(\mathbf{f}) = \frac{|\mathbf{S}_x|}{\|\mathbf{f}_x\|}. \quad (13)$$

**Material Edge Weighting Scheme.** Finally, the shadow-shading-specular variant and quasi-invariant can be constructed by projecting the derivative on the hue direction:

$$\mathbf{H}_x = (\mathbf{f}_x \cdot \hat{\mathbf{b}}) \hat{\mathbf{b}}, \quad (14)$$

$$\mathbf{H}_x^t = \mathbf{f}_x - \mathbf{H}_x, \quad (15)$$

where  $\hat{\mathbf{b}}$  is the hue direction, which is perpendicular to the previous two directions:

$$\hat{\mathbf{b}} = \frac{\hat{\mathbf{f}} \times \hat{\mathbf{c}}^i}{\|\hat{\mathbf{f}} \times \hat{\mathbf{c}}^i\|}. \quad (16)$$

The resulting quasi-invariant  $\mathbf{H}_x^t$  does not contain specular or shadow-shading edges, and can be used to assign higher weights to material edges as follows:

$$w_{s,\text{material}}(\mathbf{f}_x) = \frac{|\mathbf{H}_x^t|}{\|\mathbf{f}_x\|}. \quad (17)$$

To evaluate the influence of the edge-type classifier on the color constancy results, one additional experiment in section 5 will make use of a different specular weighting scheme, based on geometric features similar to [26].

### 3.3 Iterative Weighted Grey-Edge

In equations (8) and (16), it can be derived that the specular and shadow-shading-specular variants and quasi-invariants are dependent on the color of the light source (the shadow-shading variant and quasi-invariant are not). The underlying assumption that the scene is viewed under a white (or neutral) light source [23] is obviously not met for the images in the used data sets *prior* to applying color constancy.

However, *after* the proposed algorithm is applied, the illuminant should be neutral, at least in theory. Hence, we propose to first correct the input image  $\mathbf{A}$  with an estimated illuminant  $\mathbf{I}$ . Then, using this color corrected image  $\mathbf{B}$ , we can compute a weighting scheme  $\mathbf{W}$ , which in turn is used by the weighted Grey-Edge algorithm to compute an updated estimate of the illuminant  $\mathbf{I}$ . After some iterations, the illuminant estimate will approximate a white light source, at which point the accuracy will no longer increase and the method has converged. Consequently, we propose to iteratively apply the weighted Grey-Edge algorithm, where a new instantiation of the weighting scheme is computed every iteration based on the color corrected image at each iteration. The iterative weighted Grey-Edge is given by the following lines of pseudo-code:

---

#### Algorithm 1 Iterative Weighted Grey-Edge

---

**Input:**

input image:  $\mathbf{A}$   
 initial illuminant estimate:  $\mathbf{I}$   
 stopping criterion:  $\mathbf{C}$

**Method:**

**while** ( $\neg \mathbf{C}$ ) **do**  
 $\mathbf{B} = \text{color\_correct}(\mathbf{A}, \mathbf{I})$   
 $\mathbf{W} = \text{compute\_weighting\_scheme}(\mathbf{B})$   
 $\mathbf{I}' = \text{weighted\_GreyEdge}(\mathbf{B}, \mathbf{W})$   
 $\mathbf{I} = \mathbf{I} * \mathbf{I}'$   
 $\mathbf{C} = \text{update\_stopping\_criterion}()$   
**end while**

---

For sake of clarity, we will not change the *type* of weighting scheme  $\mathbf{W}$  (e.g. specular or shadow-based) throughout the iterations. Further, the initial illuminant estimate  $\mathbf{I}$  can either be a white light source ( $(1, 1, 1)^T$ ) or it can be the result of any color constancy algorithm. Finally, the stopping criterion  $\mathbf{C}$  can be defined as a fixed number of iterations, or as some measure of convergence (e.g. the distance between a white light source and the illuminant  $\mathbf{I}'$  at the end of each iteration is below some threshold).

## 4 PERFORMANCE USING DIFFERENT EDGE TYPES

In this section, the aim is to analyze which edge types have the most influence on the accuracy of the illuminant estimation. To this end, a spectral data set is used first to generate different edges types under controlled circumstances. On this data set, the two different edge-based color constancy algorithms, i.e. the Grey-Edge and the derivative-based Gamut mapping approach, are evaluated.

To evaluate the performance of color constancy algorithms, the angular error  $\epsilon$  is widely used [28]. This measure is defined as the angular distance between

the actual color of the light source  $e_l$  and the estimated color  $e_e$ :

$$\epsilon = \cos^{-1}(\hat{e}_l \cdot \hat{e}_e), \quad (18)$$

where  $\hat{e}_l \cdot \hat{e}_e$  is the dot product of the two normalized vectors representing the true color of the light source  $e_l$  and the estimated color of the light source  $e_e$ . To measure the performance of an algorithm on a whole data set, the median angular error is reported.

#### 4.1 Spectral data

The first experiments are performed using the spectral data set introduced by Barnard et al. [29]. This set consists of 1995 surface reflectance spectra and 287 illuminant spectra, from which an extensive range of surfaces (i.e. *RGB*-values) can be generated using eq. (1). As discussed before, for these experiments, the following types of surfaces are created:

- Material surface  $\mathbf{m}_i$ :

$$\mathbf{m}_{ik} = \int_{\omega} I_k(\lambda) \rho(\lambda) S_i(\mathbf{x}, \lambda) d\lambda. \quad (19)$$

- Intensity shadow surface  $\mathbf{p}_i$ :

$$\mathbf{p}_{ik} = \int_{\omega} \frac{I_k(\lambda)}{\phi} \rho(\lambda) S_i(\mathbf{x}, \lambda) d\lambda. \quad (20)$$

- Colored shadow surface  $\mathbf{q}_i$ :

$$\mathbf{q}_{ikk'} = \mathbf{p}_{ik} + \eta \int_{\omega} I_{k'}(\lambda) \rho(\lambda) S_i(\mathbf{x}, \lambda) d\lambda, \quad (21)$$

- Specular surface  $\mathbf{h}_{ik}$ :

$$\mathbf{h}_{ik} = \mathbf{m}_{ik} + \gamma \int_{\omega} I_k(\lambda) \rho(\lambda) d\lambda, \quad (22)$$

- Interreflection surface  $\mathbf{r}_{ijk}$ :

$$\mathbf{r}_{ijk} = \int_{\omega} I_{k^*} \rho(\lambda) S_i(\mathbf{x}, \lambda) d\lambda, \quad (23)$$

where  $I_{k^*} = I_k(\lambda) + \theta I_{k'}(\lambda) S_j(\mathbf{x}, \lambda)$ . Further, the subscripts  $i$  and  $j$  denote different surface reflectance spectra,  $k$  and  $k'$  denote different illuminant spectra. The random variables  $\phi$  and  $\gamma$  are uniformly distributed between 1 and 4, and  $\eta$  and  $\theta$  are random variables uniformly distributed between 0 and 0.25.

