Illuminant retrieval for fixed location cameras

Joanna Marguier and Sabine Süsstrunk
School of Computer and Communication Sciences, Ecole Polytechnique Fédérale de Lausanne (EPFL), Switzerland

Abstract

Fixed location cameras, such as panoramic cameras or surveillance cameras, are very common. In images taken with these cameras, there will be changes in lighting and dynamic image content, but there will also be constant objects in the background. We propose to solve for color constancy in this framework. We use a set of images to recover the scenes' illuminants using only a few surfaces present in the scene. Our method retrieves the illuminant in every image by minimizing the difference between the reflectance spectra of the redundant elements' surfaces or, more precisely, between their corresponding sensor response values. It is assumed that these spectra are constant across images taken under different illuminants. We also recover an estimate of the reflectance spectra of the selected elements. Experiments on synthetic and real images validate our method.

Introduction

Color constancy and illuminant retrieval have been an important subject of research for over 30 years and many approaches have been proposed. Mathematically, the problem of retrieving an illuminant spectra from triplets of RGB values is ill-posed, and it is necessary to introduce constraints and assumptions to reduce the number of unknowns in the equations. This can be done, for example, by introducing assumptions on the image content [10], statistics [1], the surface properties such as specular reflections [11], or by using redundancy in images ([6], [3], [12], and [17]).

We propose to solve for color constancy in the particular case of fixed location cameras, such as security or panoramic cameras. In this case, we have scene objects that will be present in every image, despite changing dynamic scene content and illumination. We will use these elements to reduce the dimensionality of the illuminant retrieval problem. We assume that the reflectances of those static objects, though unknown, remain constant.

For a set of N images, we select several elements in each scene and compute a set of values sampling all possible metameric pairs of illuminant and reflectance spectra corresponding to the pixel values of each patch. From these reflectance spectra, we compute sensor responses. By matching the sensor responses, we deduce the illuminants corresponding to the N images and we also obtain an estimate of the reflectances of the selected objects. We express all spectral quantities as sums of a limited numbers of basis functions and use a linear image formation model. We conducted experiments with synthetic and real images. Good results for illuminant retrieval are observed whenever the images present a sufficient variety of illuminants.

This article is structured as follows: in the next section we present background information on color constancy, then we describe our approach in detail, report the results, and conclude the article.

Background

There are many approaches to color constancy, here we discuss the algorithms most closely related to our method. Several methods were proposed to retrieve the illuminant in a scene using linear models for the reflectance and illuminant spectra. The dimension of linear problems can be reduced by expressing the illuminant and reflectance spectra as weighted sums of basis functions ([9], [2], [16]). Determining the spectra reduces to determining these weighting coefficients.

Buchsbaum assumes the average reflectance of a scene to be constant to retrieve the illuminant [1]. In [7], Gershon et al. modify this method by adding information about the properties of surfaces likely to appear in a scene. The scene illuminant is found by mapping scene color values onto an *ideal material space* formed by the descriptors of "all" surface reflectances. In [14], Maloney and Wandel assume the number of sensors to be superior to the number of degrees of freedom of the reflectance descriptors. In this framework, they retrieve the illuminant by inversion of a linear model. If we apply this approach to the case of an RGB camera, we have three sensors and can thus approximate the reflectances by a sum of only two functions, which is not accurate enough. It has indeed been shown in [2] and [13] that most natural and man-made surfaces have reflectance spectra that can be represented by a linear combination of 6 to 8 basis functions. In [20], D'Zmura and Iverson present a two-stage-recovery bilinear model for the retrieval of both illuminant and reflectance spectra. They investigate the solvability condition of a model using several surfaces imaged under different illuminants. They combine the information on all surfaces viewed under all illuminants into one system of linear equations. They first recover the reflectance descriptors by inverting this linear system and then use the resulting descriptors to compute the illuminants. In our approach, we solve in parallel several linear systems, one for each individual surface viewed under each test illuminant. We force the resulting reflectance spectra to be equal, which allows us to deduce the illuminants.

