Spectral imaging based on 2D diffraction patterns and a regularization model

MAX NIMMER,[1,2,*] **GABRIELE STEIDL,**[2] **RAINER RIESENBERG,**[1] **AND ANDREAS WUTTIG**[1]

[1]*Leibniz Institute of Photonic Technology, Albert-Einstein-Str. 9, 07745 Jena, Germany*
[2]*Department of Mathematics, Technische Universität Kaiserslautern, Gottlieb-Daimler-Str. 47, 67663 Kaiserslautern, Germany*
[*]*nimmer@mathematik.uni-kl.de*

Abstract: We present a variational approach to reconstruct a multispectral image from an array of diffraction patterns captured in one monochromatic image. The mathematical model relies on the superposition of wavelength dependent calibration diffraction patterns and a spectral regularization with second order differences. The necessary preprocessing steps which influence the mathematical setting are discussed in detail. For computing a minimizer of our model we propose an active set method with model specific modifications. We validate our approach by experimental results for spectra within the range of 400 nm to 700 nm.

1. Introduction

Multispectral imaging is an important tool in various applications, as both a higher spectral resolution in the visible light range than in RGB images as well as additional wavelength bands in the UV and NIR may offer valuable information. In recent years, there has been an increasing demand for handheld snapshot spectral imaging devices for mobile applications such as point-of-care medical diagnostics, agriculture, environmental control and food evaluation [1–3]. Several approaches have been proposed to tackle this problem. Modern monochromatic sensors provide a high number of pixels, thus it is reasonable to trade some of it for spectral information. The most straightforward method is similar to the Bayer pattern in RGB cameras. Directly in front of the sensor an array of spectral filters is employed and the resulting image is processed into a multispectral image [4]. The usually used Fabry-Perot resonators are hard to manufacture and require precise alignment with the sensor pixels which limits the adaptability of this approach.

Another approach is the insertion of a pinhole array in front of a prism with re-imaging optics as demonstrated in [5]. The spectral image can then be obtained by remapping the sensor pixels into a spectral data cube. The pinhole array allows the use of an interchangeable objective to change the field of view and prevents crosstalking of neighboring spectral pixels. A disadvantage is that the optical components necessary for the prism prohibit a compact design.

Other approaches are computationally more complex. Coded aperture snapshot spectral imaging (CASSI) [6] demultiplexes a measured image by solving a large optimization problem which assumes the sparsity of the spectral image in the gradient domain. The method reconstructs the whole spectral image at once and the experimental setup requires a coded aperture, a prism and additional lenses.

In [7] the possibility of reconstructing multispectral images with a lensless setup with a strongly scattering medium by a deconvolution algorithm is shown and the performance for simple scenes with collimated light is demonstrated. The concept is based on a deconvolution algorithm which relies on the orthogonality of the point spread functions which limits the performance in more complex scenes and the expandability to more wavelength bands.

The high-sensitivity color imaging device in [8], which can also be used as a snapshot multispectral imager, is based on a diffractive filter array in front of the sensor. Their transparent

diffractive filters create a diffraction pattern on the sensor which varies by wavelength, so that the spectrum of incoming light can be calculated by solving an inverse problem. This allows for pixelwise reconstruction, but the crosstalking of the array cells poses a problem. The authors propose two methods to deal with crosstalking. In [8] they undersample the image, which reduces the spatial resolution, but allows the reconstruction of spatially smooth scenes with collimated light. In [9] they propose a computational way to tackle crosstalking. It reduces the influence of crosstalking, but good reconstructions are still only possible for simple scenes which are recorded under circumstances very similar to the ones in their calibration procedure. Furthermore, the number of calibration images in this approach is the product of the number of spatial pixels and the number of wavelength bands, which limits the expandability to higher dimensions.

The reconstruction of spectral data in [7, 9, 10] is computationally difficult as their approaches involve both a spatial and spectral multiplexing. The inversion of this process relies on a low cross-correlation, or even orthogonality, of the diffraction patterns which is usually not given in practice. These approaches do not allow pixelwise reconstruction of spectra as all spectral pixels are coupled in their respective optimization problem.

In this paper, we propose a method for spectral imaging based on diffraction patterns which allows for compact devices, but requires an appropriate mathematical reconstruction model which is the main focus of this paper. In our approach the crosstalk of spatial pixels is practically eliminated at hardware level, so that we only have spectral multiplexing. We capture a monochromatic image of an array of diffraction patterns from which pixelwise reconstruction of spectral data is possible which we prefer to reconstructing the whole spectral cube at once by the following reasons: First, the spatial image structure does not influence the reconstructed spectra. Second, from a computational viewpoint, we can solve many small optimization problems so that pixelwise reconstruction allows for parallelization. Third, errors in the recorded data such as oversaturated pixels from unwanted reflections only lead to reconstruction errors in the corresponding pixel. Fourth, pixelwise reconstruction allows to choose a region of interest and to reconstruct only that part of the image which saves computation time.

We employ a cell-based diffractive structure composed of an array of currently binary transmissive diffractive optical elements (DOEs) which were optimized for the task. Compared to alternatively possible phase DOEs, they allow to eliminate some technological uncertainties and costs for the initial demonstration of the sensor principle. Further, our setup includes a pinhole array to select well-defined spatial image positions having the advantage of eliminating crosstalk and allowing the use of an interchangeable objective to change the field of view. In contrast to [5] we do not use re-imaging optics, so that our device is more compact. All components are easy to manufacture and our system is robust to inaccuracies due to a calibration procedure.

As a consequence of the setup we have to reconstruct the spectral information from the diffraction patterns. We propose a mathematical model which assumes that the recorded diffraction pattern is related to the superposition of wavelength dependent calibration diffraction patterns and includes prior information by Tikhonov type regularization based on second order differences. In this context we give details on the necessary data processing steps which allow us to extract consistent reflectance spectra. We apply an efficient algorithm to solve the resulting partly nonnegative least square problem. In order to demonstrate the capability of our system we present results from both simulated and experimental data.

