A Method for Assessing Spectral Image Utility

Marcus S. Stefanou, *Member, IEEE*, and John P. Kerekes, *Senior Member, IEEE*

***Abstract*—The utility of an image is an attribute that describes the ability of that image to satisfy performance requirements for a particular application. This paper establishes the context for spectral image utility by first reviewing traditional approaches to assessing panchromatic image utility and then discussing differences for spectral imagery. We define spectral image utility for the subpixel target detection application as the area under the receiver operating curve summarized across a range of target detection scenario parameters. We propose a new approach to assessing the utility of any spectral image for any target type and size and detection algorithm. Using six airborne hyperspectral images, we demonstrate the sensitivity of the assessed image utility to various target detection scenario parameters and show the flexibility of this approach as a tool to answer specific user information requirements. The results of this investigation lead to a better understanding of spectral image information vis-à-vis target detection performance and provide a step toward quantifying the ability of a spectral image to satisfy information exploitation requirements.**

***Index Terms*—Hyperspectral imagery, object detection, spectral image analysis, spectral image utility.**

I. Introduction

ADVANCES IN electro-optical spectral imaging—the process of collecting spatially coregistered images in multiple spectral bands—have led to increased spectral and spatial resolution. The finer sampling and improved signal-to-noise characteristics of future imaging spectrometers promise enhanced performance in applications such as land cover classification, mineral exploration, atmospheric characterization, and spatially unresolved object detection. However, without a means of quantifying the usefulness of a spectral image for a particular application, we cannot fully appreciate the additional information inherent in this higher resolution spectral imagery. Toward this end, we seek a spectral image utility metric that quantifies the ability of an image to satisfy application-specific performance measures. We postulate that the ability to consistently assess spectral image utility is an important step toward advancing our understanding of spectral image information content, and is a topic worth considering now, before the next generation of imaging spectrometers becomes fully operational.

A. Image Utility Definition and Motivation

The term "image utility" quantifies the ability of an image to satisfy performance requirements for a well-defined task. More generally, it describes the ability of the image to deliver information about the object being imaged. Thus, any metric of image utility is fundamentally an application performance measure. The subtle difference between an image utility metric and an application performance measure is based on one's perspective: image utility focuses on labeling a specific image's usefulness in a particular application, whereas a performance measure focuses on describing the performance of the application across many images. While image utility has been studied in the context of panchromatic imagery and predicting image utility, there has been no systematic exploration of the *assessment* of spectral image utility.

There is a compelling need for a spectral image utility metric and the capability to assess images using such a metric. Such a metric could be applied to many images to build a catalog of utility-labeled images covering many image acquisition and application-specific scenarios. The robust ability to assess the image utility for a given application would create the foundation for an image archive indexing scheme. An archive of spectral images, each with a utility label, would then facilitate sensor design trade studies and provide a basis for guiding future image collection activities. These important capabilities, which use current images to optimize future sensor designs and imagery collection activities to acquire the most useful images, will only be possible with a reliable utility metric and a consistent method to assess the utility of many images.

B. Panchromatic Image Utility and Its Assessment

In the optical remote sensing community, the traditional measures of image utility center on the analysis of panchromatic photographs and digital images by human image analysts. The task is to interpret the image and extract meaningful information by using spatial detail in the image. The national imagery interpretability rating scale (NIIRS) is the most widely used image utility measure in this context. It was developed in the early 1970s as a response to the inability of simple measures of physical image quality such as scale and resolution to successfully communicate image interpretability [1]. The NIIRS has ten levels, each defined by specific interpretation tasks which require better spatial resolution with increasing task difficulty. A higher NIIRS represents a more useful image from the perspective of the image analyst extracting information. NIIRS was developed by a group of image analysts who rated a large sample of interpretation tasks in terms of relative difficulty and validated the scale on many images [2]. Using NIIRS, image utility for panchromatic imagery is assessed when image analysts assign a NIIRS rating to an image using their judgment as to the level of interpretability of the particular image. NIIRS has proven to be extremely valuable tool for

Manuscript received November 12, 2007; revised February 26, 2008 and July 18, 2008. First published February 13, 2009; current version published May 22, 2009. This work was supported in part by the U.S. Government.

The authors are with the Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, NY 14623 USA (e-mail: mstefanouafit@gimail.af.mil).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2008.2006364

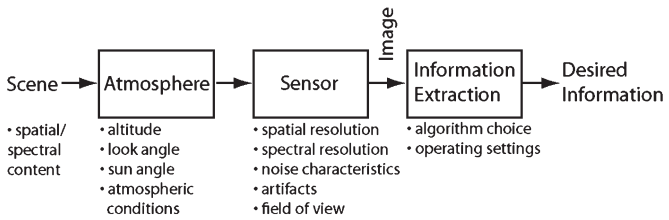

Fig. 1. Image chain and factors that influence image utility.

providing a shorthand description of image interpretability, facilitating communication between image analysts, defining image requirements, selecting and tasking imaging systems, providing quality control feedback to operational systems, and specifying the performance of new imaging systems [3].

C. Spectral Image Utility

Although very helpful in assessing the utility of panchromatic images, the NIIRS does not capture the radiometric and spectral aspects of spectral imagery. This inadequacy is primarily due to the fundamental differences in exploiting the information inherent in each type of image. The information in panchromatic images is extracted by the perception of spatial detail and patterns by a human observer. By contrast, spectral images are processed by computer algorithms operating on the spectrum associated with each pixel individually and exploiting the statistical or subspace characteristics of the image pixel vectors to uniquely identify the material in the corresponding ground resolution cell. Thus, the interaction between the spectral, radiometric, and spatial nature of a spectral image contributes to the ultimate utility of a spectral image [4].

