

Spatio-Spectral Image Restoration

Christoph Godau and Philipp Urban; Institute of Printing Science and Technology, Technische Universität Darmstadt; Magdalenenstr. 2, 64289 Darmstadt, Germany

Abstract

The recovery of spectral reflectances from camera responses is usually composed of several distinct operations. We propose a new approach that connects edge-preserving denoising, deblurring and spectral reconstruction. Each of the steps is based on actual physical properties of the camera system that can be obtained using established methods, and the filter eliminates the need for manual sharpening of the images. Results on both real images and synthetic data show significant improvements over previous methods for spectral reconstruction.

Introduction

The recovery of spectral reflectances from camera responses is usually composed of several distinct operations. Denoising, deblurring and spectral reconstruction are performed in separate steps.

Ideally, the camera image acquisition process can be modeled as a series of linear operations. Light is reflected on a surface, passes through the optical system of the camera and is discretely sampled on the sensor. In the most general case, light, surface reflectance, and Point Spread Function (PSF) are spatially varying. Our goal is to recover the original reflectance of the surface. In a controlled environment, e.g. when capturing artwork in a museum, we can assume the light source and distribution, the PSF and the sensor sensitivities to be known.

Images can be blurred because the optical properties of the filters commonly used in multi-spectral imaging can differ significantly, causing problems with focusing. At the moment, standard practice is the manual application of sharpening as a post-processing step after reconstruction has been performed. Noise is introduced because filters absorb a large percentage of the light, requiring high integration times. Denoising is therefore essential, especially for darker colors and filters with narrow spectral bands.

In general, there exists an infinite set of metameric reflectances that lead to the same camera response under given viewing conditions. The inversion of the image formation process is therefore inherently ill-posed, i.e. there is no unique solution. A variety of constraints can be used to regularize the problem, typically imposing smoothness or *a priori* knowledge about the reflectance spectra encountered in a particular application.

In this paper, we describe an image reconstruction approach that combines denoising, deblurring, and reconstruction of spectral reflectances into a single processing chain. Each step makes use of actual physical properties of the camera system.

We provide an overview of relevant previous approaches to spectral image reconstruction in the next Section and then describe our approach to inverting the image formation process using a Wiener filter. Results on both real images and synthetic data are presented and discussed in the last two Sections.

Related Works

Since the field of image restoration is extremely broad, we restrict our overview to approaches that are most relevant to our work. In particular, we focus on spectral reconstruction methods based on Wiener filters. Wiener filters have long been a popular tool in image reconstruction and have been successfully used for spectral reconstruction, denoising and deblurring [1]. The linear nature of the Wiener filter makes it possible to combine a chain of operators into a single filter, a property that we exploit to combine deblurring and reconstruction.

The spatio-spectral Wiener filter (SSW) of Murakami *et al.* [2] assumes a stationary image, where the image spatial autocorrelation does not depend on the position. This is efficient for denoising, but leads to additional blurring in the final reconstruction results. Still, the spectral reconstruction results show lower errors than for separate denoising and reconstruction.

Urban *et al.*'s [3] edge-preserving spatio-spectral Wiener (EPSSW) filter avoids this blurring by calculating noise distributions in the local neighborhood of each pixel based on bilateral filters, but the obtained images still require additional sharpening due to the inherent blur of the imaging system.

Our work can be seen as a combination and extension of these two methods, which are therefore described in detail in the next section. We propose a way to connect spatio-spectral image reconstruction and deblurring. The computational power available today makes it relatively straightforward to combine these steps, even though the complexity and memory requirements can increase significantly compared to solving each problem on its own.

Methodology

Spectral reconstruction is a highly underconstrained problem; typically an infinite number of solutions are possible. Additional information, for example on the general type of reflectance spectra encountered in an application, can significantly improve results. We use the following *a priori* information for our n -channel system, all of which can be acquired using established methods as described below:

- Covariance of the reflectance spectra \mathbf{K}_r
- Noise variance \mathbf{K}_ε
- Camera spectral sensitivities $\mathbf{D} = (\mathbf{d}_1, \dots, \mathbf{d}_n)^T$
- System PSF \mathbf{P} of size $m \times m$, m odd
- Illuminant spectral power distribution (SPD) \mathbf{p}

The covariance of reflectance spectra can be determined from a low-resolution sampling of typical reflectance spectra encountered in the scene. Databases of such spectra exist or can be created for specific applications such as capturing natural scenes, paintings or prints [4]. Noise variance can be estimated from the actual images [5] or uniform patches. The camera PSF can be

determined using random noise targets [6, 7]. Illuminant SPD and camera sensitivities are measured using standard devices or target-based methods [8, 9].

