Fourier Spectral Filter Array For Optimal Multispectral Imaging

Jie Jia, Student Member, IEEE, Kenneth Barnard, Senior Member, IEEE, and Keigo Hirakawa, Senior Member, IEEE

Abstract—Limitations to existing multispectral imaging modalities include speed, cost, range, spatial resolution, and application-specific system designs that lack versatility of the hyperspectral imaging modalities. In this article, we propose a novel general-purpose single-shot passive multispectral imaging modality. Central to this design is a new type of spectral filter array (SFA) based not on the notion of spatially multiplexing narrowband filters, but instead aimed at enabling single-shot Fourier transform spectroscopy. We refer to this new SFA pattern as “Fourier SFA” and we prove that this design solves the problem of optimally sampling the hyperspectral image data.

Index Terms—Demosaicking, hyperspectral imaging, multispectral imaging, spectral filter array.

I. INTRODUCTION

A spectrum of electromagnetic waveform is useful for making inferences on material or chemical properties of the objects. Spectroscopy enables chemical and material identification, provides minimally-invasive alternative to biopsy, automates geophysical surveying, and boosts the accuracy of object detection/recognition tasks. These discriminating features are not available to conventional colorimetric sensors that only capture metameric color information (i.e. RGB).

By hyperspectral imaging, we refer to the dense sampling over a specified wavelength/wavenumber range measured at densely sampled two dimensional spatial locations as pixels. Such data is useful for making complete spectral characterization of objects in a complex scene. Spatial samples also provide contextual and positional information that may further strengthen the evidence of detected objects or enable higher level tasks such as recognition and tracking. Hyperspectral imaging is particularly beneficial to applications such as remote sensing, where the lack of spatial resolution may limit the ability to recognize objects by their shapes.

Owing to high hardware complexity, however, it is difficult to attain high spectral and spatial resolution in practice. As described in Section III, conventional hardware use temporal or spatial multiplexing which is expensive and slow. Such configuration precludes imaging of dynamic scenes containing moving objects, or scenarios where the image sensor is not stationary. A handful of alternative solutions enabling single-shot hyperspectral imaging capabilities have been proposed in the past. For example, CTIS (computed tomography imaging spectrometer) make joint spatial-spectral measurements with the help of dispersive optical elements; complete spatial-spectral data is reconstructed in postprocessing [3]. The usefulness of CTIS depends on its application—its minimum and maximum spectral resolutions occur when viewing a spatially uniform and a point source objects, respectively [4]. Another system known as CASSI (coded aperture snapshot spectral imager) use a combination of coded aperture and dispersive elements to spatially and spectrally modulate the spectral image onto a two dimensional detector surface [5]. However, false spectral features are known to be introduced into the computationally reconstructed spectra that may hamper target detection [6]. CASSI’s multishot variant improves spectral recovery, though it comes at the sacrifice of the acquisition speed [7]. A more recent treatment of CASSI was reported in [8]. Similarly, spatial occlusion masks in PMVIS (prism-mask multispectral video imaging system) on a prism perform spatial-spectral modulation of the spectral image, which is recorded on a two dimensional detector [9]. The light efficiency of PMVIS is poor due to the occlusion mask, and this system has an inherent tradeoff between spatial and spectral resolution. The occlusion mask must be located close to the scene, also [9]. Although recent advances in PMVIS have addressed these issues to an extent [10], [11], these solutions require large depth of field to reduce optical distortions and blurring stemming from prism and occlusion masks.

Multispectral imaging is an alternative sensing modality to hyperspectral imaging with a compromise—N predetermined filters make finite spectral measurements instead of the spectral continuum—in order to increase spatial and temporal resolutions. The motivation is that greater-than-three spectral samples would provide an advantage over color imaging in terms of material and chemical identification, but sacrifices in spectral resolution would give advantage over hyperspectral imaging in terms of cost, speed, and in spatial resolution [12]-[14]. In the visible range, multispectral imaging is seen as a modality that bridges the gap in the three-color imaging and hyperspectral imaging.

Owing to the fact that only a few spectral samples at specific wavelengths are needed to make inferences on material properties, the traditional thinking behind multispectral imaging systems focuses on determining the spectral characteristics of narrowband filters optimized for specific target detection/recognition tasks [15]-[17]. However, whether a particular wavelength is useful is task- and material-dependent, meaning that it is difficult to choose a small number of wavelengths

J. Jia and K. Hirakawa are with the Department of Electrical and Computer Engineering, University of Dayton, Dayton, OH, 45469 USA (e-mail: {jiaj01, khirakawa}@udayton.edu). K. Barnard is with the Sensor Directorate, Air Force Research Laboratory, Dayton, OH, 45433 USA (e-mail: {kenneth.barnard}@us.af.mil). The preliminary results of this work appeared in [1], [2].

Manuscript received September ??, 2014.
provenably useful for detecting/recognizing a wide range of materials. In light of this fact, a general-purpose multispectral sensing modality is yet to emerge as a cost-effective alternative to hyperspectral imaging or as an alternative to application-specific multispectral imaging that is more versatile.

Nevertheless, appealing to the successes of the color filter array (CFA) in conventional cameras [18], [19], spectral filter arrays (SFA) have gained some attention as a relatively new form of multispectral imaging modality [14]–[17], [20]–[26]. At the sacrifice of spatial resolution, $N$ predetermined narrowband spectral filters are spatially multiplexed over the pixel array, meaning only one spectral sample is available at each pixel location. Subsequent interpolation step (called demosaicking) recovers the spectral image from the SFA subsampled sensor data. This hardware configuration enables single-shot image capture (i.e. without temporal multiplexing), which allows the image sensor to operate without requiring that the sensor remain stationary (i.e. mounted on a tripod or a gimbal) and is necessary for capturing scenes containing moving objects.

Considerable challenges still remain for narrowing the gap between the conceptual design of SFA and the reality, however. In particular, the spatial-spectral undersampling cause severe aliasing that has adverse effects on hyperspectral image representations. Aliasing limits the usefulness of the acquired data—not simply by reducing the spatial and spectral resolutions, but also by introducing spatial-spectral registration ambiguity in the measurements that make it impossible to tell apart image details such as the smooth and sharp transitions; or to disambiguate peaks in the spectrum from their frequency folded versions. Owing to the limitations of the narrowband filters, the SFA configuration also does not preclude the imaging system from being application specific—efforts thus far to broaden the SFA applications have been largely ad-hoc.

Proposed herein this article is a new type of SFA based not on the notion spatially multiplexing narrowband filters, but instead aimed at enabling single-shot Fourier transform spectroscopy—we refer to this new SFA pattern as “Fourier SFA.” Specifically, we first propose the notion of multispectral imaging based on Fourier transform spectroscopy, and prove that this broadband filtering approach is free of spectral aliasing contaminations even when spectrally undersampled. We then develop the Fourier SFA by replacing the notion of spatial multiplexing in traditional SFA designs with spatial frequency multiplexing, and prove that the proposed design maximize the spatial resolution of the resultant multispectral imaging sensor by minimizing the risks of spatial aliasing. In other words, Fourier SFA emerges as a logical and straightforward solution to the multispectral image acquisition task when interpreting multispectral imaging as the optimal spatial-spectral sampling problem of hyperspectral data.

Measurements taken with a multispectral imaging apparatus devised in this manner form a representation of the spatial-spectral image signal in its entirety—rather than measuring signal centered around a particular wavelength as existing multispectral imaging systems are designed to do—increasing the likelihood of detecting, tracking, discriminating, and recognizing multiple objects based on their spectral properties.

As such, the Fourier SFA represents a concrete step towards application-agnostic multispectral imaging with the versatility and generality of the hyperspectral imaging system.

**II. Spatial-Spectral Model**

We introduce a form of Fourier analysis for a hyperspectral image signal capable of representing diverse classes of scenes and objects. Using a “collection of Fourier transforms” as illustrated in Figure 1, our model gives rise to a precise the notion of joint spatial-spectral bandwidth and provides a foundation for rigorously analyzing the “resolution” of a sensor making measurements on a spectral image signal. This new analytical framework offers a principled perspective on the sacrifices made by the multispectral imaging modalities, and suggests ways to explore and exploit inefficiencies in existing image acquisition hardware.

