

Image Segmentations for Through-the-Wall Radar Target Detection

CHER HAU SENG, Student Member, IEEE
University of Wollongong
Australia

MOENESS G. AMIN, Fellow, IEEE
FAUZIA AHMAD, Senior Member, IEEE
Villanova University

ABDESSELAM BOUZERDOUM, Senior Member, IEEE
University of Wollongong
Australia

Detection of stationary targets using pixel-wise likelihood ratio test (LRT) detectors has been recently proposed for through-the-wall radar imaging (TWRI) applications. We employ image segmentation techniques, in lieu of LRT, for target detection in TWRI. More specifically the widely used between-class variance (BCV) thresholding, maximum entropy segmentation, and K-means clustering are considered to aid in removing the clutter, resulting in enhanced radar images with target regions only. For the case when multiple images of the same scene are available through diversity in polarization and/or vantage points around a building structure, we propose to use image fusion, following the image segmentation step, to generate an enhanced composite image. In particular, additive, multiplicative, and fuzzy logic fusion techniques are considered. The performance of the segmentation and fusion schemes is evaluated and compared with that of the assumed LRT detector using both electromagnetic (EM) modeling and real data collected in a laboratory environment. The results show that, although the principles of segmentation and detection are different, the segmentation techniques provide either comparable or improved performance over the LRT detector. Specifically, in the cases considered, the maximum entropy segmentation produces the best results for detection of targets inside building structures. For fusion of multiple segmented images of the same scene, the fuzzy logic fusion outperforms the other methods.

Manuscript received November 21, 2011; revised May 31 and September 21, 2012; released for publication December 20, 2012.
IEEE Log No. T-AES/49/3/944618.

Refereeing of this contribution was handled by R. Narayanan.

The work by M. G. Amin and F. Ahmad was supported in part by U.S. Army Research Laboratory Contract W911NF-07-D-0001 and by ONR Grant N0014-10-1-0455. A. Bouzerdoum is supported in part by the Australian Research Council.

The work by C. H. Seng was completed during his tenure as a visiting research associate at the Center for Advanced Communications, Villanova University, Villanova, PA.

Authors' addresses: C. H. Seng and A. Bouzerdoum, School of Electrical, Computer and Telecommunications Engineering, University of Wollongong, Wollongong, NSW 2522, Australia; M. G. Amin and F. Ahmad, Radar Imaging Laboratory, Center for Advanced Communications, Villanova University, 800 Lancaster Ave., Villanova, PA 19085, E-mail: (fauzia.ahmad@villanova.edu).

0018-9251/13/\$26.00 © 2013 IEEE

I. INTRODUCTION

Imaging of building interiors has been a subject of interest in many applications related to rescue missions, homeland security, and defense [1–7]. Indoor images are typically characterized by the presence of both spatially extended targets, like exterior and interior walls, and compact, point-like targets, such as humans. Also, near-field operations give rise to point spread functions that vary in range and cross-range. Accordingly, the same target can have different spatial distribution, depending on its position.

Detection of stationary targets in through-the-wall radar imaging (TWRI) using statistical detectors based on likelihood ratio tests (LRT) has been discussed in [8], [9]. Specifically, a Neyman-Pearson (NP) test was used in [8] to detect targets in indoor radar images by defining pixel-wise null and alternative hypotheses, coupled with a user-defined false alarm rate (FAR). However, in this test, the exact statistics of the radar images need to be known a priori. As this information is target and scene dependent, the NP test was therefore extended in [9] to iteratively adapt the test parameters to the radar image statistics. In order to improve and optimize the parameter estimates, morphological filtering in the image domain was introduced as a preprocessing stage in [10]. By adapting the target detector's parameters to the changing characteristics of the preprocessed radar images, a more robust detection can be produced. To date, both NP and Bayesian tests have been employed for the detection and fusion of multi-view and multi-polarization through-the-wall radar (TWR) images [9, 11]. Generally, the statistical detectors that incorporate the LRT generate a binary mask that depicts the target locations in the image.

Although the LRT approach has been successfully applied to some scenes and several scenarios, it assumes particular probability density functions (pdfs) for all targets in the scene. The assumed pdfs may prove unsuitable for complex scenes that are acquired using real data or data from numerical electromagnetic (EM) modeling. The need to define appropriate pdfs for all targets and clutter in the image, and to specify an appropriate FAR by the operator presents a shortcoming of the LRT detector, since in most cases, neither the pdfs nor the FAR is known a priori. In this paper, we propose the use of image segmentation methods as an alternative to target detection. In lieu of the LRT, we apply image segmentation methods that exploit information obtained from image intensities and histograms to separate targets from clutter, and thereby enhance the image quality.

There exist many different types of image segmentation methods [12], image thresholding methods and region-based methods, to name a few. Region-based methods, such as the watershed, are

not suitable for TWR images as the radar returns are very sparse and target regions tend to be very small. One of the most basic thresholding methods is to set a threshold at the valley of a bimodal intensity histogram [13]. However, TWR image histograms generally do not exhibit a traceable valley. There have been some automatic threshold selection techniques proposed in order to overcome this difficulty. Such techniques select an optimal threshold by a discriminant criterion so as to maximize the separability of the resultant classes. We propose to examine the three most commonly used automatic image thresholding methods, namely, between-class variance (BCV) thresholding [14], entropy-based segmentation [15], and K-means clustering [16] for TWR image segmentation. The segmentation process associated with these techniques produces threshold values, thus generating binary masks that are similar to that of the LRT detector. Using the generated binary image, the original image is masked, producing an enhanced image with the target regions only.

For the detection of targets in multiple images, the proposed image segmentation-based target detector involves a two-step process that predicates on the availability of TWR images acquired from different viewing angles and for different polarizations. Instead of performing the fusion and detection at the same time, like in the case of the LRT detector, image segmentation techniques are first applied to individual radar images to separate each image into several regions. The segmented results are then fused using pixel-wise addition, multiplication, or through the use of fuzzy logic [17]. The latter has proven to be a valuable tool since it deals with the fusion problem at a local level, i.e., different parts of the image can be fused differently, depending on the relative presence of the target and clutter. Reversing this sequence of operation, by applying fusion prior to segmentation, provides inferior results and does not fully utilize the potential enhancement of individual images before combinations. It is noted that image segmentation methods have been applied for object detection in synthetic aperture radar (SAR) aerial images [13, 18, 19, 20] and ground penetrating radar (GPR) images [21]. They have also been considered a preprocessing stage for change detection [22, 23] and object classification [24–26] in radar images. It is also noted that a similar segmentation-fusion approach has been proposed in [27] to improve the classification of SAR images. However, our work differs by applying the segmentation and fusion methods as an alternative to the LRT detector for both single and fused images, in the specific area of TWRI.

In this paper, we investigate the performances of the LRT detector and the proposed method for indoor radar imaging, both with and without image fusion. Both target detection schemes are evaluated and compared using real two-dimensional

(2D) polarimetric and multi-view images collected at the Radar Imaging Lab, Center for Advanced Communications, Villanova University. These methods are also applied to numerical EM modeling data. For both real and modeled data, different scenarios that are representative of a variety of possible indoor scenes are constructed so as to strengthen our findings and conclusions. In order to compare methods within the same framework, the LRT detector is cast as another form of the image segmentation method with a corresponding threshold.

The performance is assessed in terms of image enhancements after the separation of target from clutter in individual images. The results show that in most cases considered, the entropy-based segmentation technique outperforms other image segmentation techniques, including the LRT method, by providing higher clutter suppression, while successfully maintaining the target regions. The BCV thresholding and K-means clustering methods provide similar results to those of the entropy-based segmentation when considering target regions; however, their respective clutter levels are much higher. Nevertheless, the proposed segmentation-based detectors are more advantageous than the LRT detector since they do not require a predefined pdf and FAR.

As for multiple image target detection, the performance of the proposed detection scheme is compared with that of the LRT detector that jointly utilizes multi-view and multi-polarization signatures. The corresponding results show that the entropy-based segmentation method, coupled with the pixel-wise fuzzy logic fusion, outperforms the LRT detector by providing higher clutter suppression, while keeping the number of missed detections relatively low. In essence the entropy-based segmentation method can be generally viewed as most suitable for target detection in TWRI.

The remainder of the paper is organized as follows. Section II details the TWR image formation process, where the wideband delay-and-sum beamforming is reviewed. Section III describes the LRT target detection method proposed in [9]. The alternative target detection method based on image segmentation techniques is then discussed in Section IV. Section V evaluates the performance of the LRT detector and image segmentation-based methods using real and numerical modeling data, and Section VI concludes the paper.

II. BEAMFORMING

In this section, we present the fundamental equations describing the delay-and-sum beamformer, considered in this paper, for imaging through walls. We consider a 2D SAR system in which a single antenna at one location transmits the signal and

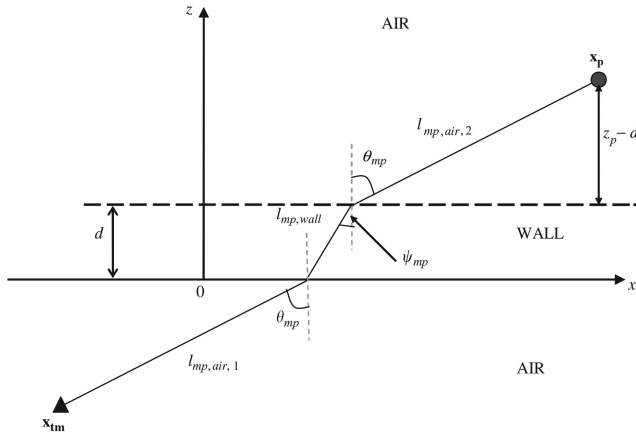


Fig. 1. Geometry on transmit for two-dimensional imaging.

receives the radar return, and then moves to the next location and repeats the same operation in a plane parallel to the front wall [7, 28].

Consider the scenario where the wall is located in the xy -plane and has a thickness d and a dielectric constant ε . Assume that there are M monostatic antenna locations, with a standoff distance z_{off} from the wall. The region to be imaged is located beyond the wall along the positive z -axis, as shown in Fig. 1. Let the transceiver, placed at the m th location $\mathbf{x}_{tm} = (x_{tm}, y_{tm}, -z_{\text{off}})$, illuminate the scene with a wideband signal $s(t)$. For the case of a single point target located at $\mathbf{x}_p = (x_p, y_p, z_p)$, the signal measured by the m th transceiver is given by $a(\mathbf{x}_p)s(t - \tau_{m,p})$, where $a(\mathbf{x}_p)$ is the complex reflectivity of the point target and $\tau_{m,p}$ is the propagation delay encountered by the signal as it travels from the m th transceiver to the target at \mathbf{x}_p , and then back to the same transceiver. The delay $\tau_{m,p}$ is given by

$$\tau_{m,p} = \frac{2l_{mp,\text{air},1}}{c} + \frac{2l_{mp,\text{wall}}}{v} + \frac{2l_{mp,\text{air},2}}{c} \quad (1)$$

where c is the speed of light and $v = c/\sqrt{\varepsilon}$ is the signal propagation speed through the wall. The variables $l_{mp,\text{air},1}$, $l_{mp,\text{wall}}$, and $l_{mp,\text{air},2}$ represent the traveling distances of the signal before, through, and beyond the wall, respectively, from the m th transceiver to the target at \mathbf{x}_p . This process is evaluated for each transceiver location until all M locations have been exhausted.

Although the corresponding M outputs can be processed to generate a three-dimensional (3D) image of the scene, we restrict ourselves to B-scan (cross-range versus downrange at a fixed height) images. This is because the images presented in this paper are either B-scan images or cross-range versus downrange images obtained with one-dimensional (1D) antenna arrays. The region of interest at height \bar{y} is divided into a finite number of pixels in downrange and cross-range, represented by the z - and x -coordinates, respectively. The complex composite signal corresponding to the image of the

pixel located at $\mathbf{x}_q = (x_q, \bar{y}, z_q)$ is obtained by applying time delays and weights to the M measurements, and then summing the results. The output for a single target case is given by

$$r_q(t) = \sum_{m=1}^M w_m a(\mathbf{x}_p) s(t - \tau_{m,p} + \tau_{m,q}). \quad (2)$$

Here, w_m and $\tau_{m,q}$ are, respectively, the weight and the focusing delay applied to the output of the m th transceiver. The focusing delay is given by (1) with the target voxel subscript p replaced by the focusing pixel subscript q . The complex amplitude image value for the pixel located at \mathbf{x}_q is obtained by passing the signal $r_q(t)$ through a filter matched to the transmitted pulse and sampling the output of the filter at time $t = 0$ as follows

$$I(\mathbf{x}_q) = \sum_{m=1}^M w_m a(\mathbf{x}_p) s(t - \tau_{m,p} + \tau_{m,q}) * h(t) \big|_{t=0} \quad (3)$$

where $h(t) = s^*(-t)$ is the impulse response of the matched filter. The process described is performed for all N pixels in the region of interest at height \bar{y} to generate the composite image of the scene. The general case of multiple targets can be obtained by the superposition of target reflections.

It is noted that the output of the delay-and-sum beamformer is a complex valued image. The magnitude image, corresponding to the complex amplitude image, is given by

$$\check{I}(\mathbf{x}_q) = |I(\mathbf{x}_q)|. \quad (4)$$

For target detection in the image domain, we deal with the magnitude image only.

III. LIKELIHOOD RATIO TESTS DETECTOR

In this section, we review the general image domain-based framework for statistical target detection based on the LRT, which utilizes multiple images corresponding to different viewing angles and/or polarization. Specifically, we consider the iterative version of the pixel-wise NP detector introduced in [9] that adapts the test parameters to the radar image statistics.

Let \check{I}_j , $j = 1, 2, \dots, J$ be the set of acquired magnitude images corresponding to a total of J viewing angles and/or polarization. The pixel-wise NP test is given as

$$\prod_{j=1}^J \frac{P_r(\check{I}_j | H_1)}{P_r(\check{I}_j | H_0)} \geq \gamma \quad (5)$$

where H_0 and H_1 denote, respectively, the null (target absent) and alternative (target present) hypothesis. The functions $P_r(\cdot | H_0)$ and $P_r(\cdot | H_1)$ are the conditional pdfs under the null and alternative hypothesis, respectively, and the parameter γ is the likelihood

ratio threshold, which can be obtained by specifying a desired FAR α as

$$\alpha = \int_{\gamma}^{\infty} P_{\ell}(\ell | H_0) d\ell \quad (6)$$

where $P_{\ell}(\ell | H_0)$ denotes the distribution of the likelihood ratio under the null hypothesis.

Let $\hat{\theta}_{H_0}^0$ and $\hat{\theta}_{H_1}^0$ denote the initial estimates of the parameter vectors θ_{H_0} and θ_{H_1} describing the pdfs under H_0 and H_1 , respectively. Given a FAR α , a binary image B_{NP}^1 , where superscript 1 represents the first iteration, can be obtained by evaluating (5). In order to enhance and optimize the estimation of the noise and test pdf parameters, morphological filtering is employed to obtain the binary image B_{MF}^1 (see [10] for details). This image can be used as a mask on the original set of images in order to obtain the revised parameter estimates $\hat{\theta}_{H_0}^1$ and $\hat{\theta}_{H_1}^1$. These revised parameters are then fed back to the NP test to obtain an improved detection result. The iteration stops when convergence is achieved. Figure 2 shows the block diagram of the iterative target detection approach.

It is noted that the final output of the LRT detector described above is a single binary image that indicates the presence of the targets. For a more detailed description of the LRT detector, the reader is referred to [8]–[11].

IV. IMAGE SEGMENTATION-BASED TARGET DETECTION

Unlike the LRT detector, which performs simultaneous detection and fusion of a set of magnitude images, the proposed method is a two-step process. First, an image segmentation technique is applied to the individual input images. Then, the segmented images are fused to generate an enhanced single composite intensity image that has high target intensities and low clutter and noise levels. Figure 3 shows the block diagram of the proposed image segmentation-based target detector.

It should be noted that when there is only one input image, only the image segmentation method will be applied. For the case of multiple input images, both image segmentation and image fusion methods will be applied for detection.