Camera Model

We represent the reflectance of a pixel as a column vector \mathbf{r} . Using the common simplified linear camera model not accounting for blurring, responses are calculated as

$$\mathbf{c} = \mathbf{DLr} + \boldsymbol{\varepsilon} = \boldsymbol{\Omega}\mathbf{r} + \boldsymbol{\varepsilon} \quad (1)$$

where $\boldsymbol{\varepsilon}$ is additive noise, \mathbf{L} is a diagonal matrix containing the SPD of the illuminant \mathbf{p} as diagonal elements and \mathbf{c} is the vector of camera responses (three-dimensional for RGB cameras, six-dimensional in our system). The camera sensitivities, reflectance spectra and illuminant SPD are discretely sampled at wavelengths $\lambda_1, \dots, \lambda_l$. The system matrix $\boldsymbol{\Omega}$ is the combination of camera sensitivities and illuminant.

Figure 1 illustrates the steps involved in the image formation process (including blurring). The original multi-spectral image is blurred by the system PSF, sampled to camera responses using the spectral sensitivities, and noise is added to each channel. The result is a blurred and noisy camera image that is used as input for our algorithm, which attempts to invert the process.

The standard Wiener filter (W)

The basic reflectance estimating Wiener filter \mathbf{W} for this model is the matrix

$$\mathbf{W} = \mathbf{K}_r \boldsymbol{\Omega}^T (\boldsymbol{\Omega} \mathbf{K}_r \boldsymbol{\Omega}^T + \mathbf{K}_\varepsilon)^{-1} \quad (2)$$

and the Wiener reflectance estimation is then $\hat{\mathbf{r}} = \mathbf{W}\mathbf{c}$.

The spatio-spectral Wiener filter (SSW)

Image noise can significantly degrade the performance of Wiener reconstruction. To reduce noise we can make use of the sensor responses in a local window around each pixel. The reflectance vector \mathbf{r}_s for the central pixel c in a window of size $w \times w$ is formed by appending the reflectances of all pixels in the window:

$$\mathbf{r}_s = \left(\mathbf{r}_1^T \mathbf{r}_2^T \cdots \mathbf{r}_c^T \cdots \mathbf{r}_w^T \right)^T \quad (3)$$

and the camera response vector \mathbf{c}_s is constructed in the same manner by appending the camera responses for all pixels in the window.

The spatio-spectral system matrix $\boldsymbol{\Omega}_s$ is a block-diagonal matrix containing the system matrix once for each pixel in the window along the diagonal:

$$\boldsymbol{\Omega}_s = \mathbf{I}_{w^2} \otimes \boldsymbol{\Omega} \quad (4)$$

where " \otimes " is the Kronecker product and the camera response for the window can then be calculated as $\mathbf{c}_s = \boldsymbol{\Omega}_s \mathbf{r}_s + \boldsymbol{\varepsilon}$.

Integrating stationary local image correlation for denoising, the reflectance estimating spatio-spectral filter proposed by Murakami *et al.* [2] is calculated as

$$\mathbf{W}_{SSW} = \mathbf{Q} \mathbf{K}_{rx} \boldsymbol{\Omega}_s^T [\boldsymbol{\Omega}_s \mathbf{K}_{rx} \boldsymbol{\Omega}_s^T + \mathbf{I}_{w^2} \otimes \mathbf{K}_\varepsilon]^{-1} \quad (5)$$

where \mathbf{I}_{w^2} is an $w^2 \times w^2$ -dimensional unity matrix, \mathbf{K}_{rx} is the spatio-spectral $lw^2 \times lw^2$ covariance matrix, \mathbf{Q} is an $l \times lw^2$ matrix that retains the values for only the central pixel of each window, and l is the number of discrete wavelengths as defined in Equation 1. Assuming the spectral and spatial correlations are separable, we can write \mathbf{K}_{rx} as

$$\mathbf{K}_{rx} = \mathbf{K}_x \otimes \mathbf{K}_r \quad (6)$$

where \mathbf{K}_x is an $w^2 \times w^2$ stationary spatial correlation matrix for the pixels in an $w \times w$ window. The calculations are described in detail by Murakami *et al.* [2]. This method significantly reduces the impact of noise but assumes a stationary image correlation which blurs edges in noisy images. This is addressed by the denoising method partly described below. Furthermore, the SSW method does not consider the system's PSF.

