MWIR Hyperspectral Imager

Jonathan M. Mooney
Rome Laboratory
Hanscom AFB, MA 01731
http://192.52.112.85/ero/eroi.htm
email: jon@max.rl.plh.af.mil

ABSTRACT

A spectral imager constructs a three dimensional (two spatial and one spectral) image from a series of two dimensional images. This paper discusses a technique for spectral imaging that multiplexes the spatial and spectral information on the focal plane, then demultiplexes the resulting imagery to obtain the spectral image. The resulting spectral image consists of 100x100 spatial pixels and 25 spectral bands. The current implementation operates over the 3-5pm band, but can easily be applied to other spectral regions. This approach to spectral imaging has high optical throughput and is robust to focal plane array (FPA) nonuniformities. A hardware description, the mathematical development and experimental results are presented.

1. INTRODUCTION

Here we investigate the applicability of spectral imaging to mid-wave infrared target detection. We are interested in spectral discrimination rather than quantitative analysis. From previous experience we know that infrared scenes consist of low contrast details on a large background pedestal, and that focal plane array (FPA) nonuniformities and low optical throughput often limit the performance of infrared spectral imagers. Our approach is to incorporate both a throughput and a multiplex advantage into a spectral imager, and at the same time address the effect of FPA nonuniformity on spectral imaging.

Typical state-of-the-art spectral imagers image the spectrum of a slit onto a 2-dimensional focal plane array; the slit is scanned over the object to create a 3-dimensional spectral image [1]. Since only a slit is imaged at a time, the out-of-slit photons are unutilized, hence these spectral imagers are inherently inefficient. Further, pixel-to-pixel nonuniformities on the FPA can corrupt the spectral imagery and reduce the ultimate performance.

The spectral imager we study images the entire scene over all of the spectral bands, and is similar to previously reported devices [2, 3, 4, 5]. Figure 1 is a schematic representation of our approach. It consists of a telescope, a direct vision prism and a camera. A direct vision prism consists of two pieces of glass that are arranged such that one wavelength passes through undeviated while other wavelengths are deviated along a line. The direct vision prism is mounted in a bearing so that it can be rotated around the optical axis. As the prism is rotated the spectral features in the image will trace out circles where the radius of the circles are wavelength dependent. To obtain a spectral image, several frames of camera data are obtained with the prism in different angular orientations, then a computer is used to extract the spectral image from the measured data. Since all available photons are utilized, we anticipate that the spectral imager performance will reflect the increase in efficiency. The actual performance, and potential multiplex advantage (or disadvantage), depend on how well the inversion algorithms perform. We will also demonstrate that the imager is robust to FPA nonuniformities.

If in Fig. 1 the field stop were replaced by a slit and the bearing prism mount were replaced by a linear scan mechanism, then the modified system would be a standard scanned-slit spectral imager. In either case the throughput is determined by the *f/#* and the field stop area. Since the field stop area is much larger for the angularly multiplexed configuration and *f/#'s are* unchanged, the angularly multiplexed configuration has higher throughput. The throughput advantage is given by the ratio of the field stop areas which is a factor of 100 when our implementation is compared to a single-pixel-width slit instrument.

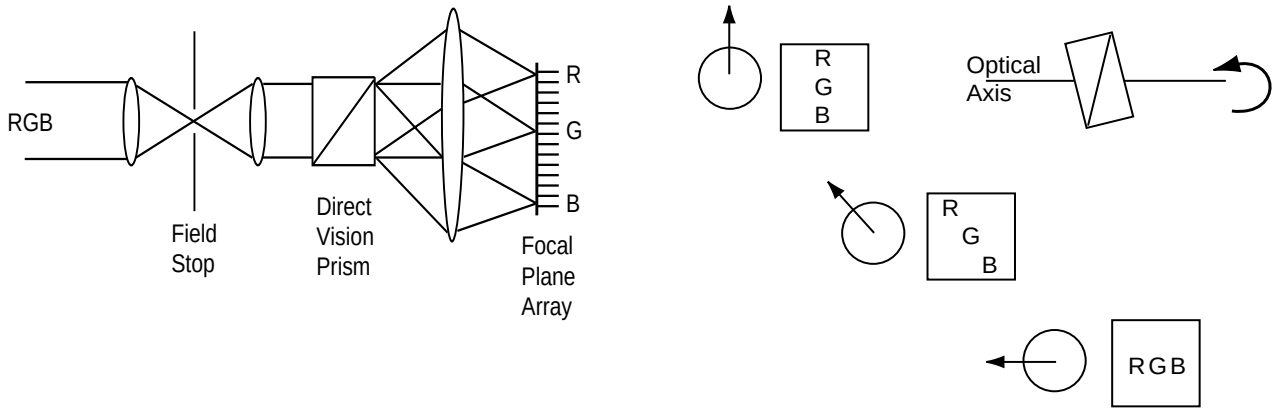

Figure 1: A schematic representation of the spectral imager and its operation. On the left the direct vision prism is shown spreading the red, green and blue light across the FPA. On the right the effect of the prism rotation on the overlapping red, green and blue images is illustrated. The circle and arrow represent the prism orientation.

In what follows we will give a description of the inversion algorithms and the hardware implementation, and illustrate the performance on measured data. We will show that our approach to spectral imaging is mathematically similar to the problem of limited angle tomography in medical imaging, and that this fact complicates both the data processing and the system analysis. These complications limit the extent to which we can completely describe the spectral imager in this paper, and expand the areas of future work. The following is an overview of our approach to spectral imaging.

2. THE INVERSION ALGORITHM

Other work in tomographic spectral imaging has relied heavily on the tomographic reconstruction techniques used in medical imaging. Here we develop an inversion technique that is appropriate for spectral imaging. We note in passing that this spectral imaging problem is very similar to that of ectomography [6, 7, 8]. In both cases all of the projections fall on the same plane and both suffer from limited viewing angle (the concept of limited viewing angle is discussed in the following section). The inversion technique developed here has the advantage of being a direct inverse.

Assume that we want to construct a three-color image of a scene $c(x, y, \lambda)$. We divide the scene into three spectral bands, $c_1(x, y), c_2(x, y)$ and $c_3(x, y)$ (e.g. red, green and blue in Fig. 1). The action of the prism-camera combination on the spectral image is to disperse each of the colors relative to one another. For example, with reference to Fig. 1, $c_2(x, y)$ (green) is undeviated, while for the first prism orientation $c_1(x, y)$ and $c_3(x, y)$ (red and blue) are deviated up and down respectively. As the prism is rotated the angular orientation of the deviation will change while its magnitude remains unchanged.

Since an image displacement can be represented as a convolution with a displaced delta function, the data that is recorded by the computer will be [9]

$$\begin{aligned} r_1(x,y) &= p_{1,1}(x,y) ** c_1(x,y) + p_{1,2}(x,y) ** c_2(x,y) + p_{1,3}(x,y) ** c_3(x,y) \\ r_2(x,y) &= p_{2,1}(x,y) ** c_1(x,y) + p_{2,2}(x,y) ** c_2(x,y) + p_{2,3}(x,y) ** c_3(x,y) \\ r_3(x,y) &= p_{3,1}(x,y) ** c_1(x,y) + p_{3,2}(x,y) ** c_2(x,y) + p_{3,3}(x,y) ** c_3(x,y), \end{aligned} \tag{1}$$

where $r_m(x, y)$ is the data recorded for frame m at location (x, y), $p_{m,n}(x, y)$ is the point spread function of spectral band n recorded on image m, and $**$ indicates a two dimensional convolution. Each of the $r_1(x, y), r_2(x, y)$ and $r_3(x, y)$ correspond to the recorded image for a different angular orientation of the direct vision prism.