**A. Radiometry and Definitions**

For a detector at the image plane of an optical system, the measured flux is found by integrating the spectral radiance weighted by the spectral response of the system and scaled by radiometry constants. $\bar{X}: \mathbb{R} \rightarrow \mathbb{R}^+$ is the spectral radiance at the sensor, $\lambda \in \mathbb{R}$ is the wavelength of the light. Considering just a single on-axis detector element, the measured flux $Y \in \mathbb{R}^+$ can be written as

$$Y = \frac{\pi AT}{4(M + 1)^2(F)^2 + 1} \int_{\mathbb{R}} \bar{\varphi}(\lambda)\bar{\eta}(\lambda)\bar{X}(\lambda) d\lambda,$$

where $A \in \mathbb{R}^+$ is the area of the detector, $T \in \mathbb{R}^+$ is the detector integration time, $M \in \mathbb{R}^+$ is the magnification of the imaging system, $F \in \mathbb{R}^+$ is the infinite conjugates f-number of the optical system, $\bar{\varphi} : \mathbb{R} \rightarrow [0, 1]$ is the spectral transmission of the optics, and $\bar{\eta} : \mathbb{R} \rightarrow [0, 1]$ is the quantum efficiency of the detector. Note that the spectral dependencies due to the atmospheric transmittance and path radiance are
contained within the spectral radiance term and may include strong absorption bands depending on the wavelength range. For off-axis detector elements, the measured flux is weighted by \(\cos^4(\theta)\), the angle between the detector surface normal and the principal ray to the detector. Integrating with respect to wavelength gives \(Y\) in electrons.

Combining spectrally dependent terms in the integrand and ignoring the scaling factor, the signal at the detector can be represented as an inner product:

\[
Y = \langle \bar{Q}, \bar{X} \rangle_\lambda := \int \bar{Q}(\lambda) \bar{X}(\lambda) d\lambda,
\]

where the spectral transmittance and quantum efficiency terms have been combined into \(\bar{Q}(\lambda) : \mathbb{R} \to [0, 1]\) having a support generally defined by the quantum efficiency.

Equation (2) provides a measure of an optical signal for a broadband spectral band imaging system. Introducing an additional spectral filter \(\bar{S} : \mathbb{R} \to [0, 1]\) in the optical path allows flexibility to alter the spectral response of the system:

\[
Y = \langle \bar{S}, \bar{Q}, \bar{X} \rangle_\lambda := \int \bar{S}(\lambda) \bar{Q}(\lambda) \bar{X}(\lambda) d\lambda,
\]

where \(\bar{S}(\lambda)\) is the filter transmittance as a function of wavelength. The quantum efficiency of a photon detector is limited to a certain spectral range determined by detector materials, meaning \(\bar{S}(\lambda)\) needs precise specification only over a finite range \(\pm\lambda \in (\lambda_{\text{min}}, \lambda_{\text{max}})\). For measuring a narrowband spectral sample, for example, \(\bar{S}(\lambda)\) is supported over a range \(\pm\lambda \in \{\lambda_i + (-\epsilon, \epsilon)\}\) where \(\lambda_i\) is the \(i\)th center band wavelength and \(2\epsilon\) is the passband width.

It is also common to represent spectral radiance in terms of wavenumber \(\sigma = 1/\lambda \in \mathbb{R}\), especially in Fourier transform spectrometers. The mappings \(X \mapsto \bar{X}, \bar{Q} \mapsto \bar{Q}, \bar{S} \mapsto \bar{S}\) between wavelength representations \(X, \bar{Q}, \bar{S}\) and the corresponding wavenumber representations \(\bar{X}, \bar{Q}, \bar{S}\) take the form

\[
\bar{X}(\sigma) = \frac{X(1/\sigma)}{\sigma^2}, \quad \bar{Q}(\sigma) = \bar{Q}(1/\sigma), \quad \bar{S}(\sigma) = \bar{S}(1/\sigma).
\]

The spectral response \(\bar{Q}\) and filter transmittance \(\bar{S}\) are bounded in the range \([0, 1]\) due to the way the transformations \(\bar{Q} \mapsto \bar{Q}\) and \(\bar{S} \mapsto \bar{S}\) are defined. The Jacobian term \(\sigma^{-2}\) in \(\bar{X}(\sigma)\) ensures that the following integral is identical to the sensor measurement in (3):

\[
Y = \langle \bar{S}, \bar{Q}, \bar{X} \rangle_\sigma := \int \bar{S}(\sigma) \bar{Q}(\sigma) \bar{X}(\sigma) d\sigma.
\]

In practice, the detector material responds over a finite spectral range \(\pm\sigma \in (\sigma_{\text{min}}, \sigma_{\text{max}})\) only, and thus we cannot sense \(\bar{X}(\sigma)\) outside of this range. Hence in the remainder of this article, we restrict the support of \(\bar{X}(\sigma)\) to \(\pm\sigma \in (\sigma_{\text{min}}, \sigma_{\text{max}})\) without loss of generality. The spectral response of Silicon CMOS sensors is approximately 350nm to 1000nm (28,500cm\(^{-1}\) to 10,000cm\(^{-1}\)), with the visible wavelength range extending from 380nm to 740nm (26,300cm\(^{-1}\) to 13,500cm\(^{-1}\)).

The spectral radiance \(\bar{X}(\sigma)\) is interpreted as the power spectral density of the optical signal and arises in Fourier transform spectrometers where the directly measured quantity is an interferogram or autocorrelation function. Invoking the Wiener-Khinchin theorem and performing an inverse Fourier transform leads to an autocorrelation as a function of a distance, i.e. the physical optical path difference (OPD) in one arm of an interferometer, given by

\[
X(\zeta) = F^{-1}\{\bar{X}(\sigma)\} := \frac{1}{2\pi} \int \bar{X}(\sigma) e^{-j\sigma\zeta} d\sigma,
\]

where \(F^{-1}\) denotes an inverse Fourier transform with respect to \(\sigma\) and \(\zeta \in \mathbb{R}\) is the OPD. See Figure 1. The autocorrelation is measured as an interference of two optical fields at a detector, one temporally delayed in time with respect to the other (as indicated by the OPD) [27].

B. Fourier Analysis

Extending the above analysis to discrete detectors in a focal plane array where spectral radiance is defined at pixel locations \(n = (n_0, n_1) \in \mathbb{Z}^2\) leads to the notion of a hyperspectral image. We update the definitions of \(X, \bar{X}, \bar{X}, S, S\) above to include \(n\) in its domain—for example, \(\bar{X} : \mathbb{Z}^2 \times \mathbb{R} \to \mathbb{R}^+\) is a space-wavenumber signal where \(X(n, \sigma)\) is the spectral radiance associated with sensor coordinate \(n\). This is very high dimensional data, presenting a challenge for making dense measurements on this space. On the other hand, filters may be applied to all pixels \((S(\sigma)\) and \((S(\lambda))\) or to individual pixels in a form of a filter array \((\bar{S}(n, \sigma)\) and \((\bar{S}(n, \lambda))\). An ideal spectral response \(\bar{Q}\) or \(\bar{Q}\) would not depend on pixel index \(n\), since all detectors are assumed identical. Spatial nonuniformity correction in real sensors removes any spatially-dependent response as well as \(\cos^4(\theta)\) signal roll off at the edges of the field of view.

The spectral radiance representations of spatial-wavenumber \((\bar{X}(n, \sigma))\) or spatial-wavelength \((\bar{X}(n, \lambda))\) signals are the forms of hyperspectral image we are most familiar with (and we are after). However, we propose another canonical representation. Consider the space-OPD signal \(X : \mathbb{Z}^2 \times \mathbb{R} \to \mathbb{R}^+\—\)we refer to this interferometric space of pixels \(n\) and OPD \(\zeta\) as the physical domain. We define a two dimensional discrete space Fourier transform \(\bar{X}\) of \(X\) as follows:

\[
\bar{X}(\omega, \zeta) = \bar{F}\{X(n, \zeta)\} := \sum_n X(n, \zeta) e^{j\omega^T n},
\]

where \(\bar{X} : (\mathbb{R}/2\pi)^2 \times \mathbb{R} \to \mathbb{C}\) are the Fourier coefficients and \(\omega = (\omega_0, \omega_1)^T \in (\mathbb{R}/2\pi)^2\) is the two dimensional spatial
frequency. Consider the composite mapping $\hat{X} \mapsto \hat{X} = \hat{X} \mapsto X \mapsto \tilde{X}$ defined by the relation

$$\hat{X}(\omega, \zeta) = \hat{F} \circ \hat{F}^{-1}(\hat{X}(n, \sigma))$$

$$:= \frac{1}{2\pi} \sum_{n} \int_{\mathbb{R}} \hat{X}(n, \sigma)e^{i\omega n + i\sigma d\sigma}. \quad (8)$$

As illustrated in Figure 1, (8) is a type of three dimensional Fourier transform, in the sense that the inverse Fourier transform is symmetrical to Fourier transform. It is also a mixture of continuous and discrete time Fourier transforms, which is unusual. To highlight a few symmetries of $\hat{X}(n, \sigma)$, $X(n, \sigma)$, $\hat{X}(\omega, \zeta)$.

$$X(n, \zeta) = X(n, -\zeta) \quad \text{(proof: } \hat{X}(n, \sigma) \in \mathbb{R}) \quad (9)$$

$$\hat{X}(n, \sigma) = \hat{X}(n, -\sigma) \quad \text{(proof: } (9) \text{ and } X(n, \zeta) \in \mathbb{R}) \quad (10)$$

$$\hat{X}(\omega, \zeta) = \hat{X}(\omega, -\zeta) \quad \text{(proof: } (9) \text{)} \quad (11)$$

$$\hat{X}(\omega, \zeta) = \hat{X}(\omega, \zeta)^* \quad \text{(proof: } \hat{X}(n, \sigma) \in \mathbb{R} \text{ and } X(n, \zeta) \in \mathbb{R}) \quad (12)$$