A. Image Segmentation

Image segmentation based on intensity and histogram information is a simple technique that involves the basic assumption that the objects and the background in the sensed image have distinct gray level distributions [29]. Since objects in remotely sensed imagery are often homogeneous, threshold values separating two or more regions in the gray level histogram can be obtained. Threshold

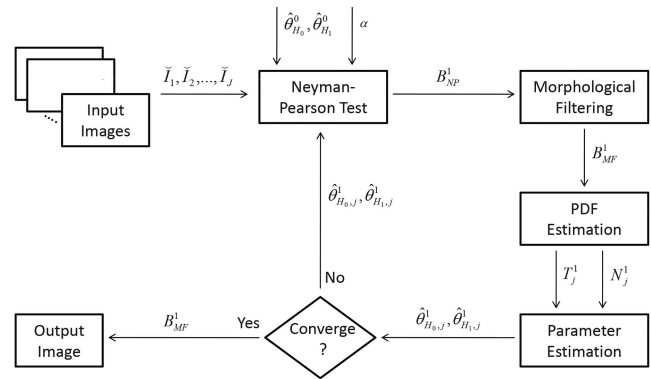


Fig. 2. Block diagram of iterative statistical target detector.

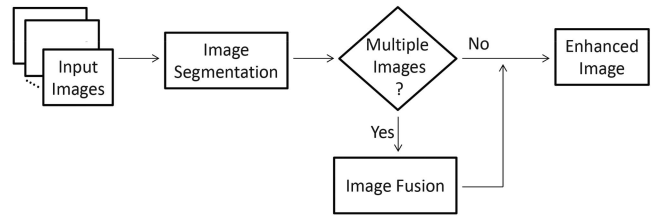


Fig. 3. Block diagram of proposed image segmentation-based target detector.

selection methods can be classified into two groups, namely, global methods and local methods. A global thresholding technique separates the entire image into target and background regions with a single threshold value, whereas local thresholding methods partition the given image into a number of subimages and determine a threshold for each of the subimages separately. As global thresholding methods are computationally less intensive, they have been more popular for radar image analysis [30]. In this paper, we consider two of the commonly applied global thresholding methods, namely, BCV thresholding [14] and entropy-based segmentation [15], as candidates for the image segmentation step of the proposed detection scheme. Since image segmentation can also be viewed as the partitioning of the observed intensities into groups, we also consider the application of K-means clustering [16].

Consider the histogram of an input image as a discrete pdf $\rho(i)$:

$$\rho(i) = \frac{f_i}{N} \quad (7)$$

with $\rho(i) \geq 0$ and $\sum_{i=0}^{L-1} \rho(i) = 1$, where f_i is the frequency of intensity level i and N is the total number of pixels in the image. Each pixel in the image assumes an intensity level from the set $\{0, 1, \dots, L-1\}$, where L denotes the number of intensity levels or histogram bins.

1) *BCV Thresholding*: The BCV thresholding method segments an image into two regions by determining a threshold value T_{BCV} that maximizes

the sum of class variances:

$$T_{BCV} = \arg \max_d \{p_{r1}(d)[m_{r1}(d) - m_i]^2 + p_{r2}(d)[m_{r2}(d) - m_i]^2\} \quad (8)$$

where m_i is the mean image intensity, r_1 and r_2 are the two regions of the image histogram relative to the intensity level d , $p_{r1}(d)$ and $p_{r2}(d)$ are the respective region probabilities, which are expressed as

$$p_{r1}(d) = \sum_{i=0}^d \rho(i) \quad (9)$$

$$p_{r2}(d) = \sum_{i=d+1}^{L-1} \rho(i) \quad (10)$$

and $m_{r1}(d)$ and $m_{r2}(d)$ are the means of the respective regions, which are given by

$$m_{r1}(d) = \sum_{i=0}^d \frac{i \cdot \rho(i)}{p_{r1}(d)} \quad (11)$$

$$m_{r2}(d) = \sum_{i=d+1}^{L-1} \frac{i \cdot \rho(i)}{p_{r2}(d)}. \quad (12)$$

All values of $d = 1, 2, \dots, L-2$ are considered and the corresponding functions (8)–(12) are evaluated. The intensity value d that produces the maximum sum of the class variances is chosen as the threshold value T_{BCV} .

2) *Entropy-Based Segmentation*: Similar to the BCV method, the entropy-based segmentation decides on the threshold value in an iterative fashion. Instead of maximizing the sum of class variances, the entropy-based segmentation maximizes the sum of class entropies. Based on the information derived from the image histogram, the entropy of two regions is maximized using the following equation:

$$T_H = \arg \max_d \{H_{r1}(d) + H_{r2}(d)\} \quad (13)$$

where $H_{r1}(\cdot)$ and $H_{r2}(\cdot)$ are the respective region entropies. Let p_i be the probability of intensity level i and $P_d = \sum_{i=0}^d p_i$ be the total probability. The entropy of each region can be expressed as

$$H_{r1}(d) = - \sum_{i=0}^d \frac{p_i}{P_d} \ln \frac{p_i}{P_d} \quad (14)$$

$$H_{r2}(d) = - \sum_{i=d+1}^{L-1} \frac{p_i}{P_d} \ln \frac{p_i}{P_d}. \quad (15)$$

Given that the entropy for a region can also be calculated as

$$H_d = - \sum_{i=0}^d \rho(i) \ln \rho(i) \quad (16)$$

the total entropy of the image can be expressed as

$$H_{\text{tot}} = - \sum_{i=0}^{L-1} \rho(i) \ln \rho(i). \quad (17)$$

Thus, (14–15) can be simplified as follows

$$\begin{aligned} H_{r1}(d) &= - \sum_{i=0}^d \frac{p_i}{P_d} \ln \frac{p_i}{P_d} \\ &= - \frac{1}{P_d} \left[\sum_{i=1}^d \rho(i) \ln \rho(i) - P_d \ln P_d \right] \\ &= \ln(P_d) + \frac{H_d}{P_d} \end{aligned} \quad (18)$$

$$\begin{aligned} H_{r2}(d) &= - \sum_{i=d+1}^{L-1} \frac{p_i}{P_d} \ln \frac{p_i}{P_d} \\ &= - \frac{1}{1 - P_d} \\ &\quad \times \left[\sum_{i=d+1}^{L-1} \rho(i) \ln \rho(i) - (1 - P_d) \ln(1 - P_d) \right] \\ &= \ln(1 - P_d) + \frac{H_{\text{tot}} - H_d}{1 - P_d}. \end{aligned} \quad (19)$$

Iterating d from 1 to $L-2$, the intensity value d that produces the maximum sum of the distribution entropies is chosen as the threshold value T_H .

3) *K-Means Clustering*: Clustering methods partition the observed intensities into classes, and can also be used for segmenting images. Also known as unsupervised classification, the classes are generally unknown and are explored based on the data by using a similarity measure. Given N pixels the K-means clustering method partitions the pixels into K clusters by minimizing the sum of the within-cluster variances (WCSS) [16]:

$$\text{WCSS} = \sum_{k=1}^K \sum_{i=1}^N \|v_i^k - \mu_k\|^2 \quad (20)$$

where v_i^k is the i th sample of the k th class with centroid μ_k . The pseudocode for the K-means clustering is given as follows

- 1) Initialize the number of classes K and centroids μ_k .
- 2) Assign each pixel to the group whose centroid is the closest.
- 3) After all the pixels have been assigned, recalculate the centroids.
- 4) Repeat steps 2 and 3 until the centroids no longer change.

Although the K-means is computationally very efficient, the major disadvantage is the need to specify the number of classes a priori. In the absence of this

knowledge, one may resort to measures that could estimate the number of classes automatically [31]. This approach is not considered here. Instead, we set the number of classes as two to ensure consistency with the other image segmentation methods.

For the case of only one input image, the binary image produced by the image segmentation methods is masked on the original input image to produce an enhanced image with the target regions only.

B. Image Fusion

When there are multiple input images, the segmented images are then fused to produce a single enhanced image. Here, we consider the commonly used image fusion techniques in TWRI, such as the pixel-wise additive fusion [32], multiplicative fusion [7], and fuzzy logic image fusion [17], as candidates for the image fusion step of the proposed detection scheme. The fuzzy logic scheme has proven valuable, since it adaptively fuses different parts of the image, depending on the relative presence of the target and clutter. However, as the fuzzy logic fusion approach is implemented through a fuzzy inference system that formulates a mapping from two inputs to one output, its use is restricted to a set of no more than two enhanced images. In general the number of images acquired from different viewing angles and for different polarizations can, however, exceed this limit. To overcome this limitation, we implement the two-stage fuzzy fusion proposed in [33], which fuses the outputs of the additive and multiplicative fusion processes. It is noted that, prior to image fusion, the segmented images are first normalized to ensure that they all have the same dynamic range.

In the two-stage image fusion, arithmetic fusion methods for TWRI proposed in [7], [32] are first applied. The images are fused through the additive and multiplicative fusion methods, given as

$$\hat{I}_A(m,n) = \sum_{j=1}^J \check{I}_j(m,n) \quad (21)$$

$$\hat{I}_M(m,n) = \prod_{j=1}^J \check{I}_j(m,n) \quad (22)$$

where $\check{I}_j(m,n)$ is the j th normalized segmented image and $\hat{I}_A(m,n)$ and $\hat{I}_M(m,n)$ are the final images resulting from the additive and multiplicative fusion, respectively. Then, a second fusion stage that exploits the capabilities of both additive and multiplicative fusion is applied, where the fuzzy logic approach is used to fuse the outputs of the arithmetic fusion. The readers are referred to [17] for a detailed description of the fuzzy-logic-based fusion approach.

V. PERFORMANCE EVALUATION

We evaluate both the LRT detector and the proposed detection method using real 2D polarimetric and multi-view images collected at the Radar Imaging Lab, Center for Advanced Communications, Villanova University. The techniques are also applied to numerical EM modeling data provided by the U.S. Army Research Lab. We first compare the performance of the image segmentation methods in terms of isolation of target regions in each individual image. The LRT detector with $J = 1$ is also applied as an image segmentation method to each individual image for comparison. The binary images generated by both image segmentation and the LRT detector are used as a mask on the original image to produce a corresponding enhanced image.

Next, we compare the performance of the various image fusion techniques to determine their suitability for the proposed detection method. It is noted that image fusion is performed only on the images that were enhanced by the image segmentation techniques in the previous step. Fusion of the images produced by the LRT detector is not performed. Instead, the output of the LRT detector with $J > 1$ was used for comparison. In other words the various combinations of the image segmentation and fusion techniques for the proposed detection method are compared with the LRT detector.

We note that the frequency bands of operation considered in these examples fall within the 0.5 to 3.5 GHz range. This frequency range is most amenable to signal propagation through various wall types [1]. Furthermore, all radar images presented in this section, other than the binary ones, are plotted on a 35 dB log scale, and the vertical and the horizontal axes represent the downrange and cross-range, respectively, with units in meters. The FAR for the LRT detector is fixed at 2.5%.

A. Experimental Setup

Two real data scenes and one numerical EM modeling scene are considered. The first real data scene consists of calibrated targets, with data acquired from a single viewpoint using multiple polarizations. The second real data scene is a populated scene with both calibrated targets and objects typically found in an office. The data corresponding to the populated scene are acquired from multiple views with a single polarization. The numerical EM modeling data consist of both multi-view and multi-polarization images.

1) *Calibrated Scene*: Both copolarization (HH and VV) and cross-polarization (HV and VH) data sets were collected from a calibrated scene containing a sphere, a top hat, a vertical dihedral, two dihedrals rotated at 22.5 and 45 deg, respectively, and three trihedrals, all placed at different downrange, cross-range, and elevations, as shown in Fig. 4. For

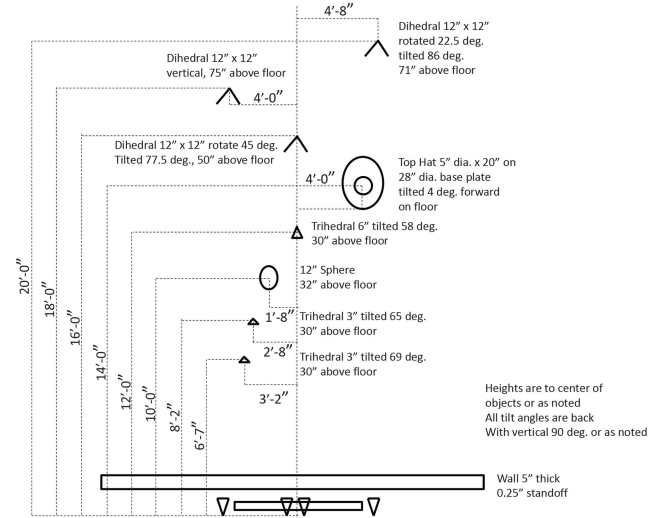


Fig. 4. Calibrated scene showing scene imaged through nonhomogenous plywood and gypsum wall (left), and schematic diagram of scene (right).

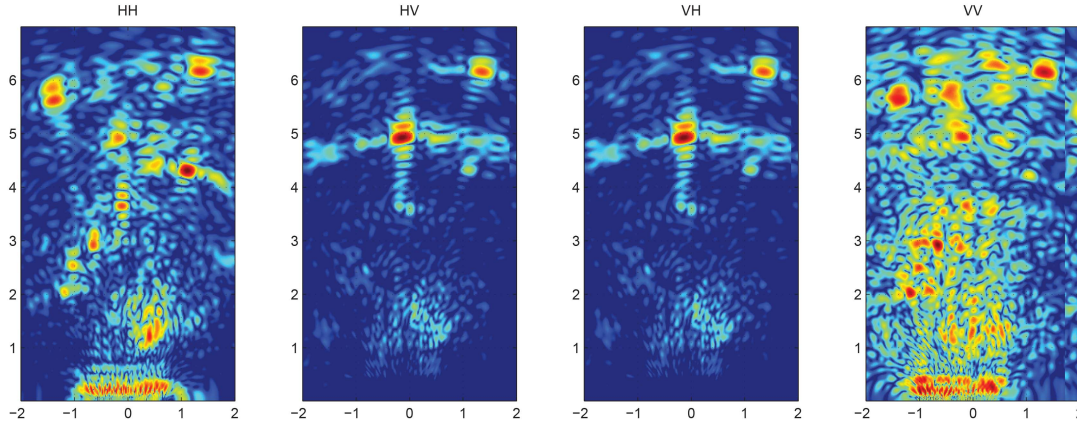


Fig. 5. Images acquired from calibrated scene.

each polarization setting the scene was imaged with a 1 GHz bandwidth stepped-frequency signal centered at 2.5 GHz. Two horn antennas, model H-1479 by BAE Systems, were mounted side-by-side on a field probe scanner, one oriented for horizontal polarization and the other for vertical polarization. The setup was used to synthesize a 57-element linear array with an inter-element spacing of 22 mm. The transmit power was set to 5 dBm. Data were collected through a 127-mm thick nonhomogeneous plywood and gypsum board wall, positioned 1.27 cm in downrange from the front face of the antennas, as shown in Fig. 4. More detailed information about the experimental setup is provided in [34], and the electrical properties of the wall material are described in [35]. The four polarimetric images acquired from the scene are shown in Fig. 5. We note that because of the nonhomogeneous nature of the wall, the correction for the wall effects in the beamforming process has not been applied. It can be observed from Fig. 5 that only two targets are detected in the HV and VH images. This is because the rotated dihedrals

produce a stronger cross-polarization return. Since the HV and VH images are almost identical, the performance of the various methods is evaluated only for the HH, HV, and VV images of the calibrated scene.