Since the focus is on edge-based color constancy, the following transitions (i.e. edges) between surfaces are generated:

- Material edge:  $\mathbf{m}_{ik} - \mathbf{m}_{jk}$ .
- Intensity shadow edge:  $\mathbf{m}_{ik} - \mathbf{p}_{ik}$ .
- Colored shadow edge:  $\mathbf{m}_{ik} - \mathbf{q}_{ikk'}$ .
- Specular edge:  $\mathbf{m}_{ik} - \mathbf{h}_{ik}$ .
- Interreflection edge:  $\mathbf{m}_{ik} - \mathbf{r}_{ijk}$ .

A material edge is generated by taking the difference between two different material surfaces,  $\mathbf{m}_i - \mathbf{m}_j$ . The difference between a material surface  $\mathbf{m}_i$  and the same surface under a weaker light source results in an intensity shadow edge,  $\mathbf{m}_i - \mathbf{p}_i$ . A colored shadow edge is defined as the difference between a material surface  $\mathbf{m}_i$  and a colored shadow surface,  $\mathbf{m}_i - \mathbf{q}_i$ . A specular edge is defined as the difference between

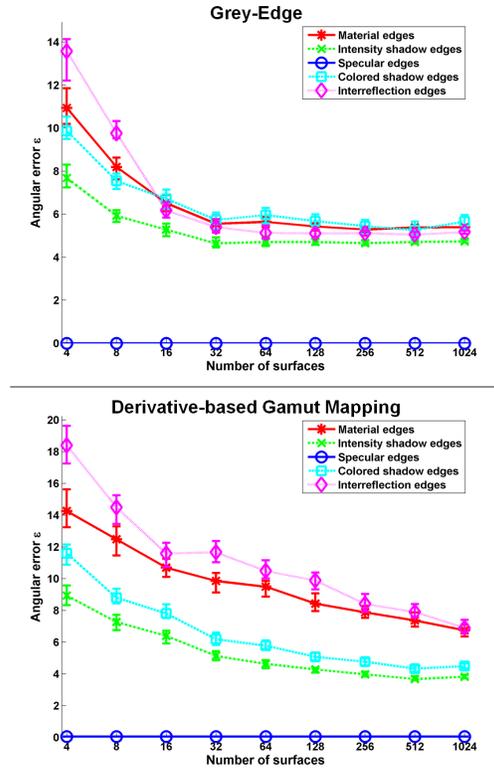


Fig. 1. Median angular error of the Grey-Edge, top figure, and the Derivative-based Gamut mapping, bottom figure, including a 95% confidence interval, using several different edge types.

a material surface  $\mathbf{m}_i$  and the bright version of the same material,  $\mathbf{m}_i - \mathbf{h}_i$ . Finally, an interreflection edge is defined as the difference between a material surface  $\mathbf{m}_i$  and an interreflection surface  $\mathbf{r}_{ij}$  where surface  $\mathbf{m}_i$  interreflects onto a second surface  $\mathbf{m}_j$ , hence  $\mathbf{m}_i - \mathbf{r}_{ij}$ . Note that these edges can be considered to be step edges. In real-world scenes, transitions are likely to be more gradual. However, for the purpose of the analysis performed in sections 4.2 and 4.3, these edges are used to give a best-case relative assessment of algorithm performance, comparing the different edge types under the same conditions. Further, we would like to note that the intrinsic properties of the used data set (i.e. the average of all surfaces is not grey) is a potential cause for error, but to avoid confusion we will ignore this in the remainder of this section.

#### 4.2 Different number of edges

In the first experiment, the performance of two edge-based color constancy algorithms is analyzed with respect to different edge types. Using the spectral data set, a number of random surfaces are created, including  $n$  material surfaces,  $n$  intensity shadow surfaces,  $n$  colored shadow surfaces,  $n$  specular surfaces and  $n$  interreflection surfaces, resulting in a total of  $5n$

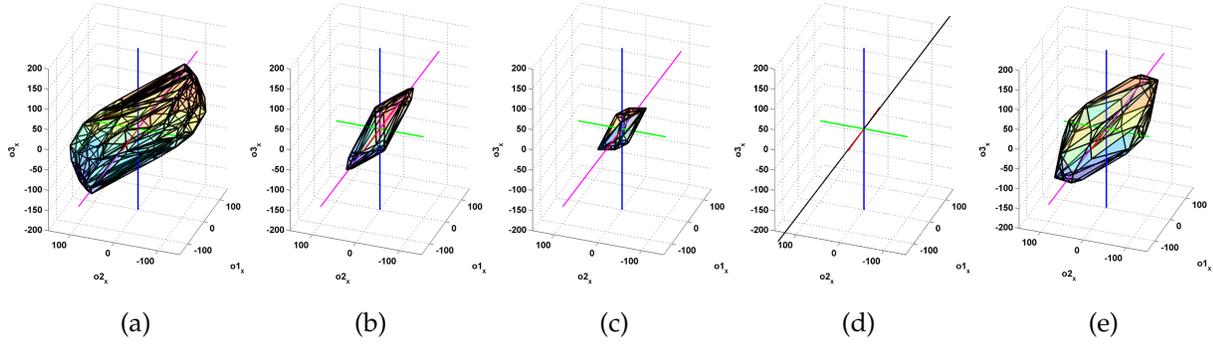


Fig. 2. Gamut in opponent color space of several edge types put under one illuminant which is specified by the fourth axis. Shown are material edges in figure (a); intensity shadow edges in figure (b); colored shadow edges in figure (c); specular edges in figure (d); interreflection edges in figure (e).

surfaces. Note that to create these surfaces, the same illuminant is used. Using these surfaces,  $n$  material edges,  $n$  intensity shadow edges,  $n$  colored shadow edges,  $n$  specular edges and  $n$  interreflection edges are created. Two edge-based color constancy algorithms are evaluated (the Grey-Edge algorithm and the derivative-based Gamut mapping) by gradually increasing the number of edges. For each value of  $n$  ( $n = \{4, 8, 16, 32, 64, 128, 256, 512, 1024\}$ ), the experiment is repeated 1000 times.

In figure 1 (top graph), the median angular error for the Grey-Edge algorithm is shown differentiated by these five edge types. The angular error when using material edges is significantly higher than when using intensity shadow edges. As expected, color constancy based on specular edges results in a close to ideal performance. Further, the performance using the colored shadow edges and the interreflection edges is similar to the performance when using the material edges. The performance of the derivative-based Gamut mapping, see figure 1 (bottom graph), shows a similar trend. Using specular edges results in near-perfect color constancy, and intensity shadow edges are more favorable than the three other types of edges.

