

Spectral image recovery from Spectral Filter Array cameras using LMMSE

Prakhar Amba, *Laboratoire de Psychologie et Neurocognition, CNRS UMR 5105, Univ. Grenoble Alpes, Grenoble, France*
email: prakhar.amba@univ-grenoble-alpes.fr

David Alleysson, *Laboratoire de Psychologie et Neurocognition, CNRS UMR 5105, Univ. Grenoble Alpes, Grenoble, France*

Abstract

A hyperspectral camera can record a cube of data with both spatial 2D and spectral 1D dimensions. Spectral Filter Arrays (SFAs) overlaid on a single sensor allows a snapshot version of a hyperspectral camera. But acquired image is subsampled both spatially and spectrally, and a recovery method should be applied. In this paper we present a linear model of spectral and spatial recovery based on Linear Minimum Mean Square Error (LMMSE) approach. The method learns a stable linear solution for which redundancy is controlled using spatial neighborhood. We evaluate results in simulation using gaussian shaped filter's sensitivities on SFA mosaics of upto 9 filters with sensitivities both in visible and Near-Infrared (NIR) wavelength. We show by experiment that by using big neighborhood sizes in our model we can accurately recover the spectra from the RAW images taken by such a camera. We also present results on recovered spectra of Macbeth color chart from a Bayer SFA having 3 filters.

Introduction

A light source or its reflection on a surface is an emitter of electromagnetic radiation which can be measured as a spectral function. Devices like a spectroradiometer employ a diffraction grating which disperses the spectrum linearly, which is then measured by a sensor, like CCD array. This is however a point measurement. Using a scanning technique (spatial or spectral), it is possible to record a cube of spectra, with two spatial dimensions and third in wavelength. Such a system is called a hyperspectral camera. Hyperspectral imaging, measures a spectral signature of scene and has applications in different fields like agriculture, food fraud detection, machine vision, autonomous vehicles, remote sensing, forensics, etc. A scanning solution (linear scan or filter wheel) might not be ideal solution for cases where time is a constraint or a portable solution is preferred. Hagen et al. [12] reviewed several snapshot spectral imaging system. Camera systems based on Spectral Filter Arrays (SFAs) (referred as Spectrally Resolving Detector Arrays by Hagen et al.) have recently made commercial appearance. IMEC [2] and Silios [3] have both launched cameras with such SFAs. Silios offers cameras with 9 or 16 narrow band filter mosaics, IMEC has up to 32 different filters. SFAs subsample the scene photographed both spatially (by nature of mosaic) and spectrally (by filters).

SFAs are an advancement of the Color Filter Arrays (CFAs) employed in digital cameras to capture Red, Green and Blue color by spatial sub-sampling. One need to inverse the spatial sub-sampling to recover the full resolution color image and this is known as demosaicing and has been extensively studied in literature [4, 6, 14]. Demosaicing algorithms makes use of the fact that RGB color channels are heavily correlated as there is overlap in the spectral sensitivities of the filters. However, they are limited to only spatial recovery of tri-stimuli values and not spectral.

Previously, this problem has been studied in the context of

estimating hyperspectral images from a single RGB image as input using sparse dictionary [8]. The same was also subject of competition NTIRE 2018 [9] where most of solutions used Deep Neural Networks. In this case, the authors recover spectra from already demosaiced images. Demosaicing is known to create false colors and artifacts in areas of high frequency content, which would be further propagated in the spectral recovery. Therefore, a real application should start with the RAW images. Also for SFAs, spectral correlations between the filters are not as straightforward as they are designed to be narrow-band and there are certain manufacturing limitations.

In this paper we will study the problem of inverting the sub-sampling induced by SFA (both spatial and spectral) and recovering the full resolution cube, i.e. hyperspectral image (spatial dimensions of size of sensor and across a range of wavelengths) from the RAW image captured by the camera. We do not limit this work to the visible domain and present a general framework for both visible and Near-Infrared domain. Our algorithm is based on finding a linear solution to the inverse problem by learning a mapping between the cross correlation of the RAW image and its corresponding spectral image taken on a test hyperspectral image database. We use the least square estimate to find the inverse operator which minimizes the mean square error between the original and the estimated spectra. We further use a neighborhood window to stabilize our solution. We describe our model mathematically in the next section. We prefer a linear solution due to its simplicity and computational efficiency as we are processing megapixels of data over a large range of wavelength points.