The spatial Fourier representations of $\hat{X}(\omega, \zeta)$ in (7) is unconventional, but are highly useful for understanding the content of the hyperspectral image $X(n, \sigma)$. One key observation is the notion of the spatial-spectral bandlimitedness. That is, we define the support $\Omega \subset \{\mathbb{R}/2\pi\}^2 \times \mathbb{R}$ of the hyperspectral image $\hat{X}(\omega, \zeta)$ as follows:

$$\Omega = \{ (\omega, \zeta) \in \{\mathbb{R}/2\pi\}^2 \times \mathbb{R} \mid |\hat{X}(\omega, \zeta)| > \tau \}, \quad (13)$$

where $\tau > 0$ is a threshold or a noise floor—portion of the signal whose Fourier magnitude falls below $\tau$ is declared unobservable and insignificant. Figure 2 show the support $\Omega$ for a typical hyperspectral image $\hat{X}$. The cardinality of $\Omega$ is small compared to $\{\mathbb{R}/2\pi\}^2 \times \mathbb{R}$, meaning Fourier transformation $\hat{X} \mapsto \hat{X}$ attains high energy compaction—majority of Fourier coefficients are insignificant as a result. It is also clear from Figure 2 that spatial ($\omega$) and temporal ($\zeta$) bandwidth are not separable. Indeed, the spatial bandwidth of $\hat{X}(\omega, \zeta)$ depends on OPD $\zeta=\hat{X}(\omega, 0)$ is a spatially broadband signal, while $\hat{X}(\omega, \zeta)$ for a large $\zeta$ is spatially narrowband. Conversely, coherence length of light depends on the spatial context. The coherence length of light in spatially lowpass (flat or smooth) regions of the image ($\hat{X}(0, \zeta)$) is far longer than the coherence length of light in spatially highpass regions ($\hat{X}(\omega, \zeta)$ for large $\omega$).

### III. Analysis of Hyperspectral Imaging Modalities

Owing to the energy compaction of the $\hat{X} \mapsto \hat{X}$, majority of Fourier coefficients $\hat{X}$ are insignificant. As such, an ideal sensing hardware would make measurements only on $\Omega$ and would ignore the insignificant coefficients in $\Omega^c$ (complement of $\Omega$). Below, we analyze various passive hyperspectral imaging modalities to understand their effectiveness and inefficiencies. We examine each hyperspectral imaging hardware to understand their potential use for multispectral imaging. By no means do the authors claim that device categories considered here represents a comprehensive set of available devices. We study strictly from the perspective of the spatial-spectral model, and largely ignore device-specific optimizations and other design benefits such as complexity and speed.

#### A. Narrowband Filters

Consider the process of making dense spectral sampling by a collection of narrowband filters $\{\tilde{S}_1, \ldots, \tilde{S}_K\}$ placed in the optical path of the light. This is typically implemented as a time multiplexing, where one filter is used at a time, and the total of $K$ images with the $K$ filters need to be acquired in succession, as illustrated in Figure 3(a). Because samples with $\tilde{S}_k$ are captured at a different time instance than samples for $\tilde{S}_k$, both the sensor and the scene must be absolutely stationary in order to prevent temporal artifacts. Alternatively, filters and pixel arrays can be arranged in vertical columns, as illustrated in Figure 3(b). By scanning the scene, this arrangement allows capture of multiple spectral samples simultaneously, albeit it works only with a few filters (i.e. $K$ small); and the scanning also requires stationary sensor and scene.

Suppose for the moment that $\tilde{S}_k(\sigma) = \delta(\sigma - \sigma_k) + \delta(\sigma + \sigma_k)$ is an ideal narrowband filter, where $\delta(\cdot)$ is an impulse function and the center bands $\{\sigma_1, \ldots, \sigma_K\}$ are evenly spaced by $2\epsilon$ (i.e. $\sigma_k+2\epsilon = \sigma_{k+1}$). Then the measurements of $\hat{X}(n, \sigma)\tilde{Q}(\sigma)$ at $\sigma \in \{\sigma_1, \ldots, \sigma_K\}$ take the form:

$$Y_k(n) = \langle \tilde{S}_k, \tilde{Q} \rangle_{\sigma} = 2 \hat{X}(n, \sigma_k)\tilde{Q}(\sigma_k). \quad (14)$$

Then the samples $\{Y_1, \ldots, Y_K\}$ collectively represent a sampled signal

$$\hat{X}_{NB}(n, \sigma) = \sum_{k=1}^{K} \frac{Y_k(n)}{2\tilde{Q}(\sigma_k)} \cdot \{\delta(\sigma - \sigma_k) + \delta(\sigma + \sigma_k)\} \quad (15)$$

whose spatial Fourier coefficients $\hat{X}_{NB}(\omega, \zeta)$ take the form

$$\hat{X}_{NB}(\omega, \zeta) = \frac{1}{2\epsilon} \sum_{\nu \in \mathbb{Z}} \hat{X}(\omega, \zeta + \nu/2\epsilon). \quad (16)$$

Under this scenario, the aliasing occurs when the supports of $\hat{X}(\omega, \zeta)$ and $\hat{X}(\omega, \zeta + 1/2\epsilon)$ overlap (i.e. co-set $\Omega$ and $(0, 1/2\epsilon) + \Omega$ are mutually exclusive). This is illustrated in Figure 3(c), where the sampling interval $2\epsilon$ was too large to avoid aliasing. Practical narrowband filters are not impulse functions, but rather they are supported over a narrow range (say $\pm \sigma + (-\epsilon, \epsilon)$). This has the effect of attenuating $\hat{X}(\omega, \zeta)$ for large $\zeta$ (i.e. lowpass filter in $\zeta$ direction). Though it effectively acts similarly to an anti-aliasing filter, it does not prevent aliasing entirely due to its gradual cut-off response. Figure 3 makes it clear that sampling along $\sigma$ axis must be very dense in order to overcome aliasing. As a hyperspectral imaging device, it requires a large number of filters (i.e. $K$ large) which is unattractive for capturing a scene with motion. This design is even more undesirable for application-agnostic multispectral imaging systems since it is highly unlikely that a signal sampled with only a few spectral filters (i.e. $K$ small) will be free of aliasing. We conclude that narrowband filter-based multispectral imaging pays penalties not only for discarding many spectral samples, but also for the aliasing that contaminates the few samples that were actually acquired.
The sensor measurement on this combined light is

\[ Y_k(n) = \langle \tilde{Q}, X(n, \zeta_k) \cdot 2|\cos(\zeta_k \sigma) \rangle \sigma \]

\[ = X(n, \zeta_k) \int_{\sigma} \tilde{Q}(\sigma)(e^{i\zeta_k \sigma} + e^{-i\zeta_k \sigma})d\sigma \]

\[ = X(n, \zeta_k) Q(\zeta_k) + Q(-\zeta_k) = X(n, \zeta_k) \frac{Q(\zeta_k)}{2\pi} = X(n, \zeta_k) \frac{Q(\zeta_k)}{\pi}, \]

(18)

B. Fourier Transform Spectroscopy

Fourier transform spectroscopy is a technique to reconstruct the spectrum of light \( \hat{X}(n, \sigma) \) from interferometric measurements of the field coherence using the relations in (5) and (6). OPD refers to a difference in length between two optical fields from the same source travel prior to interfering at the detector. Figure 4 shows the classical Michelson interferometer which splits the incoming optical field into multiple beams using a beam splitter. Distance to mirrors A and B controls the degree of OPD; the combined fields interfere to realize a measurement of the form (5). Similarly to the narrowband filters, however, this method has the disadvantage that both the sensor and the scene have to be stationary.