2. Principle

The basic layout of our spectral imager is shown in Fig. 1(a). An objective with an f-number of 8 images the scene onto the pinhole array in order to discretize the scene and prevent crosstalking in the subsequent steps. The pinholes are in the imaging plane of the objective and have a pitch of 71.5 μm along with a diameter of 4 μm. The pinhole array is on one side of a 400 μm thick glass chip, while on the opposite side of it there is a DOE array which is aligned with the pinhole

array. Each of the square DOEs has an edge length of 50 µm, which corresponds approximately to the size of the light cone created by the pinhole. We experimentally verified that the chosen DOE pitch of 71.5 µm practically eliminates crosstalk. Both sides of the pinhole-DOE chip are produced lithographically. The DOE is optimized for an air distance of 200 µm to the image sensor, however due to our currently limited alignment capabilities it is instead in a distance of between 300 µm and 500 µm to the image sensor. These dimensions were chosen to prevent crosstalking of the pinhole-DOE pairs while allowing good dispersion properties of the DOE. Due to our current non-optimal alignment, especially the dispersion in the NIR range suffers. As this design adds only a chip directly in front of the sensor the resulting camera is as compact as a standard camera. The prototype used in our experiments is shown in Fig. 1(b). We note that the distance of the pinholes determines the spatial sampling, while the pinhole size determines the amount of light available per spatial pixel. This means that the signal-to-noise ratio in each pinhole-DOE unit cell depends on the pinhole size and not on the total amount of light blocked by the pinhole array. The size of each pinhole is related to the sensor pixel size and the intended spectral resolution, it should be around one to two times the sensor pixel size. If it is chosen too small, the sensor pixel signal-to-noise ratio suffers. If it is chosen too large the spectral information might be undersampled and the highest spectral resolution might not be utilized. This connection is similar to the one in [5] with the exception that we cannot easily increase the size of the pinholes, since a certain spatial coherence of the light cone produced by it must be maintained.

(a)

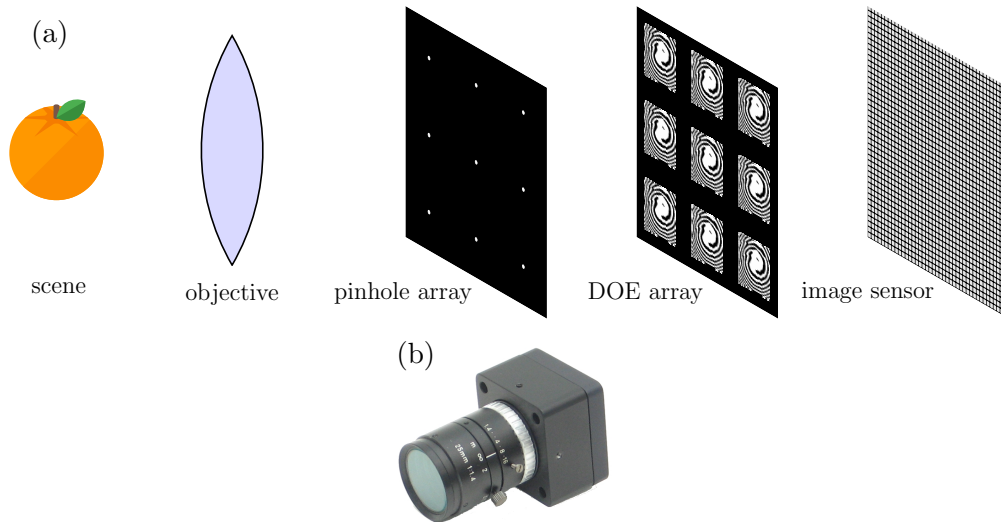scene   objective   pinhole array   DOE array   image sensor

(b)

Fig. 1. (a) Schematic of our multispectral imaging device consisting of objective, pinhole array, DOE array and image sensor. (b) Photography of the compact device used to capture the experimental data. The whole setup has the form factor of a small C-Mount camera, in this case a $40 \times 40 \times 27$ mm camera case with a 25 mm objective.

Each pinhole-DOE-pair creates a square diffraction pattern on the sensor in which the spectral information is encoded. We reconstruct the spectral abundances as follows: For each wavelength whose abundance we want to reconstruct, we record the calibration diffraction pattern specific to this wavelength, see Fig. 2(b) for examples at one pinhole-DOE pair. Note that the calibration diffraction patterns for the same wavelength are different for each pinhole-DOE pair due to optical path distortions and manufacturing inaccuracies. These calibration images can then be used to reconstruct the spectrum of light coming through a pinhole, since the diffraction pattern of light

with a certain spectrum is approximately the sum of the wavelength-specific diffraction patterns weighted by the abundances.
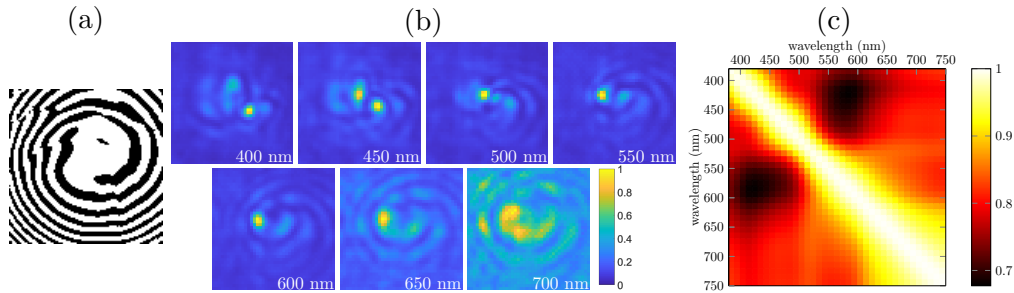

Fig. 2. (a) Optimized DOE, where white corresponds to transmissive areas and black to blocked light. (b) Measured calibration diffraction patterns for 400 nm to 700 nm in 50 nm steps for a central pinhole-DOE pair. All images normalized by maximum intensity. (c) Cross-correlation of recorded calibration patterns of different wavelength divided by their respective norms for a central pinhole-DOE pair.

Our ability to get useful reconstructions depends on the dissimilarity of the diffraction calibration patterns for the considered wavelengths at each pinhole-DOE pair. The similarity can be measured by the cross-correlation. In order to achieve best possible results, the DOE was optimized taking the cross-correlation matrix of the corresponding (simulated, see also Subsection 5.1) calibration patterns at a central pinhole-DOE pair into account. The concrete optimization procedure is beyond the scope of this paper. The resulting DOE is shown in Fig. 2(a). The cross correlation of the experimentally measured diffraction patterns for different wavelengths can be found in Fig. 2(c). The diagram shows a decorrelation of the calibration patterns which becomes worse above 600 nm due to the dispersion properties of the DOE.

3. Mathematical modeling

In the following, we describe the mathematical model and the computational steps for the reconstruction of the spectra from the diffraction patterns. Due to the acquisition of the diffraction patterns in the monochromatic image with suppressed crosstalking, we can use a pixelwise approach. Our variational model is based on the assumption that the measured diffraction pattern in one pixel can be approximated by the linear combination of basic wavelength dependent diffraction patterns. Moreover, it incorporates the non-negativity of the wavelength abundances and their „smoothness" along the spectral dimension within the regularizer of the model. Finally, this results in a partly non-negative least squares problem. We apply an active set method with model specific modifications to find a high precision minimizer.

