

Modified light field architecture for reconfigurable multimode imaging

Roarke Horstmeyer, Ravindra Athale, Gary Euliss
The MITRE Corporation, 7515 Colshire Dr., Mclean, VA 22102

ABSTRACT

Light field cameras can simultaneously capture the spatial location and angular direction of light rays emanating from a scene. By placing a variable bandpass filter in the aperture of a light field camera, we demonstrate the ability to multiplex the visible spectrum over this captured angular dimension. The result is a novel design for a single-snapshot multispectral imager, with digitally reconstructed images exhibiting reduced spatial resolution proportional to the number of captured spectral channels. This paper explores the effect of this spatial-spectral resolution tradeoff on camera design. It also examines the concept of utilizing a non-uniform pinhole array to achieve varying spectral and spatial capture over the extent of the sensor. Images are presented from several different light field – variable bandpass filter designs, and limitations and sources of error are discussed.

Keywords: Light field, aperture filter, spectral imager, pinhole array, PSF engineering

1. INTRODUCTION

A light field camera is a device capable of capturing the radiance of light rays as a function of two spatial dimensions and two angular dimensions. Early research in the area of integral photography led to camera designs that could record these four dimensions on a two dimensional plane of film^{1, 2}, by placing, for example, an array of pinhole lenses or a lenticular array at the image plane of the system. Since the arrival of digital sensors, a number of novel light field camera architectures have been proposed that work in conjunction with digital reconstruction methods for applications like depth estimation³, digital refocusing⁴⁻⁷, and glare reduction⁸.

In a similar line of research, integral photography techniques were adopted over one hundred years ago to achieve color imaging. Initial attempts used a pinhole array placed over film at the image plane along with either a prism⁹ or color filters^{10, 11} in the pupil plane of a camera. Berthon¹² introduced the use of a lenticular array at the image plane with a red, green and blue filter in the aperture. This design eventually became the Kodak Kodacolor film process, which was used extensively before the advent of color emulsion film¹³.

Digital sensors have allowed for these early color imaging approaches to find a new place in the evolution of camera technology. Levoy et al.¹⁴ recently suggested that various filters placed in the objective aperture of a light field microscope could allow for the modulation of the spectrum, polarization state, or intensity profile of light across a single snapshot. This concept was recently implemented in a conventional camera using a pinhole array at the image plane and an array of spectral, polarization and neutral density filters placed in the aperture of the lens¹⁵. Work presented in this paper will be an extension of this previous multimodal light field architecture, but will focus on the insertion of a continuously variable wide band spectral filter into the lens aperture, effectively transforming the design into a multispectral imager. In the next section, spectral imaging will be examined using previous examples, which will be followed by a brief introduction to light field imaging concepts. The tradeoff between spectral and spatial resolution for the imager presented in this paper will then be discussed, and the idea of a variable resolution multispectral imager will be introduced. Preliminary results will be presented demonstrating single-shot spectral data cube capture.

2. BACKGROUND

Using a conventional camera to simultaneously capture the two dimensional image of a scene along with its color spectrum presents a dimensionality mismatch, in which three dimensions of information must be recorded on a two

dimensional sensor. A variety of techniques, which could be described as multiplexing schemes, have been suggested to overcome this dimensionality mismatch. In the simplest case, the spectral dimension is mapped onto the time axis (i.e., time-division multiplexing). An early example of this approach is the first color motion picture camera, which had a rotating color filter wheel to capture sequential frames from behind a different filter¹⁶. A more modern example of time division multiplexing is a pushbroom hyperspectral imager such as HYDICE¹⁷, which captures and combines sequential frames of one spectral and one spatial dimension to reconstruct the three dimensional data cube.

Spatial multiplexing is another technique used to address the dimensionality mismatch. For example, the common Bayer filter is a spatially interlaced distribution of filters placed at the image plane that captures three spectral channels per image. Multi-aperture imaging systems like PERIODIC¹⁸ and TOMBO¹⁹ spatially multiplex at the aperture level to record separately filtered images on a common focal plane array (FPA).

Another approach to single shot spectral imaging is through the use of code division multiplexing. Coded aperture snapshot spectral imaging²⁰ (CASSI) combines either one or two dispersive elements with a coded aperture to compress the three dimensional data cube down to two dimensions. Image and spectrum reconstruction is then based upon a sparse-scene estimation algorithm. Alternatively, a computed tomography imaging spectrometer²¹ (CTIS) uses a unique computer generated hologram to create an array of diffraction images with different amounts of dispersion. Each of the images in the array is a unique projection of the spectral data cube, and an expectation-maximization algorithm can be used to reconstruct the data cube from these multiple two dimensional projections.

The approach we describe in this paper measures spectral diversity by multiplexing over the angular dimension of light focused from the camera lens to the FPA. This is achieved by adding both a filter with a spectral transmission profile that varies continuously along one direction in the lens aperture, along with a light field configuration that captures a mix of spatial and angular measurements. While quite similar to work presented by Horstmeyer et al.¹⁵ using a discrete filter array, implementation with a continuous filter in the aperture allows more flexibility in spectral image analysis. More opportunity is also provided by this approach to explore the camera's tradeoff between spectral and spatial resolution.

