SAR image dataset of military ground targets with multiple poses for ATR

Carole Belloni, Alessio Balleri, Nabil Aouf, Thomas Merlet, Jean-Marc Le Caillec

Carole Belloni, Alessio Balleri, Nabil Aouf, Thomas Merlet, Jean-Marc Le Caillec, "SAR image dataset of military ground targets with multiple poses for ATR," Proc. SPIE 10432, Target and Background Signatures III, 104320N (5 October 2017); doi: 10.1117/12.2277914

SPIE.

Event: SPIE Security + Defence, 2017, Warsaw, Poland

Proceedings Volume 10432, Target and Background Signatures III, 11-14 September 2017, Warsaw, Poland.
DOI:10.1117/12.2277914

SAR Image Dataset of Military Ground Targets with Multiple Poses for ATR

Carole Belloni[1,2], Alessio Balleri[1], Nabil Aouf[1], Thomas Merlet[3], and Jean-Marc Le Caillec[2]

[1]Signals and Autonomy Group, Centre of Electronic Warfare, Information and Cyber (CEWIC), Cranfield University, Defence Academy of the UK, Shrivenham, SN6 8LA
[2]IMT Atlantique, Brest, France
[3]Thales Optronique, Elancourt, France

ABSTRACT

Automatic Target Recognition (ATR) is the task of automatically detecting and classifying targets. Recognition using Synthetic Aperture Radar (SAR) images is interesting because SAR images can be acquired at night and under any weather conditions, whereas optical sensors operating in the visible band do not have this capability. Existing SAR ATR algorithms have mostly been evaluated using the MSTAR dataset.[1] The problem with the MSTAR is that some of the proposed ATR methods have shown good classification performance even when targets were hidden,[2] suggesting the presence of a bias in the dataset. Evaluations of SAR ATR techniques are currently challenging due to the lack of publicly available data in the SAR domain. In this paper, we present a high resolution SAR dataset consisting of images of a set of ground military target models taken at various aspect angles, The dataset can be used for a fair evaluation and comparison of SAR ATR algorithms. We applied the Inverse Synthetic Aperture Radar (ISAR) technique to echoes from targets rotating on a turntable and illuminated with a stepped frequency waveform. The targets in the database consist of four variants of two 1.7m-long models of T-64 and T-72 tanks. The gun, the turret position and the depression angle are varied to form 26 different sequences of images. The emitted signal spanned the frequency range from 13 GHz to 18 GHz to achieve a bandwidth of 5 GHz sampled with 4001 frequency points. The resolution obtained with respect to the size of the model targets is comparable to typical values obtained using SAR airborne systems. Single polarized images (Horizontal-Horizontal) are generated using the backprojection algorithm.[3] A total of 1480 images are produced using a 20° integration angle. The images in the dataset are organized in a suggested training and testing set to facilitate a standard evaluation of SAR ATR algorithms.

Keywords: SAR, ISAR, ATR, Dataset, MGTD

1. INTRODUCTION

SAR images are an attractive option for target classification as they can be collected from an airborne platform flying over the surveyed area in darkness and under all weather conditions. Automatic target recognition on SAR images is still under development because of the complexity of the task and the lack of readily available SAR data.

As SAR ATR algorithms are currently developed, the availability of datasets that allow a fair evaluation of the performance is essential. If various ATR methods could be tested on the same datasets, the results could be better compared and improvements could be established more effectively. A dataset of this kind should represent a variety of operating conditions, such as different target viewing angles, different target configurations, layovers or occlusions induced by the surrounding environment, as well as target movements effects. With a more realistic dataset the likelihood of over fitting is decreased because the images are less similar with each other. The database that has been most often used to test SAR ATR algorithms is the MSTAR.[1] The problem with the MSTAR is that only 3 targets are recommended for use in ATR with only a single sequence each used for

Further author information: (Send correspondence to Carole Belloni.)
Carole Belloni: E-mail: c.d.belloni@cranfield.ac.uk
Alessio Balleri: E-mail: a.balleri@cranfield.ac.uk

Target and Background Signatures III, edited by Karin U. Stein, Ric Schleijpen,
Proc. of SPIE Vol. 10432, 104320N · © 2017 SPIE CCC
code: 0277-786X/17/$18 · doi: 10.1117/12.2277914

training.[4] The independence between training and testing sets in the MSTAR has also been questioned,[2] but is essential to achieve reliable results in any recognition problems. As a result of these limitations, ATR algorithms could have shown artificially increased results and their performance not accurately reported.