Consider making measurements in the physical space with OPD \( \zeta \in \{\zeta_1, \ldots, \zeta_K\} \). The interference of the combined fields at the sensor is \( X(n, \zeta_k) \). This corresponds to a projection of \( \hat{X}(n, \cdot) \) to a sinusoid \( \cos(\zeta_k \sigma) \) in the following sense:

\[ \tilde{F}\{X(n, \zeta_k) \cdot (\delta(\zeta - \zeta_k) + \delta(\zeta + \zeta_k))\} = X(n, \zeta_k) \cdot 2\cos(\zeta_k \sigma). \]

(17)

The sensor measurement on this combined light is

\[ Y_k(n) = \langle \tilde{Q}, X(n, \zeta_k) \cdot 2|\cos(\zeta_k \sigma) \rangle \sigma \]

\[ = X(n, \zeta_k) \int_{\sigma} \tilde{Q}(\sigma)(e^{i\zeta_k \sigma} + e^{-i\zeta_k \sigma})d\sigma \]

\[ = X(n, \zeta_k) Q(\zeta_k) + Q(-\zeta_k) = X(n, \zeta_k) \frac{Q(\zeta_k)}{2\pi} = X(n, \zeta_k) \frac{Q(\zeta_k)}{\pi}, \]

(18)

where \( X(n, \zeta_k) = \frac{Y_k(n)}{Q(\zeta_k)/\pi} \). Then the samples \( \{Y_1, \ldots, Y_K\} \) collectively represent a signal

\[ X_{FTS}(n, \zeta) = \sum_{k=1}^{K} Y_k(n) \cdot \{\delta(\zeta - \zeta_k) + \delta(\zeta + \zeta_k)\}. \]

(19)

In the discrete spatial Fourier transform domain,

\[ \hat{X}_{FTS}(\omega, \zeta) = \sum_{k=1}^{K} \tilde{F}\{X(n, \zeta_k) \cdot (\delta(\zeta - \zeta_k) + \delta(\zeta + \zeta_k))\} \]

\[ = \sum_{k=1}^{K} \hat{X}(\omega, \zeta_k) \cdot (\delta(\zeta - \zeta_k) + \delta(\zeta + \zeta_k)). \]

(20)

Here, \( \hat{X}(\omega, \zeta_k) \) is a “slice” \( \Omega_k := \{\omega, \zeta \in \Omega | \zeta = \zeta_k\} \) on the support \( \Omega \). See Figure 5.

Key advantage of Fourier transform spectroscopy is that measurements are free of aliasing contaminations and resolution-independent because \( \Omega_k \) and \( \Omega_{k'} \) are mutually exclusive when \( k \neq k' \). [30] This property is highly attractive for multispectral imaging modalities because there is no aliasing penalty for undersampling when the number of spectral samples are low (i.e. \( K \) small). Moreover, each successive measurement contains information orthogonal to the previous sample, meaning \( \{\bigcup_{k=1}^{K} \Omega_k\} \) and \( \Omega_{K+1} \) are mutually exclusive (i.e. \( \{X(n, \zeta_1), X(n, \zeta_2), \ldots\} \) is an innovation sequence, in the sense of Kalman filtering). Hence optimal multispectral Fourier transform spectroscopy chooses sample OPDs \( \zeta \in \{\zeta_1, \ldots, \zeta_K\} \) that maximizes the cardinality of \( \Omega_k \). From Figures 2 and 5, it is obvious that \( \zeta_k = (k-1)\cdot\Delta\zeta \) for some sampling interval \( \Delta\zeta \) is optimal (more on \( \Delta\zeta \) in the next section).
Conversely, we conclude that Fourier transform spectroscopy is highly inefficient as a hyperspectral imaging device (as opposed to multispectral imaging). For a large number of spectral samples (i.e., \( K \) large), the cardinality of \( \Omega_{K+1} \) is very small since \( \zeta_{K+1} = K \cdot \Delta \zeta \) is large. As such, sample of \( \zeta_{K+1} \) provides very small refinement to \( \bigcup_{k=1}^{K} \Omega_k \) with little noticeable difference in the spectral domain representation \( \tilde{X}(n, \sigma) \). Hence measurements for high spectral resolution devices are wasteful.

Figure 6 compares the effects of undersampling with narrowband filters and with Fourier transform spectroscopy (full experiment setup is described in Section V). Using the hyperspectral image data shown in Figure 6(a), we simulated a theorized narrowband filter responses (where the widths of the filter is \( 2\epsilon \)) and OPD samples. The six equally spaced narrowband filter responses are interpolated using spline interpolation over the visible range (400nm-800nm), as shown by pink curves in Figure 6. We also compared to the five narrowband filter responses (as optimized by [22]) interpolated using PCA vectors trained over the dataset in [29] using non-negative projection [28]. When the ground truth spectrum is smooth or broadly shaped (e.g. Figure 6(b,c,e)), light spectra reconstructed from all multispectral imaging modalities matched the ground truth light spectrum well.

However, of utmost importance to the applications of hyperspectral imaging (detection/recognition/tracking of targets/materials/chemicals) are the spectral features such as narrow peaks and sharp transitions (e.g. Figures 6(d,f,g,i,j)). For instance, chlorophyll absorption is characterized by a narrow small peak in 500nm-550nm and a sharp climb near 700nm (see Figure 6(j)), which can be used to detect vegetation in remote sensing. Examining Figures 6(d,f,g,i,j), interpolated narrowband filter samples yielded oversmoothed spectra that poorly approximates concaved spectra shapes. More over, the spectral narrow peaks are shifted (e.g. Figures 6(d,g,j)) because they tend to peak at sample wavenumbers \( \pm \sigma \in \{ \sigma_1, \ldots, \sigma_K \} \) instead of at the local maxima of the light spectra. By contrast, the locations of peaks can be inferred more accurately from OPD samples, although a few extra peaks stemming from Gibbs phenomenon are evidenced in Figure 6(d,g,j) as well.
The spatial frequency response of this sampling scheme is

\[ \hat{X}(n, \sigma) = \sum_{(m, \zeta) \in MZ^2} \frac{\hat{X}(m, \zeta) \delta(n - m, \sigma - \zeta)}{ | \text{det}(M) | } \]  

(21)

C. Spatial Arrangement of Spectral Samples

Fourier analysis in Section II-B inspires us to devise a sampling scheme that takes advantage of the structure in \( \Omega \). Consider, for example, a joint spatial-spectral lattice sampling, as illustrated in Figure 7(b). Comparing to the implied sampling grid of Figure 3 illustrated in Figure 7(a), lattice in Figure 7(b) makes very few measurements. Let \( M \in \mathbb{R}^{3 \times 3} \) be the generator matrix and \( MZ^3 \) is the sampling lattice. Then the captured image is

\[ \tilde{X}_{\text{Lat}}(n, \sigma) = \sum_{(m, \zeta) \in MZ^3} \tilde{X}(m, \zeta) \delta(n - m, \sigma - \zeta) \]  

(22)

Similarly to Figures 3(c) and 5, \( \tilde{X}_{\text{Lat}} \) is a linear combination of shifted versions of \( \tilde{X} \), as illustrated in Figure 7(c). Aliasing occurs if generator matrix \( M \) is chosen so that co-sets \( \{ \theta, \nu \} + \Omega \) are mutually exclusive for all \( \{ \theta, \nu \} \in \{2\pi M^{-T} \mathbb{Z}^2 \} \cap \{[-\pi, \pi]^2 \times \mathbb{R} \} \) (as shown by example in Figure 7(c)).

Compared to hardware configurations in Figures 3 and 4, fewer measurements are required to overcome aliasing. Figure 7 only shows one spatial dimension \( (n_0, n_1) \in \mathbb{Z} \). Lattice sampling design in two dimensions \( (n = (n_0, n_1) \in \mathbb{Z}^2) \) will require even fewer spatial-spectral samples.

IV. FOURIER SPECTRAL FILTER ARRAY

Recall that majority of Fourier coefficients \( \hat{X}(\omega, \zeta) \) are insignificant. An \textit{ideal} sensing hardware would make measurements only on \( \Omega \) and would ignore insignificant coefficients in \( \Omega^c \). Optimal multispectral imaging device would maximize the overlap between the acquired data and \( \Omega \) with the fewest number of spectral measurements. Consider Figure 8, where a hyperspectral image signal is represented in space-wavenumber \( (\tilde{X}(n, \sigma)) \), physical \( (X(n, \zeta)) \), and spatial frequency \( (\hat{X}(\omega, \zeta)) \) domains. Qualitatively, it is clear from Figure 8(a) that \( \hat{X}(n, \sigma) \) is a high bandwidth data for all values of \( \pm \omega \sigma \in (\sigma_{\min}, \sigma_{\max}) \). It is difficult to determine what narrowband filter sets \( \{\sigma_1, \ldots, \sigma_K\} \subset (\sigma_{\min}, \sigma_{\max}) \), if any, would yield alias-free measurements of \( \tilde{X}_{NB}(\omega, \zeta) \) over its support \( \Omega \), as our analysis in Section III-A already proved.

On the other hand, as evidenced by Figure 2 and confirmed by Figure 8(b), the representation in \( X(\omega, \zeta) \) enjoyed rapid decay in signal strength as \( \zeta \geq 1 \) increases. Sampling in the increasing order of \( \zeta \) will naturally maximize the overlap with acquired data and \( \Omega \). Fourier transform multispectroscopy is also ideal for making a few alias-free spectral samples without the penalty of aliasing, as discussed in Section III-B.

By comparison, consider Figure 8(c). The support of the spatial Fourier transform \( \hat{X}(\omega, \zeta) \) decreases in the increasing order of \( \zeta \). From multispectral imaging perspective, it is desirable to design a device that would sample in this domain directly, where the number of measured samples would scale proportionally with the diminishing support of \( \hat{X}(\omega, \zeta) \). Since non-significant samples in \( \hat{X}(\omega, \zeta) \) belong to \( \Omega^c \) and do not need to be measured, the goal is to find a way to repurpose \( \Omega^c \) for measuring “modulated” \( \Omega \). In a sense, this was accomplished by hardware configuration in Figure 7(b-c)—this method uses the fewest overall number of measurements. However, it is still not a single-shot solution that we are after.

