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 $K_r$
- Noise variance $K_e$
- Camera spectral sensitivities $D = (d_1, \ldots, d_n)^T$
- System PSF $P$ of size $m \times m$, $m$ odd
- Illuminant spectral power distribution (SPD) $p$

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 \( r \). Using the common simplified linear camera model not accounting
for blurring, responses are calculated as

\[
c = DLr + \varepsilon = \Omega r + \varepsilon
\]  

(1)

where \( \varepsilon \) is additive noise, \( L \) is a diagonal matrix containing the
SPD of the illuminant \( p \) as diagonal elements and \( c \) 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 \( \lambda_1, \ldots, \lambda_l \). The system matrix \( \Omega \) 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 \( W \) for this model is the matrix

\[
W = K_{\varepsilon} \Omega^T (\Omega K_{\varepsilon} \Omega^T + K_{r})^{-1}
\]  

(2)

and the Wiener reflectance estimation is then \( \hat{r} = Wc \).

**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 reflec-
tance vector \( r_w \) for the central pixel \( c \) in a window of size \( w \times w \)
formed by appending the reflectances of all pixels in the win-
dow:

\[
r_w = \begin{bmatrix} r_1^T r_2^T \cdots r_c^T \cdots r_w^T \end{bmatrix}^T
\]  

(3)

and the camera response vector \( c_w \) is constructed in the same man-
ner by appending the camera responses for all pixels in the win-
dow.

The spatio-spectral system matrix \( \Omega_{s} \) is a block-diagonal matrix
containing the system matrix once for each pixel in the win-
dow along the diagonal:

\[
\Omega_{s} = I_{ws} \otimes \Omega
\]  

(4)

where ‘\( \otimes \)’ is the Kronecker product and the camera response for
the window can then be calculated as \( c_{s} = \Omega_{s} r_{w} + \varepsilon \).

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

\[
W_{SSW} = Q K_{\varepsilon} \Omega_{s}^T [\Omega_{s} K_{\varepsilon} \Omega_{s}^T + I_{ws} \otimes K_{\varepsilon}]^{-1}
\]  

(5)

where \( I_{ws} \) is an \( w^2 \times w^2 \)-dimensional unity matrix, \( K_{\varepsilon} \) is the
spatio-spectral \( lw^2 \times lw^2 \) covariance matrix, \( Q \) 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 \( K_{\varepsilon} \) as

\[
K_{\varepsilon} = K_{s} \otimes K_{r}
\]  

(6)

where \( K_{r} \) 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 (EPSW)**

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 \( \hat{c}(i, j) \) for the noise reduced pixels and the corre-
sponding posteriori noise covariance matrix \( \hat{K}_{c}(i, j) \). 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

\[
W_{d}(i, j) = \hat{K}_{c}(i, j) \hat{K_{c}}(i, j)^{-1} + K_{r}^{-1},
\]  

(7)

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

\[
\hat{c}(i, j) = W_{d}(i, j)c(i, j) - \hat{c}(i, j) + \tilde{c}(i, j),
\]  

(8)

\[
\hat{K}_{c}(i, j) = K_{c}(i, j) - W_{d}(i, j)K_{c}(i, j),
\]  

(9)

where \( \tilde{c}(i, j) \) is the estimate of the noise-free camera response
with corresponding covariance matrix \( K_{c}(i, j) \). Both are com-
puted by a bilaterally weighted sum of pixels in the \( w \times w \) window
\( F_{ij} \) centered at pixel \((i, j)\)

\[
\tilde{c}(i, j) = \sum_{(k,l) \in F_{ij}} c(k,l)w_{bilateral} [(k,l), (i,j)],
\]  

(10)

\[
K_{c}(i, j) = C(i,j)C(i,j)^T
\]  

(11)

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

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

(12)

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

(13)

where \( c(k,l) \) and \( c(i,j) \) are the corresponding noisy sensor
responses and the parameters \( \sigma_{spatial}^2 \) and \( \sigma_{range}^2 \) control the decay of
the two weight factors.
For derivations and more details on the calculations refer to the original paper by Urban et al. [3].

We use $\hat{e}(i, j)$ and $K_{s}(i, j)$ to attenuate the influence of noise in the subsequent spatio-spectral restoration.

