# KRISM – Krylov Subspace-based Optical Computing of Hyperspectral Images

# VISHWANATH SARAGADAM and ASWIN C. SANKARANARAYANAN, Carnegie Mellon University, USA



Fig. 1. Hyperspectral imagers resolve scenes at high spatial and spectral resolutions. We propose a novel architecture called KRISM that provides the ability to implement two operators: a spatially-coded spectrometer and a spectrally-coded spatial imager. By iterating between the two, we can acquire a low rank approximation of the hyperspectral image in a light efficient manner with very few measurements. The left image shows optical schematics for implementing the two operators. On the right, we show a hyperspectral image of a scene illuminated with a compact fluorescent lamp (CFL) acquired using our lab prototype. The proposed method enables high spatial and spectral resolution as observed in the zoomed-in image patches and CFL peaks, respectively.

We present an adaptive imaging technique that optically computes a lowrank approximation of a scene's hyperspectral image, conceptualized as a matrix. Central to the proposed technique is the optical implementation of two measurement operators: a spectrally-coded imager and a spatially-coded spectrometer. By iterating between the two operators, we show that the top singular vectors and singular values of a hyperspectral image can be adaptively and optically computed with only a few iterations. We present an optical design that uses pupil plane coding for implementing the two operations and show several compelling results using a lab prototype to demonstrate the effectiveness of the proposed hyperspectral imager.

# CCS Concepts: • Computing methodologies $\rightarrow$ Computational photography; Hyperspectral imaging;

Additional Key Words and Phrases: Krylov subspaces, optical computing, coded apertures

#### **ACM Reference Format:**

Vishwanath Saragadam and Aswin C. Sankaranarayanan. 2019. KRISM — Krylov Subspace-based Optical Computing of Hyperspectral Images. *ACM Trans. Graph.* 38, 5, Article 148 (October 2019), 14 pages. https://doi.org/https://doi.org/10.1145/3345553

#### **1** INTRODUCTION

Hyperspectral images (HSIs) capture light intensity of a scene as a function of space and wavelength and have been used in numerous vision [Kim et al. 2012; Pan et al. 2003; Tarabalka et al. 2010], geoscience and remote sensing applications [Cloutis 1996; Harsanyi

https://doi.org/https://doi.org/10.1145/3345553

and Chang 1994]. Traditional approaches for hyperspectral imaging, including tunable spectral filters and pushbroom cameras, rely on sampling the HSI, i.e., measuring the photon counts in each spatio-spectral voxel. When imaging at high-spatial and spectral resolutions, the amount of light in a voxel can be quite small, thus requiring long exposures to mitigate the effect of noise.

HSIs are often endowed with rich structures that can be used to alleviate the challenges faced by traditional imagers. For example, natural scenes are often comprised of a few materials of distinct spectra and further, illumination of limited spectral complexity [Lee et al. 2000; Parkkinen et al. 1989]. This implies that the collection of spectral signatures observed at various locations in a scene lies close to a low-dimensional subspace. Instead of sampling the HSI of the scene one spatio-spectral voxel at a time, we can dramatically speed-up acquisition and increase light throughput by measuring only projections on this low-dimensional subspace. However, such a measurement scheme requires a priori knowledge of the scene since this subspace is entirely scene dependent. This paper introduces an optical computing technique that identifies this subspace using an iterative and adaptive sensing strategy and constructs a low-rank approximation to the scene's HSI.

The proposed imager senses a low-rank approximation of a HSI by optically implementing the so-called Krylov subspace method [Golub and Kahan 1965]. We show that this requires two operators: a spatially-coded spectrometer and a spectrally-coded spatial imager; when we interpret the HSI as a 2D matrix, these two operators correspond to left and right multiplication of the matrix with a vector. The two operators are subsequently used in an iterative and adaptive imaging procedure whose eventual output is a lowrank approximation to the HSI. The proposed imager is adaptive, i.e., the measurement operator used to probe the scene's HSI at a given iteration depends on previously made measurements. This is a marked departure from current hyperspectral imaging strategies

Authors' address: Vishwanath Saragadam; Aswin C. Sankaranarayanan, Carnegie Mellon University, Electrical and Computer Engineering, Pittsburgh, PA, 15213, USA, saswin@andrew.cmu.edu.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s). © 2019 Copyright held by the owner/author(s). 0730-0301/2019/10-ART148

### 148:2 • Vishwanath Saragadam and Aswin C. Sankaranarayanan

where the signal model is merely used as a prior for recovery from non-adaptive measurements [Arce et al. 2014].

*Contributions.* We propose an optical architecture that we refer to as KRylov subspace-based Imaging and SpectroMetry (KRISM) and make the following three contributions:

- *Optical computation of HSIs.* We show that optical computing of HSIs to estimate its dominant singular vectors provides significant advantages in terms of increased light throughput and reduced measurement time.
- Coded apertures for resolving space and spectrum. Sensing architectures typically used in spectrometry and imaging are mutually incompatible due to use the of slits in spectral imaging and open apertures in conventional imaging. To mitigate this, we study the effect of pupil plane coding on the HSI and propose a coded aperture design that is capable of simultaneously achieving high spatial and spectral resolutions.
- Optical setup. We design and validate a novel and versatile optical implementation for KRISM that uses a single camera and a single spatial light modulator (SLM) to efficiently implement spatiallycoded spectral and spectrally-coded spatial measurements.

The contributions above are supported via an extensive set of simulations as well as real experiments performed using the lab prototype.

*Limitation.* The benefits and contributions described above come with a key limitation. Our method is only advantageous if there are a sufficient number of spectral bands and the hyperspectral image is sufficiently low rank. If we only seek to image with very few spectral bands or if the scene is not well approximated by a low-rank model, then the proposed method performs poorly against traditional sensing methods.

# 2 PRIOR WORK

*Nyquist sampling of HSIs.* Classical designs for hyperspectral imaging based on Nyquist sampling include the tunable filter — which scans one narrow spectral band at a time, measuring the image associated with spectral bands at each instant — or using a pushbroom camera — which scans one spatial row at a time, measuring the entire spectrum associated with each pixel on the row. Both approaches are time-consuming as well as light inefficient since each captured image wastes a large percentage of light incident on the camera.

*Multiplexed sensing.* The problem of reduced light throughput can be mitigated by the use of multiplexing. One of the seminal results in computational imaging is that the use of multiplexing codes including the Hadamard transform can often lead to significant efficiencies either in terms of increased SNR or faster acquisition [Harwit and Sloane 1979]. This can either be spectral multiplexing [Mohan et al. 2008] or spatial multiplexing [Sun and Kelly 2009]. While multiplexing mitigates light throughput issues, it does not reduce the number of measurements required. Sensing at high spatial and/or spectral resolution still requires long acquisition times to maintain a high SNR. Fortunately, HSIs have concise signal models that can be exploited to reduce the number of measurements.



Fig. 2. HSIs, interpreted as a matrix, are often low rank. We validate this observation by plotting accuracy in terms of peak SNR (PSNR) as a function of the rank of the approximation. We do this for many commonly used HSI datasets and observe that the PSNR is higher than 40dB for a rank 10 approximation across all datasets.

Low-rank models for HSIs. There are many approaches to approximate HSIs using low-dimensional models; this includes group sparsity in transform domain [Rasti et al. 2013], low rank model [Golbabaee and Vandergheynst 2012; Li et al. 2012], as well as low-rank and sparse model [Saragadam et al. 2017; Waters et al. 2011]. Of particular interest to this paper is the low-rank modeling of HSIs when they are represented as a 2D matrix (See Figure 2). These models have found numerous uses in vision and graphics including color constancy [Finlayson et al. 1994], color displays [Kauvar et al. 2015], endmember detection [Winter 1999], source separation [Hui et al. 2018], anomaly detection [Saragadam et al. 2017], compressive imaging [Golbabaee and Vandergheynst 2012] and denoising [Zhao and Yang 2015]. Chakrabarti and Zickler [2011] also provide empirical justification that HSIs of natural scenes are well represented by low dimensional models.