### 4.3 Gamuts of different edge types

To analyze the high error for the material edges and to explain the differences in performance of the other types of edges, we will go into detail on the underlying assumptions. First, consider the assumption of material edge-based color constancy:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_1^N \text{Rand}_{i,j} (|\mathbf{m}_{ik} - \mathbf{m}_{jk}|) = a \int_{\omega} I_k(\lambda) \rho(\lambda) d\lambda, \quad (24)$$

where  $\mathbf{m}_{ik}$  is a material surface as defined by eq. (19),  $N$  is the number of edges,  $\text{Rand}_{i,j}(\cdot)$  is a function that randomly selects two surfaces  $i$  and  $j$  and  $a$  is a scalar value. Further, a material edge (the difference between

two material surfaces) is computed as:

$$\mathbf{m}_{ik} - \mathbf{m}_{jk} = \int_{\omega} I_k(\lambda) \rho(\lambda) (S_i(\lambda) - S_j(\lambda)) d\lambda. \quad (25)$$

Substituting eq. (25) into eq. (24) results in the following underlying assumption:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_1^N \text{Rand}_{i,j} (|S_i(\lambda) - S_j(\lambda)|) = a. \quad (26)$$

Under the assumption that the surface reflectances are normally distributed with mean  $\mu$  and variance  $\sigma^2$ , then subtracting two surfaces results in a new random variable with larger variance ( $2\sigma^2$ ). On the other hand, consider the assumption of (intensity) shadow edge-based color constancy we have:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_1^N \text{Rand}_i (|\mathbf{m}_{ik} - \mathbf{p}_{ik}|) = a \int_{\omega} I_k(\lambda) \rho(\lambda) d\lambda, \quad (27)$$

where  $\mathbf{p}_{ik}$  is an intensity shadow surface as defined by eq. (20). Substituting eq. (20) into eq. (27) results in the following underlying assumption:

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_1^N \text{Rand}_i (|S_i(\lambda)|) = a. \quad (28)$$

Under the same assumption that the surface reflectances are normally distributed with mean  $\mu$  and variance  $\sigma^2$ , it can be observed that the variance of the intensity shadow edges ( $\sigma^2$ ) is lower than the variance of the material edges ( $\sigma^2 < 2\sigma^2$ ). This implies that a larger number of (different) material edges is required to obtain the same accuracy as shadow-edge-based color constancy.

This analytical derivation is supported by the empirical distribution of these two edge types. For the ease of illustration of the physical properties of edge types, the edges are converted to the opponent color

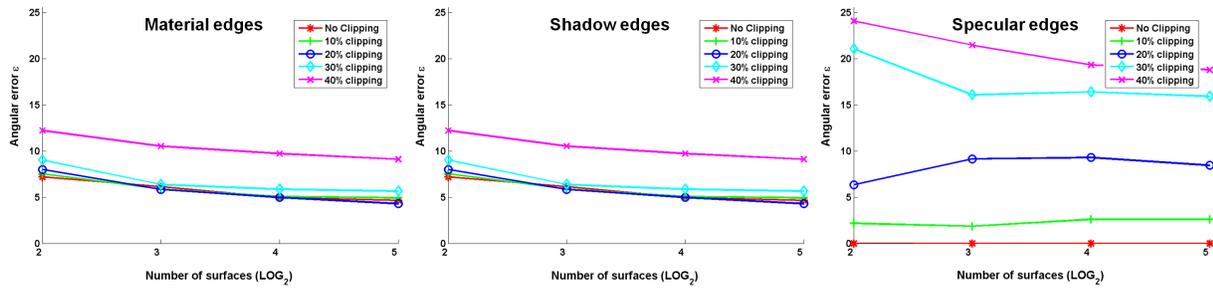


Fig. 3. Mean angular error using material edges, shadow edges and specular edges, for different clipping values.

space:

$$O1_x = \frac{R_x - G_x}{\sqrt{2}}, \quad (29)$$

$$O2_x = \frac{R_x + G_x - 2B_x}{\sqrt{6}}, \quad (30)$$

$$O3_x = \frac{R_x + G_x + B_x}{\sqrt{3}} \quad (31)$$

where  $R_x$ ,  $G_x$  and  $B_x$  are derivatives of the  $R$ ,  $G$  and  $B$  channels, respectively.

The distribution of edge responses in the opponent color space is shown in figure 2. From these graphs, it can be derived that the variation in edge color is much higher for the material edges, figure 2(a), than for shadow edges, figures 2(b) and (c), which is also analytically derived. Further, the intensity shadow edges are more directed towards the color of the light source (shown by the fourth axis) than the colored shadow edges. The shape of the gamut of the color shadow edges, which appears to be less directed towards the color of the light source than other edge types, can be explained by the influence of the second light source. The gamut of interreflection edges, figure 2(e), is similar to the material edges. Finally, specular edges, figure 2(d), all align perfectly with the color of the light source (shown by the fourth axis).

These graphs show that it is beneficial to use edges that are aligned with the color of the light source. The specular edges are all distributed on the diagonal representing the color of the light source, and near-perfect color constancy can be obtained using these edges. This observation is in accordance to pixel-based highlight analysis, where highlights contain valuable information about the color of the light source [15, 16]. Shadow edges are distributed denser around the color of the light source than material edges and interreflection edges, resulting in a higher performance.

**Color clipping.** In practice, pixel values are often bound to a certain maximum value. This effect is called color clipping. Since the specular surfaces have the highest  $RGB$ -values, these surfaces (and consequently the specular edges) risk to be affected by color clipping. To analyze this effect, a second experiment is performed where the generated  $RGB$ -values are

color clipped at a gradually decreasing value. The results of this experiment for the Grey-Edge algorithm are shown in figure 3. The derivative-based Gamut mapping reveals a similar trend (not shown here). The performance using the specular edges immediately starts to decrease significantly. The performance using the material and the shadow edges is less affected; the angular error does not significantly increase until 40% of the total number of surfaces are clipped. The effects of color clipping cause the gamuts of the specular edges to shift towards the intensity axis ( $O3_x$ ), hence the estimate of the illuminant will be biased towards white. Color clipping is an often occurring phenomena and cannot be prevented in practice. To alleviate the effects of color clipping on the performance of any color constancy algorithm, pixels that are potentially color clipped are often discarded. Practically, this means that pixels with a maximum response in either of the three color channels are not considered while estimating the illuminant.

To conclude, from an analytical approach, it can be derived that using specular edges for edge-based color constancy results in a close to ideal performance, because the specular edges align with the color of the light source. However, in practice, color clipping may eliminate the advantages of specular edges and cause a decrease in performance. Shadow edges contain more variation than specular edges but are still aligned with the color of the light source. Consequently, the performance of edge-based color constancy using shadow edges degrades only slightly with respect to using highlights. However, as material edges vary even more, their performance degrades even more. Although interreflection edges vary less than material edges, they are not aligned with the color of the light and hence their performance is the worst.

