

# The bright-chromagenic algorithm for illuminant estimation

Clément Fredembach<sup>1,2,+</sup> and Graham Finlayson<sup>1</sup>

1) School of Computing Sciences, University of East Anglia, Norwich, UK

2) School of Computer and Communication Sciences, Ecole Polytechnique Federale

de Lausanne (EPFL), Lausanne, Switzerland

+ corresponding author

`clement.fredembach@epfl.ch`, `graham@cmp.uea.ac.uk`

## Abstract

This paper proposes a new algorithm for illuminant estimation based on the concept of chromagenic color constancy, where two pictures are taken from each scene: a normal one and one where a colored filter is placed in front of the camera.

The basic formulation of the chromagenic algorithm has inherent weaknesses, namely, a need for perfectly registered images and occasional large errors in illuminant estimation. Our first contribution is to analyze the algorithm performance with respect to the reflectances present in a scene and demonstrate that fairly bright and de-saturated reflectances (e.g., achromatic and pastel colors) provide significantly better chromagenic illuminant estimation.

We thus propose the bright-chromagenic algorithm and show that it not only remedies the large error problem but also allows us to relax the image registration constraint. Experiments performed on a variety of synthetic and real data show that the newly designed bright-chromagenic algorithm significantly outperforms current illuminant estimation methods, including those having a substantially higher complexity.

**keywords:** color constancy, illuminant estimation, chromagenic.

## I. INTRODUCTION

The human visual system is, to a certain extent, color constant <sup>[1],[2],[3]</sup>; that is, it discounts the color of the illumination. This is why, for example, snow always appears white, no matter under which illuminant it is observed.

However, it has proven difficult to emulate color constancy in imaging workflow. This is not only a problem in image reproduction but also for a variety of computer vision tasks, such as tracking <sup>[4]</sup>, indexing <sup>[5]</sup>, and scene analysis <sup>[6]</sup>, where stable measures of reflectance are sought or assumed for objects in a scene.

Solving for color constancy is a two-step process. First, the color of the prevailing illuminant is estimated. At a second stage, the color bias due to illumination is removed. This second part is, in fact, relatively easy <sup>[7]</sup> and so most color constancy algorithms focus on the illuminant estimation problem.

Numerous algorithms for illuminant estimation have been proposed and can broadly be categorized in two groups. Algorithms in the first group make simple assumptions about the scene being observed. MaxRGB assumes that a maximally reflective patch exist in the image. Grey World assumes the average reflectance in a scene is grey <sup>[8]</sup> or some sort of grey <sup>[9],[10]</sup>. Another group of algorithms comprises more sophisticated approaches such as neural networks <sup>[11]</sup>, color by correlation <sup>[12]</sup>, a bayesian method that correlates the RGBs in the image with plausible RGBs under various illuminants to find the best illuminant, and Gamut Mapping methods <sup>[13],[14]</sup>. Generally, the most complex algorithms perform better, but at the expense of a (much) greater computational complexity.

The chromagenic algorithm makes a different assumption altogether. Like stereo, where two images are used to recover 3-d position of points in the scene, and photometric vision, where two polarizing filters in opposite directions can be used to identify and remove specular highlights <sup>[15]</sup>, two images are used: one normal and one where a colored filter is placed in front of the camera. The filter is chosen so that the relationship between filtered and unfiltered RGBs depends strongly on illumination. Such a filter is called a *chromagenic* filter.

The standard chromagenic color constancy algorithm <sup>[16]</sup> works in two stages: the *training* stage is a pre-processing step where the relationship between filtered and unfiltered RGBs is calculated using a given filter, camera sensitivities, and a number of candidate lights. Then, those relations are *tested* on other images in order to estimate the actual scene illuminant. While the general outcome of the algorithm shows good performance, two problems usually remain: for certain combinations of reflectances and illuminants, the error between the estimated and actual light can be large. Also, in order to achieve good performance, one has to compare RGBs transitions that occur between *identical* reflectances. It is, in essence, a pixel-level method and so the algorithm can fail when the pair of images is not exactly registered.

Our approach, first mentioned by Fredembach and Finlayson <sup>[17]</sup>, starts by asking the question that if everything else (lights, camera sensitivities and the chromagenic filter) is equal, what is the influence of different scene reflectances on the transforms and the estimation error? To answer this question, we first select 287 typical lights and 1995 reflectances from the Simon Fraser database <sup>[18]</sup>, and a filter from a set of 53 Wratten photographic filters. Using these elements, synthetic images, composed of a randomly selected illuminant and of 1-8 distinct reflectances, are created. Testing the algorithm on these images allows evaluating which RGBs exhibit a good (very low errors) or bad (very high errors) estimate of the illuminant. The result show that achromatic reflectances yield lower errors than strongly chromatic ones. This outcome is, however, not sufficient to allow the algorithm to work on real images. Indeed, image registration is still a problem and noise levels in dark achromatic pixels can impact performance.

However, independently of the image capture conditions, a white reflectance (which is achromatic) will still be the brightest value after filtering. Restricting the algorithm to test correspondences between the “whitest” RGBs of both filtered and unfiltered images will effectively use the fact that achromatic reflectances are more reliable for the chromagenic algorithm, while focusing on parts of the image that have a high signal to noise ratio. Practically, we average the RGB values of the brightest 1-3% of the original image and compare them to the brightest 1-3% values of the filtered image.

By using the brightest reflectances only, we arrive at a similar conclusion than Tominaga et al. <sup>[19]</sup>, where it was first surmised that brighter reflectances would perform better based on signal-to-noise ration considerations and then proceeded to show that this was indeed the case. In contrast, we start with all possible reflectances and look for the a set that is more reliable. Desaturation, not brightness, is actually the main correlation with good-performing reflectances. We chose bright-achromatic reflectances as a target because they are reliably found in images, are easy to identify, permit to solve the registration constraint, and are far from the worst performing colors (dark, saturated colors).

This modified algorithm is tested on four different image databases and its results are compared to those achieved by both simple and complex illuminant estimation algorithms. Results show that that our algorithm significantly outperforms other available methods, according to the median angular error and the Wilcoxon sign test, error measures recommended by Hordley and Finlayson <sup>[20]</sup>.

The rest of this paper is organized as follows. Section 2 reviews the mathematical bases of the chromagenic theory and discusses how the filter and the camera sensors affect the chromagenic performance. In Section 3, we propose our new algorithm, the bright-chromagenic algorithm, based on a detailed reflectance error analysis. Section 4 presents comparative experiments on four different image sets that range from purely synthetic to real images. The paper concludes in Section 5.

## II. THE CHROMAGENIC ALGORITHM

Chromagenic color constancy is performed using two images of the same scene: a normal image and one where a colored filter is placed in front of the camera. The relationship between those two images is then used to estimate the scene illuminant.

The idea of using colored filters to improve vision tasks is not new. In optometry, chromagen lenses are used to subjectively improve the quality of color vision for color-blind observers <sup>[21]</sup> and special colored filters are also used to improve the reading speed of some dyslexic patients <sup>[22]</sup>. In the case of color constancy, pairs of images taken with and without flash can be used to estimate the original scene

illuminant [23],[24],[25], though using an additional light source such as a flash requires a “controlled” environment: the object must not be too close or too far, and the ambient light must not be too dim or too bright.

The rest of this section analyses the chromagenic theory, first by detailing the chromagenic illuminant estimation procedure and then by discussing specifically designed filters and sensors.

