Scene Analysis and Detection in Thermal Infrared Remote Sensing Using Independent Component Analysis

Bernard R. Foy and James Theiler
Los Alamos National Laboratory
Los Alamos, NM 87545

ABSTRACT

Independent Component Analysis can be used to analyze cluttered scenes from remote sensing imagery and to detect objects. We show examples in the thermal infrared spectral region (8-12 µm) using both passive hyperspectral data and active multispectral data. The examples are from actual field data and computer simulations. ICA isolates spectrally distinct objects with nearly one-to-one correspondence with the independent component basis functions, making it useful for modeling the clutter in typical scenes. We show examples of chemical plume detection in real and simulated data.

Keywords: remote sensing, hyperspectral imaging, lidar, ICA, chemical sensing

1. INTRODUCTION

Hyperspectral and multispectral imaging in the infrared region are widely used for landcover classification of natural and urban scenes. [1] In the thermal or longwave IR region of the spectrum (approx. 8-12 µm), objects are identified by their thermal emission, rather than their reflective properties as in the visible region (0.4-0.7 µm). [2] Independent Component Analysis (ICA) offers a comparatively new way to extract spectral features in an unsupervised manner, without knowing *a priori* what the spectra should look like. In some applications, one searches for an object with a known spectral signature against an unknown cluttered background. ICA offers a way to model the cluttered background so that the desired signature can more easily be discriminated from the clutter. Unsupervised analysis of this sort is similar to Blind Source Separation (BSS), for which ICA has proved very successful. [3,4]

ICA has previously been applied to hyperspectral and multispectral remote sensing problems in the vis-SWIR spectral regions. [5-8] It has also been implemented in chemical analysis and quantification for complex mixtures in analytical chemistry applications. [9] We have a particular interest in remote chemical detection using both passive and active (lidar) sensors in the thermal IR. In these applications, the chemical has a known spectral signature and overlays the background clutter as a plume or cloud in an unknown spatial region of the scene. It is important to model the clutter as accurately as possible even though identification of the clutter components is not the primary goal. The plume modifies the background by absorption or emission of radiation in selected pixels, producing "on-plume" and "off-plume" regions of the scene.