3. APPROACH

Various detailed descriptions of light-field imaging can be found in previous literature³⁻⁵. The approach considered here is most closely represented by designs described in Raskar et al.⁸ and Horstmeyer et al., in which an array of pinhole lenses is placed in the back focal plane of the field lens of what is otherwise a conventional imaging architecture (Fig. 1).

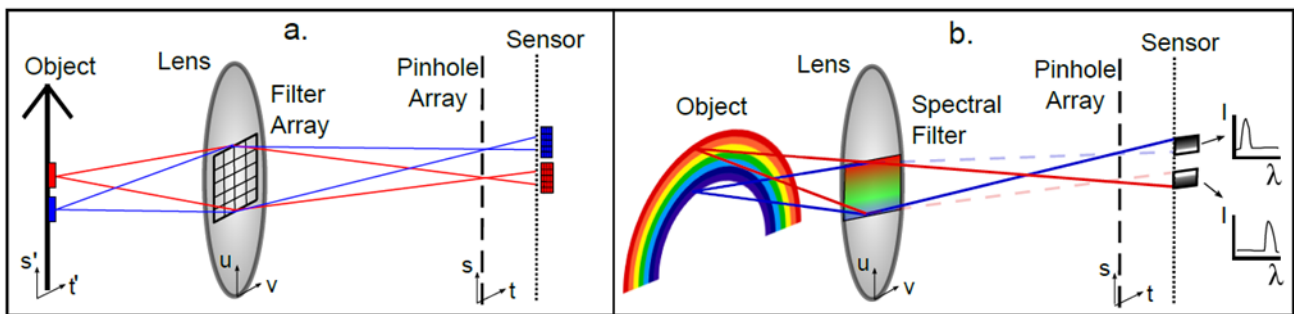


Figure 1. (a) A diagram of a conventional light field imaging system using a pinhole lens array with a filter array inserted into the aperture. An object is imaged by a lens in (u,v) coordinates onto a pinhole array in (s,t) coordinates. Each pinhole in the array effectively images the lens aperture, providing a four dimensional intensity measurement $I(s,t,u,v)$ at each pixel on the sensor. (b) Diagram of a snapshot spectral imager based on a light-field imaging architecture. Instead of the filter array in (a), a linear continuous spectral filter can be placed in the lens aperture. Each pinhole at a specific (s,t) coordinate images the (u,v) plane, providing a one-dimensional spectrum for different object points. In this setup, spatial resolution will be determined by the number of pinholes in the array, while spectral resolution will be given by the number of sensor pixels under each pinhole along the s dimension.

Pinhole lenses offer a low cost alternative to lenslet arrays at the expense of optical efficiency. Each pinhole lens in the array forms an image of the field lens aperture on the FPA. We adopt the convention of defining the field lens aperture as the (u, v) plane and the pinhole array as the (s, t) plane. Taking a geometric view of the system, a light ray from an object point (s', t') will pass through a point in the (u, v) plane and a point in the (s, t) plane and be detected at a pixel in the FPA. Each point on the FPA thus corresponds to a unique point in (s, t, u, v) space. Each point in object space will be mapped into multiple points in (s, t, u, v) , where (s, t) will correspond to a single object point location and (u, v) the point at which the associated light ray intersects the aperture. Following acquisition, the raw light-field (intensity) image can be represented by $I(s, t, u, v)$, and a set of images $I_{u,v}(s, t)$ can be constructed corresponding to the light collected by each individual region of the field lens.

Multiplexed filtering in the light-field domain has been previously demonstrated¹⁵ by placing an array of discrete filters in the (u, v) plane as illustrated in Fig. 1.a. Suppose that instead of discrete filters, a continuously variable spectral filter is placed in the aperture, as shown in Fig. 1.b. Each point in (u, v) then corresponds to a specific wavelength $\lambda(u, v)$ as determined by the spectral filter. The light incident on the focal plane at (s, t, u, v) will be proportional to the $\lambda(u, v)$ spectral component of the light reflected from a particular (s', t') coordinate. Therefore, the set of images $I_{u,v}(s, t)$ becomes $I_{\lambda(u,v)}(s, t)$, a three-dimensional multispectral data cube. In practice, the finite extent of pixels in the FPA will place a limit on the (u, v) , and hence the $\lambda(u, v)$, resolution of the imaging system. In the case of the one-dimensional spectral filter in Fig. 1.b, the wavelength is independent of u and $\lambda(u, v) \rightarrow \lambda(v)$. The spatial resolution of this particular system is determined by the number of pinhole lenses in the array, while the spectral resolution is a function of the number of pixel rows occupied by each pinhole lens-formed image.

