Target Detection in High-Resolution 3D Radar Imagery

A. Buck[a], J. M. Keller[a], M. Popescu[b], D. Sheen[c], and R.H. Luke[d]

[a] University of Missouri, Electrical and Computer Engineering Department, Columbia, MO
[b] University of Missouri, Health Management and Informatics Department, Columbia, MO
[c] Pacific Northwest National Laboratory, Richland, WA
[d]U.S. Army RDECOM CERDEC, Night Vision and Electronic Sensors Directorate, Fort Belvoir, VA, USA

ABSTRACT

Recently, the Stalker system has been developed as a high-resolution three-dimensional radar imaging system for the detection of concealed roadside explosive hazards. This system has shown considerable capability in distinguishing between true targets and false alarms using conventional processing techniques such as RX filtering on 2D projections of the data. In this paper, we develop an extension of these methods for use with 3D radar imagery. We show several different prescreening approaches for automatically marking potential target locations and describe an evaluation program called the Tiger scorer. We tested our approach on data collected at an arid U.S. Army test site.

Keywords: Target detection, 3D radar imagery, voxel processing

1. INTRODUCTION

In this paper, we present a processing methodology for detecting potential targets in 3D radar imagery. The Stalker system has recently been developed by the Pacific Northwest National Laboratory (PNNL) as a high-resolution three-dimensional radar imaging system specifically for the detection of concealed roadside explosive hazards[1]. The Stalker data is unique to other sensing modalities in its ability to reconstruct a volumetric image representation of the scene, resulting in a dense cube of image voxels (Figure 1). This is in contrast to other 3D sensing modalities such as LIDAR, which produces an irregular point cloud describing the shape and reflectivity of surface objects. Our previous work with LIDAR[2] showed promising results for unoccluded targets, however the nature of the laser-based sensor prevented the detection of most concealed targets. The Stalker radar system has the ability to penetrate light vegetation and reveal targets that may be missed with other sensing modalities.

Our main contribution in this work is the development of several different prescreening methods for processing 3D radar imagery such as the data produced by the Stalker system. We consider both 2D and 3D approaches using energy-based detection and size-contrast filtering. We examine each of these methods independently and propose a simple fusion technique for improving the performance. The results are evaluated in a blind test using a program we refer to as the Tiger scoring program, which is designed specifically for the task of roadside target detection. The outline of the remainder of this paper is as follows. Section 2 describes our processing methodology including a description of the various prescreeners and the Tiger scorer. Section 3 presents the Stalker dataset used in our experiments and Section 4 shows the results of our experiments. Finally, Section 5 gives our conclusions.

2. METHODOLOGY

2.1 Overview

Our objective in this work is to develop an algorithm that can take beamformed data from the Stalker system and produce predictions of likely target locations within the data. Figure 2 shows an overview of our processing workflow. The first step after importing the Stalker data is to decide whether to use traditional two-dimensional processing techniques or to operate in the native 3D space. Many algorithms have been optimized for two-dimensions, so it can sometimes be useful to restrict the processing to 2D projections of the data. In this case, we use a maximum-value top-down projection that we can easily map into real-world coordinates for scoring. Other projections could be used (e.g. to

extract two-dimensional features), however in this work we consider only the top-down projection for use with the 2D prescreeners. Alternatively, if the computational resources are available, we can take advantage of the full richness of the dataset and compute alarm locations in the original 3D space. For scoring purposes, the 3D alarm coordinates are reduced to real-world horizontal coordinates, ignoring any elevation.

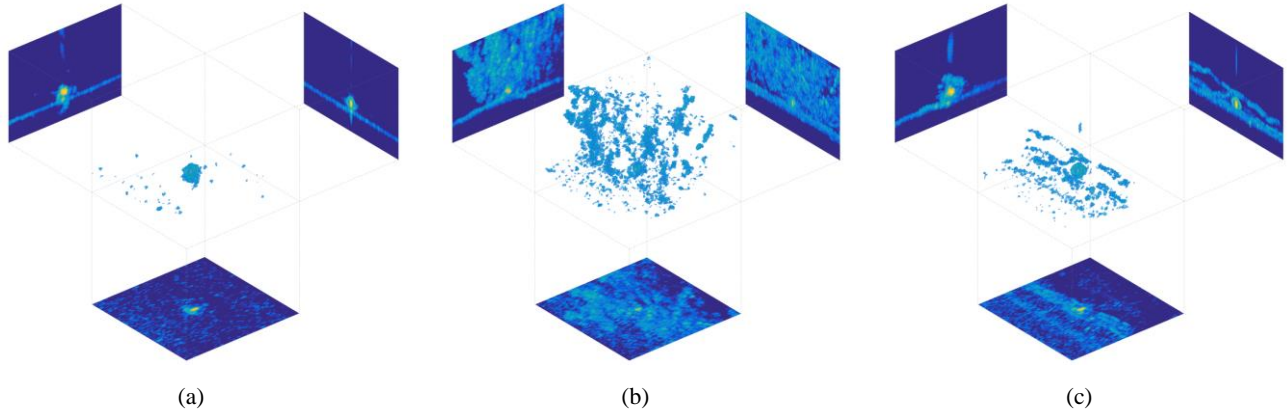<div align="center">(a)   (b)   (c)</div>