Bayesian denoising for edge-preserving spatio-spectral Wiener filtering (EPSSW)

We closely follow the denoising method proposed by Urban *et al.* [3]. Based on bilateral weightings accounting for spatial distance and range difference, we compute *maximum a posteriori* (MAP) values $\hat{\mathbf{c}}(i, j)$ for the noise reduced pixels and the corresponding *posteriori* noise covariance matrix $\hat{\mathbf{K}}_\varepsilon(i, j)$. This information is then used in all subsequent steps which reduces the impact of noise by making use of reasonable assumptions on prior noise statistics.

According to Urban *et al.* [3], the noise-reducing Wiener filter for a pixel position (i, j) is given as

$$\mathbf{W}_d(i, j) = \hat{\mathbf{K}}_\varepsilon(i, j) [\hat{\mathbf{K}}_\varepsilon(i, j) + \mathbf{K}_\varepsilon]^{-1}, \quad (7)$$

and its noise reduced value $\hat{\mathbf{c}}(i, j)$ with covariance matrix $\hat{\mathbf{K}}_\varepsilon(i, j)$ of the remaining noise are given by

$$\hat{\mathbf{c}}(i, j) = \mathbf{W}_d(i, j) [\mathbf{c}(i, j) - \bar{\mathbf{c}}(i, j)] + \bar{\mathbf{c}}(i, j). \quad (8)$$

$$\hat{\mathbf{K}}_\varepsilon(i, j) = \hat{\mathbf{K}}_\varepsilon(i, j) - \mathbf{W}_d(i, j) \hat{\mathbf{K}}_\varepsilon(i, j), \quad (9)$$

where $\bar{\mathbf{c}}(i, j)$ is the estimate of the noise-free camera response with corresponding covariance matrix $\hat{\mathbf{K}}_\varepsilon(i, j)$. Both are computed by a bilaterally weighted sum of pixels in the $w \times w$ window \mathbf{F}_{ij} centered at pixel (i, j)

$$\bar{\mathbf{c}}(i, j) = \sum_{(k,l) \in \mathbf{F}_{ij}} \mathbf{c}(k, l) w_{\text{bilateral}}[(k, l), (i, j)], \quad (10)$$

$$\hat{\mathbf{K}}_\varepsilon(i, j) = \mathbf{C}(i, j) \mathbf{C}(i, j)^T \quad (11)$$

where $\mathbf{C}(i, j)$ is an $n \times w^2$ matrix containing the weighted pixels $\mathbf{c}(k, l) w_{\text{bilateral}}[(k, l), (i, j)]$ in \mathbf{F}_{ij} , with $(k, l) \in \mathbf{F}_{ij}$. The weight $w_{\text{bilateral}}$ is the product of spatial w_{spatial} and range w_{range} weights defined as

$$w_{\text{spatial}}[(k, l), (i, j)] = \exp\left(-\frac{(i-k)^2 + (j-l)^2}{2\sigma_{\text{spatial}}^2}\right) \quad (12)$$

$$w_{\text{range}}[(k, l), (i, j)] = \exp\left(-\frac{\|\mathbf{c}(k, l) - \mathbf{c}(i, j)\|_2^2}{2\sigma_{\text{range}}^2}\right) \quad (13)$$

where $\mathbf{c}(k, l)$ and $\mathbf{c}(i, j)$ are the corresponding noisy sensor responses and the parameters $\sigma_{\text{spatial}}^2$ and σ_{range}^2 control the decay of the two weight factors.