4. SYSTEM DESIGN AND TRADEOFFS

A thorough analysis of the design space for a light field camera with a two dimensional filter array in the aperture can be found in Horstmeyer et al.¹⁵, where an optimal pinhole size, filter size and array geometry is determined. That analysis will now be extended to take into account both the continuous nature of the spectral filter as well as its response along a single dimension. Design limitations due to physical optics effects will also be discussed.

Figure 2.a shows a simplified diagram of the camera layout. As described before, each pinhole lens in the (s, t) plane will geometrically project a separate image of the variable filter onto the FPA in the (x, y) plane. The object plane for

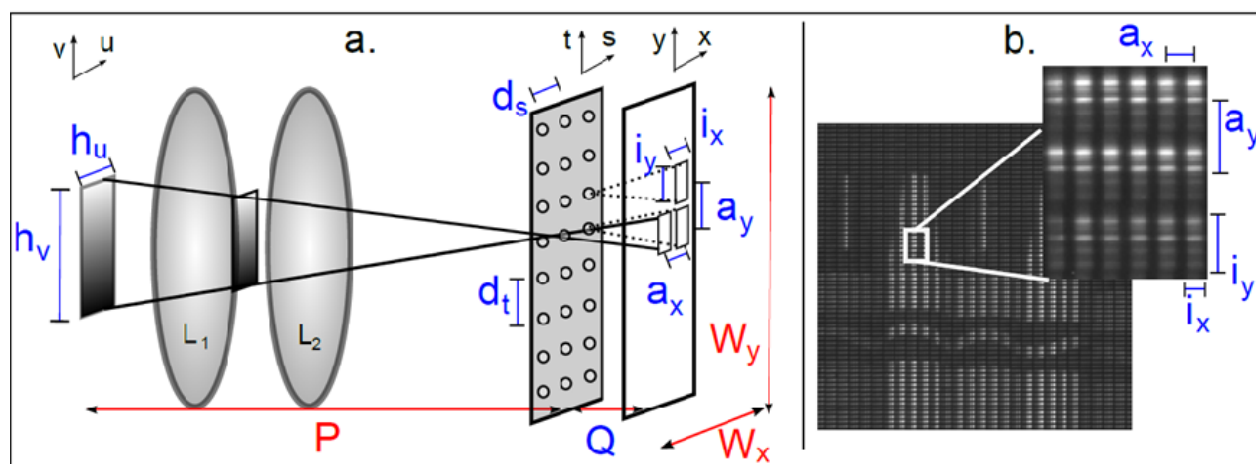


Figure 2. Diagram of the experimental setup with relevant fixed (red) and variable (blue) distances labeled. (a) Each pinhole lens in (s, t) images the virtual image of the continuous filter in (u, v) onto the FPA in (x, y) . Pinhole pitches in (s, t) are chosen to tile the pinhole lens images as close to one another as possible without overlapping. (b) A portion of a raw image of crayons from the setup with a particular region of 3×6 pinhole lens images enlarged. Each vertical strip in this region is a spectral filter image corresponding to a unique object location. To make the most efficient use of the FPA, parameters should be chosen so that (a_x, a_y) and (i_x, i_y) are equal.

these images will be a virtual image plane (u, v) slightly behind the aperture plane of a compound lens, due to the intermediate image formed by L_2 . Given a rectangular filter, the size (h_u, h_v) and distance P of the virtual image can be determined using the focal length of L_2 , or measured by experiment. The extent of each pinhole image along x and y can be expressed as

$$(i_x, i_y) = \frac{Q}{P}(h_u, h_v). \tag{1}$$

Figure 2.b displays a portion of a raw data image with the size (i_x, i_y) labeled. To maximize the amount of data recorded on the sensor, the pinhole images are formed as close together as possible without overlapping. This is accomplished by selecting the pitch of the pinhole array along s and t (which is the pitch of the pinhole lens images) to match the size of each image:

$$(d_s, d_t) = (a_x, a_y) = (i_x, i_y). \tag{2}$$

This design constraint is analogous to the f-number matching of the main lens and lenslets in other light field camera designs^{5, 14}.

The number of pinholes (n_x, n_y) that can fit in an array over a FPA of size (W_x, W_y) is then

$$\begin{aligned} n_x &= \frac{W_x}{a_x} = \frac{W_x P}{h_u Q} \\ n_y &= \frac{W_y}{a_y} = \frac{W_y P}{h_v Q} \end{aligned} \tag{3}$$

which gives the spatial resolution of the camera for a single spectral channel. For calculating spectral resolution, it must be noted that the variable filter in the aperture is one dimensional, so the spectrum is mapped only along one dimension of the FPA (the y -axis in the Fig. 2.a). Measurements will be integrated along x for each pinhole lens image to improve

