Anomaly Detection Using Simulated MTI Data Cubes Derived from HYDICE Data

Mary M. Moya, John G. Taylor, Brian R. Stallard, and Sheila E. Motomatsu

Sandia National Laboratories*
P.O. Box 5800
Albuquerque, NM 87185-0845

Abstract

In this work, we quantify the separability between specific materials and the natural background by applying Receiver Operating Curve (ROC) analysis to the residual errors from a linear unmixing. We apply the ROC analysis to quantify performance of the Multi-spectral Thermal Imager (MTI). We describe the MTI imager and simulate its data by filtering HYDICE hyperspectral imagery both spatially and spectrally and by introducing atmospheric effects corresponding to the MTI satellite altitude. We compare and contrast the individual effects on performance of spectral resolution, spatial resolution, atmospheric corrections, and varying atmospheric conditions.

1 INTRODUCTION

The U.S. Department of Energy is funding the development of the Multispectral Thermal Imager (MTI), a satellite-based multi-spectral (MS) thermal imaging sensor scheduled for launch in Oct. 1999. MTI is a research and development (R&D) platform to test the applicability of multispectral and thermal imaging technology for detecting and monitoring signs of proliferation of weapons of mass destruction. During its three-year mission, MTI will periodically record images of participating government, industrial and natural sites in fifteen visible and infrared spectral bands to provide a variety of image data associated with weapons production activities. The MTI satellite will have spatial resolution in the visible bands that is five times better than LANDSAT TM in each dimension and will have five thermal bands. It also has the capability of collecting two MS images of the same ground location from both nadir and 45 degrees off-nadir to facilitate measurement and correction of atmospheric effects. To test the utility of the MTI data for detecting sub-pixel materials, we have modeled MTI data and implemented unconstrained linear unmixing to quantify the separability between naturally occurring background materials and materials of interest to nuclear production activity, such as camouflage and military vehicle metals and paints. The work complements that reported by Shen in [8] by providing performance over a wide range of detection thresholds for comparing the entire ROC. In a separate activity, Los Alamos National Laboratories is developing algorithms for measuring the temperature of bodies of water using MTI thermal bands.

2 DATA MODELING

We model the spectral and spatial characteristics of the MTI visible, near infra-red (NIR), and short-wave infra-red (SWIR) bands using HYDICE hyperspectral data from the Western Rainbow collection[1]. Table 1 summarizes characteristics of the MTI sensor, which

---

* Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

[1] We gratefully acknowledge CMO/SITAC for providing HYDICE data to us.

has 15 spectral bands, an altitude of 565 km, and swath width of 12 km. HYDICE data used in our tests were collected at an altitude of 4990ft to yield ground sample distance (GSD) of 0.77m. To simulate MTI point spread function, we apply an 8x8 spatial low-pass box filter to the HYDICE data. To simulate the MTI GSD, we subsample each low-passed HYDICE image by a factor of 6.5. To simulate the MTI spectral characteristics, we apply the MTI spectral filter functions to the HYDICE spectra. We simulate the additional atmospheric effects at the MTI satellite altitude with an ELM-like scaling of the radiance and path radiance using a MODTRAN atmospheric model [1]. In our unmixing tests, we used MTI bands A, B, C, D, E, G, I, and O. We did not use the five thermal bands because we did not have accurate estimates of the object temperatures to predict material emissivity characteristics. We also omitted bands F and H since they cover water absorption bands.

MTI Band Name	Spectral Range (nm)	Band Width (nm)	Ground Sample Distance (m)	Point Spread Width (m)
A	450 - 520	70	5	5.5
B	520 - 600	80	5	5.7
C	620 - 680	60	5	6
D	760 - 860	100	5	6
E	860 - 900	40	20	<20
F	910 - 970	60	20	<20
G	990 - 1040	50	20	<20
H	1360 - 1390	30	20	<20
I	1550 - 1750	200	20	<20
O	2080 - 2350	270	20	<20
J	3500 - 4100	600	20	<20
K-N	4870 - 10700	200 - 600	20	<20

Table 1. MTI Sensor and Data Characteristics

Figure 1 shows Band 30 of a HYDICE image of the Malpais site, extracted from Run08, frames 33-36, which was used in these tests. To compare the effects of reduced spatial resolution versus reduced spectral resolution, we applied the unmixing algorithm to four different data sets: 1) the original HYDICE data, 2) HYDICE data filtered to MTI bands, 3) HYDICE data filtered and resampled to match the MTI spectral resolution, and 4) data derived from HYDICE that simulated both MTI spectral and spatial characteristics.

With these four data sets, we tested the spectral separability using both raw radiance spectra and apparent reflectance spectra, which were corrected for effects of the atmosphere and solar source using an Empirical Line Method (ELM) with twelve calibration panels in the Western Rainbow image.

We also simulated the effects of a mid-latitude winter atmosphere with 50km visibility at MTI altitude and another atmosphere with 23km visibility and twice the water vapor at MTI altitude. To introduce these atmospheric effects into the image data, we estimated the effects of the two atmospheres at 565km altitude using the MODTRAN model. We scaled the HYDICE radiance and path radiance, which was estimated with an ELM-like algorithm, accordingly.

To aid in determining target detection and false alarm rates, we created masks for the image pixels. Four target masks indicate the size of the target objects: 1) Large targets are larger than 5m in both dimensions, 2) Long targets are larger than 5m in only one dimension, 3)

DISCLAIMER

Medium targets are larger than 2m and 5m or smaller in both dimensions, 4) Small targets are 2m or smaller in either dimension. Around each target mask, we include a region of border pixels that are treated as "don't care" pixels. We also include objects suspected of being man-made, but having no associated ground truth, as "don't care". For these "don't care" pixels, we do not have accurate information about target identity, size, or position. Thus, we do not want to unfairly penalize the detection algorithm by forcing it to detect objects that contain little or no target material or to reject unknown man-made objects that may differ significantly without our knowledge. We ignore these "don't care" pixels in determining the detection and false alarm rate statistics. In determining the detection rate, we use only those pixels that we know to contain target material. Undetected pixels within the target masks count as missed detections. All those pixels not belonging to any target mask or to the "don't care" region comprise the background mask, which identifies those pixels that should be recognized as part of the natural background. Detected pixels within the background mask count as false alarms.