LMMSE Model for Spectral Recovery

Figure 1, shows the basic principle of the Linear Minimum Mean Square Error (LMMSE) approach to recover hyperspectral image from the SFA RAW image. Let Z be the hyperspectral image in form of reflectance data cube of size $H \times W \times P$. Meaning we have $H \times W$ spatial data for P different wavelengths. This reflectance data rendered by an illuminant gives us the radiance cube which can be captured by a SFA camera. This results in image of size $H \times W$, containing pixels with several spectral sensitivities according to the arrangement of filters placed on the SFA. We can simulate the RAW image X , captured by this camera by applying the spectral sensitivities of the filters F on the radiance cube and then spatially sampling it according the mosaic arrangement M . This is the forward image formation model. The reverse model is to estimate reflectance cube \hat{Z} from this X , such that the MSE between Z and \hat{Z} is minimized [18]. A direct simple pseudo-inverse of MFL is not good as the problem is under-determined because one is trying to estimate P values for a single spatial pixel. To overcome this problem it has been proposed to learn the inverse model indirectly [15] on several exemplar taken from images in a hyperspectral image database. It

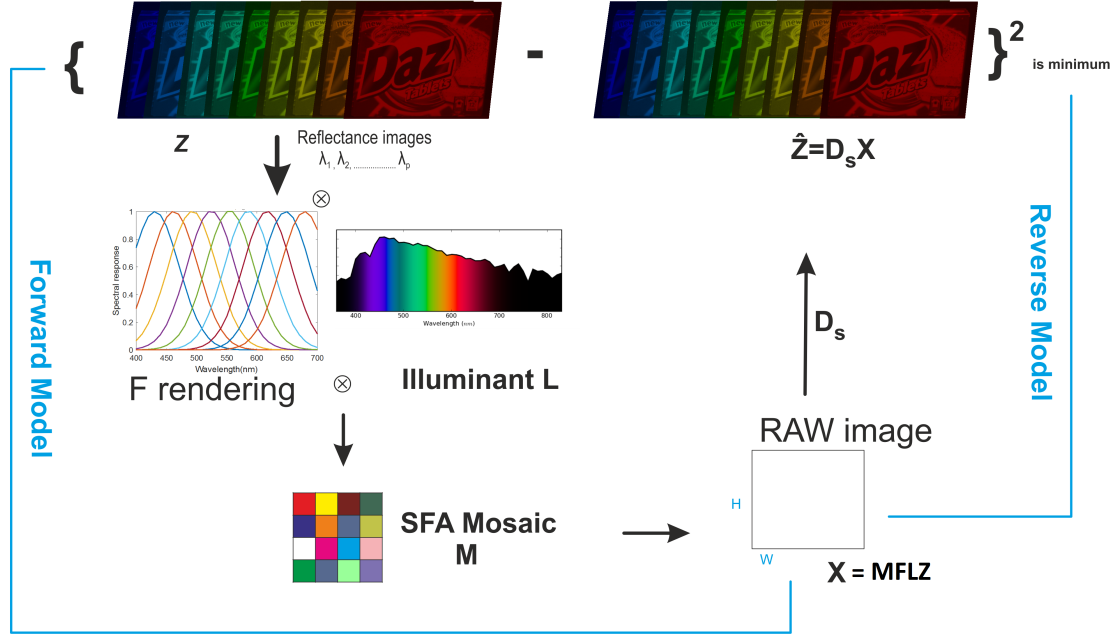


Figure 1. LMMSE model for spectral recovery. Z is the hyperspectral image data, reflectance. The image formation model is as follows. The reflectance image is rendered under an illuminant L to get the radiance image which is then rendered by spectral filters' sensitivity to get the multi-spectral images. This is further spatially subsampled by the Mosaic's M to obtain the RAW image X , such that $X = MFLZ$. \hat{Z} is the estimate of this data obtained from the RAW image by applying the D_s operator. The LMMSE model wants to design this D_s , such that the MSE between Z and \hat{Z} is minimized. Here the reflectance image shown is from the Finlayson hyperspectral database [11]. Notations in the figure do not account for neighborhood nor the figure show the NIR component.