If we adopt the perspective that image utility is another name for the performance metric of a specific application, then we arrive at a spectral image utility metric by using the measures of performance routinely employed to evaluate spectral processing algorithm performance. The simplest method to assess spectral image utility is by applying an information exploitation algorithm to the image. Thus, performance measures such as probability of detection (PD) for target detection tasks and classification accuracy for classification applications can be used for quantifying spectral image utility. We note that there is no analogous NIIRS for spectral imagery (aside from the multispectral image interpretability rating scale [5], which is an extension of NIIRS using human image analysts to multispectral images and does not apply to the higher dimensionality hyperspectral imagery).

While much research effort has gone toward evaluating spectral algorithm performance, an open and interesting research topic is the general *assessment* of image utility for a particular spectral image. We note that the assessed utility will be influenced by factors at every step of the image chain, from the scene being imaged to the information to be extracted from the image. Fig. 1 shows some of these factors in a simple manner, but in general, they may be lumped into two groups: those pertaining to the acquisition of an image (scene, atmosphere, and sensor) and those associated with extracting information from the image (information extraction approach and desired information). Note that an image represents a sampling of the infinite image acquisition space. Our perspective in this research is that we begin with an image and therefore do not have control over the parameters that govern image acquisition. Our control begins with the information extraction processes, and it is here that we seek to define the concept of image utility based on a specific application.

D. Objectives

Our goal is to develop a method to assess the utility of any spectral image for subpixel target detection and apply the methodology to real airborne hyperspectral images. We choose subpixel target detection because it is a challenging application with performance metrics that are readily adaptable to an image utility metric. Thus, referring to Fig. 1, we take an image, which represents a unique and fixed combination of scene, atmosphere, and sensor parameters, and apply our image utility assessment methodology to the image in order to label it with a utility metric. The generation of our utility metric rests on our ability to control the parameters in the information extraction block of Fig. 1 and then summarize their effect on subpixel target detectability. The rest of this paper is organized as follows: Section II discusses the theory and implementation of our spectral image utility assessment approach, Section III presents the assessed utility of six hyperspectral images collected by four airborne imaging spectrometers, and Section IV offers concluding remarks and future direction.

II. Approach

A. Methodology for Assessing Spectral Image Utility

In target detection applications, the performance of an algorithm operating on a specific image for a given target is typically described using the receiver operating characteristic (ROC) curve. The ROC curve expresses the PD for a range of probabilities of false alarm. The probabilities plotted in the ROC curve are derived from the output of the target detection algorithm. The ROC curve will serve as the point of departure for developing our spectral image utility metric.

In order to generate the ROC curve, the output of the detection filter is needed for both target absent and target present cases. However, the target present case will be unattainable unless we have ground truth about the target locations in the image. Furthermore, if the target is only present in a small number of image pixels, then the estimate of the PD will result in larger ROC curve confidence intervals [6]. Finally, if the target is not present in the image, then we do not have a target present case and cannot evaluate detection performance. Given our goal of assessing the utility of *any* image, these constraints must be considered the norm. In order to overcome these challenges, we adopt ideas from [7] and [8] to *fractionally implant* a target spectrum in every spatial pixel of the image in order to generate the target present case. We call this method of assessing image utility for subpixel target detection applications the *target implant method*. The resulting ROC curve is a summary of the overall detectability of the implanted target across the entire image. The target implant method offers a flexible means to assess the performance of detection algorithms for a

wide variety of targets and enables the calculation of a spectral image utility metric for images with no targets, limited targets, or no target ground truth.

Fig. 2 shows an overview of the target implant method for assessing spectral image utility in the target detection task. The problem is posed as a binary hypothesis test. The top left part of Fig. 2 illustrates the operations necessary to obtain the filter output for each of the hypotheses (cases). Let the spectral image have K spectral channels, so that a pixel at a given spatial location in the image is represented as the $K \times 1$ vector $\mathbf{x}$. We create the target absent case by applying the detector operator D to each pixel of the image

$$y_{\mathrm{TA}} = D(\mathbf{x}). \tag{1}$$

The scalar y_{TA} is the output in the target absent case at one pixel location. The target present case is created by applying the detector to each pixel location of the original image in which the $K \times 1$ target vector has been implanted, resulting in the scalar output y_{TP}

$$y_{\mathrm{TP}} = D(\mathbf{x}_{\mathrm{TP}}). \tag{2}$$

The implantation of the target is accomplished in a fractional manner using the prescribed subpixel fraction f in each pixel of the image as

$$\mathbf{x}_{\mathrm{TP}} = f\mathbf{t}' + (1-f)\mathbf{x}. \tag{3}$$

The $K \times 1$ implanted target pixel vector $\mathbf{x}_{\mathrm{TP}}$ is created by first realizing a $K \times 1$ random vector $\mathbf{t}'$ from a normal random process described by the statistical parameters of the $K \times 1$ target mean vector $\mathbf{t}$ and $K \times K$ covariance matrix Σ_T. These parameters are drawn from a reference library created by careful collection of known target pixel vectors from 210 band HYperspectral Digital Imagery Collection Experiment (HYDICE) imagery exhibiting unimodal normal statistics [9]. This target vector variability is a departure from the target implant method described in [8], which treats the target deterministically rather than stochastically. This modification has been adopted to better reflect the variable nature of target vectors of the same type. The random target vector $\mathbf{t}'$ is then mixed *fractionally* with every data pixel to the specified subpixel mixing fraction f. The subpixel mixing fraction may be given a physical meaning because it represents the area ratio of target to image pixel area projected to the ground. If we assume a square target and square image pixels, then the fraction may be expressed in terms of the linear dimension of the target l and the linear dimension of the ground resolved distance (GRD)

$$f = \frac{l^2}{\mathrm{GRD}^2}. \tag{4}$$

The right side of Fig. 2 shows the target absent and target present probability density functions (PDFs) $\hat{p}(y_{\mathrm{TA}})$ and $\hat{p}(y_{\mathrm{TP}})$, respectively, which are estimated from the detector output histograms for the entire image. Integrating the PDF between a given threshold value T and infinity yields a probability. The integral of the target absent PDF is called the probability of false alarm (PFA), representing the probability