Each of the $p_{m,n}(x,y)$ is a displaced delta function, and is given by

$$p_{m,n}(x,y) = \delta(x-(n-n_o)\Delta x\cos(\phi_m), y-(n-n_o)\Delta x\sin(\phi_m)), \tag{2}$$

where Δx is the pixel dimension on the FPA, n_o is the index of the undeviated color, and ϕ_m is the angular orientation of the prism on frame m. As the prism is rotated the imagery in spectral band n_o is unchanged, while the imagery from spectral band n orbits about that of n_o on a circular path with radius $(n-n_o)\Delta x$. An implicit assumption in this development is that the unique spectral bands are dispersed by integer numbers of pixels on the focal plane, and that the pixels on the focal plane are square.

We wish to solve the system of equations Eq.1 for $c_n(x,y)$. First we take the spatial two dimensional Fourier transform of each equation.

$$\begin{aligned} R_1(\xi,\zeta) &= P_{1,1}(\xi,\zeta)C_1(\xi,\zeta) + P_{1,2}(\xi,\zeta)C_2(\xi,\zeta) + P_{1,3}(\xi,\zeta)C_3(\xi,\zeta) \\ R_2(\xi,\zeta) &= P_{2,1}(\xi,\zeta)C_1(\xi,\zeta) + P_{2,2}(\xi,\zeta)C_2(\xi,\zeta) + P_{2,3}(\xi,\zeta)C_3(\xi,\zeta) \\ R_3(\xi,\zeta) &= P_{3,1}(\xi,\zeta)C_1(\xi,\zeta) + P_{3,2}(\xi,\zeta)C_2(\xi,\zeta) + P_{3,3}(\xi,\zeta)C_3(\xi,\zeta), \end{aligned} \tag{3}$$

where we use capital letters to indicate that we have Fourier transformed, and ξ and ζ are the spatial frequency variables.

The form of Eq. 3 can be reduced to

$$\begin{bmatrix} R_1(\xi,\zeta) \\ R_2(\xi,\zeta) \\ R_3(\xi,\zeta) \end{bmatrix} = \begin{bmatrix} P_{1,1}(\xi,\zeta) & P_{1,2}(\xi,\zeta) & P_{1,3}(\xi,\zeta) \\ P_{2,1}(\xi,\zeta) & P_{2,2}(\xi,\zeta) & P_{2,3}(\xi,\zeta) \\ P_{3,1}(\xi,\zeta) & P_{3,2}(\xi,\zeta) & P_{3,3}(\xi,\zeta) \end{bmatrix} \begin{bmatrix} C_1(\xi,\zeta) \\ C_2(\xi,\zeta) \\ C_3(\xi,\zeta) \end{bmatrix}, \tag{4}$$

which is simply a large array of vector-matrix multiplications

$$\vec{R}(\xi,\zeta) = \boldsymbol{P}(\xi,\zeta)\vec{C}(\xi,\zeta). \tag{5}$$

Note that our large system of equations has been block-diagonalized; we have a vector-matrix multiplication at each spatial frequency rather than one large vector-matrix multiplication. This block-diagonalization is a key computational consideration for our hardware realization and subsequent data processing. For the hardware implemented, we solve for 50 240×240 spectral images from a sequence of 80 240×240 input images. Hence, we invert 57,600 50×80 matrices rather than one 4,608,000×2,880,000 matrix. Once completed the inversion need not be repeated unless the system used to acquire the data is changed (provided the inverse is stored in some useful way).

For the continuous case each of the $P_{m,n}(\xi,\zeta)$ is given by

$$P_{m,n}(\xi,\zeta) = e^{-2\pi i(\xi\cos(\phi_m)+\zeta\sin(\phi_m))(n-n_o)\Delta x}. \tag{6}$$

$P_{m,n}(\xi,\zeta)$ can be rewritten as

$$P_{m,n}(\xi,\zeta) = P_{m,n}(\vec{\nu}) = e^{-2\pi i\vec{\nu}\cdot\vec{\rho}}, \tag{7}$$

where $\vec{\rho}$ is the vector $((n-n_o)\Delta x\cos(\phi_m),(n-n_o)\Delta x\sin(\phi_m))$. It can also be rewritten as

$$P_{m,n}(\xi,\zeta) = [e^{-2\pi i(\xi\cos(\phi_m)+\zeta\sin(\phi_m))\Delta x}]^{(n-n_o)}, \tag{8}$$

where it is evident that each $\boldsymbol{P}(\xi,\zeta)$ is Vandermonde [10]. If the algorithm is to be implemented on a computer, the expression for $P_{m,n}(\xi,\zeta)$ is complicated by the fact that $\vec{\rho}$ may not fall exactly on a pixel center. We address this issue by assuming that the four pixels nearest $\vec{\rho}$ are affected bilinearly weighted relative to their distance from $\vec{\rho}$. This is equivalent to assuming that the optical point spread function is a uniform square the same size as a pixel and that the pixels have 100% active area.

Before we address the inversion of $\boldsymbol{P}(\xi,\zeta)$ it is worth considering its symmetries. If we assume that $\phi_m = 2\pi m/M$, that $0 \le m < M$, M is divisible by 4 and $0 \le n < N$ and that m, M, n, n_o and N are integers, then the following symmetry properties hold:

$$P_{m,n}(\xi,\zeta) = \begin{array}{l} P^*_{m,2n_o-n}(\xi,\zeta) \\ P^*_{m+M/2,n}(\xi,\zeta) \\ P^*_{m,n}(-\xi,-\zeta) \\ P_{m+M/4,n}(\zeta,-\xi) \\ P_{M/2-m,n}(-\xi,\zeta) \end{array} . \tag{9}$$

These symmetry properties can be used to reduce the computer memory and computational requirements of the algorithm.

To this point we have neglected the effects of noise. If we assume that the noise is additive, then Eq. 5 becomes

$$\vec{R}(\xi,\zeta) = \boldsymbol{P}(\xi,\zeta)\vec{C}(\xi,\zeta) + \vec{N}(\xi,\zeta), \tag{10}$$

where $\vec{N}(\xi,\zeta)$ is the 2D discrete spatial Fourier transform of the additive noise vector.

In order to solve for $\vec{C}(\xi,\zeta)$ we need to invert $\boldsymbol{P}(\xi,\zeta)$. Even though efficient algorithms exist for inverting the Vandermond matrix $\boldsymbol{P}(\xi,\zeta)$, two factors lead us to use the singular value decomposition (SVD) [11]: $\boldsymbol{P}(\xi,\zeta)$ is rectangular, and for small values of ξ and ζ it will be singular.

