Hyperspectral image quality for unmixing and subpixel detection applications

John P. Kerekes* and Daniel S. Goldberg
Digital Imaging and Remote Sensing Laboratory, Chester F. Carlson Center for Imaging Science,
Rochester Institute of Technology, 54 Lomb Memorial Drive, Rochester, NY, USA 14623

ABSTRACT

The quality of remotely sensed hyperspectral images is not easily assessed visually, as the value of the imagery is primarily inherent in the spectral information embedded in the data. In the context of earth observation or defense applications, hyperspectral images are generally defined as high spatial resolution (1 to 30 meter pixels) imagery collected in dozens to hundreds of contiguous narrow ($\lambda/\Delta\lambda$~100) spectral bands from airborne or satellite platforms. Two applications of interest are *unmixing* which can be defined as the retrieval of pixel constituent materials (usually called endmembers) and the area fraction represented by each, and *subpixel detection*, which is the ability to detect spatially unresolved objects. Our approach is a combination of empirical analyses of airborne hyperspectral imagery together with system modeling driven by real input data. Initial results of our study show the dominance of spatial resolution in determining the ability to detect subpixel objects and the necessity of sufficient spectral range for unmixing accuracy. While these results are not unexpected, the research helps to quantify these trends for the situations studied. Future work is aimed at generalizing these results and to provide new prediction tools to assist with hyperspectral imaging sensor design and operation.

Keywords: Hyperspectral imaging, target detection, unmixing, image quality

1. INTRODUCTION

The quality of remotely sensed hyperspectral images is not easily assessed visually, as the value of the imagery is primarily inherent in the spectral information embedded in the data. In the context of earth observation or defense applications, hyperspectral images are generally defined as high spatial resolution (1 to 30 meter ground pixels) imagery collected in dozens to hundreds of contiguous narrow ($\lambda/\Delta\lambda$~100) spectral bands from an airborne or satellite platform. These data are often represented visually in a cube with a natural color image on the face and the depth representing the vector of spectral information collected for each pixel.

Hyperspectral images have many applications, but two of particular interest are uniquely enabled by their high spectral resolution and dimensionality. It is generally recognized that nearly all real imagery at the meter scale considered in this work have pixels that are mixtures of multiple materials or surface cover types of interest. This naturally leads to the application of *unmixing* which can be defined as the retrieval of pixel constituent materials (usually called endmembers) and the area fraction represented by each for every pixel in the image[1]. The relatively coarse spatial resolution also leads to the application of *subpixel detection*, which is the ability to detect the presence of spatially unresolved materials or objects in a pixel using the spectral information[2]. It is in the context of these applications that the quality of a hyperspectral image can be studied.

Image quality in general has been studied extensively, with much of the work focusing on evaluations made by human observers. This is because most images are produced for human viewing and the quality of an image in many instances is based on subjective interpretation. Quantitative metrics documenting the effects of image compression on image quality have been developed, as in these cases there is a reference (the original) to which compare the compressed versions. In all cases, image quality is an important topic in the development and engineering of imaging systems. Quantitative metrics of image quality can be used to make design choices and optimize performance of the system. In the context of remote sensing, a robust image quality metric can be used to index an image archive and, together with a predictive model, can be used to arbitrate amongst competing requests for imaging limited areas during airborne or satellite collections. The desire for a quantitative metric with a predictive capability has motivated much of our work in this area.

*kerekes@cis.rit.edu

Image Quality and System Performance X, edited by Peter D. Burns, Sophie Triantaphillidou,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 8653, 865304 • © 2013 SPIE-IS&T
CCC code: 0277-786X/13/$18 • doi: 10.1117/12.2001693

In this paper we review previous publications in hyperspectral image quality, address the system parameters that can affect quality, and present results of empirical and model-based analyses that provide some insight into the nature of hyperspectral image quality.

2. RELATED WORK IN HYPERSPECTRAL IMAGE QUALITY

As mentioned above, the value of hyperspectral imagery is most often extracted using automated algorithms rather than through visual interpretation. Also, there are many different applications for hyperspectral imagery in addition to the unmixing and subpixel detection applications mentioned earlier. These include surface cover type mapping, material identification and characterization, and even atmospheric parameter retrieval. This diversity of uses and the plethora of analysis algorithms has lead many to view the topic of hyperspectral image quality to be overwhelmingly complex and has stymied research on the topic. Unfortunately, relatively little work has been done in the area.