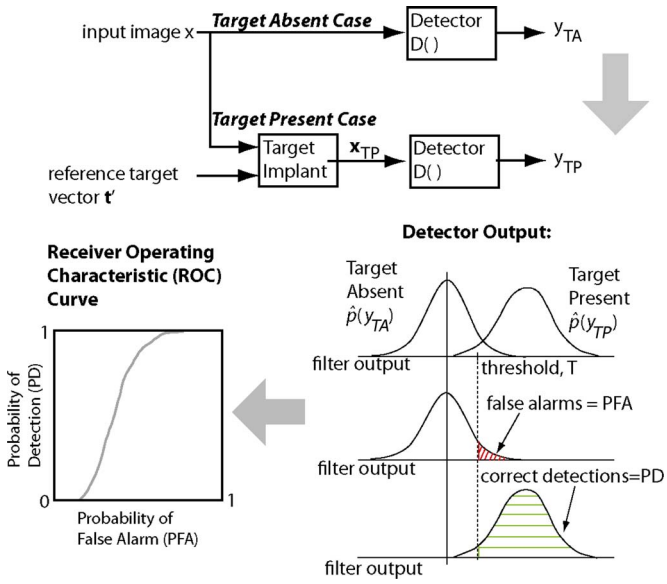

Fig. 2. Target implant method.

that a pixel will be classified as a target when it is really not, and the integral of target present PDF is called the PD, representing the probability that a pixel is correctly classified as a target

$$\mathrm{PFA} = \int_T^\infty \hat{p}(y_{\mathrm{TA}})dy_{\mathrm{TA}} \qquad \mathrm{PD} = \int_T^\infty \hat{p}(y_{\mathrm{TP}})dy_{\mathrm{TP}}. \tag{5}$$

In target detection, it is the relative relationship of these two probabilities, which expresses the goodness of a particular detector or detection scenario, with the most desirous situation being one in which a high PD is achieved at low PFA. The ROC curve, shown on the bottom left side of Fig. 2, captures this relationship. The ROC curve is obtained by plotting PD against PFA for each threshold setting. A ROC curve produced by a perfect detector would consist of a PD of one for all PFAs, a situation created by target absent and present PDFs with no overlap. At the other extreme, a ROC curve associated with a useless detector would consist of a straight line between the PD, PFA pairs (0,0) and (1,1) and correspond to complete overlap of the target absent and present PDFs.

B. Subpixel Target Detection Algorithms

Obviously, a key parameter in the assessed utility of a spectral image is the specific detection operator $D(\cdot)$. The choice of specific detector is driven by user requirements, and performance will be determined by target and background variability, pixel composition (pure or mixed), and how the detector accounts for these factors [10]. In the spirit of affording maximum flexibility to the image analyst in defining utility, we incorporate three detectors into the assessment methodology. The detectors require little *a priori* knowledge about the image, such as knowledge of the background signature, but do require that a reference library containing a target vector is available [11].

The first detector is the spectral matched filter (SMF), because it is linear, reliable, and simple. The SMF is similar to the constrained energy minimization filter of [12] but differs in

that it uses the covariance matrix and subtracts the image mean vector $\boldsymbol{\mu}$ from the target and background vectors. This detector is constructed with the $K \times K$ inverse image covariance matrix $\boldsymbol{\Sigma}^{-1}$ and a $K \times 1$ target mean vector $\mathbf{t}$ from the spectral library

$$D_{\mathrm{SMF}}(\mathbf{x}) = \frac{(\mathbf{t}-\boldsymbol{\mu})^T \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})}{(\mathbf{t}-\boldsymbol{\mu})^{\mathrm{T}} \Sigma^{-1}(\mathbf{t}-\boldsymbol{\mu})}. \tag{6}$$

The filter operates on each image pixel (the $K \times 1$ pixel vector $\mathbf{x}$ with the data mean vector subtracted) and creates a scalar result representing the relative degree to which the pixel vector matches the target vector.

The second detector is the nonlinear adaptive coherence/cosine estimator (ACE), which is derived from the generalized likelihood ratio test approach [13]–[15]

$$D_{\mathrm{ACE}}(\mathbf{x}) = \frac{\left((\mathbf{t}-\boldsymbol{\mu})^{\mathrm{T}} \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)^2}{\left((\mathbf{t}-\boldsymbol{\mu})^{\mathrm{T}} \Sigma^{-1}(\mathbf{t}-\boldsymbol{\mu})\right)\left((\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} \Sigma^{-1}(\mathbf{x}-\boldsymbol{\mu})\right)}. \tag{7}$$

The numerator of this detector is the squared Mahalanobis distance between the demeaned image pixel and the demeaned mean target vector. Both the SMF and ACE employ this distance, but the action of the denominator leads to different interpretations for each detector. In [13], it is shown that, in a whitened space, the SMF represents the distance between an image pixel vector and the target subspace whereas the ACE may be thought of as the angle between them.

Unlike SMF and ACE, the third detector was not designed to account for the variability of the image background or target. It treats the image and target vectors deterministically and assumes that they are spectrally pure. It is also simple and may be applied without any estimation of image statistics. It is the popular spectral angle mapper (SAM) expressed as

$$D_{\mathrm{SAM}}(\mathbf{x}) = \frac{\mathbf{t}^{\mathrm{T}}\mathbf{x}}{(\mathbf{t}^{\mathrm{T}}\mathbf{t})^{1/2}(\mathbf{x}^{\mathrm{T}}\mathbf{x})^{1/2}}. \tag{8}$$

The output of the detector is the cosine of the angle between the image test pixel and the mean target vectors. Note that the image mean vector is not subtracted as in SMF and ACE. We expect that SAM will have difficulty in our application of subpixel target detection but include it as a reference baseline.

