Nonnegative Matrix Factorization for Efficient Hyperspectral Image Projection

Alexander S. Iacchetta[a], James R. Fienup[a], David T. Leisawitz[b], Matthew R. Bolcar[b]
[a]Institute of Optics, Univ. of Rochester, 275 Hutchison Rd., Rochester, NY, USA 14627-0186
[b]NASA Goddard Space Flight Center, 8800 Greenbelt Rd., Greenbelt, MD, USA 20771-2400

ABSTRACT

Hyperspectral imaging for remote sensing has prompted development of hyperspectral image projectors that can be used to characterize hyperspectral imaging cameras and techniques in the lab. One such emerging astronomical hyperspectral imaging technique is wide-field double-Fourier interferometry. NASA's current, state-of-the-art, Wide-field Imaging Interferometry Testbed (WIIT) uses a Calibrated Hyperspectral Image Projector (CHIP) to generate test scenes and provide a more complete understanding of wide-field double-Fourier interferometry. Given enough time, the CHIP is capable of projecting scenes with astronomically realistic spatial and spectral complexity. However, this would require a very lengthy data collection process. For accurate but time-efficient projection of complicated hyperspectral images with the CHIP, the field must be decomposed both spectrally and spatially in a way that provides a favorable trade-off between accurately projecting the hyperspectral image and the time required for data collection. We apply nonnegative matrix factorization (NMF) to decompose hyperspectral astronomical datacubes into eigenspectra and eigenimages that allow time-efficient projection with the CHIP. Included is a brief analysis of NMF parameters that affect accuracy, including the number of eigenspectra and eigenimages used to approximate the hyperspectral image to be projected. For the chosen field, the normalized mean squared synthesis error is under 0.01 with just 8 eigenspectra. NMF of hyperspectral astronomical fields better utilizes the CHIP's capabilities, providing time-efficient and accurate representations of astronomical scenes to be imaged with the WIIT.

Keywords: nonnegative matrix factorization, wide-field imaging interferometry testbed, hyperspectral image projector, spatio-spectral, double-Fourier, NMF, WIIT, CHIP

1. INTRODUCTION

Hyperspectral imaging is an important modality for the identification and classification of objects and materials within a scene. In particular, the astrophysics community relies on hyperspectral measurements to make inferences about, and develop models of, distant astronomical objects. Although ground-based interferometric measurements are in widespread use among astrophysicists, single-aperture telescopes remain the main source of space-based hyperspectral measurements. There will come a time, however, when a monolithic aperture will be unable to meet the demands for high spatial resolution astronomical imagery due to excessive cost, weight, and size, especially at infrared wavelengths. This is when space-based observatories will likely employ wide-field spatio-spectral, or double-Fourier, interferometry. The mathematical framework for and fairly detailed descriptions of a spatio-spectral interferometer have already been published[1-5], so the following will be only a brief introduction into double-Fourier interferometric imaging.

Spatio-spectral interferometry is an extension of Fourier transform imaging spectroscopy[6] (FTIS) using aperture synthesis to obtain higher spatial resolution than with conventional FTIS. In FTIS, an input beam is split into two arms of an interferometer before being recombined and focused onto an array detector with an imaging lens. One arm of the interferometer has a fixed optical path, while the other has a scanning mirror that can vary the optical path difference (OPD) between the two arms of the interferometer, resulting in a measured datacube consisting of two spatial dimensions and one delay dimension. A hyperspectral image cube is then obtained by taking the Fourier transform over the delay dimension of the datacube. Spatio-spectral interferometric imaging, however, requires two apertures with a vector separation, called the baseline, instead of a single aperture followed by a beam splitter. The optical path following one of the mirrors is kept fixed while the other can vary, resulting in a measured datacube for each baseline separation. Taking the Fourier transform of these datacubes over the delay dimension results in a set of high-pass filtered images, where the passband in the spatial frequency domain is related to both the baseline and wavelength[7]. A single high-

spatial-resolution hyperspectral image can be recovered from the set of high-pass filtered images using an image synthesis algorithm, such as the algorithm developed by Lyon *et al.*[4,5].

Although the theoretical groundwork for double-Fourier interferometry has already been established[1-5], the space-based spatio-spectral interferometric imaging technique needs further characterization before an interferometric observatory, such as the NASA proposed Space Infrared Interferometric Telescope (SPIRIT)[8], ever becomes a reality. This prompted NASA to build the Wide-field Imaging Interferometry Testbed (WIIT)[9-20]. The WIIT was developed by NASA to further the maturation of wide-field spatio-spectral interferometry in order to meet the demands of future astronomical hyperspectral imaging missions. The testbed is a scale model of a space-based wide-field spatio-spectral interferometer, but operating at visible wavelengths, for which measurements are limited primarily by detector photon noise. The light source for the interferometer, which has been integrated into the WIIT, is a Calibrated Hyperspectral Image Projector (CHIP)[20].