This paper presents a new dataset dedicated to SAR ATR. The experimental setup and all the environmental factors are presented in section 3.1. The algorithm used to create the ISAR images is discussed in section 3.3. Guidelines are also given to correctly select the sequences to form independent and varied training and testing sets in section 4.2. In section 4.3, the best variant to choose as a confuser is discussed. The confuser is a target that is not present during training but only during testing to confuse the algorithm. Following this discussion, an evaluation method is suggested, so that all algorithms are evaluated in the same way.

Table 1: Summary of MGTD (Military Ground Target Dataset) dataset

Frequency range	X-Band: 13 GHz to 18 GHz sampled with 4001 frequency points
Resolution	3.0 cm (range)×3.3 cm (cross-range)
Target class / variant number	2 classes (T64, T72) and 4 variants (T64n, T64f, T64s, T72))
number of images	1512 images from 26 distinctive sequences (Training: 1152. Testing: 360.)
Recommended confuser	T64f
Training/Testing set	Training: 11° and 12° depression angle. / Testing: 13° depression angle

2. REASONS FOR CREATING A NEW DATASET

Algorithms performing ATR rely on a training set to recognize targets in a testing set. To ensure a fair analysis of the performance of these algorithms and avoid any possible bias in the results, training and testing images should be taken from independent sets of data.

It has been shown that the MSTAR[1] contains data with a high degree of correlation between images in the training and testing set due to the presence of correlated background.[2] Indeed, it has been demonstrated that the recognition rates of algorithms tested on the MSTAR are high even when the target to recognize is artificially hidden. Moreover, data released to the public and included in the guidelines[4] contains only two targets (BMP, T72) with a complete training and testing sets, and one, the BTR60, with only one sequence for each set. We define a sequence as a group of images obtained from one single experiment. Images from a single sequence have thus identical environment factors except from their orientation.

To avoid correlation between the training and testing set in the presented dataset, The images belonging to the training and testing set are separated so that the sets of environmental conditions are different for training and testing images. Each target has also training images with varied environmental conditions so that the algorithm can be more resilient.

3. DATASET ACQUISITION

3.1 EXPERIMENTAL SETUP

3.1.1 CONFIGURATION OF THE EXPERIMENT

The antenna is placed 5 m away from the target on an adjustable height tripod to allow measurements or depression angles between 11° and 13°. The target is on a rotating turn-table and high range resolution profiles are measured every predefined angular step. A single Horizontal-Horizontal polarization is used. To avoid any ring due to unwanted movements of the setup, a latency period is introduced after each rotation step to ensure the target is still before each measurement. The emitted signal is a stepped frequency waveform spanning a bandwidth of 5 GHz between 13 GHz and 18 GHz, with 4001 frequency points. The signal is generated and acquired using an Anritsu vector network analyzer (VNA). A piece of radiation absorbent material is laid in front of the turn-table (as shown in Fig.2c) in order to prevent unwanted multipath effects from the turn-table.