2. PASSIVE SENSING DATA

In order to investigate the utility of ICA as a clutter modeling technique, we apply it to an example of data acquired at Los Alamos National Laboratory. [10] The sensor looks horizontally at a natural scene with canyons, mesa, mountains, and sky at the top (Fig. 1). The atmospheric path varies from ~2 km to ~20 km for the topographical features. At each pixel in Fig. 1, the radiance averaged over all wavelengths is plotted. There are 300 X 85 pixels in the data, with 116 spectral channels from 750 cm^{-1} to 1248 cm^{-1}. We performed ICA on this data using the FastICA algorithm of Hyvärinen et al. [3,4] The resulting IC images are shown in Fig. 2. (Grayscale value corresponds to the "ICA signal." [3]) Principal component (PC) images are shown in Fig. 3. In the ICA, dimension reduction was first performed by PCA to 20 variables. Options chosen for the calculation are symmetric decorrelation, nonlinear function g(y)=y $exp(-y^2)$, and stabilization turned on, [3] although these do not appreciably affect the final result.
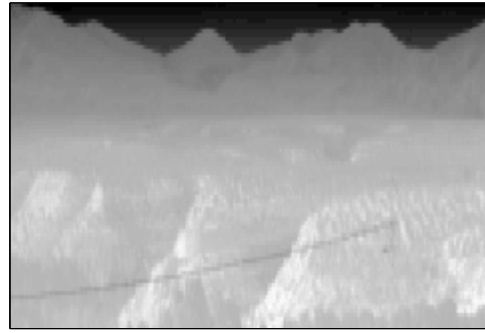
Figure 1. Image from a thermal infrared hyperspectral sensor. The mean value over all wavelengths is plotted for each pixel. The image size is 300 X 85, and there are 116 spectral channels.
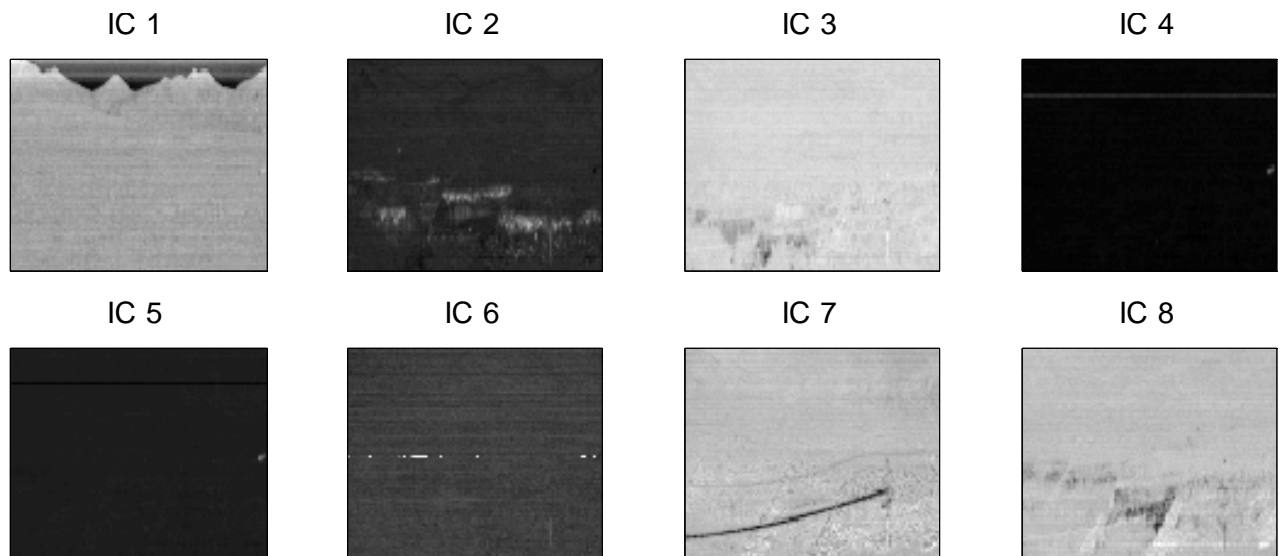
IC 1   IC 2   IC 3   IC 4

IC 5   IC 6   IC 7   IC 8

Figure 2. Independent component images for the data of Fig.1.

Image of PC 1   Image of PC 2   Image of PC 3   Image of PC 4

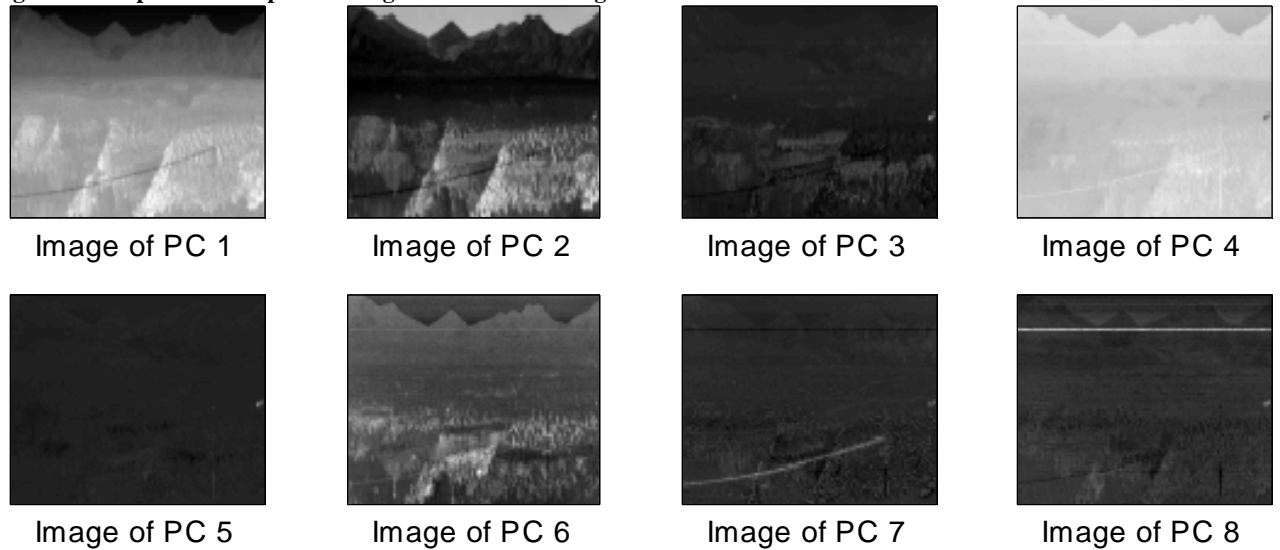Image of PC 5   Image of PC 6   Image of PC 7   Image of PC 8