3 METHOD

We assume that each MTI pixel contains multiple material spectra. If the materials are flat on the ground and produce Lambertian reflections, we can assume that the individual material spectra will mix together linearly. We form a model of the naturally occurring background in a subset of an image by identifying spectral endmembers associated with one or more soil types, one or more vegetation types, and shade. Then, when we apply spectral mixture analysis (SMA) [2,5,6] to the remaining image using a linear combination of these naturally occurring endmembers, any material not well modeled by the linear combination, appears as an error residual. This residual indicates the presence of materials with spectra distinct from the background. Thresholding the residual image yields detection and false alarm rates for SMA anomaly detection. For each test image, we produce Receiver Operating Curves (ROC) indicating the SMA target detection and false alarm performance over a range of thresholds. For most of the data processing, we apply routines available in ENVI, developed by Research Systems, Inc. [3] The following paragraphs describe the process for identifying appropriate endmembers.

We applied a Minimum Noise Fraction {MNF} transformation [3,4] to the HYDICE data to reduce the spectral dimensionality for endmember identification and to whiten the noise. We identified a small spectrally self-similar image region and assumed that the pixel to pixel apparent reflectances were correlated with one another while the pixel-to-pixel noise was uncorrelated. Estimating of the noise covariance allowed us to formulate a transformation to whiten the noise. A subsequent principal component transformation produced a series of components with decreasing signal-to-noise ratio. We retained the first eighteen MNF components for endmember identification.

We applied the Pixel Purity Index algorithm [3] to the reduced MNF data in a region of the image containing only vegetation, soil, and shade. It performs repeated projections of the 18-D data onto random 1-D vectors. It records the frequency with which individual pixel spectra project to 1-D extremes. We assume that those pixels with relatively large numbers of extreme projections appear on the convex hull of the high-dimensional data. We visually examine the convex hull pixel spectra in an N-Dimensional Visualizer [3] to identify likely endmembers, which are the purest pixels in the HYDICE image. From these endmembers, we empirically select a subset of four to six endmembers that form a comprehensive linear model of the vegetation/soil/shade region of interest. Then, we apply SMA to the entire image to detect anomalies.

For these tests, we extracted two sets of endmembers. We extracted the first endmember set, EM1, from the upper left-hand corner of the HYDICE image. Figure 2 shows the

EM1 endmember spectra, which represent two vegetation pixels, one soil pixel, and a shade pixel. We used these endmember spectra to perform SMA on both the 0.77m and 5.0m HYDICE images to test the effect of lower spatial resolution on separating the materials of interest from the background. To test the effects of the lower MTI spectral resolution, we first filtered the endmembers to MTI bands, producing EM1-M, and then applied SMA to the image with MTI spectra at 0.77m GSD. To test the combined effects of lower spatial and lower spectral resolution, we applied SMA to the simulated MTI image at 5.0m GSD using EM1-M. We also attempted to extract new endmembers from the MTI data, but doing so showed no performance improvements in the SMA performance.

Since ELM correction produces significant changes in spectral shapes, we found that unmixing any ELM-corrected data set required a new endmember set. We extracted endmember set, EM2, shown in Figure 3, from the ELM-corrected HYDICE image. Apparently, the corrections introduced more spectral distinctiveness into the data because this set of endmembers required two distinct soil spectra, as well as two vegetation spectra, and a shade spectra. A single soil spectra was incapable of representing soil near the target array. We also filtered EM2 to the MTI bands to produce EM2-M for application to the images with MTI spectra.

After each unmixing analysis with the appropriate endmember set, we plotted histograms of the SMA residual errors using both the target masks and the background masks. The range of overlap between the target and background histograms determined the range of thresholds for calculating the ROC's. To generate a ROC for each residual error image, we applied a range of thresholds to the error image and counted the number of detections within the target masks and the number of false alarms within the background mask. We generated target masks for those images with lower spatial resolution in the following way. We applied the 8x8 box filter to the high resolution mask, thresholded the smoothed mask at 0.2, and subsampled the result by a factor of 6.5. This method indicates which of the low resolution pixels could be expected to have 20% or greater content of target material..

4 RESULTS

Figures 4 through 11 show the ROCs for the image data with different spectral and spatial resolutions, both uncorrected and ELM-corrected. Figures 4 through 7 show the results for uncorrected image data. Figures 8 through 11 show the results for corrected image data. Figures 12 through 15 show the effects of the atmosphere with 50km visibility at MTI altitude. Figures 16 through 19 show the effects of the atmosphere with 23km visibility and twice the water vapor at MTI altitude. Figures 4, 5, 8, 9, 12, 13, 16, and 17 show the results at 0.77m GSD. Figures 6, 7, 10, 11, 14, 15, 18, and 19 show the results at 5.0m GSD. Figures 4, 6, 8, 10, 12, 14, 16, and 18 show the results with HYDICE spectra. Figures 5, 7, 9, 11, 13, 15, 17, and 19 show the results with MTI spectra. The dashed lines correspond to the Small-size targets, the dash-dot lines correspond to Medium-size targets, the dotted lines correspond to Large-size targets, and the solid lines correspond to Long targets, as defined in Section 2. The small and long targets do not appear on the results for 5.0m GSD, because none of these targets filled more than 20% of a pixel. In the high spatial resolution images, the long targets have the best detection rate. This result probably occurs because of our conservative selection of pixels for the long targets, i.e. the border regions for the long targets may have been larger than necessary.

5 DISCUSSION

Figure 4 shows that the target materials are at least 80% separable from the background with 5% false alarm rate (FAR) using HYDICE spectra at 0.77m GSD. Figure 5 shows that using MTI spectra reduces the target separability to a minimum of 70% at 5% FAR.

Using lower spatial resolution has a mixed effect on the separability of the target from the background. We lose the ability to detect the small and long, skinny targets, but improve the ability to detect the large and medium-sized target, as shown in Figure 6. At lower spatial resolutions, more material spectra linearly mix to form a single pixel. While pixels from large and medium size targets tend to mix with other target pixels (since we ignore the pixels with target content less than 20%), small patches of anomalous vegetation tend to mix with other vegetation making them less anomalous, which maintains a constant detection rate, but improves the false alarm rate. Also, the lower spatial resolution provides improved signal-to-noise ratio. Figure 7 demonstrates the combined effect of lower spectral and spatial resolution for the simulated MTI data, where the anomaly detector loses the ability to detect the small and long targets, but yields a minimum detection rate of 67% at 5% FAR for the large and medium targets.

