Evaluation of Multispectral Plenoptic Camera

Ricoh Innovations, Inc., Menlo Park, CA, USA 94025

ABSTRACT

Plenoptic cameras enable capture of a 4D lightfield, allowing digital refocusing and depth estimation from data captured with a compact portable camera. Whereas most of the work on plenoptic camera design has been based on simplistic geometric-optics-based characterization of the optical path only, little work has been done of optimizing end-to-end system performance for a specific application. Such design optimization requires design tools that need to include careful parameterization of main lens elements, as well as microlens array and sensor characteristics.

In this paper we are interested in evaluating the performance of a multispectral plenoptic camera, i.e. a camera with spectral filters inserted into the aperture plane of the main lens. Such a camera enables single-snapshot spectral data acquisition.\(^1\)\(^-\)\(^3\)

We first describe in detail an end-to-end imaging system model for a spectrally coded plenoptic camera that we briefly introduced in.\(^4\) Different performance metrics are defined to evaluate the spectral reconstruction quality. We then present a prototype which is developed based on a modified DSLR camera containing a lenslet array on the sensor and a filter array in the main lens. Finally we evaluate the spectral reconstruction performance of a spectral plenoptic camera based on both simulation and measurements obtained from the prototype.

1. Formulation of System Model

As shown in Fig. 1 our proposed system model for a multispectral plenoptic camera includes a statistic model of the source, a geometric model of the optical response of a plenoptic camera, a detector model, and reconstruction algorithms.

1.1 Source model

In the source model the radiance reflected from the object is estimated. The inputs to the source model are irradiance of light source \(E(\lambda)\), reflectance of object \(R(\lambda)\), and light incident angle \(\theta_i\). The irradiance and reflectance are wavelength (\(\lambda\)) dependent. The radiance reflected from a Lambertian object surface is calculated as

\[ \text{Radiance} = E(\lambda) \cdot R(\lambda) \cdot \cos(\theta_i) \]

*The author performed the work at Ricoh Innovations, Inc.
Multimodal plenoptic camera
Multimodal plenoptic imaging system

• Insertion of filter module into main lens system

![Filter module](image)

Figure 2: Overview of plenoptic camera architecture with spectral filters inserted into the main lens.

\[ L(\lambda) = \frac{1}{\pi} E(\lambda) R(\lambda) \cos(\theta_i) + N_{\text{source}}, \]  

where \( N_{\text{source}} \) is the source noise due to the texture variation of the object surface.\(^5\)

1.2 Camera model

In this paper, we assume the camera to be a plenoptic camera. These cameras are designed to capture different combinations of light rays from a scene, sampling its lightfield. 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, measuring directional scene information as it is passing through the pupil plane. Whereas most plenoptic camera designs capture directional information of rays and target applications such as digital refocusing, rotation, or depth estimation, only few address capturing spectral information of the scene. Authors in\(^1,2\) modify a plenoptic camera with a filter array inserted in the optical path of the main lens, containing different spectral and polarization filters. With this modification, sampling of the spectral dimension of the plenoptic function is performed. Light rays originating from one point source enter the pupil plane at different locations, therefore, passing through different filters. Since the microlens array images the pupil plane onto the sensor, spectral scene information as it is passing through the pupil plane is captured at the sensor (Fig. 2). As a result, the plenoptic camera is turned into a single-snapshot multispectral imaging system that trades-off spatial with spectral information captured with a single sensor.

In this section we derive a model for a plenoptic camera with spectral filters inserted into the aperture of the main lens. We assume the main lens to be focused on the microlens array plane. The radiance \( L(\lambda) \) originating from an on-axis object point passed through optics is converted to a digital output signal. Considering a multispectral plenoptic camera, the light passing through its \( i \)th filter in the aperture is calculated as

\[ b_i = \int_{\Delta \lambda} \frac{\tau_o L(\lambda) A_o}{r^2} \cdot \frac{A_d \lambda t}{h \cdot c} \cdot \rho_i(\lambda) d\lambda \cdot g_e, \]  

where \( A_o \) is the area of a sub-aperture occupied by the \( i \)th filter, \( r \) is the distance between main lens and detector, \( \tau_o \) is the optical transmittance, \( \rho(\lambda) \) is the system spectral sensitivity, \( A_d \) is the pixel area, \( t \) is the exposure time, \( \Delta \lambda \) is the filter bandwidth, \( \lambda_c \) is the center wavelength, \( h \) is the Planck's constant, \( c \) is the speed of light, and \( g_e \) is the sensor gain.\(^6\)