In the remainder of this article, we provide a novel multispectral imaging design that draws from alias-free measurements of Fourier transform spectroscopy in Figure 5, multiplexing of lattice sampling in Figure 7, and single-shot image acquisition of spectral filter array design. We will develop mathematical framework to design a new SFA pattern, and techniques to analyze data acquired from such device. The resultant SFA pattern—which we refer to as Fourier SFA—differs significantly from the previously proposed patterns in [15]–[17], [21]–[24] in the sense that the proposed filters are broadband and sinusoidal, that they avoid spatial and spectral aliasing, and that demosaicking is simple and effective. This configuration comes with the benefit of being able to capture the spectral content of the image in a single shot, which is amenable to imaging dynamic scenes with moving objects.

A. Fourier Analysis Of SFA Sampling Revisited

We take another look at the Fourier analysis of SFA. Instead of focusing on reconstruction \( \tilde{X}(\omega, \zeta) \) as we have done in Section III, we examine more closely the sensor data \( Y(n) \). Define \( X_Q : \mathbb{Z}^2 \times \mathbb{R} \rightarrow \mathbb{R}^+ \) as hyperspectral image signal \( \tilde{X} \) that is weighted by the spectral response curve \( \tilde{Q} \):

\[ \tilde{X}_Q(n, \sigma) = \tilde{Q}(\sigma) \tilde{X}(n, \sigma) \]

\[ \tilde{X}_Q(\omega, \zeta) := \int_{\mathbb{R}} \tilde{Q}(\nu) \tilde{X}(\omega, \zeta - \nu) d\nu \]  

(23)

Generalizing (5), the sensor measurement is:

\[ Y(n) = (S, \tilde{X}_Q)_{\sigma} = \int_{\mathbb{R}} S(n, \sigma) \tilde{X}_Q(n, \sigma) d\sigma \]  

(24)
Taking its spatial Fourier transform, we have
\[
\hat{Y}(\omega) = \mathcal{F}\{Y(n)\} = \left(\frac{1}{4\pi^2}\right) \left\{\hat{S}(\omega, \zeta) \ast \hat{X}_Q(\omega, \zeta)\right\}_{\zeta = 0},
\]
(25)
where \(\ast\) denotes a three dimensional convolution, as follows:
\[
\hat{S}(\omega, \zeta) \ast \hat{X}_Q(\omega, \zeta) := \int_{\mathbb{R}} \int_{\{\mathbb{R}/2\pi\}^2} S(\theta, \nu) X(\omega - \theta, \zeta - \nu) d\theta d\nu.
\]
(26)

Equation (25) is a variation on the Parseval’s theorem. Its significance is that \(\hat{Y}(\omega)\) can be interpreted as \(\{\hat{S}(\omega, \zeta) \ast \hat{X}_Q(\omega, \zeta)\}\) (the modulated version of the signal \(\hat{X}_Q\)) intersected at zero OPD (\(\zeta = 0\)).

In the subsequent discussion, we make two assumptions on the spectral curve \(\tilde{Q}\) of a typical image sensor:

- **assumption 1**: \(\tilde{Q}(\sigma) > 0 \forall \pm \sigma \in (\sigma_{min}, \sigma_{max})\).
- **assumption 2**: \(\tilde{Q}(\sigma)\) is a smooth and well-behaved function.

Indeed, Figure 9 confirms that both assumptions are realistic. The first assumption enables us to recover \(\hat{X}\) from \(\hat{X}_Q\) by the precise inverse \(\hat{X}(n, \sigma) = \hat{X}_Q(n, \sigma)/\tilde{Q}(\sigma)\) if \(\hat{X}_Q\) is known because \(\tilde{Q} \neq 0\). The second assumption implies that \(Q(\zeta) = \mathcal{F}^{-1}\{\tilde{Q}\}\) enjoys rapid decay over \(\zeta\). By this fact, the support \(\Omega_Q\) of \(X_Q\) is a lightly dilated version of \(\Omega\), where the Fourier bandlimitedness structure as described in Figure 2 remains largely intact for \(\Omega_Q\) (i.e. \(\Omega_Q \approx \Omega\)).

### B. SFA as a Fourier Transform Spectroscopy

We proceed forward with the goal to recover \(\hat{X}_Q\) from the SFA sensor measurements \(Y(n) = \langle \hat{S}, \hat{X}_Q \rangle_{\sigma}\), and the latent hyperspectral image \(\hat{X}\) is subsequently reconstructed from \(\hat{X}_Q\). The main idea is to parameterize the SFA pattern in the spatial frequency domain \(\hat{S}\), contrary to the conventional approach where we specify the spectral filter transmittance \(\tilde{S}\). Drawing from the analysis [32] and optimized design [19] of CFA sampling, consider a SFA pattern \(\tilde{S} : \{(\mathbb{R}/2\pi)^2\} \times \mathbb{R} \rightarrow \mathbb{R}\),

with the form:
\[
\tilde{S}(\omega, \zeta) = \sum_{k=1}^{K} \alpha_k \{\delta(\omega - \omega_k, \zeta - \zeta_k) + \delta(\omega + \omega_k, \zeta + \zeta_k)\}
\]
(27)
where symmetries of \(\tilde{S}\) respect the requirements of (11) and (12). Though the filter parameters \(\{(\alpha_1, \omega_1, \zeta_1), \ldots, (\alpha_K, \omega_K, \zeta_K)\}\) will be specified in this spatial Fourier domain, the corresponding spectral filter

---

**Fig. 8.** Example of a hyperspectral image signal in various forms of representation.

**Fig. 9.** Example quantum efficiency curve for CCD detectors assuming the spectral transmittance is 100%. Adopted from [31].
response can also be found via the transform \( \hat{S} \mapsto \tilde{S} \)
\[
\tilde{S}(n, \sigma) = \mathcal{F} \circ \mathcal{F}^{-1}(\hat{S})
\]
\[
= \sum_{k=1}^{K} 4 \{ \mathbb{R} \{ \alpha_k \} \cos(\omega_k^T n) + \mathbb{I} \{ \alpha_k \} \sin(\omega_k^T n) \} \cos(\zeta_k \sigma). 
\]
\[\text{(28)}\]

In other words, SFA of the type (27) give rise to three-dimensional (2D space, 1D spectral) sinusoidal responses in space-wavenumber domain. As Section IV-E elaborates further, this type of filter has an efficient physical implementation using thin film optics, whose parameters are specified in the physical (spatial-OPD) domain.

The SFA sensor data \( \hat{Y} \) of image \( \hat{X} \) captured with \( \hat{S}(\omega, \zeta) \) above is
\[
\hat{Y}(\omega) = \frac{1}{4\pi^2} \sum_{k=1}^{K} \alpha_k \{ \hat{X}_Q(\omega - \omega_k, \zeta - \zeta_k) \\
+ \hat{X}_Q(\omega - \omega_k, \zeta + \zeta_k) \\
+ \hat{X}_Q(\omega + \omega_k, \zeta - \zeta_k) \\
+ \hat{X}_Q(\omega + \omega_k, \zeta + \zeta_k) \} |_{\zeta=0} \\
= \sum_{k=1}^{K} \alpha_k \hat{X}_Q(\omega - \omega_k, \zeta_k) + \alpha_k^* \hat{X}_Q(\omega + \omega_k, -\zeta_k). 
\]
\[\text{(29)}\]

As shown in Figure 10 and confirmed by (29), the intersection between the spatial-OPD modulated signal \( \hat{X}_Q(\omega - \omega_k, \zeta - \zeta_k) \) and the zero-plane \( \zeta = 0 \) is a Fourier “slice” %
\[\Omega_{Q,k} := \{(\omega, \zeta) \in \Omega_Q | \zeta = \zeta_k \} \]
— in the sense of Fourier transform spectroscopy. Thus, one can view the SFA pattern in (27) as a new technique for carrying out Fourier transform spectroscopy, where \( \hat{X}_Q(\omega, \zeta_k) \) is spatially modulated to \( (\pm \omega_k, -\zeta_k) + \Omega_{Q,k} \) and recorded in \( \hat{Y}(\omega) \).