Comparing Figures 8, 9, and 10 to Figures 4, 5, and 6, respectively, shows that applying ELM corrections to the images can slightly improve the anomaly detection performance within a single image. Of course, the corrections are imperative for the future work of applying scene-to-scene anomaly detection. Comparing Figure 11 to Figure 7 shows that the ELM correction actually degrades the anomaly detection performance at the higher threshold values, which correspond to lower false alarm rates. This probably results from the larger number of unmixing endmembers used when unmixing the MTI image with 5.0m GSD. Recall that the corrected results required five rather than four endmembers because of the increase in spectral distinctiveness introduced by the correction. However, we lose spectral distinctiveness when we use the MTI spectra and, also, when we reduce the spatial resolution. Thus, we might improve the results for 5.0m MTI data by using a smaller number of endmembers. We will test this hypothesis in future work.

Figures 12 through 16 show that additional atmospheric effects introduce only small degradations in detection performance. We suspect that the two atmospheres impact the performance so little because the major effects introduced by the high altitude atmospheres are increased water vapor absorption and additive path radiance in the visible bands only. In unmixing all HYDICE spectra, we ignore the water vapor bands. The additional water vapor absorption largely affects MTI spectra bands F and H, which we do not use. Since the additive path radiance effects the visible bands only, there is probably sufficient distinguishing spectral shape in the remaining bands to allow the appropriate unmixing.

The ability to separate medium-size targets from the background actually improves for all the data with additional atmosphere at MTI altitude. The improved performance results because without the additional atmosphere the medium target pixels must differ significantly from the background, but not enough to cause a detection; their residual errors must lie just below the detection threshold. The addition of the extra atmospheric effects must introduce enough changes in the residual errors to boost them above the detection threshold. The changes do not change the spectra of background materials sufficiently to introduce additional false alarms. Thus, despite the additional atmospheric effects, the spectra maintain their distinctiveness sufficiently to allow separation between the background materials and the materials of interest. The additional atmosphere will become more of a problem in scene-to-scene anomaly detection because of the need completely remove the atmospheric effects.

These results give indications about the theoretical separability of the target materials from the natural background, but they give no conclusions about the applicability to a deployable system because of the following problems. 1) In this theoretical study, we have identified spectrally pure endmembers in the high spatial resolution image and used these to search for anomalous materials in the lower resolution images, but, given only the lower resolution data, how could we get spectrally pure endmembers relevant to the scene of

interest? 2) Can we correct images with sufficient accuracy to apply scene-to-scene unmixing? 3) The utility of these anomaly detection techniques for real applications using hyperspectral and multi-spectral data will depend on the false alarm requirements of the application. Many automatic target recognition (ATR) applications have false alarm rate requirements far below the 5% rate of these detectors. [7]

6 FUTURE WORK

Future work will include the following efforts. Since the current MTI simulation uses 5m spatial resolution in all the bands, we will produce a more accurate simulation that uses 5m resolution in Bands A through D and 20m resolution in Bands E, G, I, and O. We will attempt to improve the performance of anomaly detection on the ELM-corrected image with MTI spectra at 5.0 m GSD by extracting a new endmember set. We will also extract endmembers from one corrected apparent reflectance image, and then apply SMA to a second corrected image to detect anomalies. We will explore the applicability of scene-to-scene anomaly detection for change detection applications, and we will investigate the reliability of classifying the detected anomaly as specific materials of potential interest to nuclear non-proliferation. We plan to develop methods of correcting the MTI data to apparent reflectance that does not depend on having calibration panels and ground truth. The proposed atmospheric correction methods will capitalize on MTI's ability to collect independent measurements of the same ground location from two different look angles, nadir and 45 degrees off nadir.

[1] Anderson, G.P., Kneizyf, Z.Y., Chetwynd, J.H., Wang, J., Hoke, M.L., Rothman, L.S., Kimball , L.M., McClatchey, R.A., FASTCODE / MODTRAN / LOWTRAN: Past / Present / Future, Eighteenth Annual Review Conference on Atmospheric Transmission Models, Jun., 1995.

[2] Boardman, J.W., and Kruse, F.A., Automated spectral analysis: A geological example using AVIRIS data, North Grapevine mountains, Nevada, Proceedings of the Tenth Thematic Conference on Geologic Remote Sensing, May, 1994.

[3] ENVI User's Guide, Research Systems, Inc., July 1995.

[4] Green, A.A., Berman, M., Switzer, P., and Craig, M.D., A transformation for ordering multispectral data in terms of image quality with implications for noise removal, *IEEE Transactions on Geoscience and Remote Sensing,* Vol. 26, No. 1, January, 1988.

[5] HYMSMO Interim Nonliteral Exploitation Report, Spectral Information Technology Applications Center, April 30, 1996.

[6] HYMSMO Interim Nonliteral Exploitation Report, Spectral Information Technology Applications Center, January, 1995.

[7] Koch, M.W., Moya, M.M., Hostetler, L.D., and Fogler, R.J. , Cueing, feature discovery and one-class learning for synthetic aperture radar ATR, *Neural Networks,* Vol. 8, No. 7/8, pp. 1081—1102, 1995.

[8] Shen, S. S., Multiband Sensor System Design Tradeoffs and Their Effects on Remote Sensing and Exploitation, Proceedings of*SPIE* , Vol. 3118, pp. 296-307, 1997.