Figure 1. Example targets as seen by the Stalker system: unconcealed target (a), target in a bush (b), and target in ironwood (c). The native Stalker data format is a 3D volumetric image, which can be challenging to visualize. The center part of each subfigure shows a 3D scatter plot of the voxels that have intensity values exceeding a given threshold. The three images around each scatterplot show the max-value projections of the data in each of the dimensions.
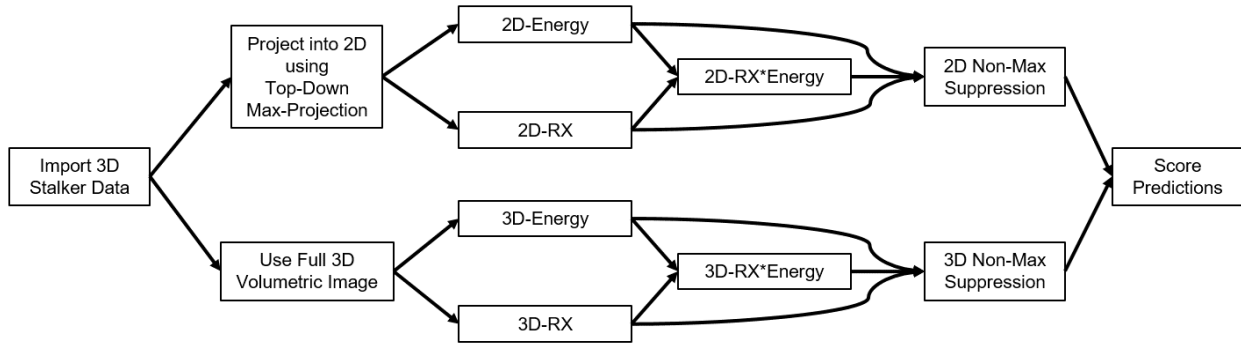

Figure 2. Overview of the processing methodology.

Ultimately, our goal is to identify the locations in the data that are most likely to represent targets. To accomplish this, we generate a large number of candidate alarm locations using a prescreening algorithm and give each one a confidence value. The purpose of the prescreener is to identify anomalous areas in the data that could possibly contain targets and would be worth investigating further, either by a human analyst or with a feature-based classifier. Two commonly used prescreening methods are energy-based prescreeners and size-contrast filters. After applying the prescreener, we use non-maxima suppression to produce a final set of alarm locations. These predictions are then scored using known ground truth.

2.2 2D Prescreeners

For the 2D processing workflow, we begin by projecting the maximum values of the 3D data along the vertical axis to produce a 2D image. This image represents a top-down view of the scene and corresponds to real-world geographic coordinates. The pixel values of the projected image represent the maximum energy intensities seen by the radar at a particular geographic location. We can use the energy values directly as the confidence of finding a target at a given location, or we can further process this image by identifying anomalies using a standard RX detector[3] as shown in Figure

3. In this work, we set the inner, guard, and outer window sizes of the RX filter as 11, 21, and 41 voxels respectively. (Note that the sizes are odd values to accommodate centering the windows on a single pixel.) Each voxel of Stalker data represents approximately 1 cm^2, so this equates to window sizes of approximately 11, 21, and 41 cm. The RX detection algorithm moves a sliding window across the entire image and computes the univariate Mahalanobis distance between the distributions of pixel intensity values within the inner window and outer annulus. The guard window serves to separate the two regions. The resulting confidence value is computed as

$$d(p, q) = \max(0, (\mu_p - \mu_q) / \sigma_q) \tag{1}$$

where μ_p and μ_q are the mean values of the inner and outer regions respectively, and σ_q is the standard deviation of the outer region. By only accepting positive values, we ensure that we are locating high-energy regions surrounded by low-energy regions and not the other way around. To improve processing time, the RX filter can be implemented using integral images[4].
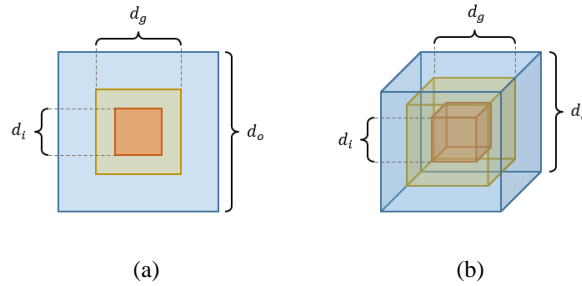

Figure 3. Description of the RX filter parameters for the 2D case (a) and 3D case (b). The inner window diameter is defined as d_i, the guard window diameter is defined as d_g, and the outer window diameter is defined as d_o.

We can apply the RX filter to the energy image to produce a confidence surface where each pixel represents the firing strength of the filter. The top-left image in Figure 4 shows an example of a top-down max-projected energy image, and the top-center image shows the result of applying the RX filter. This scene shows the berm of a gravel road travelling horizontally at the bottom of each image. A true target is located in the bush along the side of the road in the lower right corner. The energy image clearly shows the bush and road boundary, but also includes a significant amount of noise that fades away farther from the road. The filtered RX image shows several bright spots where there are local anomalies, but does not discriminate between areas based on the energy values. To compensate for this, we produce a third image, shown in the top-right of Figure 4 that multiplies the RX confidence and the original energy values together. We call this the RX*Energy image and it effectively suppresses regions of the RX image that have low energy values while retaining those with high values. Note that in the top-right image of Figure 4 nearly the entire image has been reduced to a confidence value of zero except for a few small patches around the bush (containing the target) and some regions along the roadside berm.