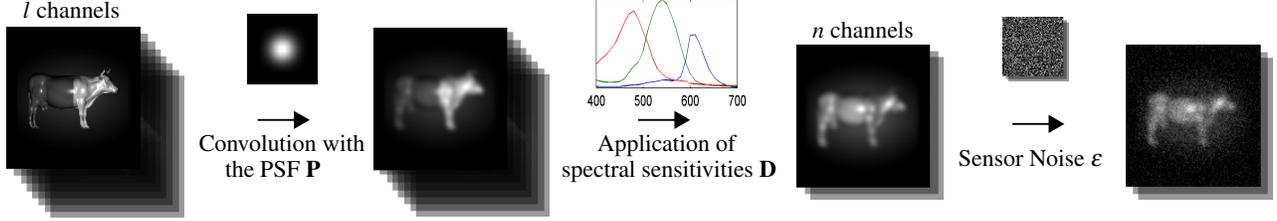


Figure 1. In the simplified linear image formation model, the original multi-spectral image (assumed to be discretely sampled at wavelengths $\lambda_1, \dots, \lambda_l$) is blurred by the system PSF, sampled to camera responses using the n spectral sensitivities, and noise is added to each channel. The result is a blurred and noisy camera image that is used as input for our algorithm, which attempts to invert the process.

For derivations and more details on the calculations refer to the original paper by Urban *et al.* [3].

We use $\hat{\mathbf{c}}(i, j)$ and $\hat{\mathbf{K}}_{\epsilon}(i, j)$ to attenuate the influence of noise in the subsequent spatio-spectral restoration.

Our method

With integration of the system PSF \mathbf{P} , we obtain a new spatio-spectral system lighting and blurring matrix Ω_{sh} that includes the PSF instead of the identity matrix used in Equation (4):

$$\Omega_{\text{sh}} = \mathbf{H} \otimes \Omega \quad (14)$$

where \mathbf{H} is the $w^2 \times w^2$ system blur matrix calculated from the PSF \mathbf{P} . \mathbf{H} allows us to write the blurred version $\hat{\mathbf{v}}$ of an input window \mathbf{v} of size $w \times w$ (in vector form) as $\hat{\mathbf{v}} = \mathbf{H}\mathbf{v}$. The complete image formation model calculating the camera responses \mathbf{c}_s for an $w \times w$ window is then written as

$$\mathbf{c}_s = \Omega_{\text{sh}} \mathbf{r}_s + \epsilon. \quad (15)$$

For a \mathbf{P} of size $m \times m$, the size w of the window is set to $2m - 1$ to ensure that the central pixel has the correct value after applying both forward and backward transformations. The camera response values are then correctly calculated for a region of size $m \times m$ in the center of the window. A larger window can be chosen to increase the denoising effect, but in most cases this should not be necessary.

We then adapt the Wiener inverse of Equation (5) to include the new system lighting and blur matrix Ω_{sh} , as well as the *posteriori* noise covariances from Equation (11) for all window pixels:

$$\mathbf{W}_{\text{Ours}} = \mathbf{Q} \mathbf{K}_{\text{rx}} \Omega_{\text{sh}}^{\text{T}} [\Omega_{\text{sh}} \mathbf{K}_{\text{rx}} \Omega_{\text{sh}}^{\text{T}} + \mathbf{N}_{\epsilon}]^{-1} \quad (16)$$

where \mathbf{N}_{ϵ} is now an $w^2 n \times w^2 n$ block-diagonal matrix, containing the *posteriori* noise covariances $\hat{\mathbf{K}}_{\epsilon}(i, j)$ for every pixel in the $w \times w$ window as block-diagonal elements.

We model the prior distribution of reflectance spectra by calculating the mean spectrum $\bar{\mathbf{r}}$ of a representative set of reflectances, similar to the one used in calculating the spectral covariance matrix. After subtracting $\bar{\mathbf{r}}$ from our reflectances, we obtain a zero-mean distribution that can be approximated by a normal distribution as assumed by the Wiener filter.