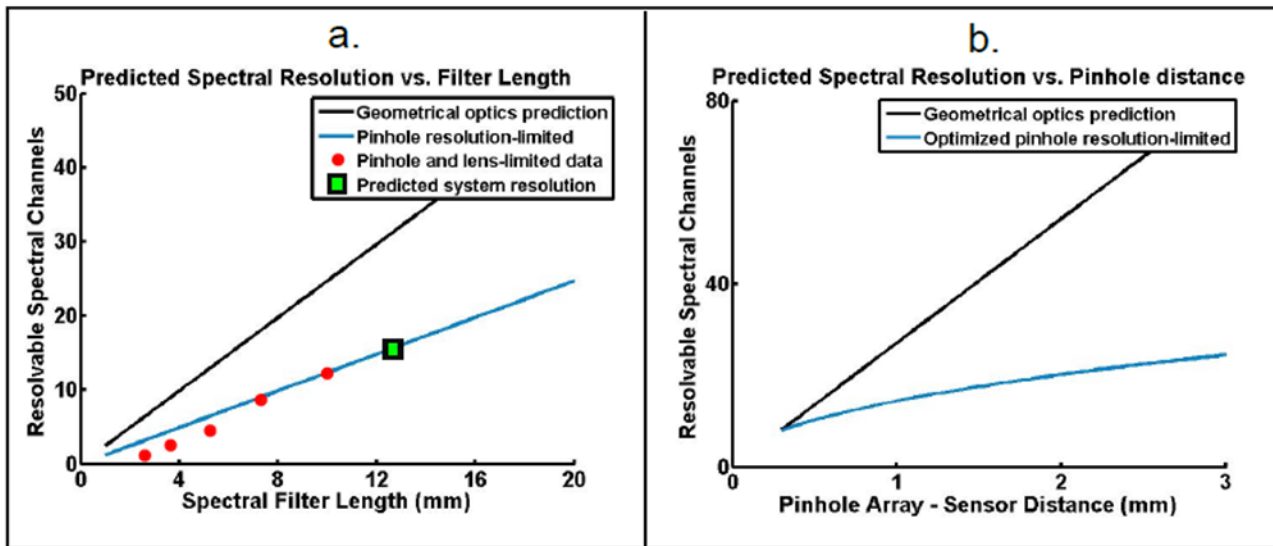


Figure 3. Plots of spectral resolution tradeoffs under geometrical and physical optics assumptions. For this design, $P=51.23$ mm, $Q=1.14$ mm, $W_y=36.1$ mm, $s_y=4008$, and $h_v=12.5$ mm. (a) Spectral resolution increases with a longer spectral filter in the field lens. Blurring from pinhole lenses cuts the number of resolvable spots on a $9 \mu\text{m}$ pixel detector approximately in half. Measured blurring from the field lens PSF further decreases spectral resolution. (b) Spectral resolution will increase with a greater pinhole distance. Blurring from pinhole lenses will decrease spectral resolution on a $9 \mu\text{m}$ pixel detector. Given pinholes with an optimized diameter that grows with image distance, resolution decrease will be quadratic.

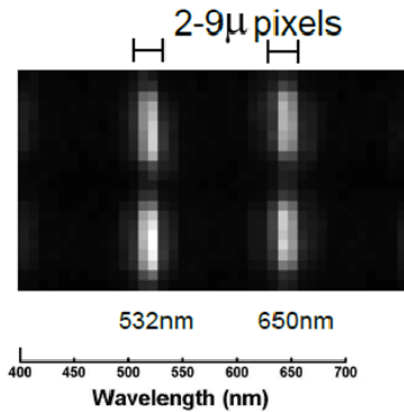


Figure 4. Raw experimental data of two laser light sources at 532 nm and 650 nm used for calibration. The two pinhole lens images shown span approximately 27 detector pixels along y and 7 pixels along x . Resolution from the pinholes is optics limited, causing the monochromatic peaks to span approximately two 9 μm detector pixels.

the signal-to-noise ratio. If the resolution of the FPA is (s_x, s_y) pixels and the system is detector-limited, then spectral resolution can be expressed as

$$\lambda_r = \frac{s_y}{n_y} = \left(\frac{s_y P}{W_y} \right) \frac{h_v}{Q}, \quad (4)$$

assuming the pinhole pitch is chosen according to Eq. 2. Here, the inverse relationship between the number of resolvable spectral values and the spatial resolution of the image, given by the number of pinholes, is clear. For a typical camera, P (which is set by the field lens focal length), the number of pixels s_y and FPA width W_y , will all be fixed. The two remaining variables that can be used to alter spectral resolution are the length of the aperture filter h_v and the pinhole array distance Q .

Separate geometrical and physical-optics based plots of the dependence of spectral resolution λ_r on the length of the spectral filter h_v are displayed in Fig. 3.a. The geometrical plot is based upon the above algebraic derivation, with a detector-limited assumption. This assumption is approximately valid if a large field lens aperture is used with a lenslet array at the focal plane.