However the work that has been done can trace its roots back to the interpretation of airborne and satellite imagery by human analysts. This lineage is in part justified by the common motivations of having a quantitative metric with a predictive model for use during design and development, during operation, and to index an image archive. The following reviews the origins of remote sensing image quality and reviews related work in its application to hyperspectral imagery.

2.1 Leachtenauer, et al, 1997[3]

This work addressed the utility of panchromatic remote sensing images as interpreted by human analysts. Versions of the National Imagery Interpretability Rating Scale (NIIRS) were developed for defense and civilian applications. Images were rated on a numerical scale of 0-9, with higher numbers indicating higher utility for the image. (The distinction between utility and quality will be addressed below). In this work, the authors developed a quantitative relationship between imaging system parameters such as ground sample distance (GSD), sensor signal-to-noise ratio (SNR), relative edge response (RER) and noise gain resulting from image sharpening. The coefficients of the functional relationship between the system parameters and the NIIRS ratings were based on a regression between analyst-rated NIIRS values and the system parameters for the corresponding image. This General Image Quality Equation (GIQE) provided a quantitative method to assess and predict the NIIRS rating for a given image.

2.2 Kerekes and Hsu, 2004[4,5]

Using the GIQE as a model, Kerekes and Hsu developed functional relationships between hyperspectral imaging system parameters and metrics for two applications: target detection and terrain classification. In particular, the system parameters used in the relationship were ground resolved distance (GRD), SNR, and the number of channels N across the solar reflective spectral region. The number of channels was used as a surrogate for spectral resolution. For target detection, they selected a 0-9 scale for a metric called the Spectral Quality Rating Scale (SQRS), modeled after NIIRS, and a Spectral Quality Equation (SQE) which for the panchromatic case of N=1 would predict values nearly equivalent to those predicted by the GIQE. For terrain classification, they used the Kappa Statistic as the quality metric, as it captures the composite accuracy of a multiclass classification confusion matrix in a single number. For both applications, the equations captured an imaginary "constant performance surface" where the quality metrics would be constant as GRD, SNR, and N trade off. The functional relationships were developed using an analytical performance prediction model[6] for the target detection application and empirical airborne hyperspectral data for the terrain classification application.

2.3 Christophe, et al, 2005[7]

This work was concerned with quantifying the effect of lossy compression on the quality of hyperspectral imagery. Thus, they were able to work with approaches that look at the difference between a reference (or original) image and a reconstructed one. They studied 15 difference quality criteria and found that 5 did a good job of capturing different types of distortions with some being more or less sensitive to particular distortion types.

2.4 Stefanou and Kerekes, 2009[8], 2010[9]

Stefanou and Kerekes explored the topic of quality and followed the notion that quality has two components: fidelity and utility. In this approach, fidelity refers to how faithfully the image reproduces the original scene, while utility quantifies the ability of an image to satisfy performance requirements for a well-defined task. Thus the work by Christophe, et al, would be characterized as addressing the fidelity of the compression process. Stefanou and Kerekes pursued the utility of hyperspectral images in the particular context of subpixel target detection. They proposed a method of implanting targets

in an image and then running a target detection algorithm to assess the utility of the image for a particular target. They then proposed a model-based approach to predicting image utility that used statistics computed over an image to predict the utility rather than assessing it through the target implant method.

3. SYSTEM PARAMETERS

Since our work pursues hyperspectral image quality in the context of a well-defined task (this is really image utility as defined by Stefanou and Kerekes), we must consider the entire imaging system, from the source of the photons to the final product. This is because the ability to accomplish a specified task will ultimately depend on all aspects of the sensing process, including the algorithms applied. The following subsections address these parameters further by dividing the hyperspectral imaging system into the scene, the sensor, and the processing components.