Compressive hyperspectral imaging. The low-rank model has also been used for compressive sensing (CS) of HSIs. CS aims to recover a signal from a set of linear measurements that are fewer than its dimensionality [Baraniuk 2007]. This is achieved by modeling the sensed signal using lower dimensional representations — low-rank matrices being one such example. The technique most relevant to this paper is that of row/column projection [Fazel et al. 2008] where the measurement model is restricted to obtaining row and column projections of a matrix. Given a matrix  $X \in \mathbb{R}^{m \times n}$ , and measurement operators  $S_{\text{row}} \in \mathbb{R}^{p \times m}$ ,  $S_{\text{column}} \in \mathbb{R}^{n \times p}$ , the measurements acquired are of the following form,

$$Y_{\rm row} = S_{\rm row}X, \quad Y_{\rm column} = XS_{\rm column}$$

When the matrix *X* has a rank *k*, it can be shown that it is sufficient to acquire *p* images and *p* spectral profiles with  $p \propto k^2$ . In contrast, the method proposed in this paper requires only a number of measurements proportional to the rank of the matrix; however, these measurements are adaptive to the scene. At an increased cost of optical complexity, adaptive sensing promises accurate results with fewer measurements than non-adaptive measurement strategies.

*Hyperspectral imaging architectures.* Several architectures have been proposed for CS acquisition of HSIs. The Dual-Disperser Coded

Method	Approach	Number of measurements	Estimation accuracy under noise	Advantages	Disadvantages
Sampling	Tunable spectral filter	$N_x N_y N_\lambda$	$\sigma \sqrt{N_x N_y N_\lambda}$	Easy calibration	Low spectral resolution; high acquisition time
	Pushbroom			High spectral resolution	Optical complexity; high acquisition time
Multiplexed	Spatial multiplexing	$N_x N_y N_\lambda$	$\sigma \sqrt{N_{\lambda}}$	- Hadamard multiplexing gain	High acquisition time
	Spectral multiplexing		$\sigma \sqrt{N_x N_y}$		
Compressive sensing	CASSI	depends on signal model		Fewer measurements	Loss in spatial/spectral resolution
	Row/column projection	$\propto k^2 (N_x N_y + N_\lambda)$	[Fazel et al. 2008]		Complex optics
KRISM (proposed method)	Optical Krylov subspace	$\propto k(N_x N_y + N_\lambda)$	prop. to model misfit + noise	Fewest number of measurements; very high light efficiency	Complex optics

Table 1. Various sensing strategies for hyperspectral imaging of  $N_x \times N_y$  spatial dimension and  $N_\lambda$  spectral bands. Noise in measurement is assumed to be AWGN with  $\sigma^2$  variance. The expressions in third column represent the number of measurements required, while those in fourth column represent the error in reconstruction.

Aperture Snapshot Spectral Imager (DD-CASSI) [Gehm et al. 2007] obtains a single image multiplexed in both spatial and spectral domains by dispersing the image with a prism, passing it through a coded aperture, and then recombining with a second prism. In contrast, the Single Disperser CASSI (SD-CASSI) [Wagadarikar et al. 2008] relies on a single prism that performs spatial coding using a binary mask followed by spectral dispersion with a prism. Baek et al. [2017] disperse the image by placing a prism right before an SLR camera. The HSI is then reconstructed by studying the dispersion of color at the edges in the obtained RGB image. Takatani et al. [2017] instead propose a snapshot imager that uses a faced reflectors overlaid with color filters. Various other snapshot techniques have been proposed which rely on space-spectrum multiplexing [Cao et al. 2016; Jeon et al. 2016; Lin et al. 2014a]. While snapshot imagers require only a single image, they often produce HSIs with reduced spatial or spectral resolutions. Data-driven approaches such as overcomplete dictionaries [Lin et al. 2014a] and convolutional neural networks [Choi et al. 2017] partially alleviate the loss in resolution by building priors for the HSI. However, they require complex optimization that can often be time consuming.

Resolution and accuracy of the HSI can be improved by acquiring multiple measurements instead of a single snapshot image. Examples include multiple spatio-spectrally encoded images [Kittle et al. 2010], spatially-multiplexed spectral measurements [Li et al. 2012; Sun and Kelly 2009] or separate spatial and spectral coding [Lin et al. 2014b]. While multi-measurement techniques overcome spatial and spectral resolution limits, the price is paid in the form of increased number of measurements and hence, reduced time resolution.

Performance of snapshot techniques can be improved by tailoring the spatial masks to a given HSI dataset [Rueda et al. 2016, 2017] or by optimizing spatial masks for sensing a selected subset of spectral bands [Arguello and Arce 2013]. Optimizing the spatial masks results in increased accuracy, but still requires long reconstruction times. A key insight into the existing methods is that the measurements are either non-adaptive and random, or adapted to a fixed signal class. In contrast, the proposed method is *adapted to the specific instance of the signal*, requires fewer measurements, and has practically no post-processing for reconstruction. Table 1 compares and contrasts various HS imaging strategies and their relative merits in terms of number of measurements and error in reconstruction. We next discuss the concept of Krylov subspaces for low-rank approximation of matrices, which motivates iterative and adaptive techniques and paves the way to the proposed method.

Krylov subspaces. Central to the proposed method is a class of techniques, collectively referred to as Krylov subpaces, for estimating singular vectors of matrices. Recall that the singular value decomposition (SVD) of a matrix  $X \in \mathbb{R}^{m \times n}$ ,  $m \le n$  is given as  $X = U\Sigma V^{\top}$ , where  $U \in \mathbb{R}^{m \times m}$  and  $V \in \mathbb{R}^{n \times n}$  are orthonormal matrices, referred to as the singular vectors, and  $\Sigma \in \mathbb{R}^{m \times n}$  is a diagonal matrix of singular values. Krylov subspace methods allow for efficient estimation of the singular values and vectors of a matrix and enjoy two key properties. First, we only need access to the matrix X via left and right multiplications with vectors, i.e., we do not need explicit access to the elements of the matrix X. Second, the top singular values and vectors of a low-rank matrix can be estimated using a small set of matrix-vector multiplications. These two properties are invaluable when the matrix is very large or when it is implicitly represented using operators or, as is the case in this paper, the matrix is the scene's HSI and we only have access to optical implementations of the underlying matrix-vector multiplications.

There are many variants of Krylov subspace techniques which differ mainly on their robustness to noise and model mismatch. The techniques in this paper are based on the so-called Lanczos bidiagonalization with full orthogonalization [Golub and Kahan 1965; Hernandez et al. 2007]. Such iterative operations to reduce the complexity of matrix-vector multiplications have found use in communication theory in the form of reduced-rank filtering [Ge et al. 2004; Tian et al. 2005] and adaptive beam forming [Ge et al. 2006]. Our goal is to leverage the benefits of iterative operations for low-rank approximation of high dimensional optical signals, in particular HSIs.

*Optical computing of low-rank signals.* Matrix-vector and matrixmatrix multiplications can often be implemented as optical systems. Such systems have been used for matrix-matrix multiplication [Athale and Collins 1982], matrix inversion [Rajbenbach et al. 1987],

# 148:4 • Vishwanath Saragadam and Aswin C. Sankaranarayanan

as well as computing eigenvectors [Vijaya Kumar and Casasent 1981]. Of particular interest to our paper is the optical computing of the light transport operator using Krylov subspace methods [O'Toole and Kutulakos 2010]. The light transport matrix T represents the linear mapping between scene illumination and a camera observing the scene. Each column of the matrix T is the image of the scene when only a single illuminant is turned on. Hence, given a vector  $\ell$  that encodes the scene illumination, the image captured by the camera is given as  $\mathbf{r} = T\boldsymbol{\ell}$ . By Helmholtz reciprocity, if we replaced every pixel of the camera by a light source and every illuminant with a camera pixel, then the light transport associated with the reversed illumination/sensing setup is given as  $T^{\top}$ . Hence, by co-locating a projector with the camera and a camera with the scene's illuminants, we have access to both left- and right-multiplication of the light transport matrix with vectors; we can now apply Krylov subspace techniques for optically estimating a low-rank approximation to the light transport matrix. This delightful insight is one of the key results in [O'Toole and Kutulakos 2010].

This paper proposes a translation of the ideas in [O'Toole and Kutulakos 2010] to hyperspectral imaging. However, as we will see next, this translation is not straightforward and requires the construction of novel imaging architectures.

# 3 OPTICAL KRYLOV SUBSPACES FOR HYPERSPECTRAL IMAGING

In this section, we provide a high-level description of optical computing of HSIs using Krylov subspace methods.

Notation. We represent HSIs in two different ways:

- $H(x, y, \lambda)$  a real-valued function over 2D space (x, y) and 1D spectrum  $\lambda$ ,
- $X \in \mathbb{R}^{N_x N_y \times N_\lambda}$  a matrix with  $N_x N_y$  rows and  $N_\lambda$  columns, such that each column corresponds to the vectorized image at a specific spectrum.

