Kernel Machines for Object Classification in High-Resolution SAR Data

Dr. W. Middelmann / U. Thoennessen
FGAN-FOM – Research Institute for Optronics and Pattern Recognition
Gutleuthausstrasse 1, D-76275 Ettlingen
GERMANY

Middelmann@fom.fgan.de / thoe@fom.fgan.de

ABSTRACT

The focus of this paper is the classification of military vehicles in high-resolution SAR images in an ATR framework. The usage of kernel machine classifiers is discussed. A new kernel machine, the relevance vector machine with integrated generator (RVMG) is introduced. Here, a single parameter controls the trade-off between speed and classification quality. Moreover classification heuristics and an adaptive feature extraction are used. These methods enable an improvement of the classification quality as well as a reduction of the computational effort. A parametrized reject criterion is presented to handle the classification of confusion objects. Therefore receiver operator characteristic (ROC) curves have been calculated. Tests have been performed using the MSTAR public target dataset and a fully polarimetric dataset from QinetiQ. An assessment of several polarimetric features has been performed.

1.0 INTRODUCTION

This paper focuses on the classification module of an ATR system of military vehicles in high-resolution SAR images. Modern digitally controlled radar systems have the ability to operate quasi simultaneously in two or more different modes. After detection of moving targets by MTI or other sensors the region of interest with the target cue can be recorded by a high-resolution spotlight SAR. Template based matching is a common approach for classification, i.e. the taken signature is matched with image catalogs of the interesting vehicles. The drawback is the high computational effort for the cross-classification. On the other hand during the last years a series of novel classification techniques – the kernel machine classifiers – have been introduced, see [1,2,3,4].

These kernel machines enable an improvement of the classification quality as well as a reduction of the computational effort. An enhanced new kernel machine, the Relevance Vector Machine with integrated Generator (RVMG) has been developed [9]. The basic idea is combining the high classification quality of the Support Vector Machine (SVM) by margin maximization and the low effort of the Relevance Vector Machine (RVM) caused by the special statistical approach. A single parameter controls the trade-off between its speed and classification quality.

Kernel machines are limited to two-class problems. Therefore, additional classification heuristics are required to solve multi-class problems. A simple but effective heuristic is proposed to handle operational classification problems of ten or more classes.

Obtaining real time capability of a classification module is possible by automatic feature extraction. We follow the common approach using adaptive Fourier-coefficients. The normalized 2D-SAR image is transformed via discrete Fourier transformation. Each single Fourier coefficient defines a weak classifier for the training dataset (two-class problem), i.e. a linear discrimination is defined. Then the best of these

Paper presented at the RTO SET Symposium on "Target Identification and Recognition Using RF Systems", held in Oslo, Norway, 11-13 October 2004, and published in RTO-MP-SET-080.

weak classifiers are chosen as feature extractors. Because of by the significantly lower dimension of the samples involved, the computational effort is now noticeably less time consuming.

Our investigations have been carried out with different datasets. Using the MSTAR public target dataset [10] ten classes were taken into consideration. The tests have shown that an RVMG with 500 Fourier coefficients yields a better classification quality (93.23%) for the ten class MSTAR problem than the nearest neighbor classifier (88.17%). Additionally the classifier (test phase) is 116 times faster than the nearest neighbor classifier that works on the original data set.

Current investigations with a nine class data set from QinetiQ deal with full polarimetric SAR data. The objective is to assess polarimetric feature extraction in combination with kernel machines to yield an even more robust ATR processing chain. The tests have shown that polarimetric features can slightly improve the classification quality. Among these the simple energy based features have proven more robust than complex ones.

Moreover, ROC curves have been calculated for the QinetiQ data set with four training classes and five classes with confusion objects. Therefore a parametrized reject criterion is proposed in this paper. It is possible to optimize the classification result with respect to the reject threshold and the kernel parameter. The ROC curves related to different polarimetric features enable an assessment relative to the False Alarm Rate (FAR).

The paper ends with the conclusion that reviews our work. Furthermore we will give attention to generalizability of SAR image catalogs for future investigations.

2.0 KEY PROPERTIES OF THE CLASSIFICATION MODULE

The proposed classification module is a part of the ATR processing chain shown in Figure 1. It is responsible for the classification of high resolution target signatures in SAR (and other) image data.

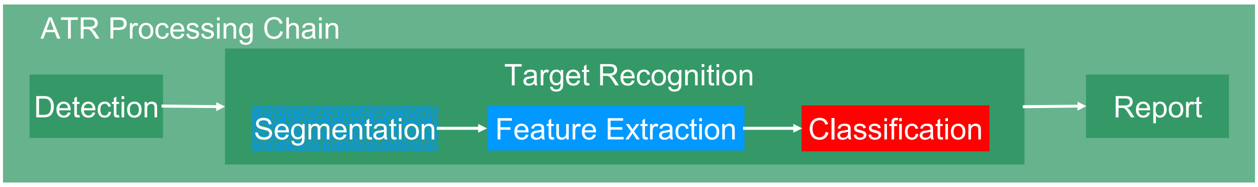


Figure 1: ATR processing chain with main modules

It has the following key properties that are dependent on several methods:

- Robust classification – kernel classifiers
- Improved applicability for multi-class problems – decision heuristic
- Real time ability – Fourier-coefficient-based feature extraction and RVMG
- Tunable trade-off between quality and FAR – RVMG
- Possible adaptation to requirements concerning classification quality, computational effort, FAR – parametrized reject criterion

The feature extraction has a very close connection to the classifier. Especially the Fourier-coefficient-based feature extraction works as a pre-classifier because of its adaptivity. Therefore the classification module may be understood as combination of the feature extraction, the kernel classifier, and the decision heuristic.