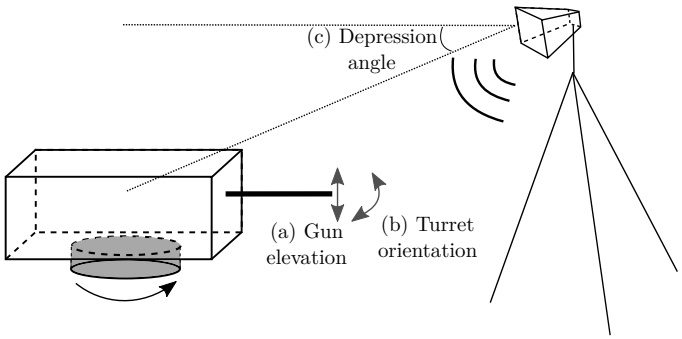


Figure 1: Sketch of the experimental setup. The antenna emits a signal towards the target placed on a turntable at each rotation step. For each sequence of measurements, at least one of the following factor is changed: the position of the gun (up/down), the orientation of the turret and the depression angle between the antenna and the target.

3.1.2 TARGETS

The presented database includes 2 classes and 4 variants. Their image representation of the target depends on the targets signature, i.e., the way it backscatters the energy sent by the radar. The representations of the classes will be easier to distinguish than the representations of the variants.

CLASSES: T64, T72

The classes categories are characterized by major signature changes. The first class is the T64 tank shown in Figs.2a, 2b, 2c The second class is the T72 tank shown in Fig.2d. If the T64 and T72 are defined as the only target, the algorithm will likely have a very high recognition rate.

VARIANTS: T64n, T64f, T64s, T72n

The variants have only minor differences with each other. The variants are at a level of differentiation beyond the classes. Among the class T64, it is possible to distinguish the following variants: "normal" T64n in Fig.2a, "side difference" T64s in Fig.2b and "flat plates" T64f in Fig.2c. The minor changes are induced by the addition or displacement of scatterers (flat plates, spheres). The class T72 has only one variant which is T72n (identical to the T72, but called T72n to be consistent with the other variants' names). The classification of variants proves to be a lot more challenging than that of the classes. We will thus refer to this target separation thereafter.

(a) T64n (b) T64s (c) T64f (d) T72n

Figure 2: The different target classes and variants.

3.1.3 ENVIRONMENTAL VARIABLES

ORIENTATION

The MGTD dataset consists of target images taken every 5° for the training set and every 10° for the test set. The training and testing sets are formed using independent image sequences collected under different environmental conditions. A further description can be found in sections 4.1 and 4.2. A smaller angular separation was chosen for the training set to allow the generation of a larger number images and hence cater for some deep learning methods that require a high number of images for training. A larger separation was adopted for the testing images in order to avoid possible correlations in the results.

DEPRESSION ANGLE

The depression angles used to generate the dataset are 11°, 12° and 13°. The depression angle is changed by adjusting the height of the antenna mount. The reflecting surfaces that are part of the target have thus virtually a new orientation and that impacts on the way the signal is backscattered and leads to a change of the target signature.

TARGET CONFIGURATION CHANGES

We define the configuration change as the displacement of an element of the target. In practice, it is the turret or gun direction that is considered as a configuration change. All orientations of the turret against its central position are included in the following set of angles: [-90; -45; -30; 0; 30; 45; 90]. The gun had only two discrete positions which are up and down.

3.2 PARAMETERS OF THE EXPERIMENT

The choice of the bandwidth determines the range resolution of the image. With a bandwidth of 5 GHz, the range resolution is:

$$\Delta r = \frac{c}{2B} = 3.0\,\text{cm}\,, \tag{1}$$

where c is the speed of light and B is the bandwidth of the signal.

An integration angle of 20° seemed realistic as it represents a 350 m synthetic antenna for an altitude of 1 km. To have a similar resolution in the cross-range given this integration angle, we choose the start frequency of 13 GHz. Consequently, the cross-range resolution is of 3.3 cm using Eq.2.

$$\Delta x = \frac{c}{2\theta_a f_{min}} = 3.3\,\text{cm}\,, \tag{2}$$

where Δx is the cross-range resolution, c is the speed of light, θ_a is the full integration angle and f_{min} is the start frequency of the signal.

This resolution is equivalent to a 17 cm resolution on a real-size tank (9.53 m T72) which is a value achievable with existing airborne SAR.