## 5 EXPERIMENTS

Experiments using the (iterative) weighted Grey-Edge method are performed on several data sets. First, experiments are performed on a data set containing indoor images that are recorded under controlled settings, see section 5.1. Then, in section 5.2, results are

All 321 images	Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$		Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$		Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$	All 321 images
$w_{s,specular}$	9.0°	5.6°	43.3°		5.6°	2.4°	43.8°		3.4°	2.1°	25.0°	$w_{s,specular}$
$w_{s,shadow}$	5.5°	3.1°	33.3°		5.5°	3.3°	33.1°		5.6°	3.3°	37.6°	$w_{s,shadow}$
$w_{s,material}$	19.3°	17.1°	56.0°		28.5°	26.4°	78.3°		13.2°	12.3°	37.6°	$w_{s,material}$
Subset A	Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$		Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$		Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$	Subset A
$w_{s,specular}$	8.9°	5.9°	43.3°		5.1°	2.0°	43.8°		3.4°	1.8°	25.0°	$w_{s,specular}$
$w_{s,shadow}$	4.8°	3.0°	24.5°		4.7°	3.0°	23.6°		4.9°	3.1°	24.5°	$w_{s,shadow}$
$w_{s,material}$	19.2°	16.4°	56.0°		30.5°	28.3°	78.3°		12.8°	11.3°	37.6°	$w_{s,material}$
Subset B	Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$		Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$		Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$	Subset B
$w_{s,specular}$	9.2°	5.6°	31.2°		6.8°	3.3°	35.1°		5.3°	3.2°	23.4°	$w_{s,specular}$
$w_{s,shadow}$	7.0°	3.6°	33.3°		7.1°	3.6°	33.1°		7.2°	3.7°	33.3°	$w_{s,shadow}$
$w_{s,material}$	19.5°	18.1°	50.2°		24.0°	20.6°	70.4°		14.2°	14.8°	32.6°	$w_{s,material}$

(a) Performance after one iteration

(b) Performance after 10 iterations

(c) Performance of theoretical scenario

TABLE 1

Mean, median and maximum angular errors on the SFU Controlled Indoor set. All results in this table are obtained using the optimal parameter settings for the proposed iterative weighted Grey-Edge (i.e.

Minkowski-norm = 7,  $\sigma = 5$  and  $\kappa = 8$ ). Table (a) shows the results of the weighted Grey-Edge for three different weighting schemes, table (b) shows the results of applying the *iterative* weighted Grey-Edge for the same three weighting schemes, and table (c) shows the results of the theoretical scenario, where we compute the weighting schemes from the color corrected images.

reported on two unconstrained data sets with natural images.

### 5.1 SFU Controlled Indoor

The first data set, denoted by *SFU Controlled Indoor* [29] consists of 31 different scenes, recorded under 11 different light sources, resulting in a total of 321 images. Two relevant subsets are distinguished to demonstrate the robustness of the proposed algorithm; one subset contains 223 images with minimal specularities (denoted *subset A*), another subset contains 98 images with non-negligible di-electric specularities (denoted *subset B*). The main difference between these two subsets is the fact that the images in subset A (some of which are flat monochromatic compositions of colored paper) contain few or no specularities, while all images in subset B contain at least *some* highlights. Some examples are shown in figure 6 (top row).

**Weighting Schemes.** Results of applying the three weighting schemes, emphasizing specular, shadow or material edges, are shown in table 1. Interestingly, results of assigning higher weights to specular edges initially, i.e. after one iteration, is worse than assigning higher weights to shadow edges. The reason for this is explained in section 3, i.e. the specular quasi-invariant assumes a neutral illumination while the shadow-shading quasi-invariant does not. Hence, the detection of highlights suffers from the colored light sources.

Running multiple iterations increases the accuracy of the specular edge detection, and hence the accuracy of the weighted Grey-Edge using the specular weighting scheme, see table 1(b). The effects of the running multiple iterations has only minor effects on the shadow weighting scheme, while the performance of the material weighting scheme considerably deteriorates, see also figure 4(a). The latter can be explained

by the misclassification rate of the specular edges: when running multiple iterations, less specular edges are misclassified as material edges. The fact that the performance of the shadow weighting scheme is not affected by running multiple iterations is expected because the shadow-shading quasi-invariant is robust to illumination changes (see section 3.2).

Results shown in table 1 are obtained using a relatively high value for  $\kappa$  ( $\kappa = 8$ ). The effect of using lower values for  $\kappa$  is shown in figure 4(b). It can be observed that especially the specular weighting scheme benefits from an increased value for  $\kappa$ , while the shadow weighting scheme is not affected at all.

**Influence of Edge Classification Accuracy.** The results of the iterative weighted Grey-Edge show that the proposed method benefits from accurate specular edge detection. An illustration is presented in figure 5, showing the results and corresponding weight maps for two individual images of the first, second and final iteration. It can be seen that the accuracy of the edge classification and the illuminant estimates increase simultaneously. These examples imply that the proposed method can benefit from accurate specular edge detection, while specular edges can be de-

Classifier	AUC	Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$
SIFT + Q.I.	0.83	5.2°	2.6°	28.5°
Quasi-Invariants	0.78	5.6°	2.9°	32.6°
Sub-optimal Q.I.	0.68	5.6°	3.2°	32.5°

TABLE 2

Results of using weighting schemes computed by applying different specular edge detection classifiers on the SFU Controlled Indoor set. The area under the ROC-curve (AUC) values are determined using cross-validation on the training set.

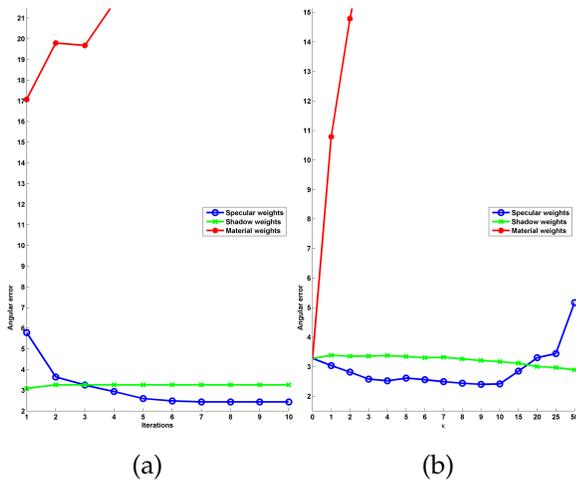


Fig. 4. The effects of applying several iterations to the different weighting schemes in figure (a). Figure (b) shows the effects of increasing the value of  $\kappa$  for different weighting schemes. Results are obtained on the SFU Controlled Indoor set.

tected more accurately on images with more accurate color constancy. To further confirm that the improved performance of the iterative weighted Grey-Edge is directly related to the accuracy of the specular edge detection, two additional experiments are performed.

As the quasi-invariants are based on neutral illumination, computing the specular weight map on the images *color corrected* with the ground truth would result in the highest edge detection accuracy. Using this weight map together with the original *uncorrected* images to estimate the color of the light source will give an indication of the potential of the proposed method. The results of this (theoretical) experiment are shown in table 1(c), where it can be observed that especially the performance of the specular weighting scheme can benefit from even more accurate specular edge detection.

