

Virtual dimensionality for hyperspectral imagery

Chein-I Chang

A versatile, novel approach enables assessment of the number of spectrally distinct signatures in data covering a large spectral range.

Estimates of how many sources are present in a given data set are very difficult to obtain in a signal/image-processing context. This is particularly challenging for hyperspectral data (defined as spanning a large fraction of the electromagnetic spectrum), because many unknown signal sources are usually uncovered on the basis of significantly improved spectral resolution. Nevertheless, several theoretical criteria have been used in conjunction with sensor arrays to **estimate the presence and number of signal sources**. However, we have shown that they are neither appropriate nor effective.^{1,2}

Virtual dimensionality (VD) provides an effective alternative.^{1,2} Its basic, underlying concept is the ‘pigeon-hole principle’.³ If we represent a signal source by a pigeon and a spectral band by a pigeon hole, we can **use a given spectral band to accommodate one particular source**.⁴ Thus, if a signal source is present in our data, we should be able to detect it in the **relevant spectral band**. In essence, this implies that we must **calculate the eigenvalues of both the data-correlation and covariance matrices**, $\{\hat{\lambda}_l\}$ and $\{\lambda_l\}$, for the l^{th} band. **A signal source is present if their difference, $\hat{\lambda}_l - \lambda_l$, is positive.**

A binary-composite-hypothesis test is formulated for each spectral band such that the null and alternative hypotheses (H_0 and H_1) represent two scenarios, $H_0: \hat{\lambda}_l - \lambda_l \equiv 0$ and $H_1: \hat{\lambda}_l - \lambda_l > 0$, corresponding to the absence and presence of a signal in the l^{th} band, respectively. We have developed the Harsanyi-Farrand-Chang (HFC) method,⁵ as well as a noise-whitened version (NWHFC),^{1,2} on the basis of Neyman-Pearson detection theory to explore how many times the test fails for all spectral bands and for a given false-alarm probability, P_F . The corresponding VD is defined as the number of failures, which indicates the **number of signal sources present in the data**. The VD is thus **completely determined by P_F** .

We have shown that VD has found many applications in hyperspectral-data exploitation.^{6,7} Here, we present a number

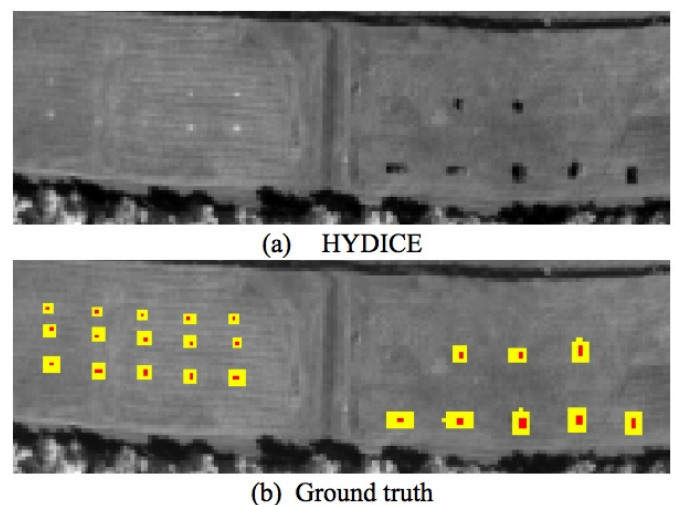


Figure 1. (a) Hyperspectral digital imagery collection experiment (HYDICE) scene containing vehicles and panels. (b) Ground-truth map of the objects' recovered spatial locations.

of illustrative examples. Because of its **use of hundreds of contiguous spectral bands**, a hyperspectral image contains vast amounts of data. **Dimensionality reduction (DR) and band selection (BS) are two major approaches that take advantage of high band correlation to remove data redundancy.** Both aim to derive how to determine the number of dimensions, q , after DR or the number of bands, \hat{q} , needed for BS. We have demonstrated that VD provides a good estimate of q .⁸ Second, we have shown that VD can also be used to estimate the \hat{q} required for BS.⁹ In addition, VD can be used to predict how many spectrally distinct signal sources are present in the data. Two applications relevant to this question, endmember extraction and linear spectral-mixture analysis (LSMA), are of particular interest. An endmember is an idealized, **pure signature specifying a spectral class**. Estimating the number of endmembers present in the data is a fundamental and crucial step in understanding hyperspectral data.^{10,11} LSMA is one of most widely used data-analysis