The effects of physical optics on resolution tradeoffs are included in Fig. 3.a using two experimental-based insights. First, a pinhole lens will typically have an associated point spread function (PSF) radius given in Young²² as

$$radius_{PSF} = 1.5 \sqrt{\lambda \left(\frac{1}{P} + \frac{1}{Q} \right)^{-1} \left(1 + \frac{Q}{P} \right)}. \quad (5)$$

This is typically larger than the size of a detector pixel. Fig. 4 displays a raw measurement from the system, where monochromatic sources are mapped by a pinhole lens to roughly two sensor pixels each. We can expect the design's actual spectral resolution to be roughly half what is predicted by Eq. 4. Second, experimental data of diffraction effects from the field lens aperture is included. The physical dimensions of the spectral filter define the aperture of the primary lens and therefore the characteristics of the associated PSF. As the filter length decreases, the associated PSF support increases for a fixed Q . This contributes to a larger blur spot on the FPA and lower spectral resolution.

The relationship between spectral resolution λ_r and the pinhole array-sensor distance Q is shown in Fig. 3.b. Eq. 4 again defines the geometrical optics-based relationship. The effect of physical optics is incorporated through the same assumptions of a finite pinhole lens and field lens PSF. In this plot, as Q is increased, an optimized pinhole size from Young will produce a quadratically varying pinhole lens PSF support.

An increase in spectral resolution is met with an inversely proportional decrease in spatial resolution. Each of these plots can be inverted using Eq. 3 to determine the expected spatial resolution variation of the system for a given h_v and Q . In general, resolution losses could be avoided by using a lenslet array instead of a pinhole array.

5. EXPERIMENTAL RESULTS

The light field camera used to obtain experimental results consisted of a Nikon 50mm f/1.8 lens and a 4008 x 2672 board level Lumenera monochrome CCD sensor with 9 μm square pixels. Pinhole arrays were printed on transparencies at 5080 dpi with approximately square holes of 50 μm width and different respective pitches. A transparency chosen to match the desired spectral resolution was sandwiched between the CCD cover glass and a second glass plate (Fig. 5.a). The thickness of the cover glass was approximately 1.14 mm and sets the distance Q between the pinhole array and the

CCD. The spectral filter was 12.5 mm long x 2.75 mm wide x 1.1 mm thick with a spectral range of 400-700 nm and a peak transmittance of approximately 30%. The filter was placed directly against the aperture stop of the field lens as shown in Fig. 5.c.

Spectral images are constructed from the raw light-field data by first determining the mapping of wavelengths onto the CCD for each sub-image formed by the pinhole lens array. Once the mapping is established, a center reference point for each sub-image along the s dimension is identified and pixels can be indexed according to wavelength. Pixels with the same wavelength index from all sub-images combine to form a spectral image

in a given waveband, with the complete set of indices forming a multi-spectral data cube. Due to the finite width of the linear filter, pixels in each sub-image are averaged along t (i.e., the direction perpendicular to spectral variation) to improve SNR. Comparison with a broadband Lambertian prior removes fixed pattern noise caused by defects in the

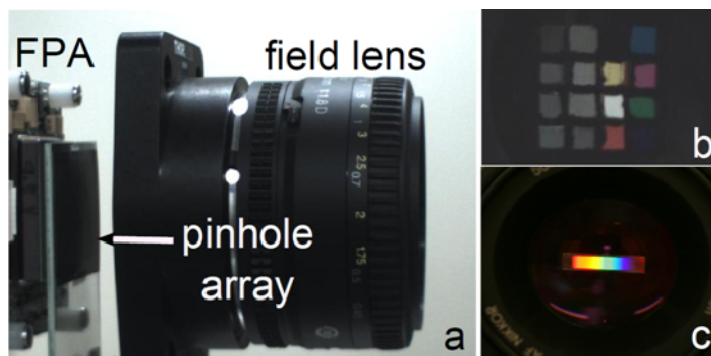


Figure 5. Images of the experimental setup. (a) A field lens images onto a pinhole array placed directly over a CCD. (b) A close-up of the 16-filter array placed in the aperture of the field lens to create the images as presented in [15] and shown in Fig. 6. (c) A close-up of the spectral filter in the field lens used to create the images in Fig. 7 and Fig. 8.

