Gaussian Mixture Model Based Features for Stationary Human Identification in Urban Radar Imagery

V. Kilaru, M. G. Amin, F. Ahmad
Radar Imaging Lab, Center for Advanced Communications
Villanova University
Villanova, PA 19085, USA

P. Sévigny, D. DiFilippo
Radar Sensing and Exploitation
Defence Research & Development Canada
Ottawa, ON, Canada K1A 0Z4

***Abstract*— In this paper, we propose a Gaussian mixture model (GMM) based approach to discriminate stationary humans from their ghosts and clutter in indoor radar images. More specifically, we use a mixture of Gaussian distributions to model the image intensity histograms corresponding to target and ghost/clutter regions. The mixture parameters, namely, the means, standard deviations, and weights of the component distributions, are used as features and a K-Nearest Neighbor classifier is employed. The performance of the proposed method is evaluated using real-data measurements of multiple humans standing or sitting at different locations in a small room. Experimental results show that the nature of the targets and ghosts/clutter in the image allows successful application of the GMM feature based classifier to distinguish between target and ghost/clutter regions.**

I. INTRODUCTION

Detection of stationary humans is one of the most challenging and important objectives in through-the-wall radar imaging (TWRI) and urban radar sensing [1]. Owing to losses in exterior-grade wall materials, high frequency radar operation is not feasible in TWRI. As such, the micro-Doppler signatures associated with biometric features, such as breathing and heartbeat, cannot be always detected and employed for identifying stationary humans behind walls, especially in SAR using moving radar platforms. Accordingly, detection of stationary humans would solely depend on effective image-domain detection techniques applied to through-the-wall radar imagery.

Image segmentation techniques have recently been considered in the literature for detecting stationary humans in TWRI [2]-[5]. These schemes were shown to be effective in distinguishing between target regions and those clutter regions which were distinct from target regions. However, in the presence of ghosts resulting from multipath propagation and clutter that closely mimics the targets in size and intensity, the performance of image segmentation methods degrades significantly.

In this paper, we investigate the use of Gaussian mixture model (GMM) based features for classifying target and ghost/clutter regions in indoor images. More specifically, we model the image intensity histograms of target and ghost/clutter regions with a mixture of Gaussian distributions [6]-[8]. GMM was considered because of its ability to form smooth approximations to arbitrarily shaped/multi-modal histograms. The mixture parameters, such as the means, variances, and the components weights, form the feature vector. *K*-Nearest Neighbor (*K*-NN) classifier is employed which classifies image regions based on closest training examples in the feature space [9], [10]. The training examples consist of the GMM-based feature vectors corresponding to known target and ghost/clutter image regions. The performance of the proposed GMM feature based detector is evaluated using real images acquired with the Multi-channel TWSAR, which is the vehicle-borne TWRI system developed by Defence Research and Development Canada (DRDC). The dataset corresponds to through-the-wall measurements of multiple humans of different heights, standing or sitting at different locations in a small room. We show that, for the specific data analyzed, the proposed scheme can successfully discriminate between target and ghost/clutter regions. We also compare the performance with the pixel-wise Likelihood Ratio Test (LRT) based approach of [11], modified to use GMMs for describing the target and clutter distributions.

The remainder of the paper is organized as follows. We present the proposed GMM based scheme for target and clutter classification in Section II. Section III briefly describes the LRT approach. Section IV evaluates the performance of the proposed and the LRT based methods using real data, and Section V concludes the paper.

II. PROPOSED TARGET AND CLUTTER CLASSIFICATION SCHEME

A. Gaussian Mixture Model

A Gaussian mixture model is a probabilistic model that assumes all the data points are generated from a mixture of a finite number of Gaussian distributions with unknown

This work is supported by DRDC under contract W7714-125544/00 l/SV.

parameters. One can think of mixture models as generalizing k-means clustering to incorporate information about the covariance structure of the data as well as the centers of the latent Gaussians.

A Gaussian mixture model is a weighted sum of M component Gaussian distributions [7], [8]

$$p(x) = \sum_{i=1}^{M} w_i p_i(x|\mu_i, \sigma_i), \quad (1)$$

where x is the image pixel intensity, w_i is the weight of the ith component, and $p_i(x|\mu_i, \sigma_i)$ is the ith component Gaussian distribution

$$p_i(x|\mu_i, \sigma_i) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left\{-\frac{1}{2\sigma_i^2}(x-\mu_i)^2\right\} \quad (2)$$

with mean μ_i and standard deviation σ_i. The component weights satisfy the constraint that $\sum_{i=1}^{M} w_i = 1$. The parameters of the GMM are collectively represented by the set $\{w_i, \mu_i, \sigma_i\}_{i=1}^{M}$. The expectation-maximization (EM) algorithm is used to maximize the likelihood estimate of the mixture parameters [8], [12].

