Automated Stationary Human Target Detector for 3-D Through-Wall Radar Imagery

P. Sévigny and J. Fournier
Defence Research and Development Canada
CANADA

pascale.sevigny@drdc-rddc.gc.ca

The use of through-wall radar imagery for remote intelligence of building interiors is promising but challenging. Relevant information about human targets and room layout features is typically buried in clutter. In this Letter, we propose a methodology for automated extraction of information about stationary human targets behind walls. Based on treatment of the individual blobs found in the imagery, the method consists of thresholding, segmentation, classification and 3-D visualization. Although further optimization of each of the steps is required, we demonstrate with two examples, including one cluttered scene, that the combination of these steps is effective at eliminating large amounts of clutter and identifying human targets behind walls.

INTRODUCTION

Through-wall radar is an important modality for remote intelligence of building interiors. The low frequency signals penetrate wall materials with limited attenuation, and scatter from objects behind walls. The scattering phenomenology is complex and includes contributions from objects of interest (humans, layout features) in addition to the contributions from furniture items, internal wall features and multi-path signals between all these contributors. Relevant information is thus buried in clutter, but can still be extracted about the location and motion of human targets behind walls [1],[2], and of room layout features [3],[4]. Although recent studies have demonstrated the potential of through-wall radar for detection of stationary human targets [5],[6], the detection task remains laborious, highlighting the requirement for automated information extraction tools.

In this Letter, we present our methodology for automated extraction of information about stationary human targets behind walls. The method is based on treatment of the "blobs", i.e. connected voxels, found in the 3-D through-wall radar imagery. It consists of thresholding, segmentation, classification, and 3-D visualization. It was developed based on an extensive set of experimental data acquisitions over a variety of buildings, under controlled but realistic conditions. The individual steps remain to be optimized, however we demonstrate with two examples how the method can be successful at detecting stationary human targets behind walls while eliminating large amounts of clutter.

