TARGET DETECTION IN SAR IMAGES USING THE MIST/AQ METHOD

by

Q. Zhang
Z. Duric
R. S. Michalski

Reports of the Machine Learning and Inference Laboratory, MLI 96-12, George Mason University, Fairfax, VA, 1996.

TARGET DETECTION IN SAR IMAGES USING THE MIST/AQ METHOD

Qi Zhang   Zoran Duric   Ryszard S. Michalski*

Machine Learning and Inference Laboratory
George Mason University

*Also with GMU Departments of Computer Science and
Systems Engineering, and
the Institute of Computer Sciences, Polish Academy of Science

P96-28
MLI 96-12

November 1996

TARGET DETECTION IN SAR IMAGES USING THE MIST/AQ METHOD

ABSTRACT

This paper describes a novel application of the MIST methodology to target detection in SAR images. Specifically, a polarimetric whitening filter and a constant false alarm rate detector are used to preprocess a SAR image; then the AQ15c learning program is applied to learn and detect targets. Encouraging and impressive experimental results are provided showing the effectiveness of the MIST/AQ method for solving this problem.

KEYWORD: Learning in vision, target detection in SAR images, MIST methodology.

ACKNOWLEDGMENTS

We would like to thank Dr. Maloof for his help in this work, Dr. Kaufman for his technical support in preparing image data, Jim Mitchell for his review. This research was conducted in the Machine Learning and Inference Laboratory at George Mason University. The Laboratory's activities are supported in part by the Defense Advanced Research Projects Agency under grant F49620-95-1-0462, administered by the Air Force Office of Scientific Research, in part by the National Science Foundation under grants DMI-9496192 and IRI-9020266, and in part by the Office of Naval Research under grant N00014-91-J-1351

1 INTRODUCTION

Vision is one of the most important senses of human beings; much of our information can be acquired only through vision. Because of this, machine vision is always a intensely studied area in artificial intelligence. Research work on vision shows that it is often necessary for a flexible and robust vision system to incorporate learning capabilities. This has become an important research direction (Michalski et al., 1992; Bhanu & Poggo, 1994). This paper is focused on detecting targets in SAR images.

A synthetic aperture radar (SAR) transmits pulses of radio waves which bounce off the targets/objects to be depicted. The scattered pulses return to the radar, where they are captured by the receiving antenna. Since a SAR is an active remote sensing system and an all-weather imaging device, it is able to provide good images of what it has detected even in fog, clouds, or darkness in which normal optical sensors are useless. Detecting and recognizing targets/objects in SAR images is of much significance both in military and civil applications.

Successful detection of targets in SAR images is difficult because there is a large amount of noise in image data. This paper presents a novel machine learning approach to this problem by using the MIST/AQ method (Michalski et al., 1996). MIST is a general framework that provides an environment for applying diverse machine learning methods to problems in computer vision. Its central idea is to combine advanced inductive learning methods with various relevant knowledge in multi-level image analysis and interpretation. In this paper, a SAR image is first processed by a polarimetric whitening filter (PWF) to improve image quality (Novak et al., 1990), and then it is input to a constant false alarm rate (CFAR) detector for screening its natural clutter and detecting potential targets within it (Ravid & Levanon, 1992). The output from the CFAR detector is used by AQ15c for training and detection. The SAR images we used were provided to us by DARPA and were collected in Stockbridge New York, by the Lincoln Laboratory at MIT. The experimental results were very encouraging and impressive.

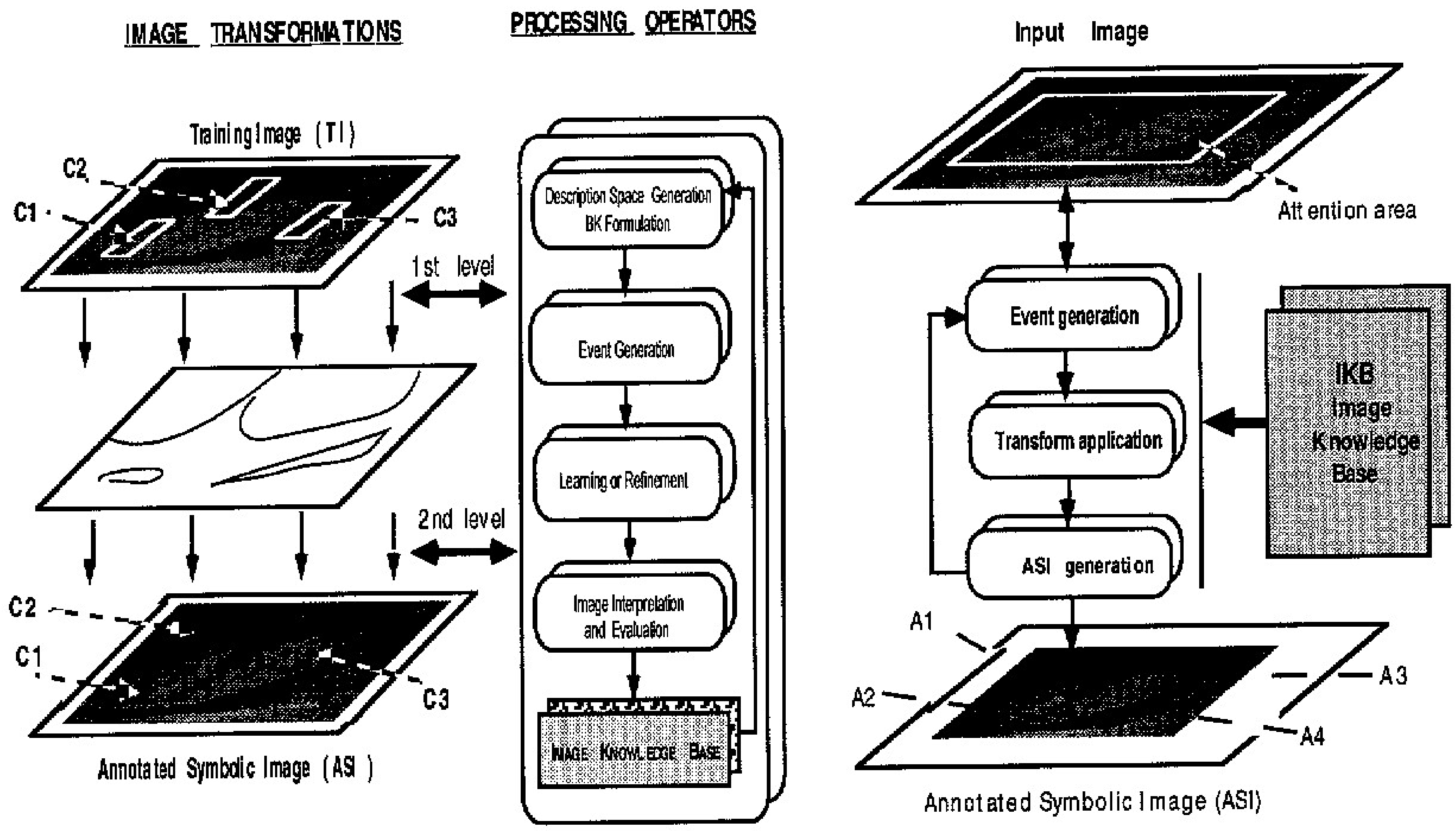
This paper is structured as follows: section 2 reviews some related work on target detection and learning, section 3 describes the MIST methodology , section 4 introduces pre-processing of a SAR image and the issues related to learning for detection, section 5 discusses experimental results and section 6 concludes this paper.