3.1. Formalization

Let each position in the grid $\mathcal{G} := \{(i, j)| i = 1, \dots, n_1, j = 1, \dots, n_2\}$ correspond to a pinhole-DOE pair. By $\boldsymbol{F}$ we denote the array of recorded diffraction pattern images $\boldsymbol{F}(i, j)$, $i = 1, \dots, n_1, j = 1, \dots, n_2$, each having a size of $s \times s$ pixels. From each of the diffraction pattern images $\boldsymbol{F}(i, j)$ we aim to reconstruct a spectrum $\boldsymbol{h}(i, j)$ with L components $h^l(i, j)$ corresponding to wavelengths λ_l. In other words, we are interested in the spectral cube with entries $h^l(i, j)$.

To this end, let $\boldsymbol{P}^l(i, j) \in \mathbb{R}^{s,s}$, $l = 1, \dots, L$ be the basic diffraction pattern at position $(i, j) \in \mathcal{G}$ corresponding to wavelength λ_l. These basic diffraction patterns are obtained by scaling wavelength depending calibration patterns within a white balance step as described

in Subsection 4.4. Note that the basic diffraction patterns $\boldsymbol{P}^l(i, j) \in \mathbb{R}^{s,s}$, $l = 1, \ldots, L$, are different in each grid point (i, j). The basic assumption is that in each grid point (i, j) a measured diffraction pattern $\boldsymbol{F}(i, j)$ can be approximated by a linear combination of the basic diffraction patterns $\boldsymbol{P}^l(i, j)$, i.e.,

$$\boldsymbol{F}(i, j) \approx \sum_{l=1}^{L} h^l(i, j) \boldsymbol{P}^l(i, j) \tag{1}$$

with $h^l(i, j) \geq 0$. This means that it is necessary to sample the spectral range densely enough in the calibration process to have for all practical purposes a complete frame of basic diffraction patterns. In practice we oversample the spectral range to justify this assumption.

3.2. Mathematical model

As described in Section 2, we can work pixelwise. For simplicity of notation, we omit the grid position in the following and simply write, e.g., $\boldsymbol{P}^l$ instead of $\boldsymbol{P}^l(i, j)$.

In order to reconstruct a spectrum $\boldsymbol{h} \in \mathbb{R}^L$ which conforms to certain prior information we propose a variational model. The general form of such a model is given by

$$\arg \min_{\boldsymbol{h}} \mathcal{D}(\boldsymbol{h}; \boldsymbol{F}, \boldsymbol{P}^1, \ldots, \boldsymbol{P}^L) + \alpha \mathcal{R}(\boldsymbol{h}), \quad \alpha > 0, \tag{2}$$

where $\mathcal{D}(\boldsymbol{h}; \boldsymbol{F}, \boldsymbol{P}^1, \ldots, \boldsymbol{P}^L)$ is a data fitting term which accounts for the calibration patterns and the measured data, $\mathcal{R}(\boldsymbol{h})$ is a regularizer which incorporates assumptions on the wanted coefficients $\boldsymbol{h}$ and α is a regularization parameter which balances the influence of both.

We reshape the calibration diffraction patterns $\boldsymbol{P}^l \in \mathbb{R}^{s,s}$, $l = 1, \ldots, L$ columnwise into vectors $\boldsymbol{p}^l \in \mathbb{R}^S$, $S := s^2$, and use them as columns of a matrix $\boldsymbol{P} := (\boldsymbol{p}^1 \ \ldots \ \boldsymbol{p}^L) \in \mathbb{R}^{S,L}$, $S > L$. The matrix $\boldsymbol{P}$ has in general a full column rank. Similarly, we reshape the recorded image $\boldsymbol{F} \in \mathbb{R}^{s,s}$ into a vector $\boldsymbol{f} \in \mathbb{R}^S$ and as before $\boldsymbol{h}$ denotes the sought spectrum. Based on our linearity assumption, we consider the data term

$$\mathcal{D}(\boldsymbol{h}, c; \boldsymbol{f}, \boldsymbol{P}) = \frac{1}{2} \|\boldsymbol{P}\boldsymbol{h} + c\boldsymbol{1} - \boldsymbol{f}\|_2^2 = \frac{1}{2} \left\| (\boldsymbol{P} \ \boldsymbol{1}) \begin{pmatrix} \boldsymbol{h} \\ c \end{pmatrix} - \boldsymbol{f} \right\|_2^2, \tag{3}$$

where $\boldsymbol{1}$ is the constant vector with S components equal to one, $\| \cdot \|_2$ denotes the Euclidean norm and the additional unknown $c \in \mathbb{R}$ appears in order to account for residual dark current and stray light in both $\boldsymbol{P}$ and $\boldsymbol{f}$. The matrix $(\boldsymbol{P} \ \boldsymbol{1}) \in \mathbb{R}^{S,L+1}$, $S > L + 1$ has full rank, so that the optimization problem of Eq. (2) with the data term in Eq. (3) and $\alpha = 0$ has a unique solution, but the reconstruction is susceptible to noise. In order to remedy this, we add a Tikhonov type regularization [11] which incorporates the prior assumption that the reflectance spectrum we reconstruct is smooth in the sense that the squared second order differences of the spectrum is small. This is the case for many natural objects such as food, plants, tissue and soil. Additionally, we add the nonnegativity of $\boldsymbol{h}$ as a constraint. This amounts to consider the following variational model:

$$\arg \min_{\boldsymbol{h} \geq \boldsymbol{0}, c \in \mathbb{R}} \frac{1}{2} \left\| (\boldsymbol{P} \ \boldsymbol{1}) \begin{pmatrix} \boldsymbol{h} \\ c \end{pmatrix} - \boldsymbol{f} \right\|_2^2 + \frac{\alpha}{2} \|\boldsymbol{\Gamma}\boldsymbol{h}\|_2^2, \quad \alpha > 0. \tag{4}$$

Here $\boldsymbol{\Gamma}$ is the matrix which models second order differences with constant step size

$$\boldsymbol{D}_2 := \boldsymbol{D}_1^{\mathrm{T}} \boldsymbol{D}_1 = \begin{pmatrix} 1 & -1 & 0 & & & \\ -1 & 2 & -1 & & & \\ & & & \ddots & & \\ & & & -1 & 2 & -1 \\ & & & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{L,L}. \tag{5}$$

and $\boldsymbol{D}_1$ is the first order difference matrix

$$\boldsymbol{D}_1 := \begin{pmatrix} -1 & 1 & & \\ & \ddots & \ddots & \\ & & -1 & 1 \end{pmatrix} \in \mathbb{R}^{L-1,L}, \tag{6}$$

which we will be used for a comparison later. In the case of non-uniform wavelength step size $|\lambda_{l+1} - \lambda_l|$ we can adapt these matrices accordingly. We can rewrite Eq. (4) as

$$\underset{\boldsymbol{h} \geq \boldsymbol{0}, c \in \mathbb{R}}{\arg \min} \frac{1}{2} \left\| \begin{pmatrix} \boldsymbol{P} & \boldsymbol{1} \\ \sqrt{\alpha}\boldsymbol{\Gamma} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{h} \\ c \end{pmatrix} - \begin{pmatrix} \boldsymbol{f} \\ \boldsymbol{0} \end{pmatrix} \right\|_2^2, \tag{7}$$

which is a suitable form for the chosen optimization algorithm.