Hyperspectral image projectors (HIPs) were developed by the Optical Technologies Division at the National Institute of Standards and Technology to solve the problem of testing and characterizing hyperspectral imaging techniques with known hyperspectral scenes[21-23]. These projectors allow for a scene to be displayed such that every object in the scene has the same arbitrary spectrum, resulting in a spatially-spectrally separable image at any given time. This is achieved using two digital light processing (DLP) units. One of the DLP units determines the gray-scale spatial distribution of the image being projected. The other unit controls the shape of the arbitrary spectrum, which is achieved by spectrally dispersing a broadband source onto the DLP such that each row of the DLP chip controls the relative strength of each wavelength bin of the output spectrum. The number of wavelength bins and the spectral range together determine the spectral resolution of the HIP's spectral output. A hyperspectral image can then be simulated with a HIP by cycling through multiple spatially-spectrally separable images that can be added together throughout the integration time of the camera to simulate the measurement of a spatially-spectrally complicated hyperspectral image. A calibrated HIP, such as NASA's CHIP, is constructed by including a fiber-coupled spectrometer into the design such that the spectral output of the projector can be monitored[20]. Combined, the WIIT and the CHIP provide a controlled means of probing the intricacies of spatio-spectral interferometry in a shot-noise-limited regime.

In this paper, we describe how the CHIP will be used to obtain interferometric measurements of realistic astronomical test scenes from the WIIT in a time-efficient manner. In Section 2 we will motivate how the CHIP can expedite the measurement of spatially-spectrally complex images in combination with the data decomposition technique known as nonnegative matrix factorization (NMF). A brief introduction to NMF and how it applies to hyperspectral image decomposition is included in Section 3. The result of decomposing an astronomically realistic test scene as a function of the number of images through which the CHIP must cycle to represent the scene is presented in Section 4, followed by concluding remarks in Section 5.

2. Time-efficient WIIT data collection using CHIP

The WIIT is invaluable for demonstrating the effectiveness of double-Fourier interferometric imaging, but the data collection process for test scenes of moderate spatial and spectral complexity can be quite lengthy. To start, there are aspects to the data collection process that are intrinsic to the technique that cannot be altered. For example, the sample spacing of the OPD dimension limits the range of spectral frequencies in the reconstructed hyperspectral image, and the range of the delay line limits spectral resolution of the recovered image. This means the number of delay line positions for a single baseline is predetermined. In the same regard, the number of baseline measurements required to fully measure the spatial frequency domain, or *u*-*v* space, is dependent on both the maximum baseline length and the size of the individual apertures of the interferometer. One could use sparse sampling to reduce total measurement time, but we are assuming a general object that may not be conducive to sparse sampling. On the other hand, an aspect of the system over which we do have control is the integration time of the camera for each delay line position. Because datacubes are collected for many baselines, and many delay line positions are required for each datacube, decreasing the integration time for each delay line position will result in a large reduction in total data collection time for the WIIT. When experimentally simulating the measurement of hyperspectral images, however, the integration time is intimately related to implementing hyperspectral scene generation with the CHIP.

Recall that the CHIP generates a hyperspectral image by cycling through multiple spatially-spectrally separable images that can be added together throughout the integration time of the camera to simulate the measurement of a spatially-spectrally complicated hyperspectral image. Imagine a simple scene with two astronomical objects having different spectra on a blank background. For this simple scene, the CHIP would have to display the image of one object with its

spectrum, followed by the image of the second object with its different spectrum. As the number of objects with different spectra gets larger, the number of images through which the CHIP has to cycle grows proportionally until the number of spectrally diverse objects is the same as the number of spectral bins of the CHIP. At that point, one could just cycle through all of the CHIP's spectral bins independently. There is another option, however, that provides a trade-off between the number of images through which the CHIP cycles and how accurately the projected image matches the original hyperspectral test scene, in a manner similar to principal component analysis (PCA). Unfortunately, PCA on an arbitrary hyperspectral image will result in eigenspectra and eigenimages, sometimes referred to as abundance maps, with negative values, which we cannot represent using the CHIP. Nonnegative matrix factorization (NMF) is similar to PCA except all values in the eigenspectra and eigenimages are restricted to be nonnegative, so it can be used to decompose hyperspectral test scenes for the CHIP. The idea for time-efficient projection using the CHIP is to decompose the test scene into as few eigenimages and eigenspectra as possible while maintaining a specified accuracy between the projected image and the original test scene.