### A. Chromagenic Illuminant Estimation

The standard chromagenic illuminant estimation algorithm proceeds as follows: Let  $S(\lambda)$  be the descriptor of surface reflectances,  $E(\lambda)$  the scene illuminant Spectral Power Distribution (SPD),  $Q_k(\lambda)$  the camera sensitivities (we consider here trichromatic cameras, so  $k = \{R, G, B\}$ ) and  $F(\lambda)$  be the transmittance of the color filter placed in front of the camera.

The sensor responses of the unfiltered,  $\underline{\rho}$ , and filtered,  $\underline{\rho}^F$ , image can be written as:

$$\rho_k = \int_{\omega} E(\lambda)S(\lambda)Q_k(\lambda)d\lambda \quad (1)$$

$$\rho_k^F = \int_{\omega} E(\lambda)S(\lambda)F(\lambda)Q_k(\lambda)d\lambda \quad (2)$$

Thus, there are, for each scene, six responses per pixel that form the input to the illuminant estimation problem.

Naively, one might expect that since images are 3-dimensional (red, green, and blue for conventional cameras), then two images taken under different conditions (in this case, using a filter) form the basis of a 6-d space. However, if the transmittance of the filter is known, so is the relationship between the two images; the two images do not provide independent information.

Let us first consider the equations of filtered and unfiltered image formation (1) and (2). The filtered image can be expressed using a second illuminant,  $E^F(\lambda)$ , equivalent to putting the filter  $F(\lambda)$  in front of the light source  $E(\lambda)$ , i.e.,  $E^F(\lambda) = F(\lambda)E(\lambda)$ .  $\underline{\rho}$  and  $\underline{\rho}^F$  can therefore be considered as the sensor responses of a single surface under two different illuminants. It has been shown [26],[27] that when the same surfaces are viewed under two lights, the corresponding RGBs can, to a good approximation,

be related by a linear transform and so we use a  $3 \times 3$  matrix to relate the RGBs captured with and without the colored filter, thus:

$$\underline{\rho}^F = \mathcal{T} \underline{\rho} \quad (3)$$

where  $\mathcal{T}$  is a  $3 \times 3$  linear transform that depends on both the chromagenic filter and the scene illuminant.

Equation (3) implies that, given the chromagenic filter and sensor responses under a known illuminant, one can predict the filtered responses. In the problem of illuminant estimation, however, only the filtered and unfiltered responses are known, not the illuminant. Moreover, the task of finding the illuminant corresponds to finding  $\mathcal{T}$ . If we know all possible illuminants *a priori* we can, for a given filter: determine the transforms  $\mathcal{T}$  for every illuminant, estimate which of these pre-computed transforms best fits the pair of filtered-unfiltered responses, and thus, determine the illuminant.

Before outlining the actual algorithm, it is worth pointing out two cases where, depending on the filter or the sensor sensitivities, chromagenic color constancy is not possible: if the filter has a neutral density or if the camera sensors behave like Dirac delta functions.

If the chosen filter has a neutral density, i.e., its transmittance does not vary across the spectrum, the relationship between filtered and unfiltered RGBs will be a constant scaling (the same for all lights). This property can be written as:

$$F(\lambda) = \alpha, \forall \lambda \quad (4)$$

where  $\alpha$  is a constant value. It follows that:

$$\underline{\rho}^F = \alpha \underline{\rho}, \forall S, E \quad (5)$$

Consequently, the 6D responses will in fact span only three dimensions and thus the chromagenic algorithm cannot solve for color constancy.

If the  $k^{\text{th}}$  sensor behaves like a Dirac function whose non-null response is at the wavelength  $\lambda_k$ ,

equations (1) and (2) become:

$$\rho_k = E(\lambda_k)S(\lambda_k)Q_k(\lambda_k) \quad (6)$$

$$\rho_k^F = E(\lambda_k)S(\lambda_k)F(\lambda_k)Q_k(\lambda_k) \quad (7)$$

It follows that  $\rho_k^F = F(\lambda_k)\rho_k$  and so the responses are, again, three dimensional and their relation depends neither on the reflectances nor on the scene illuminant.

Additionally, while not as limiting as the neutral density case, using a rank-deficient filter will deliver poor chromagenic color constancy since significant information is lost (e.g., by using a deep red filter one loses all information about the blue pixels).

Barring the cases outlined above, the transforms can be pre-computed by choosing a set of  $n$  typical scene illuminants:  $E_i(\lambda)$ ,  $i = 1, \dots, n$  and a set of  $m$  surface reflectances:  $S_j(\lambda)$ ,  $j = 1, \dots, m$  representative of the real world. For each illuminant  $i$ , we create a  $3 \times m$  matrix  $Q_i$  whose  $j^{\text{th}}$  column contains the sensor response of the  $j^{\text{th}}$  surface under the the  $i^{\text{th}}$  illuminant. We also create  $Q_i^F$ , which contains the equivalent filtered responses. For each illuminant, we can then define the transform matrix as:

$$\mathcal{T}_i = Q_i^F Q_i^+ \quad (8)$$

where  $+$  denotes the Moore-Penrose pseudo-inverse operator:  $Q^+ = Q^T(QQ^T)^{-1}$ .  $\mathcal{T}_i$  can thus be described as the transform that best maps, in a least square sense, unfiltered to filtered responses under illuminant  $i$ . Because it is a least squares fit,  $\mathcal{T}_i$  will not, in practice, map the responses without errors. What matters, however, is that the error committed when mapping responses under illuminant  $i$  is smallest when the corresponding transform  $\mathcal{T}_i$  is used.

Once the  $n$  transforms have been pre-computed, the illuminant estimation proceeds as follows: let  $Q$  and  $Q^F$  denote the  $3 \times m$  matrices of unfiltered and filtered RGBs of arbitrary reflectances under

an unknown light. For each plausible illuminant we calculate the fitting error,  $e_i$ , as:

$$e_i = \|\mathcal{T}_i \mathcal{Q} - \mathcal{Q}^F\|, \quad i = 1, \dots, n \quad (9)$$

under the assumption that  $E_i(\lambda)$  is the actual scene illuminant. The transform that minimizes the error is surmised to be the scene illuminant; the estimated illuminant is  $E_{\text{est}}(\lambda)$  where

$$\text{est} = \arg \min_i (e_i) \quad i = 1, \dots, n \quad (10)$$

It can be shown that if *both* reflectances and illuminants are *exactly* described by three basis functions each, i.e., they are three dimensional, then the chromagenic algorithm delivers perfect illuminant estimation. In natural scenes, however, these dimensions are generally higher <sup>[28],[29]</sup> and so estimation errors ensue.

While in general the chromagenic algorithm can deliver good color constancy, it has two major weaknesses: the first one is that, though good on *average*, the performance can, on occasion, be (very) poor. The second problem comes from equation (9), which implies that the fitting error for each candidate illuminant is evaluated on a per-pixel basis. For the algorithm to deliver an optimal performance, the two images therefore have to be perfectly registered, a demanding requirement when images are taken outside of the lab. Registration algorithms can be of some help, but since an exact registration at pixel-level is necessary, they may not be sufficient.

### B. The Choice of Filter and Sensors

Two important aspects of the chromagenic algorithm are the filter choice and the camera sensors, for they can greatly influence the performance of the algorithm and are the only “controllable” parameters in the image formation process.

Given a filter set (in our case the filter set is a selection of 53 Kodak Wratten filters <sup>[30]</sup>) we need, for chromagenic purposes, to select one that is non-cutoff and whose transmittance varies across the

visible spectrum. Among the then possible filters, it was shown <sup>[16]</sup> that, they all deliver a good level of performance in practice. Finlayson et al.<sup>[31]</sup> have, however, reported that designing a filter specifically for chromagenic processing gave, on average, better results than existing Wratten filters. Since the filters were only simulated, not physically created, our experiments will be carried out using an actual Wratten filter.