3.3. Optimization algorithm

There are many algorithms available for solving constrained least squares problems, see for example [12, 13] for an overview. Typically, we have a small number of variables L, one for each wavelength in the calibration set. Consequently, direct methods such as active set methods are feasible and preferable to iterative methods for solving Eq. (7), as they produce an exact solution after a finite number of steps depending on L. The standard active set method for a nonnegative least squares (NNLS) problem

$$\underset{\boldsymbol{x} \geq \boldsymbol{0}}{\arg \min} \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|_2^2 \tag{8}$$

for $\boldsymbol{A} \in \mathbb{R}^{m,n}$ with $m > n$ and $\boldsymbol{b} \in \mathbb{R}^m$ was given in [14]. We modify the algorithm in two ways to better suit our setting. First, in Eq. (4) the variables are only partly nonnegative since it contains the unconstrained variable $c \in \mathbb{R}$. This can be incorporated into the algorithm by introducing a set of indices that may be negative and keeping those in the passive set of the method. Second, since we reconstruct the spectrum in each pixel separately and neighboring pixels can be expected to have similar spectra and thereby similar passive sets, we add the possibility to start the algorithm with a specific initial passive set in order to speed up convergence. In practice, we propose after reconstructing the spectrum $\boldsymbol{h}(i, j)$, to use the optimal passive set from this pixel as the initial passive set for the reconstruction on the neighboring spectrum $\boldsymbol{h}(i, j + 1)$. Likewise we could also work row-wise. A similar idea was also mentioned in [15].

We introduce some notation for better understanding of the algorithm. With J we denote the passive set, i.e., the subset of the nonnegative variables that are greater than zero in the current iteration, while J_0 contains the indices of unconstrained variables. The active set is denoted by $\bar{J} := \{1, \ldots, n\} \backslash (J \cup J_0)$. For any index set $J = \{j_1, \ldots, j_{\bar{n}}\} \subseteq \{1, \ldots, n\}$ with $j_1 < \ldots < j_{\bar{n}}$, let $\boldsymbol{R}_J := (r_{i,j})_{i,j=1}^{\bar{n},n}$ denote the restriction matrix of $\{1, \ldots n\}$ to J, i.e., $r_{i,j_i} = 1$ and $r_{i,j} = 0$ otherwise. Then the restriction of $\boldsymbol{x} \in \mathbb{R}^n$ to its components with indices in J and the reverse zero fill in process can be written as $\boldsymbol{x}_J = \boldsymbol{R}_J \boldsymbol{x}$ and $\boldsymbol{x} = \boldsymbol{R}_J^{\mathrm{T}} \boldsymbol{x}_J$, respectively. Furthermore $\boldsymbol{A}_J = \boldsymbol{A}\boldsymbol{R}_J^{\mathrm{T}}$ denotes the matrix which contains only the rows of $\boldsymbol{A}$ with indices in J. The numerical procedure can be found in Algorithm 1. Note that we do not compute the inverse of matrices, but solve the linear systems in Steps 5, 8, 17 and 26 of the algorithm by Cholesky decomposition.

Algorithm 1 Partly NNLS

1: **Input:** A, b, J_0
2: **Output:** x
3: **Initialization:** $J, s = 0$
4:
5: $s_{J \cup J_0} \leftarrow (A_{J \cup J_0}^T A_{J \cup J_0})^{-1} A_{J \cup J_0}^T b$
6: **if** $\min s_J < 0$ **then**
7:      $J \leftarrow \emptyset$
8:      $s_{J_0} \leftarrow (A_{J_0}^T A_{J_0})^{-1} A_{J_0}^T b$
9: **else if** $\min s_J = 0$ and $J \neq \{1, \ldots, n\}$ **then**
10:      $J \leftarrow J \backslash \{i \in J : (R_J^T s_J)_i = 0\}$
11: **end if**
12: $x \leftarrow R_{J \cup J_0}^T s_{J \cup J_0}$
13: $p_{\bar{J}} \leftarrow A_{\bar{J}}^T (b - A_{J \cup J_0} x_{J \cup J_0})$
14: **while** $J \cup J_0 \neq \{1, \ldots, n\}$ and $\max p_{\bar{J}} > 0$ **do**
15:      $j \leftarrow \text{argmax}_{i \in \bar{J}} (R_{\bar{J}}^T p_{\bar{J}})_i$
16:      $J \leftarrow J \cup \{j\}$
17:      $s_{J \cup J_0} \leftarrow (A_{J \cup J_0}^T A_{J \cup J_0})^{-1} A_{J \cup J_0}^T b$
18:      $s \leftarrow R_{J \cup J_0}^T s_{J \cup J_0}$
19:      **while** $\min s_J \leq 0$ **do**
20:          **if** $\min s_J = 0$ **then**
21:              $J \leftarrow J \backslash \{i \in J : (R_J^T s_J)_i = 0\}$
22:          **else**
23:              $\alpha \leftarrow \min_{\substack{i \in J \\ s_i < 0}} \frac{x_i}{x_i - s_i}$
24:              $x \leftarrow x + \alpha(s - x)$
25:              $J \leftarrow J \backslash \{i \in J : x_i = 0\}$
26:              $s_{J \cup J_0} \leftarrow (A_{J \cup J_0}^T A_{J \cup J_0})^{-1} A_{J \cup J_0}^T b$
27:          **end if**
28:          $s \leftarrow R_{J \cup J_0}^T s_{J \cup J_0}$
29:      **end while**
30:      $x \leftarrow s$
31:      $p_{\bar{J}} \leftarrow A_{\bar{J}}^T (b - A_{J \cup J_0} x_{J \cup J_0})$
32: **end while**

---

4. Data processing

Before we discuss the results of our approach, we comment on the data acquisition and the preprocessing. In order to extract meaningful and consistent diffuse spectral reflectances we perform a flat-field correction, a data scaling and a white balance to take the experimental setup into account.

