Illuminant estimation from uncalibrated multispectral images

Jean-Baptiste Thomas
LE2I UMR6306, CNRS, Arts et Métiers,
Univ. Bourgogne Franche-Comté, F-21000 Dijon, France
Email: jean-baptiste.thomas@u-bourgogne.fr

***Abstract*—We investigate the physical validity of typical computational color constancy models for illuminant estimation of uncalibrated multispectral images. We demonstrate empirically that the assumptions may be reasonable and that we retrieve reasonably well the illumination for some images. On these images, we also have access to a good estimate of the spectral properties of the illumination while increasing the number of bands. However, some other images do not provide a very good illuminant estimation. We also show that the result depends mostly on the scene, rather than on the hypothesis made or on the number of spectral bands. Besides, the influence of the algorithm and its hypothesis is more critical for more bands than for the $3-D$ color case.**

I. Introduction

Illuminant estimation is usually performed in color within the camera domain. Common assumptions typically assume first a natural outdoor scene without any color cast. Within this condition, the assumptions can be formulated as a gray world, a maximum a posteriori, gray edges, etc. We propose to evaluate the validity of these hypothesis from a spectral perspective rather than a color one, while considering a varying multispectral sensor. We make the sensor to vary from three bands (*i.e.* a typical color camera) to an infinity of bands (*i.e.* an ideal spectroradiometer). For simplicity of presentation, we consider only equi-Gaussian sensor sensitivities. We compare the illuminant and its estimate within the camera domain. While increasing dimensionality, computation shows a surprisingly good match with some images or a surprisingly bad match with others, on the data-set we used.

In the following, we define illuminant estimation for $N-D$ sensors based on color constancy priors. Then, we explain our methodology and its limitation before to evaluate the result on images before to conclude.

II. Generalisation of illuminant estimation

Illuminant estimation is usually considered a part of the white balance process. White balance, or computational color constancy [1], is a process that simulates or approximates the human visual effect of color constancy [2], which allows to perceive the same scene when illumination changes. This often consists in a chromatic adaptation assuming Von Kries model [1].

[1]We are aware of the limit of Von Kries models, while simulating chromatic adaptation, however for a white balance perspective and purpose, the correction is often modeled simply as a diagonal matrix. So we do in the following.

It usually estimates the illumination of the scene within the sensor space, then transforms the acquired data from the scene to some data such as the scene appears simulated under a canonic or neutral illumination. In other specific cases, it aims at the recovery of the relative spectral distribution of the illuminant [3].

For an arbitrary multichannel camera (*i.e.* RGB or $\{F_1, ..., F_N\}$), we can define the acquired data by a camera channel by ρ_z, z being a given channel (i.e. $z \in \{R, G, B\}$ or $z \in \{F_1, ..., F_N\}$), such as $\rho_z = \int I(\lambda)R(\lambda)z(\lambda)d\lambda$, with $I(\lambda)$ is the illumination, $R(\lambda)$ is the reflectance of the material surface within the scene and $z(\lambda)$ the spectral sensitivity of a channel z.

Under Von Kries model of chromatic adaptation, which assumes an individual tuning of Human cones sensitivities, we can write Equation 1.

$$\begin{bmatrix} L \\ M \\ S \end{bmatrix}_{adapted} = \begin{bmatrix} K_L & 0 & 0 \\ 0 & K_M & 0 \\ 0 & 0 & K_S \end{bmatrix} . \begin{bmatrix} \rho_L \\ \rho_M \\ \rho_S \end{bmatrix} \quad (1)$$

with $K_L = \frac{1}{L_W}$, $K_M = \frac{1}{M_W}$ and $K_S = \frac{1}{S_W}$, where $\{L_W, M_W, S_W\}$ considers the response of the cones to the illumination of a white surface, therefore, the illumination itself.

The correction for an RGB sensor can be modeled likely as a linear system, such as in Equation 2:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} K_R & 0 & 0 \\ 0 & K_G & 0 \\ 0 & 0 & K_B \end{bmatrix} . \begin{bmatrix} \rho_R \\ \rho_G \\ \rho_B \end{bmatrix} \quad (2)$$

with $[RGB]^T$ the corrected values representing the ratio of R, G and B necessary to represent the object reflectance under a neutral illumination within the sensor space, with the K_z relative to an estimation of the acquisition illumination. Note that this model is largely inaccurate in RGB, and is often applied into an intermediate space (*i.e.* Bradford or such).