C. Optimal Parameter Selection

The filter parameters in (29) are \{\{\alpha_1, \omega_1, \zeta_1\}, \ldots, \{\alpha_K, \omega_K, \zeta_K\}\}. At a high level, spatial modulation frequencies \( \omega_k \), is selected to minimize the risk of aliasing. Specifically, we inherit the main advantage of the Fourier transform spectroscopy—measurements in \( \Omega_{Q,k} \) are free of OPD aliasing contaminations. However, \( \hat{Y}(\omega) \) may support multiple slices, and spatial aliasing may still occur if the support of \( \hat{X}_Q(\omega_{\pm \omega_k}, \zeta_{\pm \zeta_k}) \) and \( \hat{X}_Q(\omega_{\pm \omega_{k'}}, \zeta_{\pm \zeta_{k'}}) \) overlap for some \( k \neq k' \). That is,
\[
\{(\pm \omega_k, -\zeta_k) + \Omega_{Q,k}\} \cap \{(\pm \omega_{k'}, -\zeta_{k'}) + \Omega_{Q,k'}\} \neq \emptyset. 
\]
\[\text{(30)}\]

Leveraging the multiplexing ideas of lattice sampling in Figure 7(b-c) and the optimal CFA design of [19], we seek to find combinations of spatial modulation frequencies \( \{\omega_1, \ldots, \omega_K\} \) that are mutually far apart enough to avoid aliasing in (30), but yet close enough to densely populate the modulated Fourier slices in \( \hat{Y} \). The modulation frequency \( \omega_k \) also controls the degree of repetition there exists in the SFA pattern (as opposed to, say, randomized SFA pattern). Repetition is desirable because it promotes global processing that is easy to optimize for performance. Repetition is also desirable for fabrication, as the size of the repeating pattern controls the number of filter masks needed in fabrication. Mathematically, if \( \{\omega_1, \ldots, \omega_K\} \subset 2\pi M^{-1} \mathbb{Z}^2 \) for some a generator matrix \( M \in \mathbb{Z}^{2 \times 2} \), then the SFA pattern is guaranteed to repeat in the pixel domain according to the lattice \( M \mathbb{Z}^2 \). The number of pixels within a repeating pattern is |det(\( M \))|, and the maximum number of modulations that lattice structure \( M \) can support is also |det(\( M \))|. This implies that repetition increases in size if more OPD samples need to be acquired.

The sample OPDs \( \zeta_k \) determine the Fourier slices, and its sampling interval depends on the spectral bandwidth \((\sigma_{\min}, \sigma_{\max})\). Discussions in Section III-B already established that \( \zeta_k = (k-1) \cdot \Delta \zeta \) is optimal for maximizing the cardinality of measurement space \( \{\hat{X}_{Q,k} \} \) with the fewest number of samples (i.e. small \( K \)). At the same time, it is desirable to set the sampling interval \( \Delta \zeta \) as large as possible to take advantage of rapid decay of \( \Omega_{Q,k} \) sizes, which reduces the probability that \( (\pm \omega_k, -\zeta_k) + \Omega_{Q,k} \) and \( (\pm \omega_{k'}, -\zeta_{k'}) + \Omega_{Q,k'} \) overlap to cause spatial aliasing. Given that the \( \hat{X}_Q(\omega, \sigma) \) is bandlimited, it suffices to sample at its Nyquist rate \( \Delta \zeta = \frac{1}{\sigma_{\max}} \) to represent signal contained in \( \pm \sigma \in (\sigma_{\min}, \sigma_{\max}) \).

To see why this is the case, define OPD-sampled signal \( \hat{X}_{\Delta \zeta} : \{\mathbb{R}/2\pi^2\}^2 \times \mathbb{R} \to \mathbb{C} \)
\[
\hat{X}_{\Delta \zeta}(\omega, \zeta) = \sum_{k=-\infty}^{\infty} \hat{X}_Q(\omega, k \Delta \zeta) \delta(\zeta - k \Delta \zeta). 
\]
\[\text{(31)}\]

The corresponding representation in space-wavenumber domain is:
\[
\hat{X}_{\Delta \zeta}(n, \sigma) = \frac{1}{\Delta \zeta} \sum_{k=-\infty}^{\infty} \hat{X}_Q(n, \zeta - k \Delta \zeta). 
\]
\[\text{(32)}\]

If \( \Delta \zeta = \frac{1}{\sigma_{\max}} \), \( \hat{X}_Q(n, \zeta) \) and \( \hat{X}_Q(n, \zeta - k \cdot 2 \sigma_{\max}) \) do not
alias since

$$\left( -\sigma_{\text{max}}, \sigma_{\text{max}} \right) \cap \{ k \cdot 2\sigma_{\text{max}} + (-\sigma_{\text{max}}, \sigma_{\text{max}}) \} = \{ \}$$

$$\forall k \neq 0.$$ (33)

One can increase the sampling interval to $\Delta \zeta = \frac{1}{Q_{\text{max}}}$ if $\sigma_{\text{min}} > \frac{\sigma_{\text{max}}}{2}$, however. This condition can be met in the visible-near infrared wavelengths range, for example, when $\sigma_{\text{min}} = \frac{1}{4} 400 \text{nm}$ and $\sigma_{\text{max}} = \frac{1}{2} 400 \text{nm}$. Based on (32), $\tilde{X}_Q(n, \zeta)$ and $\tilde{X}_Q(n, \zeta - k \cdot \sigma_{\text{max}})$ do not alias since $(\sigma_{\text{min}}, \sigma_{\text{max}}) \subset \left( \frac{\sigma_{\text{max}}}{2}, \sigma_{\text{max}} \right)$ and

$$\left( \frac{\sigma_{\text{max}}}{2}, \sigma_{\text{max}} \right) \cap \{ k \cdot \sigma_{\text{max}} \} \cap \left( (-\sigma_{\text{max}}, -\sigma_{\text{max}}) \cup \left( \frac{\sigma_{\text{max}}}{2}, \sigma_{\text{max}} \right) \right)$$

$$= \left\{ \left( \frac{\sigma_{\text{max}}}{2}, \sigma_{\text{max}} \right) \cap \left( (-k+1)\sigma_{\text{max}}, (k-1)\sigma_{\text{max}} \right) \right\}$$

$$\cup \left\{ \left( \frac{\sigma_{\text{max}}}{2}, \sigma_{\text{max}} \right) \cap \left( (k+1)\sigma_{\text{max}}, (k+1)\sigma_{\text{max}} \right) \right\}$$

$$= \{ \}. \quad \text{(34)}$$

The increased sampling interval of $\Delta \zeta = \frac{1}{Q_{\text{max}}}$ is more advantageous since the cardinality of $\Omega_{Q,k}$ decays faster in $\tilde{X}_Q(\omega, \zeta)$, allowing more Fourier slices to fit into $\hat{Y}(\omega)$. On the other hand, the coefficient $\alpha_k \in \mathbb{C}$ is the Fourier strength of the modulation. Since $\alpha_k$ parameter does not affect aliasing (i.e. spatial or spectral resolution), we treat $\alpha_k$ as the flexible parameters that allow for further optimization. Large $\alpha_k$ value is desirable because it can be made robust to noise (imagine noise being added to $\hat{Y}(\omega)$ in (29)). The Fourier magnitudes $\{ \alpha_1, \ldots, \alpha_K \}$ are also subject to real-world constraints—spectral response of the filter $\tilde{S}$ must fall within the physically realizable range of $[0, 1]$. Supposing $\omega_1 = (0, 0)$ and $\zeta_1 = 0$, physical realizability condition can be satisfied by increasing $\alpha_1$ (ensures $S(n, \sigma) \geq 0$) and decreasing $\{ \alpha_2, \ldots, \alpha_K \}$ (ensures $S(n, \sigma) \leq 1$). In practice, we optimized the weights $\alpha_k \in \mathbb{C}$ based on finsearch function in Matlab based on reconstruction quality of the demosaicking algorithm outlined in Section IV-D.

Figure 11(a) shows an example arrangement of spatial modulation frequencies $\{ \omega_1, \ldots, \omega_6 \}$. All modulation frequencies occupy positions in $\frac{\pi}{2} \mathbb{Z}^2$ grid, which yields a repeating SFA pattern of size $4 \times 4$. We have $\omega_1 = (0, 0)$ and $\zeta_1 = 0$ at the spatial DC, because $\Omega_{Q,1}$ has the largest support. The Fourier slice with the next two largest support are $\Omega_{Q,2}$ and $\Omega_{Q,3}$ corresponding to $\tilde{X}_Q(\omega, 1\Delta)$ and $\tilde{X}_Q(\omega, 2\Delta)$. In our design, they appear together at $\omega_2 = \omega_3 = (\pi, \pi/2)$. This is allowed if the corresponding weights $\alpha_2 + \alpha_3$ and $\alpha_3 + \alpha_5$ (corresponding to $-\omega_2 = (\pi, -\pi/2)$) form a linearly independent samples that can be reversed later. Similarly, $\tilde{X}_Q(\omega, 3\Delta)$ and $\tilde{X}_Q(\omega, 4\Delta)$ appear together at $\omega_4 = \omega_5 = (\pi/2, \pi)$ with linearly independent weights $\alpha_4 + \alpha_5$ and $\alpha_4 + \alpha_6$. Finally, $\tilde{X}_Q(\omega, 5\Delta)$ is modulated to $\omega_6 = (\pi, \pi)$. Figure 11(b) shows the result of taking a Fourier mapping $\tilde{S} \rightarrow S$. Since this SFA has a $4 \times 4$ repeating pattern, there are at most 16 different filter types (spatially arranged as shown in Figure 14(a)) that are shown in Figure 11(b). Indeed the filters spectral transmittances exhibit strong signs of periodicity, as each filter is designed to be sum of $K$ sinusoids, where $K$ is small.