2 RELATED WORK

Target detection in SAR images is difficult because of large amounts of noise in image data. Kreithen et al. (1993) used the polarimetric whitening filter and a CFAR detector to preprocess SAR images. After grouping sets of clustered pixels (potential targets) which were seemingly from the same targets, attribute values were generated for each of them and a quadratic distance for each potential target was calculated and compared to a threshold determined previously by experiments. Obviously, a single threshold is not so understandable or flexible as symbolic knowledge descriptions. Besides, determining which pixels came from a target or not and then grouping them is often difficult or impossible due to noise. Burl et al. (1994) used the matched filter technique to detect potential volcanoes in Venus SAR images. Matched filters for each kind of volcano were constructed from training volcano examples and then were applied to scan an image pixel by pixel to locate potential volcanoes. The matched filter is possibly subject to rotation, size and other vision condition changes. Further, using matched filter to scan a whole image and computing the degrees of match with each constructed filter is time-consuming.

Application of machine learning techniques to target detection in images is relatively new. Rong and Bhanu (1996) adopted the reinforcement learning method. The training and testing were directly performed on raw FLIR (forward-looking infrared radar, not SAR) images without any transformation. They divided an image into many rectangulars which were the input unit to a learning system. Directly training and testing on raw data could consume more time and generate harder learning problems and dividing the image into small areas for training could lead to incompleteness of target information. Spence (1996) touched the problem of uncertain data in detection and devised some measurements for detection decision in neural network with which experiments were done on artificial problems.

3 MIST METHODOLOGY



(a) (b)

Figure 1. The MIST methodology (a) training mode, (b) interpretation mode.

Among the most important research goals of applying machine learning methods to vision problems is to gain better understanding of matching appropriate learning methods to appropriate vision problems. MIST (Multilevel Image Sampling and Transformation) has been developed as a general methodology for applying machine learning methods to vision problems (Michalski et al., 1996). The purpose of MIST is to provide a researcher with an environment in which a variety of machine learning methods and approaches can be flexibly applied to a wide range of vision problems.

The MIST methodology works in two modes: *training mode* and *interpretation mode* (Figure 1). In the training mode, four steps are needed and, based on training performance, possibly repeated for better results. *Event generation* is to generate examples for learning or testing. *Description space generation* means that a trainer assigns concept names to areas in training images that contain concepts or objects to be learned. In the interpretation mode, three steps are executed. *Transformation application* is to apply the Image Knowledge Base (IKB) to examples generated from testing images to produce a transformed version which, for instance, could be a segmented image. The output is annotated symbolic images (ASI). In an ASI, areas or objects that correspond to the recognized concepts in the original image are marked by symbols (e.g., colors) denoting these concepts and linked to *concept annotations* (text containing additional information about that concept, such as degree of certainty of recognition, properties of the concepts, relation to other concepts, etc.). The output ASI in one level can be the end result (one-level training) or input to later levels (multilevel training) for better results (e.g., repeated training on the same natural scene image) or for other tasks (e.g., target recognition). The image knowledge base (IKB) contains prior or learned descriptions or visual concepts, image processing operators, attribute extraction operators and background knowledge relevant to image interpretation.

The MIST methodology is an extension of the Multilevel Logical Templates methodology, originally proposed by Michalski (Michalski 1973), implemented and developed later (Channic, 1988; Bala & Pachowicz, 1991; Michalski et al., 1993). Among the advantages of this methodology are the ease of applying and testing diverse learning methods in a uniform manner, the potential for implementing advanced and complex learning processes, the natural way of interpreting images. The MIST methodology has been applied to the semantic interpretation of natural scenes (Michalski et al., 1996) and detection of blasting caps in X-ray images (Maloof et al., 1996).