C. Summary Utility Metric

The target implant method of assessing utility offers great flexibility because a wide variety of target, target implant fraction (target size), and algorithm choice combinations may be considered for a single image. Each unique combination represents a different *target detection scenario* and results in a ROC curve, which summarizes the probability of detecting the specified target over all false alarm probabilities across every pixel of the image. In this manner, each ROC curve describes the overall utility of the image for a particular target detection scenario. Because the ROC curve is generated by the target implant method which gives independent and equal consideration to every image pixel, it is a desirable means of assessing the utility of many images on a consistent basis.

Our stated goal is to assess the utility of any spectral image for the subpixel target detection application, which implies that we would like to consider more than just one specific target detection situation in arriving at a utility metric. The target detection parameters that we control in the target implant method are the specified PFA at which we want to operate the detection threshold, the particular target we seek, the amount of target implanted in each pixel (which may be considered to be the linear dimension of a square target), and the detection algorithm we use. We seek a means of summarizing a range of target detection scenario parameters into a single metric that will describe how useful a particular image is for the subpixel target detection task.

In order to summarize utility over a range of PFAs, we employ the area under a ROC curve over a PFA interval of interest. The area under curve (AUC) is a widely used figure of merit for detection performance from medical diagnostics [16]. In target detection for spectral images, the operational range of PFA will tend to be rather low, since usually, only a small number of false alarms are acceptable even with many image pixels. In order to better represent this desire to operate at low PFAs, we apply a weighting function $z(\mathrm{PFA})$ to the calculation of the AUC. A simple weighting function is a rectangular window that applies equal weight to all PFAs in a desired operating range (from the lowest achievable PFA to the specified PFA) and zero elsewhere. Any weighting function may be applied to meet the specific requirements of the particular target detection scenario, offering great flexibility. The integral version of the utility is represented as

$$\begin{aligned}&\mathrm{Utility}\,(\mathbf{t}, f, D(\mathbf{x}), \mathrm{PFA}) \\ &\quad = \frac{\int_0^1 z(\mathrm{PFA}) \cdot \mathrm{PD}\,(\mathbf{t}, f, D(\mathbf{x}), \mathrm{PFA})\ d\mathrm{PFA}}{\int_0^1 z(\mathrm{PFA})\ d\mathrm{PFA}}.\end{aligned} \tag{9}$$

Note that the utility is a function of the target type, implant fraction, detector, and PFA by virtue of PD dependence on these factors. In practice, a discrete approximation to this integral is calculated. Note also that the utility is the AUC realized for a particular detection scenario normalized by the AUC associated with a perfect detection situation (modulated by the window function). This has the effect of transforming our AUC into a number between zero and one, thus allowing us to easily compare results in a relative rather than absolute fashion, and on a scale that is bounded like NIIRS.

Using (9), we may generate the utility associated with our PFA range of interest for one target, one fraction, and one detector. If we vary the target, fraction, and detector parameters over a range to create a more robust target detection scenario, we would like to be able to summarize the utility over this range in order to make a statement about the generalized utility of the image. In order to give maximum flexibility to the user in defining utility based on his or her unique requirements, we offer the ability to condense the utilities resulting from a range of parameters to a single summary metric or to leave the ROC areas as an ensemble from which the user can select those of

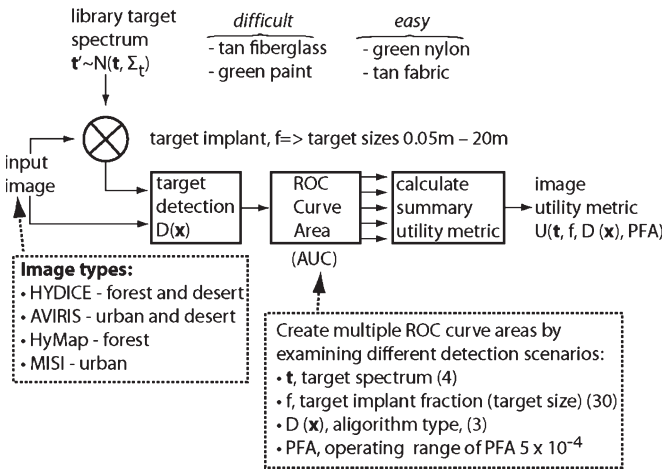

Fig. 3. Spectral image utility assessment and prediction.

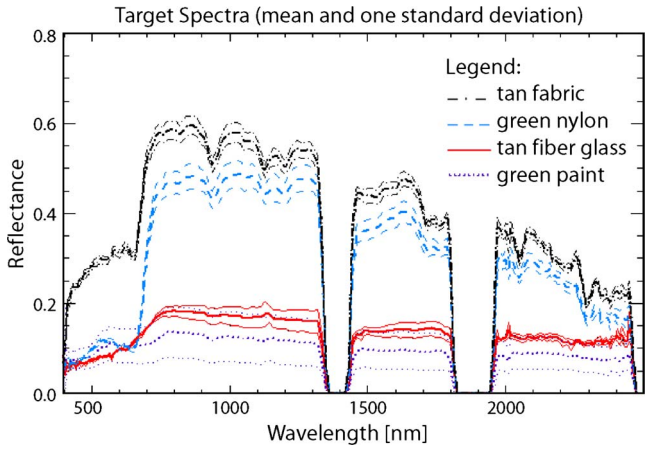

Fig. 4. Target mean spectra with ± one standard deviation.

most interest. The simplest approach to attaining a summary metric is to form a weighted summation of utilities over all L targets, M implant fractions, and N detectors. This situation for the summary utility U_{summary} is shown as

$$\text{Utility}_{\text{summary}}\left(\text{Utility}\left(\mathbf{t}, f, D(\mathbf{x}), \text{PFA}\right)\right) = \sum_{k=1}^{N} w_k \sum_{j=1}^{M} w_j \sum_{i=1}^{L} w_i \text{Utility}\left(\mathbf{t}_i, f_j, D_k(\mathbf{x}), \text{PFA}\right). \tag{10}$$

The summary utility metric is a function of the utility for the combination of detection scenario parameters, which have been explicitly indexed using i for the target, j for the implant fraction, and k for the detector. The selection of an appropriately normalized weighting function w_i, w_i, w_i to achieve the desired selectivity in targets, fractions, or detectors of interest would allow the user to tailor a utility metric best suited to his or her purposes. The final result is a single scalar between zero and one, which represents the degree of utility of a particular image in fulfilling a user-defined target detection scenario(s).

