A diffractive multispectral image sensor with on- and off-die signal processing and on-die optics in 0.18 micron CMOS

Christopher Thomas, Richard Hornsey
Dept. of Computer Science and Engineering, York University, Toronto, Canada

ABSTRACT
On-die optics have been proposed for imaging, spectral analysis, and communications applications. These systems typically require extra process steps to fabricate on-die optics. Fabrication of diffractive optics using the metal layers in commercial CMOS processes circumvents this requirement, but produces optical elements with poor imaging behavior. This paper discusses the application of Wiener filtering to reconstruction of images suffering from blurring and chromatic aberration, and to identification of the position and wavelength of point sources. Adaptation of this approach to analog and digital FIR implementations are discussed, and the design of a multispectral imaging sensor using analog FIR filtering is presented. Simulations indicate that off-die post-processing can determine point source wavelength to within 5% and position to within $\pm 0.05$ radian, and resolve features 0.4 radian in size in images illuminated by white light. The analog hardware implementation is simulated to resolve features 0.4 radian in size illuminated by monochromatic light, and 0.7 radian with white light.

Keywords: CMOS image sensors, microlenses, diffractive optics, image processing

1. INTRODUCTION
On-die optics have been used as sensor components for both imaging and non-imaging applications. Of particular interest for distributed sensing applications are systems where the entire optics and sensor package is integrated on-die. Fabrication of microlenses and high-quality diffraction gratings on-die requires extra, expensive process steps or post-processing. In a previous work, the authors presented an approach for producing diffractive optics in a commercial CMOS process by using high-level metal layers to form diffraction gratings to allow fabrication of on-die optics cheaply. A diagram of a typical sensor pixel of this type is shown in Figure 1. The devices fabricated consist of metal diffraction gratings over an array of CMOS active pixel sensors, with one grating per pixel.

Figure 1. Diffractive optics over a CMOS active pixel sensor.

In the previous work, the zone plate and pinhole gratings fabricated were found to produce large amounts of chromatic aberration and very blurry images. This work attempts to improve image quality and spectral resolving capability by applying deconvolution filtering to remove blurring and to distinguish between image contributions produced by different wavelengths of incident light.
In Section 2, the mathematical model of the imaging system is developed. The sensed image is represented as a convolved image of the scene with noise resulting from non-ideal optics and from electrical noise within the sensor. The high-level design of a fabricated sensor is also presented. In Section 3.1, Wiener filtering is applied to perform the deconvolution of simulated images. Optimization of the Wiener filter is discussed, and spatial and spectral resolution of the resulting system are assessed. Section 3.2 describes a finite impulse response (FIR) filter based on the Wiener filter that is intended to be more practical to apply in an embedded system. Section 3.3 describes an even simpler filter that is intended to be used on-die to perform approximate deconvolution of pinhole-imaged scenes. Simulated deconvolved sensor performance and conclusions are presented in Section 4.

2. SYSTEM MODEL

2.1. System Architecture

A model of the sensor system is shown in Figure 2. A scene $S$ is transformed by the optics and sensing elements into an image $I$. Post-processing is then performed to produce an image $I'$ which more closely matches the contents of scene $S$. This is done by constructing a filter which applies a transformation that is approximately the inverse of the transformation $S \rightarrow I$.

In the general case, the image $I$ can be modelled as the sum of the point spread functions produced by all of the points in $S$. This is done by constructing a filter which applies a transformation that is approximately the inverse of the transformation $S' \rightarrow I$.

To simplify the problem, the point spread functions for various regions in the scene can be approximated as displaced versions of the point spread function $P_{0,\lambda}$ of a point along the optical axis. This is represented in Equation 2, where $T(\vec{\theta})$ is a function mapping scene coordinates to image plane coordinates. The resulting image transformation is shown in Equations 3 and 4.

$$P_{\vec{\theta},\lambda}(\vec{x}) \approx P_{0,\lambda}(\vec{x} + T(\vec{\theta}))$$  \hspace{1cm} (2)

$$I(\vec{x}) = \int_{\vec{\theta}} \int_{\lambda} S_{\lambda}(\vec{\theta}) P_{\vec{\theta},\lambda}(\vec{x}) d\lambda d\vec{\theta}$$  \hspace{1cm} (3)

$$I(\vec{x}) = \int_{\lambda} \int_{\vec{y}} S_{\lambda}(T^{-1}(\vec{y})) P_{0,\lambda}(\vec{x} - \vec{y}) d\vec{y} d\lambda$$  \hspace{1cm} (4)