Figure 3. Principal component images for the data of Fig.1, ordered by decreasing eigenvalue.
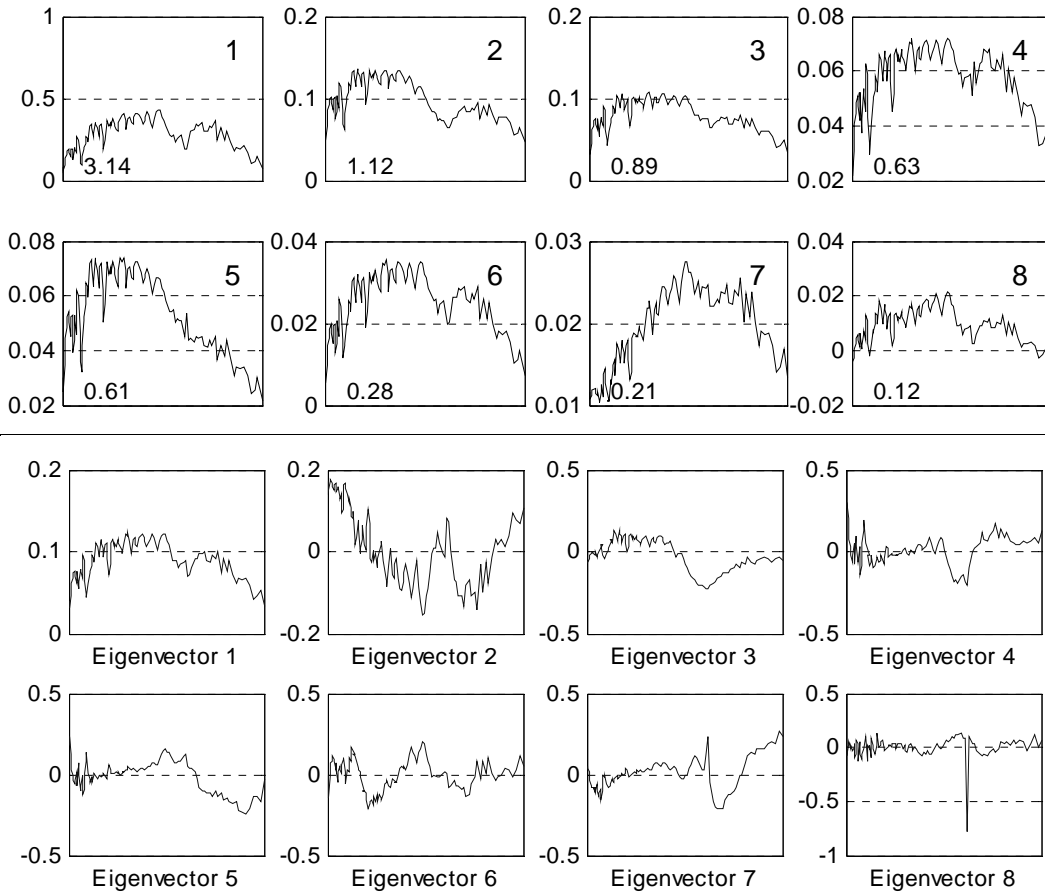
Figure 4. (Top) Independent component basis vectors for the data of Fig. 1. The vector magnitudes are listed in the lower left corner. (Bottom) Principal component eigenvectors. The x-axis ranges from 750 to 1248 cm^{-1}.

There are some interesting aspects of the IC images that merit comment. A few very distinct elements of the clutter appear mostly in one IC image. The sky is restricted to IC 1, a powerline crossing the bottom is in IC 7, the exposed rock above a roadcut is in IC 8, and another type of rock is highlighted in IC 2. A small unidentified man-made object is IC 5 as just a few pixels. These objects have either distinct spectra or a different overall radiance (due to temperature or emissivity). By contrast, PCA generally highlights these objects in multiple PC images (*e.g.* the powerline is noticeable in PC 1, 2, 7 and the sky in most of the PCs).

