Recovering Camera Sensitivities using Target-based Reflectances Captured under multiple LED-Illuminations

Philipp Urban, Michael Desch, Kathrin Happel and Dieter Spiehl

Institute of Printing Science and Technology
Technische Universität Darmstadt,
Magdalenenstr. 2, 64289 Darmstadt, Germany
urban@idd.tu-darmstadt.de
www.idd.tu-darmstadt.de/color

This paper investigates the recovery of spectral camera sensitivities combining multiple captures of a spectrally measured target under different LED illuminants. This is especially attractive since novel computer controlled LED-based viewing booths emerged on the market recently. The effective dimension of the producible spectral stimuli is higher compared to stimuli that are generated by using the same target but only a single illuminant. This allows a more accurate recovery of sensitivities. We estimate the camera sensitivities from the stimuli using a constrained maximum-a-posteriori approach that considers additional knowledge on the properties of camera sensitivities. Sensitivities of a six channel camera are recovered and show a good predicting performance if used to model camera responses.

1 Introduction

Estimating reflectances (or colors) from camera responses is required in many imaging applications, such as color reproduction or image analysis. An important class of spectral estimation as well as color correction methods relies on a detailed knowledge of the imaging system (e.g. non-linearities, spectral channel sensitivities etc.) [9, 10]. Unfortunately, in most cases such information is not provided by the camera vendor and needs to be determined in advance.

For accurately determining spectral camera sensitivities usually a monochromator is used that generates extremely narrow-band stimuli, which are distributed over the whole sensitivity wavelength-range of the camera and allow a simple sampling of the sensitivities.

Unfortunately, a monochromator is rather expensive and often not available. Therefore, many authors proposed methods that use captures of spectrally measured targets (e.g. Linear- or quadratic programming [5, 3], projection onto convex sets [6, 7], evolutionary algorithms [2] etc.). A main problem of such approaches is the relatively low effective dimension [4] of the considered stimuli if common targets (e.g. Color Checker, IT8.7/2, Esser etc.) are used and illuminated by a single light source. This is the reason why additional assumptions on the properties of common camera sensitivities are considered by the reconstruction methods, such as smoothness, positivity, boundedness or uni-modality. The smaller the effective dimension of the stimuli the larger the influence of those assumptions on the reconstruction. Furthermore, many methods require additional parameters (weighting factors) to control the importance of these assumptions compared to the residual error (see e.g. [8] for a comparison of methods). The reconstruction errors are likely to be large for a small effective dimension of stimuli, incorrect weighting factors or assumptions.

In this paper, we investigate the recovery of spectral camera sensitivities combining multiple captures of a spectrally measured target under different LED illuminants. The approach allows us to generate spectral stimuli with a high effective dimension and, as a result, a more accurate sensitivity estimation utilizing common targets. This is especially attractive since new computer controlled LED-based viewing booths emerged on the market recently.
The main features of the proposed method are:

1. The effective dimension of stimuli is rather high even though a standard LED viewing booth and targets are used.
2. No parameters are optimized to weight the influence of residual errors and assumptions.

2 Methodology

In this work we consider a discrete representation of spectra by sampling each spectrum $f(\lambda)$ at N equidistant wavelengths that cover approximately the sensitivity range of the camera. The result is a spectral vector $\mathbf{f} = (f(\lambda_1), \ldots, f(\lambda_N))^T$. A target is utilized with m different patches having the reflectances $\mathbf{r_1}, \ldots, \mathbf{r_m}$. This target is captured k times under the LED light sources with the spectral power distributions (SPD) $\mathbf{l_1}, \ldots, \mathbf{l_k}$. Considering a linear imaging system with n channels we obtain $m \cdot k$ stimuli that are captured by the camera and produce the sensor responses