Figure 1. HYDICE Image Data from Western Rainbow Collection, Extracted from Run08, Frames 33-36, Band 30

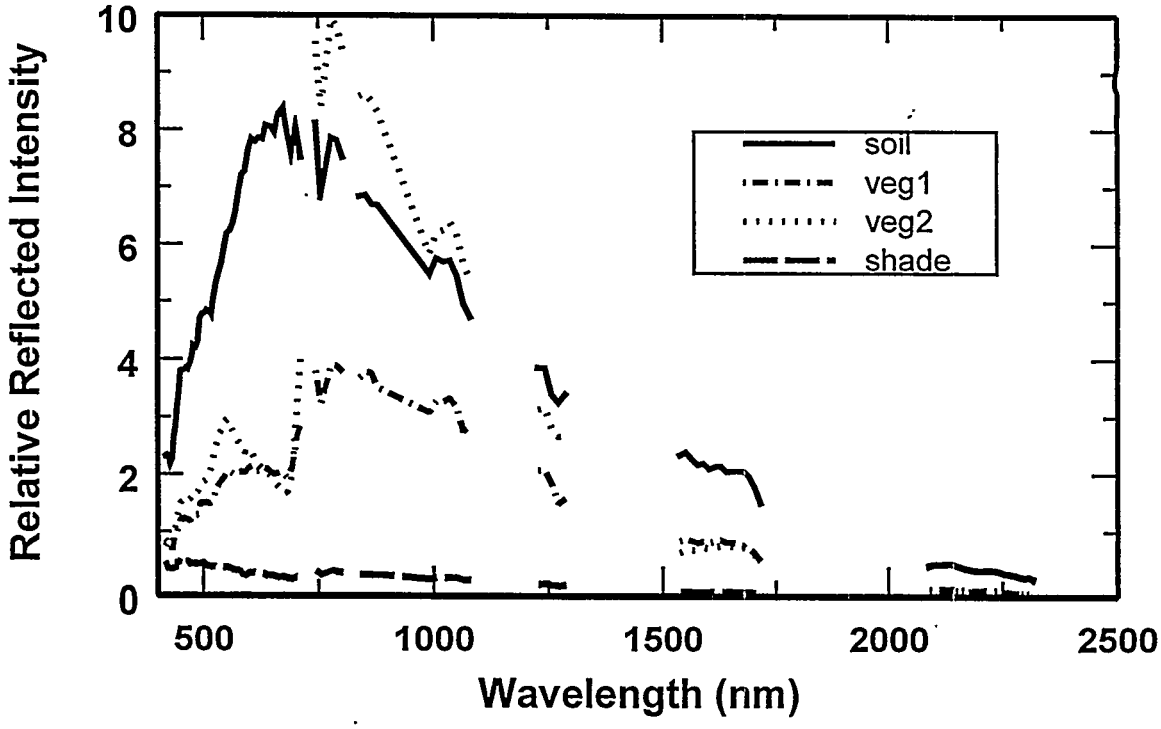

Figure 2. Endmember spectra for unmixing uncorrected images displayed with HYDICE resolution

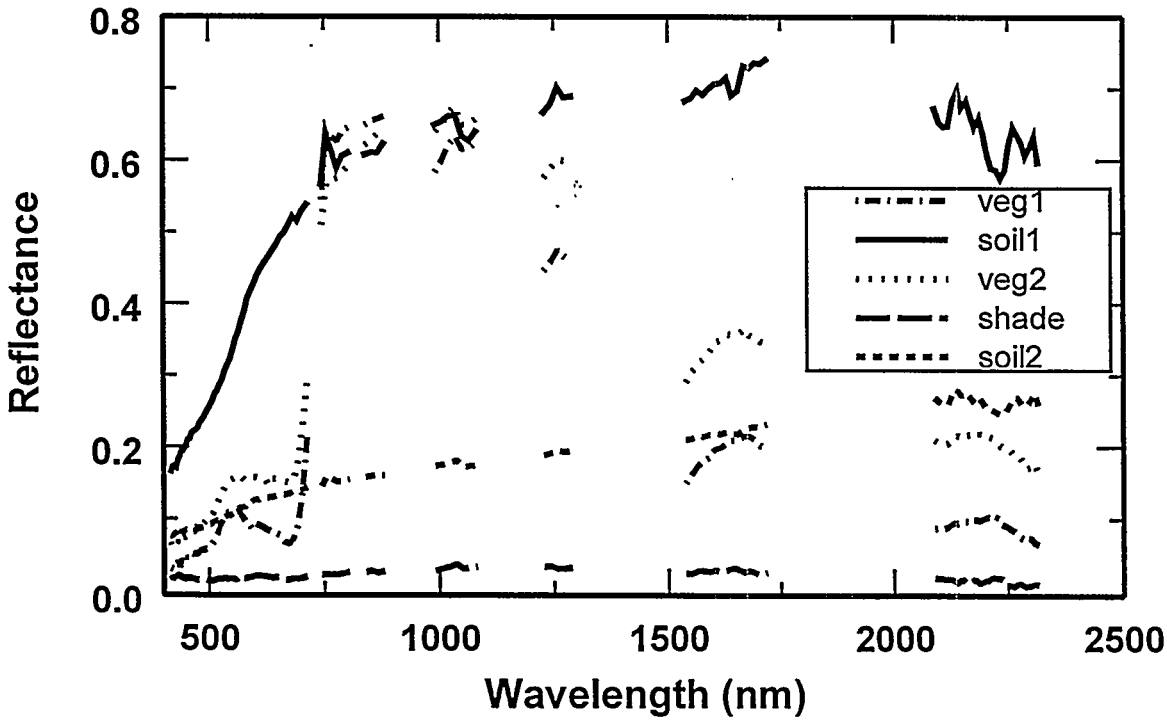

Figure 3. Endmember spectra for unmixing ELM-corrected images displayed with HYDICE resolution

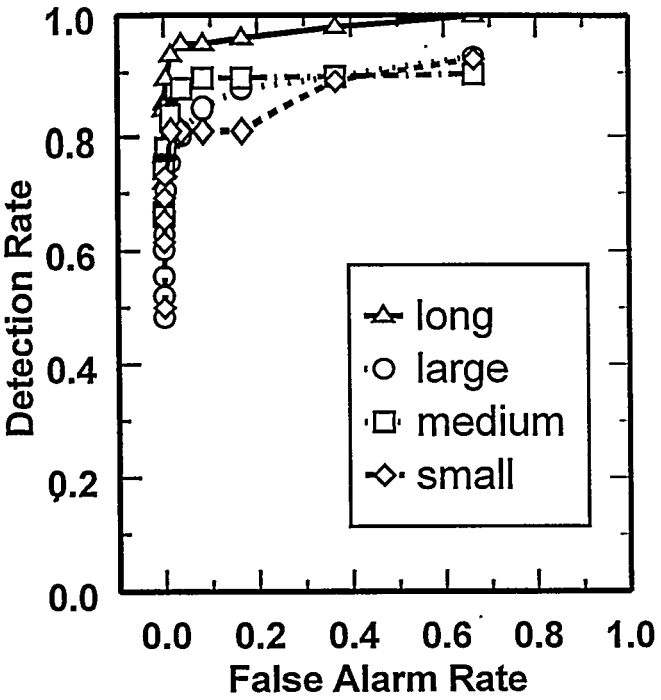

Figure 4. ROC for anomaly detection for image with HYDICE spectra, 0.77m GSD.