has also been shown that for estimating a robust inverse model, a neighborhood should be taken into the SFA RAW image [7].

Because most SFA are made on a replicated sub-pattern, called super-pixel, tiled over the surface of the sensor, we could reduced the dimension of the model to the size of a single super-pixel [17]. To account for the neighborhood subscript 1 is attached to the variable. See reference [5] for a detail description for the image unfolding with neighborhood. The model could be expressed in vector-matrix form as follows:

$$x_1 = M_1 f_1 l_1 z_1 \quad (1)$$

$$\hat{z} = D_s x_1 \quad (2)$$

$$D_s = (S_1 z_1 x_1^t) (x_1 x_1^t)^{-1} \quad (3)$$

$$D_s = (S_1 R_{z_1} l_1^t f_1^t M_1^t) (M_1 f_1 l_1 R_{z_1} l_1^t f_1^t M_1^t)^{-1} \quad (4)$$

$$R_{z_1} = \frac{z_1 z_1^t}{K} \text{ where } K \text{ is no. of exemplars learnt} \quad (5)$$

where Equation 1 states the forward model. z_1 is a matrix of column vectors where each vector corresponds to a super-pixel. Each vector in z_1 contains the unfolded reflectance for the area of the super-pixel plus the neighborhood. l_1 is a diagonal matrix containing the illuminant duplicated for the neighborhood size. f_1 is rectangular matrix having number of rows equal to the number of filters and number of column corresponding to P times the size of the neighborhood. M_1 is a subsampling matrix that selects which filter belong to which pixel, compensated for the neighborhood. So x_1 correspond to the unfolded matrix of the RAW image including the neighborhood.

Equation 2 is the reverse model that compute an estimate of z called \hat{z} . D_s is the spectral recovery operator which is defined in Equation 3, where S_1 is an operator that reduce the neighborhood on z_1 , ($z = S_1 z_1$).

Another way to express D_s is given in Equation 4 where R_{z_1} defined in Equation 5 is the correlation of reflectance over the

entire database accounting for neighborhood.

The model above shows that the R_{z_1} correlation matrix has to be learnt only once on any given hyperspectral image database. Only M_1 and S_1 should be specified for a particular SFA mosaic with particular spatial arrangement and filter sensitivities. To evaluate the performance of the spectral recovery we define a PSNR in reflectance for all the P channels, where H and W are spatial dimensions, μ_z as follows:

$$MSE = \frac{\sum \sum (\hat{z} - z)^2}{HWP}$$

$$\mu_z = 10 \log_{10} \frac{1}{MSE}$$

In the next section, we report values of μ_z , averaged for all images in that database.

Experiment

Typically SFA filters are based on Faber-Perot interferometers [13]. This provides great flexibility in design of spectral filters. Lapray et al. [13] demonstrated that these filters can be approximated by use of a Gaussian model. According to gaussian model the filter can be defined in terms of three parameters λ_μ , the wavelength of peak sensitivity, the standard deviation λ_σ , the spread of the filter and finally the amplitude A_m , the intensity factor.

$$F(\lambda) = \frac{A_m}{\lambda_\sigma \sqrt{2\pi}} e^{-\frac{(\lambda - \lambda_\mu)^2}{2(\lambda_\sigma)^2}}$$

Visible and NIR filters

Let us consider SFA having gaussian shaped spectral filters spanning both visible and NIR wavelength. For purpose of testing we consider the SCIEN database [16] which contains hyperspectral images from 414.72nm to 950.49nm. R_{z_1} was trained

