Hyperspectral Remote Sensing 
Subpixel Object Detection Performance

John P. Kerekes 
Digital Imaging and Remote Sensing Laboratory 
Chester F. Carlson Center for Imaging Science 
Rochester Institute of Technology 
54 Lomb Memorial Drive 
Rochester, New York, 14623 USA

Abstract—For nearly thirty years now, airborne and satellite hyperspectral imaging sensors have been used to collect high spatial resolution (1-30 meter) imagery of the earth’s surface in hundreds of co-registered, contiguous spectral channels. These data have been shown to enable the detection of objects smaller than a pixel due to the spectral information present. However, it is not always obvious beforehand if a given object will be detectable in a given scene, as performance has been observed to depend on many factors including illumination conditions, scene spectral complexity, target variability, sensor artifacts as well as algorithm variations. Over the past fifteen years our research has been exploring ways to predict and assess performance of hyperspectral subpixel detection. Our methods have included analytical modeling tools, empirical blind tests, and quality metrics for spectral imagery. Results of this work have confirmed the feasibility of hyperspectral subpixel objection detection and have provided tools for quantification of the performance.

Keywords-hyperspectral; target detection; sensing performance

I. INTRODUCTION

For nearly thirty years now, airborne and satellite hyperspectral imaging sensors have been used to collect high spatial resolution (1-30 meter) imagery of the earth’s surface in hundreds of co-registered, contiguous spectral channels. These data have been shown to enable the detection of objects smaller than a pixel due to the spectral information present [1]. However, it is not always obvious beforehand if a given object will be detectable in a given scene, as performance has been observed to depend on many factors [2]. These factors include illumination conditions, scene spectral complexity, target variability, sensor artifacts, and even algorithm variations.

During the past fifteen years, we have been exploring ways to predict and assess performance of hyperspectral subpixel detection. This work has been motivated by the goal of advancing the science of hyperspectral target detection through the development of quantitative modeling and assessment tools.

Four such tools for the study of hyperspectral performance are described in this paper. We begin by reviewing an analytical modeling tool, which predicts subpixel target detection performance for a given scenario comprised of a target, background, sensor, and processing algorithm. This is followed by tools for the empirical quantification and prediction of hyperspectral target detection utility, which operate on an existing hyperspectral image. The following section describes a web application that allows users to access data and score their target detection results in an unbiased manner. The last tool described in this paper is one that is not specific to hyperspectral, but rather allows the user to generate statistical confidence regions for target detection results described by receiver operating characteristic (ROC) curves. Finally, we summarize by discussing some lessons learned throughout this work.

II. HYPERSPECTRAL IMAGING

The term “hyperspectral imaging” was first coined by Alex Goetz while he was at the NASA Jet Propulsion Laboratory in the early 1980’s [3]. In this context, the term came into being to distinguish hyperspectral imagery with its hundreds of contiguous narrow spectral channels from that of multispectral imagery which typically has several widely spaced broad spectral channels. Hyperspectral imagers typically have a spectral resolution ($\lambda/\Delta\lambda$) on the order of 100. Also, this original application of the term hyperspectral imaging referred to land imaging systems with a spatial resolution of a few meters to tens of meters.

In the last decade, a number of researchers outside of the land remote sensing community have adopted the term hyperspectral to apply to optical imaging systems with even higher spectral resolution ($\lambda/\Delta\lambda \sim 1000$) [4] or even to passive microwave sensing [5], together with lower spatial resolution (several kilometers). It is unfortunate and confusing to the community to use the same term to apply to very different sensor technology and applications. In our work, we restrict the term hyperspectral imaging to apply to its original context of high resolution spectral imaging of the earth’s surface.

III. SYSTEM PERFORMANCE FORECASTING

Our initial work in this topic lead to the development of an analytical modeling approach to hyperspectral subpixel detection performance forecasting [6]. The resulting model, known as FASSP (Forecasting and Analysis of Spectroradiometric System Performance) uses first- and second-order statistics to describe surface classes (targets and backgrounds) and propagates these statistics through the effects...
of the remote sensing process using linear operations. This parametric approach leads to direct predictions of detection receiver operating characteristic curves, but without the need for large sample sizes (and correspondingly long computational times) usually necessary with Monte Carlo or image simulation approaches [7]. Initially developed for the reflective part of the optical spectrum, it was later extended to include the mid-wave and long-wave thermal infrared [8].

A recent study using this model [9] explored the limits of detectability for a subpixel target as a function of three important system parameters: sub-pixel fraction, signal-to-noise ratio, and background complexity as measured by the number of background classes. Table 1 lists the system parameters for the nominal case and Figure 1 presents the results showing the probability of detection (at a specified false alarm rate) as a function of the subpixel fraction occupied by the target.