4.1. Data acquisition

First we comment on the acquisition of the calibrated diffraction patterns $\hat{P}^l(i, j)$, $(i, j) \in \mathcal{G}$, $l = 1, \ldots, L$ in our model Eq. (1). These are obtained from dark-current-corrected data $\hat{P}_0^l(i, j)$ measured once for a chip in a laboratory setting. Note that we can extract the data for all grid positions from a single full sensor image per wavelength. The experimental circumstances like the sensor quantum efficiency and the illuminant spectrum influence the relative scaling of the diffraction patterns for different wavelengths, i.e., the ratio of their means $m_{\hat{P}_0^l} = \frac{1}{n_1 n_2 s^2} \sum_{i, j \in \mathcal{G}} \|\hat{P}_0^l(i, j)\|_1$, where $\| \cdot \|_1$ corresponds to the sum over all elements as they

are nonnegative. In order to unify the relative scaling of the calibration diffraction patterns, we set their mean to (a multiple of) the sensor efficiency and call the result $\hat{P}^l(i, j)$.

4.2. Flat-field correction

In common digital images some areas may appear darker than they actually are due to vignetting and optical path distortions. In the case of direct measurements this can be countered by dividing the measured image by a flat-field image which is an image of a scene with a uniform illumination. In our spectral imaging system, there are also two main factors contributing to undesired intensity differences in the spectral cube. For one, the illumination during the calibration process is not uniform due to the monochromator intensity distribution and a non-ideal diffusion disc. Second, we usually want to reconstruct the reflectance which is independent of the illumination of a scene while our measured data F is not. Consequently, we need to change the relative spatial scaling to ensure both an interpretability of the results as reflectance values and a spatially independent regularization. In order to include this in our model, we rescale the measured diffraction pattern as

$$F(i, j) \leftarrow \frac{\sum_{l=1}^{L} \|P^l(i, j)\|_1}{\|F_{\mathrm{fl}}(i, j)\|_1} F(i, j), \tag{9}$$

where $F_{\mathrm{fl}}(i, j)$ is the dark current corrected flat-field image diffraction pattern at position (i, j). This step is performed before the absolute scaling.

4.3. Absolute scaling

The needed value of the regularization parameter α depends on the absolute scaling of the data. We divide both P and F by their respective means $m_P := \frac{1}{L n_1 n_2 S} \sum_{l=1}^{L} \sum_{i,j \in \mathcal{G}} \|P^l(i, j)\|_1$ and $m_F := \frac{1}{n_1 n_2 S} \sum_{i,j \in \mathcal{G}} \|F(i, j)\|_1$ before finding a solution to Eq. (4). After the optimization we multiply the result by $\frac{m_f}{m_P}$ in order to make it consistent with the original data.

4.4. White balance

In order to extract the spectral reflectance, we perform a white balance similar to the one done for, e.g., line scanning cameras, see [16]. But in our case the spectral abundances are not directly given by the camera and instead have to be computed from the diffraction patterns via the given variational model. Thus we also have to incorporate the white balance in our variational model. Given the diffraction patterns $(\hat{P}^l)_{l=1}^{L}$ measured during our calibration process in the laboratory, our task is to find factors $(d^l)_{l=1}^{L}$ such that the scaled calibration data $(P^l)_{l=1}^{L} = (d^l \hat{P}^l)_{l=1}^{L}$ lead to meaningful results in our model for data recorded under given circumstances. We compute these factors as follows: Let $F_w \in \mathbb{R}^{s,s}$ be the diffraction pattern image corresponding to a "white" scene and let $(h_w^l)_{l=1}^{L}$ be the known spectral reflectance of this scene, usually assumed to be constant. Let

$$x_w := \arg \min_{x} \mathcal{D}(x; F_w, \hat{P}^1, \dots, \hat{P}^L) + \alpha \mathcal{R}(x) \tag{10}$$

be the result of our model for this scene. Then we can calculate $d^l = \frac{x_w^l}{h_w^l}$, $l = 1, \dots, L$, and use the calibration patterns $(P^l)_{l=1}^{L} = (d^l \hat{P}^l)_{l=1}^{L}$ in our model Eq. (2). In practice we average the reconstructed x_w over a high number of white pixels and calculate robust global factors d^l, i.e., the calibration patterns are scaled the same way for every unit cell.

Note that these resulting factors depend on the chosen regularizer and the regularization parameter. The intermediate result x_w is approximately the product of h_w and the lamp spectrum. In order to keep the influence of the regularizer on the result small, both of these spectra should be smooth. In experiments, we can choose the white scene, so that h_w can be assumed to be smooth in practice. The lamp spectrum sometimes cannot be chosen freely. If it is known to be

non-smooth, it is possible to use a matrix in the regularizer $\mathcal{R}$ which does not rely as much on smooth results, such as the identity $\boldsymbol{\Gamma} = \boldsymbol{I}$, for the white balance.

5. Results

In this section, we present results of our approach for both simulated and experimental data.

5.1. Simulated data

In order to verify our mathematical model and analyze its dependence on the regularization parameter α, we start with simulated data. To construct the matrix $\boldsymbol{P}$ for Eq. (7), we first simulated the diffraction patterns created by light of different wavelengths between 400nm and 700nm in 10nm steps by the angular spectrum method [17] for the given DOE in Fig. 2(a). Thus, $L = 31$. These diffraction patterns have a high spatial resolution and are downsampled to the size of $s \times s = 42 \times 42$ pixels to get $\boldsymbol{P}$. The same size of the diffraction patterns is also used for experimental data in the next subsection. The means of the columns of $\boldsymbol{P}$ are scaled according to the quantum efficiency of a CMOS sensor. Further, we assume that both $\boldsymbol{P}$ and the data $\boldsymbol{F}$ were recorded using an illuminant with the same spectrum. This implies that we do not need to perform a white balance. Note that in contrast to the experimental data we use the same matrix $\boldsymbol{P}$ in every pixel of the artificial image and do not take into account that it depends on the grid position.