Arguably, one might also design sensors for chromagenic illuminant estimation. But, pragmatically, when choosing sensors one has to consider aspects of image quality such as color rendering and image noise that strongly depend on the sensors. So, although optimal sensors for chromagenic illuminant estimation have been discussed <sup>[31]</sup>, we will work with conventional camera sensors.

### III. THE BRIGHT-CHROMAGENIC ALGORITHM

The previous section mentioned that the average performance of the chromagenic algorithm can be enhanced using specific filters and sensors. These potential improvements do not, however, address the main limitations of the chromagenic algorithm: possible large estimation errors and the need for perfectly registered images.

Looking back at equations (1) and (2), we see that the sensor responses depend on the scene illuminant, the chromagenic filter, the sensor sensitivities, and the surface reflectances. It follows that the linear transforms  $\mathcal{T}_i$  must also depend -to some degree- on those factors. Among them, the illuminant is the unknown; achievable improvements due to choosing both the filter and the sensors sensitivities were mentioned previously. The only remaining variable in the equation is therefore the scene reflectances.

Building a model based on reflectances can be difficult for one has, in general, no control over which reflectances are present in a scene. This uncertainty is the reason why simple estimation methods such as grey-world and Max-RGB are unreliable; if every scene contained at least a maximally reflective surface per color channel (e.g., at least a blue and a yellow surface), Max-RGB would be very accurate. An additional difficulty is that the input to illuminant estimation algorithms are not reflectances but

RGBs, which are composed of all of the image formation process' parameters.

To circumvent these issues, we start by quantifying the influence of reflectances on the transforms  $\mathcal{T}_i$ . We then model the best and worst performing RGBs. The results lead us to formulate the bright-chromagenic algorithm, which has three properties: it improves the average illuminant estimation performance, it reduces the maximal errors observed when the estimation is erroneous and, more importantly, it allows the algorithm to be used on unregistered images.

### A. Reflectances Analysis

For the chromagenic algorithm to work well, the transforms  $\mathcal{T}_i$  that map RGBs to their filtered counterparts should depend as much as possible on the illuminant. Here, we want to quantify the variance of the transforms when the illuminant changes and compare it to the variance observed when the illuminant is fixed but the reflectances vary.

To perform this assessment, we follow existing methodology <sup>[31],[16]</sup> in our choice of parameters. The illuminants belong to a set of 87 measured illuminant SPDs that include most common light sources. These SPDs are sampled every 10nm, from 380nm to 780nm. Barnard et al. <sup>[18]</sup> provide more details about this set, while the set itself is available online <sup>[32]</sup>. For surface reflectances, we use a synthesized set of 1995 Munsell surface reflectances<sup>[33]</sup>. The reflectances are also sampled every 10nm from 380nm to 780nm, more details about that set can be found in Parkkinen and Jaaskelainen <sup>[28]</sup>.

Concerning the choice of camera sensitivities and filter, we use the sensors of a Sony DXC-930 camera <sup>[18],[20]</sup> and a Wratten 81B filter (a yellowish filter). Both the filter and the sensor sensitivities are shown in Fig. 1.

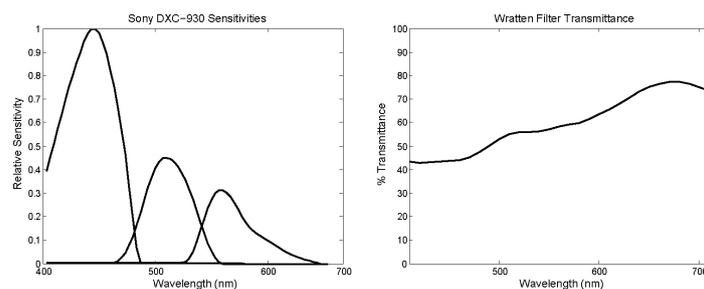


Fig. 1. Left: The Sony DXC-930 sensitivities. Right: The transmittance of the 81B Wratten filter used in the experiments.

We start by creating the transforms  $\mathcal{T}_i$  according to equation (8) by imaging all the 1995 reflectances under the 87 illuminants, and thus have 87 distinct transforms. The variability of the transforms, i.e., how differently they map the reflectances depending on the illuminant, is assessed with the inter-transform variance  $\sigma_E^2$ :

$$\sigma_E^2 = \frac{1}{87} \sum_{i=1}^{87} (\underline{t}_i - \underline{\mu}_T)^2 \quad (11)$$

where  $\underline{t}_i$  is the  $9 \times 1$  vector representation of the  $3 \times 3$  transform  $\mathcal{T}_i$  and  $\underline{\mu}_T$  is the average of all  $\underline{t}_i$ . Equation (11) quantifies the transforms variation across illuminants: the larger the variance, the more discriminative the transforms, and so the better the algorithm will perform. For a better perspective, one must compare  $\sigma_E^2$  with the variation in transforms due to choosing different reflectance sets.

Calculating the equivalent variance generated by different reflectances is somewhat more complicated. Let  $\mathbf{S}$  denote the entire reflectance set; a partition of  $\mathbf{S}$  into  $N$  subsets of equal size,  $s_j$ , can be written as:

$$\bigcup_{j=1}^N s_j = \mathbf{S} \quad (12)$$

and

$$s_j \cap s_k = \emptyset, \quad \forall j, k \in [1, N], \quad j \neq k \quad (13)$$

Let  $\mathcal{T}_i^{s_j}$  be the transform obtained with equation (8) when the subset  $s_j$  is imaged under illuminant  $i$ .

The inter-transform variance for reflectances  $\sigma_S^2$  is calculated as:

$$\sigma_S^2 = \frac{1}{87} \sum_{i=1}^{87} \sigma_{S_i}^2 \quad (14)$$

where:

$$\sigma_{S_i}^2 = \frac{1}{N} \sum_{j=1}^N (\underline{t}_i^{s_j} - \underline{\mu}_T^{S_i})^2 \quad (15)$$

In this formulation,  $\underline{\mu}_T^{S_i}$  is the mean of all  $\underline{t}_i^{s_j}$ , i.e., the mean of all subset-induced transforms under illuminant  $i$ .

An important aspect of this calculation is how to partition  $\mathbf{S}$ . On one hand, there should be enough subsets for the test to be meaningful, but since the chromagenic algorithm assumes three-dimensional signals, too small subsets will introduce large errors. While most daylight illuminants are at least 3D, this is not necessarily the case with reflectances, so, a combination of reflectances is required. Our tests, performed using subsets from 1 to 256 reflectances indicate (Fig. 2) that subsets of 16 reflectances form a good compromise. Instead of seeking every possible subset (there are  $C_{16}^{1995}$  of them), we repeat the 16 reflectances test 10'000 times and average the results over all the observations of  $\sigma_S^2$ . The results of this experiment are:  $\sigma_E^2 = 0.0306$  and  $\sigma_S^2 = 0.0753$ .