3. Hyperspectral Image Decomposition using NMF

NMF is the process in linear algebra of approximating an arbitrary nonnegative $n \times m$ matrix A by decomposing it into the product of two smaller nonnegative matrices W and H, with sizes $n \times p$ and $p \times m$, respectively, expressed as

$$A \approx WH. \tag{1}$$

The goal is to keep the size of p as small as possible while maintaining an accurate representation of the original matrix A, which is analogous to PCA with the exception that all three matrices are restricted to have nonnegative values. Most algorithms for NMF, including both the multiplicative update algorithm[24,25] and the alternating least squares algorithm[25,26], are based on iteratively reducing some cost function, usually the square of the Frobenius norm of the residual matrix:

$$\|A - WH\|_F^2 = \sum_{m,n} \left| (A - WH)_{mn} \right|^2. \tag{2}$$

The results shown later in this paper employ the alternating nonnegative least squares algorithm using projected gradient methods[26] to minimize this cost function because of its speed and consistency. Due to the imperfect data compression associated with approximating the matrix A with fewer than nm elements, all NMF optimization routines suffer from many local minima. This means that the starting values of W and H are important for finding a good solution. Instead of trying many random starting points and choosing the best result, we have found that applying nonnegative double singular value decomposition[27] to A provides an initial estimate that consistently outperforms the random guess method. The stopping criteria are related to the norm of the projected gradient as well as a maximum number of iterations. Figure 1 is an attempt to visualize how Eq. (1) applies to a hyperspectral image. The left side of Fig. 1 shows the original hyperspectral image and the right side shows how it can be approximated by summing over a sequence of spatially-spectrally separable images, represented as eigenspectra (top) and eigenimages (bottom). We can now discuss how to apply NMF to a hyperspectral image to obtain the plots and images on the right side of Fig. 1.

Adapting NMF for decomposition of a hyperspectral image into eigenspectra and eigenimages relies on array manipulations, which are simple tasks for computing languages such as Numpy/Python and Matlab. A hyperspectral image is often thought of as a three-dimensional array where two dimensions correspond to spatial coordinates and the other to the spectrum, but we can reshape the three dimensional array into a two-dimensional array such that the two spatial coordinates are collapsed down to a single dimension. This means that a hyperspectral image of size $j \times k \times l$ is reshaped to size $j \times kl$, where j indexes the spectral dimension and k and l are pixel indices. The result is an array that describes the spectrum for each image pixel and can now be assigned to the matrix A in Eqs. (1) and (2) such that $m = k$ and $n = kl$. After applying NMF to the hyperspectral matrix A, the p columns of W will be the computed eigenspectra and the p rows of H will contain the corresponding eigenimages. In order to visualize the eigenimages, we must reshape the matrix H from size $p \times kl$ to $p \times k \times l$. The eigenspectra now correspond to the spectral weightings applied to one of the CHIP's DLPs, while the eigenimages become the 8-bit gray scale images applied synchronously to the other DLP. Because spatially-spectrally complicated hyperspectral images can be decomposed into a handful of eigenspectra and eigenimages suitable as inputs for the CHIP, NMF is a natural solution to the problem of decomposing complicated hyperspectral images so that time-efficient data collection can be performed with the WIIT.

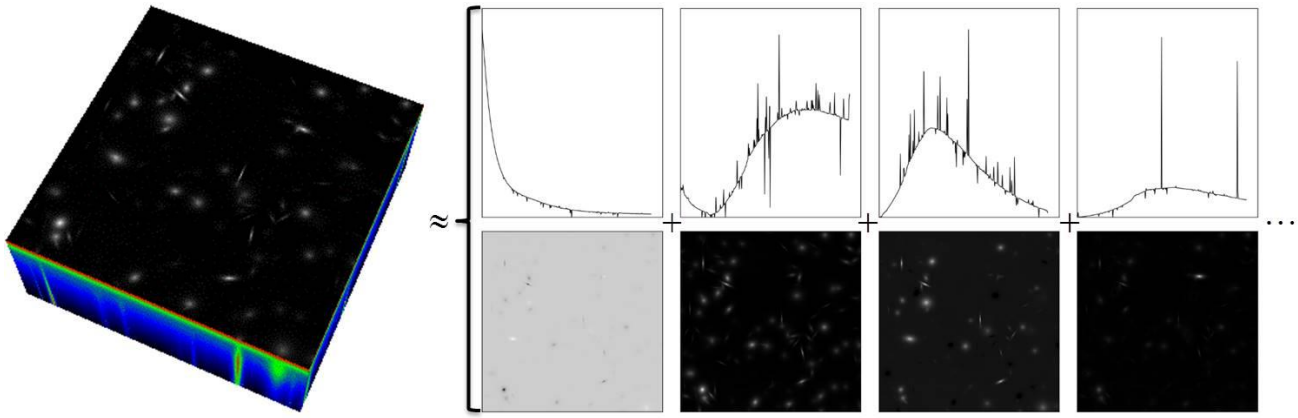

Figure 1. An illustration of Eq. (1) showing how a hyperspectral image (left) can be approximated by the sum of various spatially-spectrally separable images (right), with the eigenspectra on top and the associated eigenimages on the bottom.