ICs 4 and 6 are mostly confined to single rows of the image, which is likely traceable to a sensor artifact – a miscalibrated pixel or row of pixels on the sensor array, for example. This artifact is also detected by PCA in PC 8 and noticeable to some extent in a few other PC images. Overall, however, ICA seems to do a better job of highlighting distinct objects in separate components, which one would expect from ICA's use of statistical independence in the data. Other examples of airborne data not shown here confirm this.

The basis vectors obtained by ICA that correspond to these images also show differences from PCA, as seen in Fig. 4. The spectral shapes of the basis vectors appear to be close to physically reasonable representations of emission spectra (or brightness temperature curves) of typical scene elements such as construction materials and vegetation. Most real – world materials have slowly varying spectra from 750-1300 cm^{-1} without sharp features. (For gas-phase chemicals, by contrast, sharp features are common.) The PCA results are more difficult to interpret, in accordance with Schott's remark that "PCA equates variance with information." [1] The PC basis vectors 1-4 are actually somewhat similar to the first four ICs, but there are clearly un-physical sharp features and oscillations present in PCs 5-8. PCA provides a suitable orthogonal basis set to describe scene clutter, but individual PCs do not have a one-to-one correspondence with

actual scene elements. A similar comparison of ICA and PCA basis vectors has been reported for visible-wavelength data using cluttered natural scenes. [7]

Note that the very fine sharp structure in Fig. 4 in both IC and PC basis vectors is due to atmospheric transmission and upwelling. There is a pronounced ozone dip at 1000-1100 cm^{-1} that is repeated in most of the components. Atmospheric compensation (removal of the sharp structure by estimation of the atmospheric composition) [2] may be able to remove these features from the IC spectra. Since the sensor looks along a horizontal path and over such a large range of horizontal distances, atmospheric compensation would be very difficult.

SIMULATED PASSIVE SENSING DATA

We can obtain considerable insight into the behavior of ICA with hyperspectral infrared data by synthesizing data with cluttered scenes. By choosing the number of clutter components and their spatial location, we can visually assess the performance of the algorithm in identifying the materials present. We simulated a hyperspectral datacube with similar spectral characteristics to the field data of Fig. 1. A scene was constructed using 8 typical materials (types of asphalt, concrete, brick, roofing shingle, paint, tar paper, wood), for which spectra were taken from a library. They were arrayed in patches as shown in Fig. 5 (top). Emission spectra were calculated for a uniform scene temperature. We added realistic atmospheric transmission and upwelling to the data, although it does not affect the analysis very much. The pixels were then mixed randomly in proportions of 0-60% by adding in radiance spectra of different materials selected from the group of 8. In other words, a pixel in the first patch of the scene will have 40-100% of the character of the first material in the list, and varying amounts of the other seven materials, with amounts chosen randomly from pixel to pixel. The result is a scene with a small number of "pure" pixels and a large number of thoroughly mixed pixels. A small amount of Gaussian random noise (<1%) was added to simulate instrumental effects. [2] The synthesized datacube has 128×250 pixels and 128 wavelengths in the same spectral region as the field data of Fig. 1.

ICA was performed on the simulated data and yielded the IC images shown in Fig. 5 (middle panel). PCA yields the images in the lower half of the figure. For the ICA result, the data dimension was first reduced to 8 principal components, but the choice is not critical. For Fig. 5, the nonlinear function was chosen as $g(y) = y^3$ in the FastICA routine. [3] Within each patch of the scene, some graininess is apparent in the IC signal that reflects the random mixing of the pixels by the simulation.

Each IC image seems to be binary, highlighting one of the 8 clutter components while representing all of the others as equivalent. (The highlighted patch may be either dark or light, since the sign of the independent components is arbitrary.[3,4]) The IC images are good "detectors" of the individual components of the clutter. For the PCA images, however, there is far less selectivity. PCA detects the spatial location of each image patch quite well, but every PC image highlights multiple components of the clutter (and usually all of the components).