If a transformed version $S'$ of the scene is defined as shown in Equation 5, the image transformation of Equation 4 becomes a convolution, resulting in Equation 6:

$$S'_{\lambda}(\vec{y}) = S_{\lambda}(T^{-1}(\vec{y}))$$  \hspace{1cm} (5)

$$I(\vec{x}) = C_T \int_{\lambda} S'_{\lambda} * P_{0,\lambda} d\lambda$$  \hspace{1cm} (6)
Figure 3. a) CMOS active pixel sensor with sample-and-hold. b) CMOS active pixel sensor with deconvolution circuitry.

Restoring the original image becomes a relatively straightforward process of deconvolution, followed by transforming image coordinates to scene coordinates. The focus of this paper is on implementations of the deconvolution step.

2.2. Prototype Implementation

A demonstration implementation of an on-die diffractive optics system was designed and fabricated. This system uses several arrays of Fresnel zone plates, each tuned to a specific wavelength, to implement microlens arrays over arrays of CMOS active pixel sensors. An array of pinholes was also fabricated. The arrays are large enough (32x32) to attempt deconvolution operations on. One array has in-pixel deconvolution circuitry.

The CMOS active pixel sensor architecture used is shown in Figure 3a). Current through a photodiode is integrated on a capacitive node and sampled. Both raw and sampled photovoltages are buffered and transmitted to readout busses. The sampled photovoltage is considered to be the sensor’s output signal.

The pixels with in-pixel deconvolution circuitry are designed as shown in Figure 3b). These devices amplify the difference between neighbouring pixels, which has the effect of partially deconvolving the “top-hat” profiles produced by pinhole optics. This deconvolution approach is discussed further in Section 3.3.

The chip implementing these devices has been fabricated and is undergoing tests.

2.3. System Simulation

The method used to numerically simulate the behavior of the optical system is the same as that used for previously by the authors. Incident light is treated as the sum of several plane waves, each of which is made by the optical element before producing an interference pattern on the sensor, as shown in Figure 4. The final image is the sum of these interference patterns. The optical elements are modelled as being two-dimensional structures that either perfectly absorb or perfectly transmit light. Thickness of the diffraction gratings, multi-surface reflections off of the gratings, and polarization effects due to the gratings being made of metal are ignored. Noise due to sensor electronics is also ignored. Sample point spread functions for a Fresnel zone plate are shown in Figure 5.

2.4. Scene, Image, and Noise Spectra

The spectral contents of the scene are assumed to be unconstrained - that is, it can contain any distribution of spatial frequencies, and so should be modelled as white noise. However, the optics focusing light from the scene impose a low-pass filter on the scene’s spectrum, due to diffraction through the optics’ aperture causing unrecoverable loss of spatial information. The resulting model of the scene’s spatial frequency spectrum is a top-hat distribution with a cutoff frequency that depends on wavelength and aperture size, as shown in Equation 7.

\[
|S_\lambda(\vec{\omega})| = \begin{cases} 
1 & |\vec{\omega}| \leq \omega_{c\lambda} \\
0 & |\vec{\omega}| > \omega_{c\lambda}
\end{cases}
\] (7)

The spectrum of the transformed scene \(S'\) depends on the nature of the transforming function \(T\) used in Equations 2 through 5 in Section 2.1. If \(T\) is approximated as being a scaling function, as expressed in Equation
In both cases, the mismatch between the approximate model and reality is treated as a source of noise. 

\[ |S'_\lambda(\omega)| = \begin{cases} K_T^2 |\omega| & |\omega| \leq \left(\frac{\lambda}{\lambda_T}\right) \omega_{e\lambda} \\ 0 & |\omega| > \left(\frac{\lambda}{\lambda_T}\right) \omega_{e\lambda} \end{cases} \]  

Devising a useful closed-form expression for the spectrum of the image \( I \) requires making approximations. If the approximation of convolving by a wavelength-specific point spread function is made, as in Equation 6, the image spectrum \( I(\omega) \) can be expressed as in Equation 10. If the point spread function is approximated by a single function \( P_0 \) that does not vary with wavelength, the image spectrum described in Equation 11 is derived.