Nevertheless, our proposal is that we can use the basic methods usually used to fit this model, by extending a similar model to a multispectral sensor based on N arbitrary channels, $N \in [0, \infty]$. Consequently, we can use similar algorithms and image statistics as used in RGB illuminant estimation.

For the sake of simplicity of evaluation and visualization, we limit the present study to equi-Gaussian sensors (See Figure III-A). We can then retrieve the relative shape of the

2015 Colour and Visual Computing Symposium (CVCS)
978-1-4799-1765-5/15/$31.00

illumination in the sensor domain from any given image. In the case of an equi-Gaussian sensor, when N reaches infinity, we are in an approximation of the case of spectral measurement. If this evaluation is positive, we would not only validate these models, but also be able to perform automatic multispectral image calibration, which might be of benefit in different applications of computer vision techniques.

Considering a multispectral camera based on N filters, $\{\rho_z, z \in [1, N] \; and \; z \in \mathbf{Z}\}$, the previous model is generalized by Equation 3:

$$\begin{bmatrix} C_1 \\ C_2 \\ . \\ . \\ . \\ C_N \end{bmatrix} = \begin{bmatrix} K_1 & 0 & . & . & . & 0 \\ 0 & K_2 & & & & 0 \\ . & & . & & & . \\ . & & & . & & . \\ . & & & & . & . \\ 0 & & . & . & . & K_N \end{bmatrix} . \begin{bmatrix} \rho_1 \\ \rho_2 \\ . \\ . \\ . \\ \rho_N \end{bmatrix} \quad (3)$$

with the C_i the corrected values.

If we are in the case of a spectroradiometer with narrow spectral bands, the components of this diagonal transform matrix are the inverse of the illumination spectral sensitivities sampled with the spectroradiometer. In the case of a generic sensor, the inverse of this matrix components are a projection of the illumination into the sensor domain. When the image is corrected by this model, spectral data are normalized correctly, free from illumination, *i.e.* calibrated.

From this model, the problem of expressing the illumination is typically the same as for RGB images.

We consider a single class of very simple algorithms in this study since our goal is to present the method and to evaluate it empirically on a reduced data-set of images and hypothesis. The concept can, more or less easily, be extended to more sophisticated methods. However, the evaluation proposed in this work, particularly the visualization as spectra curves is a correct approximation limited to narrow band multispectral sensors, such as non-overlapping equi-bandpass filters filling the whole range of visible or as we use, equi-Gaussian filters not too much overlapping.

A general writing of the models investigated is the minkowski $p-$norm of the image [4], or the minkowski $p-$norm of the jth derivative of the image [5].

In the remaining of this section, the illuminant estimation processes considered are described. In the formula, α_i denotes the normalised estimate of the illuminant through channel i, $max_N(\alpha_i)$ is a normalizing factor.

A. Gray world

This algorithm is based on the assumption that the average of every channel corresponds to an average achromatic luminance (*i.e.* the Minkowski $1-$norm).

We define the multispectral gray world such as in Equation 4, where $card$ denotes the cardinality:

$$\alpha_i = \frac{\frac{1}{card(\rho_i)} \sum_{image} \rho_i}{max_N(\alpha_i)} \quad (4)$$

.

Gray world algorithm will be referred as A_1 in the following.

B. White patch

This is probably what is done intuitively in many multispectral image application where no reflectance reconstruction is performed. For color images, it is based on Retinex theory, which argues that the perceive white is based on the maximum excitation of the cones (*i.e.* the Minkowski $\infty-$norm).

We define the multispectral white patch such as in Equation 5:

$$\alpha_i = \frac{max_{image}(\rho_i)}{max_N(\alpha_i)} \quad (5)$$

.

White patch algorithm will be referred as A_2 in the following.

A modified white patch algorithm may correct saturation effects on the sensor. We would then averaged the 10% of the pixels of maximum values by channel. This will be referred as A_3 in the following. Note that the data-set used is not very likely to be saturated.

C. Gray edges

This algorithm is an extension of the gray world assuming that edges only are gray (*i.e.* the Minkowski $1-$norm of the first order derivative of the image). Our implementation of this algorithm is using standard implementation of Canny edge detection by channel. While we consider which pixels to average, we consider that an edge detected in one band is enough to consider that this is an edge in the image (*i.e.* logical OR).

Gray edge algorithm will be referred as A_4 in the following.