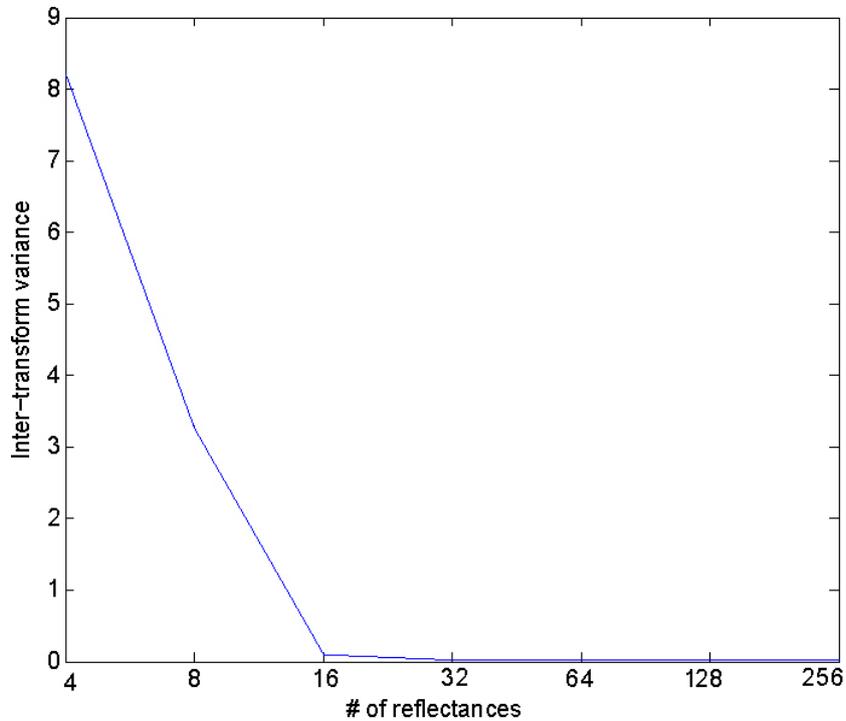


Fig. 2. *Inter-transform reflectance variance,  $\sigma_S^2$ , using a synthetic test with multiple reflectances from the Munsell set under a given illuminant. The values shown are the average over 10'000 tests.*

Based on these results, we can conclude that the linear transforms used in the chromagenic algorithm vary significantly with the reflectances used in training. It follows that there will also be a significant variability in testing. A subset of reflectances that are better suited to illuminant estimation must therefore exist, and the performance will increase the more of these “good” reflectances are present in the scene.

### B. Modeling In- and Outliers

Estimation accuracy will partly depend on the set of tested reflectances: which ones are “better”? Because there is a multitude of illuminant-reflectance combinations, we will analyze the performance of the algorithm on scenes with a single reflectance (where good estimates can still be obtained [16]). For each scene we estimate the RGB of the light and calculate the angle to the actual RGB of the illuminant.

The transforms  $\mathcal{T}_i$  are calculated as before, thus creating 87 of them. The test set for this experiment consists of all possible single reflectance-illumination combinations. A larger set of 287 illuminants is used, together with the 1995 reflectances, i.e., there are  $\sim 570,000$  pairs of filtered and unfiltered RGBs. The larger illuminant set used in testing covers the same gamut as the 87 training lights; the chromagenic algorithm will select one of the 87 lights as the scene illuminant.

Figure 3 shows the angular errors (i.e., the angle between the RGBs of the estimated and actual illuminant) for all the 570,000 pairs. The angular errors range from 0 to 42 degrees, with a mean of 9.3 and a median of 5. For this particular dataset, our experiments indicate that an angular error of 3 degrees or less is necessary for acceptable color cast removal. From these results, it follows that one needs reduced overall and, especially, maximum errors -an angular error of 42 degrees is equivalent to mistaking green for yellow.

To understand what is happening, we look at the RGBs that comprise the top and bottom 20% of the error, which corresponds to angular errors of 0-2.3 degrees and 20.5-42 degrees for each group, respectively. A brightness-saturation (i.e., S and V of the HSV space) scatter plot of these RGBs is shown in Fig. 4a for the highest errors and Fig. 4b for the lowest ones, we also analyzed the behavior of the algorithm with respect to hue but found no significant hue dependency. It is clear that low errors correlate with fairly de-saturated RGBs (pastel tones and achromatic) whereas high errors correlate with dark and saturated RGBs.

This finding is corroborated by the result of Fig. 5, which displays elliptical summaries for the high and low error sets. The ellipses, each of which accounts for 90% of its respective data, overlap little.

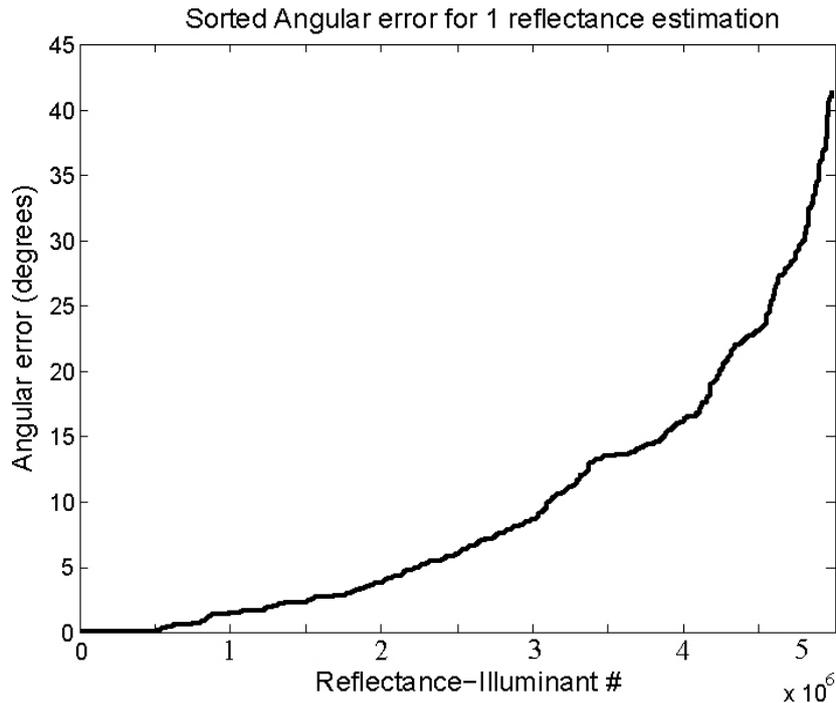


Fig. 3. Sorted errors for the single reflectance test. The mean error value is 9.3 degrees and the median 5 degrees.

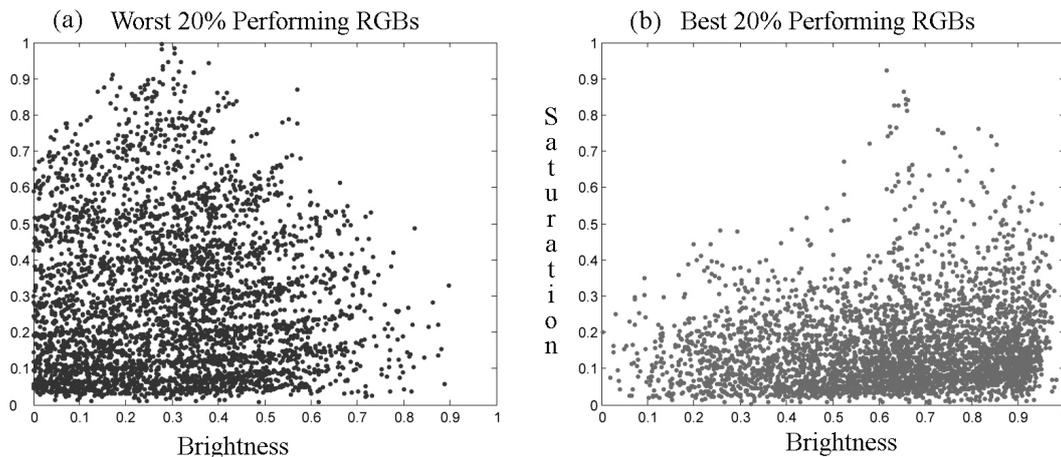


Fig. 4. (a) Brightness-Saturation scatter plot of the 20% worst performing RGBs. (b) Brightness-Saturation scatter plot of the 20% best performing RGBs.