Knowing that the model-tanks described in section 3.1.2 are 1.7 m long, an estimation of 3 m for the size of the total rotating scene seems adequate. We calculate here the maximum range size of the scene.

$$W_r = \frac{c}{2\Delta f} = 120\,\text{m}\,, \tag{3}$$

where W_r is the maximum range size of the scene, c is the speed of light, Δf is the frequency step. We calculate then the maximum cross-range size of the scene, that must satisfy the same criteria as the range size of the scene..

$$W_x = \frac{c}{2\Delta\theta f_{min}} = 3.3\,\text{m}\,, \tag{4}$$

where W_x is the maximum cross-range size of the scene, c is the speed of light, $\Delta\theta$ is the step angle and f_{min} is the start frequency of the signal.

Reflecting the equipment constraints and the previous estimation, we choose a rotation step of 0.2° that allows a 3.3 m maximum scene size (Eq.4), which is large enough to contain the tank model.

3.3 IMAGE GENERATION

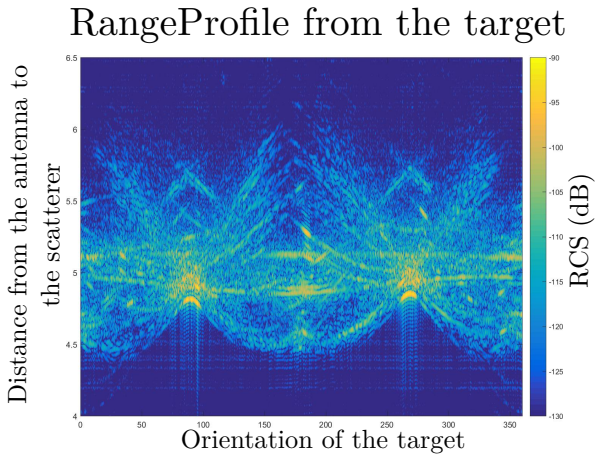


Figure 3: Radar cross section (RCS) of the target for each orientation of the target per range cell.

The target range profiles in Fig.3 are obtained by taking the ifft of the VNA output . For each target aspect angle, the backscattered signal amplitude is plotted for every range cell to obtain the sinogram of the main target scatterers. The range cells that contain the target are thus selected. Results show that strongest echoes from the target are between 4.25 m and 6.25 m. The maximum size of 3 m is consistent with the target size.

The returned frequency signal from the VNA is compensated for the delay induced by the extra travelled distance from the center of rotation to each pixel. This is done for each pixel of the image, for each frequency and pulse as in Eq.5. This relies on the widely used assumption that scatterers are isotropic and that the backscattered signal has the same intensity for each pulse.

$$S_c(f_k, \tau_n) = S(f_k, \tau_n) e^{\frac{4j\pi f_k \Delta R(\tau_n)}{c}}, \tag{5}$$

In the equation, S_c is the compensated signal for the range cell τ_n and f_k is the frequency bin, S is the measured frequency signal and $\Delta R(\tau_n)$ is the distance between the test pixel and the center of rotation. The pixel intensity is obtained as

$$I = \frac{1}{N_p K} \sum_{n=1}^{N_p} \sum_{k=1}^{K} S(f_k, \tau_n) e^{\frac{4j\pi f_k \Delta R(\tau_n)}{c}}, \tag{6}$$

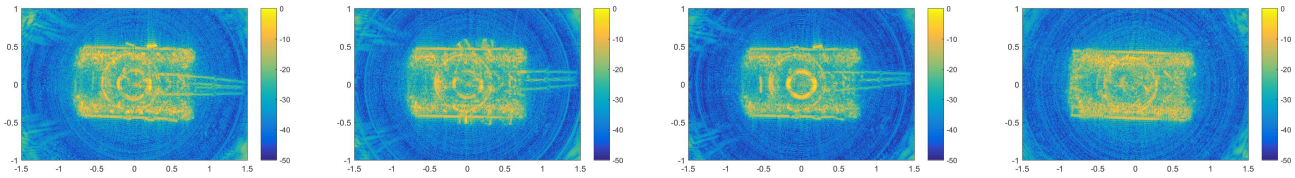
The backprojection method[3] is used to generate the images from the radar data.