Other approaches using pairs of images of one scene under different illuminants have been proposed. For example, the flash-no flash method uses two images of one scene taken with and without flash ([3], [12], and [17]). Knowing the flash spectra, the authors can retrieve the non-flash illuminant. The chromagenic method [6] offers a similar approach. A scene is imaged with and without a colored filter placed in front of the camera. Knowing the filter transmittance, the authors derive the illuminant of the unfiltered scene among a set of test illuminants.

The main difference between our method and approaches using pairs of redundant images such as the flash-no flash or chromagenic approach is that we only use the redundancy of reflectances rather than information on the difference between two images (the flash spectral power distribution in the flash-no flash approach or the colored filter transmittance in the chromagenic

approach). Another difference is the number of input images required to estimate the illuminants: these methods, when considering uniform illuminants, retrieve one illuminant using two images, an original and a modified image; whereas our method allows retrieving N illuminants in N input images. Also, our input images do not need to represent the same scene, but only to contain several objects present in every scene. Moreover, we obtain an estimate of the reflectance of these objects.

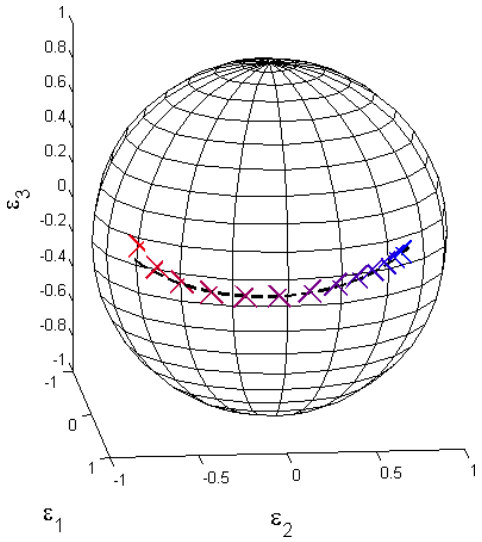

Figure 1. *When represented on a sphere, daylight illuminants can be approximated by a quadratic function (dashed line). The 13 daylight illuminants D40 to D100 are represented, from left to right, by the 'x'.*

Our approach

We express the daylight and reflectance spectra as weighted sums of 3 and 8 basis functions, respectively. The daylight illuminants are approximated as

$$E(\lambda) = \sum_{i=1}^{3} \varepsilon_i \mathscr{E}_i(\lambda)$$

where $\varepsilon_1 = 1$. $\mathscr{E}_1(\lambda) = \overline{\mathscr{E}}$ and $\mathscr{E}_i$ are, respectively, the mean and first two principal components of the standard CIE daylight illuminants D40 to D100. Using only three basis functions allows us to represent any illuminant after normalization by a point on a sphere. When plotting the CIE daylight illuminants D40 to D100, we see that the points lie on a curve that can be approximated by a quadratic function (see Figure 1). We can thus represent any illuminant using only one parameter, the azimuthal angle θ. The reflectances are expressed as

$$S(\lambda) = \sum_{i=1}^{8} \sigma_i \mathscr{S}_i(\lambda)$$

where the basis functions $\mathscr{S}_i(\lambda)$ are obtained by principal component analysis of the Munsell, MacBeth, and natural reflectance spectra [21].

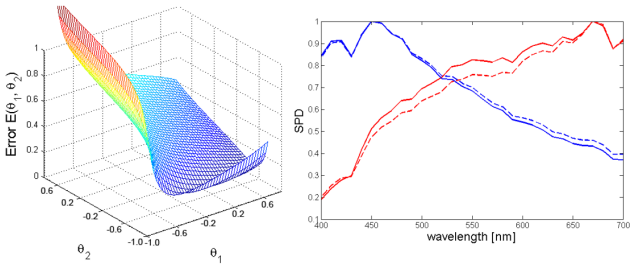

Figure 2. *Left: Error $\mathbb{E}(\theta_1, \theta_2)$ as a function of the two azimuthal angles. Right: Real illuminants (dashed lines) and retrieved illuminants (solid lines).*