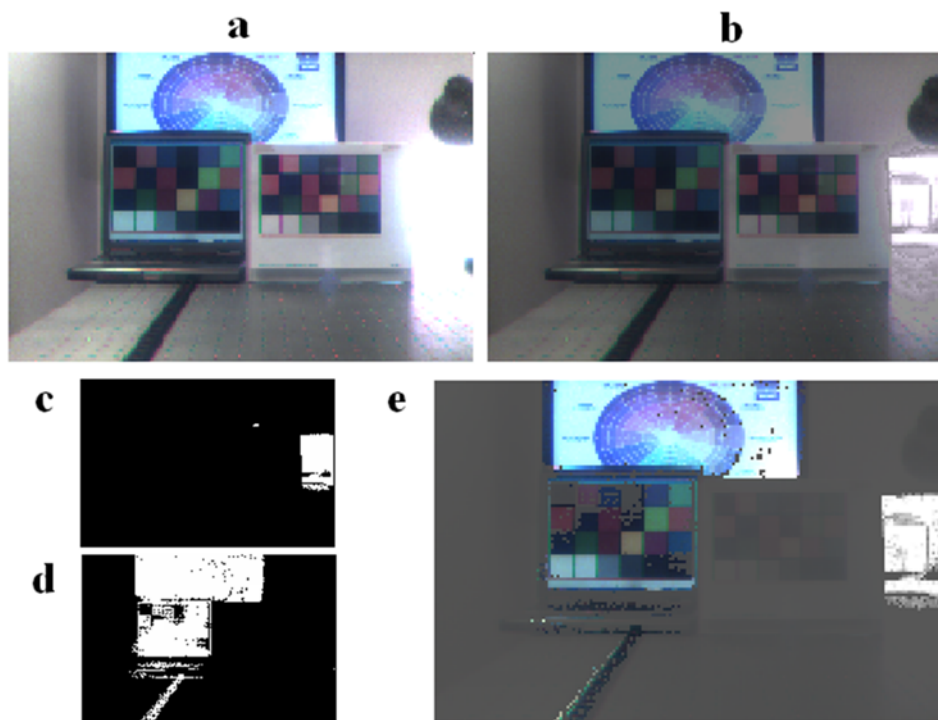


Figure 6: From [15], an example application of placing an array of sixteen filters in the aperture plane of a light field camera. Pinhole pitch was made with a symmetric $200 \mu\text{m} \times 200 \mu\text{m}$ pitch to match the square filter array. The resolution of each image is 177×117 , and a Lambertian prior was used for filter density comparison. (a) Three synthetic images (each an $I_{u,v}(s, t)$ map with (u, v) corresponding to either a red, green or blue filter) are combined to create an RGB image. Note that areas under the lamp are saturated. (b) Three synthetic images corresponding to neutral density filtered (u, v) coordinates are added to create a simple color high dynamic range image. (c) A binary map created by the near-IR synthetic image with a threshold of 50%. (d) Synthetic images from behind polarizer filters oriented at 0° , 45° , 90° , and 135° are combined to create a linear degree of polarization map with a 0.35 threshold. (e) Image (b) combined with binary maps (c), (d) to create a region-of-interest color HDR image.

pinhole array as well as the effects of non-uniform transmission across the filter.

To highlight the flexibility of the presented camera design, an example implementation of a filter array placed in the aperture plane is presented in Fig. 6. The filter array used to create these images is pictured in Fig. 5.b. A more specific look at this pupil plane filter array configuration is presented in Horstmeyer et al.¹⁵, but its response can be thought of as essentially a discrete form of the continuous spectral filter with the addition of polarization and neutral density measurements.

Inserting the continuous spectral filter into the aperture stop of the camera and a 100 μm x 250 μm pitched pinhole array at the focal plane, we are able to capture data displayed in Fig. 7. This displays an RGB image of different colored crayons at its center, created by splitting up and summing the 27 measured spectral channels into three blocks of 9 channels each. The images are scaled appropriately to account for the difference in pitch along the two axes of the pinhole array. Every spatial measurement in this 141 x 235 pixel image contains 27 spectral measurements. Four of these spectra, corresponding to pixels from different colored crayons, are displayed on the sides of the image.

With the same spectral filter and pinhole array setup, the four different crayon images in Fig. 8 were also captured. Each of these images is a sum over all 27 spectral channels and was captured under different types of illumination. The top left uses an incandescent light bulb, top right an incandescent bulb with florescent room lights, bottom left a bright-white (3500K) compact florescent light (CFL), and bottom right the CFL with florescent room lights. Note that although the grayscale images appear quite similar, the spectrum of an example point on a teal crayon changes dramatically.