2) *Populated Scene*: Multi-view vertical polarization data sets were also collected from a populated scene, containing a vertical dihedron, a sphere, a table with metal legs, and a chair, each placed at different downrange, cross-range, and height, as shown in Fig. 6. A stepped-frequency signal, consisting of 801 frequency steps of size 3 MHz, covering the 0.7–3.1 GHz band, was used for data collection. The transmit power was set to 5 dBm. A quad-ridge horn antenna, model ETS-Lindgren 3164-04, was used as the transceiver and was mounted on a field probe scanner to synthesize a 57-by-57 element planar array with an inter-element spacing of 22 mm. Imaging was performed through a 140-mm thick solid concrete block wall from two vantage points, namely the front and the side views. The reader is referred to [7] for more details of the

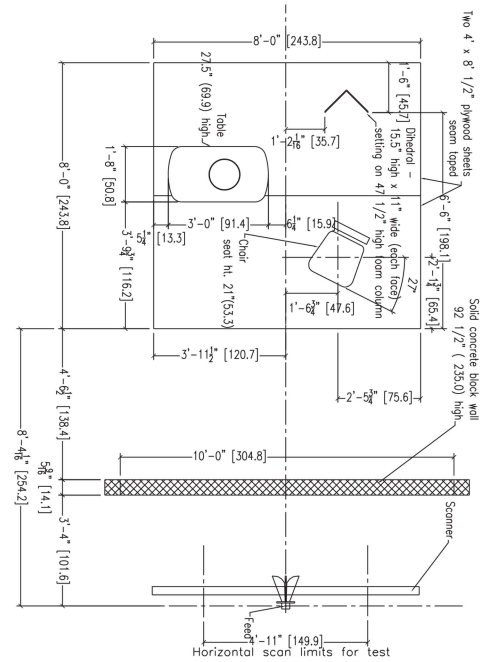


Fig. 6. Populated scene showing scene imaged through homogenous concrete wall from two vantage points (left) and schematic diagram of scene (right).

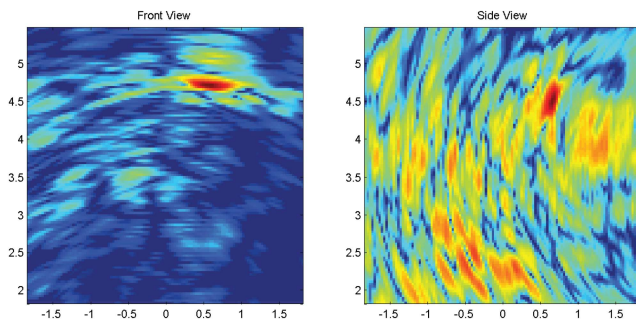


Fig. 7. Registered input images obtained from dihedral's elevation.

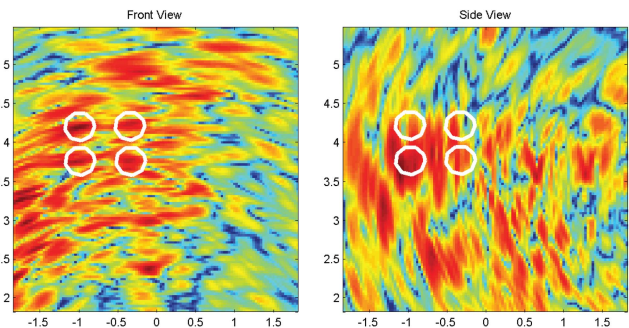


Fig. 8. Registered input images obtained from table's elevation.

experimental setup and to [35] for a description of the electrical properties of the wall material. The data were processed to produce B-scan images of size 117×117 , corresponding to the front and side views, at the heights of the dihedral and the table. Since the wall is homogeneous, the beamforming process accounted for the wall effects on signal propagation. Figure 7 shows the registered input images from the front and side views, at the dihedral's elevation. The dihedral elevation represents an example of a scene with high signal-to-clutter ratio. Figure 8 shows the images corresponding to the table elevation, which represent the case of low signal-to-clutter ratio. Although the metal legs of the table are present in the images (indicated by white circles), it is difficult to discern the target presence without any prior knowledge of the scene.

3) *EM Modeling of Complex Scene:* A complex room, constructed using a 200-mm thick exterior brick wall with glass windows and a wooden door, and having an interior room, with a 50-mm thick

sheetrock wall and a door, was simulated using numerical EM modeling software. In addition to wooden furniture, namely a bed with a generic fabric mattress, a couch with generic fabric cushions, a bookshelf, a dresser, and a table with four chairs, four humans were also present at different positions in the scene. Both copolarization (VV) and cross-polarization (HV) data sets were collected from the left and bottom views. For more details on EM modeling, the reader is referred to [36]. The schematic of the complex scene is shown in Fig. 9 and the four input images corresponding to the scene are provided in Fig. 10. No correction for the wall effects was applied during the beamforming process. The complex scene is an example that consists of targets with both high and low signal-to-clutter ratio. For instance, the exterior walls have high returns that overwhelm the returns from the interior walls and humans. Hence, the exterior walls have high signal-to-clutter ratio, while the interior walls and humans have low signal-to-clutter ratio.

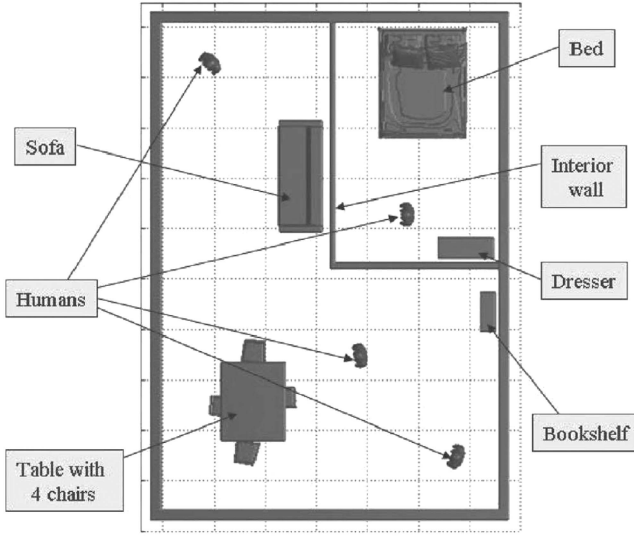


Fig. 9. Schematic of complex scene.

B. Performance Measure

The resulting images corresponding to the aforementioned scenes are assessed both qualitatively and quantitatively. Visual inspection is used to assess how well the targets are maintained in the image with respect to the ground truth. The image enhancements, in terms of clutter suppression, are assessed using the improvement factor in the target-to-clutter ratio, denoted as IF. Let $\mathcal{P}_{\mathcal{R}, \check{I}}$ denote the average power of region \mathcal{R} in image \check{I} . The IF is given as

$$\text{IF} = 10 \log_{10} \left[\frac{\mathcal{P}_{\mathcal{R}_t, \check{I}_e} \times \mathcal{P}_{\mathcal{R}_c, \check{I}_i}}{\mathcal{P}_{\mathcal{R}_t, \check{I}_i} \times \mathcal{P}_{\mathcal{R}_c, \check{I}_e}} \right] \quad (23)$$

where \check{I}_i is the input image and \check{I}_e is the enhanced image. $\mathcal{P}_{\mathcal{R}_t, \check{I}_m}$ can be expressed as

$$\mathcal{P}_{\mathcal{R}_t, \check{I}_m} = \frac{1}{N_{\mathcal{R}_t}} \sum_{x_q \in \mathcal{R}_t} (\check{I}_m(x_q))^2 \quad (24)$$

where $\check{I}_m(\mathbf{x}_q)$ is the q th pixel of region \mathcal{R}_t in image \check{I}_m , and $N_{\mathcal{R}_t}$ and $N_{\mathcal{R}_c}$ are the number of pixels in the target region \mathcal{R}_t and clutter region \mathcal{R}_c , respectively. The predefined target and clutter regions for each of the scenes are provided in Figs. 11–14. It should be noted that the target mask for the EM modeling, complex scene, as shown in Fig. 14, only includes

the targets of interest, which are the room layout and human targets. We consider furniture reflections as unwanted returns, and accordingly, they are treated as clutter.

C. Image Segmentation

In this section, the performance of the image segmentation techniques and LRT method with $J = 1$ are compared. For each individual image the respective methods are used to obtain a binary mask that depicts the target locations. After obtaining an enhanced image by applying the binary mask to the original input image, the IF is calculated.

1) *Calibrated Scene*: The results of applying the image segmentation methods to the images of the calibrated scene in Fig. 5 are presented in Fig. 15. It can be observed that the LRT detection method applied to the individual images does not yield image enhancements. While the targets are retained, the noise and clutter present in the original images persist even after application of the LRT detector. As for the BCV thresholding, the clutter is reduced in comparison with that of the LRT detector. It is also observed that the K-means clustering method produces a similar result to that of the BCV thresholding. Although there are some missed detections, visually, the entropy-based segmentation outperforms all the other methods by producing images with low clutter levels.

The IFs for the results obtained from the calibrated scene are provided in Table I. With the exception of the cross-polarization case, where the BCV thresholding and the K-means clustering removed all the clutter to obtain the highest IF, it is evident that the entropy-based segmentation generally outperforms the other methods by producing enhanced images with high IF, albeit with some missed detections.

2) *Populated Scene*: Figs. 16 and 17 show the results of applying the image segmentation methods to the populated scene images in Figs. 7 and 8, respectively. At the dihedral level, it can be observed from Fig. 16 that the BCV thresholding and K-means clustering have the best performance in terms of enhancing the images from the front view. Although the entropy-based segmentation and LRT detector are able to detect the dihedral, the clutter levels are higher in comparison. However, both BCV thresholding

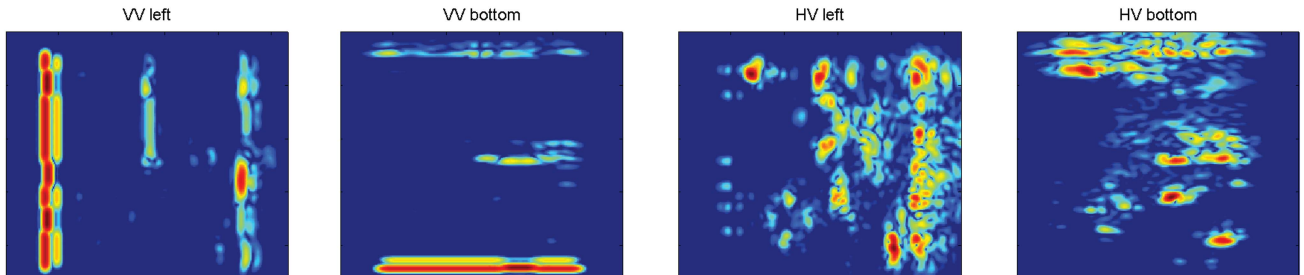


Fig. 10. Images of complex scene produced by numerical EM modeling.

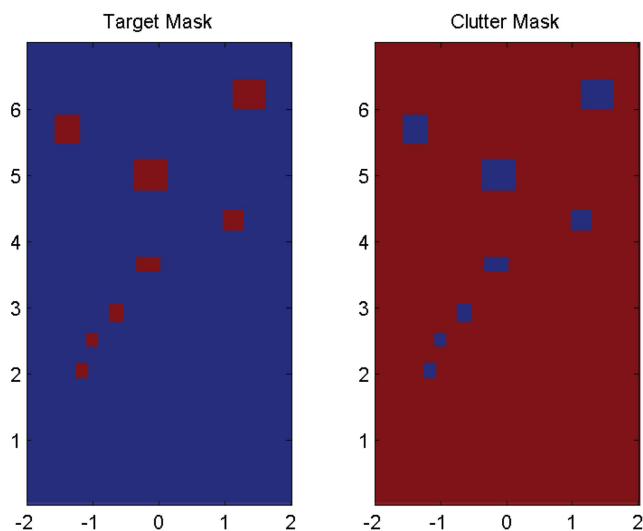


Fig. 11. Target and clutter mask of calibrated scene.

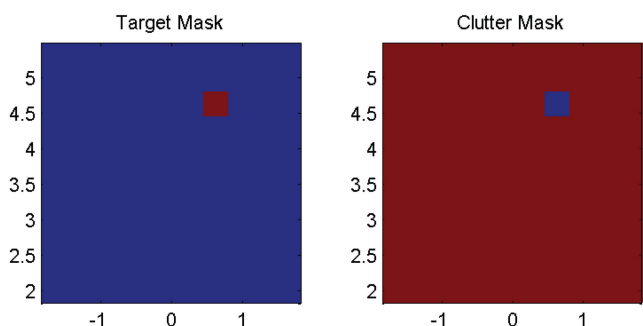


Fig. 12. Target and clutter mask of populated scene at dihedral's elevation.

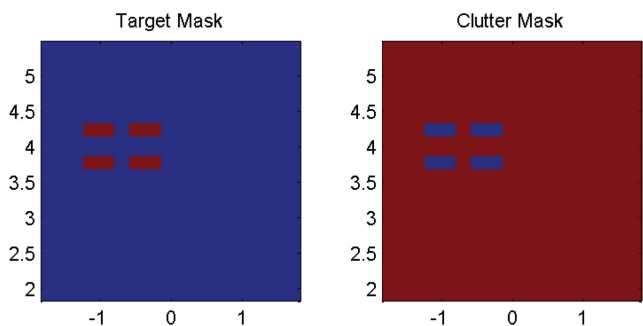


Fig. 13. Target and clutter mask of populated scene at table's elevation.

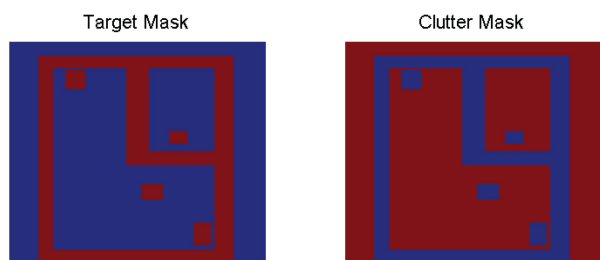


Fig. 14. Target and clutter mask of complex scene.

TABLE I

Improvement Factor in Target-to-Clutter Ratio of Calibrated Scene Images after Enhancements through Target Detection

	HH	HV	VV
LRT Detector 2.5% FAR	1.962	2.5124	3.0411
BCV Thresholding	2.5501	17.7063	2.1898
Entropy-Based Segmentation	4.3516	16.2246	6.3979
K-Means Clustering	2.571	17.7063	2.2344

TABLE II

Improvement Factor in Target-to-Clutter Ratio of Populated Scene Images after Enhancements through Target Detection (Dihedral Level)

Dihedral Level	VV Front View	VV Side View
LRT Detector 2.5% FAR	4.4331	8.6943
BCV Thresholding	8.2646	1.4556
Entropy-Based Segmentation	6.7664	11.2706
K-Means Clustering	8.2162	1.5151

TABLE III

Improvement Factor in Target-to-Clutter Ratio of Populated Scene Images after Enhancements through Target Detection (Table Level)

Table Level	VV Front View	VV Side View
LRT Detector 2.5% FAR	5.6081	8.5922
BCV Thresholding	1.3736	1.3044
Entropy-Based Segmentation	6.1886	3.0734
K-Means Clustering	1.3682	1.3237

and K-means clustering do not perform well in enhancing the images acquired from the side view. Although the LRT detector also does not perform well, comparatively less clutter is retained. Thus, the LRT produces an output image with a higher IF than the BCV and K-means for the side view. It is observed that the entropy-based segmentation outperforms all the other methods in the side view image by successfully detecting the dihedral and producing an output image with the least amount of clutter. The corresponding IFs are tabulated in Table II.

Figure 17 shows the image enhancement results at the table's elevation. It can be observed that all the methods under investigation are able to maintain the targets, except for the LRT detector that fails to retain all four legs of the table (depicted in white circles). While two targets are detected by the LRT detector in the front view image, it detects only one target in the side view image. Amongst the image segmentation techniques that are able to maintain the four targets, it can be observed from Table III that the entropy-based segmentation produces enhanced images with the highest IF. This is because the entropy produces an image with less false detections than those of the BCV and K-means.