$$\mathbf{c^{i,j}} = \begin{pmatrix} c_1^{i,j} \\ \vdots \\ c_n^{i,j} \end{pmatrix} = \delta^{i,j} \cdot \begin{pmatrix} \mathbf{r_i}^T D(\mathbf{l_j})\mathbf{s_1} \\ \vdots \\ \mathbf{r_i}^T D(\mathbf{l_j})\mathbf{s_n} \end{pmatrix}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, k \tag{1}$$

where D is an operator that converts an N-dimensional vector into a $N \times N$-dimensional diagonal matrix and $\mathbf{s_1}, \ldots, \mathbf{s_n}$ are the unknown spectral channel sensitivities. The factor $\delta^{i,j}$ depends on the radiance of the jth LED and the measurement geometry for patch i. The geometry is defined by the distance and angle to the camera and illuminating LED.

Unfortunately, the measurement geometry differs usually between patches of the target and from the geometry that is used to spectrophotometrically measure $\mathbf{r_i}$ (typically circumferential 45°/0°). Please note in this regard that for accurately modeling the imaging system using a constant factor $\delta^{i,j}$ according to eq. (1) the patches have to be nearly Lambertian surfaces.
Fortunately, we do not need to determine $\delta^{i,j}$, $i = 1, \ldots, m$, $j = 1, \ldots, k$ if we transform eq. (1) into chromaticities (as done e.g. by Ebner [2]). For channel $x = 1, \ldots, n$ the resulting equations become

$$\frac{c_x^{i,j}}{\sum_{y=1}^{n} c_y^{i,j}} = \frac{\mathbf{r_i}^T D(\mathbf{l_j})\mathbf{s_x}}{\sum_{y=1}^{n} \mathbf{r_i}^T D(\mathbf{l_j})\mathbf{s_y}}, \quad i = 1, \ldots, m, \quad j = 1, \ldots, k. \tag{2}$$

These equations can be rearranged into a simple linear equation system $\mathbf{\Lambda s} = 0$, where $\mathbf{s} = (\mathbf{s_1}^T, \ldots, \mathbf{s_n}^T)^T$ and $\mathbf{\Lambda}$ is a $(k \cdot m \cdot n) \times (n \cdot N)$ dimensional matrix that depends on $\mathbf{c^{i,j}}$, $\mathbf{r_i}$, $\mathbf{l_j}$ ($i = 1, \ldots, m$, $j = 1, \ldots, k$). In a real application we have to deal with noise ϵ, so that our equation system is not homogeneous but has the form

$$\mathbf{\Lambda s} = \epsilon. \tag{3}$$

In this paper we assume that the spectral sensitivity $\mathbf{s_x}$ for channel $\mathbf{x}$ is smooth and the correlation between sensitivities corresponding to different wavelengths can be modeled by a Toeplitz matrix, i.e.

$$\mathbf{K_x} = \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{N-1} \\ \rho & 1 & \rho & \cdots & \rho^{N-2} \\ \rho^2 & \rho & 1 & & \vdots \\ \vdots & \vdots & & \ddots & \rho \\ \rho^{N-1} & \rho^{N-2} & \cdots & \rho & 1 \end{pmatrix} \tag{4}$$

where the correlation coefficient ρ is selected to 0.99. Furthermore, we assume that there is no correlation between spectral sensitivities of different channels, which does not mean that they cannot overlap. With this assumption the correlation matrix $\mathbf{K}$ for $\mathbf{s}$ is block diagonal and has the form

$$\mathbf{K} = \mathbf{I_n} \otimes \mathbf{K_x} \tag{5}$$

where $\mathbf{I_n}$ is the $n \times n$ dimensional identity matrix and "$\otimes$" denotes the Kronecker matrix product.
We want to select the most likely of those sensitivities that satisfy eq. (3) and have no negative parts. This requires the solution of a constrained maximum-a-posteriori problem with the likelihood model shown in eq. (3) and a prior distribution induced by the correlation matrix $\mathbf{K}$ (see eq. (5)). Solving such a problem is especially convenient if we assume all distributions to be Gaussian and allow uncorrelated zero-mean noise ϵ that is statistically independent from the sensitivities. In this case, the posterior distribution is also Gaussian and we can formulate a quadratic programming (QP) problem as follows