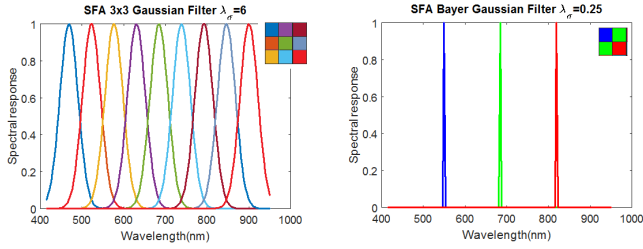


Figure 2. Simulated Gaussian Filters on the Visible+NIR wavelength

on 4 landscape images without the polarizing filter. The images provided are in radiance and not reflectance, therefore we attempt to recover radiance directly. The procedure we follow is as described earlier, we simply replace LZ by the radiance values provided in the images. This can be done by considering L to be 1 and Z to be radiance (instead of reflectance). We keep this consideration only for this sub-section.

We simulated gaussian shaped filter across wavelength range of 414.7 to 950.5nm. Figure 2, shows the simulated filters for a λ_σ of 6 for 3x3 SFA and λ_σ of 0.25 for a Bayer SFA. λ_μ were chosen to ensure the filters were equidistant. Same amplitude A_m for chosen for all filters, regularized to ensure output images are not saturated.

Figure 3 shows reconstructed spectra for a particular pixel of the Stanford Tower image for two different neighborhood sizes for a SFA of 3x3 size having 9 filters (Figure 2). It can be seen that a bigger neighborhood size allows a better recovery of the signal. In terms of average μ_z , its is 36.15 dB for neighborhood 7, compared to 33.50 dB for neighborhood 1.

Figure 4 and 5 shows the recovered radiances for 4 different pixels points in the Stanford Tower image for SFAs of 3x3 ($\lambda_\sigma=6$) and Bayer [10] ($\lambda_\sigma=0.25$) type (Figure 2). We chose to present for very narrow band Bayer type to highlight a bad case. We see that for the 3x3 SFA, the recovered radiance follows the measured (from original test images) more closely than the Bayer SFA. In terms of μ_z , its is 36.15 dB for 3x3 SFA, compared to only 31.08dB for Bayer arrangement.

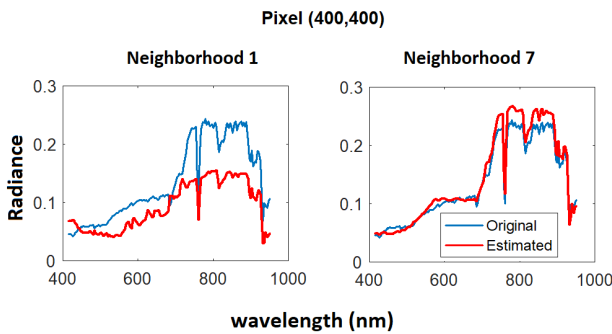


Figure 3. Effect of neighborhood window in aiding recovery of spectrum more accurately for Stanford Tower image from SCIEN for SFA 3x3 $\lambda_\sigma=6$.

For 3x3 SFAs, increasing λ_σ [0.25 to 10] first improves PSNR, μ_z and then it falls again. It suggests that very narrow-band filters are not the best option as the LMMSE algorithm exploits cross correlation between spectral channels. However very large λ_σ also degrades μ_z , as individual differences are reduced. Our algorithm takes only 1.2s to process Stanford Tower image, which has a resolution of 0.8 MP spatially and 148 wavelength bands.