Note that the selection of the number of mixture components is important. With too many components, the mixture may over fit the data and yield poor interpretations, while with too few components, the mixture may not be flexible enough to approximate the true underlying data structure. The number of mixture components was empirically determined to be nine for the experimental TWRI data considered in this paper.

B. Feature Vectors

The parameter sets $\boldsymbol{X}_m = [w_m, \mu_m, \sigma_m]^T,\ m = 1, 2, \cdots, 9$ of all nine Gaussian components are arranged in a tall vector in ascending order of their means, i.e., $\boldsymbol{Y} = [\boldsymbol{X}_1^T, \boldsymbol{X}_2^T, \ldots, \boldsymbol{X}_9^T]^T$. Therefore, the length of the resulting feature vector $\boldsymbol{Y}$ is 27. The feature vectors were computed for several image regions, which were known to correspond to stationary humans and ghost/clutter based on the ground truth. These sets of extracted feature vectors form the training set for the classification problem at hand.

C. K-Nearest Neighbors Classifier

We consider the K-NN classifier, which is commonly used in learning and classification [9], [10]. Therein, an object is classified based on the "distance" of its features from those of its neighbors (i.e. the training set), with the object being assigned to the class most common among its K nearest neighbors. Euclidean distance is the commonly used distance metric, defined between two feature vectors $\boldsymbol{Y}_i$ and $\boldsymbol{Y}_j$ as

$$D = \left\| \boldsymbol{Y}_i - \boldsymbol{Y}_j \right\| \quad (3)$$

with $\|\cdot\|$ denoting the l_2 norm. If $K = 1$, the algorithm simply becomes the nearest neighbor algorithm and the object is assigned to the class of its nearest neighbor. If $K > 1$, the object is assigned to the class of the majority of its K nearest neighbors. Typically, K is chosen to be odd when the number of classes is two to resolve any ties. A higher K increases the classification accuracy but at the expense of computational time.

D. Proposed Scheme

To summarize, the proposed algorithm consists of the following steps:

Step 1: Extract the candidate image region to be classified from the intensity image.

Step 2: Obtain the Gaussian mixture parameters for the extracted region, and generate the corresponding test feature vector.

Step 3: Compute the Euclidean distance between the test feature vector and each of the training feature vectors.

Step 4: Classify the candidate region as target or ghost/clutter depending on the class most common among its K nearest neighbors.

It is assumed that the training set of feature vectors has been obtained prior to the testing phase using Steps 1 and 2 above for known target and ghost/clutter regions.

III. Modified LRT Approach

In the image-domain based LRT approach proposed in [11], a pixel-wise hypothesis test is formulated and a Neyman Pearson test [13] is applied. The target and clutter are assumed to follow Gaussian and Weibull distributions in [11]. However, these assumptions were shown to produce inferior results for images acquired with DRDC's Multi-channel TWSAR system [14]. We, therefore, modify the non-adaptive LRT approach of [11] by using GMMs for both target and clutter distributions.

We define the pixel-wise null and alternate hypotheses as

$$\begin{aligned} H_0&: \text{clutter present at the pixel under considertion} \\ H_1&: \text{target present at the pixel under consideration} \end{aligned} \quad (4)$$

Assuming the data to be independent and identically distributed, the LRT is given by

$$LR(x) = \frac{p(x|H_0)}{p(x|H_1)} \underset{H_0}{\overset{H_1}{\gtrless}} \gamma \quad (5)$$

where $p(x|H_0)$ and $p(x|H_1)$ are the conditional probability density functions, given the null (ghost/clutter class) and alternate (target class) hypothesis, respectively, which are expressed as

$$p(x|H_0) = \sum_{n=1}^{N} w_n p_n(x|\mu_n, \sigma_n), \quad (6)$$

$$p(x|H_1) = \sum_{l=1}^{L} w_l p_l(x|\mu_l, \sigma_l), \quad (7)$$

with N and L being the respective number of mixture components under the null and alternate hypotheses. The parameter γ in (5) is the LRT threshold, which maximizes the probability of detection while controlling the probability of

(a)

(b)

(c)

Figure 1. (a) TWSAR System, (b) Troop Shelter building (small room indicated by the dashed square), (c) Scene with four human occupants.

false alarm. Given the image statistics $p(x|H_0)$ and $p(x|H_1)$ and the threshold γ, the output binary pixel value can be computed as

$$\varphi(x) = \begin{cases} 1, & LR \geq \gamma \\ 0, & LR < \gamma \end{cases} \tag{8}$$

Using the Neyman-Pearson theorem, the false alarm rate α can be fixed by evaluating

$$\alpha = \int_{\gamma}^{\infty} f_L(L|H_0) dL \tag{9}$$

where $f_L(L|H_0)$ is the probability density function of the likelihood ratio under the null hypothesis.