Continued on next page

techniques in remote sensing. It assumes that data vectors are linearly mixed using a set of p basic, **spectrally distinct signal signatures**, $\{m_1, m_2, \dots, m_p\}$, that are a priori known to be present in the data. Unfortunately, in many real applications knowledge of $\{m_j\}_{j=1}^p$ is either too difficult to obtain or may not be reliable. We have demonstrated that VD then provides an effective means of estimating p and lays the foundation for unsupervised LSMA.^{12–14}

To demonstrate the applicability of the VD, we used several well-known hyperspectral image-data sets for illustrative experiments. Figure 1 shows a hyperspectral digital imagery collection experiment (HYDICE) scene and its ground-truth counterpart covering 200×64 pixels, acquired using 210 spectral channels ranging from 0.4 to $2.5\mu\text{m}$ with spatial and spectral resolutions of approximately 1.56m and 10nm , respectively. The ground-truth map in Figure 1(b) highlights **five distinct signatures**, specifying the **15 panels on the left side of the scene**. On the right, two different types of vehicles with sizes of 4×8 and $6 \times 3\text{m}^2$ are recovered in the bottom row, while three objects are shown at the top. Thus, this particular scene reveals at least ten distinct man-made signatures in addition to background signals.

Figure 2 shows three different image scenes acquired by the Airborne Visible IR Imaging Spectrometer (AVIRIS) using 224 spectral channels ranging from 0.4 to $2.5\mu\text{m}$ with spatial and spectral resolutions of approximately 20m and 10nm , respectively. Figure 2(a) shows 145×145 -pixel representations of the well-studied Purdue Indiana Indian Pine test site,¹⁵ covering an area of mixed agriculture and forestry in northwestern Indiana.¹⁶ The ground-truth image shows 16 distinct object classes and an additional background class. Figure 2(b) shows detection of five different types of minerals in another well-known AVIRIS scene.¹⁷ We used two types of Cuprite images, ‘Reflectance’ and ‘Radiance.’ Figure 2(c) is an image scene of pixels in the Lunar Crater Volcanic Field, located in northern Nye County, Nevada. There are five targets of interest, highlighted by the radiance spectra of red oxidized basaltic cinders, rhyolite, playa (dry lake), vegetation, and shade.¹⁸

The VD values estimated using the HFC and NWHFC methods for the three AVIRIS image scenes in Figure 2 are tabulated in Table 1. VD provides reasonable estimates of the number of signatures present in the data sets when the false-alarm probability rate is limited to $P_f \leq 10^{-3}$. In addition, the NWHFC method is generally more effective than the HFC approach.

Estimating signal sources in hyperspectral imagery is a very challenging problem to which VD provides an effective solution. A number of alternative methods based on different criteria have also been developed^{19–22} (for example, to estimate the number of anomalies in anomaly detection). Their