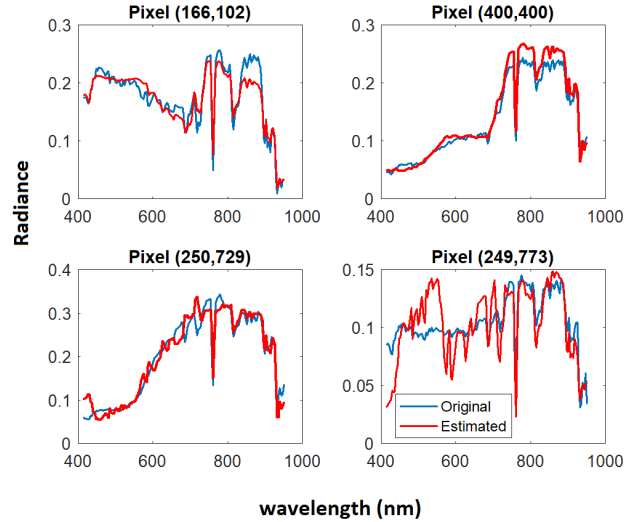


Figure 4. Recovered Radiance from 4 different pixel points on the Stanford Tower image from SCIEN for SFA 3x3 $\lambda_\sigma=6$, using a neighborhood of 7. $\mu_z=36.15$ dB

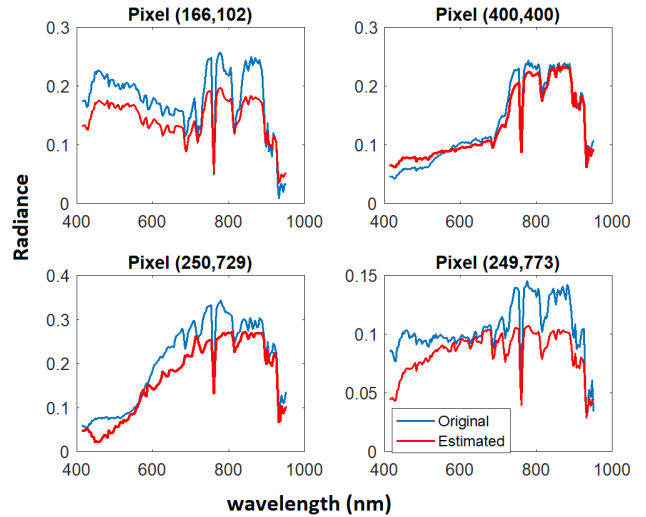


Figure 5. Recovered Radiance from 4 different pixel points on the StanfordDish image from SCIEN for SFA Bayer $\lambda_\sigma=0.25$, using a neighborhood of 7. $\mu_z=31.08$ dB. We choose to present this to purposefully highlight a bad case.

Visible filters

To test the spectral recovery into the visible domain, we consider the Finlayson hyperspectral database [11]. We considered SFAs having Bayer arrangement, 2x2 (3 filters). Spectral sensitivities were again gaussian in shape but in this case the wavelength range is restricted from 400nm to 700nm. We found that having more filters improved μ_z . Figure 6 shows the spectra of color patches from the Macbeth colorchecker chart [1], (a) as recorded in the image from the Finlayson database and (b) as recovered from the SFA RAW image considering a Bayer mosaic, having gaussian shaped equidistant filters of $\lambda_\sigma = 4$. These measurements are average of 100 pixels of each color patch. It can be observed that the measured spectra seems to follows the recorded. However it is not possible to qualify the accuracy as this would depend on the application. One could measure the MSE between the two spectra, but we feel the need for a metric which can reflect the quality of the measurement in terms of

labels like useful or not. Spectral recovery are for applications which require distinguishing two spectral signatures, for example like food fraud. If one has a sample where plastic rice is mixed with real rice grains and one needs to distinguish the two. Our spectral measurement should be different enough to allow detection of two. In this paper we remain application blind and therefore choose to present results as it is.

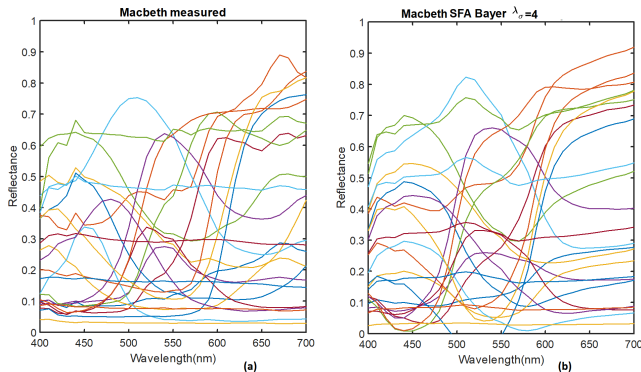


Figure 6. Spectra of the color patches from the Macbeth color checker chart. (a) as measured in the Finlayson database image, (b) as recovered from SFA RAW image shot with a Bayer arrangement.