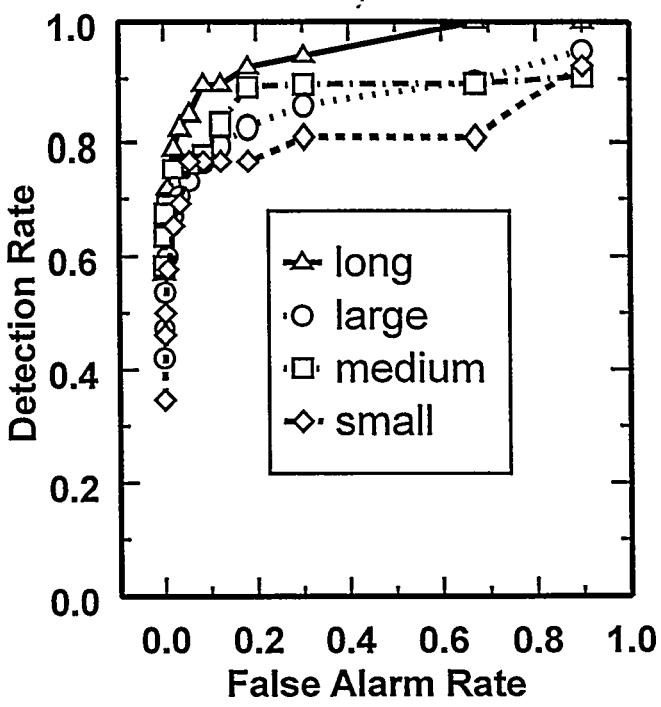

Figure 5. ROC for anomaly detection for image with MTI spectra, 0.77m GSD.

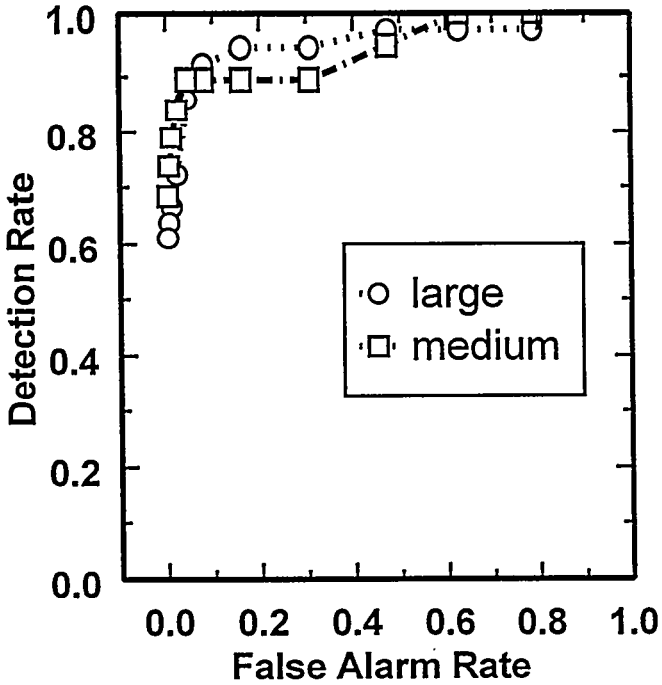

Figure 6. ROC for anomaly detection for image with HYDICE spectra, 5.0m GSD.

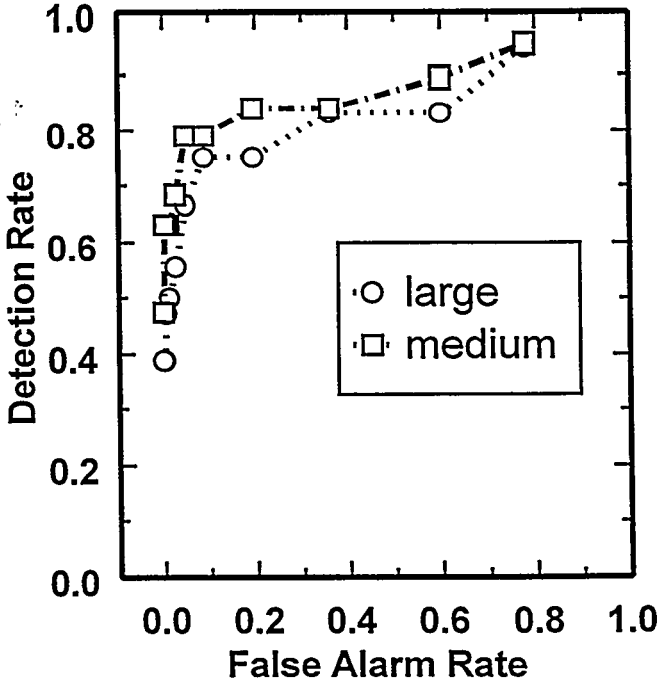

Figure 7. ROC for anomaly detection for image with MTI spectra, 5.0m GSD.