Importantly, the set of good-performing reflectances is quite large and diverse. The main characteristic they have in common is their low saturation value. Before proceeding further, one must decide how these well-performing reflectances will be chosen, and several parameters have to be taken into account. The first one is that the good reflectances should be easy to pick out, so to provide stability to the algorithm's estimation. The second concern is that the easy-to-pick-out good reflectances should be reliably present within natural scenes. Finally, one wants to avoid dark colors because of their lower signal-to-noise ratio. Considering the required parameters, we chose to use bright-achromatic

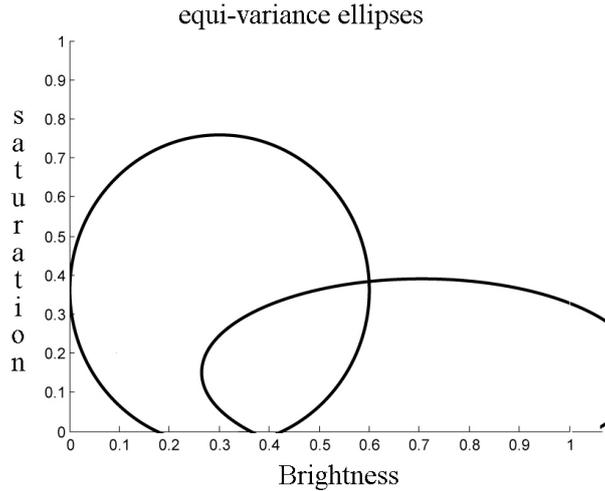


Fig. 5. *equi-variance ellipses of both sets, each containing 90% of their respective data, showing they are mostly disjoint.*

reflectances in our algorithm.

The idea of using brighter reflectances has been discussed before, notably by Tominaga et al.<sup>[19]</sup> where brighter reflectances were preferred because of their higher signal-to-noise ratio that would provide better illuminant classification. We reach a similar conclusion, but for different reasons. Indeed, the good-performing reflectances, for chromagenic color constancy, are primarily desaturated colors and are not specifically brightness-related. Indeed, we use the “brightest” colors because their RGB values are most different from the colors that perform worst, i.e., a simple thresholding scheme enables us to reliably find suitable RGBs for accurate illuminant estimation. Moreover, using the brightest surfaces has the additional advantage of permitting to remove the image registration constraint, a point we will come back to in the rest of the paper.

Assuming a uniform distribution of colors in an image, we propose that it is easy to find RGBs and their filtered counterparts that belong to this preferred set. We simply look for a small percentage of the brightest image regions. We therefore modify the chromagenic algorithm formulation so that only bright image pixels are considered.

### C. The algorithm.

Thus, the **Bright-Chromagenic** algorithm is defined as:

---

**Pre-processing:** For a database of  $m$  lights  $E_i(\lambda)$  and  $n$  surfaces  $S_j(\lambda)$  calculate  $\mathcal{T}_i \approx \mathcal{Q}_i^F \mathcal{Q}_i^+$  where  $\mathcal{Q}_i$  and  $\mathcal{Q}_i^F$  represent the matrices of unfiltered and filtered sensor responses to the  $n$  surfaces under the  $i$ th light and  $+$  denotes a *pseudo-type-inverse*

**Operation:**  $P$  surfaces in an image yield  $3 \times P$  matrices  $Q$  and  $Q^F$ . From these matrices a certain percentage of the brightest pixels is chosen, giving the matrices  $\mathcal{Q}$  and  $\mathcal{Q}^F$ , where the brightest pixels are the ones with the largest  $R^2 + G^2 + B^2$  value. Then the estimate of the scene illuminant is  $\rho_{E_{est}}$  where

$$est = \arg \min_i (err_i) \quad (i = 1, 2, \dots, m)$$

and

$$err_i = \|\mathcal{T}_i \mathcal{Q} - \mathcal{Q}^F\|$$

---

Because we are proposing to look only at bright image responses, the transform matrices might be calculated using a least-squares estimator where bright values are weighted more strongly. This is what is meant by a *pseudo-type-inverse*. However, subsequent experiments did not indicate any tangible benefits from building transforms using only the bright image RGBs. So, for the experiments presented in the next section, the conventional (unweighted) Moore-Penrose inverse is used.

The bright-chromagenic formulation is robust since it does not make assumptions about which reflectances might or might not be present in the scene, i.e., if there are no bright reflectances in the image, the bright-chromagenic algorithm will still have an equivalent performance to the original chromagenic algorithm.

Because we “exclude” -select them only if no other are available- the worst performing RGBs, we expect the bright chromagenic algorithm to significantly reduce the worst errors.

Moreover, if the scenes admit a diversity of reflectances, then it follows that (if the filter does not vary

too drastically across the spectrum) the brightest unfiltered RGBs will be mapped onto the brightest filtered RGBs. If one is relatively conservative with the number of bright pixels used to estimate the illuminant (we typically use the top 1-3% of the brightest pixels<sup>1</sup>), the bright-chromagenic algorithm will then be able to estimate illuminants even when the images are not registered. Both these properties are verified in our experiments. We stress here that no assumptions are made on the scene content; the brightest pixels are sought whatever they may be, i.e., we do not need the presence of a white-like surface in the scene.

#### IV. EXPERIMENTS

In this section, we analyze the performance of the bright-chromagenic algorithm, and compare it to various other illuminant estimation methods on four datasets of increasing difficulty: synthetic reflectances and lights, real reflectances and synthetic lights, real images taken in a controlled environment, and real-world images taken with “unknown” filter and camera. For the first three tasks, the bright-chromagenic algorithm is compared to other illuminant estimation methods whose results on the same datasets have been reported (we took the results directly from the referenced published papers). Concerning the fourth dataset, the comparison was carried out with readily implementable algorithms.

The algorithms are evaluated according to the framework of Hordley and Finlayson [20] where it was shown that, if one wants to summarize the performance of an illuminant estimation algorithm over a dataset, one should prefer the median angular over the mean or Root Mean Square error. The mean error is also reported for the SFU dataset because of its prevalence in the literature.

Angular error is an intensity independent error measure that is widely used in the literature [18],[34],[20]. It is the measure between the sensor responses of a white reflectance under both the estimated and actual scene illuminant. Let these responses be  $\rho_{E_{\text{est}}}$  and  $\rho_E$  respectively, the angular error  $e_{\text{Ang}}$  is

<sup>1</sup>To avoid saturated pixels, we use the brightest pixels of *non-maximal* value (that is, the ones for which R, G, and B < 255)

calculated as:

$$e_{\text{Ang}} = \arccos\left(\frac{\rho_{\text{E}}^{\text{T}} \rho_{\text{E}_{\text{est}}}}{\|\rho_{\text{E}}\| \|\rho_{\text{E}_{\text{est}}}\|}\right) \quad (16)$$

The use of a median statistic permits to assess if the difference of performance between two algorithms is statistically significant at chosen confidence level. That significance is given by using the Wilcoxon unranked sign test [35] at a 95% confidence level.

To simplify the writing, we will use the following notations:  $S_M$  is the set of 1995 synthesized Munsell reflectances [33],  $E_{87}$  and  $E_{287}$  are the sets of respectively 87 and 287 illuminants [32].

### A. Synthetic Reflectances and Lights

The test on synthetic images is run according to the testing protocol proposed by Barnard et al. [18], detailed hereafter.

**Training:** The linear transforms are created by imaging the whole of  $S_M$  under  $E_{87}$ , thus generating 87 transforms. We use the Sony-DXC camera sensitivities and the 81B Wratten filter, whose transmittance is shown in Fig. 1.

**Testing:** 1000 images containing  $n$  different reflectances are generated, where  $n = \{1, 2, 4, 8, 16, 32\}$ ; these reflectances are randomly taken from  $S_M$ . We then illuminate these images with one light taken at random from  $E_{287}$ ; sample filtered and unfiltered images are shown in Fig. 6.