4 LEARNING TO DETECT TARGETS

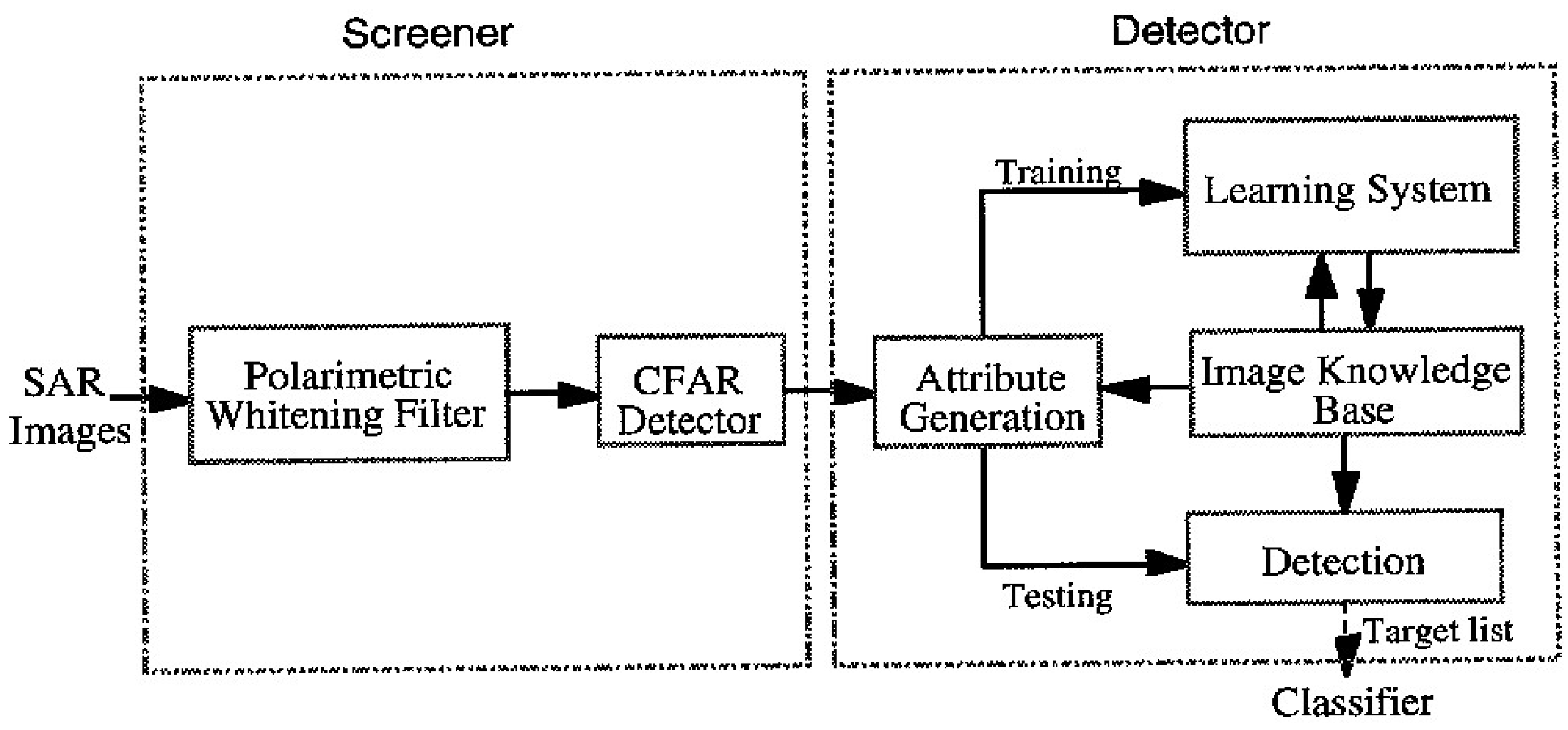


Figure 2. The system architecture.

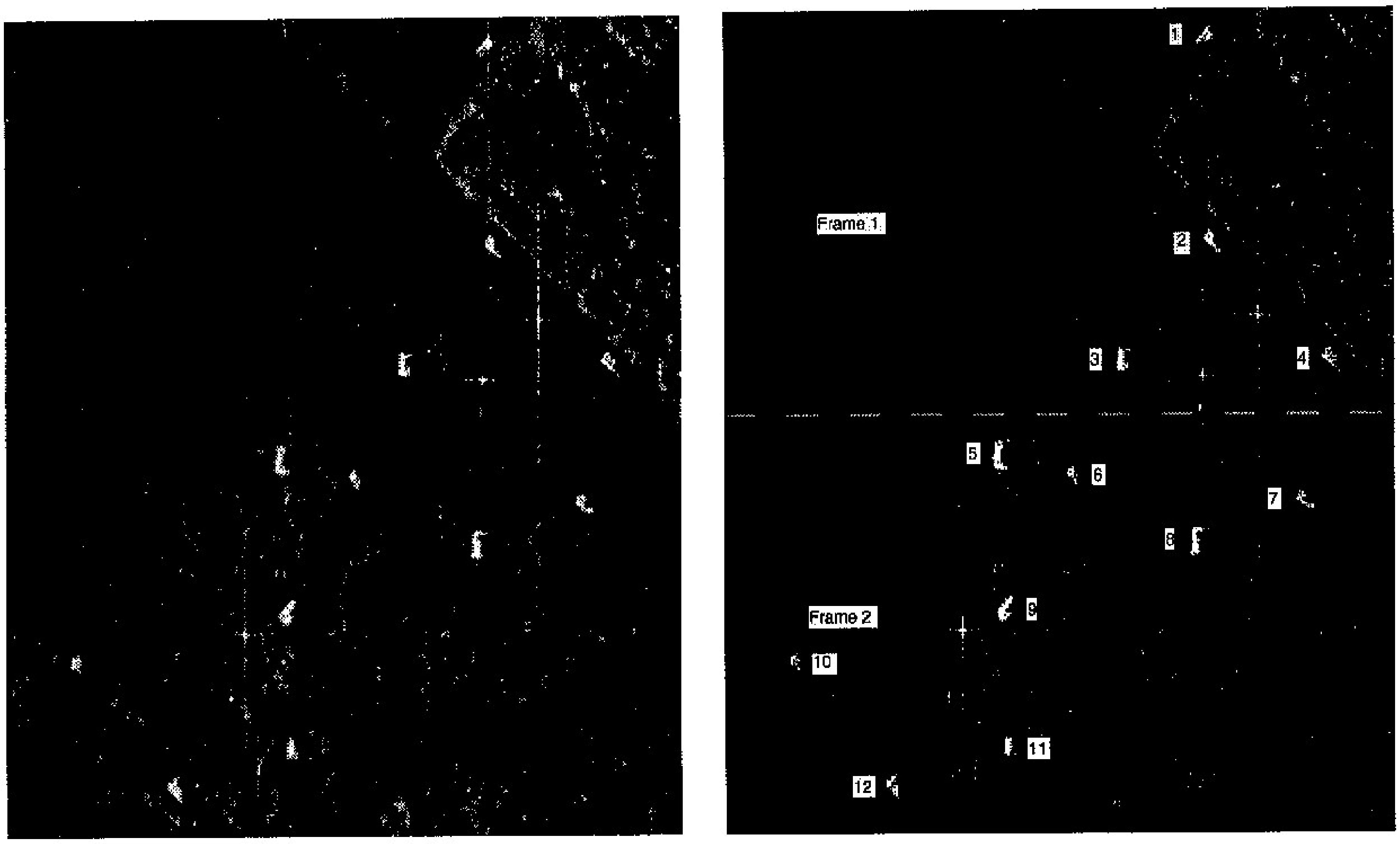
Guided by the MIST methodology, we designed a MIST/AQ system which consists of three sequential phases: enhancing image quality by a PWF, screening by a CFAR detector, and target learning and detection by AQ15c (Figure 2). The first two stages provide the preprocessing needed for further target detection and recognition.

4.1 Polarimetric Whitening Filter

A synthetic aperture radar can transmit pulses in either horizontal (H) or vertical (V) polarization and receive in either H or V, with the resultant combinations of HH (horizontal transmit, horizontal receive), VV, HV, or VH. Among these combinations, usually HH, HV and VV are used. The technique called the polarimetric whitening filter (PWF) (Novak et al., 1990) combines information in the HH, HV, and VV channels and improves image quality in two aspects: minimization of the amount of speckle (noise) and sharpening the edges of image objects. The power (i.e., the value of radar returns) of the PWF image (i.e., after the PWF processing) is determined by:

$$y(i,j) = |HH|^2 + \left|\frac{HV}{\sqrt{\varepsilon}}\right|^2 + \left|(VV - \rho^* \sqrt{\gamma} HH)\cdot(\gamma(1-|\rho|^2))^{-\frac{1}{2}}\right|^2$$

$$\varepsilon = \frac{E(|HV|^2)}{E(|HH|^2)},\quad \gamma = \frac{E(|VV|^2)}{E(|HH|^2)},\quad \rho = \frac{E(HH\cdot VV^*)}{\sqrt{E(|HH|^2)\cdot E(|VV|^2)}},$$



(a) (b)

Figure 3. An exemplary (a) PWF image and (b) CFAR image containing targets: 2,11: M60-tank; 3,7,9: M48-tank; 4,10: M84-APC; 6: M113-APC; 12: M59-APC; 5,8: M55-howitzer; 1: unknown.

where *(i, j)* is the pixel coordinates, X^* stands for a complex conjugate of X and *E()* represents the expected value or ensemble average. In our implementation of this filter, ε=0.19, γ=1.03, $\rho\sqrt{\gamma}$=0.53. Figure 3a is an exemplary PWF image.