The LRT based approach operates on a pixel-by-pixel basis and, thus, the step in (8) generates a decision for each individual pixel. In order to arrive at a single decision for the entire candidate image region, two counters are set up. Every time a pixel is assigned to a particular class, the corresponding counter is incremented. At the end of the process, the candidate image region is assigned to the class of the counter with the largest count.

IV. Experimental Results

A. System Parameters and Experimental Setup

We evaluate the proposed scheme using real three-dimensional (3D) images collected with the experimental through-the-wall multiple-input multiple-output radar testbed developed by DRDC [15]. The radar is installed inside a vehicle with its two transmit antennas and an eight-element receive array mounted on the side of the vehicle, as shown in Fig. 1(a). The antenna elements are compact Y-shaped printed bowtie antennas and, when used in the vertical polarization, have approximately 60º beamwidth in the elevation direction and 150º beamwidth in the azimuth or horizontal direction [16]. The receive array has an inter-element spacing of 15 cm, and the two transmit antennas are separated by 1.2 m. The transmit and receive array antennas have a horizontal spacing of 2 m. A frequency-modulated continuous wave (FMCW) signal covering the 0.8 to 2.7 GHz frequency band is used as the transmit signal. A switch is used to alternate the radar transmissions between the two transmit antennas, and the eight-channel radar receiver digitizes the eight received signals for each radar transmission.

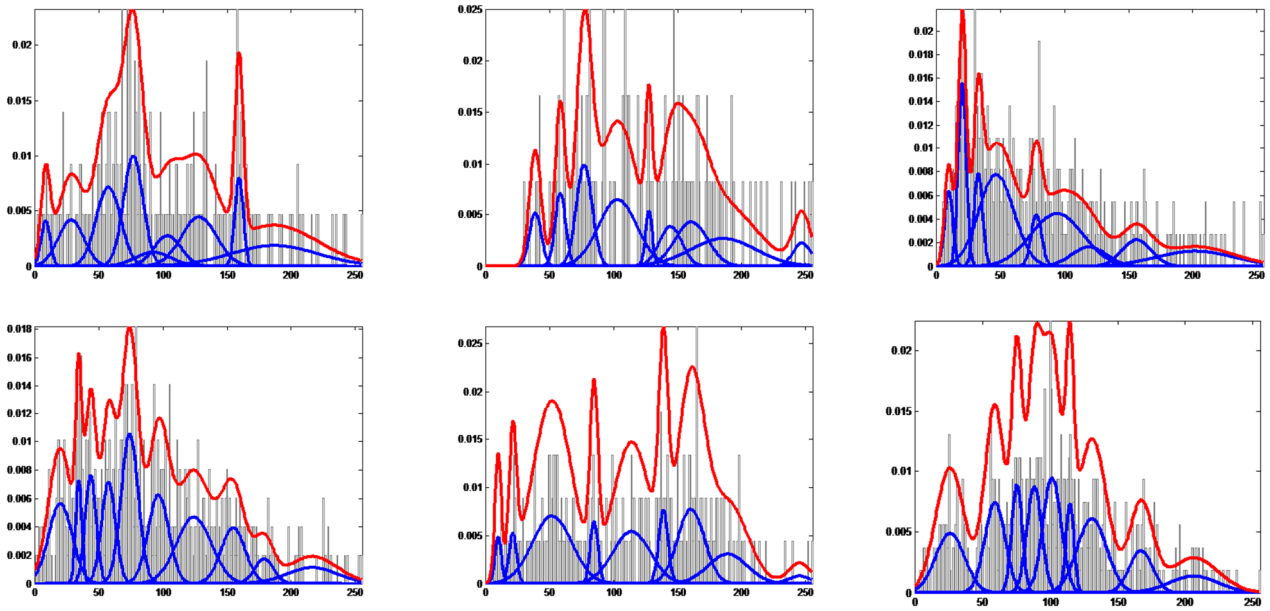

Figure 2. Target and ghost/clutter histograms and GMM fits. Top row corresponds to the targets and the bottom row to the ghosts/clutter (first two in each are from the first dataset and third from the latter measurements.) The red curve depicts the Gaussian fit and the blue curves show the individual Gaussian components. The horizontal and vertical axes represent the pixel value and the probability, respectively.