We classify algorithms A_1 and A_4 to be compliant with hypothesis of gray world H_1. We classify A_2 and A_3 to be compliant with hypothesis of maximum intensity H_2.

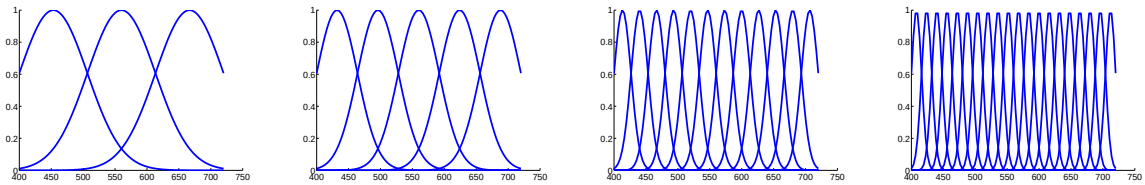
III. Evaluation

A. Simulation

We use reflectance images from Foster *et al.* multispectral image database [6] virtually illuminated by a $D65$ CIE-illuminant to simulate an image of radiance. The sensor is simulated with equi-Gaussian filters with not-too-much overlapping. We used a standard deviation of half the ratio between the spectral range and the number of channels. Examples of sensors are shown for 3, 5, 12 and 20 bands in Figure III-A. Three bands sensor is chosen for the typical color case. Five bands are chosen for it is a typical number of bands for multispectral sensors or colorimeters [7]. Twelve bands are chosen for they should provide the best spectral reconstruction while used in multispectral sensors [8], [9]. Twenty bands are chosen for this sensor would approach the accuracy of a spectroradiometer without involving too much computation time.

In the simulation, the spectral dimension is sampled with 100 data between 400 and 720 nm.

We chose the 8 scenes from Foster *et al.* multispectral image database [6]. They are referred from I_1 to I_8 in the following (see Figure IV to IV).



(a) 3 bands sensor (b) 5 bands sensor (c) 12 bands sensor (d) 20 bands sensor

Fig. 1. Simulation of equi-Gaussian sensors.

B. Index of goodness

Usually, the goodness of computational color constancy and illumination estimation may be evaluated by the angle between the estimation and the actual illuminant. This is generally performed on $3-D$ data, however, Equation 6 generalizes the hyper-angle definition to the $N-D$ case.

$$hyper-angle = \arccos \frac{\mathbf{I}.\hat{\mathbf{I}}}{\|\mathbf{I}\|.\|\hat{\mathbf{I}}\|} \tag{6}$$

where boldfaces denote vectors representing the illuminants. $\mathbf{I}$ is the normalized vector of $D65$ directly acquired by the sensor, while $\hat{\mathbf{I}}$ is the normalized estimate based on image statistics. When the angle is zero, the vectors are colinear.

Spectra may be compared with a Goodness of fit indicator (GoFC) rather than the sRMSE, especially for normalised spectra. In this case, we use Equation 7.

$$GoFC = \frac{\| \sum_i I(F_i).\hat{I}(F_i)\|}{\| \sum_i I(F_i)^2\|^{\frac{1}{2}}.\| \sum_i \hat{I}(F_i)^2\|^{\frac{1}{2}}} \tag{7}$$

which value closer to one is the spectra are similar.

No comparison can be made between results of different number of bands with these indicators. This would be meaningless to compare data in different spaces of different dimensions. However, we can compare, within one sensor, the results between images and estimation algorithms.

IV. Results and analysis

Results are shown for different sensor dimensionality in Tables I, II, III and IV. These tables permits to compare algorithms versus images each for a given sensor. It is very interesting to notice that both indicators are mostly agreeing on the rank of the evaluation. In general, results show that the number of bands does not really influence which algorithm is the best. It is rather the scene that may help to predict the result.

We observe a good correlation between angle and GoFC.

It is interesting to notice that algorithms A_1 and A_4 are giving similar results for the same image and algorithms A_2 and A_3 are giving similar results for a same image. This is not critically surprising as they are based on similar ground hypothesis. It is not surprising that the modified white patch does not improve much the estimation as these images are not that saturated. On scene I_5 it even impairs the results by smoothing differences between channels due to the relatively high brightness of the scene. It is, however, surprising that it seems that edges are not that gray after all.

For some scene we can obtain a very good estimation of the illuminant spectral properties. For some others the estimate seems critically bad.