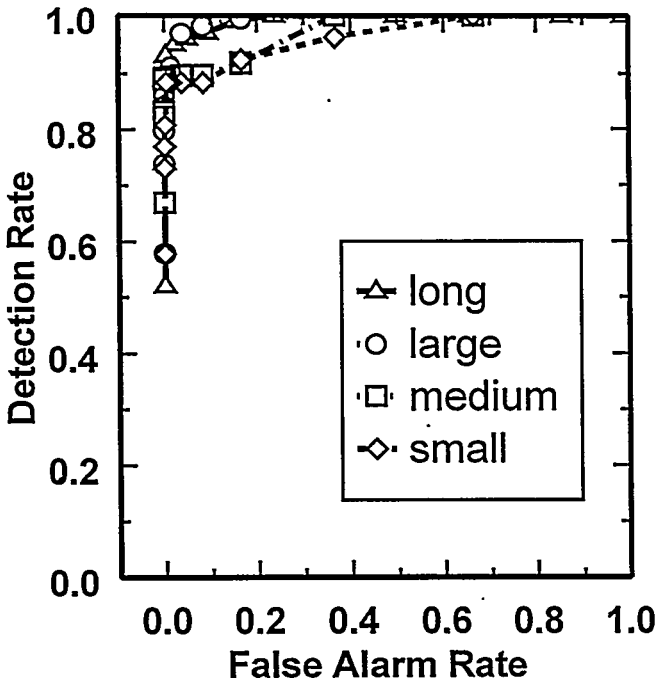

Figure 8. ROC for anomaly detection for ELM-corrected image with HYDICE spectra, 0.77m GSD.

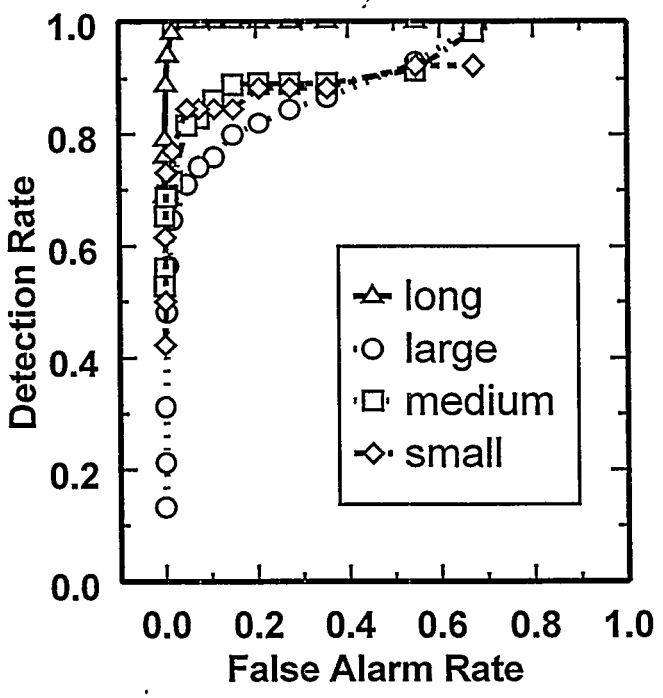

Figure 9. ROC for anomaly detection for ELM-corrected image with MTI spectra, 0.77m GSD.

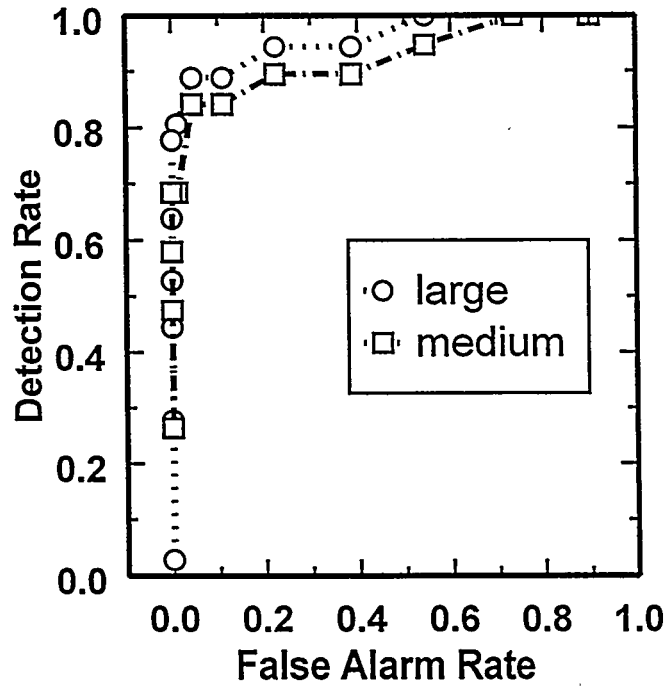

Figure 10. ROC for anomaly detection for ELM-corrected image with HYDICE spectra, 5.0m GSD.

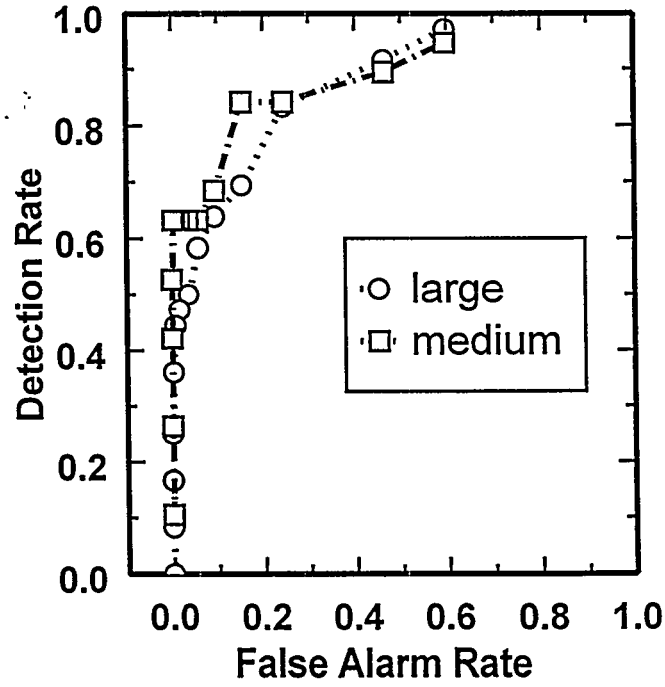

Figure 11. ROC for anomaly detection for ELM-corrected image with MTI spectra, 5.0m GSD.