$$\text{minimize} \quad \mathbf{s}^T(\mathbf{K} - \mathbf{W\Lambda K})^{-1}\mathbf{s} \tag{6}$$

$$\text{subject to} \quad \mathbf{s} \geq 0 \tag{7}$$

$$\|\mathbf{s}\|_1 = 1 \tag{8}$$

where $\mathbf{W} = \mathbf{K}\mathbf{\Lambda}^T(\mathbf{\Lambda}\mathbf{K}\mathbf{\Lambda}^T + \sigma_\epsilon^2\mathbf{I_{Nn}})^{-1}$ is the Wiener filter matrix, σ_ϵ^2 is the noise variance and $\mathbf{I_{Nn}}$ is the $(N \cdot n) \times (N \cdot n)$ dimensional identity matrix. The matrix $\mathbf{K} - \mathbf{W}\mathbf{\Lambda}\mathbf{K}$ in the objective function is the covariance matrix of the posterior distribution. The solution of the QP problem are sensitivities $\mathbf{s} = (\mathbf{s_1}^T, \ldots, \mathbf{s_n}^T)^T$ that maximize the density of the posterior distribution, are non-negative (see constraint (7)) and normalized to one (see constraint (8)). Without this normalization the resulting sensitivities would be zero. As a result we obtain only relative sensitivities. However, this is completely sufficient for estimating reflectances because they are typically normalized to the reflectance spectrum of a captured white reference (i.e. the scaling factor is eliminated anyway).

3 Experimental Setup

The proposed method was used to determine the sensitivities of a six-channel ($n = 6$) modified *Sinar 54H*-based camera that was setup in our institute for capturing artwork. The RGB-based camera back of the Sinar 54H camera has micro-positioning capabilities: the sensor is moved four times to position each color of the Bayer patterned *color filter array* (CFA) over every pixel's spatial location. Therefore, demosaicing is not necessary. By mounting a two-stage filter wheel equipped with a blue and a yellow filter in front of the lens the camera allows us to capture six channel (12 bit) 22 mega pixel images. The concept is adapted from a setup developed at the Munsell Color Science Laboratory (Rochester Institute of Technology) [1].
To generate LED illuminants we used JUST NORMLICHT's *LED Color Viewing Light* booth that utilizes six different LED ($k = 6$). A color checker chart with $m = 24$ test patches was spectrally characterized using X-Rite's EyeOne Pro spectrophotometer and placed in the viewing booth. For each of the six LED illuminants the target was captured by the camera in an unidirectional 0°/45° geometry. During each capture we measured the radiance of the target's white patch in a similar geometry using a Konica-Minolta *CS1000* spectroradiometer. From this measurement and the patch's reflectance we calculated the illuminant's SPD. The experimental setup is shown in Figure 1(a). The stimuli $\mathbf{r_i}^T D(\mathbf{l_j})$ were calculated by multiplying the SPD of any LED illuminant $\mathbf{l_j}$, $j = 1, \ldots, 6$ with the reflectance spectra $\mathbf{r_i}$, $i = 1, \ldots, 24$ of the Color Checker. The resulting 144 spectral stimuli are shown in Figure 1(b). The effective dimension of those stimuli as defined by Hardeberg [4] for an accumulated energy of 99% is 15. Using only one illuminant, e.g. the broadband white LED, the effective dimension would be only 9.
One limiting factor in our setup is the wavelength range of the spectrophotometer (X-Rite's EyeOne Pro) that was used to measure the reflectance spectra of the Color Checker. Since this range spans only 380nm - 730nm it does not cover the whole wavelength range of our camera system that is approximately 380nm - 760nm. As a consequence the spectral stimuli that would be used for calculation are insufficient for reconstructing the sensitivities. For this reason we mounted an interference UV/IR cutoff filter in front of the lens and treated it as part of the camera. The filter is used typically for imaging applications and passes only visible light in the wavelength range from 400nm - 700nm. Therefore, spectra were also sampled in the wavelength range from 400 nm - 700 nm in 5 nm steps resulting in $N = 61$.
The camera responses used to recover the sensitivities were obtained by averaging all pixel responses of a target patch. Since many thousand pixels are used for each patch we could reduce the influence of noise drastically. Nevertheless, we set the noise variance not