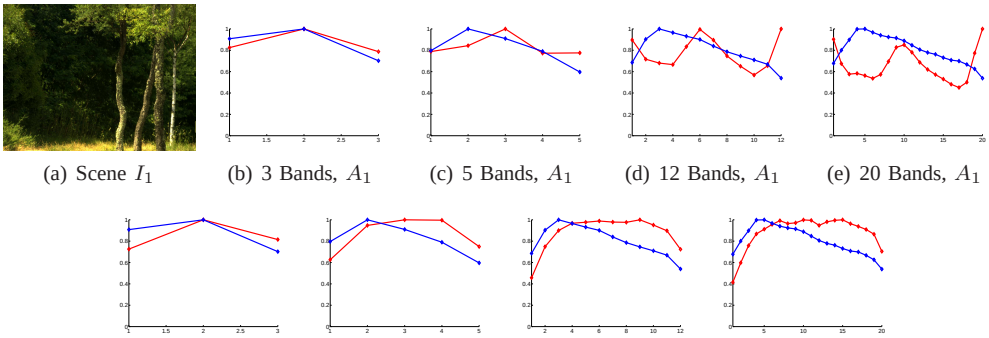
We can classify scenes subjectively by their compliance with illuminant evaluation hypothesis H_1, H_2, both of them or none of them, as in Table V.

TABLE I. Estimation of illuminant by scene, by algorithm for a 3 bands sensor.

Image VS Algorithm		A_1	A_2	A_3	A_4
I_1	GoFC	0.9969	0.9902	**0.9981**	0.9882
	hyper-angle	0.0784	0.1404	**0.0617**	0.1540
I_2	GoFC	0.9554	0.9568	**0.9951**	0.9542
	hyper-angle	0.2999	0.2951	**0.0992**	0.3038
I_3	GoFC	0.9762	0.9853	**0.9928**	0.9741
	hyper-angle	0.2185	0.1719	**0.1198**	0.2280
I_4	GoFC	0.9222	**0.9996**	0.9846	0.9307
	hyper-angle	0.3970	**0.0280**	0.1755	0.3746
I_5	GoFC	0.9707	**0.9992**	0.9900	0.9697
	hyper-angle	0.2427	**0.0401**	0.1418	0.2466
I_6	GoFC	**0.9991**	0.9974	0.9927	0.9988
	hyper-angle	**0.0418**	0.0719	0.1211	0.0485
I_7	GoFC	0.9992	**0.9998**	0.9910	0.9975
	hyper-angle	0.0407	**0.0212**	0.1345	0.0708
I_8	GoFC	0.9828	**0.9942**	**0.9942**	0.9810
	hyper-angle	0.1856	0.1081	**0.1073**	0.1954

TABLE II. Estimation of illuminant by scene, by algorithm for a 5 bands sensor.

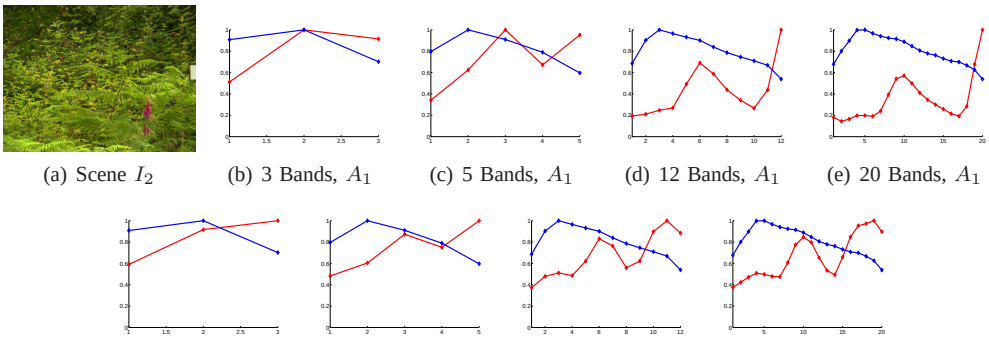
Image VS Algorithm		A_1	A_2	A_3	A_4
I_1	GoFC	0.9907	0.9870	**0.9925**	0.9808
	hyper-angle	0.1363	0.1616	**0.1225**	0.1965
I_2	GoFC	0.9249	0.9369	**0.9842**	0.9245
	hyper-angle	0.3900	0.3572	**0.1780**	0.3911
I_3	GoFC	0.9653	0.9775	**0.9890**	0.9631
	hyper-angle	0.2642	0.2124	**0.1484**	0.2725
I_4	GoFC	0.8921	0.9992	0.9902	0.9033
	hyper-angle	0.4688	0.0391	0.1401	0.4435
I_5	GoFC	0.9608	**0.9989**	0.9867	0.9597
	hyper-angle	0.2810	**0.0463**	0.1631	0.2849
I_6	GoFC	**0.9987**	0.9957	0.9906	0.9984
	hyper-angle	**0.0502**	0.0929	0.1370	0.0569
I_7	GoFC	0.9986	**0.9995**	0.9869	0.9965
	hyper-angle	0.0520	**0.0328**	0.1619	0.0840
I_8	GoFC	0.9768	**0.9919**	0.9898	0.9745
	hyper-angle	0.2156	**0.1271**	0.1428	0.2265