D. Experimental Design

We implement our utility assessment methodology in the manner shown in Fig. 3. The input image is assumed to be in the reflectance domain in the visible/short-wave infrared portion of the spectrum (0.4–2.5 μm). Any spectral image may be input, provided that we know the spectral response function of the sensor that acquired the image so that we can match the reference library target spectrum to that of the image. Six images, representing a range of spatial, spectral, and radiometric resolutions, were chosen from four different sensor types. Four of the images were from the same type of sensor (two from each) in order to also include different scene content for the same sensor type.

For this paper, we specify the four parameters that correspond to the target detection scenario as follows: four targets, 30 subpixel target implant fractions, three detection algorithms, and a specified PFA range from the lowest realizable PFA to a specified PFA of 5×10^{-4}. Thus, for each image, we have 360 utilities that represent a sampling of the infinite target detection scenario space. The four library target spectra were chosen to be representative of both easy and difficult target types. The tan fiberglass and green paint spectra are categorized as "difficult" because they are spectrally flat, dark, and have a relatively large amount of class variability. The tan fabric and green nylon are "easy" because they have large spectral contrast and relatively low class variability. Fig. 4 shows the mean spectra of the targets along with ± one standard deviation. The target implant fractions were chosen in a manner that provided 30 samples across the range from 0.06% to 100%, although the primary consideration was to select fractions that corresponded to a range of target sizes relative to the GRD of the particular image. The smallest target size considered was 0.05 m, and the largest was 20 m. The three detectors considered were the SMF, ACE, and SAM.

E. Image Preprocessing and Descriptions

There are several considerations and preprocessing steps required before employing the image utility assessment method described earlier. First, the image must be in reflectance space in order to employ the reference library target spectra descriptions, which are given in reflectance. Although we could use a radiative transfer model to forward propagate the target into radiance space and then operate on a radiance image, all of the parameters needed to do the forward propagation may not be available in general. Second, we need to operate on those spectral channels free of atmospheric absorption bands, as these may cause spurious responses to the detector. Thus, the spectral channels affected by atmospheric absorption bands must be removed by either using information included with the image header file or estimated directly from the image spectra. Third, we do not remove or limit physically implausible reflectance pixel vectors resulting from atmospheric compensation inaccuracies, although we could employ such preprocessing if needed. Fourth, we are not concerned about targets that may be in the image. If they are present, it will decrease the utility, since they will artificially inflate the target absent detector output, but we cannot assume *a priori* knowledge of targets in the image.

Six spectral images were considered for utility assessment. The two 320 × 960 spatial pixel HYDICE images were

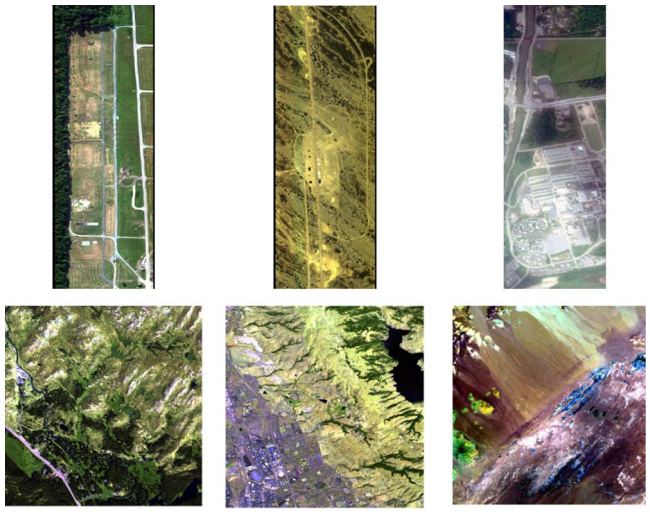

Fig. 5. (Top left) HYDICE FR I run 05. (Top center) DR run 03. (Top right) MISI RIT. (Bottom left) HyMap Cooke City. (Bottom center) AVIRIS Moffett Field. (Bottom right) Lunar Lake.

collected in 1995 during very well-controlled experiments at a sensor altitude of 5000 ft. Feature space transformations removed atmospheric absorption bands, resulting in 144 bands for the Desert Radiance II (DR II) run 03 image and 145 bands for the Forest Radiance I (FR I) run 05 image out of the original 210. Both images were collected with 0.85-m GRD and an average spectral resolution of 10 nm, covering a spectral range from 397–2496 nm. Atmospheric compensation using in-scene calibration panels and the empirical line method (ELM) produced reflectance images. The FR I image is shown on the top left part of Fig. 5. Two airborne visible/infrared imaging spectrometers (AVIRISs) were collected at a sensor altitude of 20 000 m for a 20-m GRD. Retained over the range of 370–2506 nm at a 10-nm average spectral resolution were 183 of the 224 bands. The bottom center part of Fig. 5 shows a scene of Moffett Field, CA. The other AVIRIS image is a desert scene from Lunar Lake, NV. Both AVIRIS images are 512×614 pixels, and they were atmospherically compensated to reflectance values. A 512×512 pixel subset of an image from the HyMap hyperspectral scanner of Cooke City, MT, was collected in 2006 at an altitude of 12 400 ft above sea level for a GRD of 3 m. The sensor has 126 spectral bands, covering the 454–2496-nm spectral region with a 15-nm average spectral bandwidth. Ninety six of the 126 bands are retained, and the image was atmospherically compensated using the HYCORR algorithm. The HyMap image is shown in the bottom left part of Fig. 5. The sixth image is 332×1272 pixels in spatial extent, and it was collected by the modular imaging spectrometer instrument (MISI) over the Rochester Institute of Technology (RIT) campus, Rochester, NY. The GRD is approximately 3 m and of the 35 channels in the 408–738-nm spectral range; 31 are used for this analysis. The spectral resolution is 10 nm, and atmospheric compensation was accomplished via the ELM.