to zero but to $\sigma_\epsilon^2 = 5 \cdot 10^{-7}$ to allow for measurement errors of the stimuli (Lambertian surface assumption etc.). Please note that it is unlikely that those errors follow a normal distribution and the maximum-a-posteriori approach might provide not the optimal reconstruction result in such cases.

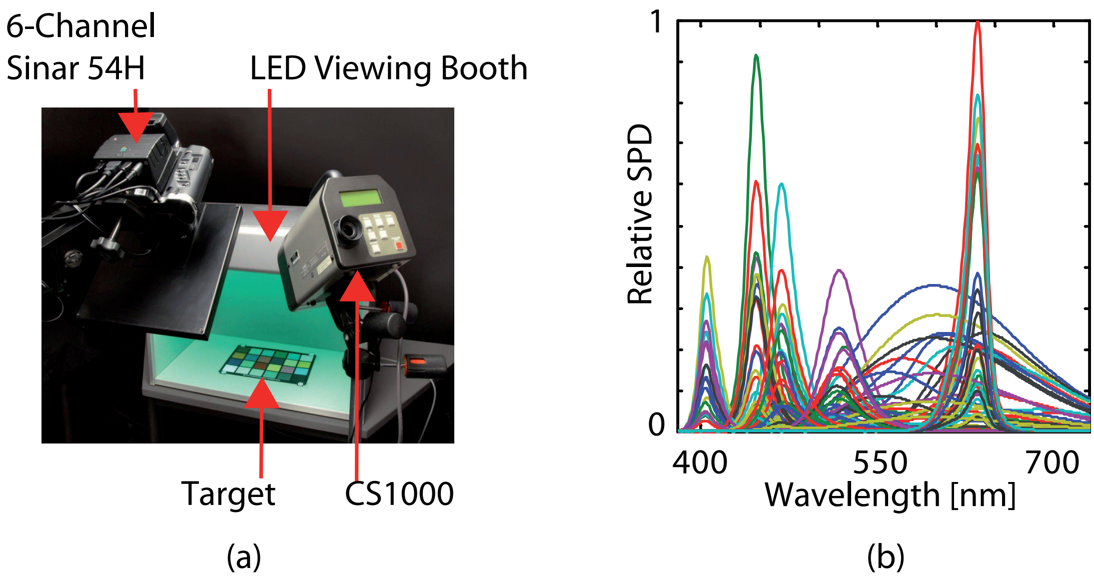

Figure 1: (a) Experimental setup (b) Stimuli of the Color Checker for the six LED-illuminants

4 Results and Discussion

The calculated channel sensitivities of the modified Sinar 54H camera are shown in Figure 2(a). Since we do not know the real channel sensitivities we can only validate our results by comparing real camera responses with predicted responses calculated using eq. (1). For this comparison we need to normalize each prediction $\hat{\mathbf{c}}$ to the magnitude of the corresponding real response $\mathbf{c}$ as follows

$$\hat{\mathbf{c}}^{\text{new}} = \frac{\|\mathbf{c}\|_1}{\|\hat{\mathbf{c}}\|_1}\hat{\mathbf{c}}. \tag{9}$$

The predicted and measured digital counts are then compared by calculating the difference $|\hat{c}_i^{\text{new}} - c_i|$ for each channel $i = 1, \ldots, 6$.