The final spectral reconstruction is then calculated from the noise reduced camera responses $\hat{\mathbf{c}}_s = (\hat{\mathbf{c}}_1^{\text{T}}, \hat{\mathbf{c}}_2^{\text{T}}, \dots, \hat{\mathbf{c}}_{w^2}^{\text{T}})^{\text{T}}$ (note that $\hat{\mathbf{c}}_i$ is defined in Equation (8)) as

$$\hat{\mathbf{r}} = \mathbf{W}_{\text{Ours}} [\hat{\mathbf{c}}_s - (\mathbf{I}_{w^2} \otimes \Omega) \bar{\mathbf{r}}_s] + \bar{\mathbf{r}} \quad (17)$$

where $\bar{\mathbf{r}}_s = (\bar{\mathbf{r}}^{\text{T}}, \dots, \bar{\mathbf{r}}^{\text{T}})^{\text{T}}$ is an lw^2 -dimensional vector.

Results

We compare our method to a basic Wiener reconstruction (W), the spatio-spectral reconstruction (SSW) of Murakami *et al.* [2] and the edge-preserving spatio-spectral filter (EPSSW) of Urban *et al.* [3]. We use the following three images in our experiments:

Cow: a segment of the public domain spectral metamerism test target METACOW [10] (Figure 3). The image has been designed so that the front and rear part of the cow are metameric under D_{65} , but differ in color under most other lights, such as Illuminant A.

Squares: a synthetic image with 9 colored rectangles generated from Munsell reflectance spectra (Figure 5).

Girl: a real-world image from the Joensuu spectral database [11] (Figure 4).

The camera images are modeled by taking an original illuminated spectral image, blurring it using the system point spread function, applying the spectral sensitivities of the camera and adding noise as described in Equation (15) and illustrated in Figure 1.

Figure 2 shows the spectral sensitivities of the camera used in the simulations and the system point spread function, as well as an example of a simulated camera image. The PSF is similar in shape and scale to point spread functions measured for high-quality cameras using random noise targets [7, 6]. For the sake of simplicity we use the same stationary PSF for all channels, although the model allows for using a different spatially varying PSF for each channel.

The reconstruction results for the "Cow" image are shown in Figure 3. An example of standard Wiener reconstruction is included for this image, but the reconstruction results for the front and rear are clearly not metameric under D_{65} . Since all other methods perform significantly better, we do not show the standard Wiener results for the other images.

Figure 4 shows the reconstruction results for the "Girl" image, rendered under illuminant D_{65} . The synthetic "Squares" image in Figure 5 contains the sharpest edges of all test images. The previous methods do not account for blurring, the reconstruction results are therefore just as blurred as the camera image.

The table below shows the results for different noise levels, using a constant PSF. For all three images, our method leads to a significantly smaller RMSE (root mean square error). The Signal-to-Noise Ratio (SNR) for an image I with mean signal power \bar{I} and

mean noise power $\bar{\epsilon}$ is calculated as $\text{SNR} = 20 * \log_{10}(\bar{I}/\bar{\epsilon})$.

RMSE for increasing noise levels

SNR	Image	W	SSW	EPSSW	Ours
42	cow	0,028	0,025	0,025	0,016
	squares	0,035	0,035	0,033	0,012
	girl	0,046	0,046	0,045	0,038
32	cow	0,036	0,028	0,027	0,019
	squares	0,037	0,036	0,033	0,015
	girl	0,048	0,047	0,046	0,039
22	cow	0,045	0,036	0,035	0,030
	squares	0,040	0,039	0,035	0,024
	girl	0,058	0,049	0,047	0,044

In all cases the standard Wiener filter (W) performs worst, the spatio-spectral Wiener filter (SSW) shows lower errors especially for higher noise levels, and the edge-preserving spatio-spectral Wiener filter (EPSSW) obtains the best results among all previous methods. The edge-preserving properties of the EPSSW filter only lead to small improvements due to the lack of sharp edges in the blurred images. As expected, our method shows clear improvements for blurred images, since none of the previous methods take blurring into account. Only for very high noise levels ($\text{SNR} < 20$) the deconvolution can lead to strong artifacts, and the error can be larger than for the other methods. If the PSF is not used in the reconstruction process, no deblurring is performed and our method performs identically to EPSSW.