**Influence of Different Edge Classifiers.** For the next experiment, the system proposed in [26] is adapted to specular edge detection. Patches from the same data set as used in [26] are manually selected and annotated. Note that these patches have no overlap with the color constancy data set. Using features derived from these patches, a classifier (SVM) is learned that is able to distinguish patches containing specular edges from patches without specular edges. By applying this classifier to a full image using a sliding window approach, a posterior probability is obtained for every block of pixels in the image. Finally, a smoothing filter is applied to reduce the inherent uncertainty of the block-based detection result. These smoothed posterior probabilities are directly used as weights in eq. (7).

Three different classifiers are learned, each with a different accuracy. The accuracy is measured as the

*area under the ROC-curve (AUC)*, and is determined using cross-validation on the training patches. The first classifier uses the SIFT-feature in combination with the quasi-invariants ( $AUC = 0.83$ ), the second classifier uses merely the quasi-invariants ( $AUC = 0.78$ ) and the third classifier uses the quasi-invariants with sub-optimal SVM-parameters ( $AUC = 0.68$ ). Results of the weight maps computed using these three classifiers are summarized in table 2. It can be seen that the classifier with the highest accuracy, i.e. the combination of SIFT and the quasi-invariants, results in the best color constancy performance. Moreover, the median angular errors of the sub-optimal classifier using the quasi-invariant features is outperformed by the optimal classifier using the same features.

Finally, to verify whether the edge detection accuracy of the training set corresponds to the accuracy on the test images, we manually annotated highlights in the SFU Controlled Indoor set. Using these manually labelled highlights and the output of the three classifiers, we are able to relate the classifier performance to the color constancy output. The accuracy of the classifier is dependent on the threshold on the classifier posterior probability, e.g. a low threshold will classify many edges as highlight and consequently result in a high recall but low precision, while a high

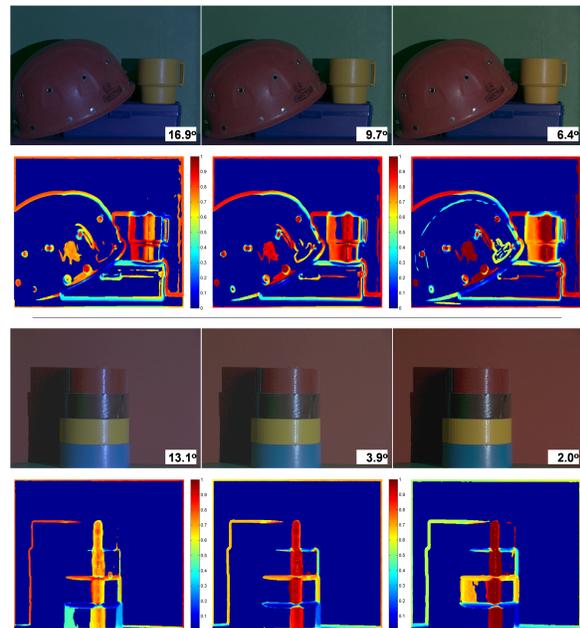


Fig. 5. Two results of running multiple iterations of the specular-weighted Grey-Edge. The first and second images show the result after the first and second iteration, the third image shows the result after the algorithm converged. The weight maps used in the corresponding iterations are color coded, such that dark red indicates a high weight and dark blue indicates a low weight.

Method	Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$
Grey-World ( $e^{0,1,0}$ )	9.8°	7.0°	37.3°
White-Patch ( $e^{0,\infty,0}$ )	9.1°	6.5°	36.2°
Using <i>I.I.C.</i> Space	15.5°	8.2°	80.9°
Gamut mapping ( $\sigma = 4$ )	4.6°	3.3°	27.1°
general Grey-World (cross-val.)	5.4°	3.5°	28.8°
1 <sup>st</sup> -order Grey-Edge (cross-val.)	5.7°	3.4°	31.8°
2 <sup>nd</sup> -order Grey-Edge (cross-val.)	5.3°	2.9°	27.6°
Proposed: Iterative (cross-val.)	5.6°	2.9°	46.7°

TABLE 3

Comparison to other algorithms on the SFU Controlled Indoor. Note that cross-validation results are obtained using three-fold cross-validation (the reported results are averaged over 100 repetitions).

threshold will classify few edges as highlight resulting in a low recall but high precision. The F-measure is often used to summarize the trade-off between precision and recall, so we use this measure here. For each method we select the threshold which yields the highest F-measure. The results of the three classifiers are 0.960 (SIFT-feature in combination with the quasi-invariants), 0.729 (quasi-invariants) and 0.513 (quasi-invariants with sub-optimal SVM-parameters). These results are in accordance with the color constancy performance in table 2 and confirm that the proposed method benefits from more accurate specular edge detection.

**Comparison to State-of-the-Art.** Results of state-of-the-art methods are shown in table 3. It should be noted that the proposed method using specular weights, as well as several other state-of-the-art methods contain parameters that considerably influence the results. Therefore, we report the performance of applying cross-validation to determine the parameter settings automatically (denoted *cross-val.* in table 3). The automatic selection of parameters is performed by three-fold cross-validation, where the reported results are the average of 100 runs.

It can be derived that the proposed method is comparable to state-of-the-art. For the gamut mapping, the implementation of [13] is used. All images recorded under one light source (*syl-50MR16Q*) are used to construct the canonical gamut, where we make sure that when testing images from a particular scene (e.g. *ball*), the corresponding training image of that scene (e.g. *ball* under *syl-50MR16Q*) is not used for computation of the canonical gamut. The mean angular error of the proposed method is slightly higher than the gamut mapping and the general Grey-World, but the median is considerably lower. Further, using the Wilcoxon sign test at 95% confidence level [28] (results are not visualized here), it is found that the proposed method is significantly better than all methods except the the general Grey-World. Finally, we would like to note that the proposed method outperforms the unweighted Grey-Edge for all choices for  $\sigma$  and

Minkowski-norm, provided the appropriate value for  $\kappa$  is selected.

## 5.2 Uncontrolled data sets

The next experiments are performed on two uncontrolled data sets containing a variety of scenes. The first data set, denoted *SFU grey-ball*, consists of 15 clips with a total of 11,346 images [30]. These images are stored in a non-linear device-RGB color space (NTSC-RGB), so to create linear images we applied gamma-correction with  $\gamma = 2.2$  and recomputed the ground truth using these linear images. Further, since the correlation among the images is rather high, video-based analysis was applied to decorrelate the visual content of the data set [31], resulting in a smaller but uncorrelated data set containing 1135 images (the reported results are obtained on the test set, consisting of 70% of the uncorrelated data, as indicated by [31]). The second data set, denoted *Color-checker*, contains 568 images [32]. This data set uses a Macbeth Color Checker that is carefully placed within the scene to capture the ground truth. Note that the latter of the two data sets does not need additional processing to acquire linear images. Examples of these data sets are shown in figure 6 (middle and bottom rows).

