

Optimization of Spectrally Coded Mask for Multi-modal Plenoptic Camera

Kathrin Berkner and Sapna Shroff

Ricoh Innovations, Inc., 2882 Sand Hill Rd, Suite 115, Menlo Park, CA 94025-7054
{berkner,sapna}@rii.ricoh.com

Abstract: We introduce a framework to optimize the spatial layout of a spectral filter mask inserted into the aperture of a plenoptic camera. The optimization merit function evaluates end-to-end system performance, including spectral crosstalk at the sensor caused by lens aberrations and diffraction effects. We use the example of a Bayer pattern to illustrate the effects of optimized filter partitions on the system performance.

© 2011 Optical Society of America

OCIS codes: 110.1758, 110.4234, 070.0070

1. Introduction

Plenoptic camera architectures are designed to capture different combinations of light rays from a scene. Whereas most of these architectures target applications such as digital refocusing, rotation, or depth estimation [1], only few address capturing spectral information of the lightfield. Authors in [2, 3] modify a plenoptic architecture with a filter array inserted in the pupil plane of the main lens, containing different spectral and polarization filters. Light rays originating from one point source enter the pupil plane at different locations, therefore, passing through different filters. The microlens array mounted close to the sensor images the pupil plane onto the sensor, producing an image of the spectral scene information as it is passing through the pupil plane. Such single-snapshot multispectral imaging system trades off spatial with spectral information captured with a single sensor. It does not require image registration of spectral bands and enables faster data acquisition in a video mode. The spectral mask, however, introduces additional diffraction effects. Chromatic aberration may introduce out of focus blur at the microlens and sensor. This could be overcome to some extent by using achromatic lenses. It is difficult, however, to manufacture achromatic microlens arrays. In an aberrated system it may be possible to design the filter mask in a way to compensate for the aberrations in order to achieve desired spectral performance of the imaging system.

In this paper we introduce a framework to optimize the spatial layout of a spectral filter mask in a plenoptic imaging system (Fig. 1(I)) with respect to a merit function that includes application-specific spectral performance of the end-to-end imaging system. We use the example of a Bayer pattern to illustrate the effects of optimized and non-optimized filter partitions on the system performance.

2. Optimization framework for spatial layout of a filter mask in a spectrally coded plenoptic camera

Binary aperture masks have been researched for use in traditional camera architectures to facilitate joint reconstruction of intensity and depth [1]. Here we introduce a spectral aperture mask and evaluate its performance in terms of spectral distribution and crosstalk of the captured sensor data. We define a *spectral aperture mask* as a partition \mathcal{P} of the aperture into a set of non-overlapping cells $c_i, i = 1, \dots, M$. Each cell has a spectral response function $\rho_i(\lambda)$ where λ is the wavelength. First we model light passing through the filter array considering only a single spectral response ρ^* . For this case the spectral mask reduces to a binary mask with aperture transmission function given by

$$t_{\rho^*}(u, v, \lambda) = \begin{cases} \rho^*(\lambda), & \text{if } \exists i \in \{1, \dots, M\} \text{ such that } (u, v) \in c_i \in \mathcal{P} \text{ with } \rho^* = \rho_i \\ 0, & \text{otherwise} \end{cases},$$

where u, v are the spatial coordinates in the aperture plane. In order to evaluate aberration and diffraction effects of the lens systems on spectral data captured at the sensor, we analyze the wavefront arriving at the sensor after passing through main lens, filter mask, and microlens using wave optics analysis and the operator notation introduced in [4]. For the scope of this paper, we assume an on-axis point source being imaged onto the