The FastICA algorithm provides several choices for the nonlinear function $g(y)$ that is used in the approximation of non-Gaussianity of the data. A full explanation is provided by Hyvärinen *et al.* [3,4] If we choose $g(y) = tanh(y)$, we obtain a result with some interesting differences, seen in Fig. 6. In this case, ICA again does a good job of highlighting the components of the clutter, but it highlights them in pairs instead of singly. In each case, the highlighted components are given large and small values of the IC signal, while the rest of the scene is assigned a middle value. One particular clutter component is repeated in each image. Only 7 ICs are needed to encompass the 8 components, so the eighth IC is just noise. This still stands in stark contrast to the PCA result (Fig. 5, bottom panel). We have seen other variations depending on the optimization algorithm, highlighting three components in a single IC image.

Albedo map

Image of PC 1   Image of PC 2   Image of PC 3   Image of PC 4

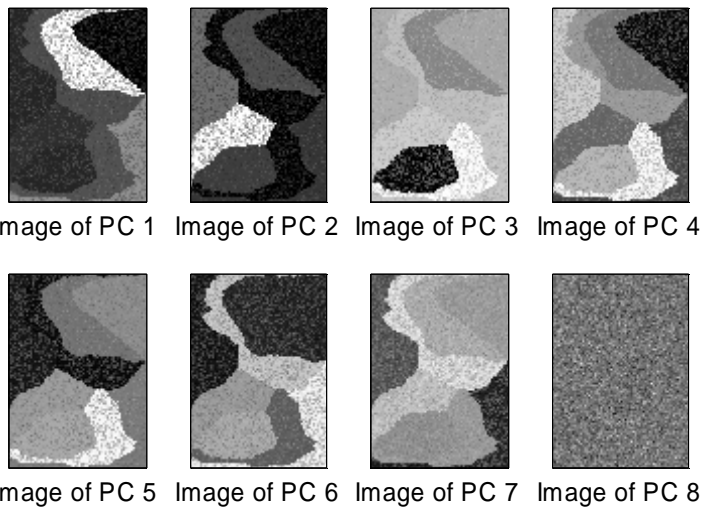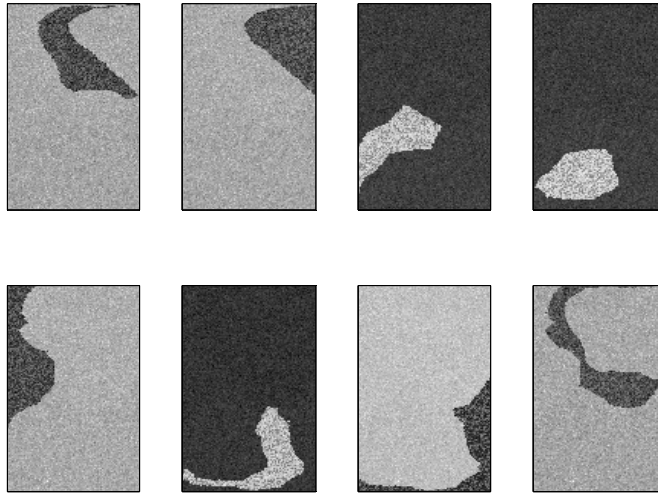Image of PC 5   Image of PC 6   Image of PC 7   Image of PC 8

Figure 5. (top) Map of material locations used to construct synthetic scene. The gray scale is simply qualitative and does not refer to radiance. (middle) IC images for the synthetic data. The graininess is caused by random mixing at each pixel. The nonlinear function used was $g(y) = y^3$. (bottom) PC images of the same data.
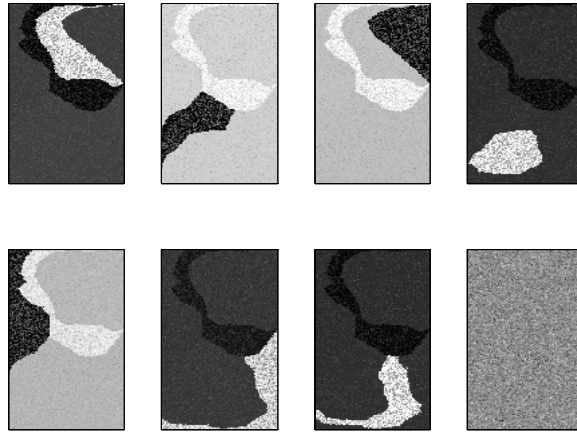
Figure 6. IC images similar to Fig. 5 (middle panel), but using the nonlinear function g(y) = *tanh*(y). Note how the distinct patches in the scene are highlighted in pairs with opposite signs.