3.3.1 TARGETS ISAR REPRESENTATION

To illustrate the major and minor signature changes between classes and variants respectively, the full 360° SAR images from each target generated with the backprojection algorithm are shown in Fig.4. There are minor signature changes between Figs.4a, 4b, 4c. These minor changes are composed of the side flaps of the target in Fig.4b or the reflecting plate at the back of the target in Fig.4c compared to the bare T64n seen in Fig.4a . However, there is a clear distinction from all the T64 variants to the other class T72 in Fig.4d. The 360° shows that the confusion between Fig.4d and one of Figs.4a-4c is a lot less likely than between the Figs.4a-4c.

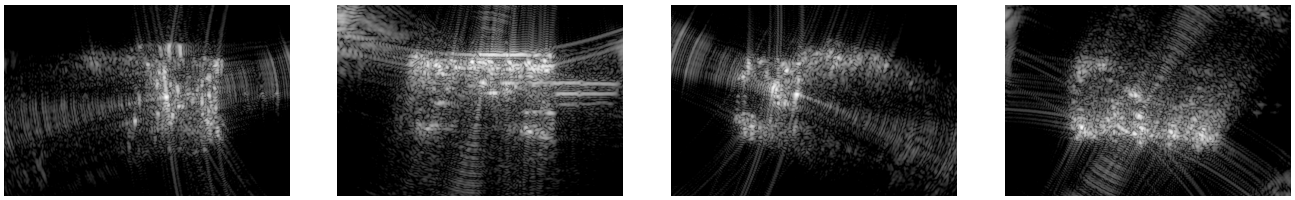
3.3.2 INFLUENCE OF THE ENVIRONMENTAL VARIABLES ON THE ISAR IMAGES ORIENTATION

One of the limitations of SAR images is that small changes in the orientation of the target can result in vastly different signatures.[5] Fig.5 shows different images from one unique sequence. The target is the T64n and apart



(a) T64n (b) T64s (c) T64f (d) T72n

Figure 4: SAR images of the different target classes and variants with a 360˚integration angle.



(a) Target orientation: 10˚ (b) Target orientation: 80˚ (c) Target orientation: 160˚ (d) Target orientation: 240˚

Figure 5: Impact of the target's orientation on the SAR images.

from the orientation, all the environmental parameters are constant. For each aspect angle, only the scatterers visible to the radar with a high enough RCS will be included in the SAR image. Fig.5b shows the ISAR image of the T64n target relative to an aspect angle of 80˚ obtained by processing the echo signals from 70˚ to 90˚ aspect angle. Results show that the gun is only visible around 80˚ and disappears in Figs.5a, 5c and 5d even though all the environmental conditions remained the same.

DEPRESSION ANGLE

The representation of scatterers in the produced ISAR image can be different with various depression angle. In the most extreme cases, it can be seen at a certain depression angle and disappear at another. Figs. 6a and6b show the ISAR images of the T64n with the same orientation and the same target configuration at different depression angles. As the depression angle increases, the scatterers have a new orientation relative to the antenna and the gun is less noticeable from the resulting ISAR image.

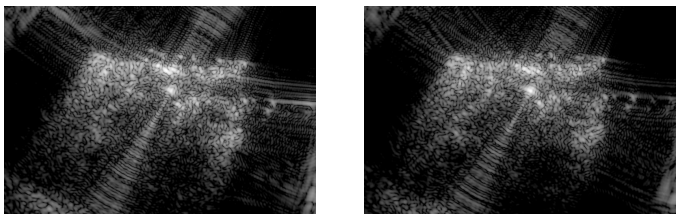
TARGET CONFIGURATION CHANGES

Fig.7 images from the T64n at the same aspect angle but with a different turret position. Results show that the signature changes as a function of the position of turret, and this should be taken into account when ATR is performed.