Discussion

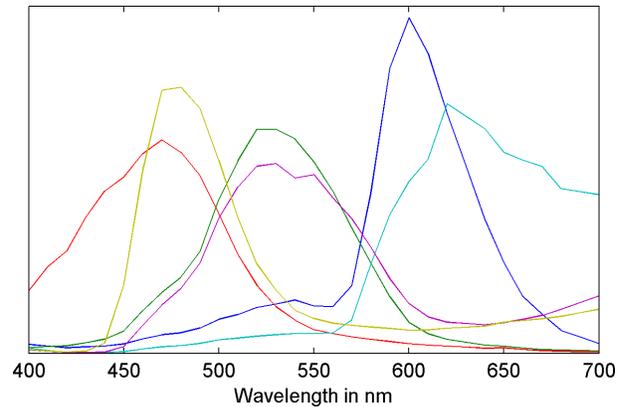
The integration of deconvolution, denoising and spectral reconstruction works well in our simulations, and is based on actual measured properties of the imaging system. On both synthetic and test images, the obtained reconstruction results exhibit a lower RMSE and are visually superior to the previous approaches when the camera images are blurred. If no deblurring is applied, our approach is identical to the current state-of-the art, the edge preserving Wiener filter of Urban *et al.* [3].

For images with a very strong noise component, the deconvolution can lead to artifacts. Under controlled conditions this is rarely of concern, the noise levels used in our experiments are already higher than those typically encountered in high-quality spectral camera systems. For other applications it is possible to reduce the deconvolution component of the operator, avoiding the occurrence of artifacts while still providing some amount of deblurring. A reference implementation of the algorithm is available on the authors homepage ¹.

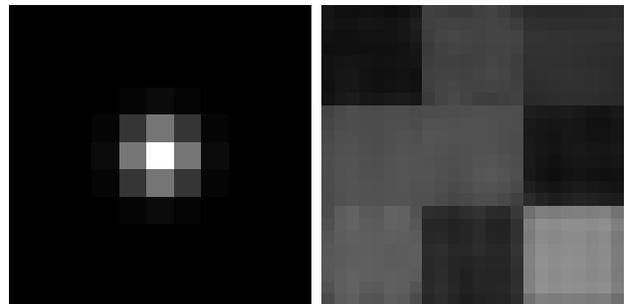
References

- [1] MR R Banham and AK K Katsaggelos. Digital image restoration. *IEEE Signal Processing Magazine*, 14(2):24–41, 1997.
- [2] Yuri Murakami, Ken Fukura, Masahiro Yamaguchi, and Na-gaaki Ohya. Color reproduction from low-SNR multispectral images using spatio-spectral Wiener estimation. *Optics express*, 16(6):4106–20, March 2008.

¹Address removed for anonymous submission.



(a) The spectral sensitivities of our 6-channel test system.



(b) The system PSF.

(c) A simulated camera image, showing the first channel of the "squares" image.

Figure 2. Our simulated test system consists of a 6-channel camera with sensitivities as shown in 2a and a single PSF that is used to blur all channels 2b. The result is a blurred and noisy image for each channel 2c.