In a first stage we look on the prediction accuracy of camera responses for the training stimuli, i.e. for the 144 Color Checker-based stimuli that were used to calculate the sensitivities. Figure 2(b) shows the measured vs. predicted digital counts (DC). The points are lying on a line showing a good prediction performance. The mean, 95th-percentile and maximum absolute differences for all channels between real and predicted digital counts are shown in table 1 (please note that the channels have a 12 bit depth). We also show the difference as the percentage of the channel's maximum digital count under same illuminant (usually for the white patch).

In terms of absolute and relative errors the prediction performance is good for most applications. These results are not astonishing since the validation is performed on the

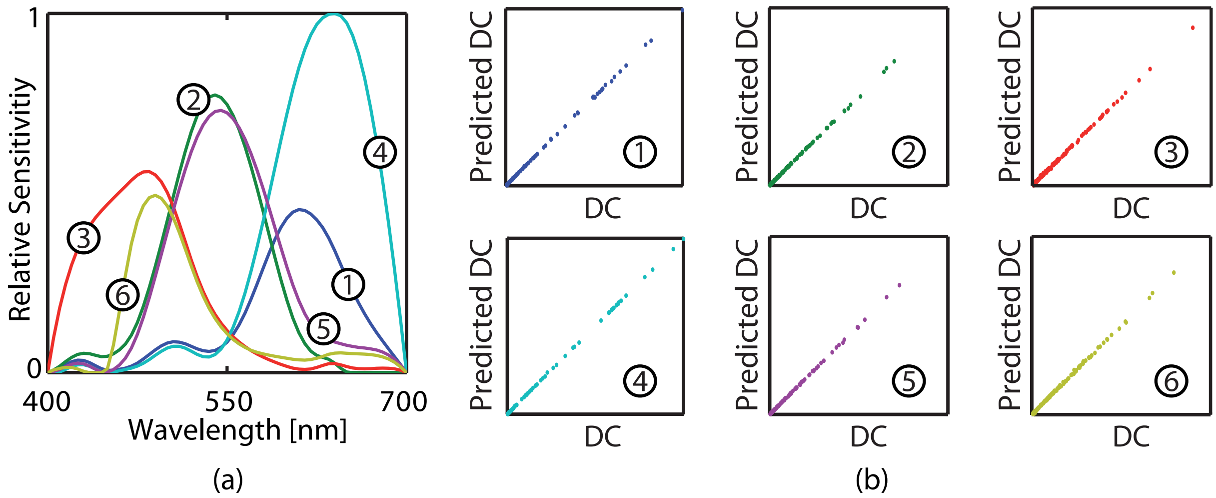

Figure 2: (a) Recovered channel sensitivities of the modified Sinar 54H camera (b) Measured vs. predicted digital counts (DC) for the 144 training stimuli. The numbers correspond to the channel sensitivities shown in (a)

	absolute[a]			**relative to max DC [%]**		
channel	**mean**	**95th**[b]	**max**	**mean**	**95th**[b]	**max**
1	1.96	6.77	11.40	0.17	0.57	0.97
2	2.02	5.56	13.23	0.14	0.39	0.92
3	3.34	9.08	13.85	0.32	0.86	1.32
4	3.07	9.27	26.33	0.12	0.36	1.02
5	1.96	5.48	12.53	0.12	0.34	0.77
6	1.96	4.65	8.72	0.30	0.71	1.34

[a] 12 bit depth per channel

[b] 95th percentile = value below which 95% of observations falls

Table 1: Prediction performance of camera responses for 144 training stimuli (Color Checker under six LED illuminants).

training stimuli. More interesting is to investigate the prediction performance on new stimuli. For this reason we captured the Esser TE221 test chart (IEC 61966-8) consisting of 264 color patches and 19 gray patches using the same setup. The reflectance spectra of the Esser target have a much larger spectral dimensionality than the spectra of the Color Checker. For the six LED illuminants a total of 1698 stimuli were calculated as described above and used to predict the camera responses. The results are shown in table 2 and as histograms in figure 3.