(a) Scene I_1 (b) 3 Bands, A_1 (c) 5 Bands, A_1 (d) 12 Bands, A_1 (e) 20 Bands, A_1

(f) 3 Bands, A_2 (g) 5 Bands, A_2 (h) 12 Bands, A_2 (i) 20 Bands, A_2

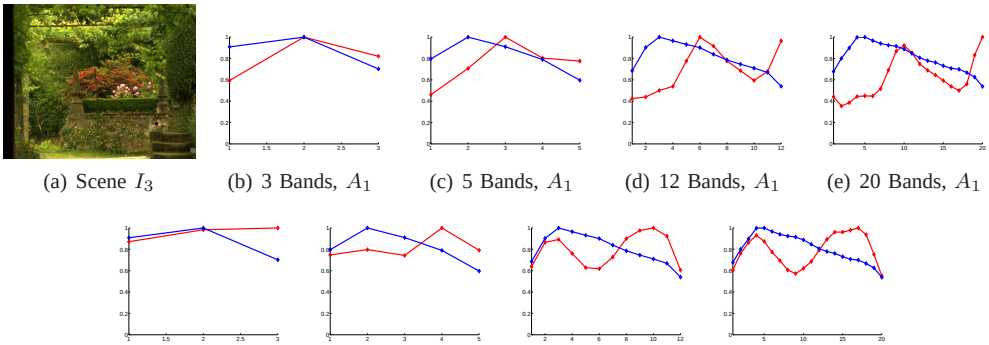
Fig. 2. Simulation results for scene I_1. In blue, the illuminant captured by the sensor is shown, in red, its estimation.



(a) Scene I_2 (b) 3 Bands, A_1 (c) 5 Bands, A_1 (d) 12 Bands, A_1 (e) 20 Bands, A_1

(f) 3 Bands, A_2 (g) 5 Bands, A_2 (h) 12 Bands, A_2 (i) 20 Bands, A_2

Fig. 3. Simulation results for scene I_2. In blue, the illuminant captured by the sensor is shown, in red, its estimation.



(a) Scene I_3 (b) 3 Bands, A_1 (c) 5 Bands, A_1 (d) 12 Bands, A_1 (e) 20 Bands, A_1

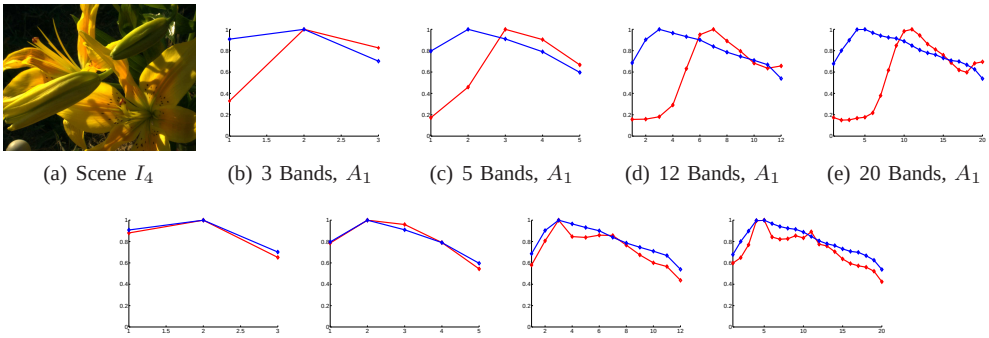
(f) 3 Bands, A_2 (g) 5 Bands, A_2 (h) 12 Bands, A_2 (i) 20 Bands, A_2

Fig. 4. Simulation results for scene I_3. In blue, the illuminant captured by the sensor is shown, in red, its estimation.