The goal is to optically build the following two operators:

• Spectrally-coded imager I – Given a spectral code  $\mathbf{x} \in \mathbb{R}^{N_{\lambda}}$ , we seek to measure the image  $\mathbf{y} \in \mathbb{R}^{N_x N_y}$  given as

$$\mathbf{y} = \mathcal{I}(\mathbf{x}) = X\mathbf{x}.$$
 (1)

The image **y** corresponds to a grayscale image of the scene with a camera whose spectral response is **x**.

 Spatially-coded spectrometer S − Given a spatial code x̃ ∈ ℝ<sup>N<sub>x</sub>N<sub>y</sub></sup>, we seek to measure a spectral measurement ỹ ∈ ℝ<sup>N<sub>λ</sub></sup> given as

$$\widetilde{\mathbf{y}} = \mathcal{S}(\widetilde{\mathbf{x}}) = X^{\top} \widetilde{\mathbf{x}}.$$
(2)

The measurement  $\tilde{y}$  corresponds to the spectral measurement of the scene, where-in the spectral profile of each pixel is weighted by the corresponding entry in the spatial code  $\tilde{x}$ .

Since the two operators correspond to left and right multiplication of a vector to the HSI matrix X, we can implement any Krylov subspace technique to estimate the top singular vectors and values.

Number of measurements required. To obtain a rank-k approximation of the matrix X, we would require at least k spatiallycoded spectral measurements – each of dimensionality  $N_{\lambda}$ , and k spectrally-coded images – each of dimensionality  $N_x N_y$ . Hence,

$$\frac{k(N_x N_y + N_\lambda)}{N_x N_y N_\lambda} = k \left(\frac{1}{N_\lambda} + \frac{1}{N_x N_y}\right).$$
(3)

For low-rank HSIs, we can envision dramatic reductions in measurements required over Nyquist sampling especially when sensing at high spatial and spectral resolutions (see Table 1).

*Challenges in implementing operators I and S.* Spatially-coded spectral measurements have been implemented in the context of compressive hyperspectral imaging [Sun and Kelly 2009]. Here, light from a scene is first focused onto an SLM that performs spatial coding, and then directed into a spectrometer. For spectral coding at a high-resolution, we could replace the sensor in a spectrometer with an SLM; subsequently, we can form and measure an image of the coded light using a lens. However, high-resolution spectrometers invariably use a slit aperture that produces a large one-dimensional blur in the spatial image due to diffraction. We show in Section 4 that simultaneous spatio-spectral localization is not possible with either a slit or an open aperture. This leads to the design of optimal binary coded apertures which enable high spectral and spatial resolutions. Subsequently, in Section 6, we present the design of KRISM and validate its performance in Section 7.

# 4 CODED APERTURES FOR SIMULTANEOUS SENSING OF SPACE AND SPECTRUM

In this section, we introduce an optical system capable of simultaneously resolving space and spectrum at high resolutions.

# 4.1 Optical setup

The ideas proposed in this paper rely on the optical setup shown in Figure 3 which is a slight modification of a traditional spectrometer. An objective lens focuses a scene onto its image plane, that we denote as P1. This is followed by two 4f relays with a coded aperture placed on the first pupil plane, P2, and a diffraction grating placed at the plane marked as P3. We are interested in the intensity images formed at the planes marked at the "rainbow plane" P4 and the "spatial plane" P5, and their relationship to the image formed on P1, the coded aperture, and the grating parameters.

We assume that the field formed on the plane P1 is *incoherent* and, hence, we only need to consider its intensity and how it propagates, and largely ignore its phase. Let  $H(x, y, \lambda)$  be the intensity of the field as a function of spatial coordinates (x, y) and wavelength  $\lambda$ . Let a(x, y) be the aperture code placed at the plane P2,  $v_0$  be the density (measured in grooves per unit length) of the diffraction grating in P3, and f be the focal length of the lenses that form the 4f relays. The hyperspectral field intensity at the plane P4 is given as

$$F_4(x, y, \lambda) = \frac{1}{\lambda^2 f^2} a^2 (-x + f \lambda v_0, -y) S(\lambda),$$
(4)

where  $S(\lambda)$  is the scene's overall spectral content defined as

$$S(\lambda) = \int_x \int_y H(x, y, \lambda) dx dy.$$

2019-10-21 09:24. Page 4 of 1-14.

the number of measurements required by the approach is proportional to  $k(N_x N_y + N_\lambda)$  and, over traditional Nyquist sampling, it represents a reduction in measurements by a factor of

ACM Transactions on Graphics, Vol. 38, No. 5, Article 148. Publication date: October 2019.



Fig. 3. Schematic diagram of simultaneous spatio-spectral measurements with a coded aperture. The diffraction grating disperses light along x-axis. The image of the scene is formed on plane P1. The coded aperture is placed in P2, which introduces a diffraction blur in spatial plane P3, and dictates the spectral profile formed on the plane P4. A slit or an open aperture on P2 is not a good choice for simultaneously high spatial and spectral resolution. Instead, we rely on design of a novel pupil aperture that enables simultaneous high spatial and spectral resolution.

The intensity field at the spatial plane P5 is given as

$$F_5(x, y, \lambda) = H(x, y, \lambda) * \left| \frac{1}{\lambda^2 f^2} A\left(-\frac{x}{\lambda f}, -\frac{y}{\lambda f}\right) \right|^2,$$
(5)

where A(u, v) is the 2D spatial Fourier transform of the aperture code a(x, y), and \* denotes two-dimensional spatial convolution along x and y axes. These expressions arise from Fourier optics [Goodman 2005] and their derivation is provided in the supplemental material.

*Image formed at the rainbow plane P4.* A camera with spectral response  $c(\lambda)$  placed at the rainbow plane would measure

$$I_{R}(x,y) = \int_{\lambda} a^{2}(-x + f\lambda v_{0}, -y) \frac{1}{\lambda^{2} f^{2}} S(\lambda)c(\lambda)d\lambda$$
  

$$\propto a^{2}(-x, -y) * \left(S\left(\frac{x}{fv_{0}}\right)\widetilde{c}\left(\frac{x}{fv_{0}}\right)\right), \qquad (6)$$

where  $\tilde{c}(\lambda) = c(\lambda)/\lambda^2 f^2$ . Here, the dimensionless term  $fv_0$ , that scales of the spectrum  $S(\cdot)$ , indicates the resolving power of the diffraction grating. For example, we used a focal length f = 100 mm and a grating with groove density  $v_0 = 300$  grooves/mm for the prototype discussed in Section 6; here,  $fv_0 = 30,000$ . This implies that the spectrum is stretched by a factor of 30,000. Therefore, a 1 nm of the spectrum maps to  $30 \ \mu$ m, which is about 6-7 pixel-widths on the camera that we used. The key insight this expression provides is that the image  $I_R$  is the convolution of the scene's spectrum — denoted as a 1D image — with the aperture code  $a(\cdot, \cdot)$  (see Figure 4). This implies that we can measure the spectrum of the scene, albeit convolved with the aperture code on this plane; this motivates our naming of this plane as the rainbow plane.

*Image at the spatial plane P5.* A camera with the spectral response  $c(\lambda)$  placed at the spatial plane P5 would measure

$$I_{S}(x,y) = \int_{\lambda} \left( H(x,y,\lambda) * \left| \frac{1}{\lambda^{2} f^{2}} A\left(-\frac{x}{\lambda f},-\frac{y}{\lambda f}\right) \right|^{2} \right) c(\lambda) d\lambda$$
(7)

 $I_S$  is a "spatial image" in that spectral components of the HSI have been integrated out. Hence, we refer to P5 as the spatial plane. Figure 4 shows the image formed at P5 for different choices of the coded apertures, including slits and open apertures.



Fig. 4. We implemented the setup shown in Figure 3 to verify the effect of different pupil codes. The scene consists of a resolution chart illuminated by two distinct narrowband light sources. An open aperture leads to sharp spatial images, but the spectrum is blurred. On the other hand, a slit offers high spectral resolution, but the spatial image is blurred. Optimized codes offer invertible spectral blur, and at the same time, invertible spatial blur.

Implementing KRISM operations. The derivation above suggests that we get a spatial image of the scene formed at the spatial plane P5 and a spectral profile at the rainbow plane P4. We can therefore build the two operators central to KRISM by coding light on one of the planes while measuring it at the other. For the spectrally-coded imager I, we will place an SLM on the rainbow plane P4 while measuring the image, with a camera, at P5. For the spatially-coded spectrometer S, we place an SLM on P3 — which is optically identical to P5 — while measuring the image formed at P4.