As expected the prediction errors are larger than the errors for the training stimuli. However, the mean errors are still small and below 0.5% of the maximum camera response. Even the 95th percentiles are around 1% of the maximum responses. This indicates that the sensitivities could be used to model the camera quite accurately for most stimuli. In contrast to predictions of the test stimuli the maximum errors are quite high (especially for the third and sixth channel). In such cases estimating reflectances would probably lead to large spectral root mean square and color errors. Apart from inaccurate sensitivity

	absolute[a]			relative to max DC [%]		
channel	**mean**	**95th**[b]	**max**	**mean**	**95th**[b]	**max**
1	2.31	8.37	29.45	0.20	0.73	2.57
2	3.14	12.09	55.33	0.24	0.92	4.19
3	4.42	12.71	74.92	0.45	1.28	7.55
4	6.77	30.25	91.00	0.26	1.14	3.44
5	2.83	10.61	55.25	0.19	0.71	3.71
6	2.43	7.51	34.26	0.39	1.22	5.54

[a]12 bit depth per channel

[b]95th percentile = value below which 95% of observations falls

Table 2: Prediction performance of camera responses for 1698 test stimuli (Esser TE221 under six LED illuminants).

estimation for distinct wavelengths (e.g. due to a still insufficient spectral variability of test stimuli for those wavelengths) one possible explanation of the relatively large maximum errors could be the Lambertian surface assumption that is not always valid for color targets. As a result the stimuli calculated by multiplying the reflectance spectra (measured by a 45°/0° circumferential geometry) with the SPD of the illuminant (measured by the spectroradiometer) are not the stimuli that are captured by the camera. Further investigations of the targets are required to validate this hypothesis.

5 Conclusion

We investigated the reconstruction of camera sensitivities using captured stimuli generated by a standard reflectance target sequentially illuminated using multiple LEDs. Such stimuli have a higher effective dimension than stimuli generated with the same target using only a single illuminant. For a computational reconstruction method a high spectral dimension of stimuli is beneficial to reduce the influence of assumptions on the reconstruction.

To allow for different measurement geometries we transform the camera responses into chromaticities assuming a Lambertian surface of the target patches. The sensitivity reconstruction is performed by solving a constrained maximum-a-posteriori problem resulting in positive sensitivities that accurately predict the camera responses for the test stimuli. We tested the approach using a six channel camera, a novel LED viewing booth and a Color Checker target. The reconstructed sensitivities were validated by predicting camera responses from stimuli that were generated using an Esser TE221 target illuminanted by multiple LED illuminants. The mean prediction error was below 0.5% of the maximum camera response and the 95th percentile around 1%. Those values are reasonable for many color correction or spectral reconstruction methods. The maximum errors show a larger disagreement (up to 7.5% of the maximum camera response) which might result from an inaccurate sensitivity estimation for particular wavelengths due to a still insufficient spectral variability of test stimuli for those wavelengths. Another explanation could be an invalid Lambertian surface assumption for distinct patches of the target.