3.1 Scene Parameters

It is obvious the content of a scene will affect the utility of an image. If the task is to find a green cotton panel and the scene is entirely made up of a brown desert, then the scene will make this task fairly easy. On the other hand, if the background is a green forest, the task will be more difficulty. Beyond this obvious aspect, other scene parameters will play into many hyperspectral imaging tasks. These include the scene complexity (number of distinct surface classes), the sun angle (time of day), the presence of haze, fog, or clouds, and the size and variability of objects of interest.

3.2 Sensor Parameters

In many discussions of image quality (or utility), the characteristics of the imaging system come to mind as the first aspect affecting the quality (or utility) of an image. These characteristics include the optical design parameters such as aperture size, focal length, and focal plane array size, as well as derived values such as spatial and spectral resolution, and noise levels. Sensor artifacts such as dead detectors, band misregistration, and calibration errors (spectral and radiometric) also can affect image utility.

3.3 Processing Parameters

This component of an imaging system is often not recognized as contributing to image quality. However, since we are focused on the utility of the image for a task, the processing analysis cannot be separated from the rest of the imaging system. Parameters in this component include the specific algorithm sequence applied and their selectable thresholds or attributes. Differences in utility can even be attributed to differences among human analysts who may process an image with various techniques.

4. EMPIRICAL ANALYSES

In this section we report preliminary results from an empirical airborne hyperspectral collection using targets specifically designed to quantitatively assess spectral unmixing accuracy. These data were analyzed using different spectral band combinations and the results interpreted relative to the system parameters and the notions of utility described above.

4.1 Ground Test Targets

In the application of linear spectral unmixing, each pixel in a scene is decomposed into fractional abundances of constituent components, or endmembers. In natural scenes it is difficult to know the true abundance in a pixel and therefore difficult to quantitatively asses performance. This motivated the design of special test targets. The fractional abundances of different materials present in the targets were calculated using an unconstrained least squares regression technique. Whole pixel samples of the target materials were provided to simplify the endmember extraction step.

Two square unmixing targets were designed: a yellow repeating pattern of two different materials with a 3:1 ratio, and a blue repeating checkerboard pattern with two different materials and a 1:1 ratio. The yellow target was 16ft (4.9m) on a side while the blue target was 24ft (7.3m) on a side. The repeating pattern was smaller than the width of an imaged pixel to yield mixed pixels when captured by the system. Due to the repeating design, the expected ratio is contained in a single pixel no matter where the target falls on the pixel array. The yellow target was expected to show 75% and 25% fractional abundances for a target pixel while the blue target was expected to show 50% and 50% fractional abundances for a target pixel. Six whole pixel targets were also included in the ground tests to be used in the endmember selection process. The details of this experiment will be presented in a future publication[10]. The arrangement of the targets is shown below in Figure 1.

Figure 1. Image of the ground targets. The yellow unmixing target is in the upper left while the blue unmixing target is in the upper right. Four of the whole pixel targets are shown in the foreground.

4.2 Airborne HSI Data

SpecTIR flew their ProSpecTIR-VS hyperspectral instrument in an experiment conducted near Rochester, New York in September 2012[11]. The ProSpecTIR-VS instrument has two sensors, which cover two different spectral regions. The first covers the visible and near-infrared (VNIR) wavelengths of 400 to 1000nm in 126 bands and the second covers the short-wave infrared (SWIR) wavelengths of 1000 to 2500nm in 234 bands. Combined, the sensors provide 360 contiguous spectral bands with spectral resolutions of 2.9nm and 8.5nm for the VNIR and SWIR sensors respectively. SpecTIR flew the instrument so that it obtained a ground resolution of approximately 1m. SpecTIR imaged the ground test targets shown in the image on the left in Figure 2.