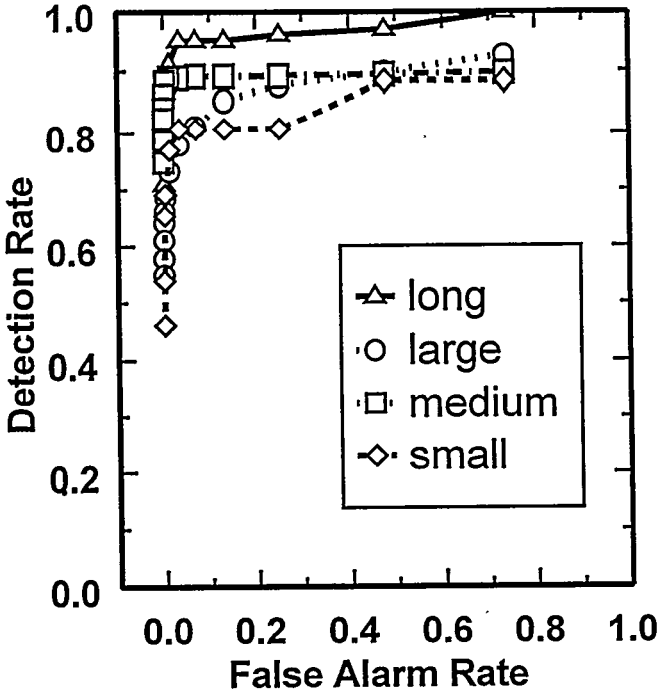

Figure 12. ROC for anomally detection in HYDICE image, 0.77m GSD, MTI Altitude.

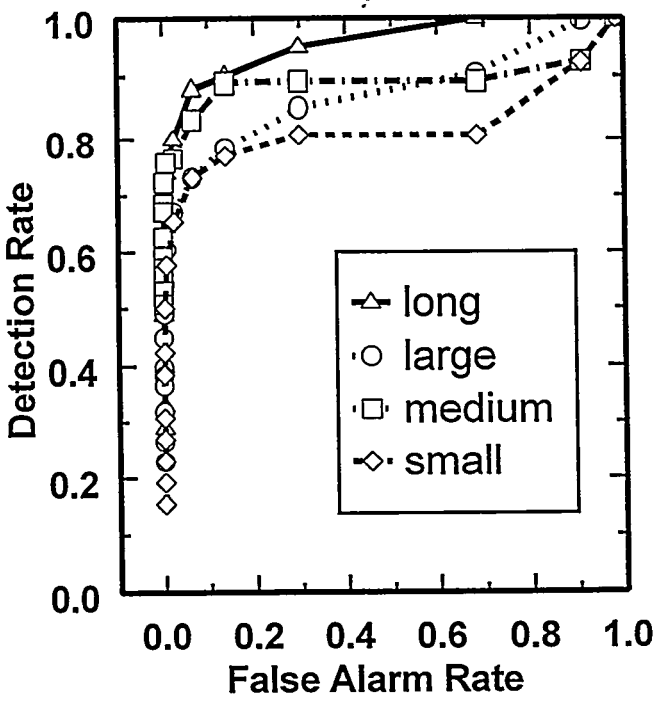

Figure 13. ROC for anomaly detection in MTI image, 0.77m GSD, MTI altitude.

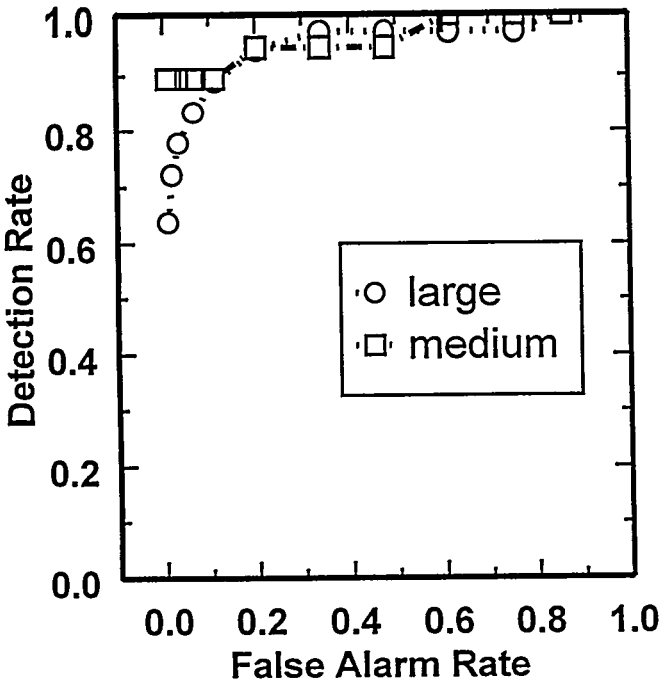

Figure 14. ROC for anomaly detection in HYDICE image, 5.0m GSD, MTI altitude.

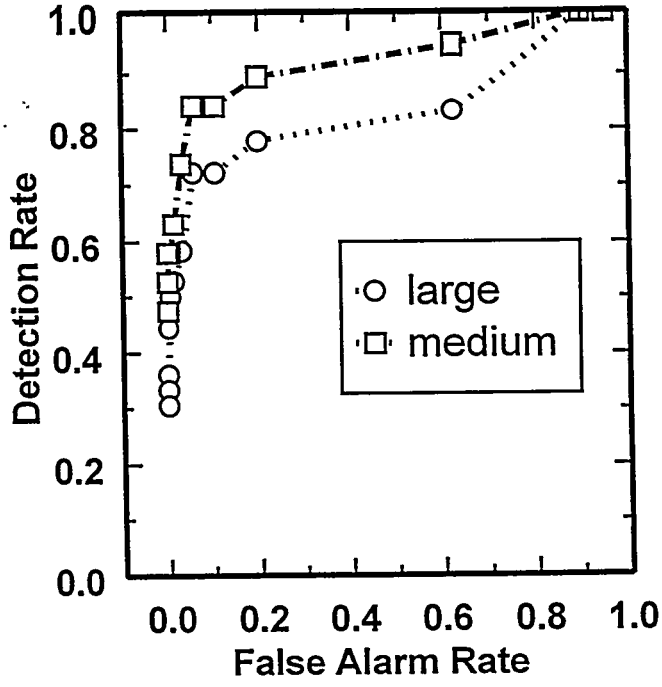

Figure 15. ROC for anomaly detection in MTI image, 5.0m GSD, MTI altitude.