We implemented this filter in our experiments without reducing image resolution; this is crucial for target detection and recognition. Note that there is a lot of other objects in PWF images, like trees, grass, etc., which are not targets and need to be eliminated.

4.2 CFAR Screening

Various constant false rate alarm (CFAR) algorithms (e.g., Ravad & Levanon, 1992; Armstrong et al., 1991; Wang et al., 1994) take a SAR image as input and perform a *screening* function, i.e., detecting potential targets in SAR images by examining intensities of radar returns and thereby providing a massive reduction of natural clutter (grass, trees etc.). Figure 3b is the CFAR image (i.e., processed by a CFAR detector) of Figure 3a. Note that due to noise inherent in SAR images, there are many false alarms which passed the screening of the CFAR detector. Determining which pixels came from a target and then grouping them in Figure 3b is not easy because pixels of false alarms are often connected and pixels from the same target can be isolated (also see Figure 6b).

4.3 Learning and Detecting

4.3.1 Definitions of attributes

In our work, the pixel is the operation unit in learning and detection. We consider each pixel that passed the CFAR detector as an example. For each example, produce a set of attribute values defined on a circular area (centered on the pixel or example of interest) from either its CFAR or PWF image, both of which can provide descriptive information about the example (Figure 4a). Two circular areas were used in this work, one with the diameter being 21 pixels and the other 31 pixels. We imported or designed the following attributes:

(1) Power: sum of intensities of all pixels in a circle of a PWF image.
(2) Deviation: standard deviation of the intensities of all pixels in a circle of a PWF image.
(3) Fractal dimension: a number within 0.0 and 2.0 which attempts to capture the spatial distribution of the brightest pixels of detected objects/targets in a CFAR image. It complements the standard deviation attribute, which depends only on the intensities and not on their spatial distribution. For a definition, see (Novak et al., 1993).
(4) Area size: the number of detected pixels in a circle of a CFAR image.
(5) Weighted-rank fill ratio: sum of the intensities of the top N (5% in our experiments) brightest pixels among all the pixels in a circle of a PWF image divided by the power of this circle.

The following examples give the reader a taste of what examples look like. Note that they are raw data (attributes ending in "1" represent attributes defined on a circle with ta diameter of 21 pixels, those ending in "2" are for circles with a 31-pixel diameter):

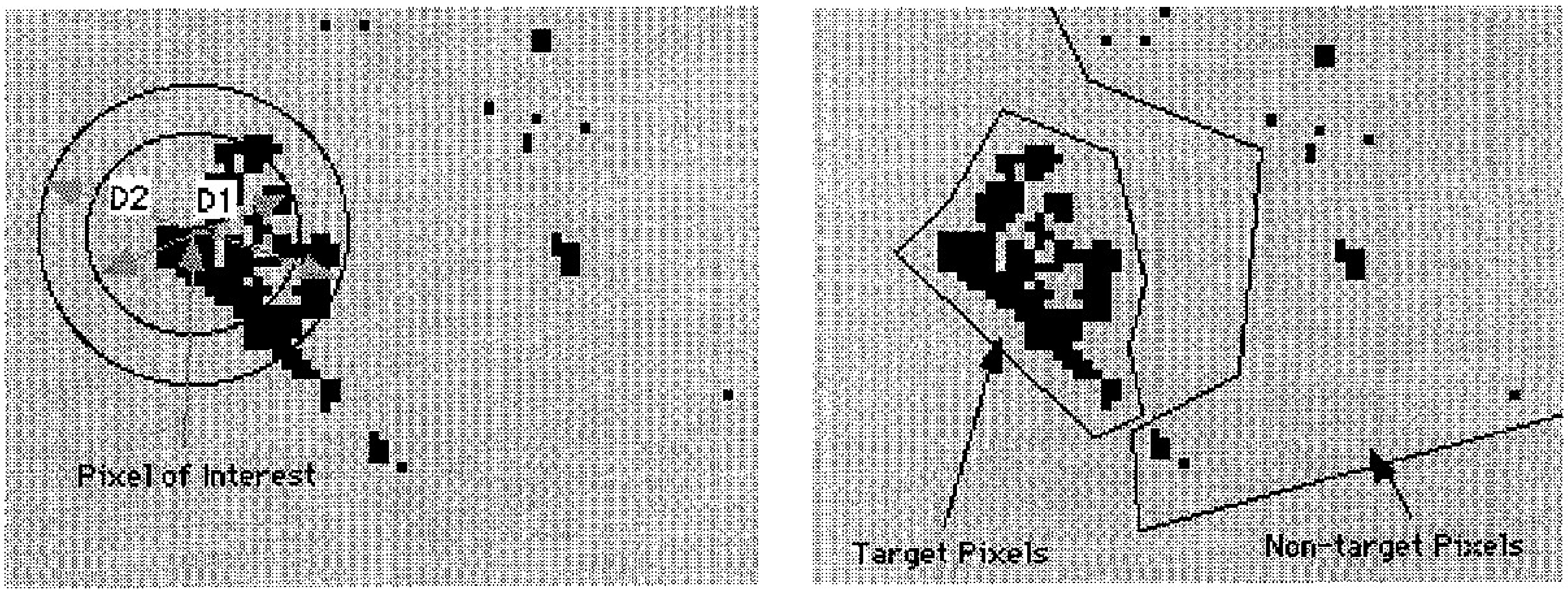


Figure 4. Enlarged target 4 in Figure 3b. (a) Attributes of a pixel are defined on circular areas; (b) Classification of target and non-target examples.

Target example

[power1=1091.466] [sd1=9.48] [fractal1=1.53] [area1=104] [wrfr1=0.54]
[power2=1498.467] [sd2=6.64] [fractal2=1.64] [area2=160] [wrfr2=0.555]

Nontarget example

[power1=206.816] [sd1=0.493] [fractal1=0.925] [area1=19] [wrfr1=0.154]
[power2=355.294] [sd2=0.424] [fractal2=0.938] [area2=23] [wrfr2=0.170]

4.3.2 Classification of training examples and discretization

Before learning, training examples must be classified as target and non-target examples. This is not easy, since it is impossible to accurately decide whether the pixels on the border of a target are target examples or non-target examples. Thus, there might be some misclassified examples for training. We adopted a simple rule: pixels connected (8-connectivity) to a target are target examples and otherwise non-target examples (Figure 4b).