We selected these six images because they represent a sampling of the image acquisition parameters discussed in Fig. 1. The six images have varying scene contents, ranging from relatively simple and uniform desert scenes to more complicated and cluttered urban scenes. Three different GRDs are represented, each exhibiting a different degree of spectral signature mixing and background variability. There are four spectrometers represented, with different image formation approaches (line scanner and pushbroom), signal-to-noise ratios (from high values for AVIRIS to relatively lower for MISI), and image artifacts. The images were atmospherically compensated using different methods, each of which introduces its own source of noise to the observed pixel vectors. While [17] presents a more comprehensive examination of more images and image characteristics, this paper deals with spectral images that span a range of image acquisition parameters.

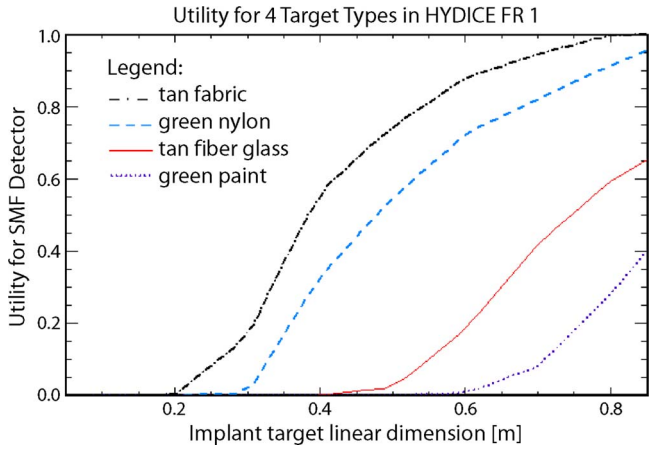

Fig. 6. Utility of HYDICE FR I for four target types using the SMF detector.

III. Results and Discussion

We emphasize two major themes with these results. First, we illustrate the varying degree to which image utility is dependent on target detection scenario parameters. Second, we show that the methodology we propose for assessing image utility provides a flexible tool that may be adopted to meet the specific requirements of any user.

The first target detection scenario parameter that we control is the target type. Fig. 6 confirms our intuition about "easy" and "difficult" targets by plotting the utility associated with the SMF detector for the four targets as a function of the target size for HYDICE FR I image. It is apparent that the green nylon and tan fabric are easier to detect and therefore produce a higher utility for this particular image. The first significant utility for detecting these targets occurs at 0.2–0.3 m, while it does not occur until 0.5–0.7 m for the tan fiberglass and green paint target types. Thus, the choice of target type creates a significant difference in the utility assessed for a particular image. We have chosen to portray the utility as a function of target implant size rather than implant fraction because of the more intuitive nature of the linear dimension rather than area fraction, assuming square targets and image pixels. All utilities in this and subsequent plots and tables have been calculated for PFAs of less than 5×10^{-4}.

If we examine detector algorithm choice on the same image using only the tan fabric target, we see that the detector plays a very large role in the assessed utility. Fig. 7 shows that this image has the highest utility when the target detection scenario includes the ACE detector and the lowest when using the SAM.

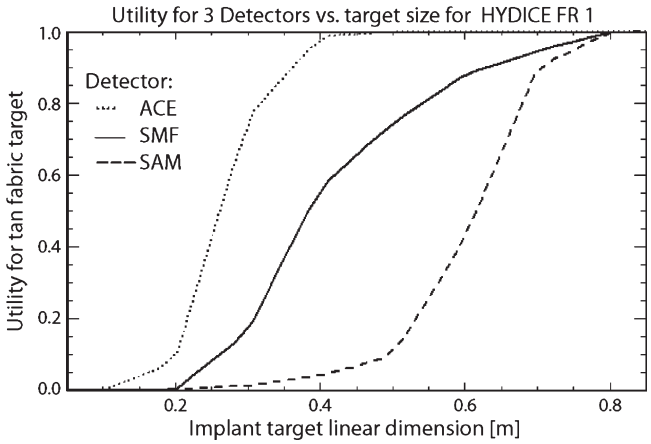

Fig. 7. Utility of HYDICE FR I for three detectors using one target.

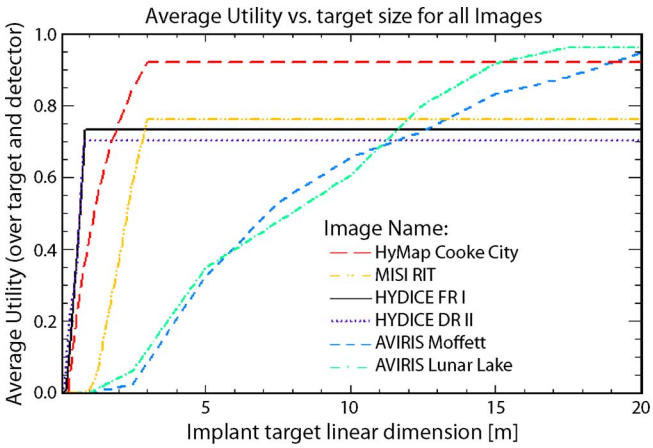

Fig. 8. Summary utility for six images plotted against target size.

While this is not surprising, the point to note is that the choice of algorithm has a substantial impact on the assessed utility of the image. It is a factor that outweighs the particular target type.