SVD factors an M×N matrix into a product of an M×N column-orthogonal matrix $\boldsymbol{U}$, an N×N diagonal matrix $\boldsymbol{W}$ with real non-negative elements, and an N×N orthogonal matrix $\boldsymbol{V}^\dagger$,

$$\boldsymbol{P} = \boldsymbol{U}\boldsymbol{W}\boldsymbol{V}^\dagger. \tag{11}$$

Since $\boldsymbol{U}$ and $\boldsymbol{V}$ are column orthonormal, we find

$$\boldsymbol{U}^\dagger\boldsymbol{U} = \boldsymbol{V}^\dagger\boldsymbol{V} = \mathbf{1}, \tag{12}$$

where the dagger indicates Hermitian adjoint (transpose conjugate). The inverse of the diagonal matrix $\boldsymbol{W}$ is illustrated by

$$\begin{bmatrix} 1/w_{1,1} & & \\ & \ddots & \\ & & 1/w_{n,n} \end{bmatrix} \begin{bmatrix} w_{1,1} & & \\ & \ddots & \\ & & w_{n,n} \end{bmatrix} = \begin{bmatrix} & \mathbf{1} & \end{bmatrix}. \tag{13}$$

We find that for an M×N matrix $\boldsymbol{P}$ we can make a pseudoinverse $\boldsymbol{V}\boldsymbol{W}^{-1}\boldsymbol{U}^\dagger$, such that

$$\mathbf{1} = \boldsymbol{V}\boldsymbol{W}^{-1}\boldsymbol{U}^\dagger\boldsymbol{P}. \tag{14}$$

We estimate $\vec{C}(\xi,\zeta)$ by $\tilde{C}(\xi,\zeta)$ where

$$\tilde{C}(\xi,\zeta) = \boldsymbol{V}(\xi,\zeta)\boldsymbol{W}^{-1}(\xi,\zeta)\boldsymbol{U}^\dagger(\xi,\zeta)\vec{R}(\xi,\zeta), \tag{15}$$

where only the matrix $\boldsymbol{W}(\xi,\zeta)$ is potentially singular, which occurs when one of its diagonal elements is zero. In this application the noise in the recorded image data $r(x,y,\phi_m)$ is effectively divided by $w_{i,i}(\xi,\zeta)$, so if the value of any of the $w_{i,i}(\xi,\zeta)$ is small, the noise will be greatly amplified. To circumvent this problem we replace $\boldsymbol{W}^{-1}(\xi,\zeta)$ with the matrix $\tilde{\boldsymbol{W}}^{-1}(\xi,\zeta)$, which is also diagonal with elements

$$\tilde{w}_{i,i}(\xi,\zeta) = \frac{w_{i,i}(\xi,\zeta)}{w^2_{i,i}(\xi,\zeta) + \epsilon^2}, \tag{16}$$

where the optimum value of ϵ is data, noise and spatial frequency dependent.

The effect of singular matrices can be further investigated by substituting for $\vec{R}(\xi,\zeta)$ in Eq. 15; we also drop the explicit (ξ,ζ) reference to simplify the expression.

$$\tilde{C} = \boldsymbol{V}[\tilde{\boldsymbol{W}}^{-1}\boldsymbol{W}\boldsymbol{V}^{\dagger}\vec{C} + \tilde{\boldsymbol{W}}^{-1}\boldsymbol{U}^{\dagger}\vec{N}]. \tag{17}$$

The columns of $\boldsymbol{V}$ that correspond to the small values of $w_{i,i}$ define the null space of our imaging system. From Eq. 17 we see that in the null space the scene information is lost and the noise is nulled. The next two sections analyze the null space and describe our treatment of it.

Before we examine the null space we introduce a notational variation that will be useful later. From the proceeding development, we find it possible to summarize Eq. 1 (with additive noise) as

$$\vec{r} = \mathcal{F}^{-1}\mathcal{P}\mathcal{F}\vec{c} + \vec{n}, \tag{18}$$

where $\mathcal{F}$ is the 2D Fourier transform operator for the lexicographically ordered inputs $\vec{c}$ and outputs $\vec{r}$, and $\mathcal{P}$ is the block-diagonalized assembly of matrices $\boldsymbol{P}$. We can extend Eq. 18 to include the SVD of $\boldsymbol{P}$

$$\vec{r} = \mathcal{F}^{-1}\mathcal{U}\mathcal{W}\mathcal{V}^{\dagger}\mathcal{F}\vec{c} + \vec{n}. \tag{19}$$

We will find Eq. 19 useful in our discussion of filling in the missing cone.

3. THE CONE OF MISSING INFORMATION

In general, when one of the $w_{i,i}(\xi,\zeta)$ values becomes too small, there is insufficient information to determine how the measured signal should be distributed among the spectral bands; information has been lost. Both Descour [4] and Bernhardt [5] have demonstrated that the missing information is similar to the missing cone in some medical imaging techniques [12] (the illustrations in Ref. [4] are particularly illuminating).

The origin of the missing cone is best understood in the context of the Radon transform combined with the central slice theorem [13]. From the central slice theorem, we know that the 2D Fourier transform of a projection is equal to a slice through the origin of the 3D Fourier transform of the spectral image. By rotating the prism, we obtain different slices through the origin each at a different angle about the optical axis, but with the same angle with respect to the optical axis. We can use the slices to reconstruct the object in the Fourier domain. In the limit of an infinite number of projections, the projections will define a cone. Since none of the projections provide any information about the image inside this cone, it is referred to as a 'missing cone'.

Figure 2 illustrates that the image through the direct vision prism is analogous to a projection through the pseudo-volume defined by two spatial and one spectral dimension. From Fig. 2 we can also identify the cone (triangle in the figure) of missing information. To estimate the impact of the missing cone on our imaging system, we calculate the fractional volume of the missing cone; which is a consequence of the projection angle θ. Our interpretation of the projection angle is governed by two assumptions: 1) adjacent spectral bands are displaced by one pixel on the focal plane, and 2) the volume object elements (voxels) are assigned a cubic shape. The first assumption is easily justified since it is consistent with scanned slit configurations. The second assumption is completely arbitrary, but has the benefit that the Fourier transform of the object will be a cube. If we accept these two assumptions, then it follows that the missing cone half angle is 45 degrees.

Another perspective of the missing cone is obtained by analyzing the singular values of $\mathcal{P}$. When the singular values are small, information is lost. We expect the small singular values to define a conical shape. Figure 3 is a graphical representation of the values of $w_{i,i}(\xi,\zeta)$ illustrating where in the frequency domain information is lost. The dark regions in Fig. 3 are due to the missing cone. While the physical interpretation is identical to that in [12], the mathematical treatment here differs from that commonly found in the literature since this treatment uses a hybrid of Fourier and SVD determined basis functions. Figure 3 also supports our conclusion that the cone angle is 90 degrees: at the highest chromatic frequency the cone diameter covers +/- Nyquist, as we expect for a 90 degree cone angle.