2019-10-21 09:24. Page 5 of 1-14.

ACM Transactions on Graphics, Vol. 38, No. 5, Article 148. Publication date: October 2019.

Effect of the aperture code on the scene's HSI. Introducing an aperture code a(x, y) on the plane P2 can be interpreted as distorting the scene's HSI in two distinct ways. First, a spectral blur is introduced whose point spread function (PSF) is a scaled copy of the aperture code a(x, y). Second, a spatial blur is introduced for each spectral band whose PSF is the power spectral density (PSD) of the aperture code, suitably scaled. With this interpretation, the images formed on planes P4 and P5 are a spectral and spatial projection, respectively, of this new blurred HSI. Our proposed technique measures a low-rank approximation to this blurred HSI and we can, in principle, deblur it to obtain the true HSI of the scene. However, the spatial and spectral blur kernels may not always be invertible. As we show next, the choice of the aperture is critical and that traditional apertures such as a slit in spectrometry and an open aperture in imaging will not lead to invertible blur kernels.

#### 4.2 Failure of slits and open apertures

We now consider the effect of the traditional apertures used in imaging and spectrometry - namely, an open aperture and a slit, respectively - on the images formed at the rainbow and the spatial planes. Suppose that the aperture code a(x, y) is a box function of width W mm and height H mm, i.e.,

$$a(x, y) = \operatorname{rect}_W(x) \operatorname{rect}_H(y).$$

Its Fourier transform A(u, v) is the product of two sincs

$$A(u, v) = \operatorname{sinc}(Wu)\operatorname{sinc}(Hv).$$

The spatial image  $I_S$  is convolved with the PSD  $|A(u, v)|^2$  scaled by  $f\lambda$ , so the blur observed on it has a spatial extent of  $f\lambda/W \times f\lambda/H$  units. Suppose that f = 100 mm and  $\lambda = 0.5\mu$ m, the observed blur is  $50/W \times 50/H \ (\mu m)^2$ . The rainbow plane image  $I_R$ , on the other hand, simply observes a box blur whose spatial extent is  $W \times H$  mm<sup>2</sup>. Armed with these expressions, we can study the effect of an open and a slit apertures on the spatial and rainbow images.

Scenario #1 — An open aperture. Suppose that W = H = 10 mm, then we can calculate the spatial blur to be  $5\mu$ m in both height and width and hence, we can expect a very sharp spatial image of the scene. The blur on the rainbow image has a spread of 10 mm; for relay lenses with focal length f = 100mm and grating with groove density  $v_0 = 300$  grooves/mm, this would be equivalent of a spectral blur of  $10,000/30 \approx 333$  nm. Hence, we cannot hope to achieve high spectral resolution with an open aperture.

*Scenario* #2 — A *slit*. A slit is commonly used in spectrometers; suppose that we use a slit of width  $W = 100\mu$ m and height H = 10mm. Then, we expect to see a spectral blur of  $100/30 \approx 3.3$  nm. The spatial image is blurred along the y-axis by a  $5\mu$ m blur and along the x-axis by a  $50/0.1 = 500\mu$ m blur; effectively, with a  $5\mu$ m pixel pitch, this would correspond to a 1D blur of 100 pixels. In essence, the use of a slit leads to severe loss in spatial resolution.

Figure 4 shows images formed at the rainbow and spatial planes for various aperture codes. This validates our claim that conventional imagers are unable to simultaneously achieve high spatial and spectral resolutions due to the nature of the apertures used. We next design apertures with carefully engineered spectral and spatial blurs, which can be deblurred in post-processing.

#### 4.3 Design of aperture codes

We now design an aperture code that is capable of resolving both space and spectrum at high-resolutions. Our use of coded apertures is inspired by seminal works in coded photography for motion and defocus deblurring [Levin et al. 2007; Raskar et al. 2006; Veeraraghavan et al. 2007].

*Observation.* Recall that the rainbow plane image  $I_R$  is a convolution between a 1D spectral profile  $s(\cdot)$  and a 2D aperture code a(x,y). This convolution is one dimensional, i.e., along the *x*-axis; hence, we can significantly simplify the code design problem by choosing an aperture of the form

$$a(x, y) = a(x) \operatorname{rect}_{H}(y), \tag{8}$$

with H being as large as possible. The choice of the rect function along the *y*-axis leads to a high light throughput as well as a compact spatial blur along the *y*-axis. For ease of fabrication, we further restrict the aperture code to be binary and of the form

$$a(x) = \sum_{k=0}^{N-1} a_k \mathbb{I}_{[k\Delta, (k+1)\Delta]}(x),$$
(9)

where  $\mathbb{I}_{[p,q]}(x) = 1$  when  $x \in [p,q]$  and zero otherwise. Hence, the mask design reduces to finding an *N*-bit codeword  $\mathbf{a} = \{a_0, \ldots, a_{N-1}\}$ . The term  $\Delta$ , with units in length, specifies the physical dimension of each bit in the code. We fix its value based on the desired spectral resolution. For example, for f = 100mm and  $v_0 = 300$  grooves/mm, a desired spectral resolution of 1nm would require  $\Delta \leq 30\mu$ m.

Our goal is to design masks that enable the following:

- *High light throughput.* For a given code length *N*, we seek codes with large light throughput which is equal to the number of ones in the code word **a**.
- *Invertibility of the spatial and spectral blur.* The code is designed such that the resulting spatial and spectral blur are both invertible.

An invertible blur can be achieved by engineering its PSD to be flat. Given that the spectrum is linearly convolved with a(x), a  $(N + N_{\lambda} - 1)$ -point DFT of the code word a captures all the relevant components of the PSD of a(x). Denoting this  $(N+N_{\lambda}-1)$ -point DFT of a as A[k], we aim to maximize its minimum value in magnitude. Recall from (7) that the spatial PSF is the power spectral density (PSD) of a(x), with suitable scaling. Specifically, the Fourier transform of spatial blur is given by  $c(\lambda f u)$ , where c(x) = a(x) \* a(-x) is the linear autocorrelation of a(x) and u represents spatial frequencies. From (9), we get,

$$c(x) = a(x) * a(-x)$$
  
= 
$$\sum_{k=-N}^{N-1} c_k \left( \mathbb{I}_{[k\Delta,(k+1)\Delta]}(x) * \mathbb{I}_{[k\Delta,(k+1)\Delta]}(x) \right), \quad (10)$$

where  $c_k$  is the discrete linear autocorrelation of  $a_k$ . Thus, it is sufficient to maximize  $c_k$  to obtain an invertible spatial blur.

We select an aperture code that leads to invertible blurs for both space and spectrum by solving the following optimization problem:

$$\max_{a_0,\ldots,a_{N-1}} \alpha \min_k \left( |\mathbf{A}[k]| \right) + (1-\alpha) \min_k c_k, \tag{11}$$

ACM Transactions on Graphics, Vol. 38, No. 5, Article 148. Publication date: October 2019.

2019-10-21 09:24. Page 6 of 1-14.



Fig. 5. Optimized codes ensure that the spectral as well as spatial blur can be deconvolved stably. We simulate the performance of optimal code on a spatial and spectral target similar to Figure 4. Spectrum was deconvolved using Wiener deconvolution, and spatial images were deconvolved using TV prior. Optimized codes offer high spatial as well as spectral resolution.

under the constraint that the elements of a are binary-valued, and  $\alpha \in (0, 1)$  is a constant. For code length N sufficiently small, we can simply solve for the optimal code via exhaustive search of all  $2^N - 1$  code words. We used N = 32 and an exhaustive search for the optimal code took over a day. The resulting code and its performance is shown in Figure 5 and 6; we used  $\Delta = 100\mu$ m and H = 6.4mm for this result. A brute force optimization is not scalable for larger codes. Instead of searching for optimal codes, we can search for approximately optimal codes by iterating over a few candidate solutions. This strategy has previously been explored in [Raskar et al. 2006], where 6 million candidate solutions are searched for a 52-dimensional code.

Figure 6 shows the frequency response of both spectral and spatial blurs for the 32-dimensional optimized code. The advantages of optimized codes are immediately evident — an open aperture has several nulls in spectral domain, while a slit attenuates all high spatial frequencies. The optimized code retains all frequencies in both domains, while increasing light throughput.