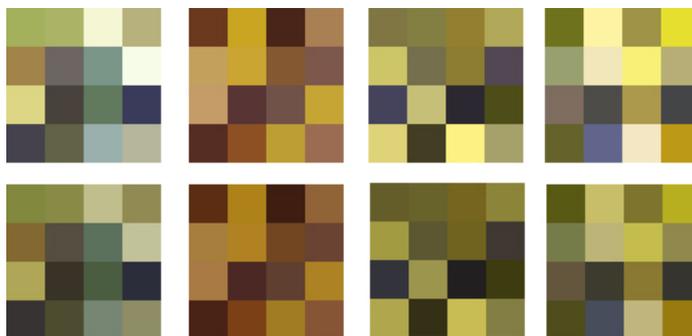


Fig. 6. Sample pairs of synthetic images with 16 reflectances under random illuminants from  $E_{287}$

The illuminant of each image is estimated using both the bright-chromagenic and the original chromagenic algorithms. For images where  $n > 4$ , the bright-chromagenic version estimates the illuminant based on the four brightest reflectances only.

# reflectances	1	2	4	8	16	32	rank
Chromagenic	6	5.2	4.5	3.5	3	2.2	2
Max RGB	9.7	7.9	6.1	4	2.9	2.6	6
Grey World	9.1	7.3	5.8	4.9	4.8	4.8	8
Database GW	9.5	6.7	4.8	3.4	2.8	2.5	4
LP GM	9.6	6.7	4.8	3.3	2.7	2.4	4
Neural Network	8.8	7.1	5	4	2.9	2.6	6
Colour by Corr.	6.9	5	3.5	3.1	2.4	2.3	2
Bright-Chroma.	6	5.2	4.5	2.8	2.1	0.9	<b>1</b>

TABLE I

*Average median angular error for 1000 tests at each complexity level. The last column is the rank, based on the 32 surfaces test, according to the Wilcoxon sign test with a confidence level of 95% (as in [20])*

The results are displayed in Table 1, where the last column indicates the ranking of the considered algorithms, according to Wilcoxon’s sign test. An algorithm is ranked better than another if its median is lower *and* if the difference is statistically significant at the 95% level. If the sign test is inconclusive, the algorithms will be ranked equally.

The results show that the bright-chromagenic algorithm performs significantly better than all other methods. Noticeably, the more complex methods form a group that is, in turn, significantly better than the simpler scene assumptions algorithms.

An additional result, shown in Fig. 7, is the reduction in maximal error achieved by the bright chromagenic algorithm. This experiment validates our selection of the bright RGBs to reduce the high maximal errors observed with the original chromagenic algorithm.

### B. Real Reflectances and Synthetic Lights

In this second test, we use spectral outdoor reflectance images measured -as opposed to synthesized- by Nascimento et al. [36]. Figure 8 shows the eight images that will be used (the images can be obtained

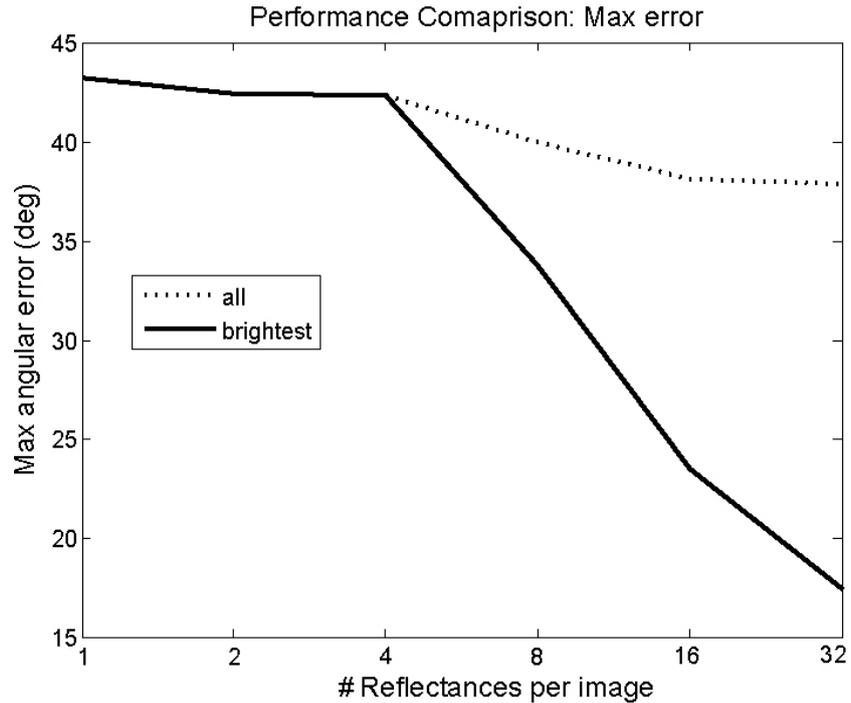


Fig. 7. Comparison of the max angular error between the original and the bright chromagenic algorithm, one can see the significant reduction achieved by selecting only the brightest RGBs. From 1 to 4 reflectances, the max error is identical since there are not enough reflectances to distinguish both algorithms.

online [37]).

Note these images measure only reflectances. To generate RGB images, we use the scene lights and camera sensors as in the previous experiment.



Fig. 8. The eight reflectance images measured by Nascimento et al. that compose the second test.

**Training:** The training step is unchanged from the previous experiment, i.e., the transforms are created on a different reflectance set than the one tested.

**Testing:** Images are created by illuminating the eight reflectance scenes with  $E_{287}$ , which generates 2,296 “half-synthetic” images. We then proceed to test our algorithm on each of those images, selecting the top 3% of the brightest, non-saturated, pixels to estimate the illuminant.

Algorithm	Median	Rank
Chromagenic	6.7	2
Max RGB	8.7	3
Grey World	13	4
Bright Chromagenic	3.5	1

TABLE II

*Summary of the results on the Nascimento images. This table displays the mean angular error values over the 2,296 images. A '+' in a row indicate that an algorithm performs significantly better (at 95% confidence level) than the one in the corresponding column, '-' and '=' are for when an algorithm performs worse or if there is no significant difference.*

Results from this dataset, Table 2, exhibit the same trend than in the all synthetic experiment, i.e., the bright-chromagenic significantly outperforms other methods.

### C. SFU dataset

The next set is the non-specular Simon Fraser University (SFU) dataset, described in detail by Barnard et al. [18].

The data set consists of 31 colorful objects captured under 11 illuminants. Figures 9 and 10 show some objects under one light and one object under all the available lights respectively. In the second case it is apparent that the images are not registered (in fact, the objects were rotated in between two pictures when creating the dataset).

This experiment differs from the previous ones because only the RGBs of the images (as opposed to reflectances) are provided. This, plus the non-registration of the image will provide a difficult test for the bright-chromagenic algorithm. The SFU dataset has been used in several illuminant estimation comparisons because ground truth is provided in addition to the images themselves. That is, both the SPD of the 11 illuminants (they are actually a subset of  $E_{87}$ ) and the camera sensitivities are given (the camera used to take the images is the Sony-DXC 930 whose sensitivities are shown in Fig. 1 and that was used in the previous tests).