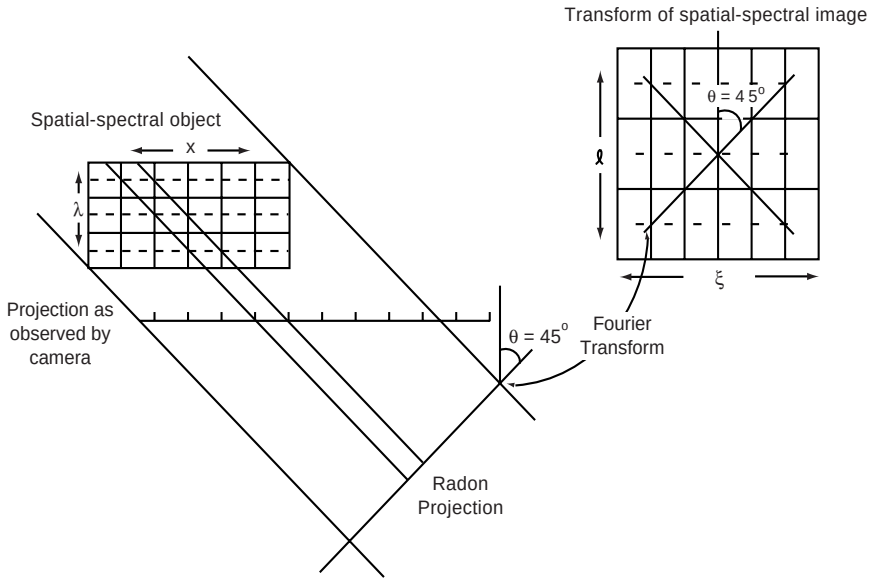

Figure 2: The correspondence of our recorded data to Radon projections. The projection angle follows from the assumptions that adjacent spectral bands are displaced by one pixel on the FPA and that the object volume elements are equilateral (square in the figure). The spectral image can be reconstructed in the frequency domain by using the central slice theorem. The projection angle defines the half-angle of the missing cone.

An intuitive understanding of the missing cone can by obtained by considering the sequence of frames recorded for a monochromatic source of infinite extent that is perfectly uniform. Since the source is uniform and of infinite extent, the sequence of frames will be unaffected by the rotation of the prism (if we consider a system with no field stop). The same sequence of recorded frames would be recorded for any wavelength monochromatic source or for any broad-band extended source that was perfectly uniform. Therefore, there is no way to algorithmically distinguish the spectral content of a spatially uniform source, and the zero spatial frequency information is lost.

Next consider a system with a slit field stop one pixel wide and with the prism dispersion oriented normal to the slit. Since the field stop is only one pixel wide, spatial and spectral information do not overlap and the slit object is easily reconstructed from one projection without a missing cone. We could also use our matrix formalism to reconstruct the object; however, in this case the reconstruction would suffer from a missing cone. We are left to resolve why the missing cone is present with one reconstruction and not with the other, from what appears to be the same set of data.

The difference arises because in the former case the information about the field stop is used in the inversion while in the latter it is not. The field stop information can be introduced into the central slice reconstruction by convolving the Fourier transform of the field stop with the projection data. The slice goes from being infinitely thin to having a finite thickness. For an infinitely thin slit the slice becomes infinitely thick, and the missing cone is filled, or for a discrete system, a slit one pixel wide would completely fill the missing cone.

The field stop $\mathcal{S}$ could be included in the matrix formalism as

$$\vec{r} = \mathcal{F}^{-1}\mathcal{U}\mathcal{W}\mathcal{V}^{\dagger}\mathcal{F}\mathcal{S}\vec{c} + \vec{n}. \tag{20}$$

where $\mathcal{S}$ defines the field stop and is a square diagonal matrix with zeros and 1s on the diagonal. Since both $\mathcal{W}$ and $\mathcal{S}$ are singular, the pseudoinverse must treat them together. This fact destroys the block diagonal structure of the matrix and makes direct inversion of the matrix intractable for most applications. However, in the next

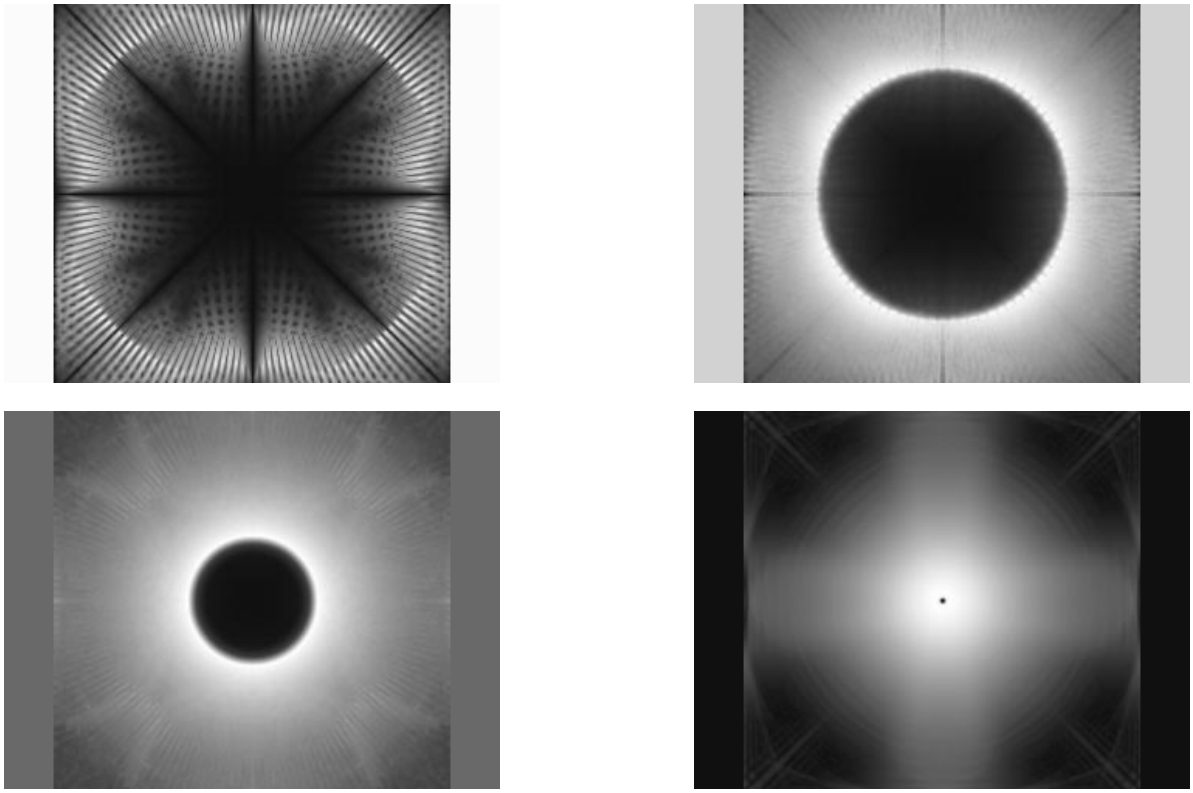

Figure 3: Illustration of the $w_{i,i}(\xi, \zeta)$ values as a function of spatial frequency. Each of these images corresponds to a horizontal slice through the 3-dimensional representation of the spatial-spectral transform image on the right side of Fig. 2. Zero spatial frequency is in the center of each image. Bright and dark regions correspond to large and small values of $w_{i,i}(\xi, \zeta)$ respectively. The sequence from left to right, top to bottom corresponds to descending spectral frequency data. The lower right image corresponds to image features that change slowly as a function of wavelength (low chromatic frequencies). The upper left image corresponds to image features that change rapidly as a function of wavelength (high chromatic frequencies). The dark region in the center of each image corresponds to the spatial-spectral frequencies where information has been lost.