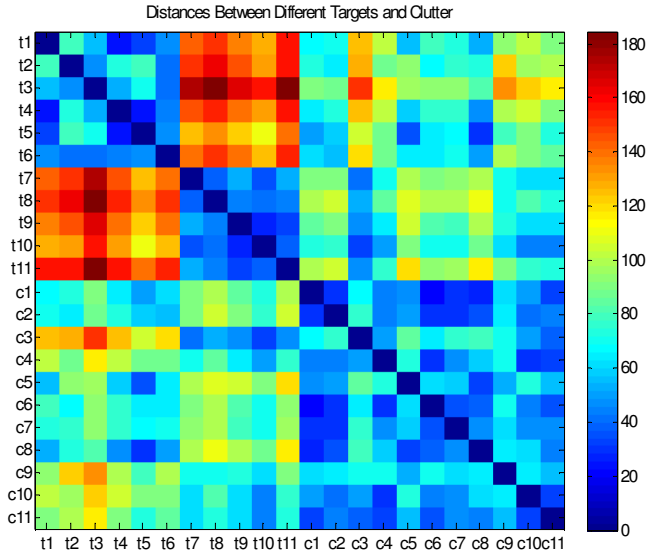

Figure 3. Euclidean distances between various target and ghost/clutter feature vectors.

A small room in the Troop Shelter building, shown in Fig. 1(b), was imaged three different times, with six, four, and one human occupant, respectively. The antennas were lowered on the van between measurements from the first scene and the latter scenes. Fig. 1(c) depicts the scene with the four human targets. The exterior walls of the building are constructed of vinyl, chip board and drywall on a 16 in. spacing wood stud frame. The raw radar data were collected while the vehicle moved along a straight path parallel to the front wall of the building, allowing 3D images to be generated in downrange, azimuth, and elevation using backprojection.

TABLE I. CONFUSION MATRIX FOR *K*-NN CLASSIFER WITH *K*=3.

Class	**Classification**	
	Target	***Ghost/Clutter***
Target	100%	0
Ghost/ Clutter	9%	91%

B. GMM Based Classification Results

Eleven target regions and eleven clutter regions were extracted from the 3D through-the-wall images. Each of the extracted regions was modeled by a mixture of nine Gaussian distributions, and the corresponding 27-element feature vector was estimated using the EM algorithm. Figure 2 depicts the intensity histograms and the corresponding GMM fits for six of the extracted target and ghost/clutter regions. The red curve depicts the Gaussian fit and the blue curves are the individual Gaussian components. Classification is performed using these feature vectors. Because of limited dataset, we used leave-one-out cross validation [17], wherein the classification is performed 22 times, using one feature vector from the dataset for testing and the remaining for training each time. In this way, all of the target and ghost/clutter regions in the dataset were used for both training and testing.

Fig. 3 depicts the distances from each target and ghost/clutter feature vector to all 22 of the target and ghost/clutter feature vectors. The distinction between the target and ghost/clutter classes is evident in Fig. 3. Further, it is observed that the target feature vectors from the first and the latter datasets are quite different. This is because when the antennas were lowered, a much higher level of background clutter was

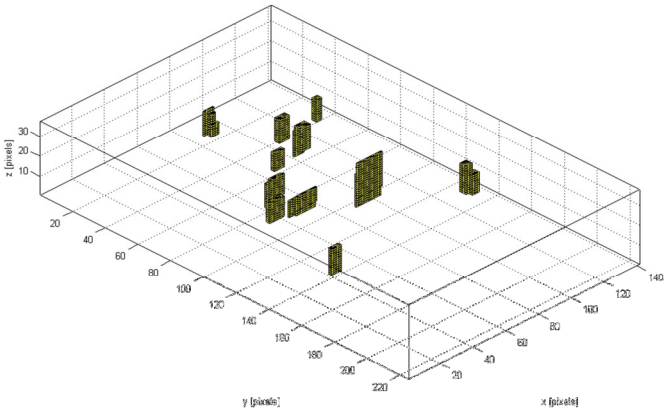

Figure 4. Result obtained using the LRT approach for the six human occupant dataset with a preset false alarm rate of 5%.

picked up, which infringed on the target regions in the corresponding images. Table I presents the confusion matrix for the *K*-NN classifier with *K*=3, which shows that the GMM feature based classifier provided a high classification accuracy with no missed detections and 9% false alarms.