To produce alarm locations from the confidence images, we apply non-maxima suppression with a suppression radius of 100 pixels (approximately 1 meter). This iterative algorithm searches for the largest confidence value in the image and marks an alarm at that location with the corresponding confidence value. All confidences within the suppression radius are then set to zero and the next highest confidence value is found in the image. This process repeats until the entire image has been searched, resulting in a dense covering of alarm locations with no two alarms closer than the suppression radius. The bottom row of Figure 4 shows the alarm locations found for this example scene using non-maxima suppression. The sizes of the circles are proportional to the alarm confidence. Note how the alarm over the true target is significantly larger than the other alarms in the RX*Energy image, although the difference is less significant in the other images.
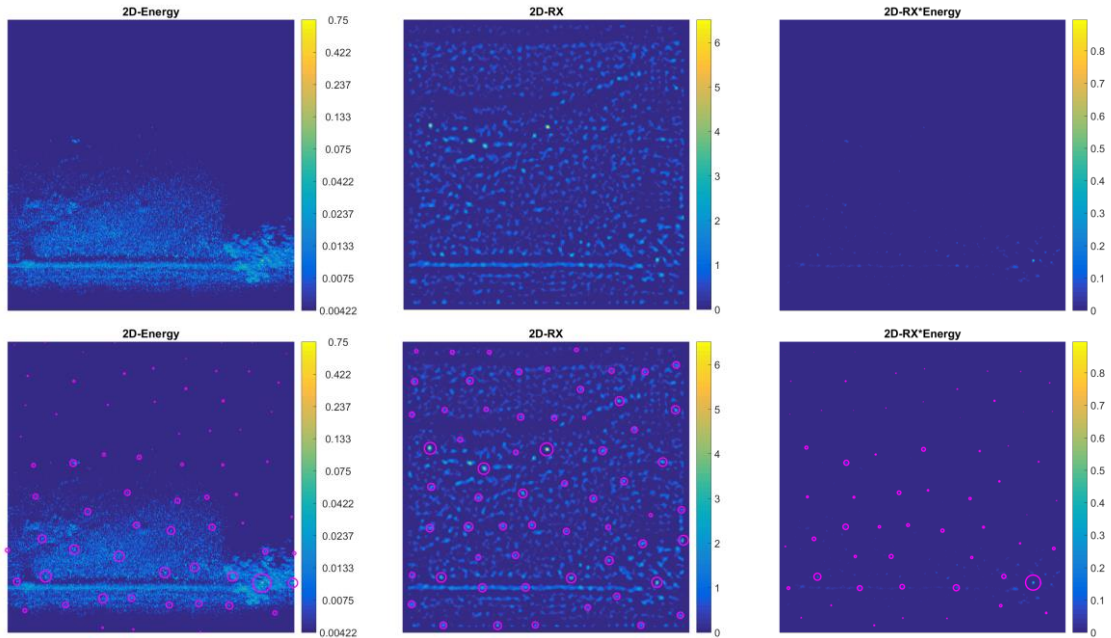
Figure 4. 2D prescreeners applied to the top-down max-projections of a frame of Stalker data. Alarm locations are circled with sizes proportional to the prescreener confidence. Note that the data is approximately 10 m^2 and a true target is located in the bush in the lower right.

2.3 3D Prescreeners

As an alternative to the 2D processing workflow, we can apply the same prescreeners to the original 3D data. In this case, we produce a confidence volume where each voxel corresponds to the confidence that a target appears at that location. We then use non-maxima suppression in all three dimensions to generate alarm locations. Note that in addition to a 2D coordinate that can be converted into latitude and longitude, the 3D prescreeners also provide elevation information about each alarm. This can provide additional context allowing improbable alarm locations to be removed, although no such filtering is done in this work. For the 3D-Energy prescreener, we simply use the beamformed Stalker data values as the confidence surface. For the 3D-RX prescreener, we extend the RX detector used in the 2D case to three dimensions, as shown in Figure 3b. For the combined 3D-RX*Energy prescreener, we multiply these two volumes together. Figure 5 shows the same scene from Figure 4 processed with the 3D prescreeners. As with the 2D-RX prescreener, the 3D-RX prescreener can be implemented using three-dimensional integral images[5].
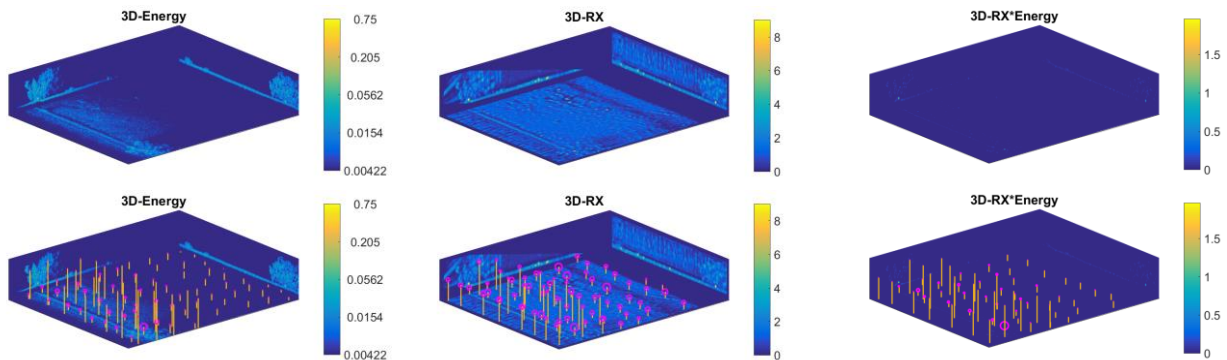