The flexibility of the target implant method of assessing utility derives from the fact that varying levels of specificity may be pursued to address specific user questions. For example, if the primary interest was in ascertaining the utility of images for detecting a certain type and size of target, then the user could focus on only those utilities associated with size and type (e.g., green vehicles of size less than 5 m). On the other hand, if the primary interest was to compare images for the utility of general target detection, then the user would seek to average the utility associated with all targets, algorithm types, and target sizes. Fig. 8 shows the utilities associated with this most general scenario in the legend and uses the plot to express the utility averaged over the four targets and three detectors as a function of target size. In calculating the summary utility metric associated with each image, we have made the assumption that targets larger than the GRD of the sensor will be treated as independent occurrences of a 100% full pixel target. In other words, beyond a target size corresponding to the GRD of the sensor, no further improvement of utility is possible. The rationale for this approach is that the detectors treat each pixel independently and are not designed to use spatial information to increase the detection. Thus, it is evident in Fig. 8 that HYDICE utility flattens out at 0.85 m, HyMap and MISI at 3 m, and AVIRIS at 20 m. The summary utility indicates that the HyMap image has the highest utility for detecting these four targets ranging in size from 0.05 to 20 m using the SMF, ACE, and SAM at a PFA of 5×10^{-4}, and the AVIRIS Moffett Field scene has the lowest. Thus, even though AVIRIS is a very high quality sensor, the large GRD and resulting mixing of spectra within each pixel make it very challenging for this sensor to detect small targets. Had we defined our target detection scenario differently, the results might be different. To illustrate this last point about the effect of how the target detection scenario is defined upon the assessed utility, consider that, if we were only interested in finding 20-m targets (under the assumption that each pixel is treated independently by the detector), AVIRIS would then provide the most useful images.

To further reinforce the point that the target detection scenario definition affects the utility metric, consider that we repeat our summary metric but, this time, averaging the target into "easy" (tan fabric and green nylon) and "difficult" (green paint and tan fiberglass) categories. Table I shows the resulting utilities over all detectors and target sizes. While it is clear that the HyMap image shows the highest utility for both types of targets, the utility of the HYDICE image, when assessed using easy targets, is very close to that of HyMap, whereas it is significantly less useful than HyMap for difficult targets.

It is informative to examine the image utility for targets of particular sizes. Table II presents the utility averaged over all targets and algorithms for targets up to four sizes.

We note that the HYDICE images have the highest utility for targets of less than 2 m while the HyMap image has the highest utility when assessed for targets of larger than 2 m. The small GRD of HYDICE images makes them useful for finding small targets. It is not clear what contributes to the high utility of the HyMap images for larger targets, but it is indicative of the fact that the image background composition is such that the four targets are very detectable using SMF, ACE, and SAM.

The superior performance of HyMap, with its larger GRD and coarser spectral resolution than HYDICE, is somewhat counterintuitive but reveals some interesting aspects of the target implant utility assessment method. Although the sensor characteristics suggest that HYDICE should provide higher utility, this ignores the important aspect of the scene content as a factor for consideration. Both of the HYDICE scenes had targets resident in them, whereas the HyMap scene did not. We have found that this utility metric reports lower utility in the case of targets present because of the false alarms created by the presence of targets in the "target absent" case. When the targets are removed from the HYDICE images using a target mask, the utility is higher than that of the HyMap image. We chose to leave the targets in the image, however, to represent the more realistic situation in that an image analyst would not have *a priori* knowledge of targets in the image.

IV. Conclusion

This paper presents a novel methodology to assess the utility of any spectral image for the subpixel target detection application. More generally, it offers a framework by which to define the image utility of spectral images and addresses

TABLE I
SUMMARY UTILITY AND TARGET DIFFICULTY

	HYDICE (0.85 m GRD)		AVIRIS (20 m GRD)		3 m GRD	
target difficulty	FR I	DR II	Moffett	Lunar Lake	HYMAP	MISI
difficult	0.49	0.44	0.44	0.47	0.78	0.49
easy	0.94	0.93	0.67	0.69	0.95	0.88

TABLE II
SUMMARY UTILITY AND TARGET SIZE

	HYDICE (0.85 m GRD)		AVIRIS (20 m GRD)		3 m GRD	
target size	FR I	DR II	Moffett	Lunar Lake	HyMap	MISI
1 m	0.29	0.30	0.00	0.00	0.17	0.00
2 m	0.53	0.52	0.00	0.01	0.39	0.08
5 m	0.65	0.63	0.09	0.11	0.68	0.44
20 m	0.71	0.69	0.56	0.58	0.87	0.68

the issue of how spectral image utility might be assessed for any application. The premise for the value of a spectral image utility metric is based on the success of NIIRS for panchromatic images. This paper addresses the fundamental difference by which information is extracted from panchromatic and spectral imagery by incorporating a means of adapting spectral processing algorithm performance measures into a measure of image utility. The resulting methodology and spectral image utility metric for the subpixel detection application are used to explore the effect of image acquisition and target detection scenario parameters on the utility of six spectral images.

This paper applies the target implant method to assess the utility for the realistic situation of dealing with images that may contain no targets, a small number of target pixels, and have no corresponding ground truth. The target implant method creates its own binary hypothesis test and considers the effect of a random target vector fractionally implanted in every image pixel as the target present case. This approach enables the investigation of the detectability of any target for which a reference library description exists at any subpixel fraction in any image. It also offers a self-contained automated method of assessing the utility of an image in a summary sense for a range of target types, size, detectors, and detection operating points. It affords the investigation of the role of the target relative to the image scene spectral content and associated atmospheric compensation and sensor parameters. Users may specify an input reflectance image and specific parameters of the target detection scenario, as well as the degree of granularity desired in the resulting image utility, thus affording wide latitude in tailoring the resulting utility metric to their specific information requirements.

Six images, four targets, 30 target sizes, and three detectors are used to sample the infinite image acquisition and target detection scenario space in assessing the utility of the images. Results confirm that the HYDICE images are the best for detecting targets of less than 2 m, and the HyMap image is best for detecting targets larger than 2 m. This is consistent with the GRD of the corresponding images. In general, the ACE detector seems to offer the most consistently high image utility across all images, and SAM offers the lowest, as expected.