As a side comment, we considered an alternate design with $\tilde{X}_Q(\omega, 1\Delta)$modulating to $\omega_2 = (\pi, \pi)$ (i.e. farthest from $\omega_1 = (0, 0)$). However, this design is inefficient compared to the shared modulation of $\tilde{X}_Q(\omega, 1\Delta)$ and $\tilde{X}_Q(\omega, 2\Delta)$ were completely eliminated.

D. Demosaicking

With no aliasing, demosaicking can be implemented as a standard linear demodulation [19], [32], [33]. Recall (29) and let $\{ \theta_1, \ldots, \theta_L \}$ be the $L = \lceil \text{det}(M) \rceil$ elements in the set $2\pi M^{-T} \mathbb{Z}^2 \cap [-\pi, \pi)^2$. In the first step of demosaicking, the sensor data $Y(n)$ is multiplied by oscillator $e^{-j\theta_i n}$ and then smoothed by a Gaussian filter $H(\omega)$:

$$Z_i(n) = H_i(n) \ast \{ Y(n) e^{-j\theta_i n} \} \quad \text{(35)}$$

In the spatial frequency domain, this is equivalent to

$$\hat{Z}_i(\omega) = \hat{H}_i(\omega) \cdot \hat{Y}(\omega - \theta_i)$$

$$= \frac{\hat{H}_i(\omega)}{2\pi^2} \sum_{k=1}^{K} \mathcal{A}_k \hat{X}_Q(\omega - \theta_i - \omega_k, \zeta_k)$$

$$+ \alpha_k^* \hat{X}_Q(\omega - \theta_i + \omega_k, \zeta_k)$$

$$\approx \frac{1}{2\pi^2} \sum_{k=1}^{K} \mathcal{A}_k \hat{X}_Q(\omega, \zeta_k) \delta(\omega_k - \theta_i) + \alpha_k \hat{X}_Q(\omega, \zeta_k) \delta(\omega_k + \theta_i),$$

where the approximation in the last step stems from the fact that Gaussian filter $\tilde{H}_i(\omega)$ is a lowpass filter (whose width is adjusted accordingly to accomplish the following task) and the assumption that $\hat{Y}(n)$ is alias-free. We summarize the relation in (36) by (37) on page 11 below. Hence we invert this process:

$$X_Q(n) = A^\dagger Z(n) \quad \text{(38)}$$

where $A^\dagger$ denotes a matrix inverse if $K = L$ and pseudo-inverse if $K < L$. Once $X_Q(n) = (X_Q(n, 0 \cdot \Delta \zeta), \ldots, X_Q(n, (K - 1) \cdot \Delta \zeta))^T$ is recovered, we can reconstruct $\tilde{X}_Q(n, \sigma)$ as

$$X_D(n, \zeta) = \sum_{k=1}^{K} X_Q(n, \zeta_k) \cdot \{ \delta(\zeta - \zeta_k) + \delta(\zeta + \zeta_k) \} \quad \text{(39)}$$

$$\tilde{X}_D(n, \sigma) = \left\{ \begin{array}{ll}
\sum_{k=1}^{K} X_Q(n, \zeta_k) \cos(\zeta_k \sigma) & \text{if } \pm \sigma \in (\sigma_{\text{min}}, \sigma_{\text{max}}) \\
0 & \text{else}
\end{array} \right.$$

$$\text{(40)}$$

Fig. 11. Example Fourier SFA design. (a) Sensor data in spatial frequency domain. (b) Spectral transmittance $\tilde{S}(n, \sigma)$. A visualization of Fourier SFA is shown in Figure 14(a).
The desired condition of \( \zeta \) can compensate for this by adjusting changes in slowly over the focal plane array, demosaicking algorithm angle of light. In hardware, one can overcome this by using a such that they can be grown as a spectral filter array. They 12. degrees of freedom to match the desired \( K \) \( K/ \) an autoregressive process where \( \alpha \) in the Fourier SFA are sums of ones shown in Figure 11(b). As described in (28), filters used use the linearly combined sinusoids to form the filters like the band filtering. The Fabry-Perot-type filter we envision would used for narrowband sampling. by combining these infinite sinusoids, which are subsequently ical usage of this technique is to form narrow spectral peaks spectral density (space-wavenumber representation). The typical usage of this technique is to form narrow spectral peaks by combining these infinite sinusoids, which are subsequently used for narrowband sampling.

The requirements for sinusoidal filters need for the Fourier SFA are very different from the requirements for the narrowband filtering. The Fabry-Perot-type filter we envision would use the linearly combined sinusoids to form the filters like the ones shown in Figure 11(b). As described in (28), filters used in the Fourier SFA are sums of \( K \) OPD samples with weights \( \alpha_k(n) \). Multilayer etalon will also accomplish this—it forms an autoregressive process where \( K/2 \) etalon layers have \( K \) degrees of freedom to match the desired \( K \) OPDs. See Figure 12.

Fabrication of Fabry-Perot-type filters are flexible enough such that they can be grown as a spectral filter array. They are also flexible for determining reflection coefficients that match \( \alpha_k(n) \). These filters, however, are sensitive to incident angle of light. In hardware, one can overcome this by using a telecentric lens system that make the incident angles uniform over the focal plane array. Alternatively, since angles vary slowly over the focal plane array, demosaicking algorithm can compensate for this by adjusting changes in \( \zeta_k \) in (39). The desired condition of \( (\sigma_{\min}, \sigma_{\max}) \subset (\frac{\sigma_{\min}}{\sigma_{\max}}, \sigma_{\max}) \) can be enforced by an additional bandpass filter applied over the entire sensor.

Efforts to enable multilayer etalon array in a working prototype is currently ongoing with our collaborators at the University of Dayton. Our initial prototype involves single Zinc Sulfate thin film with varying thickness to implement sinusoidal transmittance. Full details on the prototype reported in [30].

V. EXPERIMENTAL VERIFICATION

We simulated the proposed SFA sensor using two sets of hyperspectral data. First data set is collected using Surface Optics SOC730, a scanning device equipped with diffraction grating. It collects 128 bands over 364nm-959nm at 1024×1024 pixel resolution and 4.7nm wavelength resolution with 120ms integration time at f/11. Although the Fourier SFA results extend seemlessly to wider spectral ranges, our simulation study only used 400nm-800nm portion in order to make comparisons to narrowband SFA pattern of [22] which was designed for the visible range. Collection was conducted outdoors in August 2013 on a cloudy/sunny day. Targets used are wooden panels painted with Olympic paints. These paints are chosen to best match eight squares from the Gretag Macbeth Colorchecker. Furthermore, the spectra of these panels are measured also using an Fieldspec Pro, a calibrated contact spectrometer. Same paints were used on wooden fences to form non-achromatic edges. They are intended for testing the limits of the main assumption made in this article, namely the spatial-spectral bandlimitendness of Figure 2 which have biases towards achromatic edges. The relative spatial frequency can be increased by moving the fences farther away from the sensor. The color image shown in Figure 6(a) are shown with the empirical line method (ELM) applied using the Fieldspec Pro data as ground truth. However, the spectra in Figures 6(b-j) are shown without ELM because uncorrected light spectra is closer to what we expect the sensor to observe in real sensing modality. The second data is collected by the authors of [29] using Cambridge Research Institute Nuance FX system. The data covers 420nm-720nm with 1040×1392 pixel resolution and 10nm wavelength resolution. Since the theory developed here holds for wavenumbers, the image was converted using (4). Although this procedure technically involves interpolation, we assumed that sampling in the wavelength is dense enough that the variabilities introduced by the interpolation are negligible. Note also that the experiment setup above satisfies the desired condition \( (\sigma_{\min}, \sigma_{\max}) \subset (\frac{\sigma_{\min}}{\sigma_{\max}}, \sigma_{\max}) \).

The proposed Fourier SFA pattern is shown in Figures 13(a). The corresponding the simulated sensor data \( Y(n) \) shown in Figures 14(a) preserves high spatial details, thanks to the baseline spectra \( \hat{X}(\omega, 0) \) that is less subject to aliasing (if aliased, the spatial details would have deteriorated). This
Fig. 13. SFA patterns and the spatial frequency response of the simulated sensor data. (a) Proposed Fourier SFA pattern. (b) Log Fourier magnitude of simulated Fourier SFA sensor data of (a). (c) Narrowband SFA pattern of [22]. (d) Log Fourier magnitude of simulated Fourier SFA sensor data of (c). Simulated from hyperspectral image data acquired by Surface Optics SOC730 sensor, shown in Figure 6(a).

Fig. 14. Simulation results comparing (a,b) proposed Fourier SFA pattern against (c,d) narrowband SFA pattern of [22]. (a,c) Simulated SFA sensor data. (b,d) Demosaicking results, rendered as an RGB image based on the reconstructed hyperspectral image. Simulated from hyperspectral image data acquired by Surface Optics SOC730 sensor, shown in Figure 6(a).