4. Results for an astronomically realistic test scene

The hyperspectral test scene used for our simulations was one of a handful created by NASA to demonstrate the viability of a space-based double-Fourier interferometer for future far-infrared (FIR) missions, such as SPIRIT[8], because high resolution images of the FIR sky do not yet exist. The chosen test scene, a panchromatic image of which is shown in Fig. 2, is a deep field image comprised of many spectrally varying sources and possesses the most spatial-spectral complexity of any of the test scenes generated by NASA. This is optimal for demonstrating NMF decomposition because it will require the most eigenspectra (the largest value of p) to accurately represent the complexity of the test scene. The dimensions of the test scene are $j \times k \times l = 376 \times 375 \times 375$.

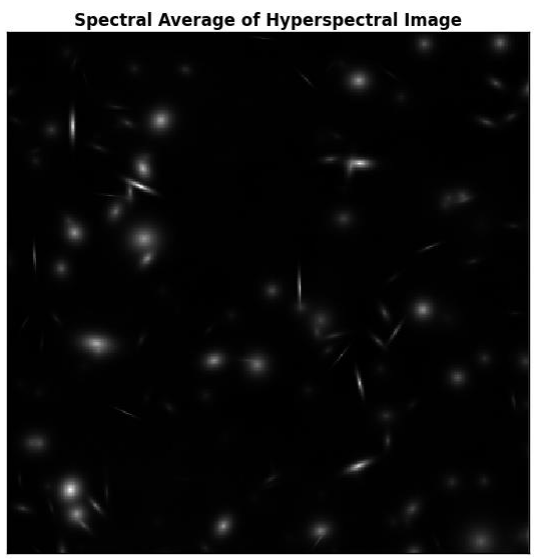

Figure 2. Spectral average of the simulated far-IR hyperspectral image to be decomposed (stretched to 0.25 power to show more faint spatial detail).

In order to perform time-efficient data collection as discussed in Section 2, we want to decompose the deep field image into as few eigenspectra as possible while maintaining adequate spatial/spectral complexity. We compare the results of the NMF estimated image with the input image using the normalized mean squared error (NMSE), which is a normalized version the cost function used by the NMF algorithm:

$$\mathrm{NMSE} = \frac{\sum_{m,n} \left(A - WH \right)_{mn}^2}{\sum_{m,n} A_{mn}^2}. \tag{3}$$

We performed NMF as described in Section 3 for integer values of p ranging from 4 to 15, and more sparsely sampled for values of p out to 100. Figure 3 is a plot of the NMSE as a function of p, showing that as few as 8 eigenspectra represents the original test scene with an NMSE < 0.01, about 27 eigenspectra represents the scene with an NMSE < 0.001, and about 58 eigenspectra produce an NMSE < 0.0001. The spectral average of the estimated hyperspectral image for p = 8 is shown in Fig. 4; it is visually similar to Fig. 3.

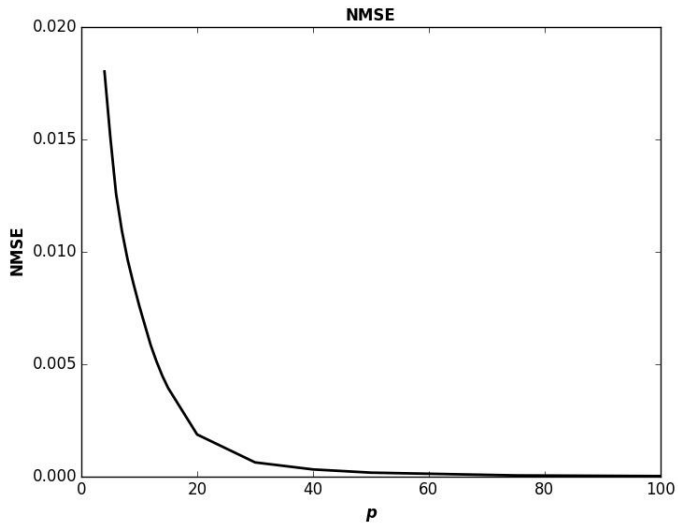

Figure 3. Plot of normalized mean squared error versus the number, p, of eigenspectra fitted by the NMF algorithm.