Considering the characteristics of SAR image data, we adopted the Chi-merge discretization scheme (Kerber, 1992) and applied it to extracted raw data from CFAR or PWF images. The following are exemplary examples after discretization:

Target examples

(1) [power1=49] [sd1=24] [fractal1=7] [area1=3] [wrfr1=8] [power2=25] [sd2=10] [fractal2=6] [area2=3] [wrfr2=4]
(2) [power1=47] [sd1=24] [fractal1=7] [area1=2] [wrfr1=9] [power2=24] [sd2=10] [fractal2=7] [area2=3] [wrfr2=5]

Non-target examples

(1) [power1=6] [sd1=0] [fractal1=0] [area1=0] [wrfr1=0] [power2=0] [sd2=0] [fractal2=0] [area2=0] [wrfr2=0]
(2) [power1=0] [sd1=0] [fractal1=0] [area1=0] [wrfr1=2] [power2=0] [sd2=0] [fractal2=0] [area2=0] [wrfr2=0]

4.3.3 Training

The SAR image acquired during a single imaging process by an airplane or satellite is called a *pass*, and a SAR image can be divided into smaller parts called *frames*. Target detection in SAR images partitions the image objects into two classes: targets and non-targets. AQ15c is our learning and inference engine. The following are exemplary AQ rules acquired by taking pixels in frame 1 of Figure 4 as training examples.

Target examples
Rule1 [sd1=13..57] & [area1=0..10] & [wrfr1=0..9] & [power2=19..99] & [sd2=7..23] & [fractal2=3..15] & [wrfr2=0..8] (t-weight:842, u-weight:107)
Rule2 [area1=10..12] & [area2=0..20] (t-weight:742, u-weight:18)
Non-target examples
Rule1 [area1=0..9] & [area2=0..14] & [wrfr2=0..3] (t-weight:1561, u-weight:890)
Rule2 [power1=0] (t-weight:632, u-weight:8)

The t-weight in the above rules represents the total number of examples covered by a rule and the u-weight the number of examples uniquely covered by that rule. The larger weights are, the more stronger the pattern in a learned rule is.

5 EXPERIMENTAL RESULTS AND DISCUSSION

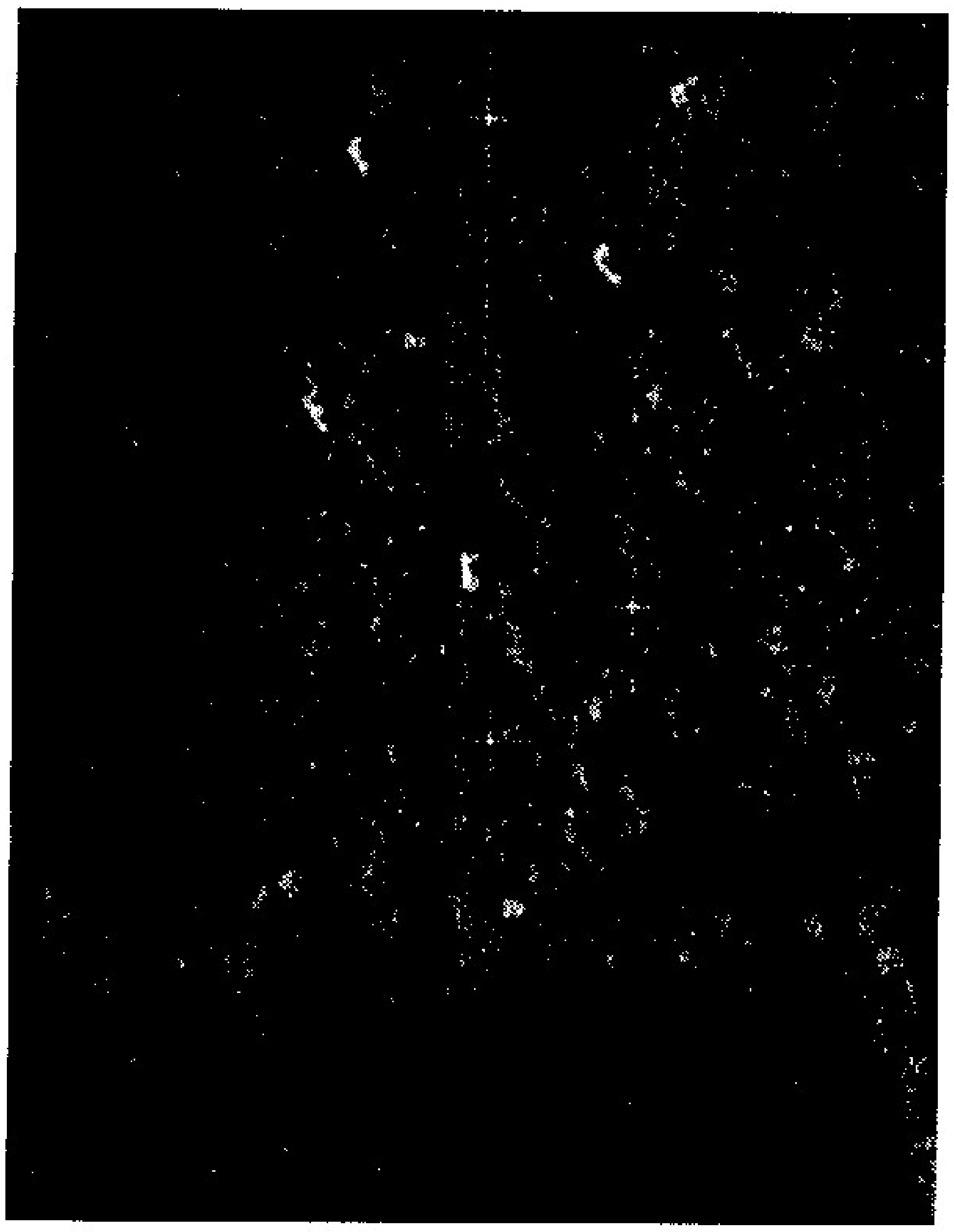
Experiments were done in fully polarimetric SAR images with 1ft x 1ft resolution. We tested all our SAR images whose ground truth is known. The experimental results are summarized in Table 1.

experiment	# training examples	# testing examples	# learned AQ rules	classification accuracy x100% (approx.)
1	3074	2121	26	98
2	3069	1706	12	98
3	9284	4915	19	98