3. CHEMICAL PLUME DETECTION WITH ICA

Both ICA and PCA provide ways to account for the clutter in hyperspectral infrared data. We can make use of these basis vectors to detect targets in the clutter that are only weakly present. By weak, we mean that they are either very rare in the scene or they only weakly modify the background radiance. A chemical plume is usually of the latter sort. A simple way to make use of the ICA basis vectors in target detection is to perform a least-squares fit of the known target signature d together with some number of IC basis vectors a_i, where the ICs account for the background radiance at each pixel as necessary. Collecting d and the a_i in the matrix $U = [d\ a_1\ a_2\ .....\ \ a_i\ 1]$ with a column of ones, we find the fitted coefficient of the chemical β_1 from ordinary least squares applied to the original data vectors x:

$$\hat{\beta} = \left(U^T U\right)^{-1} U^T x$$

We prepared a simulated datacube just as described above and overlaid a chemical plume on it. The chemical chosen (Freon 134) has chemical absorption features throughout the 750-1300 cm^{-1} region. The amount of chemical and its temperature contrast ($T_{plume} - T_{ground}$) were adjusted to make the radiance change just barely perceptible. Figure 7 contains images of the chemical coefficient, β_1. At least 6 basis vectors are needed to begin seeing the plume reasonably well. With 7 IC basis vectors, the background patches are "nulled out" almost completely and the rectangular plume can be seen equally well above all background patches. At an intermediate stage, using only a few ICs in the fit, these patches are evident in the result with high contrast, suppressing the plume almost completely.
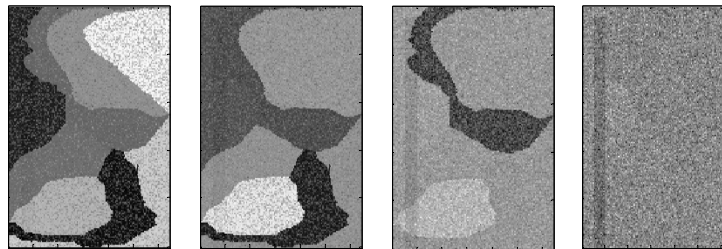

Figure 7. Detection of a weak chemical plume in the synthetic scene. The known chemical spectrum was fitted to the observed spectrum at each pixel simultaneously with IC basis vectors. The number of ICs used in the fit was 1,4,6,7 from left to right. The very weak chemical plume is apparent in the last image as a rectangular feature close to the left edge of the scene. Results using PC basis vectors were very similar.

A similar chemical detection procedure using PCA to generate basis vectors has been described previously. [11] This is effectively equivalent to orthogonal subspace projection. [12] We tried PCA on our simulation and obtained a very similar

result to Fig. 7. Matched-filter detection [13] also gives similar results, using the full scene to generate the covariance matrix for use in the filter. We are currently exploring other ways to make use of the ICA basis to effectively characterize the clutter, thereby enabling better weak target detection.

We also performed ICA on strong plumes in the scenario of Fig. 7. We verified that ICA "detects" the plume simply as one of the distinct components in the data, i.e. without even using the chemical signature. In some circumstances, the chemical plume appears in only one IC image, separated from the background clutter. However, as the strength of the plume is reduced, this approach fails. The strong plume case, with an obvious radiance signature, is far less challenging than the weak plume case. Very simple techniques such as plotting the ratio of wavelengths, using on-peak and off-peak wavelengths, will readily detect a strong plume.