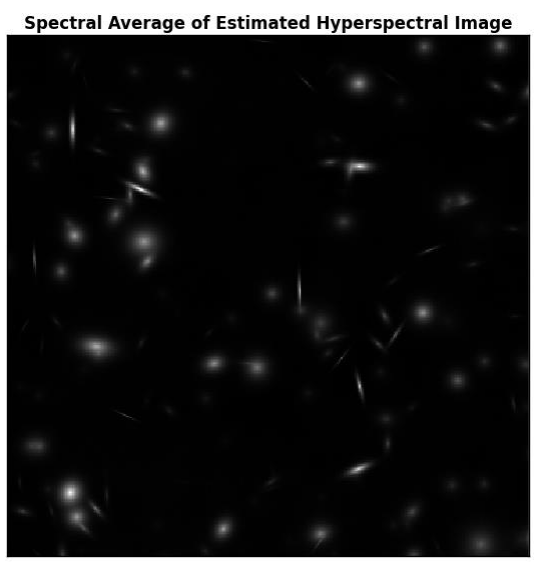

Figure 4. Spectral average of the decomposed far-IR hyperspectral image produced by NMF algorithm for p = 8 (stretched to 0.25 power).

As expected, the NMSE between the original and decomposed image monotonically decreases as the number p of recovered eigenspectra increases. We can expect that the NMSE will continue to decrease, approaching zero as either the value of p reaches the number of spectral bins in the test image, 376 in this case, or as the value of p reaches the total number sources with unique spectra. Figure 5 shows the reconstructed eigenspectra and eigenimages corresponding to

the approximated panchromatic image in Fig. 4. The first eigenspectrum and eigenimage recovered by the NMF algorithm in Fig. 5 are nearly identical for all values of p because together they describe the background radiation contributing energy to every pixel in the scene. The shape of the remaining eigenspectra and eigenimages are dependent on the value of p.