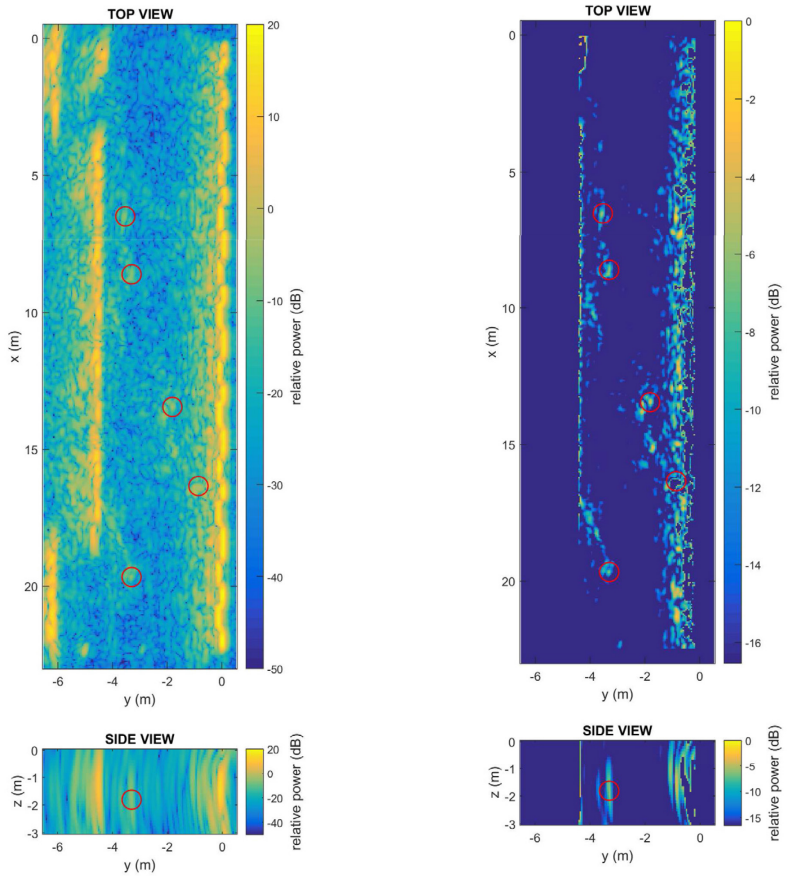
PROTOTYPE VEHICLE-BASED THROUGH-WALL RADAR AND SAR IMAGE FORMATION

The Defence Research and Development Canada (DRDC) L-band through-wall radar is mounted on a vehicle and operated as a synthetic aperture radar while the vehicle is driven in front or around a building of interest. The two transmit antennas toggling in time and the eight-element vertical receive array enable slow-time multiple-input multiple output (MIMO) operation for doubled resolution in the elevation direction. The ultrawideband radar achieves resolutions of 12 cm in the range and azimuth directions, while the elevation resolution is bound to 4 degrees, corresponding to approximately 60 cm at 10 m range. This elevation resolution is not ideal but is crucial in distinguishing between ceiling features for example and human targets located closer to the ground. The 3-D radar is supported by a line-scanning LIDAR

PUBLIC RELEASE

(LIght Detection And Ranging) system which provides 3-D high-resolution information of building exteriors, including precise location of doors and windows and of other external features on the front wall.

The through-wall radar imagery is processed using a time-domain back-projection algorithm resulting in a 3-D matrix of data (voxels) ranging over typically 60-80 dB. In Figure 1(a) are shown two slices of the 3-D SAR image of a building with a front wall made of cinder blocks. The building includes six adjacent rooms, separated by cinder block internal walls perpendicular to the radar path, as can be seen on the photograph of Figure 2. Access to each room is provided by a front door and a window, both of which are covered with a plywood panel. The radar path goes from $(x, y, z) = (-10, +10, -1.4)$ m to $(x, y, z) = (+30, +10, -1.4)$ m.



(a) Original

(b) Thresholded, wall and exterior voxels removed

Figure 1: 2-D top view and side view slices of the 3-D SAR image of the mockup motel building.

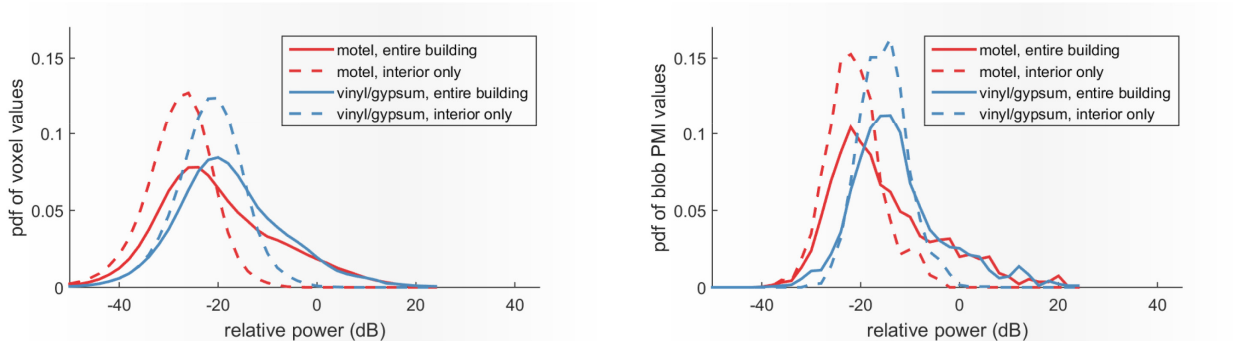
Figure 2: Photograph of the mockup motel building.

On the top view image of Figure 1(a), the path is thus at the right and goes from top to bottom. Five human targets were located in five different rooms at the time of data acquisition, and their signatures are circled on the image. The texture of the imagery is blob-like, evidenced by the granular aspect of the top view SAR image. The typical blob is elongated in the elevation direction and follows a cylindrical frame of reference with its axis along the path of the antenna phase center. This is best observed in the side view image, especially at shorter ranges from the radar. Each blob is characterized by the location and intensity of its pixel of maximum intensity (PMI), which is a local maximum of the 3-D data matrix.

THRESHOLDING AND SEGMENTATION

To efficiently process and extract information from the entire 3-D through-wall SAR image, a threshold is applied and the remaining regions of interest are segmented for further processing. This step is critical in the automation of the human target detection task in the context of images of entire buildings where the number of blobs can be relatively large.