For a single object point the intensity at each sensor location depends on the optical response function of the system. Using ray tracing through the optical system, a geometric approximation of the optical response of the system can be obtained. This optical response is not a tight point spread function forming an image of an object point as conventional imaging systems, but a wide response forming an image of the pupil. This image of the pupil covers the entire sensor area under a lenslet, so called super-pixel.\(^7\) We calculate the fraction of the number of rays that hit a sensor pixel, considering an aperture mask in the main lens containing several filters. Each sensor pixel measurement can be thought of as a linear combination of the spectral intensities...
\[ x_i = \sum_{j=1}^{N} f_{i,j}b_j, \]  
where the coefficients \( f_{i,j} \) model the fraction of the number of rays passing through \( N \) filters arriving at sensor location \( i \). Sensor measurements are modeled as

\[ x = Fb + n, \]  

where \( x = [x_1, \ldots, x_M]^T \) is a vector containing plenoptic sensor data for a super-pixel, \( F \) is the system response matrix containing the coefficients \( f_{i,j} \), \( b = [b_1, \ldots, b_N]^T \) is a vector containing the spectral intensity values, \( n \) is sensor noise. The fraction matrix \( F \) models multiplexing of the different spectral responses due to chromatic aberration of the optical system and sensor pixelation.

### 1.3 Spectral Reconstruction Algorithms

The spectral feature vector is extracted from the plenoptic sensor data vector \( x \) via a linear model defined as

\[ y = \Phi x = \Phi(Fb + n), \]  

where \( y = [y_1, \ldots, y_N]^T \) is the extracted feature vector, and \( \Phi \) is the spectral reconstruction matrix.

Three different approaches are used in this paper for spectral reconstruction. 1) Averaging contiguous sub-pixels that are considered as collecting light from the same spectral filter. 2) Extracting a single pixel in each cell of the mask that has the maximum response to the corresponding spectral filter. 3) Our novel system-dependent spectral demultiplexing algorithm. The output signal is demultiplexed based on a calibrated system response matrix \( \hat{F} \). The spectral features are extracted by taking a pseudoinverse of \( \hat{F} \), i.e., \( \Phi = (\hat{F}^T\hat{F})^{-1}\hat{F}^T \).

### 1.4 Performance Metrics

The spectral reconstruction quality is evaluated from two different perspectives. The spectral reconstruction accuracy measured as the mean-squared-error metric (Err)

\[ Err = \sqrt{E[(y - b)^2]}, \]  

where \( b \) is the spectral intensity vector and \( y \) is the the estimated spectral intensity vector at a specified object location.

The signal-to-noise ratio (SNR) of the extracted spectral signals

\[ SNR = \frac{\bar{y}}{\sigma_y}, \]  

where \( \bar{y} \) is the mean of reconstructed spectral intensity and \( \sigma_y \) is the standard deviation of the reconstructed spectral intensity.

### 2. PROTOTYPE OF MULTISPECTRAL PLENOPTIC CAMERA

A multispectral plenoptic camera prototype has been developed by modifying an off-the-shelf DSLR camera. The modified camera body is shown in Fig. 3. A microlens array is placed over the CMOS sensor using a mechanical mount. The microlens array has a lens pitch of 250 \( \mu \)m and a focal length of 1.026mm. From these specs the F number of the microlens array results in 4.17. For the main lens we chose a fixed-focus 50mm lens, inserted a filter array between its aperture and one of the two lens groups, and stopped the lens down to an F number 4.5 in order match that of the microlens array. The modified main lens is shown in Fig. 4. As shown in Fig. 4(a), an array of four different narrowband spectral filters is inserted in the main lens aperture. The narrowband filters have 20 nm bandwidth, and their spectral responses are shown in Fig. 4(b). The filter array was custom-made by Ocean Thin Films. An example of the output image from the spectrally coded plenoptic camera is shown in Fig. 5.
Figure 3: Modified DSLR camera body by placing microlens array over the CMOS sensor.

Figure 4: Main lens with spectrally coded aperture. (a) Array of narrowband spectral filters inserted in the aperture. (b) Spectral response of the four spectral filters.