4. CHEMICAL DETECTION IN LIDAR IMAGERY

Lidar images are also amenable to ICA. A more complete description of this experimental approach is in previous papers. [10,14] The lidar system generates an image of a cluttered scene that is almost the same as a passive infrared sensor, except that the image is produced by object reflectance spectra rather than thermal emission spectra. We set up a chemical plume release in a natural scene, using a very dense plume in light winds. The very strong chemical plume of SF_6 gas was detected by absorption of the CO_2 laser radiation. ICA gives images as shown in Fig. 8. Since this is an example of a strong plume with peak absorption of ~90%, ICA exhibits the plume as simply a distinct component of the data. The third IC also shows the vertical release pipe that was used to generate the plume. (Note again that components may be either light or dark since sign is arbitrary. [3,4])
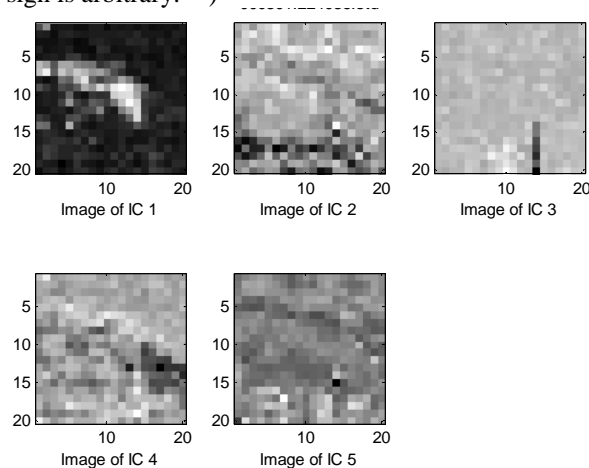

Figure 8. IC images of lidar data on a chemical plume. The first IC shows a strong chemical plume rising near center and fading out to the left, and the third shows the chemical release stack. Other ICs are associated with landscape features.

We also note that in images without a chemical plume, the ICA signal is associated with scene components in a manner quite similar to the K-means clustering approach discussed in our previous paper. [14]

5. DISCUSSION

The failure to recognize the qualitative differences between strong plume and weak plume detection has caused some confusion among researchers. When the plume is strong, it becomes observable as a major component of the scene. Simple unsupervised classification methods can then be used for detection with no need even to know the chemical signature of the plume. For example, applying ICA to a scene with a strong (and relatively large) plume often results in a "plume component" -- a single ICA component in which the plume alone is highlighted. Projection pursuit, a technique with some similarities to ICA, has been successfully used to detect a gaseous plume in hyperspectral infrared data, but the signal was in the strong-plume regime. [15] In the weak plume case, the plume is too subtle a feature for this to work. The chemical signature of the plume is crucial information in this regime, and is typically used in matched-filter detection [13] or a multiple regression of the chemical along with some set of basis vectors that account for the

background clutter. [11] A plume that is weak and present in only a small number of pixels is a problem similar to subpixel target detection for solid objects in a scene. [16] Weak plumes are expected to be associated with radiance changes of a few percent or less from on-plume to off-plume pixels. In this case, ICA is not used to *find* the plume, but to characterize the clutter in an effective manner so that matched filter approaches [13,17] can be used to more effectively locate and characterize the plume.

The tendency for ICA to produce coordinates which highlight specific spectral features of the data leads to components that act almost like a partition of the data. ICA is not a clustering algorithm, but this feature that is often exhibited by ICA solutions provides an informal sense that ICA directions are "more physical" than, for instance, the directions produced by PCA. This is not a mathematical statement, but it motivates the use of ICA in physical retrievals from remote sensing imagery.

ICA in its ordinary form is not necessarily well suited to unmixing of hyperspectral data, as others have commented. [18] The ICA basis vectors will not necessarily match the pure components. [8] We found this to be true in our simulations. However, if one wishes to use the ICA basis vectors to detect weak targets in a scene (such as a chemical plume), the basis vectors are useful even if not perfectly matched to the scene components. Whether they are more useful than PCA components remains to be seen. Variants of ICA that have been described by others [5,8] may be useful in thermal IR imagery where independence of the data is not completely satisfied.