Future work will seek to better understand the sensitivity of the assessed utility to image chain parameter variations and investigate the robustness of different image utility prediction approaches. While at this point the results are dependent on the specific images and targets under investigation, we hope that, by further exploration of a range of images and targets, we can generalize the results to classes of images and targets.

ACKNOWLEDGMENT

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the U.S. Air Force, Department of Defense, or the U.S. Government.

REFERENCES

[1] J. M. Irvine, "National imagery interpretability rating scales (NIIRS): Overview and methodology," in *Proc. SPIE—Airborne Reconnaissance XXI*, 1997, vol. 3128, pp. 93–103.

[2] J. C. Leachtenauer and R. G. Driggers, *Surveillance and Reconnaissance Imaging Systems: Modeling and Performance Prediction*. Boston, MA: Artech House, 1997, ch. 6.

[3] R. G. Driggers, P. Cox, and M. Kelley, "National imagery interpretation rating system and the probabilities of detection, recognition, and identification," *Opt. Eng.*, vol. 36, no. 7, pp. 1952–1959, Jul. 1997.

[4] D. Manolakis, D. Marden, and G. Shaw, "Hyperspectral image processing for automatic target detection applications," *MIT Linc. Lab. J.*, vol. 14, no. 1, pp. 79–116, 2003.

[5] Image Resolution Assessment and Reporting Standards (IRARS) Committee, *Multispectral Imagery Interpretability Rating Scale Reference Guide*, 1995, Washington, DC.

[6] J. P. Kerekes, "Receiver operating characteristic curve confidence intervals and regions," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 2, pp. 251–255, Apr. 2008.

[7] J. Theiler and B. Foy, "Effect of signal contamination in matched-filter detection of the signal on a cluttered background," *IEEE Geosci. Remote Sens. Lett.*, vol. 3, no. 1, pp. 98–102, Jan. 2006.

[8] C. E. Caefer, S. R. Rotman, J. Silverman, and P. W. Yip, "Algorithms for point target detection in hyperspectral imagery," in *Proc. SPIE—Imaging Spectrometry VIII*, 2002, vol. 4816, pp. 242–257.

[9] J. P. Kerekes and J. E. Baum, "Spectral imaging system analytical model for subpixel object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1088–1101, May 2002.

[10] D. Manolakis and G. Shaw, "Detection algorithms for hyperspectral imaging applications," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 29–43, Jan. 2002.

[11] P. Bajorski, "Analytical comparison of the matched filter and orthogonal subspace projection detectors for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 7, pp. 2394–2402, Jul. 2007.

[12] W. H. Farrand and J. C. Harsanyi, "Mapping the distribution of mine tailings in the Coeur d'Alene River Valley, Idaho, through the use of a constrained energy minimization technique," *Remote Sens. Environ.*, vol. 59, no. 1, pp. 64–76, Jan. 1997.

[13] D. Manolakis, "Taxonomy of detection algorithms for hyperspectral imaging applications," *Opt. Eng.*, vol. 44, no. 6, pp. 066 403-1–066 403-11, Jun. 2005.

[14] S. Kraut and L. Scharf, "The CFAR adaptive subspace detector is a scale-invariant GLRT," *IEEE Trans. Signal Process.*, vol. 47, no. 9, pp. 2538–2541, Sep. 1999.
[15] S. Kraut, L. Scharf, and L. McWhorter, "Adaptive subspace detectors," *IEEE Trans. Signal Process.*, vol. 49, no. 1, pp. 1–16, Jan. 2001.
[16] C. E. Metz, "ROC methodology in radiologic imaging," *Invest. Radiol.*, vol. 21, no. 9, pp. 720–733, Sep. 1986.
[17] M. S. Stefanou, "Spectral image utility for target detection applications," Ph.D. dissertation, Rochester Inst. Technol., Rochester, NY, 2008.

Marcus S. Stefanou (M'94) received the B.S. degree in electrical engineering from the U.S. Naval Academy, Annapolis, MD, in 1990 and the M.S. degree in electrical engineering from the U.S. Naval Postgraduate School, Monterey, CA, in 1997. He is currently working toward the Ph.D. degree in the Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, NY.

From 1990 to 2003, he was an Aircraft Maintenance Officer with the U.S. Marine Corps. From 2003 to 2005, he was a Developmental Engineer with the Sensors Directorate, Air Force Research Laboratory. He is currently a Lieutenant Colonel with the U.S. Air Force. His research interests are in hyperspectral target detection and spectral image quality.

John P. Kerekes (S'81–M'89–SM'00) received the B.S., M.S., and Ph.D. degrees in electrical engineering from Purdue University, West Lafayette, IN, in 1983, 1986, and 1989, respectively.

From 1983 to 1984, he was a Member of the Technical Staff with the Space and Communications Group, Hughes Aircraft Co., El Segundo, CA, where he performed circuit design for communications satellites. From 1986 to 1989, he was a Graduate Research Assistant with the School of Electrical Engineering and the Laboratory for Applications of Remote Sensing, Purdue University. From 1989 to 2004, he was a Technical Staff Member with the Lincoln Laboratory, Massachusetts Institute of Technology, Lexington, MA. Since 2004, he has been an Associate Professor with the Chester F. Carlson Center for Imaging Science, Rochester Institute of Technology, Rochester, NY. His research interests include the modeling and analysis of remote sensing system performance in pattern recognition and geophysical parameter retrieval applications.

Dr. Kerekes is a member of Tau Beta Phi, Eta Kappa Nu, the American Geophysical Union, the American Meteorological Society, the American Society for Photogrammetry and Remote Sensing, the Optical Society of America, and SPIE. From 1995 to 2004, he served as the Chair of the Boston Section Chapter of the IEEE Geoscience and Remote Sensing Society (GRSS). He is currently serving as the Founding Chair of the Western New York Chapter of GRSS. He was the General Cochair of the 2008 IEEE International Geoscience and Remote Sensing Symposium held in Boston, MA.