Figure 5. 3D prescreeners applied to the same scene from Figure 4. The max-projection in each dimension is shown and the alarms are displayed as stem plots in the bottom row. The size of the circles is proportional to the prescreener confidence.

2.4 Tiger scoring system

The Tiger scoring system has been developed by Night Vision and Electronic Sensors Directorate to evaluate the performance of various algorithms for roadside target detection. The system takes as input a configuration file that describes the locations of the ground truth targets, the lane boundary, and the alarm locations and confidence values that are to be scored. Additionally, the configuration file specifies a scoring radius, bias offset, and a series of flags to indicate the types of targets that should be considered. Multiple alarm files can be combined in the configuration file and the system will generate an averaged result in addition to the result of each alarm file independently. A set of figures are generated for each configuration that includes the receiver operating characteristic (ROC) curves, alarm distribution around the ground truth targets, a plot of sorted alarm confidences, and a table of the false alarm rates when each target was first detected. These plots are described in more detail as part of the results in Section 4.

To score a set of alarms, the Tiger scorer computes the distance between each alarm and the nearest ground truth target. This is accomplished using an efficient tessellation of the lane boundary and disregarding any alarms that are not within the scoring region. Alarms that fall within the scoring radius of a ground truth target are marked as positive hits and the remainder are marked as false alarms. If multiple alarms appear within the scoring radius of a single target, only the one with the highest confidence value is considered. The alarms are sorted by confidence value and the ROC curve is computed to show how the probability of detection (PD) compares to the false alarm rate (FAR) as the confidence threshold is adjusted. For a given threshold t, PD(t) is computed as the number of ground truth targets that have an associated alarm (within the scoring radius) with a confidence value greater than or equal to t. We compute FAR(t) as the number of false alarms that have confidence values greater than or equal to t, normalized by the area of the scoring region. The ROC curve characterizes the performance of the detector and it can be used to compare different algorithms. If multiple alarm files are provided (e.g. from separate passes over the same emplacement), the combined alarm confidence values are sorted and an average PD and FAR value are computed for each threshold.

3. DATA

Our method was blind tested on a dataset collected at an arid U.S. Army test site described in Table 1. These four emplacement configurations were used to collect our experimental results, but our methods were developed on an earlier independent dataset that was collected in a similar way. In this way, our results were acquired on blind data that had not been used in the development of our methods. For each configuration, the targets were positioned along the side of a road, some placed within bushes or partially buried within an embankment. The format of the data we receive from the Stalker system is a sequence of 3D beamformed intensity images. Each image in the sequence covers approximately 10 meters by 10 meters of terrain and about 2 meters of elevation at a resolution of roughly 1 cm^3. Because of the downward angle of the radar, the height of the beamformed data decreases at farther distances from the vehicle, giving a large region with missing data. This, along with the low energy values returned at beamformed locations representing empty space, results in relatively few high-energy voxels. A viewer is provided to produce max-projection images of the data in each dimension and a logarithmic scale is usually used to visualize the Stalker imagery.

Table 1. Description of the dataset.

Lane #	Emp.	Length	Runs	# of Targets
1	A	1 km	2	11
1	B	1 km	2	11
2	A	1 km	2	29
2	B	1 km	2	29

A run consists of a sequence of several 3D beamformed images with approximately 1m overlap. Our prescreening algorithms operate on each image separately and the resulting alarm locations are stored in the local Stalker coordinate system. After we have determined the alarm locations, we compute a transformation matrix for each image that maps alarms from Stalker coordinates to UTM coordinates. The alarms from each run are then grouped together for scoring.