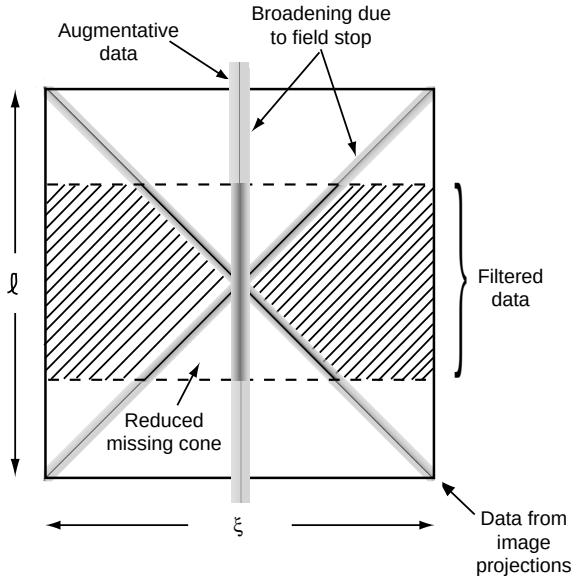

Figure 4: A schematic representation of how low-pass filtering, augmenting the data, and utilizing the field stop prior knowledge can be used to reduce the fractional volume of the missing cone. Low-pass filtering reduces the chromatic resolution, but preferentially filters out missing information, the augmentative data fills in the zero spatial-frequency axis information, and the field stop prior knowledge broadens all of the data samples.

section we will show that an iterative technique can be used to include the field stop information into the reconstructed image without loosing the block diagonal structure of the matrix.

4. FILLING IN THE MISSING CONE

The missing cone occupies 26% of the object volume in the frequency domain after the direct inverse. Clearly this is an unacceptably large fraction of missing information. In this section we discuss several methods for reducing the fraction of missing information.

The first and easiest way to reduce the missing volume is to filter the data [7]. Figure 4 illustrates the effect of reducing the chromatic resolution by a factor of 2. In this case a rectangle function filter is used to pass only the low chromatic frequency data (the figures and discussion in Ref. [7] are illuminating). The effect of this filter is to reduce the fraction of missing information to 6.5%.

The second approach is to use the field stop to fill in the missing cone. As we discussed in the previous section, direct methods for including the field stop information fail due to the large computational and memory requirements of inverting the matrix $\mathcal{WVFS}$ directly. We avoid this problem by using an iterative method similar to that in Ref. [14]. The idea is to find a solution that is consistent with both our measured data and with our knowledge of the field stop. We begin by constructing a constraint $\mathcal{K}$ such that $\mathcal{K}\tilde{c}$ is an altered version of our estimate $\tilde{c}$ that is consistent with our knowledge of the field stop. Then we transform $\mathcal{K}\tilde{c}$ into the frequency domain and replace the data values in the missing cone with the values as modified by the constraint. Then we repeat the process until some halt criterion is reached. In general the operator $\mathcal{K}$ need not be linear so that we can utilize constraints such as non-negativity.

The initial inverse gives the result

$$\tilde{c}_0 = \mathcal{F}^{-1}\mathcal{V}\tilde{\mathcal{W}}^{-1}\mathcal{U}^\dagger\mathcal{F}\vec{r}. \tag{21}$$

Subsequent iterations use the preceding results and the constraint to fill in the missing cone

$$\tilde{c}_j = \mathcal{F}^{-1}\mathcal{V}[\mathcal{L}\tilde{\mathcal{W}}^{-1}\mathcal{U}^\dagger\mathcal{F}\vec{r} + \mathcal{R}\mathcal{V}^\dagger\mathcal{F}\mathcal{K}\tilde{c}_{j-1}], \tag{22}$$

where $\mathcal{L}$ and $\mathcal{R}$ are generalizations of $\boldsymbol{L}(\xi,\zeta)$ and $\boldsymbol{R}(\xi,\zeta)$ which are diagonal with elements

$$L_{i,i}(\xi,\zeta) = \frac{w_{i,i}^2(\xi,\zeta)}{w_{i,i}^2(\xi,\zeta) + \beta^2}, \tag{23}$$

and

$$R_{i,i}(\xi,\zeta) = 1 - L_{i,i}(\xi,\zeta), \tag{24}$$

and β is a constant similar to ϵ. From Eqs. 23 and 24 we see that $\mathcal{L}$ and $\mathcal{R}$ function as a switch selecting between the left and right term in Eq. 22: large values of $w_{i,i}(\xi,\zeta)$ select the left term (measured data selected outside the missing cone), while small values of $w_{i,i}(\xi,\zeta)$ select the right term (constrained solutions selected inside the missing cone). In Fig. 4 we see that the effect of the field stop is to broaden the data samples.

The third way to fill in the missing cone is to augment the data set with an alternative spectral technique. In particular we have found it useful to re-image the entrance pupil onto the field stop, and extract the on-axis component of the missing cone from the step response at the edge of the field stop. This additional spectral information can be incorporated into the constraint $\mathcal{K}$ in Eq. 22. When combined with the field stop constraint, the information on-axis is broadened, and our image reconstruction very rapidly converges.

Bernhardt [5] has proposed a very clever technique for filling in the missing cone that utilizes a zoom lens. We did not use Bernhardt's technique because an infrared zoom lens with sufficient variation in power was not available. We believe that some relief on the zoom lens requirements can be achieved by incorporating the methods described immediately above, but a demonstration of that has been left for future work.

An additional technique for filling in the missing cone will be introduced in Section 8 when we discuss the principal component analysis of our data. We have not yet investigated its utility and postpone its discussion until after we have introduced principal component analysis. Investigation of additional methods for reducing the fractional volume of the missing cone is an area of continuing research.

5. IMPLEMENTATION

We verified the operation of the spectral imager using a PtSi infrared camera imaging over the 3-5μm band. The camera operated at $f/2.7$ and utilized a 320×244 FPA obtained from FLIR Systems. We chose this focal plane because it was non-interlaced and had a fill factor of over 80%. A Diversified Optical Products infrared direct vision prism was used as the disperser. The IR camera, drive electronics and 12-bit/pixel real-time consecutive frame data acquisition electronics were all built in-house.

The strong dependence of PtSi quantum efficiency on wavelength $\eta(\lambda)$ has both negative and positive consequences. While complicating calibration of the incident photon flux, it compensates for the large imbalance between the photon flux at short and long wavelengths. This compensated balance simplifies measurement of object emissivity, and makes PtSi an appealing candidate for spectral imaging applications where measurement of object emissivity rather than measurement of incident flux is required.

The dispersion of the prism, the focal length of the final lens and the band pass of the camera established that a broad-band point source was spread over 50 pixels ($N = 50$). In the current implementation 80 prism orientations ϕ_m are uniformly spaced over 2π ($M = 80$).

A field stop 100×100 pixels on a side was mounted at the real image in the telescope. In this infrared imaging application the field stop is warm and emissive, so it provides the prior knowledge that it is uniform, but its intensity is non-zero and the spectral distribution is unknown. Since we are most interested in spectral discrimination we chose to measure the scene spectrum relative to the field stop. We could use this instrument radiometrically by calibrating the field stop, but as yet have avoided this complication.