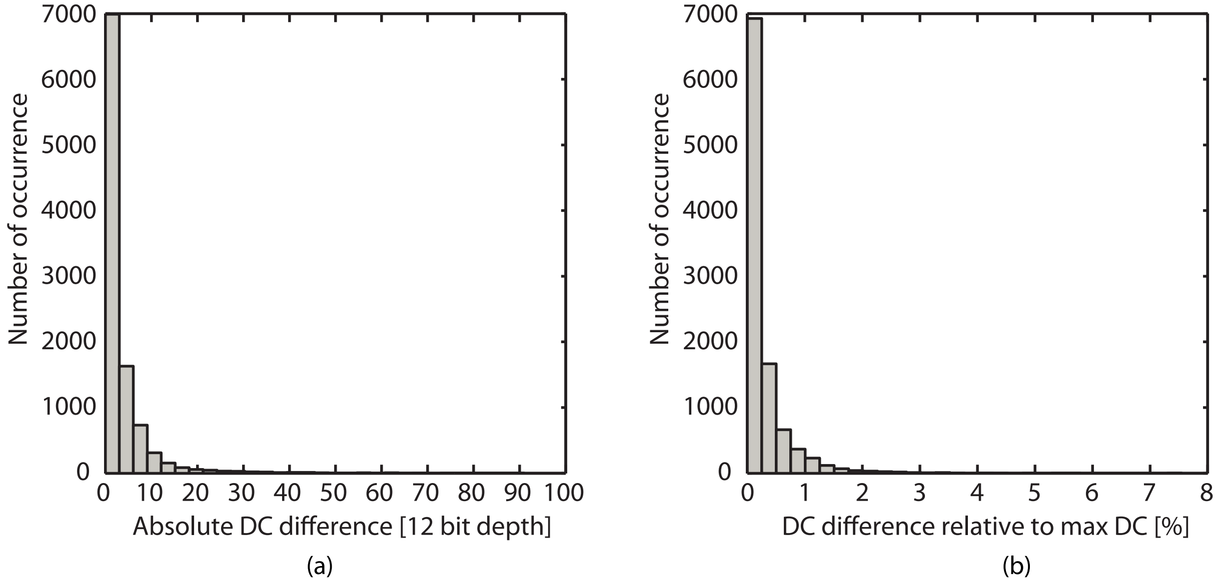

Figure 3: Histogram of differences between predicted and measured camera responses (digital counts - DC) for the 1698 test stimuli (Esser TE221 under six LED illuminants), i.e. 10188 channel values. (a) Absolute difference (b) Difference relative to the maximum digital count (white patch).

6 Acknowledgments

We thank Maria Fedutina, Tanja Kaulitz and Henri Kröling for developing the two-stage filter wheel that allows us to increase the number of camera channels of the Sinar 54H camera to six, and Karsten Rettig and Manfred Jakobi for their accurate work of manufacturing and assembling the wheel. The research was financially supported by the German Research Foundation (Deutsche Forschungsgemeinschaft).

References

[1] Berns, R. S. and L. A. Taplin: *Practical Spectral imaging Using a Color-Filter Array Digital Camera.* MCSL Technical Report, 2006.

[2] Ebner, M.: *Estimating the spectral sensitivity of a digital sensor using calibration targets.* In *Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pages 642–649. ACM, 2007.

[3] Finlayson, G. D., S. Hordley and P. M. Hubel: *Recovering Device Sensitivities with Quadratic Programming.* In *IS&T/SID*, pages 90–95, Scottsdale Ariz., 1998.

[4] Hardeberg, J. Y.: *On the Spectral Dimensionality of Object Colours.* In *CGIV*, pages 480–485, Poitiers, France, 2002. IS&T.

[5] König, F.: *Die Charakterisierung von Farbsensoren.* PhD thesis, Rheinisch-Westfälische Technische Hochschule Aachen, Germany, 2001.

[6] Sharma, G. and H. J. Trussell: *Characterization of Scanner Sensitivity.* In *IS&T/SID*, pages 103–107, Scottsdale Ariz., 1993.

[7] SHARMA, G. and H. J. TRUSSELL: *Set Theoretic Estimation in Color Scanner Characterization*. Journal of Electronic Imaging, 5:479–489, 1996.

[8] URBAN, P.: *Metamere und multispektrale Methoden zur Reproduktion farbiger Vorlagen*. PhD thesis, Technische Universität Hamburg-Harburg, Germany, 2005. BoD, ISBN 3833426659.

[9] URBAN, P. and R.-R. GRIGAT: *The Metamer Boundary Descriptor Method for Color Correction*. Journal of Imaging Science and Technology, 49:417–429, 2005.

[10] URBAN, P., M. R. ROSEN and R. S. BERNS: *Spectral Image Reconstruction using an Edge Preserving Spatio-Spectral Wiener Estimation*. Journal of the Optical Society of America A, 26(8):1868–1878, 2009.