The threshold is computed as the median of the blob PMI values. This choice of a threshold relies on empirical observations over several buildings [7]. These investigations have shown that the statistics of the voxel values vary widely from building to building and depending on the scene chosen (with or without the front wall). The statistics of the blob PMI values have been found to not depend as much on the scene content and building type. Using the median value allows the elimination of half the blobs or candidate targets while retaining the typical human target signature. Other common thresholding methods widely used in image processing rely on, for example, bimodal probability distributions [8]. Statistical methods have been used for through-wall radar image applications using assumptions on the probability distributions [9]. In our experience, these methods are not easily applicable to the experimental 3-D through-wall SAR images. Instead we use the median of the blob PMI values as a simple but robust threshold, even if it may not be optimal. In Figure 3 are shown examples of the probability distribution functions (pdf) of the voxel values and of the blob PMI values for a few different buildings and scenes.



(a) Voxel values **(b) Blob PMI values**

Figure 3: Pdf of the voxel and blob PMI values, for four 3-D SAR images of the mockup motel and a vinyl/gypsum/wood studs building.

The watershed algorithm [10] is used to automatically segment the 3-D image as a collection of distinct blobs, one for each local maximum, and a background region. The algorithm finds the boundaries between the different blobs and between blobs and background voxels. Its advantage is that any given voxel is assigned to only one blob; each blob is assigned a number of voxels which will be used later to compute features.

In addition to the thresholding and segmentation steps, we remove the voxels of high intensity belonging to the front wall and to some very bright features in the imagery. This step eliminates a number of blobs of high

intensity which would act as confusers for the classifier. Rather than using a fixed threshold, a region-growing algorithm [11] is applied with seeds at the location of the blob PMIs of strongest intensity in the imagery. Finally, the voxels outside the boundaries of the building are thresholded. The thresholded and wall voxels removed image of the building under study is shown in Figure 1(b). The thresholding step is effective in zeroing many voxels of the empty areas, however some clutter signal remains. Some of the wall signature, created by multi-path phenomena, also remains. For the scene of interest, one of the human targets is located very close behind the front wall, and its signature is blended in the wall signature. Although it is detected by the radar analyst, it is discarded by the wall voxels removal algorithm. The signatures of only four human targets are thus available for prediction by the classifier.

CLASSIFICATION

14 features are calculated for each blob or candidate target in the 3-D SAR image surviving the thresholding step. Most features are related to the shape of the blob, such as the length of its three principal axis, its meridional and equatorial eccentricities, and its extent in the range and azimuth directions. The elevation and intensity of the blob PMI are important features. Other features include the count of neighbouring blobs in a pre-defined box centered on the blob PMI. Some of the current features are redundant or correlated, and additional features are being investigated. A formal study on feature selection remains to be performed. Nonetheless, the current features are sufficient to illustrate the methodology with acceptable results.

A two-class Support Vector Machine (SVM) classifier [12] is trained using experimental data collected over buildings (distinct from the building under test) of various wall materials where one or more human targets are positioned inside the buildings. A variety of human target positions are available (sitting, standing, holding a rifle, etc.), but most are standing still. Detailed ground truth of the human target locations is available and is used by the radar analyst to associate a given human target to a blob in the imagery. The radar analyst detections are the “truth” for the classifier training and testing. Given the large class imbalance in the experimental data sets (of the order of a thousand clutter blobs (negatives) for less than ten human targets (positives)), the true and false positive rates (TPR and FPR, respectively) are used as the classifier performance metrics. In Figure 4 are reported the TPR and FPR as a function of mis-classification cost imposed to the classifier, for different training data sets. These results were obtained while testing on five different data acquisitions with different human target scenarios of the building under study. 16 human targets were present, explaining the coarse steps of the TPR curve. In comparison, there were 2892 clutter blobs. As the mis-classification cost is increased, penalizing the classification of the true positives as negatives, the true positive rate increases as expected. However, more clutter blobs (negatives) are falsely classified as positives, increasing the false positive rate. The graph also demonstrates that the choice of the training data set impacts the classifier performance. Group 2 included all data acquisitions available from three buildings with walls made of vinyl, gypsum and wood studs. These buildings do not attenuate much the radar signals, and tend to produce lots of multi-path signals due to the periodic wood stud structure. Group 1 included a sub-set of data acquisitions from Group 2, such that the proportion of human targets in the various positions is balanced. Group 3 included data acquisitions from a variety of buildings with differing wall materials including vinyl/gypsum/wood studs, cinder block, cinder block plus brick. In what follows, we use a mis-classification cost of 10^2 and the Group 1 training data set.