As the ground truth spectral image we use eight $20 \times 20 \times 31$ blocks of reflectance spectra $\boldsymbol{h} \in \mathbb{R}^{31}$ from the GretagMacBeth ColorChecker chart [18], where each block only contains one of the eight spectra. An RGB projection of the ground truth is shown in Fig. 4(a). The reflectance spectra $\boldsymbol{h}$ of the colorchecker chart are publicly available. We set pixelwise

$$\boldsymbol{F} = \sum_{l=1}^{L} \boldsymbol{P}^l h^l + \eta, \tag{11}$$

where η is a realization of white Gaussian noise of standard deviation $\sigma = 2$ (signal-to-noise ratio (SNR) 37.8 dB). Note again that $\boldsymbol{P}^l$ is both independent of the spectrum of the square and its position within the square; h^l depends only on the spectrum of the square, but not on its position within the square; and $\boldsymbol{F}$ depends both on the spectrum of the square and its position due to the different noise realizations.
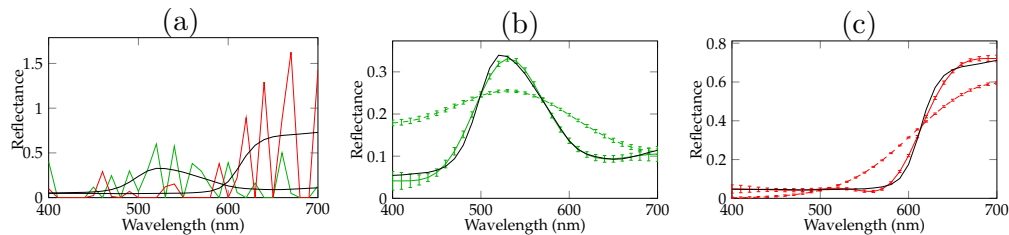

Fig. 3. Reconstructed spectra from noisy simulated data based on reflectance spectra of the GretagMacBeth ColorChecker chart. (a) Single reconstructed reflectance spectra without regularization ($\alpha = 0$) of green and red squares of the color checker along ground truth (black). (b) Mean and standard deviation of reconstructed green spectrum averaged over 400 reconstructions for $\alpha = 10^4$ (solid line) and $\alpha = 10^7$ (dashed line). (c) Same as (b) for the red spectrum.

We reconstruct the spectrum by Eq. (4) for different regularization parameter α. Figure 3(a)-(c) shows the reconstructions of the green and red spectra for different regularization

parameters. The reconstruction without regularization ($\alpha = 0$) is sparse instead of smooth and is therefore not a useful reconstruction. With an adequate regularization parameter ($\alpha = 10^4$) the reconstructed spectra are very close to the ground truth, while it is oversmoothed for the larger regularization parameter ($\alpha = 10^7$). The increased standard deviation towards the boundaries of the reconstructions are a result of a combination of reduced DOE dispersion and lower sensor efficiency.

Figure 4(b) contains the mean-square errors (MSE) defined by

$$\frac{1}{n_1 n_2 L} \sum_{i,j \in \mathcal{G}} \|\boldsymbol{h}(i,j) - \boldsymbol{h}_{\text{rec}}(i,j)\|_2^2, \tag{12}$$

with the ground truth spectrum $\boldsymbol{h}(i,j) \in \mathbb{R}^L$ in pixel (i,j) and the reconstructed spectrum $\boldsymbol{h}_{\text{rec}}(i,j)$ for different matrices $\boldsymbol{\Gamma} \in \{\boldsymbol{D}_1, \boldsymbol{D}_2, \boldsymbol{I}\}$ and different regularization parameters α.
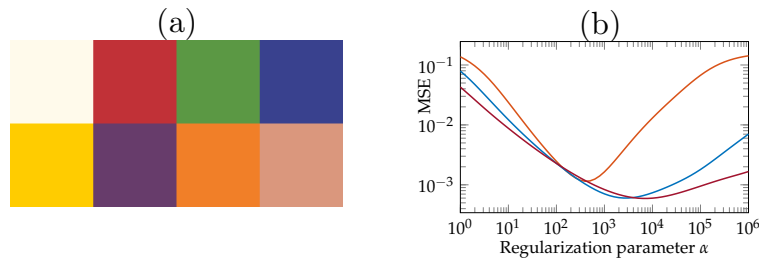

Fig. 4. (a) RGB projection of synthetic spectral reflectance data cube used as ground truth. (b) MSE of reconstructed data cube depending on regularization parameter for $\boldsymbol{\Gamma} = \boldsymbol{I}$ (orange), $\boldsymbol{\Gamma} = \boldsymbol{D}_1$ (blue) and $\boldsymbol{\Gamma} = \boldsymbol{D}_2$ (dark red).

We see that with the two regularizers based on derivatives we can reach similar minimal errors. With second order differences we can obtain a slightly smaller minimal error and the choice of the regularization parameter is more robust. The choice $\boldsymbol{\Gamma} = \boldsymbol{I}$ used, e.g., in [8, 19] leads to worse results and is more sensitive to the choice of α near the minimum. Therefore we prefer second order differences in our experimental section. In all cases the reconstruction error first decreases with increasing regularization, since the influence of the noise is suppressed, but after a certain point the reconstruction error increases due to the over-smoothing. Note that the L-curve criterion [20], which is often helpful in finding a suitable reconstruction parameter for Tikhonov-type regularization does not give good results in our application as it is known to fail for smooth solutions [21] which we expect here. In all tests with experimental data we observed that our choice of $\alpha = 400$ yielded good results which is in accordance with the results from simulated data in Fig. 4(b) when accounting for the noise reduction by the averaging process in experiments. In general, a good parameter can be found by measuring data with known reflectance spectra such as the colorchecker and adjusting the parameter until the reconstructed spectra have the desired smoothness.

5.2. Experimental data

Now we verify our approach by experimental data. In all experiments we used $\alpha = 400$ in the regularization model.

We start with measured diffraction patterns of certain squares of the Gretag MacBeth color checker [18]. We calibrated our system with a high pressure xenon arc lamp (Horiba FL-1039) and a monochromator (Horiba MicroHR), creating $L = 31$ images $\boldsymbol{P}^l \in \mathbb{R}^{42 \times 42}$ for the wavelengths from 400 nm to 700 nm in 10 nm steps. Throughout the experiment, we used an f-number of 8 and a focal length of 0.3 m which was also the object distance. The used Aptina MT9J001

CMOS sensor has a size of 3840×2748 pixels with a pixel pitch of 1.67 μm. This means that each diffraction image $F(i, j)$ has a size of about 70 μm $\times$ 70 μm and the reconstructed (diffuse) spectral reflectance cube has a size of $n_1 \times n_2 \times L = 89 \times 63 \times 31$. Note that the chosen sensor is not optimal in terms of sensitivity, but the small pixel pitch completely eliminates the risk of undersampling the data for the presented investigation of the basic sensor principle. For real applications with the currently used DOE structure a pixel pitch of around 2.5 μm is expected to work well, opening other sensor options with substantially larger sensitivity. For the white balance a PTFE sheet was used to get F_w. For the examples in Fig. 5 and Fig. 6 we employed a Carl Zeiss HAL 100 halogen lamp with heat filter as a light source and we placed a diffusion disc directly in front of the lamp to get a more homogeneous illumination. We averaged 16 images with an exposure time between 750 ms for the PTFE sheet and 1900 ms for parts of the color checker.