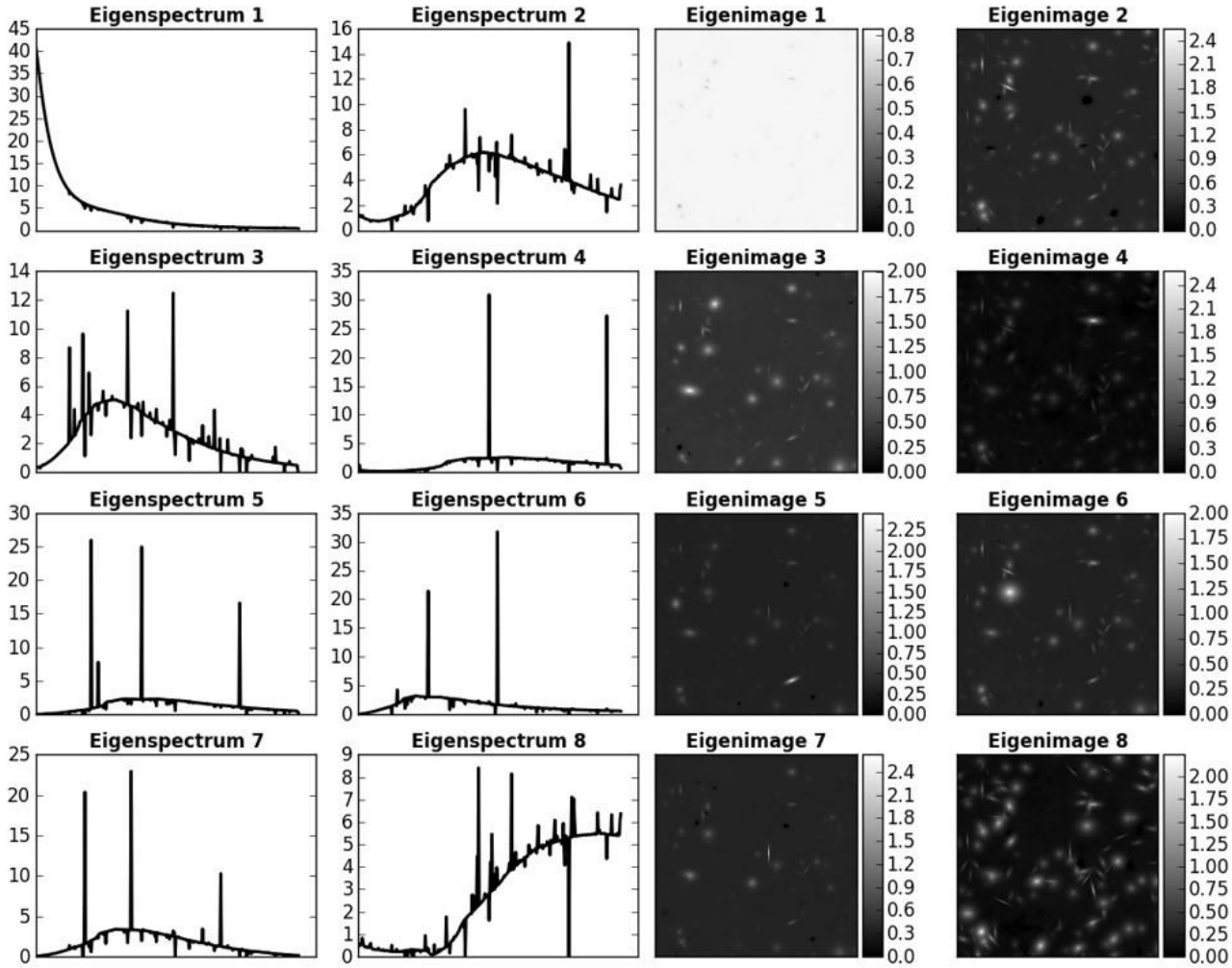

Figure 5. Eigenspectra and eigenimages (stretched to 0.3 power) produced by the NMF algorithm for $p = 8$.

Figure 6 shows the spectral average of the residual images, $A - WH$, for various values of p between 6 and 50 along with their associated NMSE values. Bright and dark features in the residuals correspond to spectral sources that are under-estimated and over-estimated, respectively. Notice that as the value of p increases, the NMF algorithm tends to improve the brightest and darkest spatial features, which contribute the most energy to the residual images. This is what we should expect because the algorithm is minimizing the Frobenius norm of the residual images, as discussed in Section 3. The remaining spatial features in the residual images likely have sharp spectral signatures not shared by the majority of astronomical sources within the scene.

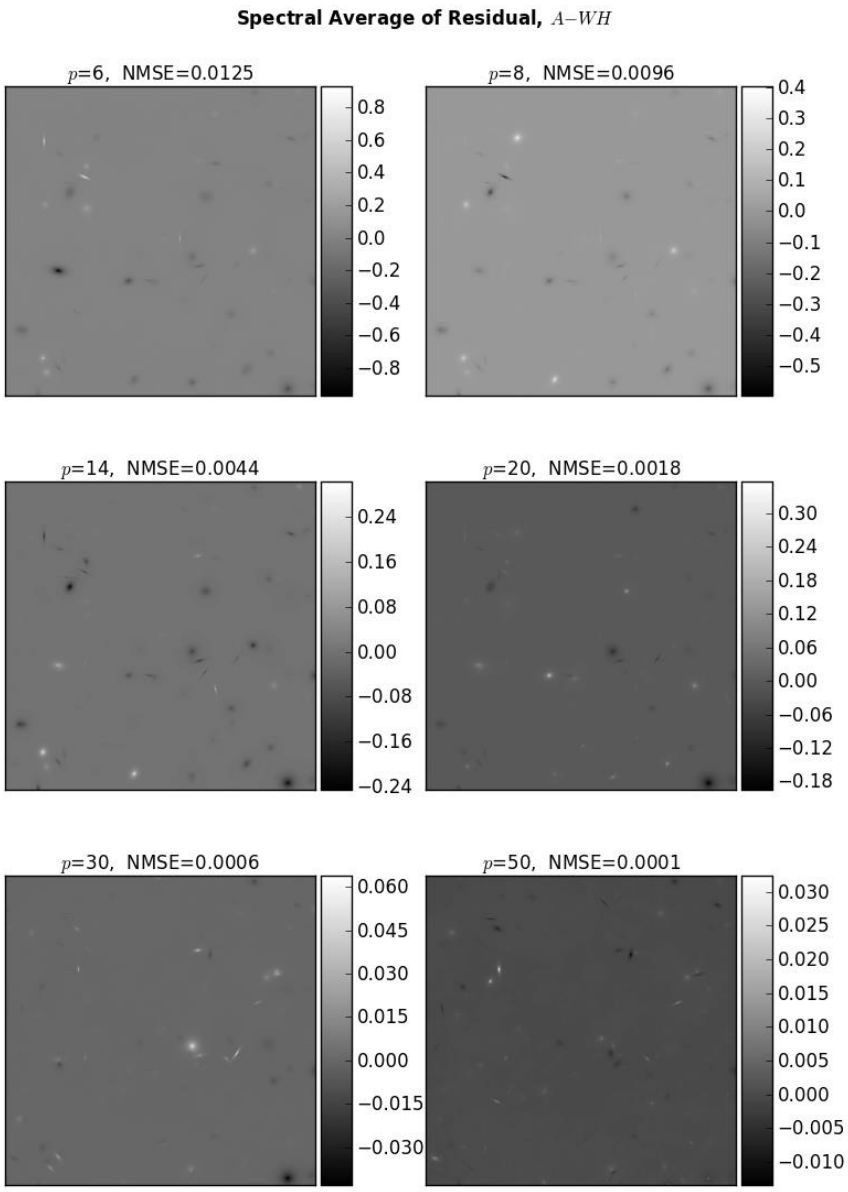

Figure 6. Spectral average of the residual hyperspectral image as the number of estimated eigenspectra varies from p = 6, 8, 14, 20, 30, 50. Note the difference in colorbar values.