4. RESULTS

We now compare the performance of the prescreeners with hand-labeled alarm locations produced by PNNL. The hand-labeled locations were marked by two human analysts with an intimate knowledge of the Stalker system. A scoring rubric was used to identify alarm locations with confidence values of 0.9, 0.6, 0.3, and 0.1. For the automated prescreeners, the alarms were generated using the methods described in Section 2 to produce a list of alarm locations and confidence values for each run. These locations are mapped to UTM coordinates and scored against known ground truth. We use the Tiger scoring system to evaluate the alarm files with a scoring radius of 0.25 meters. Each of the four lane emplacement configurations consists of two independent runs that are combined using confidence level averaging. The ROC curves are shown in Figures 6 and 7.
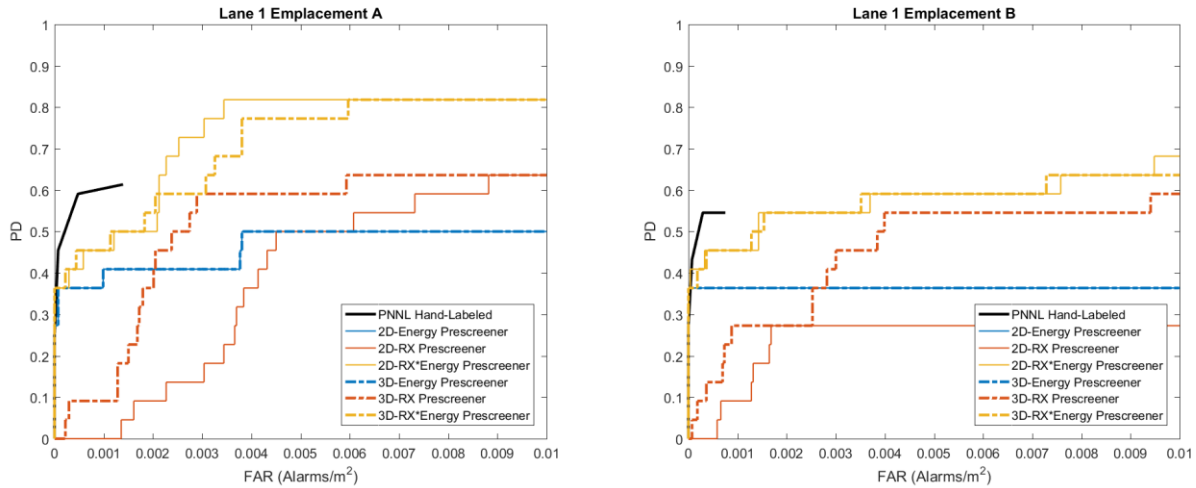

Figure 6. Prescreener comparison results on Lane 1. Each plot shows the ROC curves for the three 2D prescreeners (Energy, RX, and RX*Energy) as well as their 3D counterparts. A ROC curve showing human-level performance is shown in black.
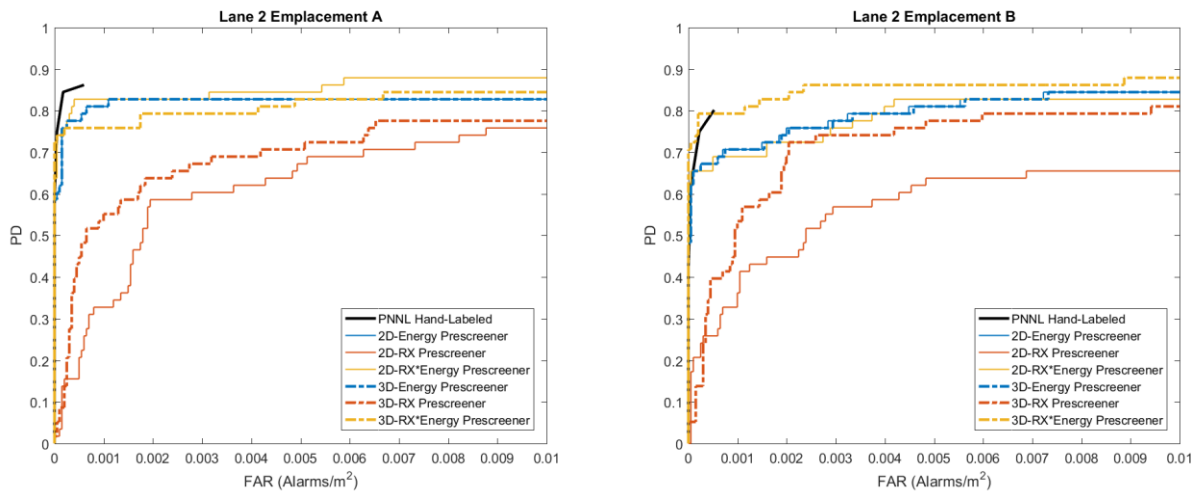

Figure 7. Prescreener comparison results on Lane 2. Notation is identical to Figure 6.

Figure 6 shows the results from Lane 1 and Figure 7 shows the results from Lane 2. The 2D prescreeners are shown as thin solid lines and the 3D prescreeners are shown as thicker dotted lines. The hand-labeled scores are shown in black. Note that the hand-labeled ROC curves appear truncated because they represent far fewer alarm locations than the ROC