Conclusion

In this paper we present a linear approach to recover hyperspectral image cube from SFA RAW images. We present results in simulation considering gaussian shape sensitivities of filters for both visible and visible+NIR wavelength range. We are able to recover the shape of the spectra quite reasonably if a large neighborhood window is use for learning the inverse model. We found that having more filters, improves the recovered spectra. For the evaluation purpose PSNR is used. PSNR is useful for comparing two different test conditions or methods. However we feel the need of a metric which can reflect whether the recovered spectra is meaningfully accurate. These results are obtained in simulation, a future work could validate the method on real RAW images shot with a camera.

References

[1] <https://xritephoto.com/colorchecker-classic>.
[2] Imec, <https://www.imec-int.com/en/hyperspectral-imaging>.
[3] Silios color shades, <https://www.silios.com/cms-cameras-1>.
[4] D. Alleysson, S. Susstrunk, and J. Héroult. Linear demosaicing inspired by the human visual system. *IEEE Transactions on Image Processing*, 14(4):439–449, 2005.
[5] P. Amba. *Learning methods for digital imaging*. PhD thesis, 2018.
[6] P. Amba, J. Dias, and D. Alleysson. Random color filter arrays are better than regular ones. *Journal of Imaging Science and Technology*, 60(5):50406–1, 2016.
[7] P. Amba, J. B. Thomas, and D. Alleysson. N-Immse demosaicing for spectral filter arrays. *Journal of Imaging Science and Technology*, 61(4):40407–1, 2017.
[8] B. Arad and O. Ben-Shahar. Sparse recovery of hyperspectral signal from natural rgb images. In *European Conference on Computer Vision*, pages 19–34. Springer, 2016.
[9] B. Arad, O. Ben-Shahar, and R. Timofte. Ntire 2018 challenge on spectral reconstruction from rgb images. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 929–938, 2018.
[10] B. Bayer. Color imaging array, July 20 1976. US Patent 3,971,065.
[11] G. D. Finlayson, S. D. Hordley, and P. Morovic. Using the spec-

tracube to build a multispectral image database. In *Conference on Colour in Graphics, Imaging, and Vision*, volume 2004, pages 268–274. Society for Imaging Science and Technology, 2004.
[12] N. A. Hagen and M. W. Kudenov. Review of snapshot spectral imaging technologies. *Optical Engineering*, 52(9):090901, 2013.
[13] P.-J. Lapray, J.-B. Thomas, P. Gouton, and Y. Ruichek. Energy balance in spectral filter array camera design. *Journal of the European Optical Society-Rapid Publications*, 13(1):1, 2017.
[14] X. Li, B. Gunturk, and L. Zhang. Image demosaicing: A systematic survey. In *Visual Communications and Image Processing 2008*, volume 6822, page 68221J. International Society for Optics and Photonics, 2008.
[15] A. Ribes and F. Schmitt. Linear inverse problems in imaging. *IEEE Signal Processing Magazine*, 25(4):84–99, 2008.
[16] T. Skauli and J. Farrell. A collection of hyperspectral images for imaging systems research. In *Digital Photography IX*, volume 8660, page 86600C. International Society for Optics and Photonics, 2013.
[17] D. Taubman. Generalized wiener reconstruction of images from colour sensor data using a scale invariant prior. In *Image Processing, 2000. Proceedings. 2000 International Conference on*, volume 3, pages 801–804. IEEE, 2000.
[18] H. J. Trussell and R. E. Hartwig. Mathematics for demosaicking. *IEEE Transactions on image processing*, 11(4):485–492, 2002.