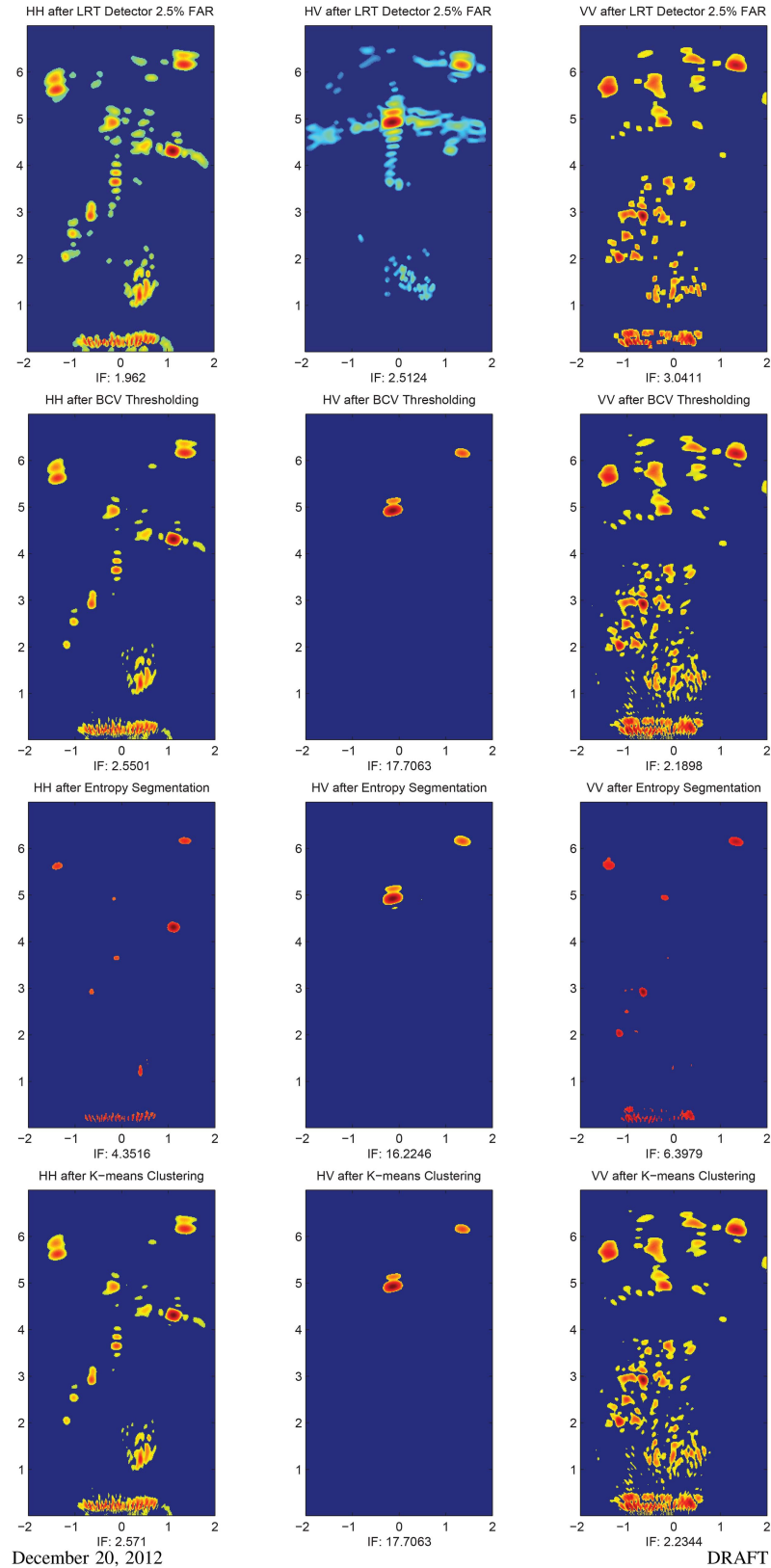


Fig. 15. Target detection in calibrated scene: from top to bottom, statistical LRT detector, BCV thresholding, entropy-based segmentation, and K-means clustering.

3) *EM Modeling of Complex Scene:* The results of applying the image segmentation methods to the complex scene images (Fig. 10) are presented in

Fig. 18. From the image enhancement results shown in Fig. 18, the LRT detection method produces the best result in detecting both exterior and interior walls

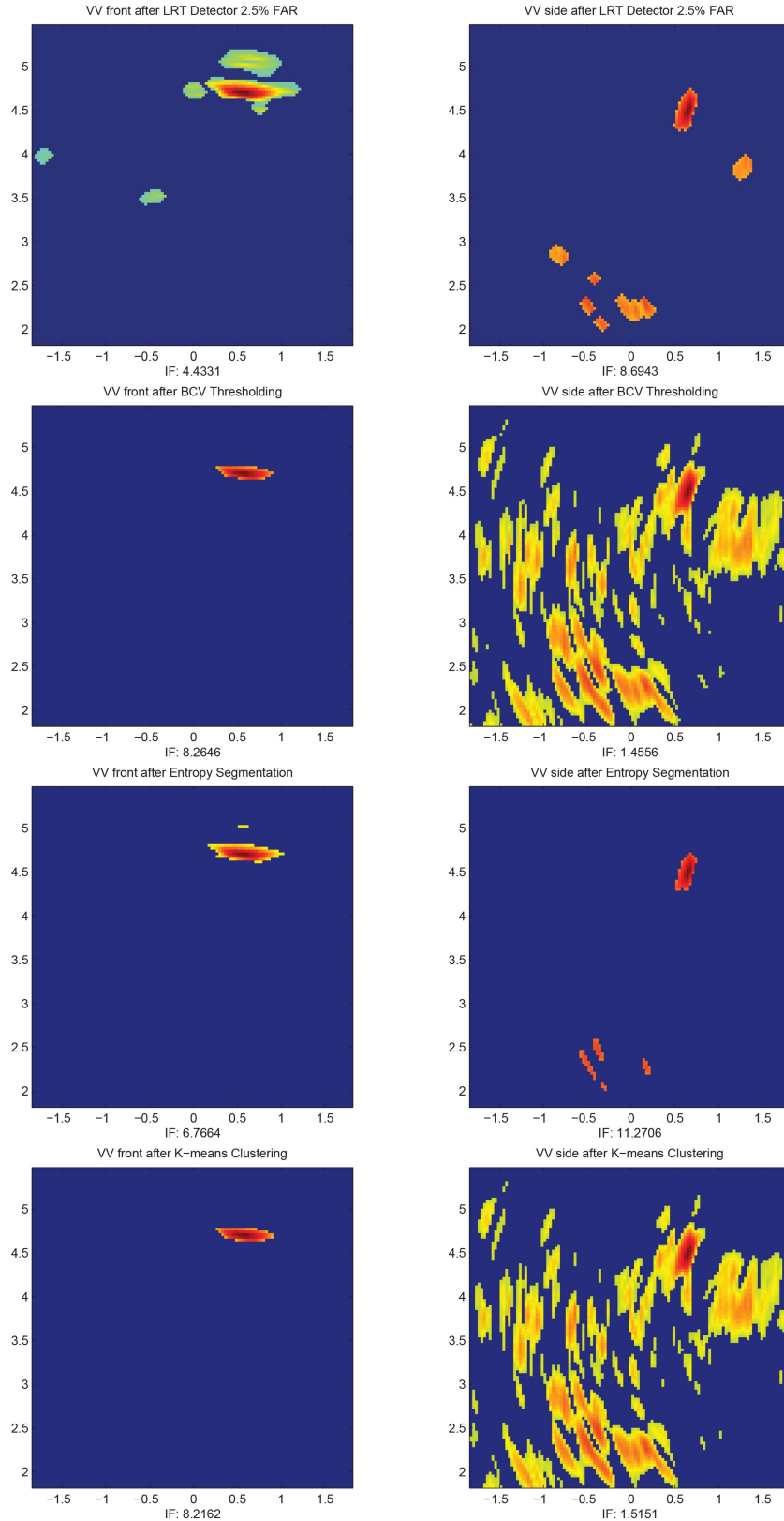


Fig. 16. Image enhancement results from populated scene (dihedral level).

in the copolarization (VV) images. However, the clutter from the original images is also maintained. The entropy-based segmentation manages to detect the interior walls with less clutter in the bottom view of

the copolarization image. However, it fails to detect the interior walls from the left view image. Although the BCV thresholding and K-means clustering have higher IFs, as shown in Table IV, both methods also

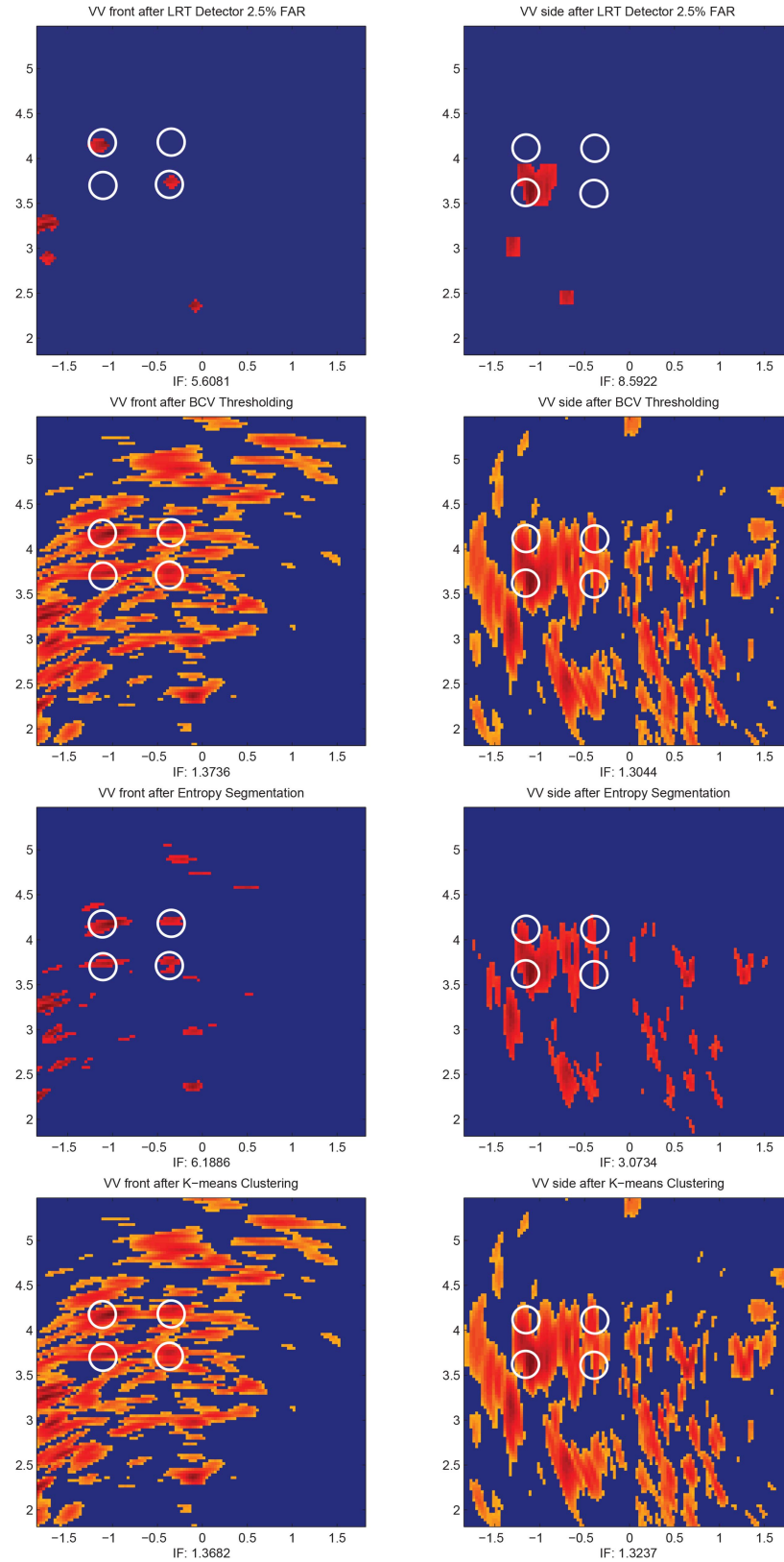


Fig. 17. Image enhancement results from populated scene (table level).

remove the interior walls and only the exterior walls are detected in the copolarization images.

For the cross-polarization (HV) images, the LRT detector does not perform well as there are missed

detections of the human targets. It is observed that the LRT detector fails to detect the targets in the left view image, while only one target is detected in the bottom view image. On the other hand the

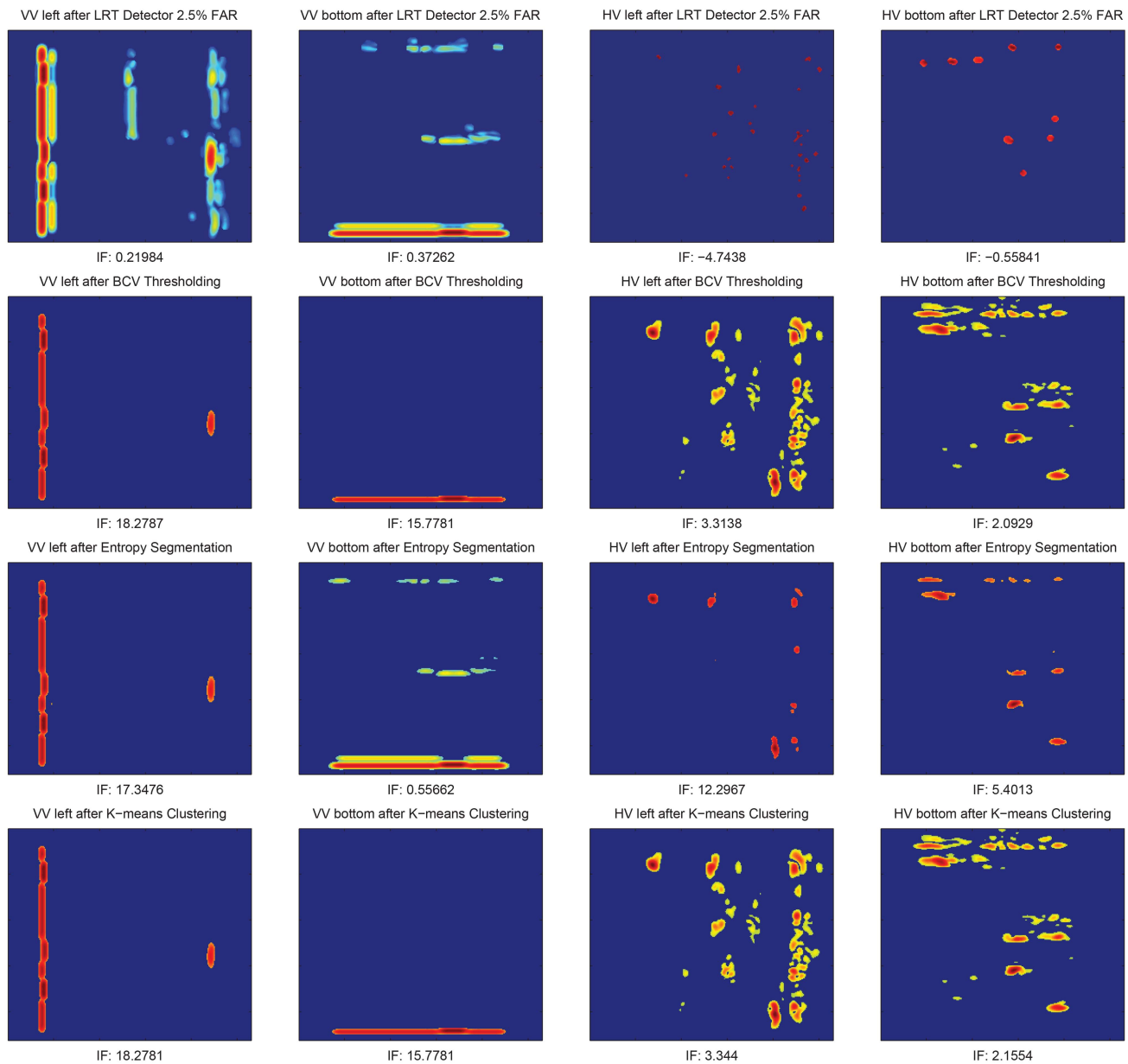


Fig. 18. Image enhancement results from complex scene.

TABLE IV
Improvement Factor in Target-to-Clutter Ratio of Complex Scene Images after Enhancements through Target Detection

	VV Left	VV Bottom	HV Left	HV Bottom
LRT Detector 2.5% FAR	0.2198	0.3726	-4.7438	-0.5584
BCV Thresholding	18.2787	15.7781	3.3138	2.0929
Entropy-Based Segmentation	17.3476	0.5566	12.2967	5.4013
K-Means Clustering	18.2781	15.7781	3.344	2.1554

BCV thresholding, entropy-based segmentation, and K-means clustering produce good detection of the human targets from the bottom view, with the entropy having the advantage of lower clutter levels. This is validated from Table IV, which shows that the entropy-based segmentation has the highest IF. However, the entropy segmentation has two missed detections in the left view image.