5. Conclusion and Future Work

NMF has been applied to an astronomically realistic test scene that will be used by NASA's CHIP and WIIT for further characterization of double-Fourier interferometric imaging. We demonstrated that the NMSE between the original and decomposed hyperspectral image decreases monotonically as a function of the value of p, dropping below 0.01, 0.001, and 0.0001 for as few as 8, 27, and 58 eigenspectra and eigenimages, respectively.

Before the eigenspectra and eigenimages can be supplied to the CHIP, however, they both must be converted to 8-bit values limited by the DLP, reducing how accurately the CHIP can represent the original test scene. As a result, it would be better to include this information into the optimization algorithm in order to reduce quantization error. The quality of the approximated image might also be improved by choosing new starting guesses for the NMF algorithm from the eigenspectra and eigenimages found for larger values of p. However, the most important aspect of future work is the measurement of the decomposed hyperspectral image experimentally with the WIIT.

Acknowledgements

This work was supported by a NASA Space Technology Research Fellowship.

References

[1] Itoh, K., and Y. Ohtsuka, "Fourier-transform spectral imaging: retrieval of source information from three-dimensional spatial coherence," J. Opt. Soc. Am. A 3(1), 94-100 (1986)

[2] Mariotti, J.-M., and S. T. Ridgway, "Double Fourier spatio-spectral interferometry – combining high spectral and high spatial resolution in the near infrared," Astron. Astrophys. 195(1-2), 350-363 (1988).

[3] Elias, N., M. Harwit, D. T. Leisawitz, and S. A. Rinehart, "The mathematics of double-Fourier interferometers," Ap. J. 657(2), 1178-1200 (2007).

[4] Lyon, R. G., S. A. Rinehart, D. T. Leisawitz, and N. Memarsadeghi, "Wide-field imaging interferometry testbed (WIIT): Image construction algorithms," Proc. SPIE 7013, 70131M (2008).

[5] Lyon, R. G., D. T. Leisawitz, S. A. Rinehart, N. Memarsadeghi, and E. J. Sinukoff, "Wide-field imaging interferometry spatial-spectral image synthesis algorithms," Proc. SPIE 8445, 84450B (2012).

[6] Bennett, C. L., M. Carter, D. Fields, and J. Hernandez, "Imaging Fourier transform spectrometer," Proc. SPIE 1937, 191-200 (1993).

[7] Iacchetta, A. S., J. R. Fienup, and D. T. Leisawitz, "Spectral optical transfer function of a wide-field double-Fourier interferometer," in *Imaging and Applied Optics* 2015, OSA Technical Digest (online), IT2A.2 (2015).

[8] Leisawitz, D. T., C. Baker, A. Barger, D. Benford, A. Blain, R. Boyle, R. Broderick, J. Budinoff, J. Carpenter, R. Caverly, P. Chen, S. Cooley, C. Cottingham, J. Crooke, D. DiPietro, M. DiPirro, M. Femiano, A. Ferrer, J. Fischer, J. P. Gardner, L. Hallock, K. Harris, K. Hartman, M. Harwit, L. Hillenbrand, T. Hyde, D. Jones, J. Kellogg, A. Kogut, M. Kuchner, B. Lawson, J. Lecha, M. Lecha, A. Mainzer, J. Mannion, A. Martino, P. Mason, J. Mather, G. McDonald, R. Mills, L. Mundy, S. Ollendorf, J. Pellicciotti, D. Quinn, K. Rhee, S. A. Rinehart, T. Sauerwine, R. Silverberg, T. Smith, G. Stacey, H. P. Stahl, J. Staguhn, S. Tompkins, J. Tveekrem, S. Wall, and M. Wilson, "The space infrared interferometric telescope (SPIRIT): High-resolution imaging and spectroscopy in the far-infrared," J. Adv. Space Res. 40(5), 689-703 (2007).

[9] Leisawitz, D. T., B. J. Frey, D. B. Leviton, A. J. Martino, W. L. Maynard, L. G. Mundy, S. A. Rinehart, S. H. Teng, and X. Zhang, "Wide-field imaging interferometry testbed I: purpose, testbed design, data, and synthesis algorithms," Proc. SPIE 4852, 255-267 (2003).