\[ I(\omega) \approx C_T \int_\Lambda S'_\lambda(\omega)P_0(\omega) \, d\lambda \]  

\[ I(\omega) \approx C_T P_0(\omega) \int_\Lambda S'_\lambda(\omega) \, d\lambda \]  

Noise contributions to the image spectrum come from two sources: electrical and optical. Electrical noise for the CMOS active pixel sensor system implementation described in Section 2.2 is dominated by fixed pattern noise and by shot noise. Fixed pattern noise and related gain variation can be greatly attenuated by correlated

---

**Figure 4.** Simulated image construction process.

**Figure 5.** Simulated point spread functions for a Fresnel zone plate, for a) on-axis, in-focus, b) off-axis, and c) out-of-focus cases.
double-sampling and by calibration performed by the host system prior to sensing.\textsuperscript{10,11} The remaining dominant electrical noise source is shot noise.

Shot noise results from statistical variation in the number of electrons present in a flow of current, or in the number of photons arriving at a photosensor. For purposes of the CMOS active pixel sensor and sample-and-hold circuit implemented for the diffractive sensing system, shot noise manifests as uncertainty in the number of electrons present in the photodiode’s parasitic capacitance and in the storage capacitors of the sample-and-hold circuit. The capacitance of each of these nodes is on the order of 10 fF, allowing the charge and number of stored electrons per node to be computed as in Equations 12 through 14. The standard deviation in the number of electrons is equal to the square root of the number of electrons, which allows calculation of the magnitude of the shot noise per node as shown in Equations 15 through 17.

\[ n_e = \frac{CV}{q_e} \]  
\[ n_e = \frac{1 \times 10^{-14} \cdot V}{1.6 \times 10^{-19}} \]  
\[ n_e = 6.3 \times 10^4 \cdot V \]  
\[ \sigma_V = \frac{\sigma_n q_e}{C} \]  
\[ \sigma_V = \frac{\sqrt{n_e} \cdot 1.6 \times 10^{-19}}{1 \times 10^{-14}} \]  
\[ \sigma_V = 4 \times 10^{-3} \sqrt{V} \]  

Two capacitive storage nodes occur in series, increasing the standard deviation by a factor of \( \sqrt{2} \). The voltage on the storage nodes is typically between 0.6 V and 3.3 V. This allows computation of the range of relative error as shown in Equations 18 through 19. Applying the most conservative (largest) error value to the spectrum \( S'_\lambda(\vec{\omega}) \) produces an approximate upper bound estimate to the shot noise, shown in Equation 20.

\[ \frac{4 \times 10^{-3}}{3.3} < \frac{\sigma_V}{V} < \frac{4 \times 10^{-3}}{0.6} \]  
\[ 2.2 \times 10^{-3} < \frac{\sigma_V}{V} < 5.2 \times 10^{-3} \]  
\[ N_{\text{sh}}(\vec{\omega}) \approx 5 \times 10^{-3} S'(\vec{\omega}) \]  

Optical noise for the sensor system implementation has two components. The first component is deviation of the point spread function \( P_{\vec{\theta},\lambda} \) from its approximation as a displaced version of \( P_{0,\lambda} \) or of \( P_0 \). This modelled by taking the magnitudes of the spectra of the point spread functions, to remove translation (which alters only the phase), and then by finding the standard deviation of these spectra magnitudes with respect to whichever point spread function is used as an approximation. For the wavelength-dependent approximation used in Equation 10, the resulting optical noise model is shown in Equation 21, with the noise contributions for each wavelength averaged. For the wavelength-independent approximation used in Equation 11, the resulting optical noise model is shown in Equation 22.

\[ N_{\text{op}}(\vec{\omega}) \approx \left( \frac{1}{|\Lambda|} \right) \int_\Lambda \left[ \left( \frac{1}{|\vec{\theta}|} \right) \int_{\vec{\theta}} |P_{\vec{\theta},\lambda}(\vec{\omega}) - P_{0,\lambda}(\vec{\omega})|^2 \ d\vec{\theta} \right] d\lambda \]  
\[ N_{\text{op}}(\vec{\omega}) \approx \sqrt{ \left( \frac{1}{|\vec{\theta}||\Lambda|} \right) \int_{\vec{\theta}} \int_\Lambda (|P_{\vec{\theta},\lambda}(\vec{\omega})| - |P_0(\vec{\omega})|)^2 \ d\lambda \ d\vec{\theta} } \]  