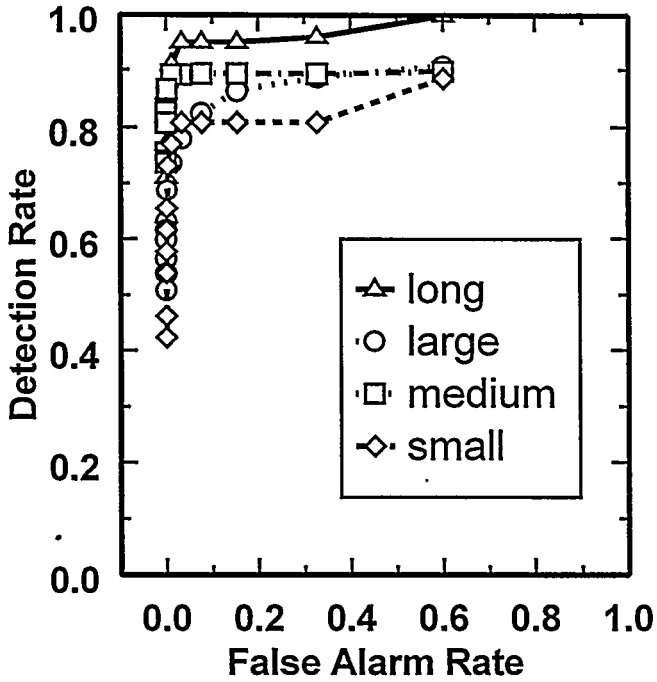

Figure 16. ROC for anomaly detection in HYDICE image, 0.77m GSD, MTI altitude with degraded atmosphere.

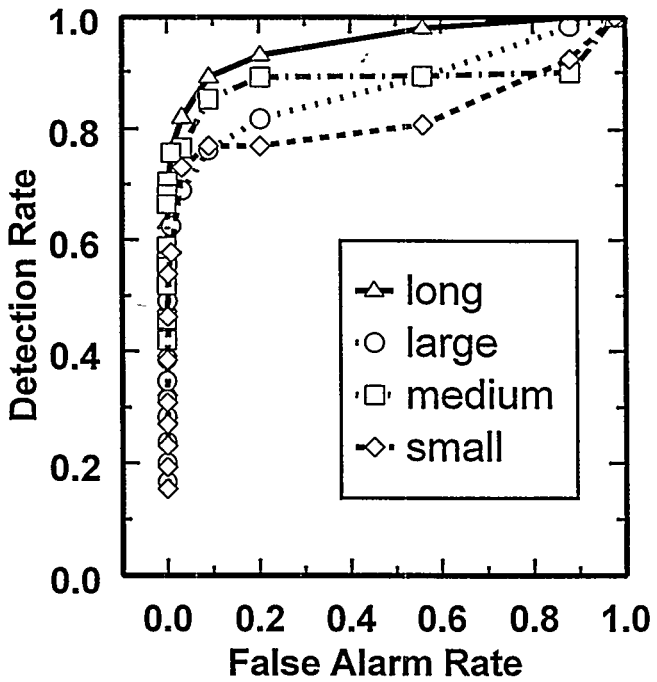

Figure 17. ROC for anomaly detection in MTI image, 0.77m GSD, MTI altitude with degraded atmosphere.

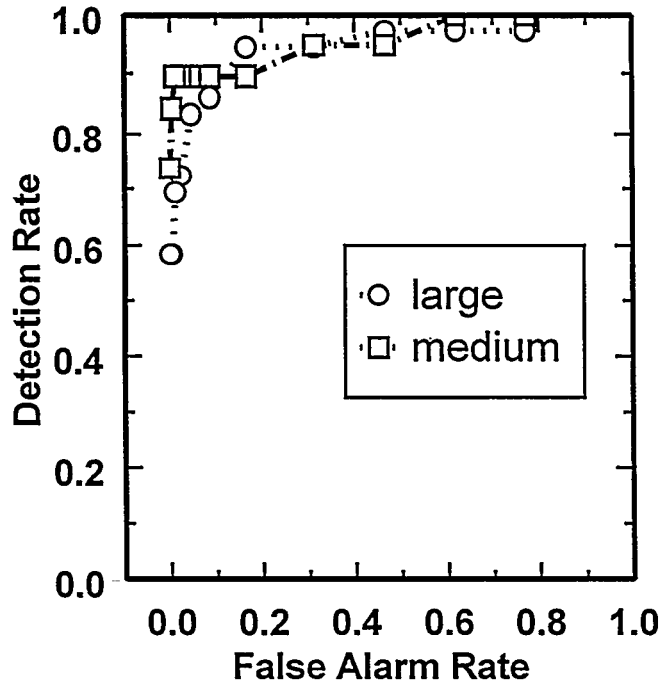

Figure 18. ROC for anomaly detection in HYDICE image, 5.0m GSD, MTI altitude with degraded atmosphere.

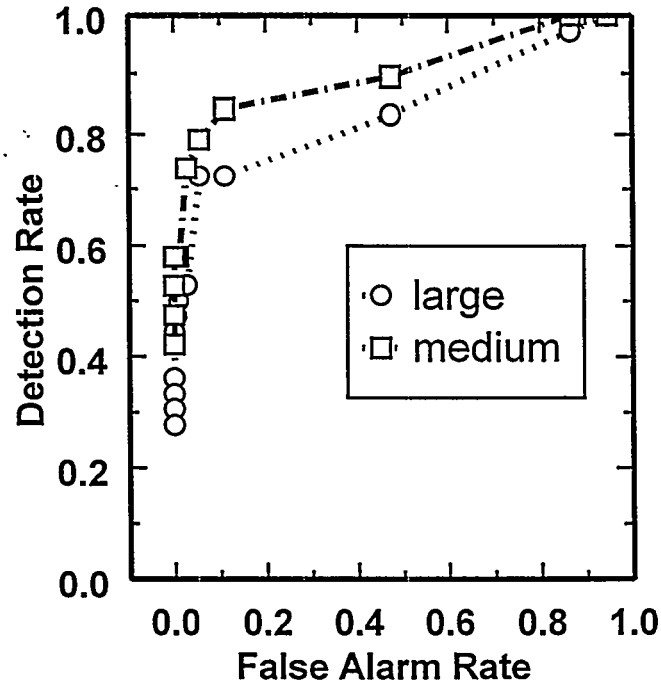

Figure 19. ROC for anomaly detection in MTI image, 5.0m GSD, MTI altitude with degraded atmosphere.