D. Fusion of Enhanced Images

After segmenting the individual images, the enhanced images are normalized and fused to compare the detection results with the LRT detector that incorporates the multi-view and multi-polarization information. Here, pixel-wise additive fusion, multiplicative fusion, and two-stage fuzzy logic fusion are used to fuse the enhanced images.

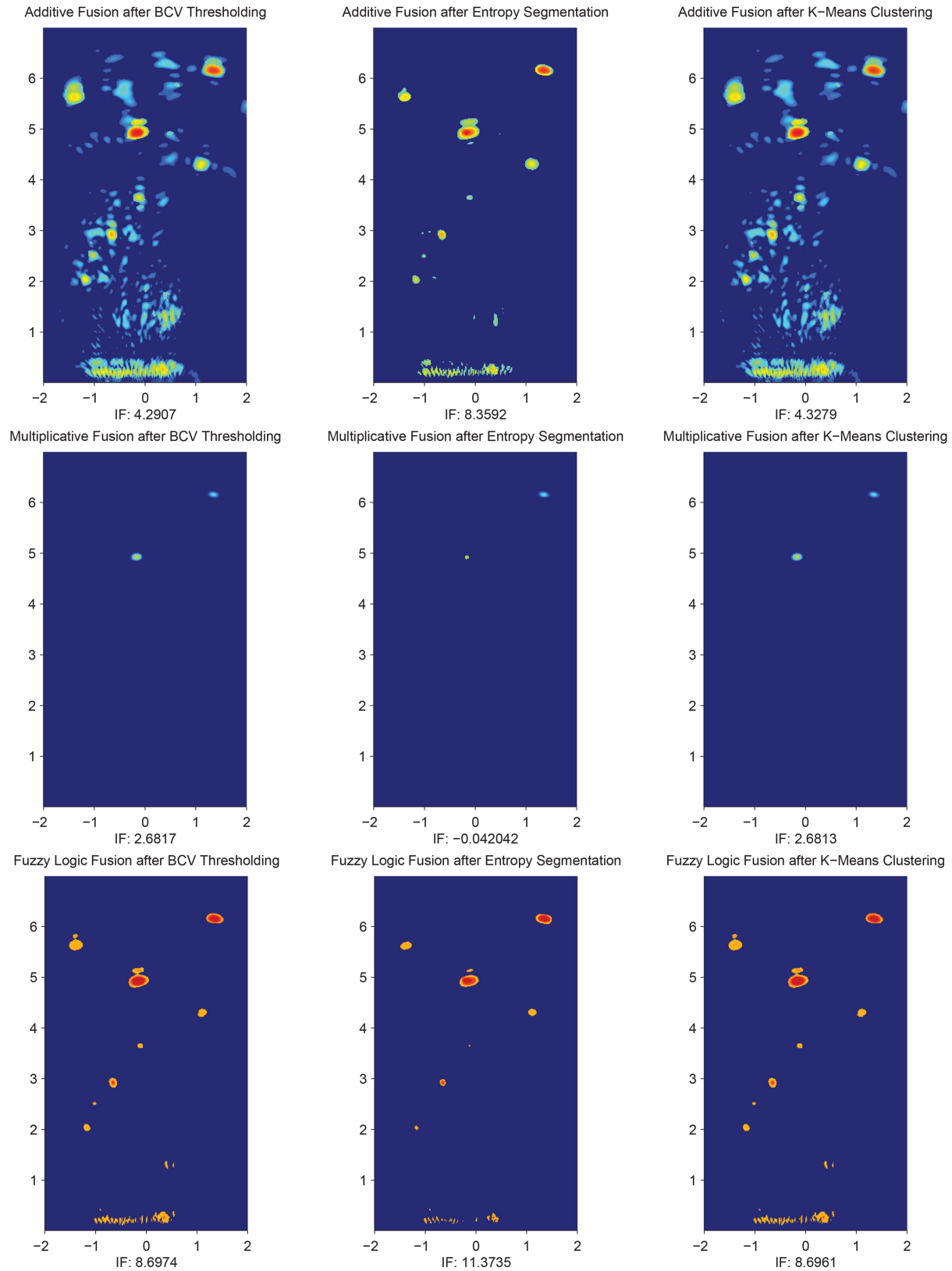


Fig. 19. Image fusion of segmented images from calibrated scene.

Through empirical observations of the three experimental data sets, the radar images are separated into four regions for the fuzzy logic fusion. The pixel intensities, ranging from 0 to 105 for Region 1, 105

to 165 for Region 2, 165 to 225 for Region 3, and the remaining 225 to 255 for Region 4, are used to form the membership functions. As for the fuzzy rules, it is generally observed that target objects tend to have

TABLE V
Fuzzy Rules

Input 1	Input 2	Output
Region 1	Region 1	Region 1
Region 2	Region 2	Region 2
Region 3	Region 3	Region 3
Region 4	Any	Region 4
Region 2	Region 1	Region 2
Region 3	Region 1	Region 2

TABLE VI

Improvement Factor in Target-to-Clutter Ratio of Fused Calibrated Scene Images

	Additive Fusion	Multiplicative Fusion	Fuzzy Logic Fusion
BCV Thresholding	4.2907	2.6817	8.6974
Entropy-Based Segmentation	8.3592	-0.0420	11.3735
K-Means Clustering	4.3279	2.6813	8.6961

high pixel intensities (Region 4) and that background noise has low pixel intensities (Region 1). Thus, input pixels with higher intensities are preferred. For instance, if an input pixel is labeled as Region 4, the output pixel will be set to Region 4 automatically. When input pixels are from different regions, such as from Region 1 and Region 3, the average value from both regions is set as the output value. When the input pixels originate from Region 1 and Region 2, the output values could be suppressed to Region 1, as in the case of [17]. However, in our experiments, we maintain the output value as Region 2 to avoid oversuppression of information. The defined set of nonoverlapping logical rules is shown in Table V.

1) *Calibrated Scene*: Fig. 19 shows the fusion results of the enhanced images from the calibrated scene (Fig. 15). We observe that the additive fusion of the enhanced images maintains all the targets, clutter, and false detections. Among the three combinations of the image segmentation methods with additive fusion, the entropy-based additive fusion produces the best result with the detection of all eight targets and an IF of 8.36 dB, as shown in Table VI. The multiplicative fusion fails to detect all eight targets because only the rotated dihedrals, which are mutual to both copolarization and cross-polarization images, survive the fusion process while the remaining targets are suppressed. The two-stage fuzzy fusion, which fuses the outputs of the additive and multiplicative fusion, produces a balanced output with high targets and low clutter levels. Between the additive and fuzzy logic fusion that maintains most of the targets, it can be observed that the entropy-segmented image, coupled with the two-stage fuzzy fusion, produces the best result with an IF of 11.37 dB.

Multi Polarization LRT Detection 2.5% FAR

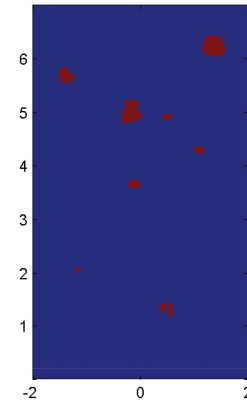


Fig. 20. Multi-polarization statistical LRT detection on calibrated scene images.

In comparison with the LRT detector that incorporates the multi-polarization signatures to produce the output shown in Fig. 20, the fuzzy fused composite image based on entropy segmentation outperforms the multi-polarization LRT detector by detecting seven targets versus the LRT detector's six. The false detections in the image produced by the LRT detector are also more significant than the fuzzy logic fusion. Thus, the entropy-based segmentation coupled with the fuzzy logic fusion outperforms the other methods by producing an image with high target and low clutter levels.

2) *Populated Scene*: Similar fusion results to that of the calibrated scene are also observed for the populated scene. As can be observed from Fig. 21 for the high signal-to-clutter ratio scenario, although the additive fusion generally retains most of the noise and clutter present in the individual images, the corresponding composite image based on the entropy-segmented images produces an output image with minimal clutter. The image segmentation processes also help improve the multiplicative and fuzzy logic fusion results by suppressing all the clutter. As shown in Table VII, both multiplicative and fuzzy logic fusion produce composite images with similar IFs. Generally, it can be observed that all three fusion methods, employing the entropy-segmented images, produce an output with low clutter levels. Comparison with the multi-view LRT detector, whose output is shown in Fig. 22, reveals that the composite images produced by the proposed methods outperform the LRT detector. This is because the LRT detector fails to distinguish between targets and clutter in its detection mask.

In the low signal-to-clutter ratio case, it can be observed from Fig. 23 that the composite images produced by the multiplicative and fuzzy logic fusion based on the entropy-segmented images clearly depict the four legs of the metal table, even though there are some false detections close to the table. The LRT detector that incorporates the multi-view information

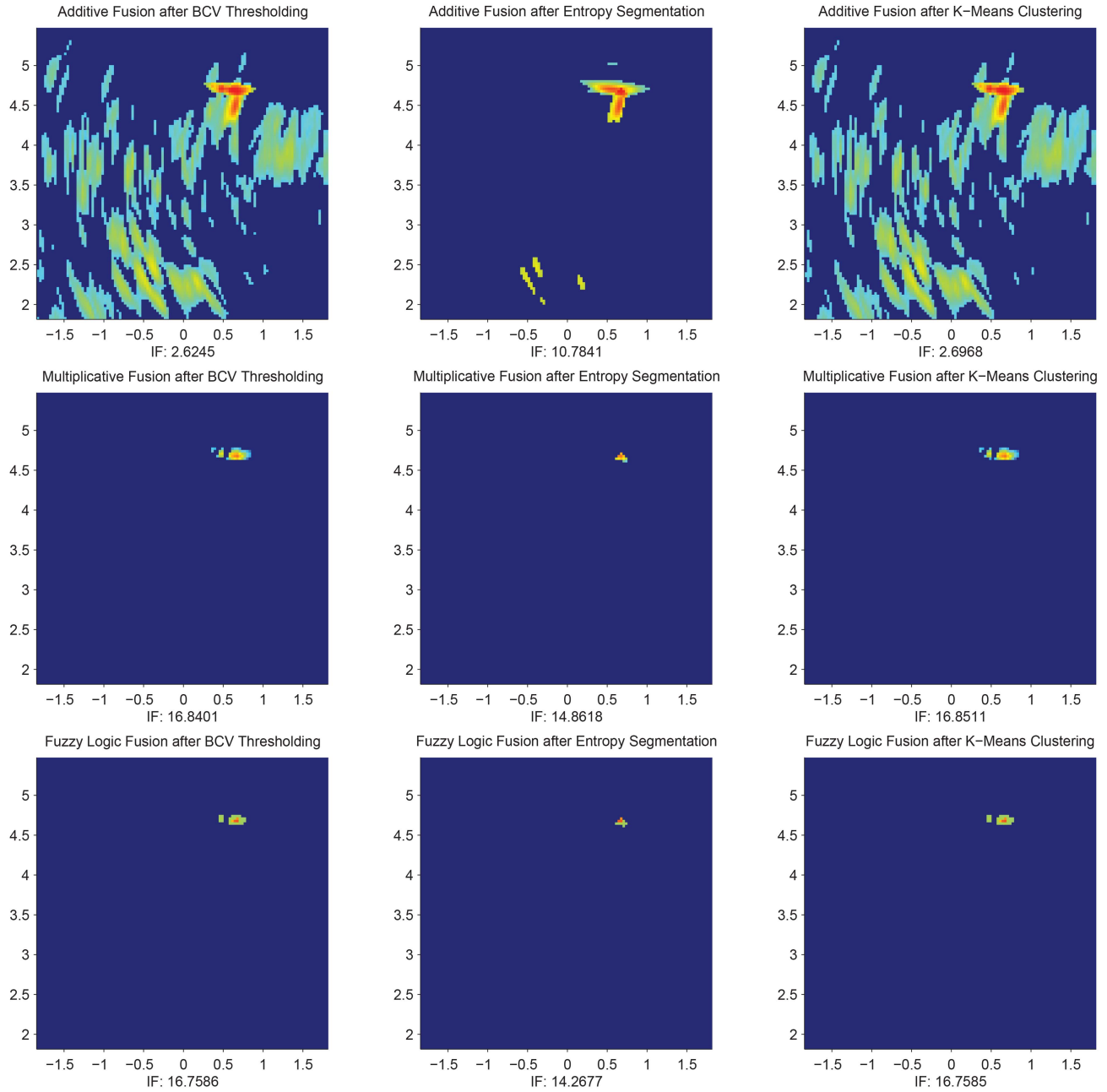


Fig. 21. Image fusion of segmented images from populated scene (dihedral level).

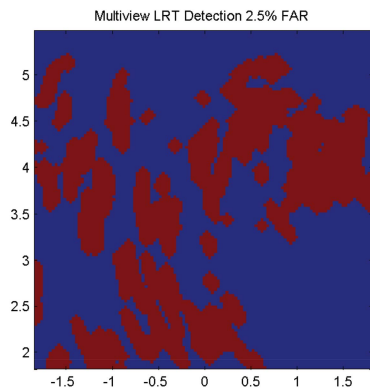


Fig. 22. Multi-view statistical LRT detection on populated scene images (dihedral level).

TABLE VII
Improvement Factor in Target-to-Clutter Ratio of Fused Populated Scene Images (Dihedral Level)

	Additive Fusion	Multiplicative Fusion	Fuzzy Logic Fusion
BCV Thresholding	2.6245	16.8401	16.7586
Entropy-Based Segmentation	10.7841	14.8618	14.2677
K-Means Clustering	2.6968	16.8511	16.7585

(see Fig. 24) misses detection of one of the table legs, but has fewer false alarms compared with that of the composite images. Table VIII shows that the fusion

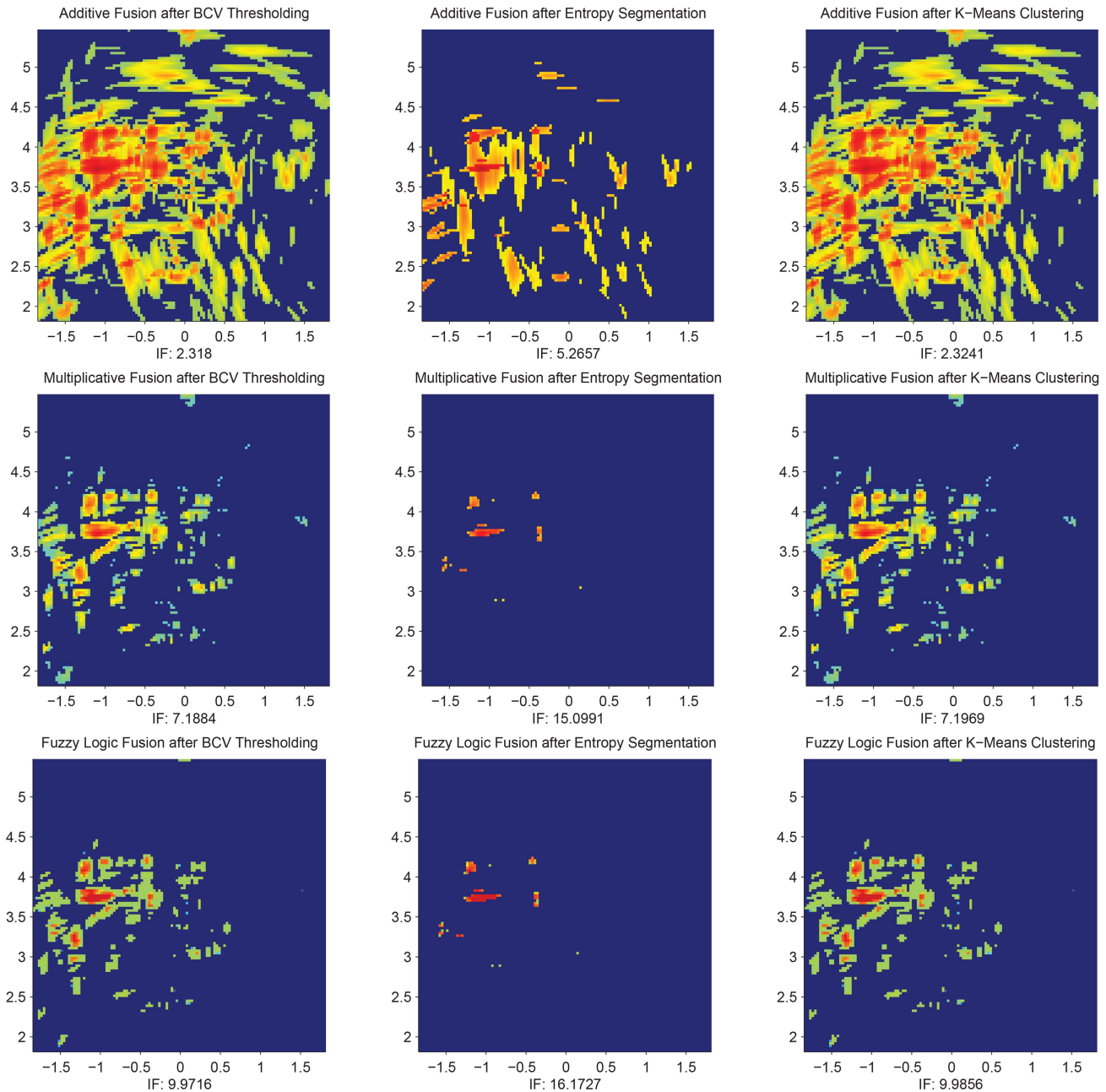


Fig. 23. Image fusion of segmented images from populated scene (table level).