Performing chromagenic illuminant estimation requires pairs of images taken with and without a

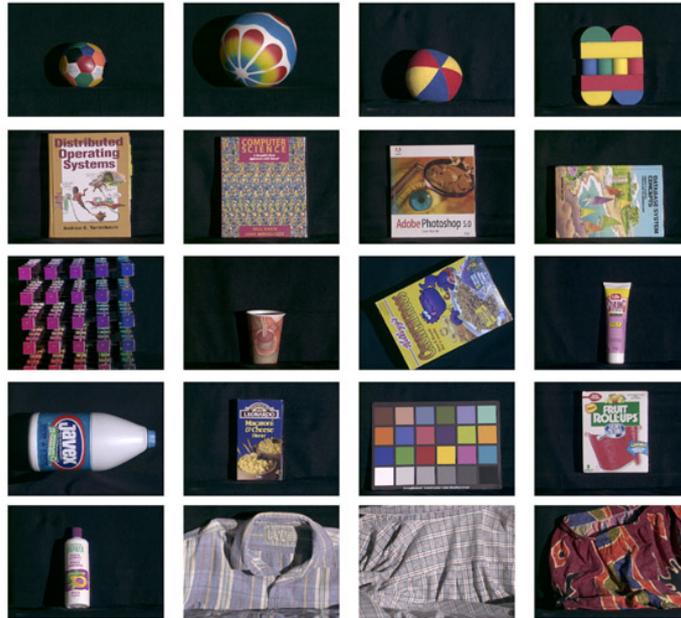


Fig. 9. Some objects of the SFU dataset under one illuminant.



Fig. 10. One object from the SFU dataset under the 11 illuminants. Note that the lights in the second row are the same than the ones of the first row with a bluish filter placed in front of them. One can also see that the images are not registered.

colored filter. However, with only image RGBs available, one cannot retrospectively model the filtered responses. As it turns out, 8 out of the 11 illuminants present in the set come in pairs: the original lamp lights and those lights filtered with a blue filter. Since the actual illuminant SPDs are known, one can derive the filter that was used by dividing the spectra of the lights by the filtered ones. The eight -two pairs of four- lights that are considered and the derived filter are shown in Fig. 11.

**Training:** The transforms  $\mathcal{T}_i$  are obtained by imaging the *synthetic* reflectances  $S_M$  under the illuminants of  $E_{87}$ . The filter is the one derived from the eight illuminants and shown in Fig. 11; the camera sensitivities are the same as in previous experiments.

**Testing:** To test the algorithm, we estimate the illuminant of all the possible pairs of images (124 pairs

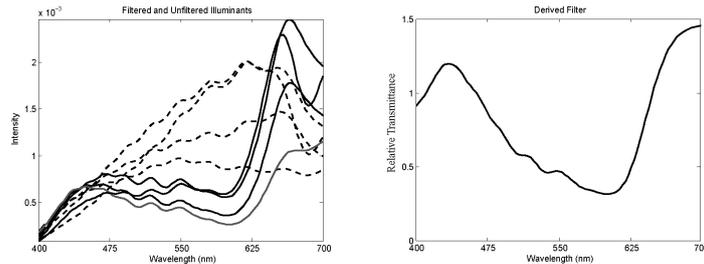


Fig. 11. *Left: The 8 light sources considered in this experiment. The dashed lines are spectra of the light sources, while the continuous ones are from the filtered sources. Right: The Filter derived from the light source data.*

Algorithm	Mean	Median	Rank
Bright-Chromagenic	4.8	3.4	<b>1</b>
Max RGB	6.4	4.1	4
Grey World	11.9	9.3	7
Database GW	10	7	5
Neural Network	8.9	7.8	5
LP Gamut Mapping	5.5	3.8	2
Colour by Corr.	6	3.6	2

TABLE III

*Mean and median angular errors over the SFU dataset. The ranks are significant at the 95% level.*

in total) using the top 3% of the brightest pixels in both filtered and unfiltered images independently. These pixels typically belong to one or two of the surfaces in the scene (we do not need a white reflectance *per se*, we simply use the brightest ones available). Since the images are not registered, it provides a test for the hypothesis that, in general, the brightest pixels in both images come from the same surfaces, and that the bright-chromagenic algorithm does away with the need for registration.

The angular errors reported in the first two columns of Table 3 show that, despite its simplicity, the bright chromagenic algorithm outperforms in terms of both mean and median angular error all other algorithms at the 95% confidence level. The original chromagenic algorithm is not shown here since its registration requirement is not fulfilled.

Perhaps the most remarkable aspect of the bright-chromagenic algorithm is that, despite modeling

the transforms on synthetic data with a filter derived from measurements, it is still able to accurately estimate the illuminants of real, significantly non-registered, images.

#### *D. Real Images*

The last experiment is designed to evaluate the performance of the bright-chromagenic algorithm *in situ*. Whereas the previous datasets were obtained in “controlled conditions” (purely synthetic data and controlled lighting environment), we use here a set of real-world images taken with a digital camera whose specifications are unknown.

#### **Chromagenic Photography:**

For the illuminant estimation to be meaningful, one must take a couple of precautions when capturing the images. The camera that was used is a Nikon D70: a Single Lens Reflex camera. The camera was set-up to capture linear Raw, unprocessed, images<sup>2</sup>. To prevent the camera from using a different white balance between filtered and unfiltered images, all images were captured with the white balance set to “daylight”.

Additional technical aspects to consider are that the image pairs should be as registered as possible (so that the original chromagenic algorithm can also be tested), and also devoid of over or under-exposed regions, which make the relationship between filtered and unfiltered pixels meaningless. To that effect, the images were captured using a tripod and a remote shutter release (to minimize registration errors) and the settings of the camera aperture and shutter speed were set to “manual” mode where we aimed to capture the entire dynamic range of the image.

For the filter, we used an actual 81B-type Wratten filter. The captured images were then exported, using Nikon capture, as 16-bits/channel linear tiff images.

The dataset consists of 86 pair of images taken under a variety of indoor and outdoor illuminants. In every scene, we placed a Macbeth color checker that is used to accurately determine the color of the prevailing light, thereby providing a ground truth to assess the accuracy of illuminant estimation

<sup>2</sup>In fact, even with Raw settings, the camera and associated software will process the image somewhat; it is however as unprocessed as one can have with a general purpose digital camera.

algorithms.

From the dataset, we then created separate training and testing sets. The training set consists of the 24 Macbeth patches present in all the images. The testing set is created by blacking the color checker from the images. Images from the original (with the color chart) and the testing set are shown in Fig. 12.

We note that, despite the precautions taken, the registration between images is not perfect and some image regions can be over-exposed. Additionally, multiple illumination is sometimes present in images, which can lead to errors when estimating the prevailing illuminant.



Fig. 12. *Top 2 rows: Examples of unfiltered/filtered image pairs from our real image set used in training. Bottom 2 rows: the testing images, where the charts have been cropped out.*

**Training:** We create 86 linear transforms, using the 24 RGBs of the color checker present in each image.

**Testing:** We estimate the illuminant for both the original and bright-chromagenic (using the top 3% brightest pixels) algorithms on the 86 pairs of images that have the color chart clipped out.

The results are shown in Table 3 and illustrate that the most accurate illuminant estimation is given by the bright-chromagenic algorithm. The results otherwise exhibit the same behavior than previous experiments.

Algorithm	Mean	Median	Rank
Bright-Chromagenic	7.09	4.15	<b>1</b>
Max RGB	7.87	7	3
Grey World	11	10.8	5
Chromagenic	7.96	5.1	2
$\ell^4$ Grey	10.3	9.6	4

TABLE IV

Mean and median angular errors over the 86 real images. The ranks are significant at the 95% level.

## V. CONCLUSION

A chromagenic illuminant estimation algorithm exploits the relationship between RGBs captured by a camera and those captured through a colored filter. Different lights induce different relationships and so, the illuminant color can be estimated by testing pre-computed relations *in situ*. While the chromagenic approach can work well, it occasionally performs poorly. Moreover, typical chromagenic camera embodiments, such as a stereo rig or multiple surveillance cameras (a filter can easily be placed over one camera) do not have pixel registration and this is assumed in the chromagenic theory.