The reflectance descriptors are computed as

$$\boldsymbol{\sigma} = \mathbf{M}^{+}\mathbf{p}^{T} \quad (1)$$

where $\mathbf{p} = (r,g,b)$ represents the mean pixel value of a patch in the image, $^{+}$ represents the Moore-Penrose pseudo-inverse, and the matrix $\mathbf{M}$ elements are given by $m_{kj} = \sum_\lambda E_{test} R_k \mathscr{S}_j$, where $E_{test}(\lambda) = \overline{\mathscr{E}}(\lambda) + \varepsilon_2 \mathscr{E}_2(\lambda) + \varepsilon_3 \mathscr{E}_3(\lambda)$ is the test illuminant spectral distribution. $R_k(\lambda)$ is the sensitivity of the kth sensor and $\mathscr{S}_j$ are the reflectance basis functions.

In order to retrieve the illuminants, we use a set of several images of the same scene and select N_p patches of uniform color present in every image, hence having the same reflectance spectra. We average the pixel values of each patch and by inverting a linear image formation model we obtain $8 \cdot N_p \cdot N$ reflectance descriptors corresponding to a set of N test illuminants, indexed by their azimuthal angle θ, and N_p patches. Rather than matching reflectance spectra directly, we compute and compare, for each patch, the corresponding sensor responses under illuminant D65. We compute the sensor responses for each test illuminant $E_{test}(\lambda)$ and patch as

$$p(\theta)_k^{D65} = \sum_\lambda R_k(\lambda) \sum_{i=1}^{8} \sigma_i \mathscr{S}_i(\lambda) E_{D65}(\lambda), \quad k = 1,2,3 \quad (2)$$

For any illuminant angle combination $\Theta_k = (\theta_1(k), \theta_2(k), ..., \theta_N(k))$, we can compute the corresponding sensor responses. Testing all illuminant combinations $\Theta_k = (\theta_1(k), \theta_2(k), ..., \theta_N(k))$ by sampling $\theta_i(k)$ is computationally heavy as the number of iterations increases exponentially with the number of images considered. Instead of testing all illuminant combinations, we find the solution by gradient descent on an error function reaching its minimum when the sensor responses match under illuminant D65, or, indirectly, when the reflectance spectra match.

According to our assumption, the reflectance spectra of each patch $S_p(\lambda) = \sum_{i=1}^{8} \sigma_i \mathscr{S}_i(\lambda)$ remain unchanged across the N images taken under N unknown illuminants.

We vary the test illuminants and compute the corresponding reflectance spectra and sensor responses for each element in every images. We force the test sensor responses to be equal. A simple least square error function is used for computational simplicity. The error function is thus defined as the Euclidean distance between the $N \cdot N_p$ sensor responses corresponding to a combination of N illuminants represented by $\Theta_k = (\theta_1(k), \theta_2(k), ..., \theta_N(k))$.

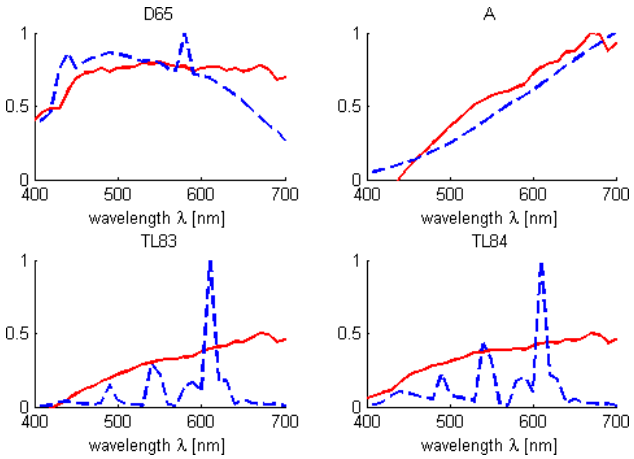

Figure 3. *This figure shows the real non-daylight (blue dashed line) and retrieved daylight illuminants (solid red line) of Fig 4*

The error function is computed as

$$\mathbb{E}(k) = \frac{1}{N_p}\frac{1}{N}\Big[\sum_{j=2,j>i}^{N}\sum_{i=1}^{N}\sum_{n_p=1}^{N_p}[p_{i,n_p}^{D65}(\Theta_k) - p_{j,n_p}^{D65}(\Theta_k)]^2\Big]^{\frac{1}{2}} \quad (3)$$

where the indices i and j run over the images and n_p runs over the N_p patches.