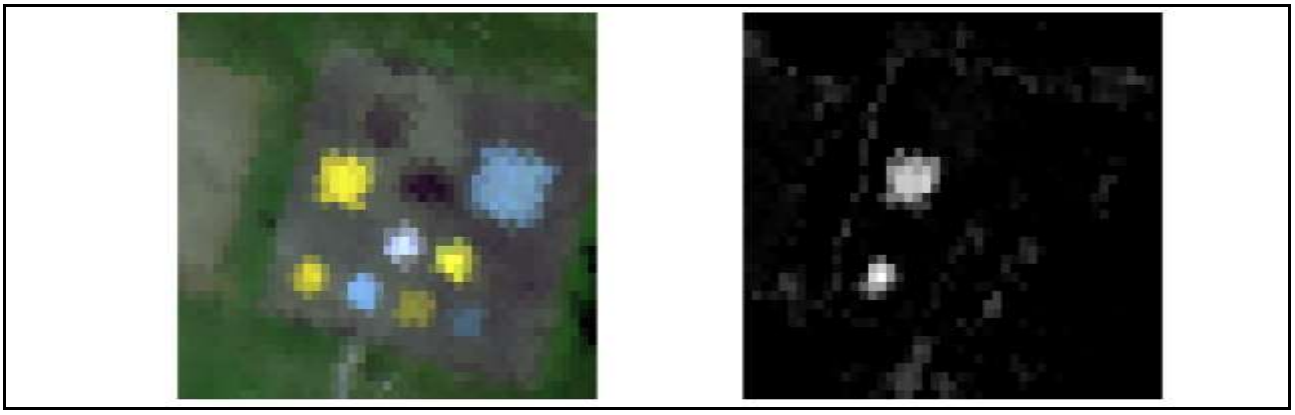

Figure 2. Natural color representation of the scene captured by the ProSpecTIR-VS instrument (left) showing the asphalt court with targets, sand court without targets, and grass. The fractional abundances for each pixel of the unmixing results for one of the two mixed materials in the yellow target is shown as a grey level intensity (right).

4.3 Results

As discussed in section 4.1, unconstrained linear spectral unmixing was performed on the data. To obtain endmember spectra, we selected from the scene single pixels each representing a pure material. Each of the six cloth materials were used along with grass, asphalt, and sand making up nine total endmembers. The unmixing results are shown in Table 1. We expected the mean fractional abundances to be 0.75, 0.25, 0.50, and 0.50 for each of the four materials respectively. To measure how sensor specifications and limitations affected the results, we tried different subsets of spectral bands (VNIR, SWIR, and the full spectral range). Next, we averaged over various target areas, starting from the full area of a target (where some material was detected) and moving in increments of one pixel toward the center of the target, to test the effect of the system's point spread function on unmixing. Finally, we down sampled the data from 360 bands to 36 spectral bands evenly spaced over the full spectral range.

Table 1. Retrieved fractions of materials found using unconstrained spectral unmixing. The mean and standard deviations are from the set of unmixed fractions from pixels in the target area. The full area includes every pixel the target touched.

		Yellow 1		**Yellow 2**		**Blue 1**		**Blue 2**	
		Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
True Fractions		0.75		0.25		0.50		0.50	
Full Range 360 Bands	Full Area	0.33	0.30	0.12	0.18	0.26	0.25	0.27	0.29
	1 pixel in	0.56	0.21	0.21	0.17	0.41	0.19	0.43	0.22
	2 pixels in	0.76	0.057	0.24	0.087	0.51	0.011	0.53	0.026
	3 pixels in	0.77	0.016	0.26	0.005	0.52	0.007	0.53	0.021
Full Range Down sampled to 36 bands	3 pixels in	0.77	0.033	0.27	0.006	0.55	0.017	0.47	0.040
VNIR Only (126 bands)	3 pixels in	0.54	0.077	0.42	0.066	0.47	0.020	0.57	0.060
SWIR Only (234 bands)	3 pixels in	0.56	0.14	0.093	0.037	0.47	0.020	0.57	0.060

4.4 Discussion

The results in Table 1 show that the expected pixel fractions were only achieved after we restricted the area for the statistics to be two to three pixels in from the edge of the target. The standard deviations drop from about 100% of the mean to about 5% of the mean by reducing the area considered. This shows that the point spread function of the system allows contributions to a pixel from up to 3 pixels away.

Next, when comparing the down sampled results with the results using all bands, there was only a modest change in the retrieved fractions. This shows that bad bands or relatively poor spectral resolution may not be as much of a hindrance in this specific application. Finally, comparing the different spectral regions (VNIR versus SWIR), it can be seen that the neither region alone did well on the yellow target, but both did reasonably well on the blue target. This result will be further investigated to better understand the difference. In summary of the spectral band comparisons, it was observed that spectral range was more important for this case than the spectral resolution or number of bands.