# 5 SYNTHETIC EXPERIMENTS

We tested KRISM via simulations on three different datasets, listed in Table 2, and compared against existing approaches. For all methods, we simulated both photon and readout noise respectively as Poisson and Gaussian random variables. All KRISM simulations were done with diffraction effects due to coded aperture.

We quantify performance through compression in measurements N/M which is ratio of number of unknowns to measurements and peak signal to noise ratio (PSNR). Given a HSI matrix **x** and its reconstruction  $\hat{\mathbf{x}}$ , we define peak SNR as

$$PNSR = 20 \log_{10} \left( \frac{\|\mathbf{x}\|_{\infty}}{RMSE(\mathbf{x}, \widehat{\mathbf{x}})} \right),$$

where RMSE is the root mean squared error defined as

$$\text{RMSE}(\mathbf{x}, \widehat{\mathbf{x}}) = \sqrt{\frac{1}{N} \sum_{n=1}^{N} (\mathbf{x}_n - \widehat{\mathbf{x}}_n)^2}.$$
 (12)

*Comparison with snapshot techniques.* Snapshot techniques such as CASSI [Wagadarikar et al. 2008] and spatial-spectral encoded CS [Lin et al. 2014a] recover HSI from a single image and hence are appropriate for video-rate hyperspectral imaging. In contrast, KRISM is *not* a snapshot technique since, at the very least it requires



(a) Freq. response of spectral blur

(b) Freq. response of spatial blur

Fig. 6. Frequency response of spatial and spectral blur for various pupil codes. Width of the slit was  $100\mu$ m, while that of open aperture was 3.2mm. The length of optimized code is 32-bits, with each bit being  $100\mu$ m wide, giving a 3.2mm wide aperture. We assume that a slit can resolve up to 1nm. In the graph, 0.5 cycles/nm corresponds to a spectral resolution of 1nm, and hence the frequency response of the slit falls off after 0.5 cycles/nm. Similarly, the maximum spatial resolution is  $15\mu$ m and hence  $f_x$  is shown till 32 cycles/mm. For spectral measurements, a slit has a flat frequency response, while an open aperture has several nulls. In contrast, an open aperture has no nulls for spatial measurements, whereas a slit attenuates high frequencies. Optimized codes have a fairly flat frequency response for spectral blur, and no nulls for spatial blur.

Dataset	Spatial resolution	#Spectral bands	Waveband (nm)
KAIST [Choi et al. 2017]	512 x 384	31	400-700
Harvard [Charkabarti et al. 2011]	696 x 520	31	400-700
ICVL [Arad and Ben-Shahar 2016]	256 x 256	260	390-1043

Table 2. Datasets used for simulations. The spatial resolution for KAIST and Hardvard datasets, and spectral resolution for ICVL dataset was reduced to keep computation tractable with competing methods.

the measurement of an image and a spectral profile. Nevertheless, we compare KRISM against snapshot techniques by varying the number of KRISM iterations. Figure 7 shows performance of these methods with varying number of measurements on KAIST and Harvard datasets. We observe that in the setting closest to snapshot mode, Choi et al. [2017] and Lin et al. [2014a] do outperform KRISM; this is to be expected since after a single iteration, KRISM provides only a rank-1 approximation. As the number of KRISM iterations are

ACM Transactions on Graphics, Vol. 38, No. 5, Article 148. Publication date: October 2019.

#### 148:8 • Vishwanath Saragadam and Aswin C. Sankaranarayanan



Fig. 7. Evaluation against snapshot techniques. We compare KRISM with varying rank against results from [Lin et al. 2014a] and [Choi et al. 2017] in terms of compression as well as accuracy. We show zoomed in image patches for each method and spectrum at pixel marked by a cross. At similar compression rates (K = 1), KRISM has lower accuracy than snapshot techniques. However, snapshot techniques require solving a complex optimization problem that can be time consuming. In contrast, KRISM requires practically no reconstruction time as the dominant singular vectors are captured directly.

increased (which allows approximations of higher ranks), KRISM performance improves. KRISM enjoys advantages when we look at computational cost for reconstruction. The reconstruction time for Choi et al. [2017] is more than 10 minutes<sup>1</sup> even with multiple GPUs, while it runs to several hours for Lin et al. [2014a]<sup>2</sup>. In contrast, KRISM requires practically no reconstruction time for recovering the HSI as we directly measure the singular vectors.

*Comparison with multi-frame techniques.* Since KRISM is essentially a multi-frame technique, we compare against multi-frame version of CASSI [Kittle et al. 2010], and spatially-multiplexed hyperspectral imager [Sun and Kelly 2009]. We simulate spatiallymultiplexed HSI imager via randomly permuted Hadamard multiplexed spectra and recover using sparsity of individual bands in wavelet domain. Note that the compression ratio is lower for Kittle et al. [2010] and Sun and Kelly [2009] since the results were inaccurate for higher compressions<sup>3</sup>. Figure 8 shows a comparison of recovered spatial and spectral images for ICVL dataset. The poor performance of Kittle et al. [2010] is due to usage of a translational mask to get multiple measurements. On the other hand, Sun and Kelly [2009] performs poorly as multiplexing is done only in the spatial domain. Performance can be improved if we multiplex in the spectral domain as well; the resulting method is the low-rank CS approach proposed by Fazel et al. [2008]. This results in an increase in accuracy with fewer measurements, as seen in Figure 8 (f). Note that CS-based techniques are based on random projections and are not adapted to the scene. In contrast, KRISM adaptively computes a low-rank approximation leading to an increase in accuracy with the same number of measurements as Fazel et al. [2008].

Based on these simulations, we conclude that KRISM is indeed a compelling methodology when spatial and spectral resolution are high - a desirable operating point in many applications. When the number of spectral bands are smaller, the gains are modest, but nevertheless present. In the next section, we provide an optical schematic for implementing KRISM.

# 6 THE KRISM OPTICAL SETUP

We now present an optical design for implementing the two operators presented in Section 3 and analyzed in Section 4. For efficiencies in implementation, we propose a novel design that combines both operators into one compact setup. Figure 9 shows a schematic that uses polarization to achieve both operators with a single SLM and a single camera. First, in Figure 9(a), an SLM is placed 2f away from the grating, and an image sensor 2f away from the SLM, implementing spectrally coded spatial measurement operator I. In Figure 9(b),

 $<sup>^1</sup>We$  used code, dataset and model from https://github.com/KAIST-VCLAB/deepcassi $^2We$  used code, dataset and overcomplete dictionary from the paper itself.

<sup>&</sup>lt;sup>3</sup>Please see supplementary for further details.

ACM Transactions on Graphics, Vol. 38, No. 5, Article 148. Publication date: October 2019.



Fig. 8. Evaluation with multi-frame techniques. We compare KRISM against spatially-multiplexed HSI [Sun and Kelly 2009], multi-frame version of CASSI [Kittle et al. 2010], and row/column CS [Fazel et al. 2008]. We show zoomed in image patches for each method and spectrum at pixel marked by a cross. Across the board, KRISM has highest accuracy with fewest measurements.

light follows an alternate path where in the SLM is 4f away from the grating; the camera is still 2f away from the SLM. This light path allows us to achieve the spatial-coded spectral measurement operator  $\mathcal{S}$ . The two light pathways are combined using a combination of polarizing beam splitters (PBS) and liquid crystal rotators (LC). The input light is pre-polarized to be either S-polarized or P-polarized. When the light is P-polarized, the SLM is effectively 2f units away from the grating leading to implementation of I, the spectrally-coded imager. When the light is S-polarized, the SLM is 4f units away, provided the polarizing beamsplitter, PBS 3 was absent. To counter this, an LC rotator is placed before PBS 3 that rotates S-polarization to P-polarization when switched on. Hence, when S-light is input in conjugation with the rotator being switched on, we achieve the operator S, a spatially-coded spectrometer. By simultaneously controlling the polarization of input light and the LC rotator, we can implement both I and S operators with a single camera and SLM pair.

Figure 10 shows our lab prototype with the entire light pathway including the coded aperture placed in the relay system between

the objective lens and diffraction grating. The input polarization is controlled by using a second LC rotator with a polarizer, placed before the diffraction grating. Finally, an auxiliary camera is used to image the pattern displayed on the SLM. This camera is used purely for alignment of the pattern displayed on the SLM. A detailed list of components can be found in the supplemental material.