curves of the automated prescreeners. The energy prescreeners in blue appear nearly identical for both the 2D and 3D versions over all lane configurations. This occurs because the high intensity peak locations used by both versions of the prescreener occur with the same horizontal coordinates. The two prescreeners will only differ when there are multiple high intensity voxels located near each other horizontally, but far enough apart vertically that the 3D version of the energy prescreener produces separate alarms. The RX prescreeners in red seem to have reduced performance compared to the energy prescreeners, particularly at very low false alarm rates. This can be explained by observing the alarms generated in Figures 4 and 5. The RX prescreeners consider only the contrast between the inner and outer regions and do not depend on the magnitude of the energy. They therefore produce many more high confidence false alarms on noisy objects in the data such as flat-faced rocks that actually have low energy values. Although the 3D-RX prescreener does not perform as well as the energy prescreener at low false alarm rates, it does outperform the 2D version for both lanes and even surpasses the energy prescreener at higher false alarm rates on Lane 1. The combined RX*Energy prescreeners in yellow outperform both of the other types of prescreeners in nearly all circumstances. However, it is less clear whether the 2D or 3D version works better on this dataset. The 2D-RX*Energy prescreener appears to perform better on Emplacement A of both lanes, but the 3D-RX*Energy prescreener does better on Emplacement B of Lane 2. The RX*Energy prescreeners compare favorably to human-level performance and even give higher PD values at higher FAR. On Emplacement B of Lane 2, the 3D-RX*Energy prescreener exceeds human-level performance.

To gain a better understanding of the performance of each method, we examine the additional output figures of the Tiger scorer. First, Figures 8 and 9 show the relative offsets of the alarms relative to the closest ground truth targets. For brevity, we show only the results from Emplacement A of both lanes. These plots are used to identify any systematic bias in the system and to observe the spread of alarms. A circle is drawn indicating the scoring radius. Only alarms that fall within this circle are marked as true targets. Higher confidence alarms are shown in red. The plots in Figure 9 show more alarms because Lane 2 contains more ground truth targets. An ideal classifier would place all alarms in a tight cluster centered at the origin and give them all high confidence, similar to the left-most plots showing the hand-labeled alarms. We notice that the 3D prescreeners tend to have a greater spread than the 2D prescreeners have. This is because 3D non-max suppression can produce alarms that appear near each other in the top-down projection, but in actuality have different elevations. Because the Tiger scorer does not evaluate elevation, these projected locations may be located near the target and contribute to the apparent spread of the 3D prescreeners.
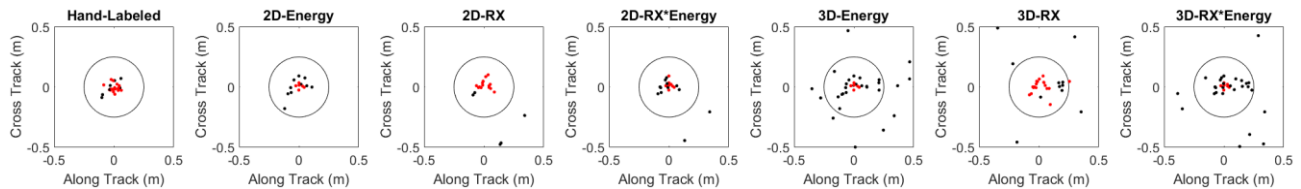

Figure 8. Alarm offsets from ground truth locations for Emplacement A of Lane 1. The circle indicates the 0.25-meter scoring radius around each target and alarms are plotted relative to the closest true target. Alarms with higher confidence are shown in red.
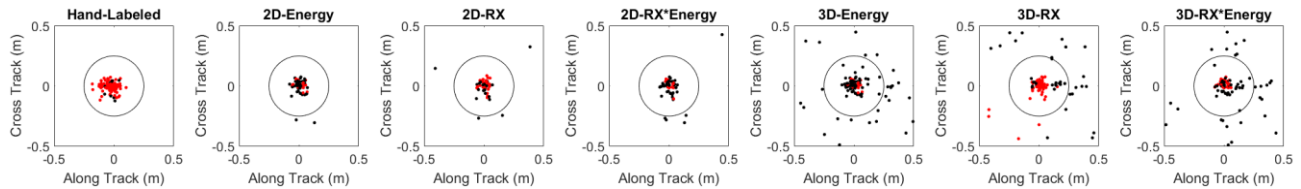

Figure 9. Alarm offsets from ground truth locations for Emplacement A of Lane 2. Notation is identical to Figure 8.

The next plots we examine are the sorted confidence plots shown in Figures 10 and 11. Again, for brevity we show only the results from Emplacement A. These plots show the magnitude of the alarm confidence values for each set of alarms in descending order. True targets are shown with a blue 'x' and false alarms are shown with a red '+'. The x-axis is given a logarithmic scale to show the highest-confidence alarms more clearly. An ideal classifier would cleanly separate

the true targets and false alarms, allowing for perfect classification with a single confidence threshold. These plots show the distribution of confidence values across the set of alarms and the overlap between the classes. We notice that the RX*Energy prescreeners have the sharpest "bend" that keeps the confidence value of false alarms low and pushes true targets to the left with higher confidence. The plain RX prescreeners clearly have a high degree of overlap between classes giving low separability.
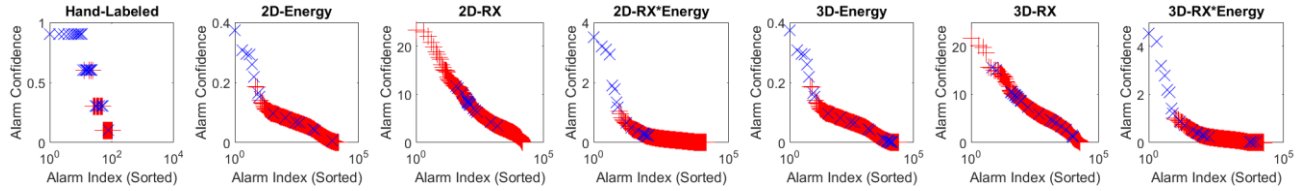