5. MODEL-BASED ANALYSES

While empirical analyses are essential for testing and verifying results, model-based analyses can allow the study of many different possible scenarios and are essential for trade-off studies. Our work uses a previously developed model[6] to study trends in the subpixel detection problem.

5.1 Scenario for Analysis

We considered the task of detecting a green vehicle in hyperspectral imagery using a known target spectrum in a spectral matched filter detection algorithm. The scene was modestly complex with 3 backgrounds (trees, grass, and a road). The reflectance statistics used in the model were obtained from empirical measurements. The sun was at a solar zenith angle of 60°, and the atmosphere was moderately hazy with 10 km visibility. We used a typical hyperspectral sensor with approximately 150 spectral bands across the visible through shortwave infrared. The parameter we studied here was sensor SNR, adjusted by adding various amounts of noise with a standard deviation that is a fraction of the mean signal level, thus achieving a constant SNR over wavelength.

5.2 Results

Figure 3 presents the results in two ways. On the left we present the receiver operating characteristic (ROC) curves when the target vehicle occupies 25% of a pixel for five different SNR levels. On the right we show a curve depicting the minimum target fill fraction as a function of SNR to achieve a designated minimum level of performance defined as P_D = 0.8 and $P_{FA} = 10^{-5}$.

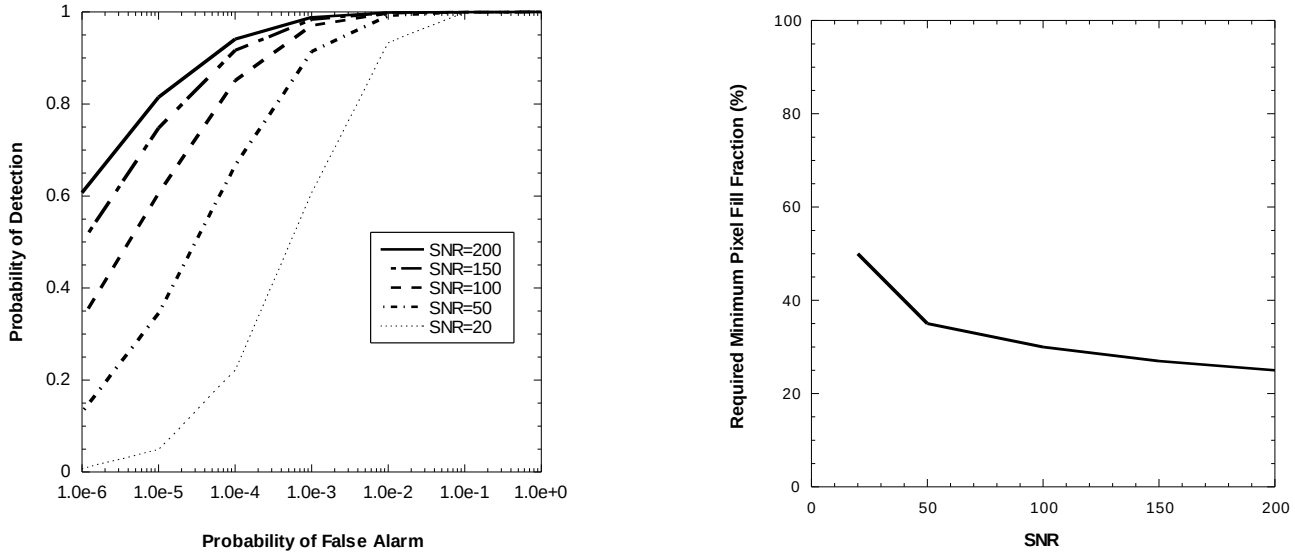

Figure 3. ROC curve for target with 25% pixel fill fraction for several levels of sensor SNR (left). Minimum pixel fill fraction necessary for good detection as a function of sensor SNR (right).