6. CONCLUSIONS

The limited examples displayed in this paper demonstrate that ICA is able to extract spectrally distinct features from cluttered scenes produced by passive and active sensors. Similar results have been reported for sensors operating in the reflective regions of the spectrum. [5-9] Our results show utility in the thermal IR region. Images of the ICA output for the extracted components show qualitative differences from PCA in the sense that individual objects are highlighted separately, at least in some cases. There are many more materials present in the data of Fig. 1 that were not separated successfully, so there are limits on this methodology. Instrument artifacts (e.g. bad pixels or faulty radiance calibration) are also revealed in the independent component images, which could be a problem or an opportunity to at least detect them.

ACKNOWLEDGEMENTS

REFERENCES

1. J. R. Schott, *Remote Sensing: the Image Chain Approach* (Oxford, 1997).
2. L. Kirkland, K. Herr, E. Keim, P. Adams, J. Salisbury, J. Hackwell, A. Treiman, "First use of an airborne thermal infrared hyperspectral scanner for compositional mapping," *Remote Sens. Environ.* **80,** 447-459, 2002.
3. A. Hyvärinen and E. Oja, "Independent component analysis: algorithms and applications," *Neur. Net.*, **13**, 411-430, 2000.
4. A. Hyvärinen , J. Karhunen, and E. Oja, *Independent Component Analysis*. New York: Wiley, 2001.
5. A. Kaarna, P. Zemcik, H. Kalviainen, and J. Parkkinen, "Compression of multispectral remote sensing images using clustering and spectral reduction," *IEEE Trans. Geosci. Rem. Sens.*, **38**, 1073-1082, 2000.
6. T. Tu, P. Huang, and P. Chen, "Blind separation of spectral signatures in hyperspectral imagery," *IEEE Proc. Vision Image Sig. Proc.*, **148**, 217-226, 2001.
7. T. Wachtler, T. Lee, and T. Sejnowski, "Chromatic structure of natural scenes," *J. Opt. Soc. A*, **18**, 65-77, 2001.
8. L. Parra, C. Spence, P. Sajda, A. Ziehe, and K. Muller, "Unmixing hyperspectral data," *Advances in Neural Information Processing Systems*, **12**, 942-948, 2000.
9. J. Chen and X. Wang, "A new approach to near-infrared spectral data analysis using independent component analysis," *J. Chem. Inf. Comp. Sci.*, **41**, 992-1001, 2001.

10. B.R. Foy, R.R. Petrin, C.R. Quick, T. Shimada, and J.J. Tiee, "Comparisons between Hyperspectral Passive and Multispectral Active Sensor Measurements," *SPIE Proc.* **4722**, 98-109, 2002.
11. A. Hayden, E. Niple, and B. Boyce, "Determination of trace-gas amounts in plumes by the use of orthogonal digital filtering of thermal-emission spectra," *Appl. Opt.*, **35**, 2802-2809, 1996.
12. J. Harsanyi and C. Chang, "Hyperspectral Image Classification and Dimensionality Reduction - An Orthogonal Subspace Projection Approach," *IEEE Trans. Geosci. Rem. Sens.*, **32**, 779-785, 1994.
13. P. Villeneuve, H. Fry, J. Theiler, W. Clodius, B. Smith, and A. Stocker, "Improved matched-filter detection techniques," *SPIE Proc.*, **3753**, 278-285, 1999.
14. B.R. Foy, B.D. McVey, R.R. Petrin, J.J. Tiee, and C.W. Wilson, "Remote Mapping of Vegetation and Geological Features by Lidar in the 9-11 μm Region," *Appl. Opt.,* **40,** 4344-4352, 2001.
15. A. Ifarraguerri and C. Chang, "Unsupervised hyperspectral image analysis with projection pursuit," *IEEE Trans. Geosci. Rem. Sens.*, **38**, 2529-2538, 2000.
16. D.W.J. Stein et al., "Anomaly Detection from Hyperspectral Imagery," *IEEE Signal Processing Magazine,* Jan. 2002, pp. 58-69.
17. C. Funk, J. Theiler, D. Roberts, and C. Borel, "Clustering to improve matched filter detection of weak gas plumes in hyperspectral thermal imagery," *IEEE Trans. Geosci. Rem. Sens.*, **39**, 1410-1420, 2001.
18. J. Nascimento and J. Dias, "Does Independent Component Analysis Play a Role in Unmixing Hyperspectral Data?," *Lecture Note in Computer Science* **2652**, 616-625, 2003.