V. Conclusion

This work demonstrates that the evaluation of the illumination on multispectral images based on computational color constancy may be feasible on similar models than used in computational color constancy. The accuracy of the result depends on the scene content and is more sensitive to the algorithm that it is in the color $(3-D)$ case, particularly when we increase the number of bands.

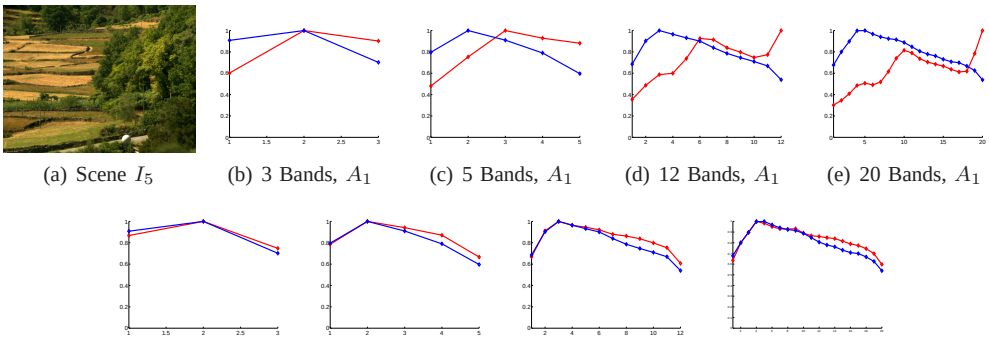
Sensor spectral sensitivities influence might be evaluated in further work. However, while changing the sensitivity of channels from equi-Gaussians to random shapes, we would get further away from spectral measurements and comparison between data of different dimensions would become extremely complex. Moreover, due to higher or different correlation between spectral sensitivities, the assumption of a similar diagonal linear Von Kries may break.



(a) Scene I_4 (b) 3 Bands, A_1 (c) 5 Bands, A_1 (d) 12 Bands, A_1 (e) 20 Bands, A_1

(f) 3 Bands, A_2 (g) 5 Bands, A_2 (h) 12 Bands, A_2 (i) 20 Bands, A_2

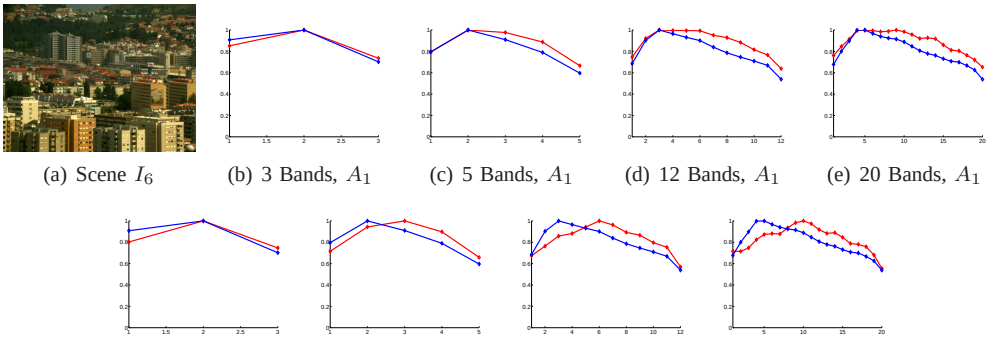
Fig. 5. Simulation results for scene I_4. In blue, the illuminant captured by the sensor is shown, in red, its estimation.



(a) Scene I_5 (b) 3 Bands, A_1 (c) 5 Bands, A_1 (d) 12 Bands, A_1 (e) 20 Bands, A_1

(f) 3 Bands, A_2 (g) 5 Bands, A_2 (h) 12 Bands, A_2 (i) 20 Bands, A_2

Fig. 6. Simulation results for scene I_5. In blue, the illuminant captured by the sensor is shown, in red, its estimation.



(a) Scene I_6 (b) 3 Bands, A_1 (c) 5 Bands, A_1 (d) 12 Bands, A_1 (e) 20 Bands, A_1

(f) 3 Bands, A_2 (g) 5 Bands, A_2 (h) 12 Bands, A_2 (i) 20 Bands, A_2

Fig. 7. Simulation results for scene I_6. In blue, the illuminant captured by the sensor is shown, in red, its estimation.