6. SUMMARY AND FUTURE WORK

This paper discussed issues associated with hyperspectral image quality. A review of previous related work is provided demonstrating a diversity of approaches to what is a challenging problem. A distinction was made between the comprehensive term quality, and the two aspects that make up the concept of quality: fidelity and utility. This work primarily considered the utility of a hyperspectral image in that it addressed system effects on two example applications: spectral unmixing and subpixel detection. A discussion of system parameters that affect hyperspectral image utility was presented, making the point that all aspects of the imaging system (scene, sensor, and processing) can affect utility in a given task.

Results were presented showing how various system parameters can affect the performance of a hyperspectral imaging system in the two example applications. An empirical analysis demonstrated the accuracy of an HSI system to retrieve the subpixel fractions of different materials. In particular, we found that good results were obtained when using spectral bands across the full reflective VNIR/SWIR spectrum, but for one target poor results were observed when just using the VNIR or just the SWIR spectral regions. We also observed the significant effects of the sensor point spread function near edges of the target, showing that edges up to 3 pixels away can affect the retrieved abundance of the target materials.

These empirical results were extended by a brief modeling study that looked at the effect of sensor SNR on the ability of an HSI system to detect subpixel targets. We found a nonlinear relationship between sensor SNR and the minimum pixel fill fraction required for a specified level of performance. This result is an example of the type of trade studies that modeling can support, and that can be used to gain insight into the anticipated performance and utility of an imaging system.

Future work will involve additional analyses of the airborne empirical data and the use of the empirical data to further validate the system model. The model will then be further exercised to extend the insights gained using the empirical data. It is envisioned that through continued development of HSI performance models and the insight gained from many

additional studies the community will advance toward a common understanding of hyperspectral image utility, fidelity, and quality.

ACKNOWLEDGMENTS

The authors would like to acknowledge the efforts of several colleagues in fabricating, deploying, and measuring ground truth for the targets used in the experiment: Bo Ding, Kim Horan, Ming Li, Kyle Ludgate, Cara Perkins, Nina Raqueno, and Jiashu Zhang.

REFERENCES

[1] Keshava, N. and Mustard, J.F., "Spectral unmixing," IEEE Signal Processing Magazine 19(1), 44-57 (2002).
[2] Manolakis, D., Siracusa, C., and Shaw, G., "Hyperspectral subpixel target detection using the linear mixing model," IEEE Trans. on Geosci. and Rem. Sens., 39(7), 1392-1409 (2001).
[3] Leachtenauer, J.C., Malila, W., Irvine, J., Colburn, L., and Salvaggio, N., "General image quality equation: GIQE," Applied Optics, 36(32), 8322-8328 (1997).
[4] Kerekes, J. and Hsu, S., "Spectral quality metrics for VNIR and SWIR hyperspectral imagery," in Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery X, Proc. SPIE 5425, 549-557 (2004).
[5] Kerekes, J. and Hsu, S., "Spectral quality metrics for terrain classification," in Imaging Spectrometry X, Proc. SPIE 5546, 382-389 (2004).
[6] Kerekes, J. and Baum, J., "Spectral imaging system analytical model for subpixel object detection," IEEE Trans. on Geosci. and Rem. Sens., 40(5), 1088-1101 (2002).
[7] Christophe, E., Leger, D., and Maihes, C., "Quality criteria benchmark for hyperspectral imagery," IEEE Trans. on Geosci. and Rem. Sens., 43(9), 2103-2110 (2005).
[8] Stefanou, M. and Kerekes, J., "A method for assessing spectral image utility," IEEE Trans. on Geosci. Rem. Sens., 47(6), 1698-1706 (2009).
[9] Stefanou, M. and Kerekes, J., "Image-derived prediction of spectral image utility for target detection applications," IEEE Trans. on Geosci. Rem. Sens., 48(4), 1827-1833 (2010).
[10] Kerekes, J., Ludgate, K., Giannandrea, A. and Raqueno, N., "SHARE 2012: subpixel detection and unmixing experiments," abstract submitted to Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX, Proc. SPIE (2013).
[11] Giannandrea, A. et al, "The SHARE 2012 data collection campaign," abstract submitted to Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XIX, Proc. SPIE (2013).