The results on these two data sets mostly agree with the previous experiments (see table 4). The error when using specular-based weights decreases as the algorithm is applied multiple iterations, while the error when using shadow-based weights roughly remains stable. Although the increase in performance on the SFU grey-ball set is only marginal with respect to the unweighted Grey-Edge (a decrease in mean angular error of approximately 2% is not perceptually significant [33]), the proposed method still performs significantly better than all other methods according to the Wilcoxon sign test (at the 95% confidence level). An overview of all results is shown in table 5. A possi-

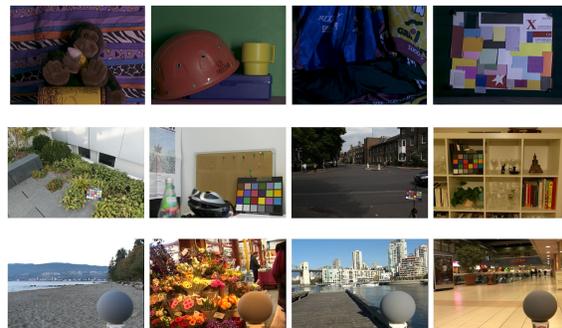


Fig. 6. Some example images of the used data sets. The top row shows some examples of the SFU Controlled Indoor set, the middle row shows examples of the color checker set and the bottom row shows some examples of the SFU grey-ball set.

SFU grey-ball	Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$		Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$		Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$	SFU grey-ball
$w_{s,specular}$	10.8°	9.4°	36.1°		10.4°	8.9°	50.4°		7.7°	6.2°	30.0°	$w_{s,specular}$
$w_{s,shadow}$	10.6°	8.6°	52.3°		10.5°	8.6°	50.1°		11.1°	9.3°	30.0°	$w_{s,shadow}$
$w_{s,material}$	19.5°	18.0°	64.5°		29.5°	29.9°	74.1°		13.7°	12.2°	43.2°	$w_{s,material}$
Color-checker	Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$		Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$		Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$	Color-checker
$w_{s,specular}$	7.6°	5.5°	36.8°		6.5°	4.7°	44.°		3.9°	3.2°	19.0°	$w_{s,specular}$
$w_{s,shadow}$	7.0°	5.2°	35.1°		7.0°	5.2°	35.1°		7.3°	5.6°	36.0°	$w_{s,shadow}$
$w_{s,material}$	19.6°	19.1°	49.4°		37.7°	37.5°	67.8°		16.5°	15.9°	38.9°	$w_{s,material}$

(a) Performance after one iteration

(b) Performance after 10 iterations

(c) Performance of theoretical scenario

TABLE 4

Mean, median and maximum angular errors on the two Uncontrolled data sets. The results on the SFU grey-ball set are obtained with  $\kappa = 10$ , the results on the color-checker set with  $\kappa = 50$  (both sets are processed with  $\sigma = 1$ ).

ble reason for the apparent lack of improvement with respect to the unweighted Grey-Edge on this data set is the number of color clipped pixels: on average, 5% of the pixels in this data set are possibly clipped (a pixel is possibly clipped if it has a maximum response in either of the three channels). The color-checker set, for instance, only consists of 0.5% color clipped pixels. Since these pixels are ignored during the estimation of the light source, they can not contribute to more accurate estimations. In fact, as was shown in section 4, the accuracy will rapidly degrade if a significant percentage of the pixels with high intensity (likely specular pixels) is ignored.

Applying the proposed method to the color checker set results in state-of-the-art performance, see table 6. The proposed method significantly outperforms the other methods according to the Wilcoxon sign test, including the gamut mapping. It should be noted that the Wilcoxon sign test is based on the number of images on which one method performs better than another. In the comparison between the proposed method and the gamut mapping, it was found that the proposed method performs better than the gamut mapping on the majority of the images, resulting in a statistical significant difference in favor of the proposed method. Running multiple iterations using the specular weighting scheme improves the performance considerably: compared to the unweighted Grey-Edge, a decrease in angular error of more than

11% can be obtained. This difference is perceptually *and* statistically significant.

Some example results of both the SFU grey-ball and the color-checker set are shown in figure 7. Overall, it can be concluded that the proposed method using specular edges improves upon the traditional Grey-Edge method. Assigning higher weights to specular edges can lead to an improvement of up to 11% (on the color-checker set). However, from the experiments in this section, it becomes clear that the proposed method introduces stronger (but not more) outliers: the maximum angular error *increases* on all data sets, but the median angular error *decreases*.

## 6 DISCUSSION

In this paper, it is shown that weighted edge-based color constancy based on specular edges can significantly improve unweighted edge-based color constancy. Further, it is shown that shadow edges contain valuable information. The reason for these conclusions can be derived as follows. Edges that are achromatic when viewed under a white light source, will accumulate in a tight gamut and assume the color of the light source when observed under colored illumination. This will increase the saturation from 0 to the saturation of the light source. Consequently, these edges are well suited for estimating the color of the light source, as all properties of the light source

Method	Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$
Grey-World ( $e^{0,1,0}$ )	13.0°	10.6°	55.2°
White-Patch ( $e^{0,\infty,0}$ )	12.3°	10.0°	38.7°
Using <i>I.I.C.</i> Space	14.0°	10.6°	72.0°
Gamut mapping ( $\sigma = 7$ )	11.9°	9.0°	39.1°
general Grey-World (cross-val.)	11.3°	9.8°	42.3°
1 <sup>st</sup> -order Grey-Edge (cross-val.)	10.7°	9.1°	48.5°
2 <sup>nd</sup> -order Grey-Edge (cross-val.)	10.8°	9.2°	46.5°
Proposed: Iterative (cross-val.)	10.4°	9.0°	50.1°

TABLE 5

Comparison to other algorithms on the SFU grey-ball set.

Method	Mean $\epsilon$	Median $\epsilon$	Max $\epsilon$
Grey-World ( $e^{0,1,0}$ )	9.8°	7.4°	46.0°
White-Patch ( $e^{0,\infty,0}$ )	8.2°	6.1°	36.3°
Using <i>I.I.C.</i> Space	9.7°	6.0°	61.3°
Gamut mapping ( $\sigma = 3$ )	7.1°	4.9°	37.0°
general Grey-World (cross-val.)	7.1°	5.4°	36.7°
1 <sup>st</sup> -order Grey-Edge (cross-val.)	7.1°	5.6°	36.3°
2 <sup>nd</sup> -order Grey-Edge (cross-val.)	7.3°	5.5°	37.3°
Proposed: Iterative (cross-val.)	6.7°	5.0°	42.1°

TABLE 6

Comparison to other algorithms on the Color Checker set.