The reconstructed reflectances in the blue, green, yellow and red square along with the ground truth can be found in Fig. 5. The mean squared error to the ground truth in these examples is between $1 \cdot 10^{-4}$ and $5 \cdot 10^{-4}$. These are slightly smaller errors than the minimum for simulated data in Fig. 4 as the noise is lower in the experimental data due to averaging. It can be seen that the reconstructed spectra are very close to the ground truth reflectance between 450 nm and 650 nm which speaks for an accurate white balance in this wavelength band and a suitable model. Below 450 nm the white balance is off due to a low signal to noise ratio caused mostly by the reduced spectral power of the illuminant, as its spectral intensity at 400 nm is only about 16% of the one at 550 nm. The lower reconstruction quality above 650 nm can mostly be attributed to the decreased dispersion of the DOE, which is indicated in the correlation matrix in Fig. 2(c).
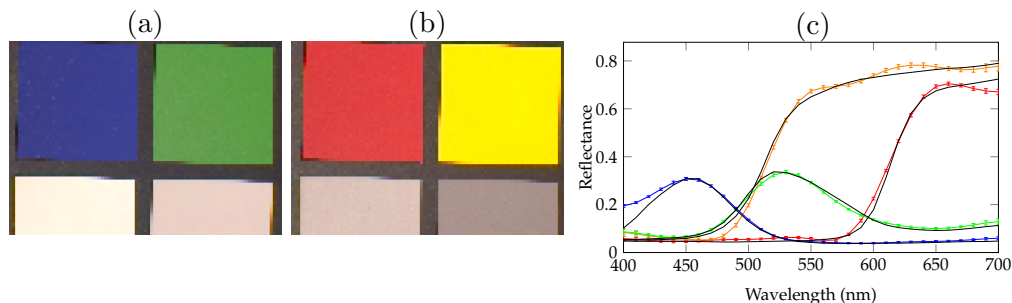

Fig. 5. (a),(b) RGB projections of reconstructed reflectance spectra of the GretagMacBeth ColorChecker chart, (c) Mean and standard deviation of reconstructed reflectances of blue, green, yellow and red square of the color checker along with the ground truth in black.

Next, we show two examples of more natural scenes. Figure 6 shows an RGB projection and several slices of the reconstructed reflectance spectral data cube of a tomato with a sticker on a small pedestal. The raw data was recorded under the same circumstances as the colorchecker in the previous example with an exposure time of 3000 ms. Note that we do not display all slices of the reconstructed data cube, but only every fifth for the sake of clarity. The slices display the behavior expected for this object, except for several oversaturated pixels in the center of the image. For these exceptional pixels reflections from the light source cause nonsensical results. The tomato is only visible in the wavelength bands starting at 600 nm. The green part of the sticker is most pronounced in the slice belonging to 550 nm and visibly present in the slices for 500 nm and 600 nm, while the blue part of the sticker is strongest at 450 nm and vanishes at 550 nm and above. Note that that 400 nm slice is noisier than all others due to the low spectral power of the illuminant.
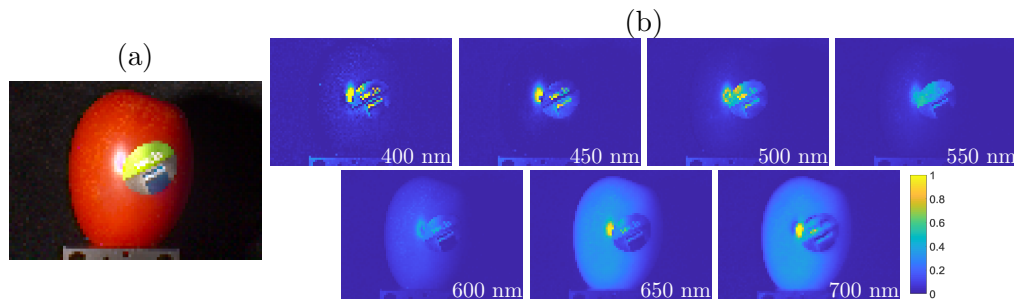
Fig. 6. Reconstructed reflectance spectral image of a tomato with a sticker. (a) RGB projection of the reconstructed reflectance spectral cube, (b) slices of the spectral data cube corresponding to different wavelength bands. The images are false color encoded to make the dynamic full range recognizable.

The example in Fig. 7 shows a fruit ensemble, which was recorded outside the laboratory in the sunlight with an exposure time of 700 ms and a focal length of 0.5 m, which was also the object distance. We performed a white balance with a white tile to take the spectrum of the sunlight into account. All other reconstruction parameters and the hardware are the same as in the examples from the laboratory. Again, there are few oversaturated pixels caused by reflected sunlight in the tomatoes and the pear. In general, the reconstructed data cube slices fit the expected spectra of the fruit. The pear if visible in all slices, but most prominent at 550 nm and above, which is expected for its light yellow color. Similarly to the tomato example from the laboratory in Fig. 6, the tomato can only be observed starting at 600 nm. In the 500 nm slice the apricots are much more pronounced than the tomatoes, but less than the pear, which fits its orange color.
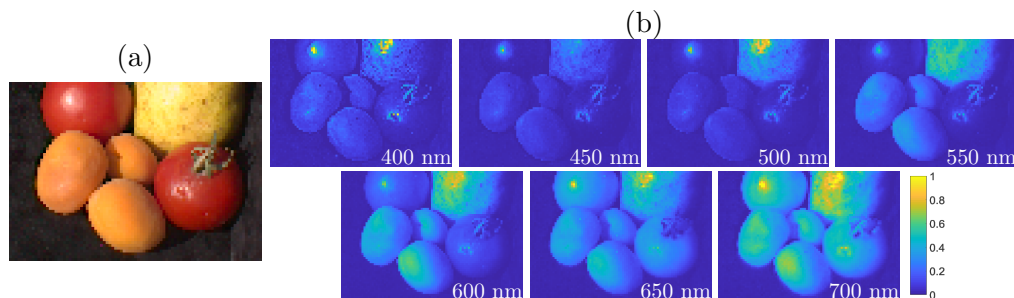

Fig. 7. Reconstructed reflectance spectral image of a fruit ensemble in the sunlight. (a) RGB projection of the reconstructed reflectance spectral cube, (b) slices of the spectral data cube corresponding to different wavelength bands. Again, the images are false color encoded to make the dynamic full range recognizable.