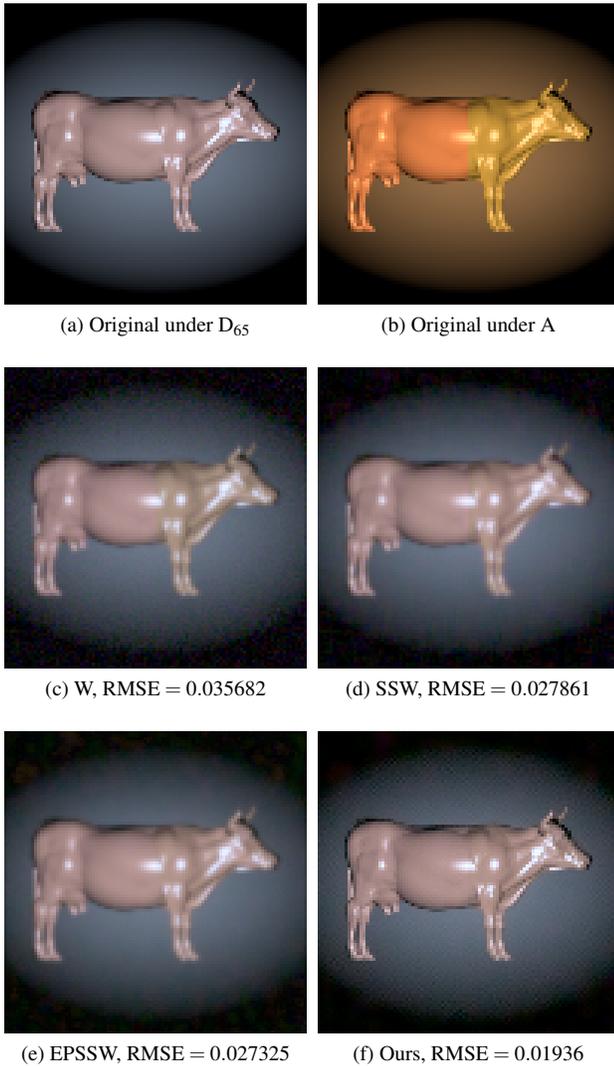


Figure 3. Reconstruction results for the "cow" image, rendered under illuminant D_{65} . It is a segment of the public domain spectral metamerism test target METACOW [10], which was designed so that the front and rear part of the cow are metameric under D_{65} .



Figure 4. Reconstruction results for the "girl" image, rendered under illuminant D_{65} . The image is available from the Joensuu spectral database [11].



Figure 5. Reconstruction results for the "squares" image, rendered under illuminant D_{65} . The 9 colored rectangles were generated from randomly selected Munsell reflectance spectra.

Free, Metameric, Extended-Dynamic-Range, Spectral Test Target for Imaging System Analysis and Simulation. In *IS&T/SID, 12th Color Imaging Conference*, pages 239–245, Scottsdale Ariz., 2004.

[11] Spectral Color Research Group. Spectral Image Database.

Author Biography

Christoph Godau received his BS in physics (2008) from the Ruhr-University Bochum, Germany and his MS in colour science (2010) from the University of Saint-Etienne, France and the University of Eastern Finland. He has previously worked as a research engineer at the KU Leuven, Belgium, and the Center for Mathematical Morphology, France. He is currently working on color imaging and colorimetry at the TU Darmstadt, Germany.

Philipp Urban received the M.S. degree in mathematics from the University of Hamburg, Germany, and the Ph.D. degree from the Hamburg University of Technology, Hamburg, Germany. He has been the head of the Color Research Group, Institute of Printing Science and Technology, Technische Universität Darmstadt, Darmstadt, Germany, since 2009. His research focuses on spectral-based acquisition, processing, and reproduction of color images considering the limited metameric and spectral gamut and low dynamic range of output devices. From 2006 to 2008, he was a Visiting Scientist at the Munsell Color Science Laboratory, Center for Imaging Science, Rochester Institute of Technology, Rochester, NY.

- [3] Philipp Urban, Mitchell R Rosen, and Roy S Berns. Spectral image reconstruction using an edge preserving spatio-spectral Wiener estimation. *Journal of the Optical Society of America A*, 26(8):1865–1875, 2009.
- [4] Oili Kohonen, Jussi Parkkinen, and Timo Jääskeläinen. Databases for spectral color science. *Color Research & Application*, 31(5):381–390, October 2006.
- [5] Noriyuki Shimano. Recovery of spectral reflectances of objects being imaged without prior knowledge. *IEEE Transactions on Image Processing*, 15(7):1848–56, July 2006.
- [6] Mauricio Delbracio, Pablo Musé, Andrés Almansa, and Jean-Michel Morel. The Non-parametric Sub-pixel Local Point Spread Function Estimation Is a Well Posed Problem. *International Journal of Computer Vision*, 96(2):175–194, June 2011.
- [7] Johannes Brauers, Claude Seiler, and Til Aach. Direct PSF estimation using a random noise target. In Francisco H. Imai, Nitin Sampat And, and Feng Xiao, editors, *Digital Photography*, page 7537. SPIE, 2010.
- [8] A. Alsam and G. Finlayson. Recovering spectral sensitivities with uncertainty. In *CGIV*, pages 22–26, Poitiers, France, 2002. IS&T.
- [9] P. Urban, M. Desch, K. Happel, and D. Spiehl. Recovering Camera Sensitivities using Target-based Reflectances Captured under multiple LED-Illuminations. In *16. Workshop Farbbildverarbeitung der German Color Group*, pages 9–16, Ilmenau, Germany, 2010.
- [10] M. D. Fairchild and G. M. Johnson. METACOW: A Public-Domain, High-Resolution, Fully-Digital, Noise-