C. Modified LRT Based Results

The output binary 3D image generated with the modified LRT approach applied to the six human occupant dataset is shown in Fig. 4 for a preset value of $\alpha = 0.05$. We observe that most of the clutter appears at the output along with the targets, thereby causing a much higher number of false alarms compared to the proposed GMM feature based technique. Similar results were observed for other values of α ranging from 0.01 to 0.2 and for the four and one human occupant cases.

V. CONCLUSION

In this paper, we presented a Gaussian mixture modeling approach for detecting stationary human targets in through-the-wall radar imagery. The problem of target discrimination from multipath ghosts and clutter was investigated by means of a feature-extraction strategy and a K-NN classifier based on the Euclidean distance metric. The mixture parameters, namely, the means, variances, and weights of the component Gaussian distributions, were used as features. The performance of the proposed scheme was evaluated using real 3D images. The results showed that the proposed scheme was successful in providing higher classification accuracy compared to the likelihood ratio test based approach.

REFERENCES

[1] M. Amin, *Through the Wall Radar Imaging*, CRC Press, 2010.

[2] M. G. Amin, P. Setlur, F. Ahmad, P. Sevigny, and D. DiFilippo, "Histogram based segmentation for stationary indoor target detection," in *Proc. SPIE*, Baltimore, MD, April 2012.

[3] A. A. Mostafa and A.M. Zoubir, "3D target detection in through-the-wall radar imaging," in *Proc. SPIE* , vol. 7697, pp. 76971F, 2011.

[4] C. H. Seng, A. Bouzerdoum, M. G. Amin, F. Ahmad, "A Gaussian-Rayleigh mixture modeling approach for through-the-wall radar image segmentation," in *Proc. IEEE Int. Conf. Acoustics, Speech Signal Process.*, 2012.

[5] C. H. Seng, M. G. Amin, F. Ahmad, and A. Bouzerdoum, "Image segmentations for through-the-wall radar target detection," *IEEE Trans. Aerosp. Electronic Syst.,* vol. 49, no. 3, pp. 1869 - 1896, 2013.

[6] K. Pearson, "Contributions to the mathematical theory of evolution," *Phil. Trans. Royal Soc.*, 185A, pp. 71-110, 1894.

[7] G. McLachlan and D. Peel, *Finite Mixture Models*, John Wiley & Sons, Inc., Hoboken, NJ, 2000.

[8] R. A. Redner and H. F. Walker, "Mixture densities, maximum likelihood and the EM algorithm," *SIAM Review*, vol. 26, no. 2, pp. 195-239, Apr. 1984.

[9] J. H. Friedman, J. L. Bentley, and R. A. Finkel, "An algorithm for finding best matches in logarithm expected time," *ACM Trans. Mathematical Software*, vol. 3, pp. 209-226, 1977.

[10] B.V. Dasarathy (Ed.), *Nearest Neighbor NN Norms: NN Pattern Classification Techniques*, IEEE Computer Society Press, Los Alamitos, CA, 1991.

[11] C. Debes, M. G. Amin, and A. M. Zoubir, "Target detection in single- and multiple-view through-the-wall radar imaging," *IEEE Trans. Geosci. Remote Sens.,* vol. 47, no. 5, pp. 1349-1361, 2009.

[12] T. K. Moon, "The expectation-maximization algorithm," *IEEE Signal Process. Mag.*, 1996.

[13] S. M. Kay, *Fundamentals of Statistical Signal Processing*, vol. 2. Englewood Cliffs, NJ: Prentice-Hall, 1998.

[14] M. G. Amin, P. Setlur, F. Ahmad**,** P. Sevigny, and D. DiFilippo, "Histogram-based segmentation for stationary indoor target detection," in *Proc. SPIE,* vol. 8361, 2012, pp. 83610M.

[15] P. Sévigny et al., "Concept of operation and preliminary experimental results of the DRDC through-wall SAR system," in *Proc. SPIE*, vol. 7669, 2010, pp. 766907.

[16] S. Raut and A. Petosa, "A compact printed bowtie antenna for ultra-wideband applications," in *Proc. European Microwave Conf.*, 2009, pp. 81-84.

[17] M. Stone, "Cross-validatory choice and assessment of statistical predictions," *J. Roy. Statist. Soc. Ser. B*, vol. 36, pp. 111–147, 1974.