The format of the FPA dictated the size of the discrete Fourier transform. We cropped the two top and bottom rows and the 40 left and right columns to implement a 240×240 Fourier transform using the Good-Thomas method described in Ref. [15].

The software to invert the data was written in-house in C for Solaris. The matrix decomposition was performed once for a given configuration and the result stored on hard disk. Subsequent inversions accessed the data on disk. The size of the file was minimized by utilizing the symmetry properties from Eq. 9. In effect the decomposition of only one octnt need be stored. The symmetry was also used to reduce the data from complex to real. The precision of the stored data was short integer (16 bits). The size of the file for the current configuration was approximately 100 Mbytes. We believe that an additional reduction by a factor of 2 is possible, but have left it for future work.

The operation of the instrument consists of: 1) digitally recording the 80 projection images, 2) digitally recording diffused image data with the prism oriented normal to one edge of the field stop (supplementary data used to fill in the missing cone), 3) performing the direct inverse, 4) performing the iterative refinement calculations. For a camera operating at 30 frames/sec. the data acquisition is completed in less than 3 seconds. The data processing is completed in under 5 minutes on a Pentium Pro 200 computer.

6. THE EFFECT OF FPA NONUNIFORMITY

Before we demonstrate the operation of the spectral imager, we would like to highlight one of its most important characteristics (especially in infrared imaging). Infrared imagery is characterized by low contrast details on a large background pedestal. Pixel to pixel nonuniformities on the FPA interact with the large background pedestal to create a spatial noise pattern in the imagery. This effect is well documented [16] and played a large role in the evolution of infrared imagers. Therefore, we expect FPA nonuniformity to play a large role in the development of infrared spectral imagers, and must consider the effect of FPA nonuniformity on our measured data.

As the prism rotates the nonuniformity is spatially fixed on the FPA and coincides with the undeviated spectral band (also fixed). Additive nonuniformities will be entirely consistent with image features in the undeviated spectral band, and will be treated as such by the inversion algorithm. Therefore, the undeviated spectral band will be corrupted by FPA additive nonuniformities, while the other spectral bands will not. The system could be designed such that the corrupted band coincides with an atmospheric absorption band, or some other band of little interest in order to minimize their impact on spectral discrimination.

The effect of multiplicative nonuniformities is dominated by their interaction with the large background pedestal. Since the background pedestal is independent of prism rotation, the bulk of multiplicative nonuniformities will again be mapped to the undeviated spectral band by the inversion algorithm, and a similar argument can be made for the spectral nonuniformity interaction with the background pedestal. We find that the vast majority of nonuniformity effects appear in the undeviated spectral band.

The fraction of the multiplicative and higher order nonuniformities that interacts with low contrast image details, the higher frequency component of the 1/f noise (drift that occurs during data acquisition), and the interaction of the spectral nonuniformities with the low contrast details will all appear as noise. These higher order terms are almost always negligible.

The inherent robustness of this technique to FPA nonuniformity and its high throughput are the two major

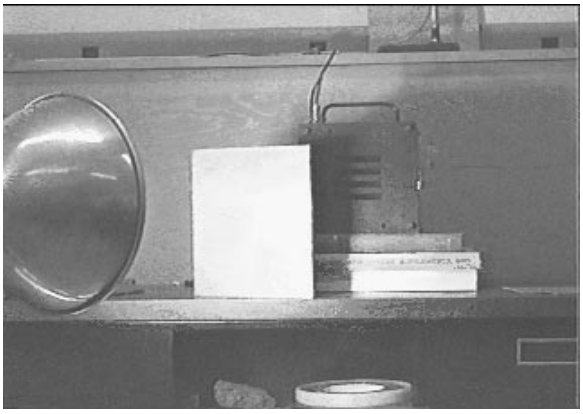

Figure 5: Visible image of the scene used for the demonstration. The reflective plate and differential source are illuminated by an incandescent lamp. Here the lamp intensity is much greater than when the experiment was performed so that the reflected light can be seen.

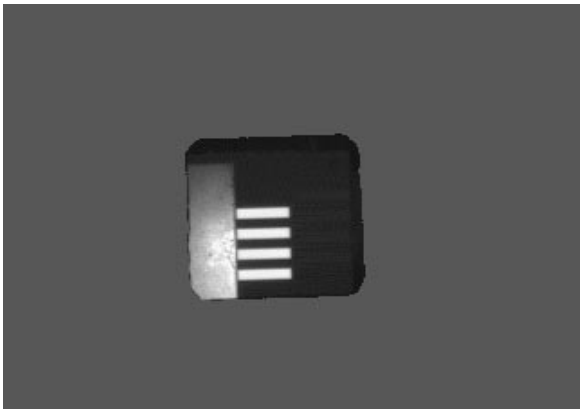

Figure 6: Broad-band infrared image of the scene used for the demonstration. Here the field stop limits the field of view to the central 100×100 pixels.

technical reasons for considering this spectral imaging approach.

7. DEMONSTRATION

To demonstrate the operation of the spectral imager we imaged a laboratory scene that contained objects with different spectral signatures. We used a controlled laboratory scene so that we could investigate the effect of the missing cone on spectral reconstruction and spectral discrimination. We note in passing that the standard technique of imaging the exit slit of a monochrometer to calibrate the spectral imager will give optimistic results since a slit object has a lot of high spatial-frequency energy that falls outside the missing cone. Our demonstration scene consists of extended objects, where a large fraction of the scene energy falls inside the missing cone, creating more of a challenge for our reconstruction algorithm.

A visible image of the scene is shown in Fig. 5. It consists of a differential source and a reflective plate. The differential source consists of two plates one in front of the other. The front plate has a four-bar pattern of holes so that the back plate can be seen through the holes. The back plate is held at a temperature elevated above that of the front plate so that in the infrared the source appears as four bright bars. The reflective plate was illuminated with an incandescent lamp that was controlled by a variable autotransformer. The lamp voltage was adjusted so that in the broad-band infrared image the signal from the reflective plate matched that of the differential source. A broad-band infrared image of the scene is shown in Fig. 6. We expect three major spectral signatures (the background, the hot bars and the reflective plate) and hope to use the differences between the spectral profiles of the plate and the bars to distinguish between them.

Figures 7 and 8 illustrate one of the projections and one of the spectral reconstructions. The spectral profiles of a pixel in the plate and of a pixel in a bar are shown in Fig. 9. Since the illumination of the plate is intended to simulate solar illumination, the plate profile is labeled 'Solar'. We see that the 'Solar' profile is larger at short wavelengths while the thermal profile is larger at long wavelengths, and that the profiles cross at approximately 4.1 μm. The spectral notch at 4.27 μm is due to absorption by CO_2 in the atmosphere.