The second component of optical noise is nonlinearity in the function \( T(\vec{\theta}) \) which maps scene coordinates to image coordinates. This function is approximated as a scalar multiplication in Equation 8. For purposes of this paper, the noise introduced by this approximation is assumed to be negligible compared to the noise introduced by point spread function distortion. This approximation is expected to be accurate for small values of \( \vec{\theta} \). Development of a more accurate model applicable over wider ranges of \( \vec{\theta} \) is a topic for future work.
2.5. Performance Evaluation

The primary figure of merit for the diffractive sensor’s behavior as an imaging device is its modulation transfer function (MTF). This is defined as the magnitude of the reconstructed scene’s frequency spectrum divided by the original scene’s frequency spectrum. For purposes of the model developed in Section 2.1, a transformed modulation transfer function \( MTF'(\bar{\omega}) \) is constructed, defined as the reconstructed image’s frequency spectrum divided by the transformed scene’s frequency spectrum, as shown in Equation 23.

\[
MTF'(\bar{\omega}) = \frac{I(\bar{\omega})}{S'(\bar{\omega})} \tag{23}
\]

If the approximation of \( T(\bar{\theta}) \) as a scalar multiplication is used, as given in Equation 8, then the MTF can be recovered via Equation 24.

\[
MTF(\bar{\omega}) = \left( \frac{1}{K_T} \right)^2 MTF' \left( \frac{\bar{\omega}}{K_T} \right) \tag{24}
\]

A derived figure of merit from the MTF is the estimated maximum angular resolution of the sensor system. This can be estimated as the angular separation corresponding to one period of the largest \( \bar{\omega} \) for which \( MTF(\bar{\omega}) \) is greater than some threshold value. Setting the threshold equal to 0.5 for a MTF normalized to have unity maximum produces the expected \( \bar{\omega}_{3dB} \) for radially symmetrical MTFs that do not contain notches (in other words, that are low pass filters). Deconvolution is expected to produce notches but not to break radial symmetry for radially symmetrical point spread functions.

The figure of merit chosen for assessing the sensor’s behavior when resolving spectra is derived from the modulation transfer function sampled for multiple wavelengths. The MTF at various spatial frequencies is plotted as a function of illumination wavelength. If spatial frequencies are chosen so that the MTF is near-unity for some wavelength at these spatial frequencies, the plots against wavelength are expected to form peaks, representing wavelength regions where the deconvolution filter produces images with high spatial resolution. The spectral resolution of the sensor is estimated as being the diameter of selected well-defined peaks, measured from the points at which the peaks cross some fraction of their maximum value (typically 0.5, as with the estimation of maximum angular resolution of the sensor).

3. FILTER IMPLEMENTATIONS

3.1. Wiener Deconvolution Filter

When the spectral characteristics of both the expected image and the system noise are known, Wiener filtering is an optimal approach to performing deconvolution. The standard form of the filter is shown in the left part of Equation 25, where \( W(\bar{\omega}) \) is the Wiener deconvolution filter, \( S'(\bar{\omega}) \) is the transformed scene’s spatial frequency spectrum, \( N_s(\bar{\omega}) \) is the spatial frequency spectrum of the noise in the transformed scene, and \( \bar{P}(\bar{\omega}) \) is the spectrum of the point spread function chosen to be representative of \( P_{\bar{\theta},\lambda} \). This produces least mean squared error when applied to the system described by the right part of Equation 25, where \( I(\bar{\omega}) \) is the image’s spatial frequency spectrum, with (scene) noise \( N_s(\bar{\omega}) \) applied before convolution. By contrast, the system described in Section 2.1 applies image noise \( N_i(\bar{\omega}) \) after deconvolution, as in Equation 26. This results in a Wiener filter having the form shown in Equation 26, which will be used for the remainder of this analysis.