In this result we observe that the target is detectable with high probability ($P_D > 0.8$) when occupying more than 20% of a pixel. As an example, this means a 5 sq. meter (or larger) target would be detectable by using a sensor with pixels having a ground area of 25 sq. meter, or approximately 5 m ground resolution. Of course, this result should be interpreted as being only valid for the scenario considered, but is illustrative of the type of analysis possible with the model.

### Table 1. Nominal System Parameter Values

<table>
<thead>
<tr>
<th>System Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Target</td>
<td>Green cotton fabric</td>
</tr>
<tr>
<td>Target fill fraction</td>
<td>varies</td>
</tr>
<tr>
<td>Background classes</td>
<td>10</td>
</tr>
<tr>
<td>Atmospheric model</td>
<td>Summer mid-latitude</td>
</tr>
<tr>
<td>Aerosol model</td>
<td>Rural haze</td>
</tr>
<tr>
<td>Meteorological range</td>
<td>10 km</td>
</tr>
<tr>
<td>Solar zenith angle</td>
<td>30º</td>
</tr>
<tr>
<td>Sensor</td>
<td>HyMap [8]</td>
</tr>
<tr>
<td>Relative calibration error</td>
<td>1%</td>
</tr>
<tr>
<td>Number of spectral bands</td>
<td>126</td>
</tr>
<tr>
<td>Detection algorithm</td>
<td>Spectral matched filter</td>
</tr>
</tbody>
</table>

IV. Spectral Image Utility Assessment

A topic of ongoing interest in the hyperspectral image analysis community is how to quantitatively assess the quality of a given image. We have explored this question from a variety of perspectives, with recent work focusing on the utility of an image, which represents one component of quality. In particular, we have focused on subpixel target detection [10].

Our proposed approach involves the use of a target library and fractionally implanting a realization of the target in every pixel of the image [11]. By using a target library with many samples, and generating different realizations according the sample statistics we achieve a realistic distribution of target samples. By fractionally implanting these samples in every pixel of a given image, we obtain a large sample size for good ROC estimation (see section VI) and we assess the utility of an image in a holistic way rather than in just a local area. The quantitative estimate of utility is then the integrated area under the ROC curve over a specified region of probability of false alarm.

As an example, consider the hyperspectral image for which the RGB is shown in Fig. 2. This image has 145 spectral bands (after bad bands removal) and ground resolution of approximately 0.85 m. Figure 3 shows the resulting utility for this image computed for four target types as a function of the linear dimension of a square realization of the target. In this example we see the image has high utility for finding tan fabrics that are greater than 0.6 m in linear dimension. The lowest utility is found for the green vehicle; not surprising given the green vegetation that dominates the image.

Figure 2. Natural color rendition of airborne hyperspectral image.

Figure 3. Spectral utility of hyperspectral image shown in Fig. 2 for four target types as a function of target size.
We have also combined this approach of empirical hyperspectral utility assessment with our system performance modeling tools to yield a spectral utility prediction tool [12]. This method uses class-specific spectral statistics derived from a given hyperspectral image to drive the analytical model described in section III.

V. UNBIASED TESTING DATA

Hyperspectral target detection algorithm development continues to be a focus of many researchers, with new approaches being published each year. In order for the authors to quantitatively compare the results of their new algorithm to others, by necessity they must know the ground truth for their image data. However, this also raises the possibility that the authors may knowingly (or unknowingly) “tune” their algorithm to achieve optimum performance on their data.

This situation suggests a need for unbiased community data and evaluation tools. We have prepared such a data set and website to support the R&D community in this way [13, 14].

Fig. 4 shows one of the hyperspectral images provided through the website. For this “self test” image we provide spectra and locations for spatially unresolved targets that were embedded in the scene at the time it was imaged by the airborne hyperspectral imager HyMap [15]. The web site also provides a second image for which we only supply the target spectra, not the locations. Users then apply their own target detection algorithm to this “blind test” image and upload their results to the web site for automatic scoring. The site then automatically calculates a score that is the number of pixels containing a detection metric equal to or greater than the value at the correct location of the subpixel target. Thus a perfect score would be 1 indicating only the correct pixel had a value sufficient for detection.

One of the interesting aspects of this web site is the publication of a “Top Ten” list showing the user-labeled algorithms that achieved the best performance amongst all uploaded results. Table II shows a recent Top 10 for one of the targets. As can be seen in the table, out of the 174 separate uploads for scoring on this target, only two have achieved a perfect score, with several others coming close with only a few false alarms. To date, over 350 users have registered on the site and several publications have used the data. An overview paper lead off a recent special session at a conference where there was much interest in these data and calls for additional data sets [16]. This continues to be a topic of interest in our research group and we are looking for additional opportunities to provide similar data.