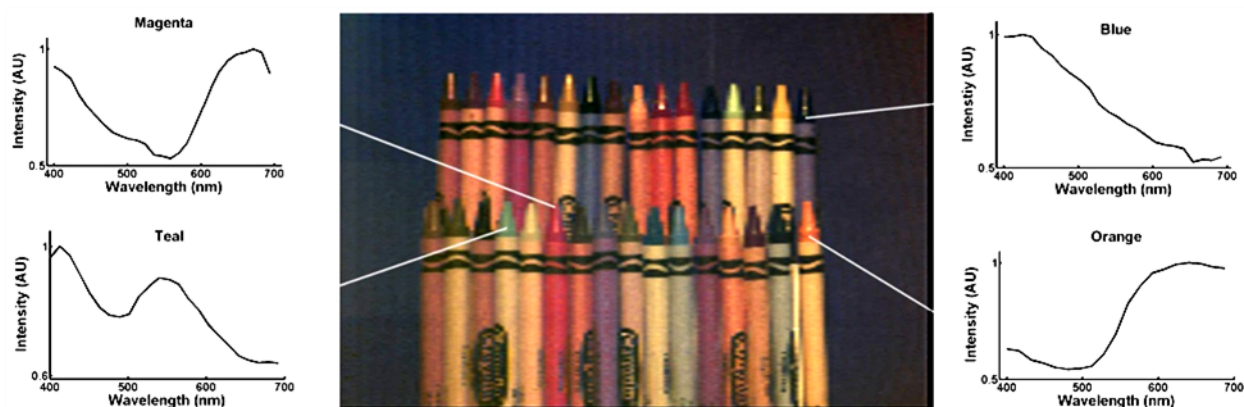


Figure 7. Results from the pinhole array – spectral filter setup of a box of crayons under lamp light after comparison to a broadband prior. The color image was created by equally dividing up and averaging the 27 measurable spectral channels for each pixel into three bins representing red, green and blue values. Four spectrum plots are shown for single pixels from various colored crayons (magenta, teal, blue and orange). For every spatial measurement, a 27-channel spectrum is captured ranging from 400 to 700 nm.

6. CONCLUSION

The addition of a spectral filter to the aperture of a light field camera offers a simple and flexible design for a single-snapshot multispectral imager. There are several parameters that can be varied within the camera to explore a tradeoff between spatial and angular resolution, each exhibiting unique tradeoffs. We found an optimal configuration by minimizing the pinhole array-detector distance and maximizing the spectral filter width.

The design presented in this paper is a simple proof-of-concept demonstration, and there are a number of improvements and future areas for exploration. First, large improvements could be made with the use of a lenslet array instead of a pinhole lens array at the focal plane, as resolution losses could be reduced significantly. Second, since the filter used was

only one dimensional, future work could include the design and analysis of an optimal two dimensional spectral filter for use in the aperture. This would offer further resolution improvements by minimizing unused detector pixels. Finally, as the spectral filter exhibits a continuous response, a curved or layered pinhole lens array implementation could offer an interesting method of producing a variation in spectral and spatial resolution across the detector for a single image.