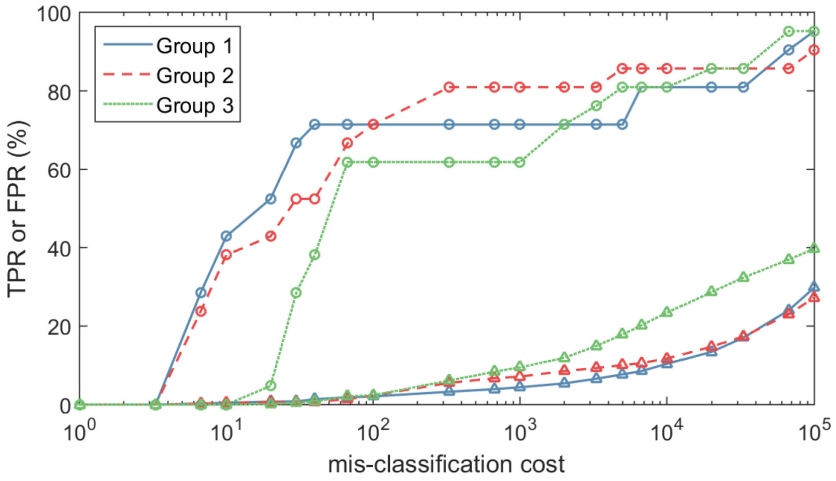
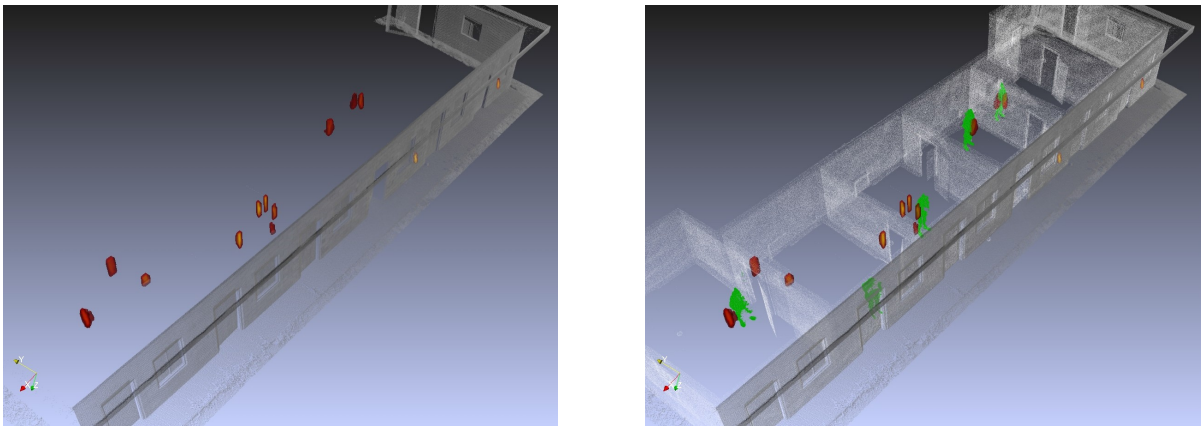


Figure 4: TPR (circles) and FPR (triangles) as a function of mis-classification cost for three different training data sets.

3-D VISUALIZATION

The output of the classifier is a predicted label for each of the candidate targets (blobs). Of high interest is the distribution of the positives relative to the 3-D scene to better interpret the classifier results, as 10 false positives closely related to the front wall is a very different outcome than 10 false positives scattered inside a room. The 3-D visualization methodology is the following. For each candidate target predicted as positive, a seed is positioned at the location of its PMI. A region-growing algorithm grows a volume from that seed, restricted to connected voxels that are within 2 dB of the blob PMI SAR intensity and within 4 pixels of the seed location. Once the volume is grown for a given positive, a new volume is started with the next positive as the seed. The result is a 3-D mask of N volumes corresponding to the N classifier positives. The mask is multiplied by the original 3-D SAR image to obtain the final 3-D classifier image. This method allows to conserve the shape and intensity of the blobs in the original 3-D SAR image while eliminating clutter. The final 3-D image can be visualized using common 3-D volume rendering methods.