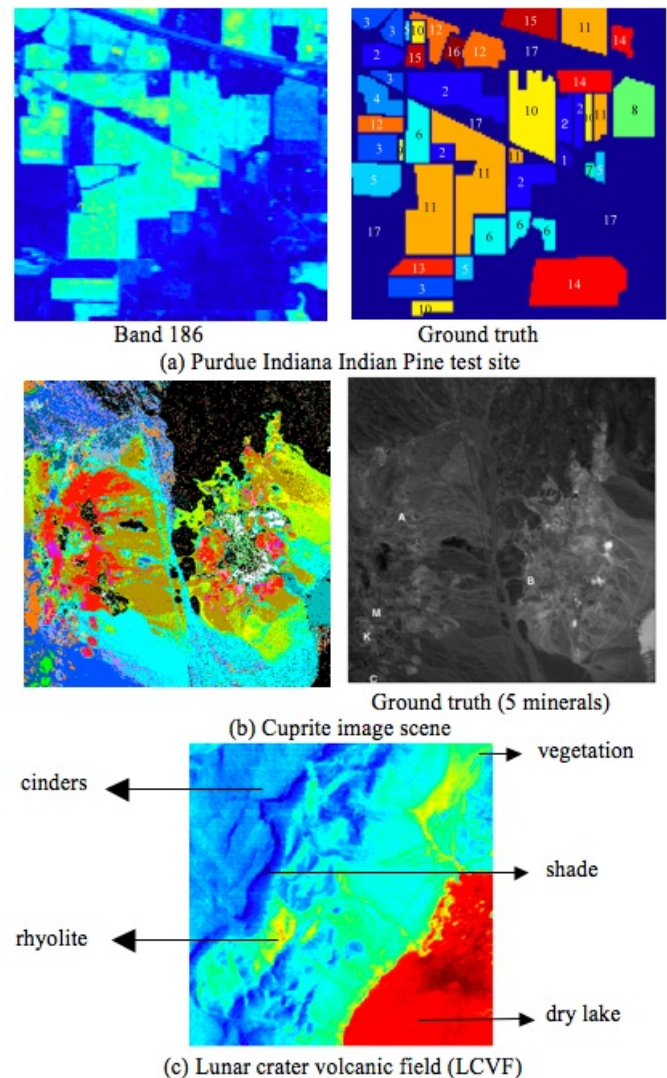


Figure 2. Airborne Visible IR Imaging Spectrometer image scenes and their ground-truth counterparts. (a) Purdue Indiana Indian Pine test site (the numbers refer to the distinct object and background classes). (b) Cuprite. A: Alunite. B: Buddingtonite. C: Calcite. K: Kaolinite. M: Muscovite. (c) Lunar Crater Volcanic Field.

performance varies for different applications and is not necessarily superior to that achieved by the HFC/NWHFC methods. Since the latter estimate VD on the basis of the pigeon-hole principle, they effectively identify spectral rather than spatial classes (as generally done in traditional spatial-domain-based analysis). Consequently, VD may not be suitable for use with applications and image data characterized by significant spatial correlations.

Continued on next page

**Table 1.** Estimated VD on the basis of the HFC and NWHFC methods.

	P_F	10^{-1}	10^{-2}	10^{-3}	10^{-4}
HYDICE	HFC	19	12	9	8
(200 × 64)	NWHFC	25	22	19	18
Cuprite	HFC	34	30	24	22
Reflectance	NWHFC	32	29	25	23
Cuprite	HFC	29	18	17	15
Radiance	NWHFC	23	20	20	19
Purdue with	HFC	72	42	34	29
background	NWHFC	19	18	18	18
Purdue without	HFC	55	33	30	24
background	NWHFC	18	18	18	15
Lunar Crater	HFC	8	6	4	4
Volcanic Field	NWHFC	20	13	11	11

Therefore, we recently extended and generalized VD, providing an **effective spectral dimensionality**. We are currently investigating a number of techniques related to this new concept, such as approaches based on linear spectral mixing and maximum orthogonal subspace projection.

Author Information

Chein-I Chang

Remote Sensing Signal and Image Processing Laboratory
University of Maryland, Baltimore County
Baltimore, MD

Chein-I Chang is a professor. He received his PhD in electrical engineering from the University of Maryland at College Park. He has published over 100 journal articles, authored a book (*Hyperspectral Imaging*), edited two books (*Recent Advances in Hyperspectral Signal and Image Processing* and *Hyperspectral Data Exploitation: Theory and Applications*), and co-edited a fourth book (*High Performance Computing in Remote Sensing*). He is currently working on his second single-author book, *Hyper*.