In this paper, a detailed error analysis demonstrated that bright pixels in images lead to smaller chromagenic estimation errors. This led to the bright-chromagenic algorithm, which bases its estimation only on a fixed percentage of the brightest pixels in the filtered and unfiltered images. Importantly, these pixels are chosen independently in each image so there is no need for image registration. Experiments on various sets of synthetic and real data demonstrate that the bright-chromagenic algorithm delivers a statistically significant better illuminant estimation than all other tested algorithms. The performance is especially promising considering that, if the camera sensitivities are known, the transforms can be pre-computed on synthetic data even without knowing the content of the test scenes, such as demonstrated on the SFU test (third experiment).

## ACKNOWLEDGMENTS

The authors are grateful to Prof. Sabine Susstrunk, EPFL, for her help and dedication in improving the written quality of this manuscript.

## REFERENCES

- [1] L. Arend and A. Reeves, "Simultaneous color constancy," *Journal of The Optical Society of America part A*, vol. 3, pp. 1743–1751, 1986.
- [2] D. Brainard, "Color constancy in the nearly natural image: 2 achromatic loci," *Journal of The Optical Society of America part A*, vol. 15, pp. 307–325, 1998.
- [3] D. Brainard and W. Freeman, "Bayesian color constancy," *Journal of The Optical Society of America part A*, vol. 14, pp. 1393–1411, 1997.
- [4] H. Jiang and M. Drew, "Tracking objects with shadows.," in *CME03: International Conference on Multimedia and Expo.*, pp. 100–105, 2003.
- [5] M. J. Swain and D. H. Ballard, "Color indexing," *International Journal of Computer Vision*, vol. 7, pp. 11–32, 1991.
- [6] G. J. Klinker, S. A. Shafer, and T. Kanade, "A physical approach to color image understanding," *International Journal of Computer Vision*, vol. 4, pp. 7–38, 1990.
- [7] P. Hubel, J. Holm, G. Finlayson, and M. Drew, "Matrix calculations for digital photography," in *Proc. the Fifth color Imaging Conference*, pp. 105–111, 1997.
- [8] G. Buchsbaum, "A spatial processor model for object colour perception," *Journal of the Franklin Institute*, vol. 310, pp. 1–26, 1980.
- [9] G. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Proc. of the Twelfth Color Imaging Conference*, pp. 37–41, 2004.
- [10] J. van de Weijer and T. Gevers, "Color constancy based on the grey-edge hypothesis," in *Proc. of the International Conference on Image Processing, Genoa*, pp. 722–725, 2005.
- [11] V. Cardei, B. Funt, and K. Barnard, "Estimating the scene illuminant chromaticity using a neural network," *Journal of The Optical Society of America part A*, vol. 19, pp. 2374–2386, 2002.
- [12] G. Finlayson, S. Hordley, and P. Hubel, "Color by correlation: A simple, unifying framework for color constancy," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1209–1221, 2001.
- [13] G. Finlayson and R. Xu, "Convex programming colour constancy," in *IEEE Workshop on Color and Photometric Methods in Computer Vision*, pp. –, 2003.
- [14] D. Forsyth, "A novel algorithm for colour constancy," *Intl Journal of Computer Vision*, vol. 5, pp. 5–36, 1990.
- [15] S. Lin and S. Lee, "Detection of specularly using stereo in color and polarization space," *computer Vision and Image Understanding*, vol. 65, pp. 336–346, 1997.
- [16] G. Finlayson, S. Hordley, and P. Morovic, "Colour constancy using the chromagenic constraint.," in *Computer Vision and Pattern Recognition (CVPR) 2005*, pp. 1079–1086, 2005.

- [17] C. Fredembach and G. Finlayson, "The bright-chromagenic algorithm for illuminant estimation," in *Proc. of the IS & T 15th Color imaging conference*, pp. 137–142, 2007.
- [18] K. Barnard, V. Cardei, and B. Funt, "A comparison of computational color constancy algorithms- part i: methodology and experiments with synthesized data," *IEEE Trans. on Image Processing*, vol. 11, pp. 972–984, 2002.
- [19] S. Tominaga, S. Ebisui, and B. Wandell, "Scene illuminant classification: brighter is better," *Journal of The Optical Society of America part A*, vol. 18, pp. 55–64, 2001.
- [20] S. Hordley and G. Finlayson, "Reevaluation of color constancy algorithm performance," *Journal of the Optical Society of America A*, vol. 24, pp. 1008–1020, 2006.
- [21] N. B. Hodd, "Putting chromagen to the test," *Optometry today*, vol. 38, pp. 39–42, 1998.
- [22] A. Wilkins, *Reading through colour: How coloured filters can reduce reading difficulty, eye strain, and headaches*. John Wiley, 2003.
- [23] J. DiCarlo, F. Xiao, and B. Wandell, "Illuminating illumination," in *Proc. of the ninth Color Imaging Conference*, pp. 27–34, 2001.
- [24] G. Petschnigg, R. Szeliski, M. Agrawala, M. Cohen, H. Hoppe, and K. Toyama, "Digital photography with flash and no-flash image pairs," *ACM Trans. on Graphics*, vol. 23, pp. 664–672, 2004.
- [25] C. Lu and M. Drew, "Practical scene illuminant estimation via flash/no-flash pairs," in *Proc. of the fourteenth Color Imaging Conference*, pp. 1–1, 2006.
- [26] L. Maloney and B. Wandell, "Color constancy: A method for recovering surface spectral reflectance," *Journal of The Optical Society of America part A*, vol. 3, pp. 29–33, 1986.
- [27] M. D'Zmura and G. Iverson, "Color constancy i: Basic theory of two-stage linear recovery of spectral descriptors for lights and surfaces," *Journal of the Optical Society of America*, vol. 10, pp. 2148–2165, 1993.
- [28] J. Parkkinen and T. Jaaskelainen, "Characteristic spectra of munsell colors," *Journal of The Optical Society of America part A*, vol. 6, pp. 318–322, 1989.
- [29] D. Judd, D. MacAdam, and G. Wyszecki, "Spectral distribution of typical daylight as a function of correlated color temperature," *Advances in Neural Information Processing*, vol. 54, pp. 1031–1040, 1964.
- [30] Kodak, *Kodak Wratten Filters 4th Edition*. Kodak Limited London, 1969.
- [31] G. Finlayson, S. Hordley, and P. Morovic, "Chromagenic filter design," in *Proceedings of the 10th AIC*, pp. 1079–1083, 2005.
- [32] K. Barnard, "Data for computer vision and computational colour vision [online]." <http://www.cs.sfu.ca/~color/data>, 2002.
- [33] J. University, "Musell colors matt [online]." [http://spectral.joensuu.fi/databases/download/munsell\\_atof.htm](http://spectral.joensuu.fi/databases/download/munsell_atof.htm), 1989.
- [34] B. Funt, K. Barnard, and L. Martin, "Is machine color constancy good enough?," in *Proc. of the Fifth Color Imaging Conference*, pp. 455–459, 1998.
- [35] R. Hogg and E. Tanis, *probability and Statistical Inference*. Prentice Hall, 2001.
- [36] S. Nascimento, F. Ferreira, and D. Foster, "Statistics of spatial cone-excitation ratios in natural scenes," *Journal of the Optical Society of America A*, vol. 19, pp. 1484–1490, 2002.
- [37] M. University, "Hyperspectral images [online]." <http://personalpages.manchester.ac.uk/staff/david.foster/>, 2002.