In Figure 5(a) is shown the classifier image for the scene under study, overlaid with the mobile LIDAR data that was collected at the same time as the radar data, and which provides context information with the exterior of the building. The classifier image consists of the colored volumes while the mobile LIDAR point cloud is displayed using gray levels. The 3-D visualization is rendered using ParaView, an open- source, multi-platform data analysis and visualization application [13]. To provide a detailed survey of the rooms including precise location of human targets and furniture, ground-truth LIDAR data was also collected in-between data acquisitions, using a fixed LIDAR system inside the buildings. The ground-truth LIDAR point cloud is overlaid in Figure 5(b) to better interpret the classifier results. The ground-truth LIDAR data is displayed as white dots for the interior walls and green dots for the human targets. We see that the four human targets are correctly predicted by the classifier (recall that the fifth human target was located too close behind the wall and was rejected by the wall removal algorithm). Some of the human target signatures are slightly shifted relative to the ground-truth LIDAR data, possibly due to human motion between data acquisitions. Two of the human target signatures include two or more closely located blobs. Of the false positives, two can be associated with the front wall, and three belong to internal walls.



(a) Mobile LIDAR point cloud and 3-D classifier image

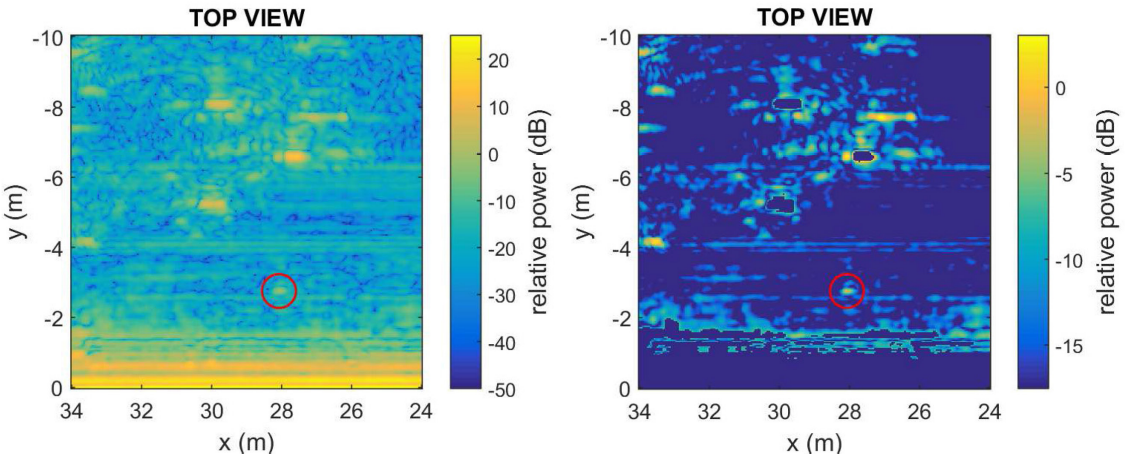
(b) Mobile LIDAR point cloud and 3-D classifier image overlaid with ground- truth LIDAR point cloud (white and green dots)

Figure 5: 3-D visualization for the mockup motel building.

EXAMPLE OF A CLUTTERED SCENE

The results above demonstrate that the combination of the thresholding and classification approaches is effective at eliminating a large amount of clutter and identifying true human targets. In this section is examined the case of a scene cluttered with furniture. The building is a large garage with a front wall made of a layer of cinder blocks and a layer of bricks. Immediately behind the wall is an empty area where one human target is positioned, and further behind are piled various furniture and storage items, such as desks, chairs, large metallic cabinets and a metallic portable staircase. Photographs of the building exterior and interior are provided in Figure 6. The top view 2-D slice of the 3-D original and thresholded SAR images are shown in Figure 7, where the radar path goes from right to left and at the bottom. The thresholded SAR image shows that an important amount of clutter remains, especially in the area filled with furniture.