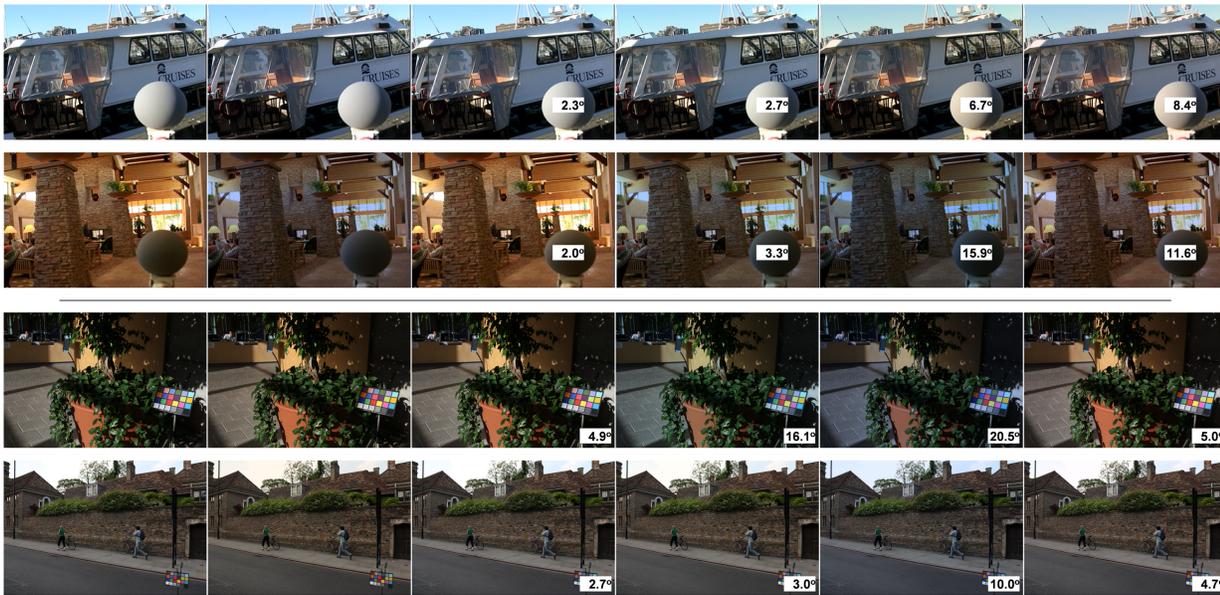


Fig. 7. Some example results of several methods, applied to the uncontrolled data sets. Note that these images are gamma-corrected for better visualization, but the estimation and correction are performed on the linear images. The value reported in the bottom right corner indicates the angular error. The top two rows show examples of the SFU grey-ball set, with from left to right, the original image, the result of correction with the ground truth, the proposed method, the 1<sup>st</sup>-order Grey-Edge, the Grey-World and using *I.I.C.* Space. The bottom two rows show examples of these algorithms applied to the color-checker set.

are contained in this edge. This was shown in section 4.

Colored edges, on the other hand, are edges that correspond to the transition from one surface (e.g. a red surface when viewed under a white light source) to another surface (e.g. a blue surface when viewed under a white light source). The saturation of this red-to-blue edge when viewed under a white light source can be an arbitrary value. Moreover, when this edge is observed under a colored light source, the saturation and color will result in gamuts with large variation that can take on unpredictable values, from which it is extremely hard (if not impossible) to estimate the color of the light source. In general, the hue of an edge is more altered by the illuminant when the saturation of that edge under a white light source is lower. More formally, a negative correlation exists between  $S_w$  and  $d_{wu}$ , where  $S_w$  is the saturation of an edge under a white light source and  $d_{wu}$  is the distance of that edge under a white light source  $w$  to that same edge viewed under an unknown light source  $u$ .

From this, it can be concluded that edges that are unsaturated under a white light source are good candidates for estimating the color of the light source. Specular and shadow edges are examples of such edges. However, specular edges are difficult to detect because of disturbances like a colored illumination. Using a specular edge detector that is dependent on the color of the illumination will introduce the necessity of running multiple iterations. The shadow

edges can be detected regardless of the color of the light source, but are of less value than the specular edges.

If we assume highlights are the most important cue for the estimation of the color of the light source, it becomes obvious that Grey-Edge-based methods gain more from this knowledge than pixel-based methods. In a typical scene, the number of pixels that coincide with the illuminant direction is vastly outnumbered by the number of pixels that do *not* coincide with this direction: most pixels are not highlights or perfect reflectances. For edges, a similar conclusion holds, although the ratio of edges that coincide with the illuminant direction versus edges that do not is likely to be larger. Indeed, uniformly colored patches will only have non-zero edge energy near the boundary of the patch, so small patches will generate relatively more edges with non-zero energy than large patches. Consequently, as highlights (or other pixels that coincide with the color of the light source) are often small regions in an image, the edges that are caused by these regions are likely to stand out among the other edges more than highlight pixels among other pixels. Note, however, even though edges are more likely to coincide with the illuminant direction than pixels, this does not mean pixels are inferior to edges at all times. Moreover, even scenes with few or no edges or pixels that coincide with the illuminant direction could result in accurate illuminant estimates: the average of a scene can still be an accurate cue

for the color of the light source if there is enough variation among the pixels or edges.

The main advantage of the proposed (iterative) weighted Grey-Edge over existing methods is that additional information, which is provided by the different edge types, is used. This information, when available, results in more accurate illuminant estimates. Moreover, the proposed method connects two theories involving color image analysis, i.e. the Grey-Edge for color constancy and the Quasi-Invariants for edge classification. The disadvantage of the proposed method is that in case of misclassification of the edge types the method may result in lesser performance. The weighted Grey-Edge inherits the weakness of the regular Grey-Edge: opaque parameter selection, i.e. the optimal parameters are difficult to select without prior knowledge of the input images.

## 7 CONCLUSION

In this paper, the influence of different edge types on the performance of edge-based color constancy is analyzed.

It is shown that weighted edge-based color constancy based on specular edges can result in accurate illuminant estimates. Using a weight map that is based on shadow edges performs slightly worse than specular edges, but considerably better than using material edges. However, the accuracy of the detection of specular edges is degraded by failing assumptions of the quasi-invariants (like the assumption of a white light source). Iteratively classifying specular edges and estimating the illuminant can considerably reduce the dependency of specular weights on the intrinsic assumption of a white light source, and hence result in a better performance. Weight maps that put more emphasis on shadow edges are not dependent on the color of the light source, and result in stable illuminant estimates.

Experiments on images that are recorded under controlled circumstances demonstrate that the proposed iterative weighted Grey-Edge algorithm based on highlights reduces the median angular error with approximately 25%. Further, in experiments on images that are recorded in an uncontrolled environment, improvements in angular error up to 25% with respect to unweighted edge-based color constancy are obtained.

## ACKNOWLEDGMENT

This work has been supported by the EU projects ERGTS-VICI-224737; by the Spanish Research Program Consolider-Ingenio 2010: MIPRCV (CSD200700018); and by the Spanish projects TIN2009-14173. Joost van de Weijer acknowledges the support of a Ramon y Cajal fellowship.