Our method

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

$$\Omega_{sh} = H \otimes \Omega$$  \hspace{1cm} (14)

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

$$e_{s} = \Omega_{sh}r_{s} + \varepsilon.$$  \hspace{1cm} (15)

For a $P$ 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 $\Omega_{sh}$, as well as the posteriori noise covariances from Equation (11) for all window pixels:

$$W_{Ours} = QK_{s}T\Omega_{sh}T\Omega_{sh}T\Omega_{sh}T + N_{\varepsilon}^{-1}$$  \hspace{1cm} (16)

where $N_{\varepsilon}$ is now an $w^2n \times w^2n$ block-diagonal matrix, containing the posteriori noise covariances $K_{s}(i, j)$ 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 $\bar{r}$ of a representative set of reflectances, similar to the one used in calculating the spectral covariance matrix. After subtracting $\bar{r}$ 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 $\hat{e}_{s} = (\hat{e}_{1}, \hat{e}_{2}, ..., \hat{e}_{n})^{T}$ (note that $\hat{e}_{s}$ is defined in Equation (8)) as

$$\hat{r} = W_{Ours}[\hat{e}_{s} - (I_{w^2} \otimes \Omega)\hat{r}_{s}] + \bar{r}$$  \hspace{1cm} (17)

where $\hat{r}_{s} = (\hat{r}_{1}, \hat{r}_{2}, ..., \hat{r}_{n})^{T}$ is an $hw^2$-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 (EPSWW) 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

\[ \text{SNR} = 10 \log_{10} \frac{\bar{I}}{\sigma^2}, \]

where $\sigma^2$ is the variance of the noise.
mean noise power $\bar{\varepsilon}$ is calculated as $\text{SNR} = 20 \times \log_{10}(\bar{I}/\bar{\varepsilon})$.

### RMSE for increasing noise levels

<table>
<thead>
<tr>
<th>SNR</th>
<th>Image</th>
<th>W</th>
<th>SSW</th>
<th>EPSSW</th>
<th>Ours</th>
</tr>
</thead>
<tbody>
<tr>
<td>42</td>
<td>cow</td>
<td>0.028</td>
<td>0.025</td>
<td>0.025</td>
<td><strong>0.016</strong></td>
</tr>
<tr>
<td></td>
<td>squares</td>
<td>0.035</td>
<td>0.035</td>
<td>0.033</td>
<td><strong>0.012</strong></td>
</tr>
<tr>
<td></td>
<td>girl</td>
<td>0.046</td>
<td>0.046</td>
<td>0.045</td>
<td><strong>0.038</strong></td>
</tr>
<tr>
<td>32</td>
<td>cow</td>
<td>0.036</td>
<td>0.028</td>
<td>0.027</td>
<td><strong>0.019</strong></td>
</tr>
<tr>
<td></td>
<td>squares</td>
<td>0.037</td>
<td>0.036</td>
<td>0.033</td>
<td><strong>0.015</strong></td>
</tr>
<tr>
<td></td>
<td>girl</td>
<td>0.048</td>
<td>0.047</td>
<td>0.046</td>
<td><strong>0.039</strong></td>
</tr>
<tr>
<td>22</td>
<td>cow</td>
<td>0.045</td>
<td>0.036</td>
<td>0.035</td>
<td><strong>0.030</strong></td>
</tr>
<tr>
<td></td>
<td>squares</td>
<td>0.040</td>
<td>0.039</td>
<td>0.035</td>
<td><strong>0.024</strong></td>
</tr>
<tr>
<td></td>
<td>girl</td>
<td>0.058</td>
<td>0.049</td>
<td>0.047</td>
<td><strong>0.044</strong></td>
</tr>
</tbody>
</table>

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 synthetically 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 Address removed for anonymous submission.
Figure 3. Reconstruction results for the “cow” image, rendered under illuminant D65. 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 D65.

Figure 4. Reconstruction results for the “girl” image, rendered under illuminant D65. The image is available from the Joensuu spectral database [11].
Figure 5. Reconstruction results for the “squares” image, rendered under illuminant D65. The 9 colored rectangles were generated from randomly selected Munsell reflectance spectra.