TABLE VIII
Improvement Factor in Target-to-Clutter Ratio of Fused Populated Scene Images (Table Level)

	Additive Fusion	Multiplicative Fusion	Fuzzy Logic Fusion
BCV Thresholding	2.3180	7.1884	9.9716
Entropy-Based Segmentation	5.2657	15.0991	16.1727
K-Means Clustering	2.3241	7.1969	9.9856

of the entropy-segmented images produces composite images with the highest IF.

3) *EM Modeling of Complex Scene*: Figs. 25 and 26 show the image fusion results of the segmented

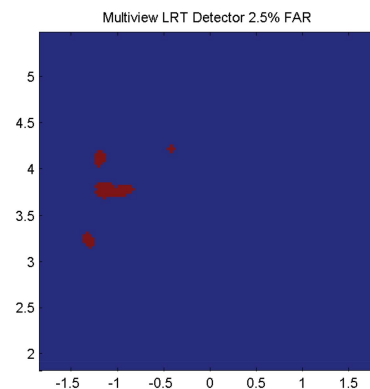


Fig. 24. Multi-view statistical LRT detection on populated scene images (table level).

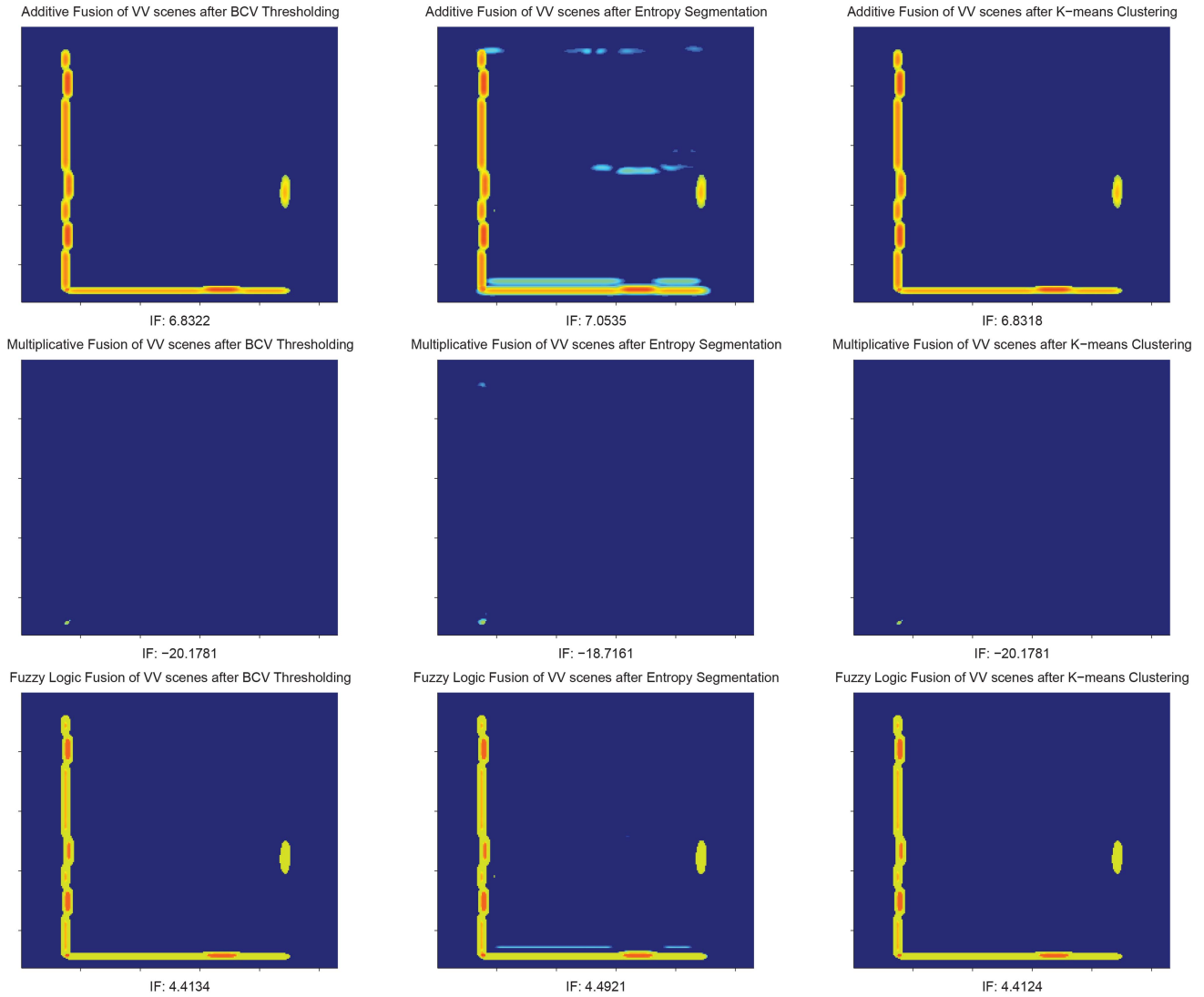


Fig. 25. Image fusion of segmented multi-view copolarization images from complex scene.

multi-view copolarization (VV) and cross-polarization (HV) images acquired from the complex scene, respectively. In the copolarization composite images, it is evident from Fig. 25 that the additive fusion based on the entropy-segmented images produces the best visual result by detecting both exterior and interior walls. It also has the highest IF, as tabulated in Table IX.

For the cross-polarization composite images (Fig. 26), the fuzzy logic fusion produces images that balance the output of the additive and multiplicative fusion. Although the IF is lower than that of the multiplicative, visually, the entropy-segmented image coupled with the fuzzy fusion produces a result with the least clutter. The additive fusion of the entropy-segmented images also produces a similar result, where the clutter level is lower compared with that of BCV thresholding and K-means clustering. When only considering the human targets, it is evident that, for the cross-polarization images, the

multiplicative fusion of the BCV thresholding and K-means clustering images produces the best result. However, it can be observed that the entropy-based segmentation, coupled with the fuzzy logic fusion, produces the best result that detects both human targets and room layout.

We observe from Fig. 27 that the multi-view LRT detector, when applied to the copolarization images, produces a similar output to that of additive fusion based on the entropy-segmented images. While the LRT detector manages to detect the four humans in the cross-polarization images, the clutter levels are higher than that of the multiplicative fusion with the image segmentation techniques.

Figures 28 and 29 show the image fusion results of the segmented polarimetric images from the left and bottom views of the complex scene, respectively. From the left view (Fig. 28), it can be observed that, while the additive fusion and fuzzy logic fusion coupled with BCV thresholding and K-means

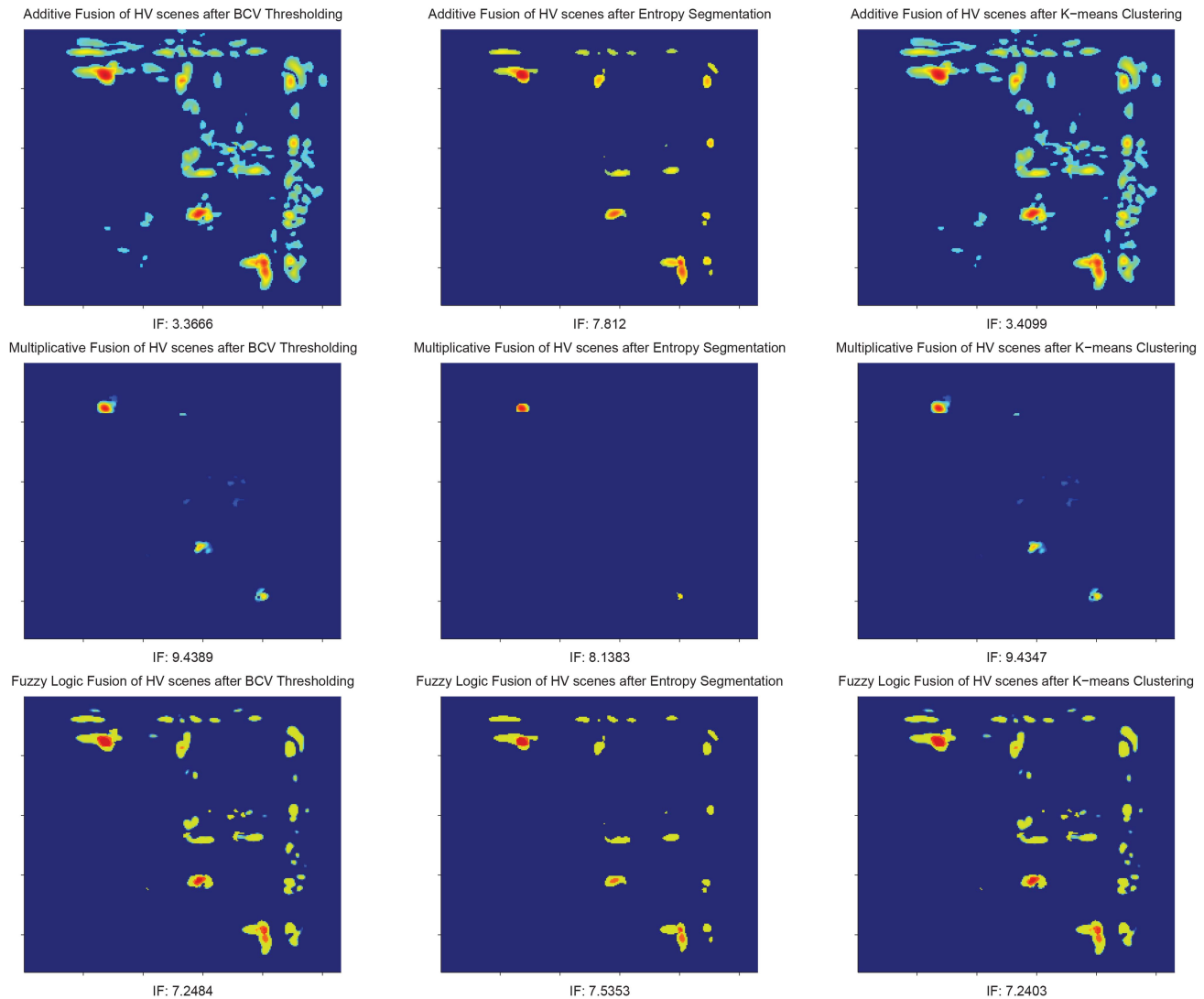


Fig. 26. Image fusion of segmented multi-view cross-polarization images from complex scene.

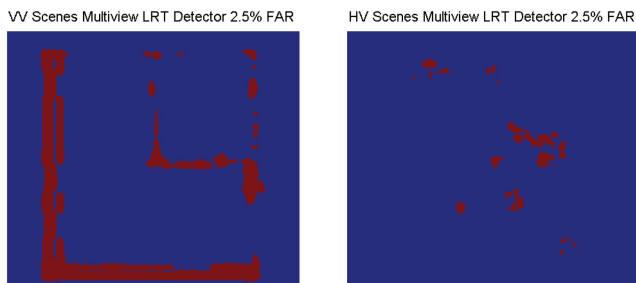


Fig. 27. Multi-view statistical LRT detection on copolarization and cross-polarization complex scene images.

clustering are able to maintain the targets, clutter levels are still quite high. Similar results are also observed for the bottom view (Fig. 29). Although the composite images based on the entropy-segmented images do not detect all the targets, they generally produce an image with high target and low clutter levels. The IFs, as tabulated in Table X, are also the highest. In both cases the multiplicative fusion fails to

TABLE IX
Improvement Factor in Target-to-Clutter Ratio of Fused Multiview Complex Scene Images

	Additive Fusion	Multiplicative Fusion	Fuzzy Logic Fusion
Copolarization (VV)			
BCV Thresholding	6.8322	−20.1781	4.4134
Entropy-Based Segmentation	7.0535	−18.7161	4.4921
K-Means Clustering	6.8318	−20.1781	4.4124
Copolarization (HV)			
BCV Thresholding	3.3666	9.4389	7.2484
Entropy-Based Segmentation	7.8120	8.1383	7.5353
K-Means Clustering	3.4099	9.4347	7.2403

detect the human targets and the room layout due to the suppression of nonmutual pixels in the enhanced images. Figure 30 shows that the multi-polarization LRT detector, which is evaluated on the polarimetric

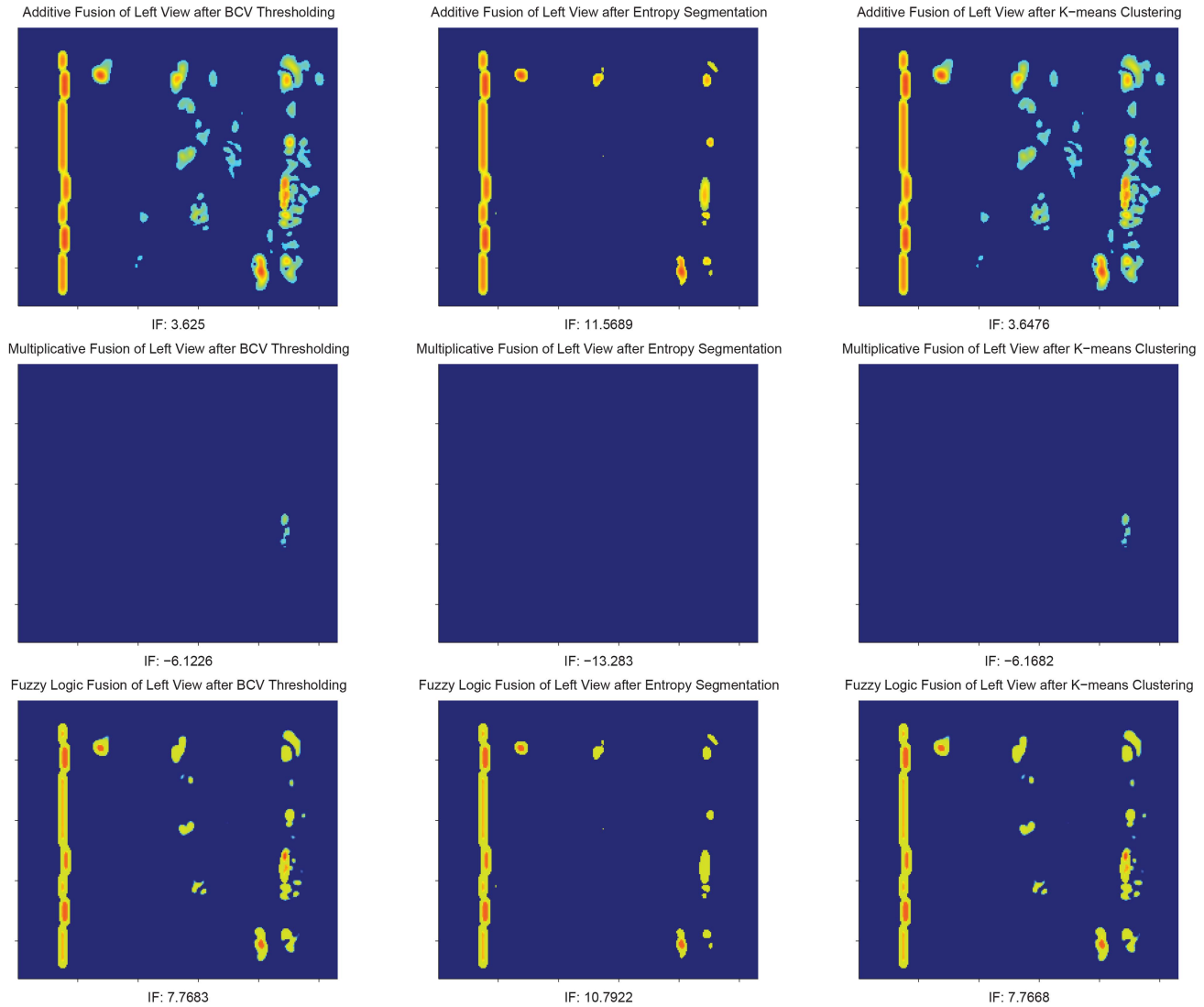


Fig. 28. Image fusion of segmented polarimetric images from left view of complex scene.

images both from the left and bottom views, does not perform well enough to detect the walls and human targets. In both cases high clutter levels are retained.