6. Discussion

Our computational approach enables us to reconstruct spectra from 2D diffraction patterns pixelwise. Although not necessary for our experiments so far, we like to mention that spatial relations could be incorporated into the variational model via additional regularizers. However, this results in higher computational costs and is beyond the scope of this paper. Our model is based on the assumption that the diffraction pattern created by light of an arbitrary spectrum diffracted by the DOE is the linear superposition of given calibration patterns. Since undersampling the spectral range leads to severe spectral artifacts due to the partly complex shape of the diffraction

patterns, we oversample the spectral range. Particularly, we pick the spectral step size $|\lambda_{l+1} - \lambda_l|$ smaller than the potential spectral resolution of the system. Initially, this leads to a bad condition number of the matrix P as its columns are almost linearly dependent. To cope with this, our model includes Tikhonov regularization based on derivatives to make the model robust to noise despite the bad condition number and to incorporate prior knowledge about the expected spectra in our model. The evaluation of simulated data demonstrates that our choice of regularizer is better suited for smooth spectra than the one used in similar approaches. These simulations also illustrate that the choice of the regularization parameter is not crucial, i.e., even if we do not choose the optimal parameter, the reconstruction quality does not suffer severely. In the data preprocessing we include a white balance procedure and a flat-field correction to ensure the consistent reconstruction of reflectance spectra under varying lighting conditions. We demonstrated the effectiveness of our approach with experimental data where we were able to reconstruct spectra close to the ground truth.

The partly NNLS algorithm presented to solve the pixel-wise inverse problem resulting from the model is numerically stable and flexible. It utilizes only matrix vector multiplications and the solution of small linear systems of a size of at most $L \times L$ which can be solved efficiently by, e.g., Cholesky decomposition. It can be adapted to arbitrary regularization matrices Γ and allows us to use the similarity of neighboring pixels to speed up the reconstruction process.

As explained in [22] there is no consistent notion of a spectral resolution for computational spectral imagers like ours as their performance depends the signal itself and, e.g., the chosen regularization parameter. Due to this we presented results from experimental data which demonstrate the capability of the setup and the reconstruction algorithm.

7. Conclusion

We demonstrated an algorithm for the pixelwise reconstruction of spectra from 2D diffraction patterns produced by a snapshot spectral imager. Compared to other imager concepts, the setup is compact, robust and easy to manufacture, but as a consequence the numerical reconstruction of the spectra becomes more sophisticated. We developed a suitable mathematical model for the produced data and implemented an efficient optimization algorithm which allows for pixelwise reconstruction of spectra. The incorporated regularization makes our approach robust to noise and we can integrate prior information about the expected spectra. We demonstrated the capability of our system with both simulated and experimental data.

References

1. J. Qin, K. Chao, M. S. Kim, R. Lu, and T. F. Burks, "Hyperspectral and multispectral imaging for evaluating food safety and quality," J. Food Eng. **118**, 157–171 (2013).
2. O. Carrasco, R. B. Gomez, A. Chainani, and W. E. Roper, "Hyperspectral imaging applied to medical diagnoses and food safety," in *Proc. SPIE 5097*, 215–222 (2003).
3. E. Boegh, H. Soegaard, N. Broge, C. Hasager, N. Jensen, K. Schelde, and A. Thomsen, "Airborne multispectral data for quantifying leaf area index, nitrogen concentration, and photosynthetic efficiency in agriculture," Remote. sensing Environ. **81**, 179–193 (2002).

4. A. Lambrechts, P. Gonzalez, B. Geelen, P. Soussan, K. Tack, and M. Jayapala, "A CMOS-compatible, integrated approach to hyper- and multispectral imaging," in *2014 IEEE International Electron Devices Meeting,* (2014), pp. 10.5.1–10.5.4.

5. A. Bodkin, A. Sheinis, A. Norton, J. Daly, S. Beaven, and J. Weinheimer, "Snapshot hyperspectral imaging: the hyperpixel array camera," in *Proc. SPIE 7334*, 73340H (2009).

6. A. A. Wagadarikar, N. P. Pitsianis, X. Sun, and D. J. Brady, "Spectral image estimation for coded aperture snapshot spectral imagers," in *Proc. SPIE 7076*, 707602 (2008).

7. S. K. Sahoo, D. Tang, and C. Dang, "Single-shot multispectral imaging with a monochromatic camera," Optica **4**, 1209–1213 (2017).

8. P. Wang and R. Menon, "Ultra-high-sensitivity color imaging via a transparent diffractive-filter array and computational optics," Optica **2**, 933–939 (2015).

9. P. Wang and R. Menon, "Computational multispectral video imaging [invited]," J. Opt. Soc. Am. A **35**, 189–199 (2018).

10. A. Wagadarikar, R. John, R. Willett, and D. Brady, "Single disperser design for coded aperture snapshot spectral imaging," Appl. optics **47**, B44–B51 (2008).

11. P. C. Hansen, *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion* (SIAM, 1998).

12. Å. Björck, *Numerical Methods for Least Squares Problems* (SIAM, 1996).

13. D. Chen and R. J. Plemmons, "Nonnegativity constraints in numerical analysis," in *The Birth of Numerical Analysis,* A. Bultheel and R. Cools, eds. (World Scientific, 2010), pp. 109–139.

14. C. L. Lawson and R. J. Hanson, *Solving Least Squares Problems* (SIAM, 1995).

15. R. Bro and S. De Jong, "A fast non-negativity-constrained least squares algorithm," J. Chemom. **11**, 393–401 (1997).

16. H. M. Stokman and T. Gevers, "Detection and classification of hyper-spectral edges," in *BMVC,* (1999), pp. 1–9.

17. J. W. Goodman, *Introduction to Fourier optics* (McGraw-Hill, 1968). Ch. 3.

18. D. Pascale, "RGB coordinates of the macbeth colorchecker," The BabelColor Co. pp. 1–16 (2006).

19. P. Wang and R. Menon, "Computational spectroscopy via singular-value decomposition and regularization," Opt. Express **22**, 21541–21550 (2014).

20. P. C. Hansen, "Analysis of discrete ill-posed problems by means of the L-curve," SIAM review **34**, 561–580 (1992).

21. P. C. Hansen, "The L-curve and its use in the numerical treatment of inverse problems," in *Computational Inverse Problems in Electrocardiology,* P. Johnston, ed. (WIT Press, 2000), pp. 109–139.

22. N. Hagen and M. W. Kudenov, "Review of snapshot spectral imaging technologies," Opt. Eng. **52**, 090901 (2013).