[10] Rinehart, S. A., B. J. Frey, D. T. Leisawitz, D. B. Leviton, A. J. Martino, W. L. Maynard, L. G. Mundy, S. H. Teng, and X. Zhang, "Wide-field imaging interferometry testbed II: implementation, performance, and plans," Proc. SPIE 4852, 674-684 (2003).

[11] Leviton, D. B., B. J. Frey, D. T. Leisawitz, A. J. Martino, W. L. Maynard, L. G. Mundy, S. A. Rinehart, S. H. Teng, and X. Zhang, "Wide-field imaging interferometry testbed 3: metrology subsystem," Proc. SPIE 4852, 827-838 (2003).

[12] Rinehart, S. A., J. T. Armstrong, B. J. Frey, J. Kirk, D. T. Leisawitz, D. B. Leviton, L. W. Lobsinger, R. G. Lyon, A. J. Martino, T. A. Pauls, L. G. Mundy, and E. Sears, "The wide-field imaging interferometry testbed: II. Characterization and calibration," Proc. SPIE 5491, 1790-1801 (2004).

[13] Thompson, A. K., A. J. Martino, S. A. Rinehart, D. T. Leisawitz, B. J. Frey, and D. B. Leviton, "Optical modeling of the wide-field imaging interferometry testbed," Proc. SPIE 6268, 62682R (2006).

[14] Rinehart, S. A., J. T. Armstrong, B. J. Frey, J. Jung, J. Kirk, D. T. Leisawitz, D. B. Leviton, R. G. Lyon, A. J. Martino, T. A. Pauls, L. G. Mundy, and A. K. Thompson, "The wide-field imaging interferometry testbed: recent results," Proc. SPIE 6268, 626836 (2006).

[15] Rinehart, S. A., J. T. Armstrong, B. J. Frey, J. Jung, J. Kirk, D. T. Leisawitz, D. B. Leviton, R. G. Lyon, S. F. Maher, A. J. Martino, and T. A. Pauls, "The wide-field imaging interferometry testbed: enabling techniques for high angular resolution astronomy," Proc. SPIE 6687, 66870F (2007).

[16] Leisawitz, D. T., A. J. Martino, A. K. Thompson, S. A. Rinehart, and B. J. Frey, "An optical model of the wide-field imaging interferometry testbed," Proc. SPIE 6687, 66870J (2007).

[17] Rinehart, S. A., D. T. Leisawitz, B. J. Frey, R. G. Lyon, S. F. Maher, and N. Memarsadeghi, "The wide-field imaging interferometry testbed (WIIT): recent progress and results," Proc. SPIE 7013, 70132S (2008).

[18] Rinehart, S. A., D. T. Leisawitz, M. R. Bolcar, K. M. Chaprnka, R. G. Lyon, S. F. Maher, N. Memarsadeghi, E. J. Sinukoff, and E. Teichman, “Recent progress in wide-field imaging interferometry,” Proc. SPIE 7734, 77342D (2010).
[19] Leisawitz, D. T., M. R. Bolcar, R. G. Lyon, S. F. Maher, N. Memarsadeghi, S. A. Rinehart, and E. J. Sinukoff, “Developing wide-field spatio-spectral interferometry for far-infrared space applications,” Proc. SPIE 8445, 84450A (2012).
[20] Bolcar, M. R., D. T. Leisawitz, S. F. Maher, and S. A. Rinehart, “Demonstration of the Wide-field Imaging Interferometer Testbed using a Calibrated Hyperspectral Image Projector,” Proc. SPIE 8445, 84452D (2012)
[21] Rice, J. P., S. W. Brown, B. C. Johnson, and J. E. Neira, "Hyperspectral image projectors for radiometric applications," *Metrologia* **43**, S61-S65 (2006).
[22] Brown, S. W., J. P. Rice, J. E. Neira, R. R. Bousquet, and B. C. Johnson, "Hyperspectral image projector for advanced sensor characterization," Proc. SPIE 6296, 629602 (2006).
[23] Rice, J. P., S. W. Brown, J. E. Neira, and R. R. Bousquet, "A hyperspectral image projector for hyperspectral imagers," Proc. SPIE 6565, 65650C (2007).
[24] Lee, D. D., and H. S. Seung, “Algorithms for non-negative matrix factorization,” Proceedings of the Neural Information Processing Systems, 556-562 (2001).
[25] Berry, M. W., M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and Applications for Approximate Nonnegative Matrix Factorization." Computational Statistics and Data Analysis 52(1), 155–173 (2007).
[26] Lin, C. J., “Projected gradient methods for nonnegative matrix factorization,” Neural Computation 19(10), 2756-2779 (2007).
[27] Boutsidis, C., and E. Gallopoulos, “SVD-based initialization: A head start for nonnegative matrix factorization,” Pattern Recognition 41(4), 1350-1362 (2008).