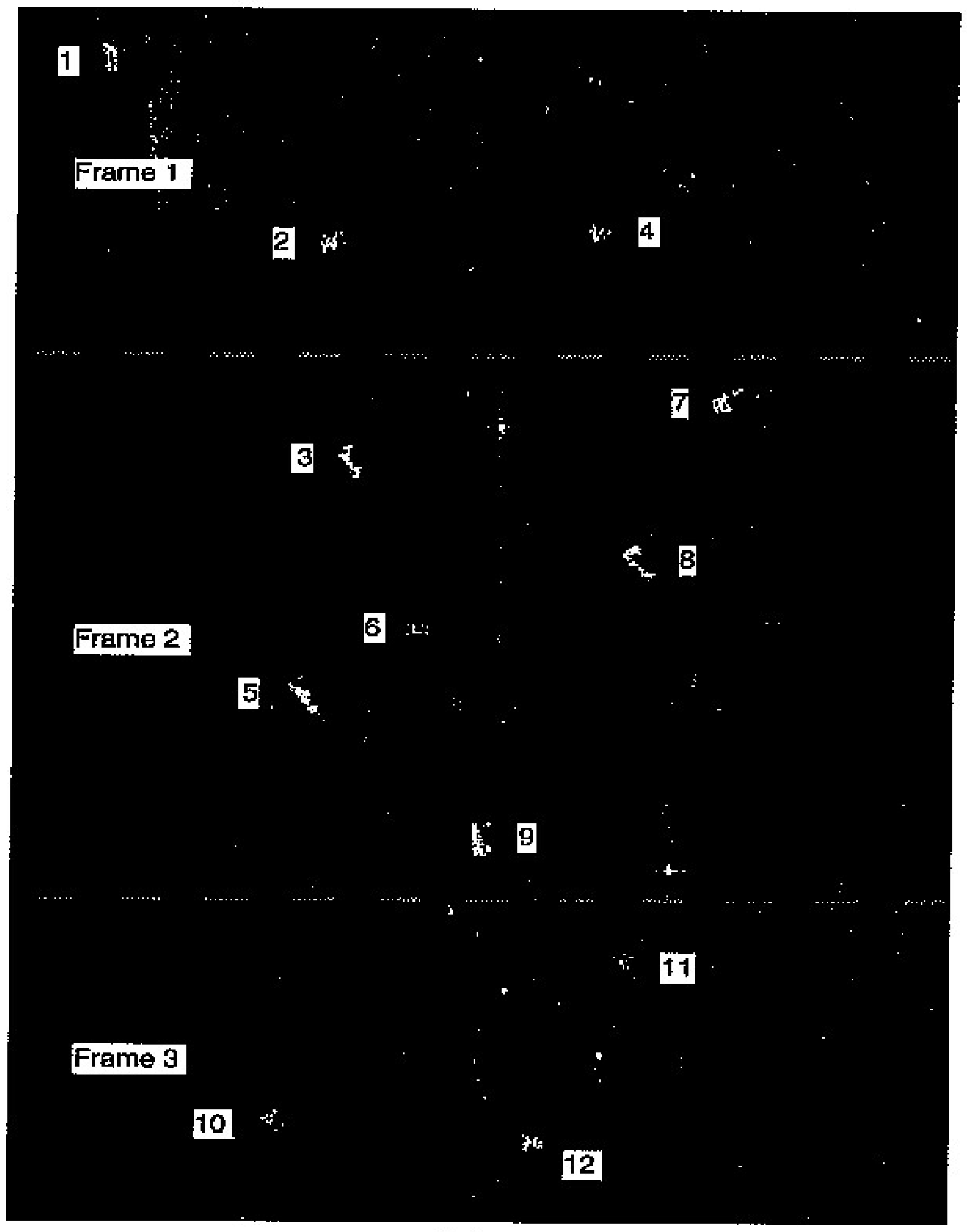
Table 1. Experimental results.

The first experiment was done from the perspective of machine learning to observe the performance of the defined attributes. It was a cross-validation experiment (Weiss & Kulikowski, 1992). All the target and non-target examples in the CFAR image in Figure 4b were put together and then randomly divided into two sets: 60% for training and 40% for testing. The accuracy in the first row of Table 1 shows that the selected attributes have discriminating power.

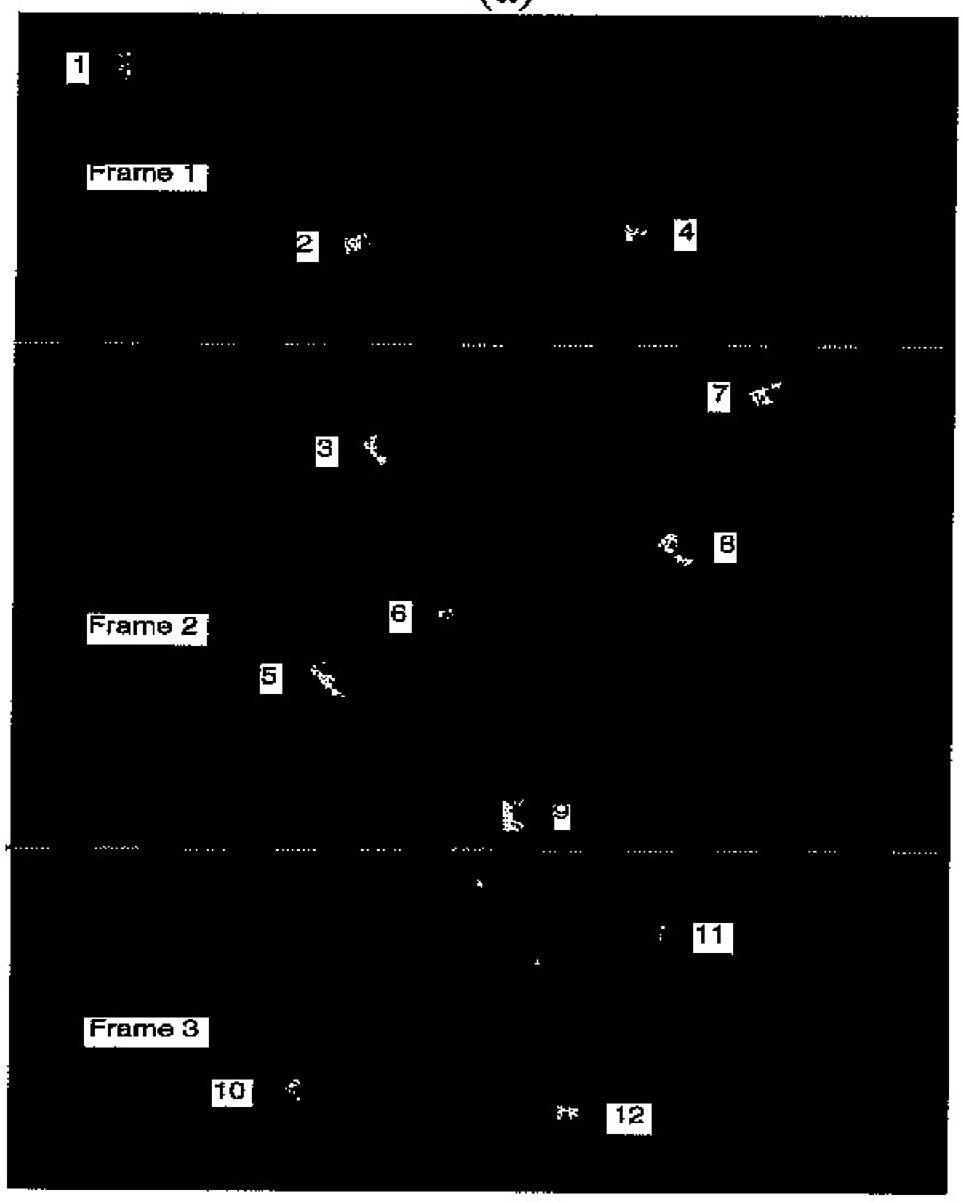
In the second experiment, training was executed on Frame 1 in Figure 3b and tested on Frame 2 in Figure 5b. The goal of this experiment was to see the performance of the rules learned in one pass but applied to a *another* place of a *different* pass. Note that targets M113-APC and M55-howitzer in Figure 5b were untrained new targets. The AQ detection results are very good, as can be seen in Figure 5c.