The N image illuminants corresponding to the N images satisfy

$$\Theta_k^{illu} = \min_k(\mathbb{E}(k)) \quad (4)$$

The minimum $\Theta_k^{Illu} = (\theta_1^{Illu}, \theta_2^{Illu}, ..., \theta_N^{Illu})$ is found by gradient descent on the error function (3). At each step, the angles $\theta_i(k)$ are updated in a way to minimize the error function. $\theta_i(k+1)$ is given by

$$\theta_i(k+1) = \theta_i(k) - \alpha\Big[\frac{\mathbb{E}(\theta_i(k)+\Delta\theta) - \mathbb{E}(\theta_i(k)-\Delta\theta)}{\Delta\theta}\Big] \quad (5)$$

The initial angles θ corresponding to the N image illuminants are initialized as random N angles laying on the curve shown in Figure 1.

This method also returns an estimate of the reflectance spectra of the redundant image elements.

Results

We conducted experiments on both synthetic and real images. The synthetic images represent a MacBeth ColorChecker built using Canon 350D sensor sensitivities and the 13 standard CIE illuminant D40 to D100. We selected 6 patches as the constant reflectance elements across the images. We ran experiments using 2 to 4 synthetic images. The angle θ indexing the illuminants follows the curve formed by CIE standard illuminants, as shown in Figure 1. It means that we not only test for the standard illuminants, but for a continuum of illuminants ranging from D40 to D100.

The error is measured as the angle between the 3 dimensional vectors representing the D65 reference and retrieved white points.

Illuminants		Angular errors	
I_1	I_2	$\Delta\varphi_1$	$\Delta\varphi_2$
D40	D50	8.0	8.4
D40	D90	0.6	0.2
D40	D100	0.8	0.9
D65	D80	2.7	2.5
D65	D90	1.8	1.6
D65	D100	1.4	1.2
D80	D90	0.7	0.6
D80	D100	0.8	0.8

Table 1: Angular errors for several choices of two images, indicated by their illuminants. We see that the angular errors can become large when the illuminants are close.

Illuminants			Angular errors		
I_1	I_2	I_3	$\Delta\varphi_1$	$\Delta\varphi_2$	$\Delta\varphi_3$
D40	D50	D60	5.2	5.3	5.5
D40	D50	D80	2.0	1.7	1.8
D40	D50	D100	0.8	0.1	0.2
D65	D75	D85	0.1	0.2	0.2
D65	D75	D100	1.0	1.0	0.8
D75	D80	D85	2.4	2.3	2.2
D80	D90	D100	1.3	1.3	1.2

Table 2: Angular errors for several choices of three images, indicated by their illuminants. We see that the angular errors can become large when the illuminants are close.

The angular error [8] is computed as

$$\Delta\varphi_{WP} = \arccos\Big(\frac{\mathbf{x}_{D65}^{ref}\cdot\mathbf{x}_{D65}^{exp}}{\|\mathbf{x}_{D65}^{ref}\|\|\mathbf{x}_{D65}^{exp}\|}\Big) \quad (6)$$

where $\mathbf{x}_{D65}$ is the white point expressed in sRGB. A perturbation of 1 is usually not noticeable, while an error up to 3 degree remains generally acceptable for such data [6].