Figure 10. Sorted confidence plots for Emplacement A of Lane 1. The alarms are sorted by confidence along the x-axis and plotted according to the confidence along the y-axis. True targets are shown with a blue 'x' and false alarms are shown with a red '+'. Note the logarithmic scale of the x-axis.
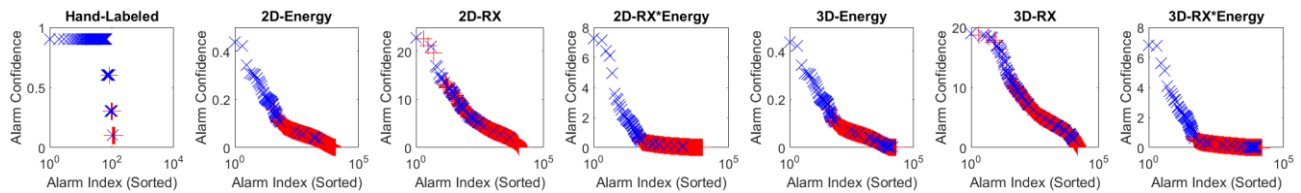

Figure 11. Sorted confidence plots for Emplacement A of Lane 2. Notation is identical to Figure 10.

The final set of output figures we examine from the Tiger scorer are image tables of the minimum FAR required for detecting each target, shown in Figure 12. Again, we show only the results from Emplacement A. The figure is separated with Lane 1 on the left and Lane 2 on the right. Each column represents a ground truth target and the gray levels indicate the minimum FAR required to detect that target. Brighter values indicate higher FAR and if a target is not detected, it is marked with an 'N/A' symbol. The results of the different methods are shown on the rows. Each of the automated prescreeners was applied to two runs and the hand-labeled results were collected from two analysts on the same two runs. The bottom row shows the approximate occlusion of each target, with darker values corresponding to less occlusion (i.e. easier to detect). Our first observation is that some targets were consistently detected across all methods while other targets were hardly detected at all. Of particular interest are the targets that were detected by only a few of the methods or were found by one of the automated prescreeners but missed by the human analysts.