VI. ROC CONFIDENCE REGION TOOL

In presenting the results of their new target detection algorithms, it is common for authors to provide ROC curves empirically estimated from data with known ground truth. This is very useful as a ROC curve is an excellent metric by which to judge the performance of a detection algorithm.

However, it is often the case the estimates for detection probability are made with a relatively small number of samples (either targets or target pixels). It is well known that estimates of probability based on small sample sizes are unreliable, and indeed the statistical literature has developed methods to assign a confidence interval to these estimates. While a confidence interval is appropriate to provide a range for a scalar value, ROC curves have point estimates that lie in a two-dimensional space. Thus, it becomes appropriate to consider a confidence region rather than interval. Our work has investigated this topic and developed an approach and online tool for the calculation of confidence regions around each point estimate on an empirical ROC curve [17, 18].

The basic idea of our approach is to assume detection events (true detections or false alarms) are statistically independent events and thus we can model the distribution of the estimates of \( P_D \) or \( P_{FA} \) as binomial. One rational basis for the assumption of independence is every pixel measured by an imaging sensor has noise associated with it, and the noise can be reasonably assumed to be independent from measurement to measurement. Once we have a statistical model for the \( P_D \) and \( P_{FA} \) estimates, we can define a boundary around the joint density function to enclose a specified probability.

Fig. 5 provides examples of the type of plots produced by the online tool [18]. The example considered here is an image containing 100,000 pixels and two different target situations. In the top plot, we show results for a case where there are 200 target pixels, while the bottom plot is for the case with only 20 target pixels. The detection and false alarm rates here are calculated on a per-pixel basis. The curves drawn around each point estimate on the plots show the 95% confidence regions around these estimates. That is, we can say with 95% confidence that the true values of the point estimates on the ROC curve lie inside those regions centered at each point.

<table>
<thead>
<tr>
<th>Rank</th>
<th>Algorithm Name</th>
<th>Algorithm Version</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>WTA</td>
<td>1.1</td>
<td>6 of 174</td>
</tr>
<tr>
<td>2</td>
<td>Improved PLMF</td>
<td>1.1</td>
<td>6 of 174</td>
</tr>
<tr>
<td>2</td>
<td>BGU Y.C - PLMF</td>
<td>1.0</td>
<td>5 of 174</td>
</tr>
<tr>
<td>2</td>
<td>BGU Y.C - F5 Full Local ACE With PF 1x1</td>
<td>1.0</td>
<td>5 of 174</td>
</tr>
<tr>
<td>5</td>
<td>BGU Y.C - F5 Full Local ACE With PF 2x2</td>
<td>1</td>
<td>6 of 174</td>
</tr>
<tr>
<td>5</td>
<td>Isaac Gerg - ARU/PSU</td>
<td>mace</td>
<td>6 of 174</td>
</tr>
<tr>
<td>5</td>
<td>Isaac Gerg - ARU/PSU</td>
<td>mace</td>
<td>6 of 174</td>
</tr>
<tr>
<td>6</td>
<td>Local ACE</td>
<td>.1 (top score mapped to single superpixel)</td>
<td>10 of 174</td>
</tr>
</tbody>
</table>

Figure 4. Airborne hyperspectral image available through blindtest website.
Lincoln Laboratory, including Jerrold Baum, Michael Griffin, Dimitris Manolakis, and Seth Orloff, as well as former students at RIT including David Snyder and Marcus Stefanou.

REFERENCES


VII. SUMMARY AND LESSONS LEARNED

In this paper we have provided introductions to several research activities focused on the development of the science of hyperspectral remote sensing, particularly in the context of subpixel object detection. Through system modeling, empirical analyses, and quantitative assessment tools this research is contributing to the understanding and quantitative evaluation of hyperspectral imaging technology.

While much of the value of this work lies in the evaluation of performance for specific observational scenarios, there are a few general lessons learned we can offer. The first is that unresolved object detection with hyperspectral imaging data is clearly feasible. However, this ability is highly dependent upon the specifics of the situation and it is not always possible. Also, one must take care in presenting target detection results based on small sample sizes and it is recommended authors provide confidence regions on empirical ROC curves.

ACKNOWLEDGMENT

The work described in this paper has benefited greatly from contributions and collaborations with former colleagues at MIT.

As can be seen in Fig. 5, the confidence regions are much smaller when using a larger number of samples to estimate the points on the ROC curve. It is also instructive to note that with only 20 pixels, the confidence region on \( P_D \) can be quite large ranging as much as ±0.2.

Figure 5. Point estimates for a ROC curve (connected by lines) with surrounding 95% confidence regions for 100,000 background pixels and a) 200 target pixels and b) 20 target pixels.