Figure 6: Photographs of the garage building.

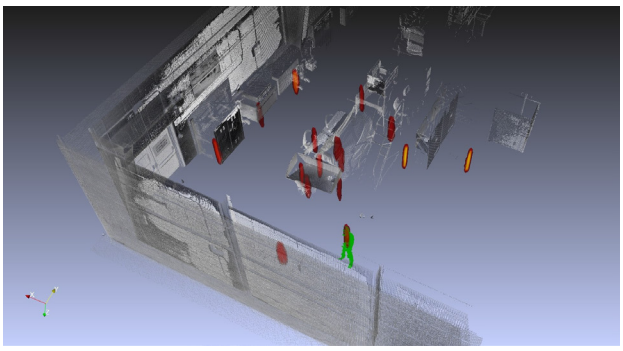


(a) Original

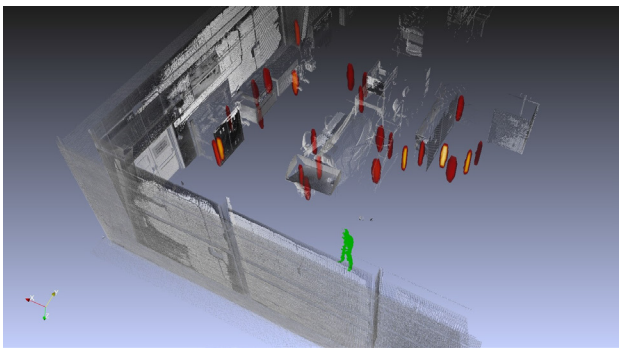
(b) Thresholded, wall voxels and exterior voxels removed

Figure 7: 2-D top view slices of the 3-D SAR image of the garage building.

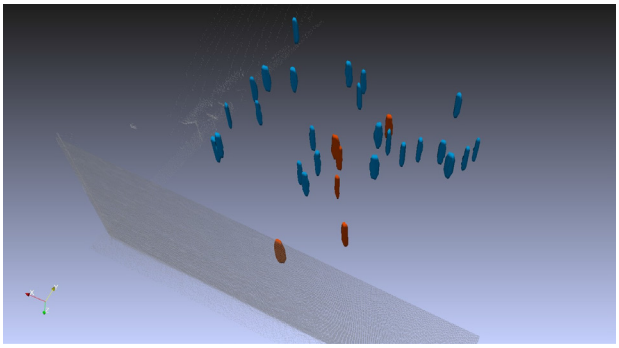
In Figure 8(a) is shown the 3-D classifier image for the scene of interest, obtained with the Group 1 training data set and mis-classification cost of 10^2. The classifier succeeds in predicting the presence of the human target. However it predicts as positives a number of candidate targets that actually belong to the furniture items. To improve this classification, another SVM classifier was trained to predict the presence of furniture only. To do so, the same training data set as before was used, but all samples were labelled as negatives (this training data set consists of buildings with rooms empty of furniture). Furniture samples were added as follows and labelled as positives. Surveying blobs and identifying them as furniture items is a tedious task. Luckily, another classifier result from a building filled with furniture resulted in a large number of false positives which could be manually assessed as belonging to furniture, thanks to the very good match with the ground-truth LIDAR data. This provided 39 furniture samples which could be labelled as positives to train a furniture classifier. The results of the furniture classifier can be seen in Figure 8(b). It detects 26 blobs in the area filled with furniture, and as desired it does not detect the human target or any front/side wall blobs. When combining the results of the two classifiers, we can provide the image of Figure 8(c) to a potential user, where the blue and red blobs represent furniture and human targets, respectively. 10 false positives of the original classifier result are now declared as furniture, and only 6 false positives remain. One of these false positives can be associated to the wall signature. More furniture-specific features and more furniture samples will be required to further improve this encouraging result.



(a) Mobile LIDAR point cloud and human target classifier image overlaid with ground-truth LIDAR point cloud (white and green dots)



(b) Mobile LIDAR point cloud and furniture classifier image overlaid with ground-truth LIDAR point cloud (white and green dots)



(c) Mobile LIDAR point cloud and combined classifier image

Figure 8: 3-D visualization for the garage building.