Using only two images allows to represent the error as a function of two variables, θ_1 and θ_2, indexing the illuminants. The estimation of the illuminant is good when the image illuminants are quite different. The error function has a strong minimum and the angular error usually remains under 1 degree. When the image illuminants are too similar, the error is sometimes larger and the result can be poor. Figure 2 shows the results using two images built using illuminants D40 and D90. The graph on the left shows the error as a function of the angles θ, while the graph on the right show the real (dashed line) and retrieved illuminants (solid line). Table 1 reports some values of angular errors for various illuminant pairs. Each computation was run three times to check for the consistency of the convergence. The direction of the shifts are the same for both illuminants.

Using three or four images gives more robust results. However, when the illuminants are close, the estimation quality decreases. See Tables 2 and 3 for examples of angular errors in sRGB for different combinations of illuminants using three and four images, respectively. The amplitude and direction of the shifts are the same for the illuminants, as it was also observed using two images. It suggests that we can exploit this constant shift by considering it as an estimation bias. Having a large database of images would ensure having a variety of illuminants and give more robust results.

Figure 4. *Row 1 shows uncorrected images and row 2 shows the corresponding corrected images. Illuminants are, from left to right, D65, A, TL83, and TL84. The red rectangles indicate the reference elements used to white balance the images. The images are rendered to sRGB using a white point preserving transform [4].*

Illuminants				Angular errors			
I_1	I_2	I_3	I_4	$\Delta\varphi_1$	$\Delta\varphi_2$	$\Delta\varphi_3$	$\Delta\varphi_4$
D40	D50	D60	D70	2.3	2.0	2.2	2.2
D40	D50	D80	D90	0.8	1.1	1.0	0.8
D40	D50	D80	D100	2.1	1.9	2.0	1.8
D40	D60	D80	D100	0.9	0.7	0.5	0.3
D55	D65	D75	D95	3.1	3.2	3.1	2.8
D65	D75	D85	D95	3.0	3.1	3.1	2.6
D65	D75	D90	D100	2.6	2.6	2.3	2.1
D75	D80	D85	D90	0.8	0.7	0.6	0.4

Table 3: Angular errors for several choices of four images, indicated by their illuminants. We see that the angular errors can become large when the illuminants are close.

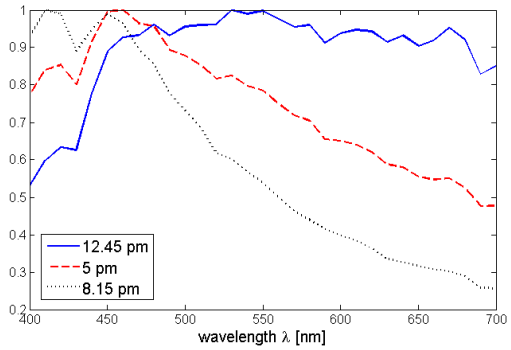

Figure 5. *This figure shows the illuminants retrieved in the images of Fig 6*

We also ran experiments using four real images taken with a Canon 350D camera in a GTI Color Matcher CMB2080 lightbooth with four different illuminants: daylight D65, incandescent A, and two fluorescent illuminants TL83 and TL84. We took the 6 elements in the scene as reference reflectances across the images. Figure 6 shows the images before (first row) and after (second row) color correction using our method. The red rectangles indicate the objects used as references.

In this case, we are approximating non-daylight illuminants with daylight illuminants. We cannot retrieve the actual illuminant spectra using daylight illuminant basis functions. For fluorescent illuminants, the actual and retrieved spectra will be very different (see Figure 3). However, the retrieved daylight illuminants can be used to white balance the images, as shown in Figure 4. The corresponding angular errors are, from left to right, $\Delta\varphi_1 = 5.4°$, $\Delta\varphi_2 = 3.7°$, $\Delta\varphi_3 = 13.4°$, and $\Delta\varphi_4 = 9.7°$. The second image, while having the lowest angular error, has a color cast too important to be correctly white balanced using a global method. Illuminant A emits very little in the short wavelengths. Scenes taken under that illuminant exhibit a deficiency in the blue channel often impossible to correct for.