E. Discussions

From the experimental results it can be generally observed that the BCV thresholding and K-means clustering yield similar segmentation performances and that their differences in the IF are almost negligible. This is because the K-means clustering's minimization of the WCSS is equivalent to the BCV thresholding's maximization of the sum of class variances when there are only two regions [14]. It can also be observed that the BCV thresholding generally has a poorer performance than the entropy-based segmentation when the image has more clutter. This is because the sparsity of the high pixel values (target pixels) in a TWR image forces the BCV thresholding method to also include low pixel values (clutter pixels) in the target class, in order to obtain

a high class variance. Unlike the BCV method the entropy-based segmentation is less affected by the sparse target pixels in the image and is only dependent on the pixel value probability. That is, low pixel values will affect thresholding if they assume similar probabilities as those of the target pixels. In the case when the scene has less clutter, it is observed that a similar performance will be produced by both entropy-based segmentation and BCV thresholding methods.

Using the calibrated scene as an example, the corresponding copolarization images contain more clutter than the cross-polarization image. Hence, in the case of copolarization images, the entropy-based segmentation outperforms the BCV, which chooses a lower thresholding value. This is evident from Fig. 31, wherein the maximum class variances for both regions in the copolarization image are skewed towards the lower pixel regions. As a result the summation of both class variances produces a threshold value that

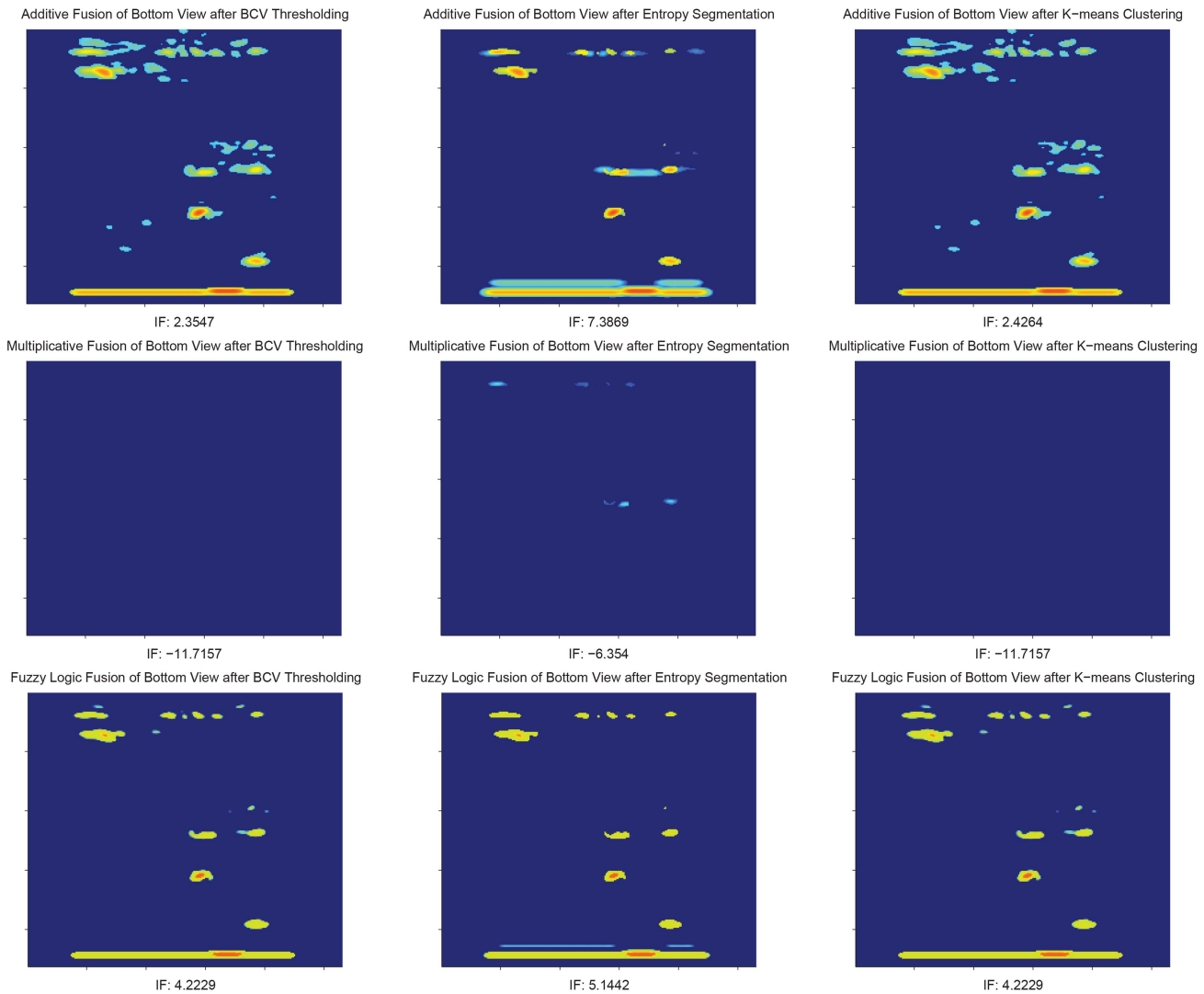


Fig. 29. Image fusion of segmented polarimetric images from bottom view of complex scene.

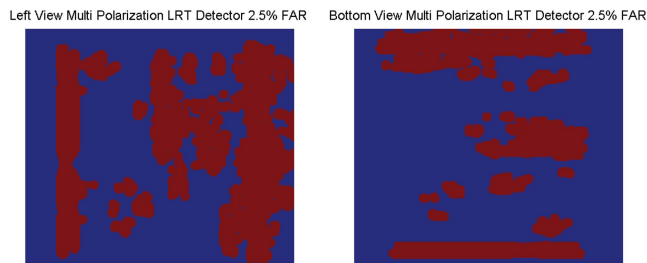


Fig. 30. Multi-polarization statistical LRT detection on copolarization and cross-polarization complex scene images.

is also located in the low pixel values region, causing the segmented image to include most of the clutter. As can be observed from Fig. 32, the respective class entropies are less affected by the skewness towards the low pixel values region. As the threshold value is increased, the entropy of one region will generally decrease, while the entropy of the other region will increase. Thus, the maximization of the sum of class entropies will produce a balanced threshold where

TABLE X
Improvement Factor in Target-to-Clutter Ratio of Fused Polarimetric Complex Scene Images

	Additive Fusion	Multiplicative Fusion	Fuzzy Logic Fusion
Left View			
BCV Thresholding	3.6250	-6.1226	7.7683
Entropy-Based Segmentation	11.5689	-13.2830	10.7922
K-Means Clustering	3.6476	-6.1682	7.7668
Bottom View			
BCV Thresholding	2.3547	-11.7157	4.2229
Entropy-Based Segmentation	7.3869	-6.3540	5.1442
K-Means Clustering	2.4264	-11.7157	4.2229

the information content in both regions is almost equal. For the case of a less cluttered image, such as the one acquired using cross-polarization, a higher threshold value is produced by the BCV method

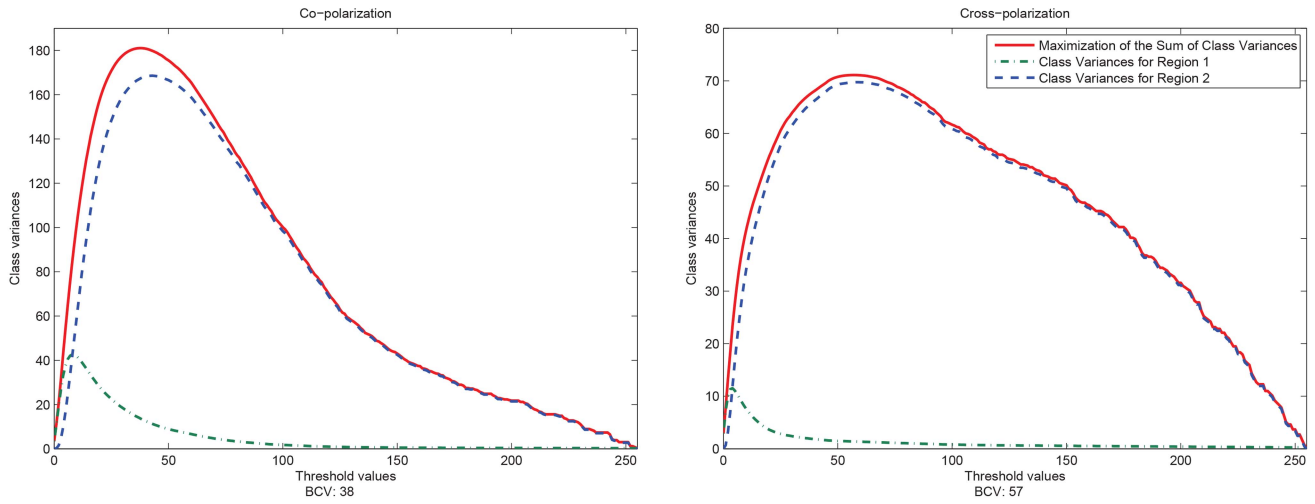


Fig. 31. Class variances with respect to threshold values for copolarization images (left) and cross-polarization image (right) of calibrated scene.

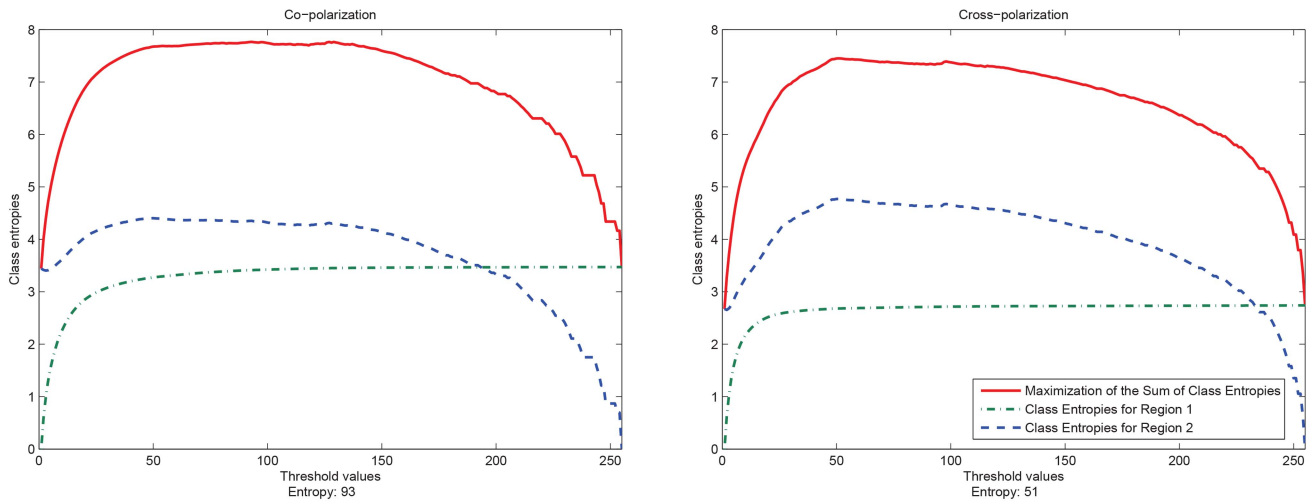


Fig. 32. Class entropies with respect to threshold values for copolarization images (left) and cross-polarization image (right) of calibrated scene.

since the maximum class variance for the second region is located at a much higher pixel value region. This dominance allows the summation of both class variances to produce a high threshold value.

The observations that 1) the entropy-based segmentation outperforms the BCV thresholding method when there is more clutter, and 2) both methods have similar performance when there is less clutter in the image, are further supported by Fig. 33. The histograms in the left column are produced from images with more clutter, and those in the right column are histograms of images with less clutter. The thresholds associated with the various segmentation methods are plotted in red, green, and black, respectively, for the entropy-based segmentation, BCV thresholding, and K-means clustering. As can be seen the threshold values corresponding to entropy are much higher in the case of high clutter and comparable with those of K-means and BCV thresholding in the case of low clutter.

In this paper, we have considered three different scenes of varying complexity in an attempt to have a broad representation of a variety of possible indoor scenes. In all cases the corresponding image histograms are uni-modal, as depicted in Fig. 33. Thus, it can be generally concluded that most TWR images have uni-modal histograms. In order to automate the target detection process for uni-modal images, an image segmentation method that is not affected by the uni-modality and sparsity of the targets needs to be taken into consideration. Through the experimental results it has been shown that the entropy-based segmentation, coupled with the fuzzy-logic-based image fusion, is the most effective and viable alternative to the LRT for target detection in TWR images.