*Calibration.* Our optical setup requires three calibration processes. The first one is camera to SLM calibration. We used an auxiliary camera (Component 12 in Figure 10) that is directly focused on the SLM for this purpose. The second one is calibration of wavelengths. We used several narrowband filters to identify the location of wavelengths. Finally, the third one is radiometric calibration. We used a calibrated Tungsten-Halogen light source to estimate the spectral response of the setup. A detailed description of the calibration procedure can be found in the supplementary material.

System characterization. Spectral resolution (FWHM) of the setup was computed using several 1nm narrowband filters across visible wavelengths. Our optical setup provided an FWHM of 2.9nm. Spatial resolution was computed by capturing photo of a Siemens star, and then deconvolving with a point-spread function obtained by capturing image of a  $10\mu m$  pinhole. The frequency that achieved 30% of the modulation transfer function, MTF30, was found to be nearly 0.4 line pairs/pixel. All computation details, as well as relevant figures, can be found in the supplementary material.

Diffraction due to LCoS pattern. Since the SLM is placed 2f away from spectral or spatial measurements, the displayed pattern introduces diffraction blur, creating a non-linear measurement system. To counter this, we add a constant offset to both positive and negative patterns, which makes the diffraction blur compact enough that the non-linearities can be neglected.

Spectral deconvolution. Measurements by our optical system return spectra at each point, convolved by the aperture code. To get the true spectrum, we deconvolved the  $k^{\text{th}}$  measured singular vector using a smoothness prior. The specific objective function we used:

$$\min_{\mathbf{v}_{k}} \quad \frac{1}{2} \|\mathbf{y}_{k} - a * \mathbf{v}_{k}\|^{2} + \eta \|\nabla \mathbf{v}_{k}\|^{2}, \tag{13}$$

where  $\mathbf{v}_k$  is the true spectrum,  $\mathbf{y}_k$  is the measured spectrum, a is the aperture code,  $\nabla \mathbf{x}$  is the first order difference of  $\mathbf{x}$ , and  $\eta$  is weight of penalty term. Solution to (13) was computed using conjugate gradient descent. Higher  $\eta$  favors smoother spectra, and hence is preferred for illuminants with smooth spectra, such as tungsten-halogen bulb or white LED. For peaky spectra such as CFL, a lower value of  $\eta$  is preferred. In our experiments, we found  $\eta = 1$  to be appropriate for peaky spectra, whereas,  $\eta = 10^3$  was appropriate for experiments with tungsten-halogen illumination. We compare performance of various deconvolution algorithms in the supplementary section.

*Spatial deconvolution*. Equation (7) suggests that the spatial blur kernel varies across different spectral bands. More specifically, the blur kernels at two different spectra are scaled versions of each other. However, we observed that the variations in blur kernels were not significant when we image over a small waveband — for example, the visible waveband of 420 – 680nm. Given this, we approximate

### 148:10 • Vishwanath Saragadam and Aswin C. Sankaranarayanan



(a) Optical setup in spatially-coded spectral measurement mode

(b) Optical setup in spectrally-coded spatial measurement mode

Fig. 9. Proposed optical setup in spectral coding (a) and spatial coding (b) mode. The optical method relies on polarization to switch between the two types of coding. When the input light is S-polarized, the LC rotator is switched off, enabling spectrally coded spatial measurements. When the input light instead is P-polarized, the LC rotator is turned on, which enables spatially coded spectral measurements. The input light polarization is controlled by a second LC rotator placed before the grating. With a novel use of LC rotators, our optical setup enables dual coding of hyperspectral scenes with a single camera-SLM pair.



Fig. 10. Photograph of our lab prototype. The optical paths for spectral as well as spatial coding shown in Fig. 9 have been overlaid for easy understanding. Components have been marked, grouped and labeled for convenience. All other relevant information is available in supplementary material.

the spatial blur as being spectrally independent, which leads to the following expression:

$$I_{S}(x,y) \propto \left[ \int_{\lambda} H(x,y,\lambda)c(\lambda)d\lambda \right] * p(x,y),$$
(14)

where p(x, y) is the spatial blur. We estimated the spatial blur kernel by imaging a pinhole and subsequently deconvolved the spatial singular vectors. We used a TV prior based deconvolution using the technique in [Bioucas-Dias and Figueiredo 2007] using the image of a pinhole as the PSF. Details of the deconvolution procedure are in the supplementary section.

# 7 REAL EXPERIMENTS

We present several results from real experiments which show the effectiveness of KRISM. We evaluate the ability to measure singular vectors with high accuracy, and high spatial and spectral resolution capabilities. Unless specified, experiments involved a capture of a rank-4 approximation of the HSI, with 6 spectral and 6 spatial measurements. Lanczos iterations were initialized with all-ones spatial image to speed up convergence. HSIs were acquired with a spatial resolution of 560 × 550 pixels and a spectral resolution of 256 bands between 400nm to 700nm, with 3 nm FWHM. For verifying spectroradiometric capabilities, we obtained spectral measurements at a small set of spatial points using an Ocean Optics FLAME spectrometer. We use spectral angular mapper (SAM) [Yuhas et al. 1992] similarity and PSNR between spectra measured by our optical setup

and that measured with a spectrometer. SAM between two vectors **x** and  $\hat{\mathbf{x}}$  is defined as  $SAM = \cos^{-1}\left(\frac{\mathbf{x}^{\top}\hat{\mathbf{x}}}{\|\|\mathbf{x}\|\|\|\hat{\mathbf{x}}\|}\right)$ .

Visualization of Lanczos iterations. Figure 11 shows iterations for the "Color checker" scene in Figure 14. The algorithm initially captures brightest parts of the image, corresponding to the spectralon, and the white and yellow patches. Consequently, by iteration 5, the blue and red parts of the image are isolated. The iterations are representative of the signal energy in various wavelengths. Maximum energy is concentrated in yellow wavelengths, due to tungstenhalogen illuminant and spectral response of the camera. This is then followed by the red wavelengths, and finally the blue wavelengths.

*Comparison of measured singular vectors.* We obtain the complete hyperspectral image through a permuted Hadamard multiplexed sampling in the spectral domain for comparison with ground-truth singular vectors. We chose a scene with four colored dice for this purpose, shown in Figure 12 (a). We then computed 4 singular vectors of spectrally Hadamard-multiplexed data. Figure 12 shows a comparison of the spatial and spectral singular vectors. The singular vectors obtained via Krylov subspace technique are close to the ones obtained through Hadamard sampling. On an average, the reconstruction accuracy between KRISM and Hadamard multiplexing was found to be greater than 30dB, while the angle between the singular vectors was no worse than 20°, with the top three singular vector having an error smaller than 8°. Hadamard sampling took



Iteration 4: spectral and spatial

Iteration 5: spectral and spatial

Iteration 6: spectral and spatial

Fig. 11. Data captured during measurement process for a rank-4 approximation of the "Color checker" scene for six iterations. A picture of the scene is shown in Figure 14. Positive part of data is shown in red and the negative part is shown in blue. KRISM alternates between acquiring spectral and spatial measurements to compute both spatial and spectral singular vectors. The first four iterations involve capturing the dominant wavelengths that includes yellow and green colors, since they have the highest magnitude. The next set of iterations capture the blue and red wavelengths.



(a) Setup with RGB image in inset

(b) First singular vector

(c) Second singular vector



Fig. 12. Comparison of singular vectors captured via spectrally Hadamard-multiplexed sensing and KRISM for the dice scene. the left image singular vector is from Hadamard multiplexed data and the right one is from KRISM. Blue represents negative values and red represents positive values. KRISM required a total of 6 spectral and 6 spatial measurements to construct 4 singular vectors. While spectral Hadamard sampling method took a total of 49 minutes, KRISM took under 2 minutes. The SAM value between the singular vectors was less than 20°.

49 minutes while KRISM took under 2 minutes for 6 spatial and 6 spectral measurements, thus offering a speedup of  $20\times$ .

*Peaky spectrum illumination.* We imaged a small toy figurine of "Chopper", placed under CFL, which has a peaky spectrum, to test high spatio-spectral resolving capability. Figure 1 shows the rendered RGB image and spectra at a representative location. Spectra at a selected spatial point, as measured by KRISM, and a spectrometer are shown as well. The SAM between spectrum measured by KRISM and that measured by spectrometer was found to be 14.7°. Notice that the location of the peaks, as well as the heights match accurately. Indeed, the chopper example establishes the high spatio-spectral resolution capabilities of KRISM.