CONCLUSION

In this Letter, we proposed a methodology for automated extraction of stationary human target information behind walls using a vehicle-based 3-D synthetic aperture through-wall radar. Rather than using all voxels in the imagery for which the statistics are typically unknown, the method is based on the individual blobs found in the imagery, identified as the local maxima in the volumetric radar data. The method involves the median of the blob PMI values as the threshold, and makes use of common image processing algorithms such as seeded region-growing and watershed for segmentation of the 3-D imagery. Classification using SVMs is performed to distinguish human target signatures from clutter. 3-D visualization tools are used to better interpret the classifier results. The method has been tested on an extensive set of experimental data acquisitions over various buildings with different wall materials. Although further optimization of each of the steps is still required, we demonstrated with two examples, including one cluttered scene, that the combination of these steps is effective at eliminating large amounts of clutter and identifying human targets.

REFERENCES

[1] Ahmad, F., Amin, M. G.: 'Through-the-wall human motion indication using sparsity-driven change detection', *IEEE Trans. Geosci. Remote Sens.*, 2013, **51**,(2), pp.881–890, doi:10.1109/TGRS.2012.2203310.

[2] Li, J., Zeng, Z., Sun, J., Liu, F.: 'Through-wall detection of human being's movement by UWB radar', *IEEE Geosci. Remote Sens. Lett.*, 2012, **9**, (6), pp.1079–1083, doi: 10.1109/LGRS.2012.2190707.

[3] de Wit, J., Anitori, L., van Rossum, W., Tan, R.: 'Radar mapping of buildings using sparse reconstruction with an overcomplete dictionary', 2011 European Radar Conference (EuRAD), Manchester, UK, October 2011, pp. 9–12.

[4] Le, C., Dogaru, T., Nguyen, L., Ressler, M. A.: 'Ultrawideband (UWB) radar imaging of building interior: Measurements and predictions', *IEEE Trans. Geosci. Remote Sens.*, 2009, **47**, (5), pp. 1409–1420, doi: 10.1109/TGRS.2009.2016653.

[5] Sévigny, P.: 'Joint through-wall 3-D radar imaging and motion detection using a stop-and-go SAR trajectory', 2016 IEEE Radar Conference (RadarConf), Philadelphia, USA, May 2016, pp. 1–5, doi: 10.1109/RADAR.2016.7485325.

[6] Xin, S., BiYing, L., Yang, Z., LanZi, Z., ZhiMin, Z.: 'Wall artifacts removal for target imaging enhancement in UWB through-the-wall radar application', *Signal Processing*, 2014, **104**, pp. 325–338.

[7] Sévigny, P.: 'Clutter reduction using an image-based thresholding approach for 3-D through-wall synthetic aperture radar', 2015, Defence R&D Canada âA˘ S¸ - Ottawa, DRDC-RDDC-2015-R278.

[8] Otsu, N.: 'A threshold selection method from gray-level histograms', *IEEE Trans. Syst., Man, Cybern.*, 1979, **9**, (1), pp. 62–66, doi: 10.1109/TSMC.1979.4310076.

[9] Debes, C., Zoubir, A. M.: 'Detection approaches in through-the-wall radar imaging' in Amin, M. G. (Ed.): 'Through-the-wall radar imaging' (CRC Press, Boca Raton, FL, 2011).

[10] Meyer, F.: 'Topographic distance and watershed lines', *Signal processing*, 1994, **38**, (1), pp. 113–125.

[11] Adams, R., Bischof, L.: 'Seeded region growing', *IEEE Trans. Pattern Anal. Mach. Intell.*, 1994, **16**, (6), pp. 641–647, doi: 10.1109/34.295913.

[12] Cristianini, N., Shawe-Taylor, J.: 'An introduction to support vector machines and other kernel-based learning methods', (Cambridge University Press, UK, 2000).

[13] Ahrens, J., Geveci, B., Law, C.: 'Paraview: An end-user tool for large-data visualization' in Hansen, C. D., Johnson, C. R. (Ed.): 'The Visualization Handbook' (Elsevier, Oxford, UK, 2005).