sensor plane with the aperture being located at a telecentric stop. For these assumptions, we use the model of a planar wave passing through a lens with focal length F_λ and aperture transmission function t_{ρ^*} described by $Q[-\frac{1}{F_\lambda}]\{t_{\rho^*}\}$, with $Q[c]\{U(x)\} = \exp^{j\frac{\kappa}{2}cx^2}U(x)$, where κ is the wave number. Free space propagation of a wavefront U over a distance z is given by $R[z]\{U(x_1)\} = \frac{1}{\sqrt{j\lambda z}} \int U(x_1) \exp^{j\frac{\kappa}{2z}(x_2-x_1)^2} dx_1$. For a plenoptic camera system, the propagation of a planar wavefront entering the pupil with aperture transmission function t_{ρ^*} and propagating through main lens and microlens of diameter d with clear aperture transmission function t_d onto the sensor is therefore given by $U_{sensor}\{t_{\rho^*}\} = R[z_2]Q[-\frac{1}{f_\lambda}]\{t_d\}R[z_1]Q[-\frac{1}{F_\lambda}]\{t_{\rho^*}\}$.

The spectral information recorded at the sensor plane is not simply a scaled image of the spectral code at the pupil plane. Lens aberrations, especially chromatic aberrations caused by wavelength-dependent magnification factors, cause the non-overlapping cells in the aperture mask to be imaged onto overlapping areas at the sensor plane. To evaluate this "spectral crosstalk" in the captured data, we first define the overlapping regions between cell images on the sensor inside a superpixel S under a microlens as $\Delta_1 = \bigcup_{(i,j), i \neq j} \{ \{(\eta, \xi) \in S \mid \int U_{sensor}\{t_{\rho_i}\}(\eta, \xi, \lambda) d\lambda \neq 0\} \cap \{(\eta, \xi) \in S \mid \int U_{sensor}\{t_{\rho_j}\}(\eta, \xi, \lambda) d\lambda \neq 0\} \}$. Evaluating the intensity value in area Δ_1 provides a measure for how much spectral crosstalk is present at the sensor. The light collected at sensor pixel location (η, ξ) , $J_{sensor}(\rho_m)(\eta, \xi)$, and the final information value corresponding to the filter ρ_m , $I(\rho_m)$, computed by integrating over pixels that are not effected by spectral crosstalk, are computed as

$$J_{sensor}(\rho_m)(\eta, \xi) = \int U_{sensor}\{t_{\rho_m}\}(\lambda, \eta, \xi) \tau(\lambda) d\lambda, \quad \text{and} \quad I(\rho_m) = \int_{S \setminus \Delta_1} J_{sensor}(\rho_m) d\mu, \quad (1)$$

where $\tau(\lambda)$ is the spectral sensitivity of the sensor and μ is an integration measure.

The cells of the aperture code can have different shapes, e.g. rectangles, hexagons, ring segments. In this paper, we assume a circular aperture of radius R and a radial filter partition into cells forming ring segments (e.g. annuli, disc sectors, or ring sectors). Each cell c_n has N rings with diameters δ_n with $n = 1, \dots, N$ and inner radii $r_{0,n} = \delta_n - r_{0,n-1}$ with $r_{0,0} = 0$ and $r_{0,N-1} + d_N = R$. The n -th ring is divided into $K(n)$ ring segments $c_{n,k}$ with angular span $\phi_{n,k}$ and angular offset $\phi_{0,n,k}$ parameterized by $\Psi_{n,k} = (\delta_n, r_{0,n}, \phi_{n,k}, \phi_{0,n,k})$ and spectral response function $\rho_{n,k}(\lambda)$ (Fig. 1(II)).

The goal is to optimize the parameters of a radial partition with respect to a end-to-end system performance metric. In this paper we consider the application of designing a RGB filter mask used for photographic image rendering for human observers. Similar to the traditional Bayer pattern, we want to capture twice as much green-filtered light arriving at the sensor than red- and blue-filtered light. That means the spectral distribution captured at the sensor, $\frac{I(\rho_m)}{\sum_n I(\rho_n)}$, should be the same as the target distribution $\alpha = \{\alpha_m\}_{m=1, \dots, 3}$, $\alpha = (0.25, 0.5, 0.25)$ for the blue, green and red color channels. We measure this similarity by $C_1(m) = \left| \frac{\alpha_m^{-1} I(\rho_m)}{\sum_n I(\rho_n)} - 1 \right|$. To estimate spectral crosstalk we define $C_2(m) = \frac{\int_{\Delta_1} J_{sensor}(\rho_m)}{\sum_n I(\rho_n)}$, and set out to minimize the cost function C defined as the weighted sum of $\|C_1\|_\infty$ and $\|C_2\|_\infty$:

$$C(\{\Psi_{n,k}, \rho_{n,k}\}_{n,k}) = w_1 \cdot \|C_1\|_\infty + w_2 \cdot \|C_2\|_\infty, \quad 0 \leq w_i \leq 1, w_1 + w_2 = 1, \quad (2)$$

subject to constraints given by the radial partition $\sum_{n=1}^N \delta_n = R, 0 \leq \delta_n \leq R, r_{0,n} = \delta_n - r_{0,n-1}, \sum_{k=1}^{K(n)} \phi_{n,k} = 2\pi, 0 \leq \phi_{n,k}, \phi_{0,n,k} \leq 2\pi$, and $h(\{\Psi_{n,k}, \rho_{n,k}\}_{n,k}) \leq 0$, where h is a nonlinear constraint function evaluating spatial extent of cell areas, such that feasible solutions to the optimization result in cells with a sensor image covering at least one sensor pixel (see also [2]). We evaluate C_1 and C_2 in the maximum norm to place significant weight on outliers.

3. Simulation results

We consider the task of designing a spectral mask for a lens system with specs described in Fig. 2 targeting a Bayer-filter-type spectral response distribution on the sensor, i.e. twice as usable green than red and blue pixels. The microlens array is positioned at the focal length of the main lens. We choose two types of aperture code partitions, the first consisting of three annuli-shaped cells $c_n, n = 1, 2, 3$ of diameter $\delta_n < R$, coded with spectral responses with center wavelengths $\lambda_1 = 450nm, \lambda_2 = 550nm$, and $\lambda_3 = 650nm$ ((a),(b) in Fig. 1(III)). In this case, we have two optimization variables δ_1, δ_2 and constraints $0 < \delta_n < R, \delta_1 + \delta_2 < R, n = 1, 2$ and $\delta_3 = R - \delta_1 - \delta_2$. The cells of the second partition are disc sectors, i.e. $c_n, n = 1, 2, 3, 4$ with $\delta_n = R$ ((c) in Fig. 1(III)), variables $\phi_{0,1}, \phi_n, n = 1, \dots, 3$ and constraints $\sum_{n=1, \dots, 4} \phi_n = 2\pi, 0 \leq \phi_n, \phi_{0,1} \leq 2\pi$.

Fig. 1(III) shows the light intensity on the sensor for different configurations of the filter mask. For all configurations, the non-optimized partitions are designed such that the area coverage of the partition cells satisfy the 1:2:1 relationship

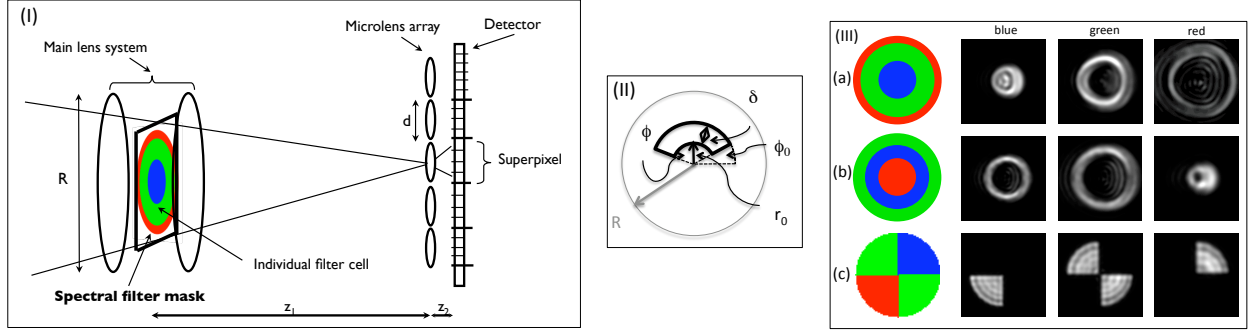


Fig. 1. (I) Plenoptic camera architecture with spectrally coded mask. (II) Partition parameterization. (III) Intensity $|J_{Sensor}(\rho_m)|^2$ at the sensor for annular and disc sector partitions.