4. DATASET USE FOR SAR ATR

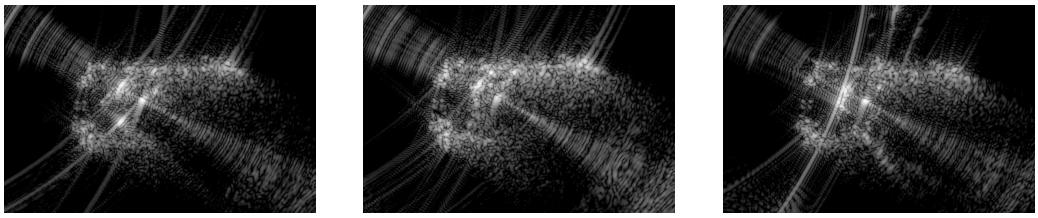
4.1 IMAGE DATASET DESCRIPTION

The total dataset consists of 26 different sequences, of which 16 are used for training and 10 for the testing. These are made out of 4 different variants with different depression angles and configurations. Each training sequence is composed of 72 images of the target obtained with a 20˚ integration angle every 5˚ to cover 360˚ of possible orientations for a total of 1152 $(72 * 16)$ images. Each sequence represent thus one unique target with different aspect angle but the same depression angle and target configuration. Each testing sequence is composed of 36 images every 10˚ for a total of 360 $(36 * 10)$ images. There are thus 1512 images in total in this dataset. Table 2 summarizes the different turret angles for each sequence taken with a specific target and depression angle. Each sequence is labelled with an identifier following this nomenclature : *x-vd* where x is the number of the sequence (1-26), v is the variant of the target (T64n, T64s, T64f, T72n) and d is the height of the antenna, linked to the depression angle of the sequence (l for 11˚, m for 12˚, h for 13˚). An example of a sequence name could be *1-T64sh*. In this case, the sequence number is 1, the target used is the T64s and the depression angle is 13˚.



(a) Depression angle: 11° (b) Depression angle: 12°

Figure 6: Impact of the depression angle on the SAR images.



(a) Turret angle: −45° (b) Turret angle: 0° (c) Turret angle: 90°

Figure 7: Impact of the configuration on the SAR images.

4.2 SEPARATION IN TRAINING AND TESTING DATASET

Extended operating conditions have been used to produce a diverse dataset in terms of target types, orientation, configuration and depression angle. This variety of conditions will be used to minimize the degree of correlation between the training set and the testing set. The sequences used for training and testing must be acquired differently. The criterion that was chosen to be different between the sets is the depression angle because the images look different when it is varied and because it leaves a reasonable number of images in both sets. The recommended training sequences are those with a depression angle of 11° and 12° consisting of 16 sequences, 4 for each target type. The testing sequences were obtained with a depression angle of 13°. There are 10 total sequences, with 2 to 3 sequences for each target. All measured configurations and relative parameters of both training and testing set are presented in Table 2.

4.3 PERFORMANCE QUANTIFICATION

4.3.1 CORRECT CLASSIFICATION RATE AND USE OF A CONFUSER

The performance evaluation only deals with target recognition. Detection is out of the scope of this paper as the images are already focused on the targets. The recognition step consists in correctly labelling targets as defined in 3.1.2. ATR methods performance are measured using the probability of correct classification P_{cc}. This is the ratio between the number of correctly labelled targets over the total number of targets. If a rejection class is used when the ATR is not able to label the target, the probability of false positive P_{fp} is also analysed. This represents the number of wrongly labelled targets over the total number of target tested. If the target is labelled as the rejection class, it is neither correctly, nor wrongly classified. The presence of a rejection class allows a reduction of false positives. These scores should be given when using the dataset separation presented in section 4.2.