Figure 13 shows three examples targets as they appear in the max-projection images of the native Stalker data. Red dotted lines indicate the extent of the scoring radius around the ground truth target, which appears as a circle in the top-down projection and as lines in the other two projections. The first example (Figure 13a) shows a partially occluded target that was detected by all of the prescreeners and the human analysts. It is easy to identify a bush in this image with a bright target at the base. Figure 13b shows a highly occluded target that was missed by most of the algorithms and it is difficult to see any targets in the imagery. Figure 13c shows a moderately occluded target that was usually missed by the analysts, but was detected by most of the automated prescreeners. Although it is hard to give high confidence to this example based on visual inspection, the prescreeners were able to identify features that make this true target location stand out from the rest of the imagery.
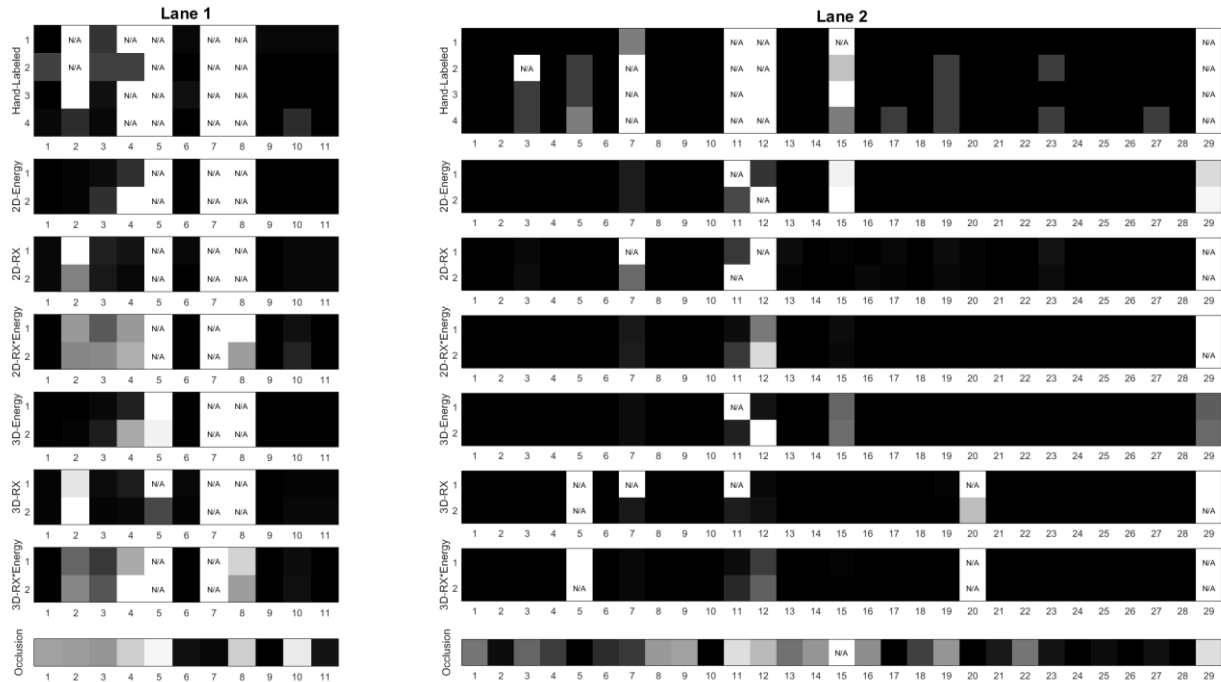
Figure 12. Image tables of the minimum FAR required for detecting each target for Emplacement A of Lanes 1 and 2. Each column represents a target and the gray level indicates the minimum FAR required for detection, with brighter values corresponding to higher FAR. An 'N/A' symbol is used to indicate targets that were not detected. The bottom row indicates the approximate occlusion of each target, with darker gray levels corresponding to less occlusion (i.e., easier to detect).
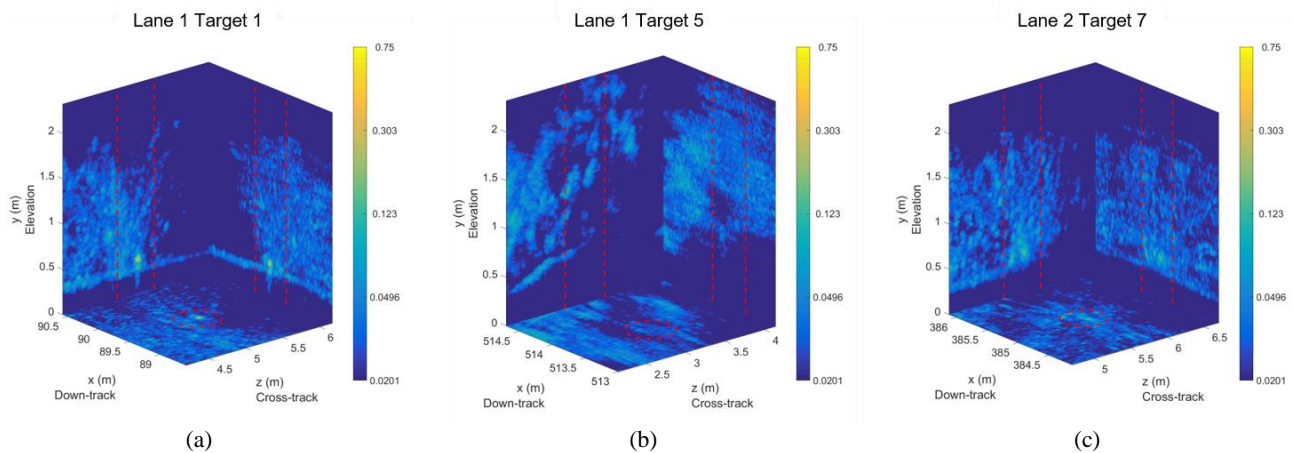

(a) (b) (c)

Figure 13. Examples of the max-projection images of the native Stalker imagery at ground truth target locations. Red dotted lines indicate the extent of the scoring radius as it is projected through all elevations.

5. CONCLUSIONS

In this paper, we presented a processing methodology for use with 3D radar imagery. We compared both 2D and 3D prescreeners using energy-based detection and size-contrast filters. Our results show that the detection rate can be improved by combining these two approaches. Using the RX*Energy prescreeners, we were able to obtain results that are competitive with human-level performance. We used the Tiger scoring system in a blind test scenario to evaluate our

results and identify targets that proved to be challenging to detect. The Tiger scorer can help in the design of additional features for classification by focusing on these difficult targets. Future work in this area will involve extracting features from the alarm locations and developing classification algorithms to improve the prescreener performance. The 3D radar imagery provided by the Stalker system offers the potential to characterize alarms by volumetric shape, which could lead to useful feature descriptors that can help reduce the number of false alarms.

ACKNOWLEDGEMENTS

REFERENCES

[1]   Sheen, D., Hall, T., McMakin, D., Jones, A., Tedeschi, J., "Three-dimensional radar imaging techniques and systems for near-field applications," Proc. SPIE 9829, Radar Sensor Technology XX, 98290V (2016); doi: 10.1117/12.2229235.

[2]   Buck, A., Keller, J. M., Popescu, M., "Improving the detection of explosive hazards with LIDAR-based ground plane estimation," Proc. SPIE 9823, 982313 (2016).

[3]   Reed, I. S., Yu, X., "Adaptive multiple-band CFAR detection of an optical pattern with unknown spectral distribution," IEEE Trans. Acoust., 38(10), 1760–1770 (1990).

[4]   Viola, P., Jones, M. J., "Robust real-time object detection," Int. J. Comput. Vis., 1–30 (2001).

[5]   Tapia, E., "A note on the computation of high-dimensional integral images," Pattern Recognit. Lett., 32(2), 197–201, Elsevier B.V. (2011).