\[
W(\bar{\omega}) = \left( \frac{|\bar{P}(\bar{\omega})|}{|\bar{P}(\bar{\omega})|^2 + \frac{1}{1 + \frac{N_s(\bar{\omega})}{S'(\bar{\omega})}}} \right) \quad I(\bar{\omega}) = \bar{P}(\bar{\omega}) \left( S'(\bar{\omega}) + N_s(\bar{\omega}) \right) \tag{25}
\]

\[
W(\bar{\omega}) = \left( \frac{|\bar{P}(\bar{\omega})|}{\bar{P}(\bar{\omega}) + \frac{N_i(\bar{\omega})}{S'(\bar{\omega})}} \right) \quad I(\bar{\omega}) = \bar{P}(\bar{\omega})S'(\bar{\omega}) + N_i(\bar{\omega}) \tag{26}
\]

As described in Section 2.4, the scene is modelled as white noise modulated by low-pass filter imposed by diffraction through the sensor’s aperture (zone plate or pinhole). The definition chosen for this spectrum is shown in Equation 9 in that section.
The spectrum \( \hat{P}(\vec{\omega}) \) of the point spread function is not modelled as a closed-form expression, but is instead evaluated by simulation using the method presented in Section 2.3. Furthermore, as the point spread function varies with both point location and wavelength, any value chosen for \( \hat{P}(\vec{\omega}) \) is an approximation. Two approaches to approximation are used in this paper. The first, shown in Equation 27, is to choose some representative wavelength \( \lambda_0 \) and scene coordinate \( \vec{b}_0 \), and to approximate all point spread functions by using the PSF for that location. The second, shown in Equation 28, is to approximate the point spread function using the average of the PSFs over the domain of applicability of the filter (denoted by \( \Theta_g \) and \( \Lambda_g \)).

\[
\hat{P}(\vec{\omega}) = P_{\vec{b}_0,\lambda_0}(\vec{\omega})
\]

\[
\hat{P}(\vec{\omega}) = \left( \frac{1}{|\Theta_g||\Lambda_g|} \right) \int_{\Theta_g} \int_{\Lambda_g} P_{\vec{b},\lambda}(\vec{\omega}) \, d\lambda \, d\vec{\theta}
\]  

Noise in the scene comes from several sources. Equation 29 represents the noise as the sum of electronics noise \( (N_e) \), variation in “good” (accepted) point spread functions \( (N_g) \), and contributions from “bad” (rejected) point spread functions \( (N_b) \).

\[
N(\vec{\omega}) = N_e(\vec{\omega}) + N_g(\vec{\omega}) + N_b(\vec{\omega})
\]

Electronics noise \( (N_e) \) is described in Section 2.4, and is modelled as being dominated by shot noise \( (N_{sh}) \) represented by the spectrum described by Equation 20. Noise corresponding to the variation in “good” point spread functions \( (N_g) \) is defined as the standard deviation of the magnitudes of all point spread functions within the domains of \( \Theta_g \) and \( \Lambda_g \), as shown in Equation 30. This is similar to the optical noise computation performed in Section 2.4, and has the effect of suppressing portions of the spectrum of \( \hat{P} \) that have high variability within the domain of accepted PSFs. The noise spectrum corresponding to contributions from “bad” point spread functions is described by Equation 32. This is the average magnitude of the spectra of point spread functions outside of the domain of acceptance, but within the domain of scene coordinates \( (\vec{\Theta}) \) and wavelengths \( (\Lambda) \) that contribute to the image recorded by the sensor. This domain of rejection is defined by Equation 31, and the average over this domain is then weighted by the relative sizes of the "accepted" and "rejected" PSF domains to normalize it with respect to \( N_g \).

\[
N_g(\vec{\omega}) = \left( \frac{1}{|\Theta_g||\Lambda_g|} \right) \int_{\Theta_g} \int_{\Lambda_g} |P_{\vec{b},\lambda}(\vec{\omega}) - \hat{P}(\vec{\omega})|^2 \, d\lambda \, d\vec{\theta}
\]

\[
\vec{\Theta}_b = \vec{\Theta} - \vec{\Theta}_g \quad \vec{\Lambda}_b = \vec{\Lambda} - \vec{\Lambda}_g
\]

\[
N_b(\vec{\omega}) = \left( \frac{1}{|\Theta_g||\Lambda_g|} \right) \int_{\vec{\Theta}_b} \int_{\vec{\Lambda}_b} |P_{\vec{b},\lambda}(\vec{\omega})| \, d\lambda \, d\vec{\theta}
\]

Sample point spread functions \( \hat{P}(\vec{x}) \) can be found in Figures 6a (representative) and 6d (averaged). With respect to the averaged \( \hat{P}(\vec{x}) \), noise due to accepted \( (N_g(\vec{x})) \) and rejected \( (N_b(\vec{x})) \) point spread functions can be found in Figures 6b and 6e, respectively. The Wiener deconvolution filter optimized for this averaged point spread function is shown in Figure 6c, with its spatial frequency spectrum shown in Figure 6f.