## REFERENCES

- [1] T. Gevers and A. Smeulders, "Pictoseek: combining color and shape invariant features for image retrieval," *IEEE TIP*, vol. 9, no. 1, pp. 102–119, 2000.
- [2] M. Fairchild, *Color Appearance Models, 2nd Ed.* Chichester, UK: John Wiley & sons, 2005.
- [3] S. Hordley, "Scene illuminant estimation: past, present, and future," *CRA*, vol. 31, no. 4, pp. 303–314, 2006.
- [4] M. Ebner, *Color constancy*, ser. Wiley-IS&T Series in Imaging Science and Technology. John Wiley & Sons, 2007.
- [5] G. Buchsbaum, "A spatial processor model for object colour perception," *J. Franklin Institute*, vol. 310, no. 1, pp. 1–26, 1980.
- [6] G. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Proc. CIC*, 2004, pp. 37–41.
- [7] E. Land, "The retinex theory of color vision," *Scientific American*, vol. 237, no. 6, pp. 108–128, December 1977.
- [8] G. Finlayson, S. Hordley, and P. Hubel, "Color by correlation: a simple, unifying framework for color constancy," *IEEE TPAMI*, vol. 23, no. 11, pp. 1209–1221, 2001.
- [9] G. Finlayson, S. Hordley, and I. Tastl, "Gamut constrained illuminant estimation," *IJCV*, vol. 67, no. 1, pp. 93–109, 2006.
- [10] D. Forsyth, "A novel algorithm for color constancy," *IJCV*, vol. 5, no. 1, pp. 5–36, 1990.
- [11] A. Chakrabarti, K. Hirakawa, and T. Zickler, "Color constancy beyond bags of pixels," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2008.
- [12] H. Chen, C. Shen, and P. Tsai, "Edge-based automatic white balancing with linear illuminant constraint," in *Visual Communications and Image Processing*, 2007.
- [13] A. Gijsenij, T. Gevers, and J. van de Weijer, "Generalized gamut mapping using image derivative structures for color constancy," *IJCV*, vol. 2–3, pp. 127–139, 2010.
- [14] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *IEEE TIP*, vol. 16, no. 9, pp. 2207–2214, 2007.
- [15] K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms; part i," *IEEE TIP*, vol. 11, no. 9, pp. 972–984, 2002.
- [16] K. Barnard and B. Funt, "Color constancy with specular and non-specular surfaces," in *Proc. CIC*, 1999, pp. 114–119.
- [17] R. 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, 2004.
- [18] K. Barnard, G. Finlayson, and B. Funt, "Color constancy for scenes with varying illumination,"

*Comp. Vis. Im. Und.*, vol. 65, no. 2, pp. 311–321, 1997.

- [19] G. Finlayson, B. Funt, and K. Barnard, “Color constancy under varying illumination,” in *Proc. ICCV*, 1995, pp. 720–725.
- [20] G. Finlayson, S. Hordley, C. Lu, and M. Drew, “On the removal of shadows from images,” *IEEE TPAMI*, vol. 28, no. 1, pp. 59–68, 2006.
- [21] J. Geusebroek, R. van den Boomgaard, A. Smeulders, and H. Geerts, “Color invariance,” *IEEE TPAMI*, vol. 23, no. 12, pp. 1338–1350, 2001.
- [22] T. Gevers and H. Stokman, “Classifying color edges in video into shadow-geometry, highlight, or material transitions,” *IEEE T. Multimedia*, vol. 5, pp. 237–243, 2003.
- [23] J. van de Weijer, T. Gevers, and J. Geusebroek, “Edge and corner detection by photometric quasi-invariants,” *IEEE TPAMI*, vol. 27, no. 4, pp. 625–630, 2005.
- [24] J. von Kries, “Influence of adaptation on the effects produced by luminous stimuli,” in *Sources of Color Vision*, D. MacAdam, Ed. MIT Press, 1970, pp. 109–119.
- [25] W. Freeman and E. Adelson, “The design and use of steerable filters,” *IEEE TPAMI*, vol. 13, no. 9, pp. 891–906, 1991.
- [26] A. Gijsenij and T. Gevers, “Shadow edge detection using geometric and photometric features,” in *IEEE International Conference on Image Processing*, 2009, pp. 1–8.
- [27] J. Zhu, K. Samuel, S. Masood, and M. Tappen, “Learning to recognize shadows in monochromatic natural images,” in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2010, pp. 223–230.
- [28] S. Hordley and G. Finlayson, “Reevaluation of color constancy algorithm performance,” *J. Opt. Soc. Am. A*, vol. 23, no. 5, pp. 1008–1020, 2006.
- [29] K. Barnard, L. Martin, B. Funt, and A. Coath, “A data set for color research,” *CRA*, vol. 27, no. 3, pp. 147–151, 2002.
- [30] F. Ciurea and B. Funt, “A large image database for color constancy research,” in *Proc. CIC*, 2003, pp. 160–164.
- [31] S. Bianco, G. Ciocca, C. Cusano, and R. Schettini, “Improving color constancy using indoor-outdoor image classification,” *IEEE TIP*, vol. 17, no. 12, pp. 2381–2392, 2008.
- [32] P. Gehler, C. Rother, A. Blake, T. Minka, and T. Sharp, “Bayesian color constancy revisited,” in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, 2008, pp. 1–8.
- [33] A. Gijsenij, T. Gevers, and M. Lucassen, “A perceptual analysis of distance measures for color constancy,” *J. Opt. Soc. Am. A*, vol. 26, no. 10, pp. 2243–2256, 2009.



on the program committee of several conferences and workshops.

**Arjan Gijsenij** is currently scientific software consultant for Alten PTS. Formerly, he was postdoctoral researcher at the University of Amsterdam. He received his M.Sc. degree in Computer Science at the University of Groningen in 2005, and obtained his Ph.D. degree at the University of Amsterdam in 2010. His main research interests are in the field of color in computer vision and psychophysics. He has published papers in high-impact journals and conferences, and served



geometric and photometric invariants. He is chair of various conferences and he is associate editor for the IEEE Transactions on Image Processing. Further, he is program committee member of a number of conferences, and an invited speaker at major conferences. He is a lecturer of post-doctoral courses given at various major conferences (CVPR, ICPR, SPIE, CGIV). He is member of the IEEE.

**Theo Gevers** is an Associate Professor of Computer Science at the University of Amsterdam, The Netherlands, where he is also teaching director of the M.Sc. of Artificial Intelligence. He currently holds a VICI-award (for excellent researchers) from the Dutch Organisation for Scientific Research. His main research interests are in the fundamentals of contentbased image retrieval, color image processing and computer vision specifically in the theoretical foundation of



**Joost van de Weijer** received his MSc (Delft University of Technology) in 1998 and his PhD (University of Amsterdam) in 2005. He was Marie Curie Intra-European fellow in INRIA Rhone-Alpes. He was awarded the Ramon y Cajal Research Fellowships in Computer Science by the Spanish Ministry of Science and Technology. He is based in the Computer Vision Center in Barcelona (Spain) since 2008. He is member of the IEEE.