*Diverse real experiments.* Figure 13 shows several real world examples captured with our optical setup, with a diverse set of objects. For verification with ground truth, we captured spectral profiles at select spatial locations. The "Dice" and "Objects" scene captures several more colorful objects with high texture. The zoomed-in pictures show the spatial resolution, while the comparison of spectra

highlights the fidelity of our system as a spectral measurement tool. "Ace" scene was captured by placing the toy figurine under CFL illuminant, which is peaky. We could not obtain ground truth with a spectrometer, as the toy was too small to reliably probe with a spectrometer. The peaks are located within 2nm of ground truth peaks, and the relative heights of the peaks match the underlying color. "Crayons" scene consists of numerous colorful wax crayons illuminated with a tungsten-halogen lamp. The closeness of spectra with respect to spectrometer readings shows the spectral performance of our setup. Finally, "Feathers" consists of several colorful feathers illuminated by tungsten-halogen lamp. The fine structure of feathers is well captured by our setup.

*Color checker.* Since our setup is optimized for viewing in 400nm-700nm, we evaluated our system on the 24-color Macbeth color chart. The Macbeth color chart consists of a wide gamut of colors in visible spectrum that are spectrally well separated, and forms a good test bench for visible spectrometry. We placed the "Color passport" and spectralon plug in front of our camera and illuminated it with a

# 148:12 • Vishwanath Saragadam and Aswin C. Sankaranarayanan



Fig. 13. Real data captured with our optical setup. We show the physical setup used for capturing the data, rendered RGB image with some interesting patches zoomed in, and spectra at some points, compared with a spectrometer. The results are promising, as the spectra is very close to spectrometer readings (PSNR > 20dB), and the spatial images are captured in high resolution.

ACM Transactions on Graphics, Vol. 38, No. 5, Article 148. Publication date: October 2019.



Fig. 14. Macbeth color chart. Spectra is shown at four locations and compared with spectrometer readings. The PSNR is 25dB or higher and the SAM between KRISM spectra and spectrometer readings is less than  $6^{\circ}$ .

tungsten-halogen DC light source. The spectralon has a spectrally flat response, and hence helps estimate the spectral response of the illuminant+spectrometer system. This enables measurement of true radiance of the color swatches. Since the spectra is smooth, we used least squares recovery of the spectrum, with  $\ell_2$  penalty on the first difference of spectral singular vectors. The captured data was then normalized by dividing spectrum of all points with the spectrum of the spectralon. Figure 14 shows the captured image against reference color chart along with spectra at select locations plotted along with ground truth spectra. On an average, the PSNR between spectra measured by KRISM and that measured by spectrometer is greater than 25dB, while the SAM is less than 6°.

#### 8 DISCUSSION AND CONCLUSION

We presented a novel hyperspectral imaging methodology called KRSIM, and provided an associated novel optical system for enabling optical computation of hyperspectral scenes to acquire the top few singular vectors in a fast and efficient manner. Through several real experiments, we establish the strength of KRISM in three important aspects: 1) the ability to capture singular vectors of the hyperspectral image with high fidelity, 2) the ability to capture an approximation of the hyperspectral image with 20× or faster acquisition rate compared to Nyquist sampling, and 3) the ability to measure simultaneously at high spatial and spectral resolution. We believe that our setup will trigger several new experiments in adaptive imaging for fast and high resolution hyperspectral imaging.

Added advantages. There are two additional advantages to KRISM. One, since we capture the top few singular vectors directly, there is a data compression from the acquisition itself. Two, the only recovery time involves deconvolution of a few spatial and spectral



Fig. 15. Capturing higher singular vectors. Since KRISM computes higher singular vectors by progressively blocking more light, photon noise dominates measurements after some iterations resulting in noisy estimates of singular vectors. The above example shows inaccurately estimated fifth singular vectors measured for "color checker" scene with our lab prototype.

singular vectors, which is significantly less than the time required for recovery of hyperspectral images from CS measurements.

*Beyond low-rank volumes.* Key to our paper is the assumption that the underlying HSI is low-rank. Sensing a high rank HSI will require several measurements which negates the benefits of KRISM. However, there are several other matrix sampling techniques that rely on row or column sensing [Hašan et al. 2007; Ou and Pellacini 2011] to capture information about high rank matrices in an efficient manner. Since the proposed setup is capable of computing arbitrary matrix-vector products, such matrix sampling techniques can be implemented efficiently.

*Effect of photon noise.* Although Krylov subspace based methods are very robust to noise [Simoncini and Szyld 2003], the quality of the singular vectors degrade as the rank of acquisition is increased (see Figure 15). This is primarily due to photon noise, as we progressively block most of the energy contained in initial singular vectors. This can be mitigated by increasing the exposure time of measurements for higher singular vectors. All said, the problem of noisy higher singular vectors exists with any kind of sampling scheme and hence needs separate attention via a good noise model.

## 9 ACKNOWLEDGEMENT

The authors thank Prof. Ioannis Gkioulekas (Robotics Institute, CMU) for valuable feedback, and Ms. Yi Hua (ECE Department, CMU) for help with making the figures. The authors acknowledge support via the NSF CAREER grant CCF-1652569, the National Geospatial-Intelligence Agency's Academic Research Program (Award No. HM0476-17-1-2000), and the Intel ISRA on compressive sensing. Vishwanath Saragadam also gratefully acknowledges support via the Prabhu and Poonam Goel fellowship.

#### REFERENCES

- Gonzalo R Arce, David J Brady, Lawrence Carin, Henry Arguello, and David S Kittle. 2014. Compressive coded aperture spectral imaging: An introduction. *IEEE Signal Processing Magazine* 31, 1 (2014), 105–115.
- Henry Arguello and Gonzalo R Arce. 2013. Rank minimization code aperture design for spectrally selective compressive imaging. *IEEE Trans. Image Processing* 22, 3 (2013), 941–954.
- Ravindra A Athale and William C Collins. 1982. Optical matrix-matrix multiplier based on outer product decomposition. Appl. Optics 21, 12 (1982), 2089–2090.