References

[1] A. Gijsenij, T. Gevers, and J. van de Weijer, "Computational color constancy: Survey and experiments," *IEEE Transactions on Image Processing*, vol. 20, no. 9, pp. 2475–2489, 2011. [Online]. Available: https://ivi.fnwi.uva.nl/isis/publications/2011/GijsenijTIP2011

[2] D. H. Brainard and L. T. Maloney, "Surface color perception and equivalent illumination models," *Journal of Vision*, vol. 11, no. 5, 2011. [Online]. Available: http://www.journalofvision.org/content/11/5/1.abstract

[3] S. Peyvandi, S. H. Amirshahi, J. Hernández-Andrés, J. L. Nieves, and J. Romero, "Spectral recovery of outdoor illumination by an extension of the bayesian inverse approach to the gaussian mixture model," *J. Opt. Soc. Am. A*, vol. 29, no. 10, pp. 2181–2189, Oct 2012. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-29-10-2181

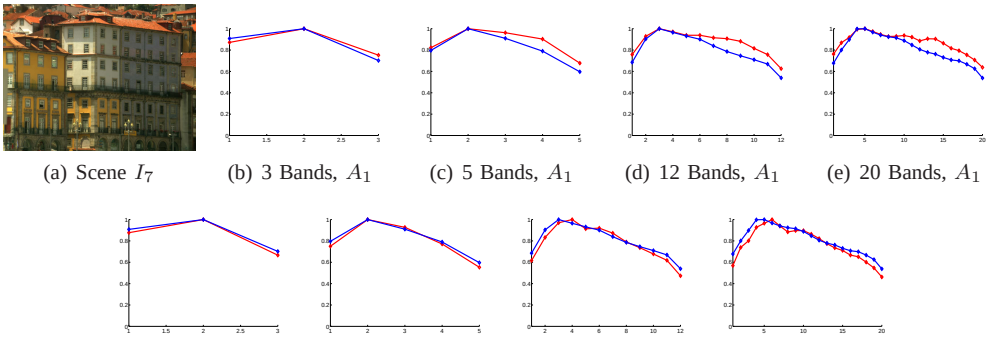
[4] G. D. Finlayson and E. Trezzi, "Shades of gray and colour constancy," in *Twelfth Color Imaging Conference: Color Science and Engineering Systems, Technologies, and Applications*, November 2004, pp. 37–41. [Online]. Available: https://ueaeprints.uea.ac.uk/23682/

[5] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-based color constancy," *Image Processing, IEEE Transactions on*, vol. 16, no. 9, pp. 2207–2214, Sept 2007.

[6] D. H. Foster, K. Amano, S. M. C. Nascimento, and M. J. Foster, "Frequency of metamerism in natural scenes," *J. Opt. Soc. Am. A*, vol. 23, no. 10, pp. 2359–2372, Oct 2006. [Online]. Available: http://josaa.osa.org/abstract.cfm?URI=josaa-23-10-2359

[7] Y. Monno, M. Tanaka, and M. Okutomi, "Multispectral demosaicking using adaptive kernel upsampling," in *Image Processing (ICIP), 2011 18th IEEE International Conference on*, Sept 2011, pp. 3157–3160.

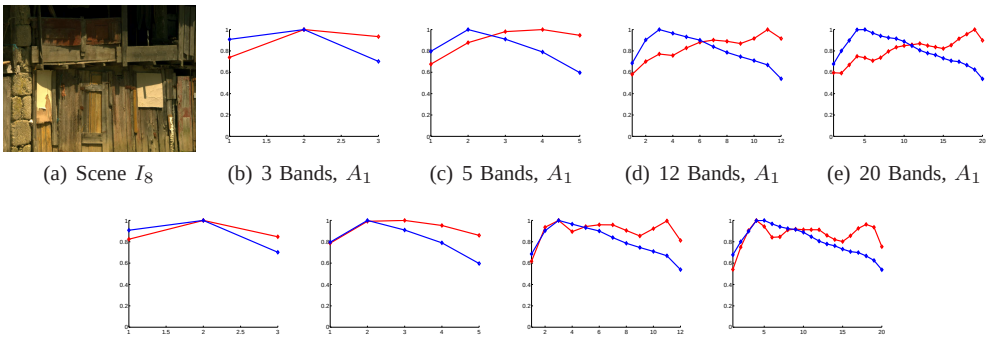
[8] X. Wang, J.-B. Thomas, J. Y. Hardeberg, and P. Gouton, "A Study on the Impact of Spectral Characteristics of Filters on Multispectral Image



(a) Scene I_7 (b) 3 Bands, A_1 (c) 5 Bands, A_1 (d) 12 Bands, A_1 (e) 20 Bands, A_1

(f) 3 Bands, A_2 (g) 5 Bands, A_2 (h) 12 Bands, A_2 (i) 20 Bands, A_2

Fig. 8. Simulation results for scene I_7. In blue, the illuminant captured by the sensor is shown, in red, its estimation.