(a)



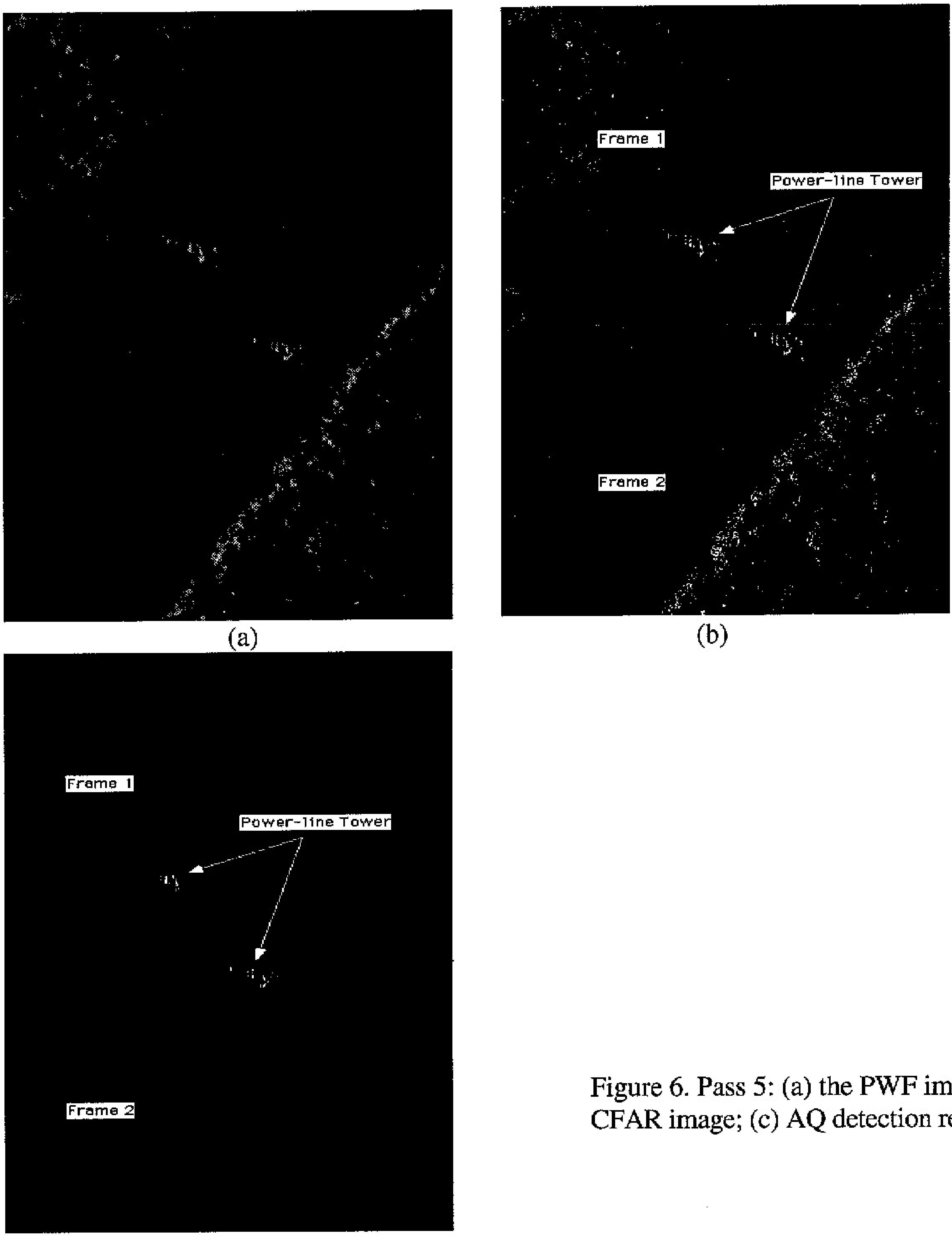
(b)



(c)

Figure 5. Pass 7: (a) PWF image, (b) CFAR image, (c) AQ detection results. Targets are: 2,11: M60-tank; 3,7,9: M48-tank; 4,10: M84-APC; 6: M113-APC; 12: M59-APC; 5,8: M55-howitzer; 1: unknown.

In the third experiment, in addition to Frame 1 in Figure 3b, the data in Frame 2 in Figure 6b was also used for training, in which there is a power-line tower. Frame 1 in Figure 6b was used for testing. As can be seen, AQ detection results are excellent (Figure 6c). Further, the rules learned in this experiment were retested in Figure 5b and the results were almost the same. This indicates that training on different targets and non-targets do not have mutually adverse effects. In fact, we did anther experiment in which only the data in Frame 1 in Figure 3b were used for training and the testing results on Figure 6b were still good (not shown here due to restriction on paper length); however, with one tower for training, the results are better. Note that immense number of false alarms in Figure 6b indicate the difficulty in determining and grouping pixels adopted by Kreithen et al. (1993).



(a) (b) (c)

Figure 6. Pass 5: (a) the PWF image; (b) the CFAR image; (c) AQ detection results.

The first two experiments indicate that AQ was able to capture the patterns among data and that learned knowledge was successfully applied to testing images, even to untrained targets. The results of the third experiment are interesting and important because it shows the necessity of learning. It can be easily seen from Figure 6b that it is virtually impossible to remove non-target examples by using vision techniques like the size filter or majority voting (Michalski et al., 1996). In these experiments, we attempted to vary vision conditions and to thus generate harder learning problems; nonetheless, the experimental results remained good, showing the effectiveness of AQ15c for detecting targets in SAR images.