#### 148:14 • Vishwanath Saragadam and Aswin C. Sankaranarayanan

- Seung-Hwan Baek, Incheol Kim, Diego Gutierrez, and Min H Kim. 2017. Compact single-shot hyperspectral imaging using a prism. ACM Trans. Graphics 36, 6 (2017), 217:1–12.
- Richard G Baraniuk. 2007. Compressive sensing. IEEE Signal Processing Magazine 24, 4 (2007), 118–121.
- José M Bioucas-Dias and Mário AT Figueiredo. 2007. A new TwIST: Two-step iterative shrinkage/thresholding algorithms for image restoration. *IEEE Trans. Image* processing 16, 12 (2007), 2992–3004.
- Xun Cao, Tao Yue, Xing Lin, Stephen Lin, Xin Yuan, Qionghai Dai, Lawrence Carin, and David J Brady. 2016. Computational snapshot multispectral cameras: Toward dynamic capture of the spectral world. *IEEE Signal Processing Magazine* 33, 5 (2016), 95–108.
- Ayan Chakrabarti and Tod Zickler. 2011. Statistics of real-world hyperspectral images. In Comp. Vision and Pattern Recognition (CVPR).
- Inchang Choi, Daniel S. Jeon, Giljoo Nam, Diego Gutierrez, and Min H. Kim. 2017. Highquality hyperspectral reconstruction using a spectral prior. ACM Trans. Graphics 36, 6 (2017), 218:1–13.
- Edward A Cloutis. 1996. Review Article Hyperspectral geological remote sensing: Evaluation of analytical techniques. Intl. J. Remote Sensing 17, 12 (1996), 2215–2242.
- Maryam Fazel, Emmanuel J Candes, Benjamin Recht, and Pablo A Parrilo. 2008. Compressed sensing and robust recovery of low rank matrices. In Asilomar Conf. Signals, Systems and Comp.
- Graham D Finlayson, Mark S Drew, and Brian V Funt. 1994. Color constancy: Generalized diagonal transforms suffice. J. Optical Society of America A 11, 11 (1994), 3011–3019.
- Hongya Ge, Ivars P Kirsteins, and Louis L Scharf. 2006. Data dimension reduction using Krylov subspaces: Making adaptive beamformers robust to model orderdetermination. In Intl. Conf. Acoustics Speech and Signal Processing (ICASSP).
- Hongya Ge, LL Scharf, and Magnus Lundberg. 2004. Reduced-rank multiuser detectors based on vector and matrix conjugate gradient Wiener filters. In Workshop on Signal Processing Adv. Wireless Communications.
- Michael E Gehm, Renu John, David J Brady, Rebecca M Willett, and Timothy J Schulz. 2007. Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics Express* 15, 21 (2007), 14013–14027.
- Mohammad Golbabaee and Pierre Vandergheynst. 2012. Hyperspectral image compressed sensing via low-rank and joint-sparse matrix recovery. In *Intl. Conf. Acoustics, Speech and Signal Processing (ICASSP).*
- Gene Golub and William Kahan. 1965. Calculating the singular values and pseudoinverse of a matrix. J. Society for Industrial and Appl. Mathematics, Series B: Numerical Analysis 2, 2 (1965), 205–224.
- Joseph W Goodman. 2005. Introduction to Fourier optics. Roberts and Company Publishers.
- Joseph C Harsanyi and C-I Chang. 1994. Hyperspectral image classification and dimensionality reduction: An orthogonal subspace projection approach. *IEEE Trans. Geoscience and Remote Sensing* 32, 4 (1994), 779–785.
- Martin Harwit and Neil J Sloane. 1979. Hadamard Transform Optics.
- Miloš Hašan, Fabio Pellacini, and Kavita Bala. 2007. Matrix row-column sampling for the many-light problem. ACM Trans. Graphics 26, 3 (2007), 26:1–10.
- Vicente Hernandez, Jose E Roman, Andres Tomas, and Vicente Vidal. 2007. Restarted Lanczos bidiagonalization for the SVD in SLEPc. Technical Report.
- Zhuo Hui, Kalyan Sunkavalli, Sunil Hadap, and Aswin C. Sankaranarayanan. 2018. Illuminant spectra-based source separation using flash photography. In Comp. Vision and Pattern Recognition (CVPR).
- Daniel S Jeon, Inchang Choi, and Min H Kim. 2016. Multisampling compressive video spectroscopy. In *Comp. Graphics Forum*.
- Isaac Kauvar, Samuel J Yang, Liang Shi, Ian McDowall, and Gordon Wetzstein. 2015. Adaptive color display via perceptually-driven factored spectral projection. ACM Trans. Graphics 34, 6 (2015), 165:1–10.
- Min H Kim, Todd Alan Harvey, David S Kittle, Holly Rushmeier, Julie Dorsey, Richard O Prum, and David J Brady. 2012. 3D imaging spectroscopy for measuring hyperspectral patterns on solid objects. ACM Trans. Graphics 31, 4 (2012), 38:1–11.
- David Kittle, Kerkil Choi, Ashwin Wagadarikar, and David J Brady. 2010. Multiframe image estimation for coded aperture snapshot spectral imagers. *Appl. Optics* 49, 36 (2010), 6824–6833.
- Te-Won Lee, Thomas Wachtler, and Terrence J Sejnowski. 2000. The spectral independent components of natural scenes. In *Intl. Workshop on Biologically Motivated Comp. Vision.*
- Anat Levin, Rob Fergus, Frédo Durand, and William T Freeman. 2007. Image and depth from a conventional camera with a coded aperture. ACM Trans. Graphics 26, 3 (2007), 70:1–10.
- Chengbo Li, Ting Sun, Kevin F Kelly, and Yin Zhang. 2012. A compressive sensing and unmixing scheme for hyperspectral data processing. *IEEE Trans. Image Processing* 21, 3 (2012), 1200–1210.
- Xing Lin, Yebin Liu, Jiamin Wu, and Qionghai Dai. 2014a. Spatial-spectral encoded compressive hyperspectral imaging. ACM Trans. Graphics 33, 6 (2014), 233:1–11.

- Xing Lin, Gordon Wetzstein, Yebin Liu, and Qionghai Dai. 2014b. Dual-coded compressive hyperspectral imaging. Optics Letters 39, 7 (2014), 2044–2047.
- Ankit Mohan, Ramesh Raskar, and Jack Tumblin. 2008. Agile spectrum imaging: Programmable wavelength modulation for cameras and projectors. In Comp. Graphics Forum.
- Matthew O'Toole and Kiriakos N Kutulakos. 2010. Optical computing for fast light transport analysis. ACM Trans. Graphics 29, 6 (2010), 164:1–12.
- Jiawei Ou and Fabio Pellacini. 2011. LightSlice: Matrix slice sampling for the many-lights problem. ACM Trans. Graphics 30, 6 (2011), 179:1–8.
- Zhihong Pan, Glenn Healey, Manish Prasad, and Bruce Tromberg. 2003. Face recognition in hyperspectral images. *IEEE Trans. Pattern Analysis and Machine Intelligence* 25, 12 (2003), 1552–1560.
- Jussi PS Parkkinen, J Hallikainen, and T Jaaskelainen. 1989. Characteristic spectra of Munsell colors. J. Optical Society of America A 6, 2 (1989), 318–322.
- Henri Rajbenbach, Yeshayahu Fainman, and Sing H Lee. 1987. Optical implementation of an iterative algorithm for matrix inversion. Appl. Optics 26, 6 (1987), 1024–1031.
- Ramesh Raskar, Amit Agrawal, and Jack Tumblin. 2006. Coded exposure photography: Motion deblurring using fluttered shutter. ACM Trans. Graphics 25, 3 (2006), 795– 804.
- Behnood Rasti, Johannes R Sveinsson, Magnus O Ulfarsson, and Jon Atli Benediktsson. 2013. Hyperspectral image denoising using a new linear model and sparse regularization. In *IEEE Intl. Geoscience and Remote Sensing Symposium*.
- Hoover Rueda, Henry Arguello, and Gonzalo R Arce. 2016. Compressive spectral testbed imaging system based on thin-film color-patterned filter arrays. *Appl. Optics* 55, 33 (2016), 9584–9593.
- Hoover Rueda, Henry Arguello, and Gonzalo R Arce. 2017. High-dimensional optimization of color coded apertures for compressive spectral cameras. In *European Signal Processing Conf.*
- Vishwanath Saragadam, Jian Wang, Xin Li, and Aswin Sankaranarayanan. 2017. Compressive spectral anomaly detection. In Intl. Conf. Comp. Photography (ICCP).
- Valeria Simoncini and Daniel B Szyld. 2003. Theory of inexact Krylov subspace methods and applications to scientific computing. SIAM J. Scientific Comp. 25, 2 (2003), 454– 477.
- Ting Sun and Kevin Kelly. 2009. Compressive sensing hyperspectral imager. In *Comp. Optical Sensing and Imaging.*
- Tsuyoshi Takatani, Takahito Aoto, and Yasuhiro Mukaigawa. 2017. One-shot hyperspectral imaging using faced reflectors. In Comp. Vision and Pattern Recognition (CVPR).
- Yuliya Tarabalka, Jocelyn Chanussot, and Jon Atli Benediktsson. 2010. Segmentation and classification of hyperspectral images using watershed transformation. *Pattern Recognition* 43, 7 (2010), 2367–2379.
- Zhi Tian, Hongya Ge, and Louis L Scharf. 2005. Low-complexity multiuser detection and reduced-rank Wiener filters for ultra-wideband multiple access. In Intl. Conf. Acoustics, Speech, and Signal Processing (ICASSP).
- Ashok Veeraraghavan, Ramesh Raskar, Amit Agrawal, Ankit Mohan, and Jack Tumblin. 2007. Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing. ACM Trans. Graphics 26, 3 (2007), 69:1–12.
- Bhagavatula Vijaya Kumar and David Casasent. 1981. Eigenvector determination by iterative optical methods. Appl. Optics 20, 21 (1981), 3707–3710.
- Ashwin Wagadarikar, Renu John, Rebecca Willett, and David Brady. 2008. Single disperser design for coded aperture snapshot spectral imaging. *Appl. Optics* 47, 10 (2008), B44–B51.
- Andrew E Waters, Aswin C Sankaranarayanan, and Richard Baraniuk. 2011. SpaRCS: Recovering low-rank and sparse matrices from compressive measurements. In Adv. Neural Info. Processing Systems.
- Michael E. Winter. 1999. N-FINDR: An algorithm for fast autonomous spectral endmember determination in hyperspectral data. In Proc. SPIE.
- Roberta H Yuhas, Alexander FH Goetz, and Joe W Boardman. 1992. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm. In Summaries of the Third Annual JPL Airborne Geoscience Workshop.
- Yong-Qiang Zhao and Jingxiang Yang. 2015. Hyperspectral image denoising via sparse representation and low-rank constraint. *IEEE Trans. Geoscience and Remote Sensing* 53, 1 (2015), 296–308.