### 3.2. Finite Impulse Response Deconvolution Filter

The Wiener filtering described in Section 3.1 can be performed in \( O(n \log n) \) time with respect to the number of pixels in the image. However, as it is a frequency-domain approach, it requires performing a Fourier transform of the entire image. This may not be practical for embedded applications. An alternative and widely-used approach is to perform finite impulse response (FIR) filtering in the spatial domain to implement an approximation of the desired filter. The FIR version of the Wiener filter of Section 3.1, \( W_{FIR}(\vec{x}) \), is constructed from the full space-domain version of the filter \( W(\vec{x}) \) by setting all elements beyond some radius \( r_{FIR} \) from the filter’s centroid to zero, as shown in Equation 33. As the truncated filter \( W_{FIR}(\vec{x}) \) has a small number of samples compared to the image as a whole, it can be convolved in the time domain relatively inexpensively, especially if dedicated embedded DSP hardware is available.
Figure 6. Wiener deconvolution filter example. Point spread functions for a) representative and d) averaged cases. Noise due to b) accepted and e) rejected point spread functions. Wiener deconvolution filter in c) spatial and f) frequency domains.

\[ W_{FIR}(\vec{x}) = \begin{cases} W(\vec{x}) & |\vec{x} - \vec{x}_0| \leq r_{FIR} \\ 0 & |\vec{x} - \vec{x}_0| > r_{FIR} \end{cases} \tag{33} \]

Two cases illustrating FIR filtering are shown in Figure 7. These images show reconstruction of images of a large checkerboard pattern and of a small checkerboard pattern, with identical simulated optics and illumination. For the small checkerboard pattern, a larger cutoff radius allows more information to be obtained, facilitating reconstruction. However, for the large checkerboard, the signal contribution from distant parts of the checkerboard overwhelms the spread information that the deconvolution filter attempts to collect, resulting in a further blurring of the image. FIR filter reconstruction behavior is assessed quantitatively in Section 3.4.

3.3. Differentiating Deconvolution Filter

While a finite impulse response filter can be implemented in an embedded system if convolution hardware is available, it still requires a substantial amount of dedicated digital image processing hardware. A smaller on-die system was implemented by picking as the deconvolution filter a very small pattern in the space domain that acted approximately as a differentiator. This impulse response pattern is shown in Equation 34, and was implemented in each pixel using four differential amplifiers. The frequency response of this filter is shown in Equation 35, with \( \omega \) normalized to the range \(-\pi \ldots \pi\), and can be seen to roughly approximate that of a differentiator.

\[ D_{FIR}(\vec{x}) = \begin{cases} 4 & \vec{x} = 0 \\ -1 & x_1 = \pm 1 \\ -1 & x_2 = \pm 1 \end{cases} \tag{34} \]

\[ D_{FIR}(\vec{\omega}) = 4 - 2 \cos \omega_1 - 2 \cos \omega_2 \tag{35} \]

A differentiator is useful as a deconvolution filter for sensors where the optical element is a wide pinhole. These optical elements have a point spread function that is approximated by a square pulse. The frequency spectrum of a square pulse is the sinc function, as shown in Equation 36. Applying a perfect differentiating
Figure 7. Reconstructed images of a large checkerboard (a, b, c) and a small checkerboard (e, f, g), built using FIR filters with varying cutoff radii. Two versions of the reconstruction filter used for the small checkerboards are shown in (d, h).