While the result is visually satisfying, the convergence of the gradient descent depends on the initial conditions and sometimes converges to an incoherent result. For this particular set of images, there are two distinct solutions, and only one is acceptable. This effect was also observed on synthetic images built with various real and synthetic, daylight and fluorescent illuminants, showing that this is not due to the use of real images. When the illuminants are far from daylight illuminants, the algorithm does not converge to a satisfying solution for all initial conditions, probably due to the presence of multiple minima.

We also ran experiments using a set of three real images taken with a Canon 350D camera under real, unknown daylight illuminants. The images were taken, respectively, at 12.45 pm, 5 pm, and 8.15 pm late June. The algorithm was run several times for slightly different choices of reference patches and converged each time to the same result. Figure 5 shows the illuminants retrieved in the images shown in Figure 6. Ten elements, displayed in the image, were selected as reference reflectances and used to retrieved the illuminants. The left image has a reddish color cast and the right image has a bluish color cast, corresponding to the retrieved illuminants. The corresponding illuminant spectra maximum shift to shorter wavelengths.

Our technique is designed to be applied to a fixed location camera. In this case, using a set of images covering a sufficient range of illuminants, we can compute the reflectance spectra and the corresponding sensor responses of the static objects in the scene and use them as references. For each new image, we can extract the pixel values of the reference objects, whose positions are known, and run our algorithm using the previously computed reference reflectances.

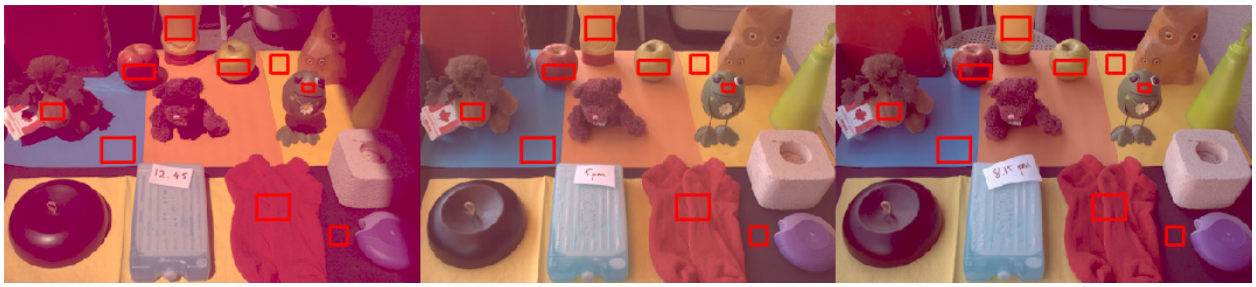

Figure 6. *This figure shows uncorrected images. The red rectangles indicate the reference elements used to retrieve the illuminants, shown in Fig 5. The images are rendered to sRGB using a white point preserving transform [4].*

Conclusion

We present an approach to solve for color constancy in images taken with fixed location cameras. Such images, though having different dynamic contents and being imaged under changing illuminants, contain several static objects in the background. We use a set of images containing redundant elements, hence having the same reflectance spectra, to retrieve the illuminant in each image of the set. The problem is solved by gradient descent. We minimize the distance between test sensor responses corresponding to the static objects present in every scene. We demonstrated the feasibility of our method on both synthetic and real images. The method will be improved by modifying the gradient descent computation, for example by using an adaptive step in the illuminant angle update. The quality and rapidity of the convergence should improve. While the method is designed to retrieve daylight illuminants, it also allows to white balance images taken under non-daylight illuminants. However, the convergence is not always guaranteed. The convergence conditions will be studied in details in order to determine the reliability of the method for the white-balancing of images taken under non-daylight illuminants. The method shows good results whenever the set of images contains a sufficient variety of illuminants.