We implement the spectral discrimination by principal component analysis. Each pixel's spectral profile is used to form the rows of a matrix which we decompose using SVD. The decomposition yields singular values, spectral eigenvectors and spatial eigenvectors. Since our demonstration scene has three components (background,

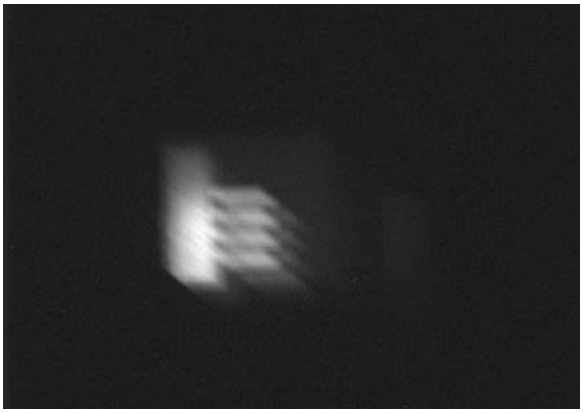

Figure 7: One of the 80 projections. Here the blurring is due to the prism dispersion.

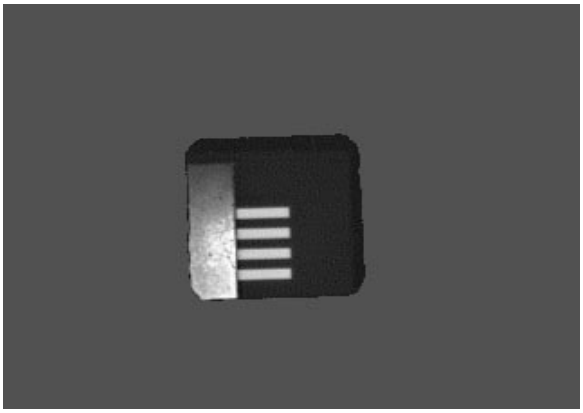

Figure 8: One of the 50 spectral reconstructions before filtering. This reconstruction corresponds to 3.8 μm.

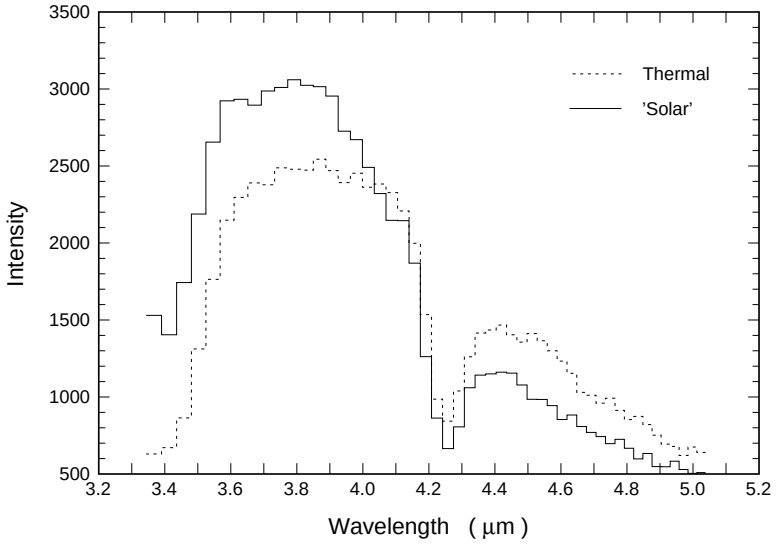

Figure 9: A comparison of single pixel spectral profiles from the reflective plate and a 'hot' bar prior to filtering. The spectral profiles shown are measured relative to the field stop, which we have assumed to be spectrally flat.

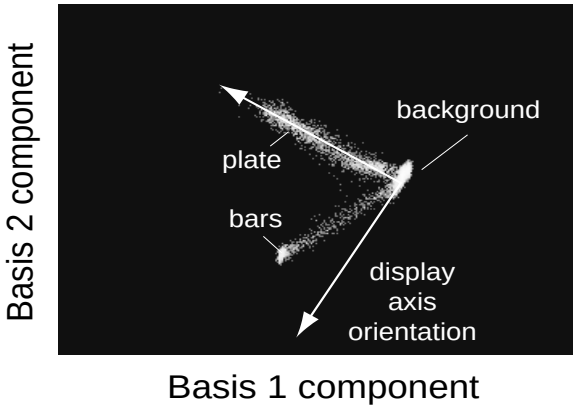

Figure 10: Scatter plot of the spatial-spectral image data using the two leading terms of the singular value decomposition. The indicated display axis corresponds to the left image in Fig. 11.

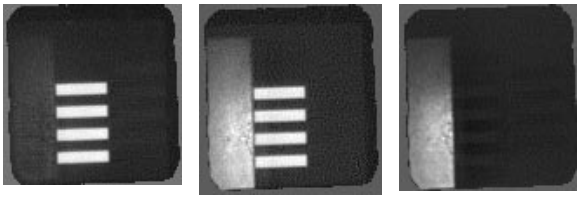

Figure 11: Spatial display of the spectrally discriminated objects. The center image is a broadband image. On the left the display axis is oriented to map the plate and background to the same level; and therefore accentuate the bars. On the right the plate is accentuated.

plate, and bars), two of each type eigenvector should be enough to describe the spectral image; the remaining eigenvectors describe the noise.

Once decomposed, each pixel in our demonstration has two statistically meaningful coefficients, one from each of the spatial eigenvectors. In Fig. 10 each pixel is represented by a bright point located using the two coefficients as cartesian co-ordinates to form a scatter plot. In the figure we see that the pixels associated with the background, plate, and bars are distinguishable from one another. The pixels that lie between the bars group and the background group are pixels that fall on a bar edge. The fraction of the pixel that is filled by the bar can be determined from its location relative to the two groups.

In Fig. 11 we have linearly combined the two spatial eigenvectors to accent the bars or the plate by rotating the display axis relative to the two orthogonal bases. The left image in Fig. 11 corresponds to the display axis as shown in Fig. 10. Both Fig. 10 and Fig. 11 show that this approach to spectral imaging is effective as a spectral discriminant.

8. ANALYSIS OF THE ITERATIVE ENHANCEMENTS

Unfortunately, we are as yet unable to quantify the performance of the spectral imager. While the high optical throughput gives us reason to be optimistic, the extent to which the missing cone will degrade performance is not clear. The nature of the missing cone implies that the performance will be data dependent, and our techniques for filling in the missing cone make the data dependence difficult to identify. In this section we address the effect of our fill-in techniques on the fidelity of our reconstructed spectral imagery.

Since the demonstration scene was designed to create data that can be represented by two eigenvectors, analysis of the remaining eigenvectors provides information about the effect of the missing cone on the reconstructed image. Further, by analyzing the singular values as a function of the iteration number we can estimate how things are improving as we iterate.

Figure 12 is a plot of five of the singular values as a function of number of iterations. The scene information is represented by the largest two singular values, the missing cone artifacts are represented by the third and lower singular values, and the noise is present in all of the eigenvectors, and becomes dominant for the smaller

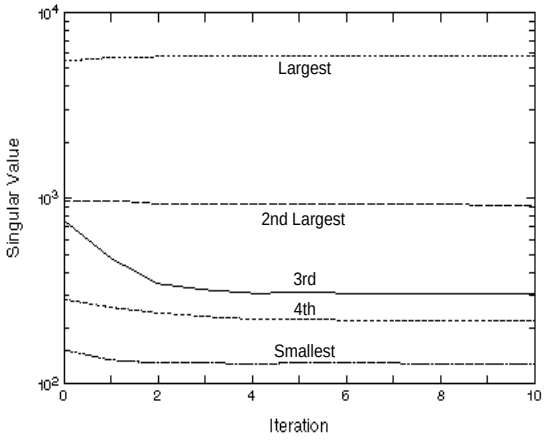

Figure 12: The four largest and smallest singular values as a function of number of iterations. The decline of the third largest singular value for the first few iterations indicates the effectiveness of the iterative techniques for filling the missing cone.