VI. CONCLUSION

In this paper, we examined image processing approaches for target detection in TWRI. Image

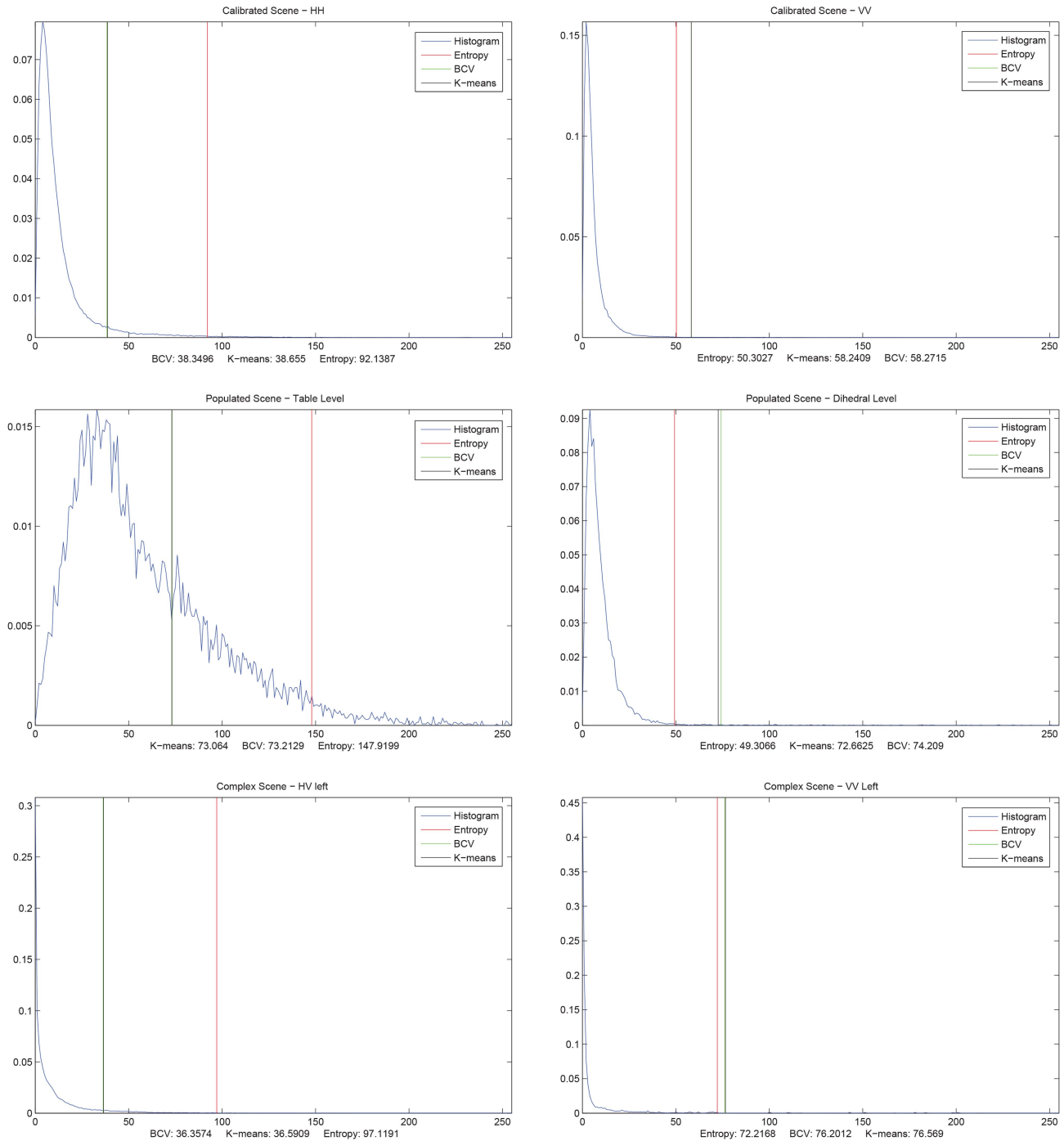


Fig. 33. Histogram of different images along with thresholds associated with the segmentation methods, showing calibrated scene (top row), populated scene (middle row), and complex scene (bottom row).

segmentation techniques were first applied to enhance the images by distinguishing between the target and clutter regions. Image fusion techniques were then used to fuse the enhanced images. Performances of various candidate segmentation and fusion techniques were evaluated using real 2D polarimetric and multi-view images, as well as numerical EM modeling data. The results showed that the entropy-based segmentation technique produced, in most of the cases considered, an output better than those produced by the other segmentation

schemes, as well as the LRT detector. Although there were some missed detections, the entropy-based segmentation consistently provided high target and low clutter levels. While the BCV thresholding and K-means clustering methods also maintained most of the targets, the clutter levels were much higher by comparison. For the fusion of multi-view and multi-polarization images, it was generally observed that the fuzzy logic fusion of the segmented images revealed important information about the indoor targets and walls.

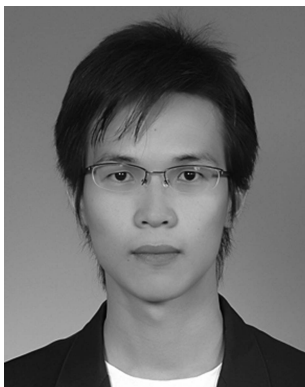
ACKNOWLEDGMENT

The authors would like to thank Dr. Traian Dogaru from the U.S. Army Research Lab for providing the numerical EM modeling data.

REFERENCES

- [1] Amin, M.
Through the Wall Radar Imaging. Boca Raton, FL: CRC Press, 2010.
- [2] Baranoski, E.
Through wall imaging: Historical perspective and future directions.
Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP 2008), Las Vegas, NV, Mar. 31–Apr. 4, 2008, pp. 5173–5176.
- [3] Gaugue, A. and Politano, J.
Overview of current technologies for through-the-wall-surveillance.
Proceedings of SPIE, 2005, pp. 1–11.
- [4] Yoon, Y., Amin, M., and Ahmad, F.
MVDR beamforming for through-the-wall radar imaging.
IEEE Transactions on Aerospace and Electronic Systems, **47**, 1 (2011), 347–366.
- [5] Lai, C-P. and Narayanan, R.
Ultrawideband random noise radar design for through-wall surveillance.
IEEE Transactions on Aerospace and Electronic Systems, **46**, 4 (2010), 1716–1730.
- [6] Ahmad, F., Amin, M., and Zemany, P.
Dual-frequency radars for target localization in urban sensing.
IEEE Transactions on Aerospace and Electronic Systems, **45**, 4 (2009), 1598–1609.
- [7] Ahmad, F. and Amin, M.
Multi-location wideband synthetic aperture imaging for urban sensing applications.
Journal of the Franklin Institute, **345**, 6 (2008), 618–639.
- [8] Debes, C., Zoubir, A., and Amin, M.
Target detection in multiple-viewing through-the-wall radar imaging.
Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2008), Boston, MA, July 7–11, 2008, pp. 1-173–1-176.
- [9] Debes, C., Amin, M., and Zoubir, A.
Target detection in single- and multiple-view through-the-wall radar imaging.
IEEE Transactions on Geoscience and Remote Sensing, **47**, 5 (2009), 1349–1361.
- [10] Debes, C., et al.
Adaptive target detection with application to through-the-wall radar imaging.
IEEE Transactions on Signal Processing, **58**, 11 (2010), 5572–5583.
- [11] Debes, C., Zoubir, A., and Amin, M.
Enhanced detection using target polarization signatures in through-the-wall radar imaging.
IEEE Transactions on Geoscience and Remote Sensing, **50**, 5 (2012), 1968–1979.
- [12] Sezgin, M. and Sankur, B.
Survey over image thresholding techniques and quantitative performance evaluation.
Journal of Electronic Imaging, **13**, 1 (2004), 146–165.
- [13] Lee, J-S. and Jurkevich, I.
Segmentation of SAR images.
IEEE Transactions on Geoscience and Remote Sensing, **27**, 6 (1989), 674–680.
- [14] Otsu, N.
A threshold selection method from gray-level histogram.
IEEE Transactions on Systems, Man and Cybernetics, **SMC-9**, 1 (1979), 62–66.
- [15] Kapur, J., Sahoo, P., and Wong, A.
A new method for gray-level picture thresholding using the entropy of the histogram.
Computer Vision, Graphics, and Image Processing, **29**, 3 (1985), 273–285.
- [16] Seber, G.
Multivariate Observations. Hoboken, NJ: Wiley, 1984.
- [17] Seng, C., et al.
Fuzzy logic-based image fusion for multi-view through-the-wall radar.
Proceedings of 2010 International Conference on Digital Image Computing: Techniques and Applications (DICTA), Sydney, Australia, Dec. 1–3, 2010, pp. 423–428.
- [18] Jansing, E., Albert, T., and Chenoweth, D.
Two-dimensional entropic segmentation.
Pattern Recognition Letters, **20**, 3 (1999), 329–336.
- [19] Wang, W., et al.
Knowledge-based bridge detection from SAR images.
Journal of Systems Engineering and Electronics, **20**, 5 (2009), 929–936.
- [20] Wang, W., et al.
Oil spill detection from polarimetric SAR image.
Proceedings of 2010 IEEE 10th International Conference on Signal Processing (ICSP), Beijing, China, Oct. 24–28, 2010, pp. 832–835.
- [21] Pasolli, E., Melgani, F., and Donelli, M.
Automatic analysis of GPR images: A pattern-recognition approach.
IEEE Transactions on Geoscience and Remote Sensing, **47**, 7 (2009), 2206–2217.
- [22] Sun, W., et al.
Research of unsupervised image change detection algorithm based on 2-D histogram.
Proceedings of 2010 IEEE 10th International Conference on Signal Processing (ICSP), Beijing, China, Oct. 24–28, 2010, pp. 686–689.
- [23] Bazi, Y., Melgani, F., and Al-Sharari, H.
Unsupervised change detection in multispectral remotely sensed imagery with level set methods.
IEEE Transactions on Geoscience and Remote Sensing, **48**, 8 (2010), 3178–3187.
- [24] Schoenmakers, R. and Vuurpijl, L.
Segmentation and classification of combined optical and radar imagery.
Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS '95), vol. 3, Florence, Italy, July 10–14, 1995, pp. 2151–2153.
- [25] Pasolli, E., et al.
Automatic detection and classification of buried objects in GPR images using genetic algorithms and support vector machines.
Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2008), vol. 2, July 7–11, 1998, pp. II-525–II-528.
- [26] Patra, S. and Bruzzone, L.
A fast cluster-assumption based active-learning technique for classification of remote sensing images.
IEEE Transactions on Geoscience and Remote Sensing, **49**, 5 (2011), 1617–1626.
- [27] Pellizzeri, T. M., et al.
Improved classification of SAR images by segmentation and fusion with optical images.
Proceedings of the IEEE Radar Conference (RADAR), Edinburgh, UK, Oct. 15–17, 2002, pp. 158–161.

- [28] Amin, M. and Ahmad, F.
Wideband synthetic aperture beamforming for through-the-wall imaging.
IEEE Signal Processing Magazine, **25**, 4 (2008), 110–113.
- [29] Pal, N. R. and Pal, S. K.
A review on image segmentation techniques.
Pattern Recognition, **26**, 9 (Sept. 1993), 1277–1294.
- [30] Sahoo, P. and Arora, G.
A thresholding method based on two-dimensional Renyi's entropy.
Pattern Recognition, **37**, 6 (2004), 1149–1161.
- [31] Pham, D., Dimov, S., and Nguyen, C.
Selection of k in k-means clustering.
Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science, **219**, 1 (2005), 103–119.
- [32] Wang, G. and Amin, M.
Imaging through unknown walls using different standoff distances.
IEEE Transactions on Signal Processing, **54**, 10 (2006), 4015–4025.
- [33] Seng, C., et al.
A two-stage image fusion method for enhanced through-the-wall radar target detection.
Proceedings of the IEEE Radar Conference (RADAR), Kansas City, MO, May 23–27, 2011, pp. 643–647.
- [34] Dilsavor, R., et al.
Experiments on wideband through the wall imaging.
Proceedings of the SPIE, vol. 5808, *Symposium on Defense and Security, Algorithms for Synthetic Aperture Radar Imagery XII*, 2005, pp. 196–209.
- [35] Thajudeen, C., et al.
Measured complex permittivity of walls with different hydration levels and the effect on power estimation of TWRI target returns.
PIER B: Progress in Electromagnetics Research B, **30** (2011), 177–199.
- [36] Dogaru, T. and Le, C.
SAR images of rooms and buildings based on FDTD computer models.
IEEE Transactions on Geoscience and Remote Sensing, **47**, 5 (2009), 1388–1401.



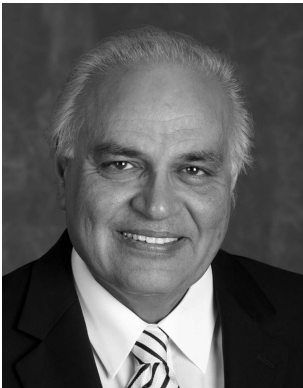
Cher Hau Seng received his B.E. (Hons) degree in computer engineering from the University of Wollongong, Australia.

He is currently working towards the Ph.D. degree at the School of Electrical, Computer, and Telecommunications Engineering, University of Wollongong. His general research interests are in the area of image processing with applications to radar and medical imaging.

Moeness Amin (F'01) received his Ph.D. degree in 1984 from the University of Colorado, Boulder in electrical engineering.

He has been on the faculty of the Department of Electrical and Computer Engineering at Villanova University since 1985. In 2002 he became the Director of the Center for Advanced Communications, College of Engineering.

Dr. Amin is a fellow of the International Society of Optical Engineering and a fellow of the Institute of Engineering and Technology (IET). He is a recipient of the IEEE Third Millennium Medal, recipient of the 2009 Individual Technical Achievement Award from the European Association of Signal Processing, recipient of the 2010 NATO Scientific Achievement Award, recipient of the Chief of Naval Research Challenge Award, recipient of the Villanova University Outstanding Faculty Research Award, 1997, and the recipient of the IEEE Philadelphia Section Award, 1997. He was a distinguished lecturer of the IEEE Signal Processing Society, 2003–2004, and is a member of the Electrical Cluster of the Franklin Institute Committee on Science and the Arts. He is a member of SPIE, EURASIP, ION, Eta Kappa Nu, Sigma Xi, and Phi Kappa Phi. Dr. Amin has over 550 journal and conference publications in the areas of wireless communications, time-frequency analysis, sensor array processing, waveform design and diversity, interference cancellation in broadband communication platforms, satellite navigations, target localization and tracking, direction finding, channel diversity and equalization, ultrasound imaging, and radar signal processing. He coauthored 20 book chapters and is the editor of the first book on through the wall radar imaging. He is a recipient of eight paper awards. He holds two U.S. Patents on smart antennas. Dr. Amin currently serves on the Overview Board of the *IEEE Transactions on Signal Processing* and the Editorial Board of the *IEEE Signal Processing Magazine*. He also serves on the Editorial Board of the *EURASIP Signal Processing Journal*. He was a plenary speaker at ICASSP 2010. Dr. Amin was the Special Session Cochair of the 2008 IEEE International Conference on Acoustics, Speech, and Signal Processing. He was the Technical Program Chair of the 2nd IEEE International Symposium on Signal Processing and Information Technology, 2002. Dr. Amin was the General and Organization Chair of both the IEEE Workshop on Statistical Signal and Array Processing, 2000 and the IEEE International Symposium on Time-Frequency and Time-Scale Analysis, 1994. He was an Associate Editor of the *IEEE Transactions on Signal Processing* during 1996–1998. He was a member of the IEEE Signal Processing Society Technical Committee on Signal Processing for Communications during 1998–2002. He was a member of the IEEE Signal Processing Society Technical Committee on Statistical Signal and Array Processing during 1995–1997. He has given several keynote and plenary talks and served as a session chair in several technical meetings.



Fauzia Ahmad (S'97—M'97—SM'06) received her M.S. degree in electrical engineering in 1996 and Ph.D. degree in electrical engineering in 1997, both from the University of Pennsylvania, Philadelphia.

From 1998 to 2000 she was an assistant professor in the College of Electrical and Mechanical Engineering, National University of Sciences and Technology, Pakistan. During 2000–2001 she served as an assistant professor at Fizaia College of Information Technology, Pakistan. Since 2002 she has been with the Center for Advanced Communications, Villanova University, Villanova, PA, where she is now a Research Associate Professor and the Director of the Radar Imaging Lab.

Dr. Ahmad has over 110 journal and conference publications in the areas of radar imaging, radar signal processing, waveform design and diversity, compressive sensing, array signal processing, sensor networks, ultrasound imaging, and over-the-horizon radar. She has also coauthored three book chapters in the aforementioned areas.





Abdesselam Bouzerdoun (M'89—SM'03) received the M.Sc. and Ph.D. degrees in electrical engineering from the University of Washington, Seattle.

In 1991 he joined The University of Adelaide, Adelaide, Australia, and in 1998, he was appointed an associate professor with Edith Cowan University, Perth, Australia. Since 2004 he has been a Professor of Computer Engineering with the University of Wollongong, Wollongong, Australia, where he also served as Head of the School of Electrical, Computer, and Telecommunications Engineering from 2004 to 2006 and has been the associate dean (research) with the Faculty of Informatics since 2007. He held a number of visiting appointments at several universities and research institutes, including Institut Galilée (Université Paris 13), LAAS-CNRS, Toulouse, France, The Hong Kong University of Science and Technology, Kowloon, Hong Kong, and Villanova University, Villanova, PA.

From 2009 to 2011 he was a member of the Australian Research Council College of Experts and was the Deputy Chair of the Engineering, Mathematics, and Informatics (EMI) panel from 2010 to 2011. He is the recipient of numerous awards and prizes; the most notable are the Eureka Prize for Outstanding Science in Support of Defence or National Security in 2011, the Chester Sall Award in 2005, and a Distinguished Researcher Award (Chercheur de Haut Niveau) from the French Ministry of Research in 2001. He has published over 280 technical articles and graduated many Ph.D. and Masters students. From 1999 to 2006 he served as Associate Editor of *IEEE Transactions on Systems, Man, and Cybernetics*. He is currently an associate editor of three international journals and a member of the governing board of the Asia Pacific Neural Network Assembly. He is also a member of the International Neural Network Society and the Optical Society of America.