(a) Scene I_8 (b) 3 Bands, A_1 (c) 5 Bands, A_1 (d) 12 Bands, A_1 (e) 20 Bands, A_1

(f) 3 Bands, A_2 (g) 5 Bands, A_2 (h) 12 Bands, A_2 (i) 20 Bands, A_2

Fig. 9. Simulation results for scene I_8. In blue, the illuminant captured by the sensor is shown, in red, its estimation.

TABLE III. ESTIMATION OF ILLUMINANT BY SCENE, BY ALGORITHM FOR A 12 BANDS SENSOR.

Image	VS Algorithm	A_1	A_2	A_3	A_4
I_1	GoFC	0.9660	**0.9835**	0.9804	0.9580
	hyper-angle	0.2616	**0.1821**	0.1984	0.2909
I_2	GoFC	0.8416	0.9261	**0.9658**	0.8444
	hyper-angle	0.5705	0.3866	**0.2621**	0.5653
I_3	GoFC	0.9409	0.9709	**0.9884**	0.9400
	hyper-angle	0.3456	0.2416	**0.1526**	0.3481
I_4	GoFC	0.8605	**0.9978**	0.9879	0.8761
	hyper-angle	0.5345	**0.0669**	0.1560	0.5030
I_5	GoFC	0.9452	**0.9988**	0.9865	0.9438
	hyper-angle	0.3326	**0.0500**	0.1645	0.3369
I_6	GoFC	**0.9983**	0.9935	0.9896	0.9980
	hyper-angle	**0.0586**	0.1144	0.1441	0.0627
I_7	GoFC	0.9982	**0.9988**	0.9864	0.9958
	hyper-angle	0.0605	**0.0483**	0.1651	0.0919
I_8	GoFC	0.9712	**0.9894**	0.9884	0.9685
	hyper-angle	0.2407	**0.1457**	0.1527	0.2516

TABLE IV. ESTIMATION OF ILLUMINANT BY SCENE, BY ALGORITHM FOR A 20 BANDS SENSOR.

Image	VS Algorithm	A_1	A_2	A_3	A_4
I_1	GoFC	0.9547	**0.9824**	0.9820	0.9482
	hyper-angle	0.3020	**0.1881**	0.1899	0.3234
I_2	GoFC	0.8116	0.9217	**0.9620**	0.8141
	hyper-angle	0.6240	0.3984	**0.2765**	0.6197
I_3	GoFC	0.9318	0.9698	**0.9887**	0.9313
	hyper-angle	0.3716	0.2465	0.1505	0.3728
I_4	GoFC	0.8541	**0.9972**	0.9880	0.8671
	hyper-angle	0.5469	**0.0754**	0.1550	0.5215
I_5	GoFC	0.9382	**0.9987**	0.9876	0.9368
	hyper-angle	0.3533	**0.0514**	0.1577	0.3575
I_6	GoFC	**0.9982**	0.9930	0.9888	0.9980
	hyper-angle	**0.0608**	0.1183	0.1495	0.0632
I_7	GoFC	0.9980	**0.9986**	0.9873	0.9956
	hyper-angle	0.0628	**0.0527**	0.1595	0.0934
I_8	GoFC	0.9695	0.9884	**0.9902**	0.9671
	hyper-angle	0.2475	0.1522	**0.1402**	0.2572

TABLE V. CLASSIFICATION OF SCENES RELATED TO THE HYPOTHESIS.

Scene VS hypothesis	$\emptyset$	H_1	H_2	$H_1 \cap H_2$
Scenes	1, 2, 3, 8	6	4, 5	7
Number of scenes	4	1	2	1
% of scenes	0.5	0.125	0.25	0.125

Acquisition," in *Proceedings of AIC Colour*, S. W. Lindsay MacDonald, Stephen Westland, Ed., vol. 4, Gateshead, Royaume-Uni, Jul. 2013, pp. 1765–1768. [Online]. Available: http://hal.archives-ouvertes.fr/hal-00847828

[9] ——, "Multispectral imaging: narrow or wide band filters?" *Journal of the international colour association*, vol. 12, pp. 44–51, Jul 2014.