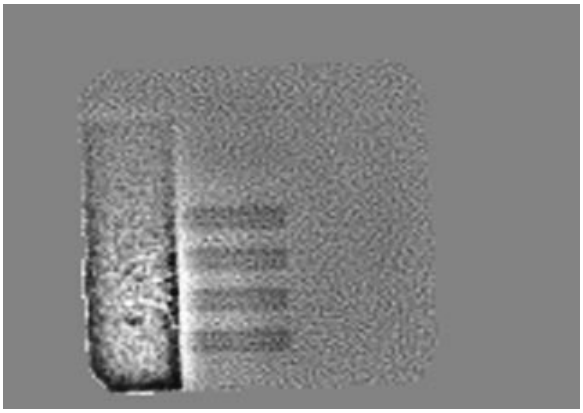

Figure 13: An enlarged view of the spatial eigenvector corresponding to the third largest singular value (artifact due to the missing cone) after 5 iterations. The ringing at the plate edges is indicative of lost information.

singular values. While the separation of the scene information from the missing cone artifacts is probably not strictly true, it appears to be a good approximation.

In Fig. 12 we see that the third largest singular value is most affected by the iterative process; its value drops by almost a factor of 3 after 4. The spatial eigenvector corresponding to the third largest singular value is shown in Fig. 13. The hypothesis that Fig. 13 represents a missing cone artifact is supported by the fact that the dark edge around the bottom part of the plate and the bright areas between the bars are inconsistent with the physical attributes of the scene.

From this analysis we conclude that the iterative procedure has reduced the effect of the missing cone by a factor of nearly 3. Given that the filtering had reduced the fractional missing cone volume to 6.5%, our estimate of the post-iteration missing fractional volume is just over 2% for this particular scene.

We could use this approach to further reduce the missing cone artifacts for this particular scene by eliminating the third and higher eigenvectors from reconstruction altogether, but then we could no longer use the principal component analysis to evaluate the results. It is possible that the SVD analysis of the iterative procedure could be used to separate those principal components that are iteration independent from those that are iteration dependent to eliminate missing cone artifacts for a general scene. We have not yet found a way of evaluating the approach, so it remains for future work.

While the iterative techniques are effective, there is still room for improvement. We hope to expand and refine the techniques used to eliminate the missing cone artifacts in future work. It is important to remember that while scenes that are spatially very uniform will be poorly reconstructed, analysis of spatially uniform scenes does not require imaging techniques (due to the lack of spatial structure), and scenes with spatial structure will be less susceptible to the missing cone.

9. CONCLUSIONS

We have described and demonstrated a high throughput approach to spectral imaging that is also robust to FPA nonuniformities. We demonstrated our approach in the mid-infrared, but it is generally applicable to any spectral band where 2-dimensional imaging arrays and direct vision prisms are available.

The high throughput and robustness to FPA nonuniformity mean that this approach can be used with nonuniform high-quantum efficiency FPAs to minimize the amount of time required to collect high fidelity hyperspectral image data.

We have also shown that this approach is subject to a cone of missing information. In particular, the low spatial-frequency/high spectral-frequency information is missing from the resulting spectral imagery which leads to limited degradation of the image quality. However; we demonstrated that by filtering and using iterative techniques, the fractional volume of missing information could be reduced to approximately 2%. In the process of analyzing our data, we demonstrated that principal component analysis is a very useful tool for identifying artifacts in image sequences.

For scenes with substantial spatial structure, this is a promising approach to spectral imaging.

10. ACKNOWLEDGMENTS

The authors thank Bill Ewing and Steve DiSalvo for fabricating the infrared camera; John Mead for fabricating the data acquisition interface; Linda Bouthillette for graphical assistance; and Jerry Silverman, Paul Pellegrini, Al Kahan and Freeman Shepherd for critical reading of this manuscript. This work was supported by AFOSR Task 2305CR under Dr. Gerry Witt.

11. REFERENCES

[1] R. Basedow, D. Carmer, and M. Anderson, "Hydice systems, implementation and performance," *Proc. SPIE*, vol. 2480, pp. 258–267, (1995).

[2] G. G. Levin and G. N. Vishnyakov, "On the possibilities of chronotomography of high speed processes," *Optics Communications*, vol. 56, pp. 231–234, December (1985).

[3] T. Okamoto, A. Takahashi, and I. Yamaguchi, "Simultaneous acquisition of spectral and spatial intensity distribution," *Applied Spectroscopy*, vol. 47, pp. 1198–1202, August (1993).

[4] M. Descour and E. Dereniak, "Computed-tomography imaging spectrometer: experimental calibration and reconstruction results," *Applied Optics*, vol. 34, pp. 4817–4826, August (1995).

[5] P. A. Bernhardt, "Direct reconstruction methods for hyperspectral imaging with rotational spectrotomography," *J. Opt. Soc. Am. A*, vol. 12, pp. 1884–1901, September (1995).

[6] P. Edholm, G. Granlund, H. Knutsson, and C. Petersson, "Ectomography - a new radiographic method for reproducing a selected slice of varying thickness," *ACTA Radiologica-diagnosis*, vol. 21, pp. 433–442, (1980).

[7] H. Knutsson, P. Edholm, G. Granlund, and C. Petersson, "Ectomography - a new radiographic reconstruction method - I. Theory and error estimates," *IEEE Trans. Biomed. Eng.*, vol. BME-27, pp. 640–648, (1980).

[8] C. Petersson, P. Edholm, G. Granlund, and H. Knutsson, "Ectomography - a new radiographic reconstruction method - II. Computer simulated experiments," *IEEE Trans. Biomed. Eng.*, vol. BME-27, pp. 649–655, (1980).

[9] J. M. Mooney, "Spectral imaging via computed tomography," *Proc. 1994 Meeting of the IRIS specialty group on passive sensors*, vol. 1, pp. 203–215, March (1994).

[10] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, pp. 52 – 54. New York: Cambridge University Press, (1988).

[11] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, pp. 60 – 72. New York: Cambridge University Press, (1988).

[12] H. H. Barrett, "Editorial: Limited-angle tomography for the nineties," *The Journal of Nuclear Medicine*, vol. 31, pp. 1688–1692, October (1990).

[13] H. H. Barrett and W. Swindell, *Radiological Imaging*, vol. 2, pp. 375 – 446. New York: Academic Press, (1981).

[14] A. Papoulis, "A new algorithm in spectral analysis and band-limited extrapolation," *IEEE Trans. Circuits and Sys.*, vol. CAS-22, pp. 735–742, September (1975).

[15] R. Tolimieri, M. An, and C. Lu, *Algorithms for Discrete Fourier Transform and Convolution.* New York: Springer-Verlag, (1989).

[16] J. Mooney, F. Shepherd, W. Ewing, J. Murguia, and J. Silverman, "Responsivity nonuniformity limited performance of infrared staring cameras," *Optical Engineering*, vol. 28, pp. 1151–1161, November (1989).