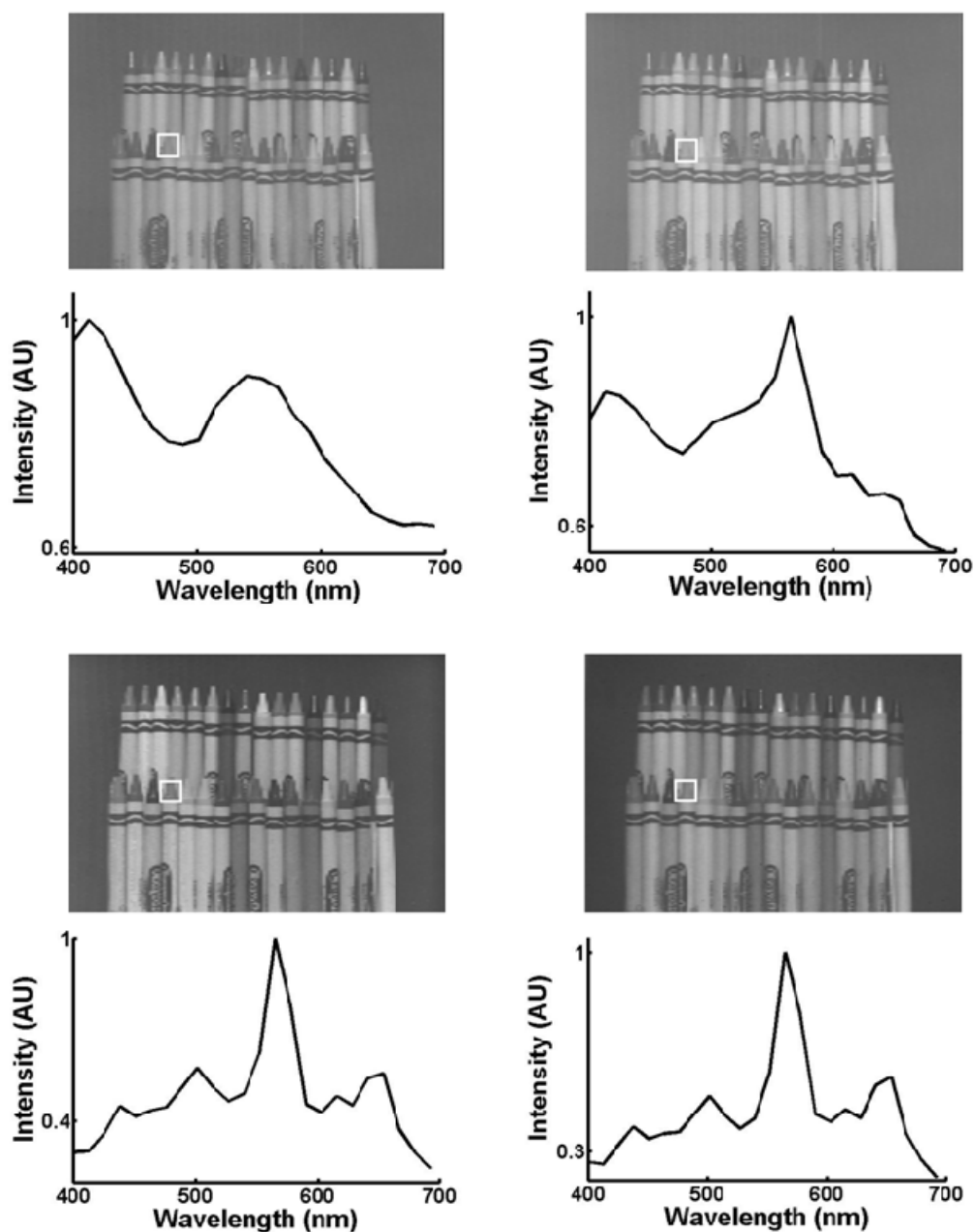


Figure 8. Four different images of a box of crayons under different types of illumination with a corresponding spectrum measurement. Clockwise from top left, each grayscale image is an integral over all 27 spectral channels for illumination from a conventional (incandescent) desk lamp, from the desk lamp and florescent room lights, from a florescent desk lamp and florescent room lights, and from a florescent desk lamp. The spectrum plot for a single pixel on a crayon (highlighted with a white box) is displayed below each image. Note how the spectrum for a single point on a crayon changes dramatically with different illumination sources, with little change in the appearance of each grayscale image.

REFERENCES

- [1] G. Lippmann. La photographie integrale. C.R. Acad. Sci. 146: 446-451 (1908).
- [2] H. E. Ives. Parallax Panoramagram. United States Patent 1918705 (1930).
- [3] E. H. Adelson and J. Y. A. Wang, "Single lens stereo with a plenoptic camera," *IEEE Trans. Pattern Anal. Machine Intell.* **14**, 99-106 (1992).
- [4] A. Isaksen, L. McMillan, and S.J. Gortler, "Dynamically reparameterized light fields," *ACM Trans. Graphics Proc. SIGGRAPH*, 297-306 (2000).
- [5] R. Ng, M. Levoy, M. Bredif, G. Duval, M. Horowitz, and P Hanrahan, "Light field photography with a hand-held plenoptic camera," Stanford Tech Report CTSR 2005-02 (2005).
- [6] T. Georgiev and C. Intwala, "Light field camera design for integral view photography," Adobe Technical Report (2006).
- [7] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan and J. Tumblin, "Dappled photography: mask enhanced cameras for heterodyned light fields and coded aperture refocusing," *ACM Trans. Graphics Proc. SIGGRAPH* **26**, 1-12 (2007).
- [8] R. Raskar, A. Agrawal, C.A. Wilson, and A. Veeraraghavan, "Glare aware photography: 4D ray sampling for reducing glare effects of camera lenses," *ACM Trans. Graphics Proc. SIGGRAPH* **27**, 56 (2008).
- [9] F. N. Lanchester. English Patent No. 16548/95 (1895).
- [10] R. E. Liesegang. *British Journal of Photography* Vol. 43: 569 (1896).
- [11] J. A. C. Branfill. *British Journal of Photography* Vol. 44: 142 (1897).
- [12] R. Berthon. English Patent No. 10611/09 (1909).
- [13] J. S. Friedman. *History of Color Photography*. Read Books: 222-250 (2007).
- [14] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," *ACM Trans. Graphics Proc. SIGGRAPH* **25**, 924-934 (2006).
- [15] R. Horstmeyer, G. Euliss, R. Athale, and M. Levoy, "Flexible multimodal camera using a light field architecture," *International Conference on Computational Photography* (2009).
- [16] G. A. Smith. Kinematograph apparatus for the production of color pictures. United States Patent 941960, 1909.
- [17] R. W. Basedow, D. C. Carmer, and M. E. Anderson, "HYDICE system: implementation and performance," *Proc. SPIE* **2480**, 258-267 (1995).
- [18] R. J. Plemmons, S. Prasad, S. Matthews, M. Mirotznic, R. Barnard, G. Gray, V. P. Pauca, T. C. Torgersen, J. van der Gracht, and G. Behrmann, "PERIODIC: Integrated Computational Array Imaging Technology," *Adaptive Optics OSA Technical Digest, CMA1* (2007).
- [19] R. Shogenji, Y. Kitamura, K. Yamada, S. Miyatake, J. Tanida, "Multispectral imaging using compact compound optics," *Opt. Express* **12**, 1643-1655 (2004).
- [20] M. E. Gehm, R. John, D. J. Brady, R. M. Willett, and T. J. Schulz, "Single-shot compressive spectral imaging with a dual-disperser architecture," *Opt. Express* **15**, 14013-14027 (2007).
- [21] Nathan Hagen and Eustace L. Dereniak, "Analysis of computed tomographic imaging spectrometers. I. Spatial and spectral resolution," *Appl. Opt.* **47**, F85-F95 (2008).
- [22] M. Young, "Pinhole Optics," *Appl. Opt.* **10**, 2763-2767 (1971).