Filter \((D_0)\) produces the spectrum in Equation 37. While phase distortion occurs and notches are present, high-frequency features are still restored, corresponding to potentially improved angular resolution. This is evaluated quantitatively in Section 3.4.

\[
P_{0,\lambda}(\vec{\omega}) = C_{1,\lambda} \left( \frac{1}{|\vec{\omega}|} \right) \sin |\vec{\omega}| \tag{36}
\]

\[
P_{0,\lambda}(\vec{\omega})D_0(\vec{\omega}) = C_{1,\lambda} \sin |\vec{\omega}| \tag{37}
\]

3.4. Filter Performance

The primary role of an imaging sensor is imaging. The simulated imaging behavior of zone plate and pinhole sensors with monochromatically illuminated scenes is shown qualitatively in Figure 8. This is quantified in Figure 9, which estimates the modulation transfer functions of the systems under test by applying the chosen deconvolution filters to the point spread functions produced by the optical systems when illuminated on-axis. As the point spread function’s shape is approximately constant for small displacements, this provides an approximation of the modulation transfer function that is valid for the region of the image near the optical axis.

The modulation transfer functions shown in Figure 9 have several features that can be leveraged for image sensing purposes. First, the part of the MTF above the noise floor (approximately \(5 \times 10^{-3}\), from Section 2.4) can usefully be boosted by deconvolution filters. This occurs, with both the FIR filters and the differentiating filter bringing parts of the spectrum in this region above the 0.5 amplitude value. This provides a noise-limited angular resolution of approximately 0.4 radian for the zone plate (40th harmonic is above the noise floor before processing), and approximately 0.7 radian for the pinhole (25th harmonic is above the noise floor before processing). Other parts of the spectrum representing higher harmonics are also boosted above 0.5 amplitude, but they remain below the noise floor (which is itself magnified by the deconvolution). Making use of these higher harmonics requires reducing shot noise, which can be accomplished by integrating the results of multiple exposures of the sensor, or by aggregating the values sensed by many pixels, or by a combination of these.
Figure 8. Raw and deconvolved images. I) Large checkerboard via Fresnel zone plate, II) small checkerboard via zone plate, III) large checkerboard via pinhole. a) Raw images, b) deconvolved via Wiener filter (9 um cutoff radius), c) deconvolved via FIR (1 um cutoff radius), d) deconvolved via differentiator.

techniques. This would provide a resolution of approximately 0.1 radian, corresponding to the 160th spatial harmonic.

The last feature of the modulation transfer functions that is of note is the sensitivity of these functions to wavelength. This was noted as both a problem and a source of sensing potential in the chip preceding the present design. This feature can be leveraged for sensing by observing that lobes of the MTF are at different locations for different illumination wavelengths. Measuring the MTF response (either directly or indirectly via spectral analysis of the illuminated scene) provides information into the spectral components of the light source. The variation of MTF magnitude with illumination wavelength was simulated and is shown in Figure 10.

The modulation transfer functions’ harmonic amplitudes are extremely variable over wavelength, and contrary to expectations expressed in Section 2.5, there are few well-defined peaks. However, the variability - changing amplitude by an order of magnitude or more over a step of 25 nm in illumination wavelength - means that the components of the modulation transfer function that are above the noise floor can be used to assess the wavelength of a point source of light with a known point spread function to within 25 nm or better. With the large number of MTF harmonics and peaks within each harmonic, spectral decomposition of incoming light is also possible given sufficient time to average out noise in sampled images. An exploration of this phenomenon is a topic for future study.
Figure 9. Modulation transfer functions for a) Fresnel zone plates and b) pinholes. I) Raw transfer functions without post-processing, II) Transfer functions after processing with a 1µm radius FIR filter, III) Transfer functions after processing with the on-die differentiator. The noise floor for the unprocessed transfer functions is at $5 \times 10^{-3}$ amplitude.

Figure 10. Dependence of the modulation transfer function on wavelength, plotted for selected spatial harmonics. a) Zone plate with large-image FIR, b) Zone plate with small-image FIR, c) Pinhole.
4. CONCLUSIONS

In conclusion, an on-die multispectral imaging system has been described, and an approach to improving the spatial and spectral resolving ability of this system has been presented. Deconvolution filters suitable for implementation on a host computer system, an embedded sensor system, and the imager die itself were evaluated. Sensor systems using these filters obtain a spatial resolution of approximately 0.1 radian under ideal conditions, or approximately 0.4 radian in situations involving polychromatic illumination, sensor noise, and scene clutter. The spectral response of the sensors simulated did not match expectations. The reasons for this are under investigation. However, the sensor geometries studied nevertheless proved capable of resolving spectral features of point sources to within 25 nm or better under low-noise conditions.

REFERENCES