References

1. C.-I Chang, **Hyperspectral Imaging: Techniques for Spectral Detection and Classification**, Kluwer/Plenum Acad. Publ., 2003.
2. C.-I Chang and Q. Du, *Estimation of the number of spectrally distinct signal sources in hyperspectral imagery*, **IEEE Trans. Geosci. Remote Sens.** **42** (3), pp. 608–619, 2004.
3. S. S. Epp, **Discrete Mathematics with Applications**, Brooks/Cole, Pacific Grove, CA, 1995.
4. C.-I Chang, *Hyperspectral imaging requires reinventing multispectral concepts*, **SPIE Newsroom**, 2008. doi:10.1117/2.1200812.1301 <http://spie.org/x32212.xml>
5. J. C. Harsanyi, W. Farrand, and C.-I Chang, *Detection of subpixel spectral signatures in hyperspectral image sequences*, **Proc. Am. Soc. Photogramm. Remote Sens. Annu. Mtg.**, pp. 236–247, 1994.
6. C.-I Chang, *Exploration of virtual dimensionality in hyperspectral image analysis*, **Proc. SPIE** **6233**, p. 623313, 2006. doi:10.1117/12.665280
7. C.-I Chang, Chapter 1. *Recent advances in hyperspectral signal and image processing*, C.-I. Chang ed., **Utility of Virtual Dimensionality in Hyperspectral Signal/Image Processing**, Trivandrum, Kerala: Research Signpost, India, 2006.
8. J. Wang and C.-I Chang, *Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis*, **IEEE Trans. Geosci. Remote Sens.** **44** (6), pp. 1586–1600, 2006.
9. C.-I Chang and S. Wang, *Constrained band selection for hyperspectral imagery*, **IEEE Trans. Geosci. Remote Sens.** **44** (6), pp. 1575–1585, 2006.
10. A. Plaza and C.-I Chang, *Impact of initialization on design of endmember extraction algorithms*, **IEEE Trans. Geosci. Remote Sens.** **44** (11), pp. 3397–3407, 2006.
11. C.-I Chang, C. C. Wu, W. Liu, and Y. C. Ouyang, *A growing method for simplex-based endmember extraction algorithms*, **IEEE Trans. Geosci. Remote Sens.** **44** (10), pp. 2804–2819, 2006.
12. X. Jiao and C.-I Chang, *Unsupervised hyperspectral target analysis*, **Proc. SPIE** **7086**, p. 70860P, 2008. doi:10.1117/12.795242
13. X. Jiao, Y. Du, and C.-I Chang, *Component analysis-based unsupervised linear spectral mixture analysis for hyperspectral imagery*, 2009.
14. X. Jiao and C.-I Chang, *Unsupervised hyperspectral image classification*, **Proc. SPIE** **6661**, p. 66610I, 2007. doi:10.1117/12.732614
15. <http://cobweb.ecn.purdue.edu/~biehl/MultiSpec/documentation.html> Documentation for MultiSpec. Accessed 4 September 2009.
16. D. A. Landgrebe, **Signal Theory Methods in Multispectral Remote Sensing**, John Wiley & Sons, New York, 2003.
17. C. C. Wu, C. S. Lo, and C.-I Chang, *Improved process for use of a simplex growing algorithm for endmember extraction*, **IEEE Trans. Geosci. Remote Sens. Lett.** **6** (3), pp. 523–527, 2009.
18. J. C. Harsanyi and C.-I Chang, *Hyperspectral image classification and dimensionality reduction: an orthogonal subspace projection approach*, **IEEE Trans. Geosci. Remote Sens.** **32** (4), pp. 779–785, 1994.
19. J. M. Dias and J. M. P. Nascimento, *Estimation of signal subspace on hyperspectral data* **5982**, p. 59820L, 2005. doi:10.1117/12.620061
20. J. M. Bioucas-Dias and J. M. P. Nascimento, *Hyperspectral subspace identification*, **IEEE Trans. Geosci. Remote Sens.** **46** (8), pp. 2435–2445, 2008.
21. X. Jiao, *Unsupervised Hyperspectral Target Analysis*, PhD thesis, Department of Computer Science and Electrical Engineering, Univ. of Maryland, 2010.
22. O. Kuybeda, D. Malah, and M. Barzohar, *Rank estimation and redundancy reduction of high-dimensional noisy signals with preservation of rare vectors*, **IEEE Trans. Signal Process.** **55** (12), pp. 5579–5592, 2007.