Figure 5: Color checker image capture by (a) regular RGB camera, and (b) multispectral plenoptic camera.
Table 1: Spectral reconstruction quality comparison.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Object Distance</th>
<th>Average [DN]</th>
<th>Single Pixel [DN]</th>
<th>Demultiplexing [DN]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Err [DN]</td>
<td>( z_1 )</td>
<td>243.48</td>
<td>7.06</td>
<td>2.80</td>
</tr>
<tr>
<td></td>
<td>( z_1 - \Delta z )</td>
<td>254.74</td>
<td>6.94</td>
<td>26.54</td>
</tr>
<tr>
<td></td>
<td>( z_1 + \Delta z )</td>
<td>234.40</td>
<td>6.95</td>
<td>16.19</td>
</tr>
<tr>
<td>SNR [dB]</td>
<td>( z_1 )</td>
<td>45.76</td>
<td>37.61</td>
<td>45.34</td>
</tr>
<tr>
<td></td>
<td>( z_1 - \Delta z )</td>
<td>45.89</td>
<td>37.64</td>
<td>45.35</td>
</tr>
<tr>
<td></td>
<td>( z_1 + \Delta z )</td>
<td>45.86</td>
<td>37.56</td>
<td>45.46</td>
</tr>
</tbody>
</table>

Figure 6: Super-pixel localization in the raw image.

3. RESULTS

Simulation experiments have been performed to evaluate the performance of a multispectral plenoptic camera using our proposed system model for certain source, optics, and sensor parameters. In the camera model a spectrally coded plenoptic camera with a \( 3 \times 3 \) square filter layout design was considered. The spectral reconstruction error and SNR of extracted features based on the three different spectral reconstruction algorithms are calculated based on Monte Carlo experiments and compared in Table 1. The ground truth is the spectral intensity integrated into one pixel corresponding to each of the filters. The results shown are the averaged values based on the reconstructed spectral intensity at 9 different wavelengths. In Table 1 the results are also shown when the object distance \( z_1 \) shows certain variability (\( \Delta z = 25 \text{mm} \)). In the simulation when the object is at distance of \( z_1 \pm \Delta z \) the system is out-of-focus, but the system response matrix calibrated at in-focus condition is still applied to reconstruct the spectral intensity. It can be seen that the average method shows very poor spectral reconstruction accuracy, due to spectral crosstalk caused by chromatic aberration originating from the lenses and pixelation from the sensor. The single pixel extraction, however, provides high reconstruction accuracy, but leads to very low SNR. The demultiplexing method presents low spectral reconstruction error and maintains high SNR. It is also noticed that the performance of demultiplexing method varies with varying object distance. This is due to the fact that the system response matrix is not characterized at the out-of-focus condition, and improvement could be achieved by calibrating the system at different configurations.

The system performance is further evaluated using our prototype. Spatial and spectral calibrations are conducted on the prototype. Spatial calibration is first performed to localize super-pixels in the raw image. Morphological filtering is applied to extract the centroid of each super-pixel. An example of super-pixel localization is shown in Fig. 6.

Spectral calibration is then performed to construct the system response matrix. To calibration the system response monochrome light corresponding to the wavelength of each spectral filter is sent through the main lens. The response of each filter can then be captured and used to form the system response matrix. Examples of the response by sending monochrome light with central wavelength of 540 nm and 650 nm are shown in Fig. 7.

Downloaded From: http://proceedings.spiedigitallibrary.org/ on 07/08/2013 Terms of Use: http://spiedl.org/terms
The single pixel extraction approach and our spectral demultiplexing method are used in the evaluation of prototype system performance. Reflectance values measured from different patches in a color checker are used as ground truth. The reconstructed spectral intensity values are normalized to reflectance based on a reference image captured on a Labsphere reflectance target. The spectral reconstruction error is computed at four different wavelengths, and by averaging the values calculated based on the white, black, white, red, orange, green, and blue patches. In this evaluation the reflectance is reconstructed from only one lenslet. The results based on single pixel approach and demultiplexing are compared in Fig. 8. It is seen that our demultiplexing approach is comparable to the single pixel extraction approach, showing, that we can compensate for the chromatic system distortions by demultiplexing the sensor data according to the lens characteristics.

The system performance is further evaluated based on the SNR. The SNR is computed based on the image of a white object captured with two different exposure settings. The results are shown in Fig. 9. It is shown that the spectral demultiplexing method reconstruct the spectral information with much higher SNR.

4. CONCLUSIONS
We have introduced an end-to-end system model for a multispectral plenoptic camera, and a novel spectral demultiplexing algorithm. We also presented our prototype of multispectral plenoptic camera, which was developed by modifying an off-the-shelf DSLR camera. Finally we compared different spectral reconstruction methods by evaluating the system performance. The system performance was evaluated based on both simulations using the system model and real images collected using our prototype. A good match between the simulation and prototype data was obtained, and the capability of our spectral demultiplexing algorithm has been demonstrated.
Figure 9: SNR comparison at two different exposure settings: (a) 1/80 sec, and (b) 1/4 sec.

REFERENCES