The ability of the ATR algorithm to use the rejection class can be further investigated by using a confuser. A confuser class is not included in the training set and thus all images that belong to it should not be recognized during testing. We recommend to use the target T64f as a confuser to evaluate the ATR algorithms. It appears indeed that T64n and T64s targets are regularly confused with T64f and thus using the T64f as a confuser makes the choice between the training targets and the rejection class the most challenging. The training images from T64f should be removed from the dataset presented in section 4.2. Images from the testing set representing the T64f target are then considered correctly labelled is the ATR algorithms classify them with the rejection class as the target is unknown to it.

Target type \ Depression angle	Training set		Testing set
	11°	12°	13°
T64n	0°, 30°	0°, 30°	−45°, −45°, −90°
T64s	−30°, 30°	−30°, 30°	0°, 0°
T64f	−30°, 30°	−30°, 30°	−90°, 0°, 0°
T72	0°, 30°	−30°, 0°	−30°, 30°

Table 2: Sequences of images and orientation of the turret of the imaged target. Training and testing separation of the sequences.

Some ATR methods do not have the ability to dismiss targets in complicated images to a rejection class. In this case, the probability of correct classification under forced-decision can still assess the algorithm performance. This is the same definition as the P_{cc} except that the target has to be classified in a native target type as the rejection class is non-existent. This score has little meaning if a confuser is introduced as it will be wrongly classified in any case.

For every testing of the ATR, the detailed confusion matrix should be given as this gives the possibility to assess, in addition to the P_{cc}, the inter-class classification performance. An inter-class classification does not take into account the variants (T64n, T64s, T64f, T72n) but only the classes (T64, T72).

4.3.2 REQUIREMENTS

To have a global picture of an ATR algorithm, it is also interesting to know its requirements. The amount of storage needed after training to describe the set of targets is the space requirement of the algorithm. The processing power required to run the ATR algorithm can be described with the total training and testing computation time while specifying the processor and graphic card if one is needed as well as the amount of random-access memory (RAM) used during the computation. This will estimate the computation load of the ATR algorithm.

5. CONCLUSION

In this paper a new dataset of SAR images is presented that can facilitate the evaluation and comparison of SAR ATR algorithms. The choice of parameters and the process leading to the acquisition of the images are explained. The operating conditions changed throughout the whole dataset to introduce variability in the dataset are also described. For each target type, different orientations, configurations and depression angles are used for each sequence. Guidelines are also provided to test fairly ATR algorithm. A suggestion of separation between the training and testing set is given. Standard indicators are recommended for a baseline evaluation that should help for future comparisons of ATR methods.

REFERENCES

[1] DARPA and AFRL, "Sensor data management system website, MSTAR database." https://www.sdms.afrl.af.mil/index.php?collection=mstar (1995). Online; accessed 2016-07-12.

[2] Schumacher, R. and Rosenbach, K., "ATR of battlefield targets by SAR classification results using the public MSTAR dataset compared with a dataset by QinetiQ UK," in [*RTO SET Symp. Target Identification and Recognition Using RF Systems*], Citeseer (2004).

[3] Gorham, L. A. and Moore, L. J., "Sar image formation toolbox for matlab," in [*SPIE Defense, Security, and Sensing*], 769906–769906, International Society for Optics and Photonics (2010).

[4] Ross, T. D., Worrell, S. W., Velten, V. J., Mossing, J. C., and Bryant, M. L., "Standard sar atr evaluation experiments using the mstar public release data set," in [*Aerospace/Defense Sensing and Controls*], 566–573, International Society for Optics and Photonics (1998).

[5] Keydel, E. R., Lee, S. W., and Moore, J. T., "Mstar extended operating conditions: A tutorial," in [*Aerospace/Defense Sensing and Controls*], 228–242, International Society for Optics and Photonics (1996).

SAR image dataset of military ground targets with multiple poses for ATR

Belloni, Carole

2017-10-05

Belloni C, Balleri A, Aouf N, Merlet T & Le Caillec JM (2017) SAR image dataset of military ground targets with multiple poses for ATR. In: Proceedings Volume 10432: Target and Background Signatures III, 2017, Warsaw, 11-12 September 2017.

https://doi.org/10.1117/12.2277914