between blue, green, and red filter areas. The annular layout with red at the outside suffers from severe spectral crosstalk caused by chromatic aberration, whereas having red inside and green outside reduces the effective spectral crosstalk due to the chromatic aberration. Aberrations of main and microlens cause further asymmetric distortions in the sensor images. Results of the optimization of all configurations are listed in the right table in Fig. 2. For the two annuli configurations, the optimization leads to significant decrease of the cost and significant adjustment of the ring diameters with (b) outperforming (a). For comparison, the disc sector configuration (c) creates the least cross-talk at the sensor and yields minimum total cost C . Its optimized version does not vary much compared to the starting configuration. Annuli configurations may have, however, advantages compared to the disc sector partition in applications that require rotationally symmetric filtering of the rays at the pupil plane or that require good depth-of-focus performance at the microlens plane [5]. Parameters in the cost function from Eq. 2 are set to $w_1 = w_2 = 0.5$ and $\alpha = (0.25, 0.5, 0.25)$. The Matlab routine *fmincon* is used to solve the constrained non-linear optimization problem.

4. Conclusions

We introduced an optimization framework for designing the spatial layout of a spectrally coded mask inserted into the pupil plane of the main lens of a plenoptic camera. The optimization merit function evaluates spectral crosstalk at the sensor as well as similarity of the captured spectral distribution to an application-specific target distribution.

Lens specs	Main lens	Microlens
F number	12.5	12.5
Focal length	50mm	1.650mm
Reference wavelength	588nm	588nm
Material	BK7	Fused Silica
Diameter	4mm	130 μ m
OPD-RMS	1/100 λ	1/20 λ

Annuli	δ_B	δ_G	δ_R		$C_1(B)$	$C_1(G)$	$C_1(R)$	$C_2(B)$	$C_2(G)$	$C_2(R)$	C
BGR non-optimized (a)	0.5	0.366	0.133		0.3399	0.6236	0.0364	0.2322	0.2728	0.8358	0.8450
BGR optimized	0.4169	0.3293	0.2539		0.2371	0.4966	0.2663	0.2891	0.3112	0.5775	0.3214
RBG non-optimized (b)	0.2071	0.2929	0.5		0.2178	0.4746	0.3076	0.5111	0.3841	0.3275	0.3708
RBG optimized	0.3084	0.2250	0.4666		0.2492	0.5038	0.2492	0.4489	0.3584	0.3586	0.2302
Disc segments	ϕ_B	ϕ_{G1}	ϕ_R	ϕ_{G2}	$C_1(B)$	$C_1(G)$	$C_1(R)$	$C_2(B)$	$C_2(G)$	$C_2(R)$	C
RGBG non-optimized (c)	1.5708	1.5708	1.5708	1.5708	0.2427	0.4978	0.2596	0.2390	0.2101	0.1697	0.1387
BGRG optimized	1.5709	1.5705	1.5709	1.5709	0.2542	0.4861	0.2597	0.2348	0.2151	0.1694	0.1369

Fig. 2. Left: Lens specs. Right: Non-optimized and optimized partition parameters and cost function values for the configurations from Fig. 1.

References

1. A. Levin et al., "Understanding camera trade-offs through a Bayesian analysis of lightfield projections," Proc. European Conference on Computer Vision, Marseille, France, October 2008.
2. R. Horstmeyer et al., "Flexible Multimodal Camera Using a Light Field Architecture," IEEE Conf. on Computational Photography, 2009.
3. D. Canvaugh et al., "VNIR hyperspectral camera system," Proc. of SPIE, vol. 6966, 2009.
4. J.W. Goodman, "Introduction to Fourier Optics," MacGraw-Hill, New York, 1986.
5. E. Tremblay et al., "Ultrathin cameras using annular folded optics," Applied Optics **14** (4) 463–471, (2007).