An outstanding aspect in our experiments is that all targets remained after AQ detection while false alarms were maximally reduced, almost to zero. Another thing worth mentioning is that even though there was noise in our training data the results were still satisfactory. A possible way of avoiding data noise and acquiring fewer rules is selecting only some typical target pixels (e.g., pixels in the center of a target) for learning. However rules learned in this way might not capture the various data patterns near or on the border of a target. Because of this, the spatial distribution (important for target recognition) of pixels of detected targets would be damaged. Further, some targets would probably disappear. Experiments proved this analysis (not shown here). In our results, the number of learned rules was small and almost all targets were well preserved in their spatial distribution of pixels.

Our system for target detection is a good combination of vision and machine learning techniques. The noise reduction by the PWF and the screening from a CFAR detector maximally reduce and improve data needed as input to a learning method, in contrast to directly performing learning and testing on raw images (Rong & Bhanu, 1996).

6 CONCLUSION AND FURTHER WORK

This paper describes a novel application of the MIST methodology to target detection in SAR images. The presented MIST/AQ system for target detection in SAR images consists of: using the polarimetric whitening filter to enhance the quality of SAR images, applying a CFAR detector to screen natural clutter to maximally reduce the unnecessary information for training, and learning and detecting targets by AQ15c. The contributions of our approach can be summarized: the pixel-based operations in our approach avoid the problem of determining and grouping pixels or building templates; utilizing vision techniques maximally reduces and cleans data necessary to learning so that learning is more likely to succeed; false alarms are maximally reduced so as to provide a classifier with a list of potential targets of good quality. Experimental results presented are very promising and clearly show the effectiveness of the MIST/AQ method for solving this problem.

Further work will focus on (1) importing more SAR images and justifying the usefulness of the used attributes and the effectiveness of the MIST/AQ method for target detection in SAR images, and (2) designing appropriate attributes for detected targets by AQ15c and applying AQ15c or more advanced AQ learning programs such as AQ17-DCI (Bloedorn & Michalski, 1991) to target recognition.

REFERENCES

Armstrong, B.C, & Griffiths, H.D., "CFAR detection of fluctuating targets in spatially correlated K-distributed clutter", *IEE Proceedings-F,* vol. 138, no. 2, pp.139-152, 1991.

Bala. J.W., & Pachowicz, P.W., "Application of symbolic machine learning to the recognition of texture concepts", *Proceedings of the 7th IEEE Conference on Artificial Intelligence Application*, Miami, Fl., 1991.

Bhanu, B & Poggio, T., "Introduction to the special section on learning in computer vision", *IEEE Transactions on pattern analysis and machine intelligence*, vol. 16, no. 9, 1994.

Bloedorn, E., & Michalski, R.S., "Constructive induction from data in AQ17-DCI: further experiments," Report MLI91-12, Machine Learning and Inference Lab., 1991.

Burl, M.C., & Fayyad, U.M., & Perona, P., & Smyth, P., & Burl, M.P., "Automating the hunt for volcanoes on Venus", *Proceedings 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Seattle, WA, pp. 302-309, 1994.

Channic, T., "TEXPERT: an application of machine learning to texture recognition", Reports of Machine Learning and Inference Laboratory, 88-27, George Mason University, 1988.

Kerber, R., "ChiMerge: discretization of numeric attributes", *Proceedings of the Tenth National Conference on Artificial Intelligence*, San Jose, CA, pp. 123-128, 1992.

Kreithen, D.E., & Halversen, S.D., & Owirka, G.J., "Discriminating targets from clutter", *The Lincoln Laboratory Journal*, vol. 6, no. 1, pp.25-52, 1993.

Maloof, M.A., & Duric, Z., & Michalski, R.S., & Rosenfeld, A., "Recognizing blasting caps in x-ray images", *Proceedings of Image Understanding Workshop*, Palms Springs, CA, 1996.

Michalski, R. S., "AQVAL/1-computer implementation of a variable-valued logic system VL_1 and examples of its application to pattern recognition", *Proceedings of the First International Joint Conference on Pattern Recognition.,* pp.3-17, 1973.

Michalski, R.S., & Mozetic, I., & Hong, J., & Larvac, N., "The multipurpose incremental learning system AQ15 and its testing application to three medical domains", *Proceedings of the 5th National Conference on Artificial Intelligence.*, 1986.

Michalski, R.S., & Rosenfeld, A. & Aloimonos, "Machine Vision and Learning: a report on the NSF/ARPA workshop on machine learning and vision", Harpers Ferry, WV., 1992.

Michalski, R.S., & Bala, J.W., & Pachowicz, P.W., "CMU Research on learning in vision: initial results", *Proceedings of Image Understanding Workshop*, Washington D.C., 1993.

Michalski, R.S., & Zhang, Q., & Maloof, M.A. & Bloedorn, E., "The MIST methodology and its application to natural scene interpretation", *Proceedings of Image Understanding Workshop*, Palms Springs, CA., 1996.

Novak, L.M., & Burl, M.C., Chaney, & R.D., & Owirka, G.J., "Optimal processing of polarimetric synthetic-aperture radar imagery", *The Lincoln Laboratory Journal,* vol. 3, no. 2, pp.273-290, 1990.

Novak, L.M, & Owirka, G.J., & Netishen, C.M., "Performance of a high-resolution polarimetric SAR automatic target recognition system", *The Lincoln Laboratory Journal,* vol. 6, no. 1, pp.11-24, 1993.

Ravid, & Levanon, N., "Maximum-likelihood CFAR for Weibull background", *IEE Proceedings-F*, vol. 139, no. 3. June, 1992, pp.256-264, 1992.

Rong, S. & Bhanu, B., "Reinforcement learning for integrating context with clutter models for target detection", *Proceedings of Image Understanding Workshop*, Palms Springs, CA, pp.1389-1394, 1996.

Spence, C., "Supervised learning of detection and classification tasks with uncertain training data", *Proceedings of Image Understanding Workshop*, Palms Springs, CA, Feb. 1996, pp.1395-1402, 1996.

Wang, Y, & Chellappa, R., & Zheng, Q., "Detection of point targets in high resolution synthetic aperture radar images", *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing,* vol. 5, pp9-12, 1994.

Weiss, S.M., & Kulikowski, C.A., *Computer systems that learn: classification and prediction methods from statistics, neural nets, machine learning and expert systems*, Morgan Kaufmann, San Mateo, CA., 1992.

Wnek, J., & Kaufman, K., & Bloedorn, E., & Michalski, R.S., "Inductive learning system AQ15c: the method and user's guide", Reports of the Machine Learning and Inference Laboratory, MLI 95-4, George Mason University, Fairfax, VA., 1995.