Fig. 15. Comparisons of ground truth spectra (blue) to Monno demosaicking result (green and pink) of Figure 14(c-d) and from Fourier transform spectroscopy (red) of Figure 14(a-b). The blue line presents the ground truth spectra. The red dash line, green dash line, and magenta dash line present the reconstruction spectra from proposed FSFA, Monno’s SFA using Cubic interpolation, and Monno’s SFA using PCA.

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>light gray</th>
<th>dark gray</th>
<th>lime green</th>
<th>flesh</th>
<th>blue</th>
<th>green</th>
<th>red</th>
<th>yellow</th>
<th>speed (sec)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>σ</td>
<td>11.575</td>
<td>17.849</td>
<td>15.446</td>
<td>18.806</td>
<td>15.6</td>
<td>11.622</td>
<td>13.322</td>
<td>19.577</td>
</tr>
</tbody>
</table>
agrees with the spatial Fourier magnitude $\|\hat{Y}(\omega)\|$ shown in Figure 13(b). Modulation is clearly evidenced by the Fourier magnitudes peaks at $\omega = (0, 0), (\pi, \pi), (\pi, \pm \pi/2)$, and $(\pm \pi/2, \pi)$. Although $\hat{X}(\omega, 0)$ has a large support, modulation frequencies $\omega_2, \ldots, \omega_5$ adequately separate $\hat{X}(\omega \pm \omega_2, 2 \cdot \Delta \zeta), \ldots, \hat{X}(\omega \pm \omega_5, 6 \cdot \Delta \zeta)$ away from $\omega_1 = (0, 0)$. We contrast this to the state-of-art narrowband SFA pattern of Monno [22], shown in Figures 13(c) and 13(d). Despite it having a fewer sampling bands (five total), its spatial Fourier magnitude shows that the risk of spatial aliasing is high. Specifically, modulations occur at $\omega = (\pm \pi/2, \pm \pi/2), (0, 0)$ and $(\pi, 0)$, which severely restrict the allowable bandwidth of the baseband at $\omega = (0, 0)$. In fact, our previous work in [19] rigorously proved that it is not possible for narrowband SFA patterns to be alias-free (see Proposition 2 in [19]; though derived for CFA patterns, results generally apply to SFA as well). In practice, aliasing may be overcome by sophisticated demosaicking (albeit the computational overhead), as evidenced by the demosaicking results of [22] shown in Figure 14(d). Besides the spatial aliasing, Monno’s SFA carries additional risk for using narrowband filters that we demonstrated in Figure 6.

The proposed demosaicking yielded satisfactory results. Figure 14(b) shows the result of demosaicking for the Fourier SFA pattern, which recovered the entire spectrum (Figure 15) but only displays a projection to RGB tristimulus values here. The overall reconstructed spatial details by and large are intact and well preserved, albeit the edges are softer compared to the original image in Figure 6(a). A small degree of artifacts are evidenced in the strong edges of the car parked to the right, likely due to saturation of the specular reflection (which break our bandlimitedness assumption). However, there are no obvious evidences of zippering artifacts that surface in presence of aliasing. There are some color artifacts (such as fence)—these are not aliasing artifacts (i.e. not zipperping) but rather arises as a result the small support of the bandpass filter used in demodulation (i.e. lacking highpass in sharp transitions). This can be improved by incorporating directional filters, but we leave adaptive filtering for future investigation.

We compared our result to the narrowband SFA pattern of [22] with five spectral measurements, shown in Figure 14(d). The demosaicking has comparable spatial reconstruction quality, with similarly softer-than-original spatial details. However, the demosaicking method in [22] is based on guided filter, which involve iterations that are slow compared to the single pass approach of the demodulation developed in Section IV-D. The computational overhead of upsampling the narrowband measurements to full spectrum using the PCA reconstruction approach of [28] as suggested by the authors of [22] is significant—orders of magnitude more complex than the single-pass Fourier transform spectroscopy approach of (40).

It is difficult to make out from Figure 14(b,d) the differences in spectral features based on their color appearances. The advantage of the Fourier transform spectroscopy is clear in the quality of reconstruction of the light spectrum, however. Figure 15 shows the reconstructed spectra of the panels seen in Figure 6(a), which is virtually identical to the spectra from Fourier spectroscopy in Figure 6(b-j). Signal-to-noise ratio (SNR) scores for the reconstruction qualities of the painted wooden panel targets are calculated using

$$SNR = 10 \log_{10} \left( \frac{\tilde{X}_D(n, \sigma)}{\tilde{X}(n, \sigma) - \tilde{X}_D(n, \sigma)} \right)^2,$$

where $\tilde{X}_D(n, \sigma)$ is reconstructed signal, and $\tilde{X}(n, \sigma)$ is the ground truth. SNR values in Table I clearly favor the proposed approach. Table II summarizes the SNR scores of the recovered spectra for the 73 hyperspectral images in [29]. Compared to the visible light spectra reconstructed from the narrowband SFA acquisition [22] using cubic interpolation, the proposed Fourier SFA quality was superior. We also show the result of recovering the visible light spectra using the principal component analysis in the manner described in [28]. The PCA vectors trained from 1257 Munsell color chips (as the authors in [22], [28] have done) did not yield acceptable results. We instead show the results of the PCA vectors learned from the dataset collected using Surface Optics SOC730. Through Table II, the PCA results improved on the cubic reconstruction, but it did not match the performance of the Fourier transform spectroscopy approach of (40).

VI. Conclusion

We proposed a novel general-purpose multispectral imaging using Fourier transform spectroscopy. Based on the spatial-spectral Fourier analysis, we proved that Fourier transform spectroscopy is an efficient multispectral sensing modality since the measured interferogram quantity is free of aliasing stemming from undersampling the spectral continuum. We also developed a Fourier spectral filter array pattern design that maximizes spatial resolution of the sensor by minimizing the risk of spatial aliasing stemming from spatial undersampling. We verified the effectiveness of the Fourier spectral filter array design by simulation and demonstrated that spatial-spectral fidelity of the spectra reconstructed from Fourier spectral filter array is superior to that of narrowband spectral filter array.

ACKNOWLEDGMENT

This work was sponsored in part by the U.S. Air Force Research Laboratory, U.S. Air Force Summer Faculty Fellowship Program, the Institute for Development and Commercialization of Advanced Sensor Technology, and the National Science Foundation. The authors thank the Surface Optics Corporation,

### Table II

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>proposed SFA</th>
<th>Monno SFA PCA</th>
<th>Monno SFA Cubic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average</td>
<td>14.55</td>
<td>14.29</td>
<td>13.08</td>
</tr>
<tr>
<td>Max</td>
<td>29.30</td>
<td>24.91</td>
<td>24.09</td>
</tr>
<tr>
<td>Min</td>
<td>1.25</td>
<td>2.36</td>
<td>1.24</td>
</tr>
<tr>
<td>Stdev</td>
<td>10.43</td>
<td>11.53</td>
<td>9.32</td>
</tr>
</tbody>
</table>

This table shows the signal-to-noise ratio values of hyperspectral image (computed in wavelength domain) recovered from demosaicking averaged over 73 image set in [29]. Principal component analysis (PCA) is highly dependent on data; PCA here is trained from data set collected using Surface Optics SOC730.
San Diego, CA for upgrading the SOC 730 hyperspectral imaging sensor; the authors of [22] for providing the source code to simulate their SFA design and demosaicking method; and the authors of [29] for providing access to the hyperspectral dataset used in our experiments.

REFERENCES


Jie Jia received the B.S. degree in telecommunication engineering from the Beijing University of Posts and Telecommunications, in 2009, and the M.S. degree in electrical engineering from the University of Dayton, Dayton, OH, in 2011, where he is currently pursuing the Ph.D. degree with the Intelligent Signal Systems Laboratory. He is a Research Assistant with the Intelligent Signal Systems Laboratory, University of Dayton. His current research interests in multispectral imaging.

Kenneth J. Barnard received the B.S. degree in electrical engineering from the University of Michigan, in 1986, and the M.S. and Ph.D. degrees in electrical engineering from the University of Central Florida, in 1990 and 1992, respectively. He is a Principal Research Engineer with the Air Force Research Laboratory, Wright-Patterson Air Force Base, Dayton, OH. His current research interests include day/night passive imaging sensors, focal plane array characterization, hyperspectral remote sensing, and imaging processing for performance enhancement and exploitation.
Keigo Hirakawa received the B.S. (Hons.) degree in electrical engineering from Princeton University, Princeton, NJ, in 2000, the M.S. and Ph.D. degrees in electrical and computer engineering from Cornell University, Ithaca, NY, in 2003 and 2005, respectively, and the M.M. (Hons.) degree in jazz performance studies from the New England Conservatory of Music, Boston, MA, in 2006. He was a Research Associate with Harvard University, Cambridge, MA, from 2006 to 2009. He is currently an Assistant Professor with the University of Dayton, Dayton, OH. He heads the Intelligent Signal Systems Laboratory, University of Dayton, where the group focuses on statistical signal processing, color image processing, and computer vision.