References

[1] G. Buchsbaum, *A spatial processor model for object colour perception*, J. Franklin Inst., vol. 310, p. 1-26, 1980

[2] J. Cohen, *Dependency of the spectral curves of the Munsell color chips*, Psychon. Sci., p. 369-370, 1964

[3] J.M. DiCarlo, F. Xiao, and B.A. Wandell, *Illuminating illumination*, Proc. of the 9th Color Imaging Conf., p. 27-34, 2001

[4] G.D. Finlayson and M.S. Drew, *Constrained least-squares regression in color spaces*, Journal of Electronic Imaging, vol. 6 (4), p. 484-493, 1997

[5] G.D. Finlayson and S. Süsstrunk, *Spherical sampling and color transformations*, Proc. of the 9th IS&T/SID Color Imaging Conf., p. 321-325, 2001

[6] G.D. Finlayson, S. Hordley, and P. Morovic, *Colour Constancy using the chromagenic constraint*, Proc. of the 2005 IEEE Conf. on Comp. Vision and Pattern Recogn., p.1079-1086, 2005

[7] R. Gershon, A.D. Jepson, and J.K. Tsotsos, *From [R, G, B] to Surface Reflectance: Computing Color Constant Descriptors in Images* Percept. Psychophys., vol. 17, p. 755-758, 1988

[8] S.D. Hordley, *Scene Illuminant Estimation: Past, Present, and Future*, Color Research & Application, p. 303-314, vol. 31 (4), 2006

[9] D.B. Judd, D.L. McAdam, and G. Wyszecki, *Spectral distribution of typical daylight as a function of correlated color temperature*, J. Opt. Soc. Am. A, vol. 54, p. 1031-1040, 1964

[10] E.H. Land, *Lightness and retinex theory*, J. Opt. Soc. Am. A, vol. 61, p. 1-11, 1971

[11] H.-C. Lee, *Method for computing the scene-illuminant from specular highlights*, J. Opt. Soc. Am. A, vol. 3 (10), p. 1694-1699

[12] C. Lu and M.S. Drew, *Practical scene illuminant estimation via flash/no flash pairs*, Proc. of the 14th IS&T/SID Color Imaging Conf., p. 84-89, 2006

[13] L.T. Maloney, *Evaluation of linear models of surface spectral reflectance with small number of parameters*, J. Opt. Soc. Am. A, vol. 3, p. 1673-1683, 1986

[14] L.T. Maloney and B.A. Wandel, *Color constancy: a method for recovering surface spectral reflectance*, J. Opt. Soc. Am. A, vol. 3, p. 29-33, 1985

[15] D.H. Marimont and B.A. Wandell, *Linear models of surface and illuminant spectra*, J. Opt. Soc. Am. A, vol. 9 (11), p. 1905-1914, 1992

[16] J.P.S. Parkkinen, J. Hallikanen, and T. Jaaskelainen, *Characteristic spectra of Munsell colors*, J. Opt. Soc. Am. A, vol. 6, p. 318-322, 1989

[17] G. Petschnigg, R. Szeliski, M. Agrawala, M.F. Cohen, H. Hoppe, and K. Toyama, *Digital photography with flash and no flash image pairs*, ACM Trans. on Graphics, vol. 23, p. 664-672, 2004

[18] B. Smith, C. Spiekermann, and R. Sember, *Numerical methods for colorimetric calculations: sampling density requirements*, Color Research & Applications, vol. 17 (6) p.394-401, 1992

[19] S. Westland, A.J. Shaw, and H.C. Owens, *Colours statistics of natural and man made surfaces*, Sensor Review, vol. 20 (1), p. 50-55, 2000

[20] M. D'Zmura and G. Iverson, *Color Constancy. I. Basic theory of two-stage linear recovery of spectral descriptions for lights and surfaces*, J. Opt. Soc. Am. A, vol. 10 (10), p. 2148-2165, 1993

[21] Spectral Database, University of Joensuu Color Group, http://spectral.joensuu.fi/

Author Biography

Joanna Marguier received her MS in physics from the Ecole Polytechnique Fédérale de Lausanne (EPFL) in 2003. She worked two years as process engineer in the Displays Department at Swatch Group's R&D Laboratories. Since 2005, she has been a Research Assistant in Color Imaging, pursuing a PhD degree in the Audiovisual Communications Laboratory, EPFL.