3.0 IMPLEMENTED METHODS

In this paragraph we introduce the methods implemented in the classifiers processing chain.

3.1 Kernel Machines

Our investigations are focused on kernel machines as classifiers. In many applications these have shown a high potential for robust classification [3,4,5,7]. Therefore the SVM (Support Vector Machine) and the RVM (Relevance Vector Machine) have been implemented.

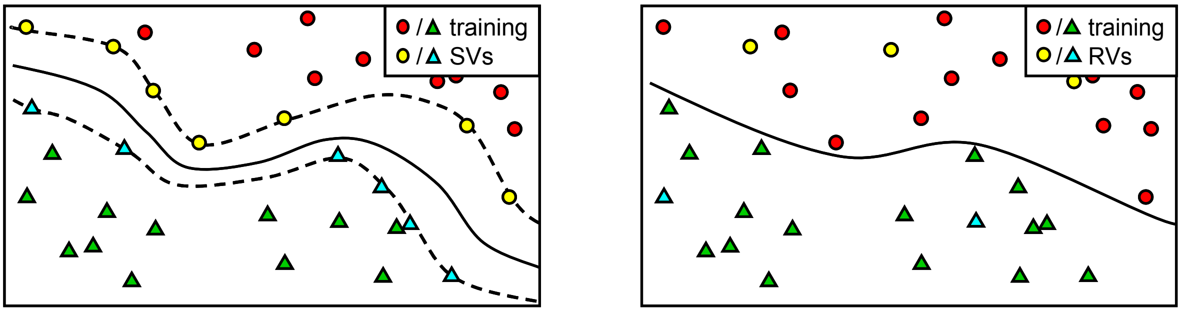


**Figure 2 left: SVM – robustness and high quality by margin maximization
right: RVM – low effort by low number of RVs**

The main characteristics of SVM and RVM are sketched in Figure 2. The SVM calculates Support Vectors (SVs) with non zero weights by margin maximization. This results in robustness and high generalization ability. In contrast, the RVM uses a special statistical approach to maximize the auto-classification quality and reduce the number of Relevance Vectors (RVs) – these are the vectors with non zero weights.

All investigated kernel machines use the RBF (Radial Basis Function) kernel, see Figure 3. As mentioned in [7] it has been shown that the RBF kernel is usually better than linear or polynomial kernels.

$$f(x) = \sum_{i=1}^{l} w_i K(x, x_i) + w_0 \qquad K(x, x_i) = \exp\left(- \frac{\| x - x_i \|^2}{\sigma} \right)$$

Figure 3: Test function of SVM and RVM, K is chosen as RBF kernel

SVM and RVM have the same test function (Figure 3). But the training – i.e. the determination of the weights – is based on different concepts. Therefore they have a very different behavior.

3.2 RVMG – Relevance Vector Machine with Integrated Generator

Customizing kernel machines results in the RVMG (RVM with integrated Generator), see [9]. A single parameter controls the trade-off between speed and classification quality of the RVMG. The basic idea is to combine the advantages of the SVM – high classification quality by margin maximization – and the RVM – low effort caused by the special statistical approach. Therefore additional points are generated in dependence on a single parameter λ controlling the relative distance of the new points to the margin, see Figure 4.

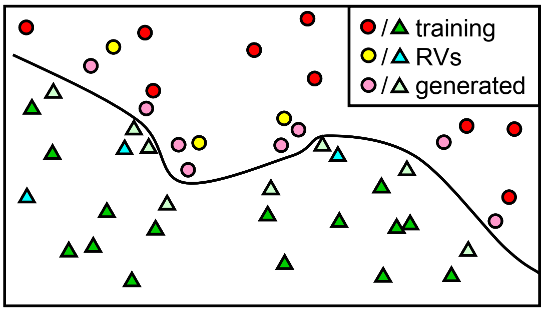


Figure 4: RVMG - Generated points establish a margin between the points of the two training classes

Using such additional points it is possible to define the RVMG as a modified RVM. The training algorithm of the RVM only uses a system matrix Φ with kernel elements K, e.g. K is the RBF kernel. It also uses class labels, so called hyper-parameters, and weights w_i, that have to be determined, see [5,6]. The rows of Φ correspond to the training vectors. The columns correspond to the basis functions. Basis functions and the RVM training dataset are independent of each other. It follows: The training dataset can be extended by generated data without any effect on the basis functions, see Figure 5.

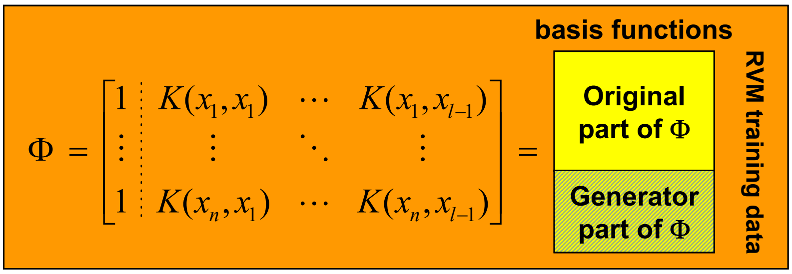


Figure 5: Extended system matrix by generated points, basis functions are not affected

This is in contrast to the original RVM where the basis functions have been set up corresponding to the training dataset. For more details refer to [9].

The RVMG has the following characteristics:

- Fortification of class boundaries
- No blur of class boundaries, no increase of class overlap
- No density increase of inner class regions
- Manageable number of additional training points
- Variable distance (controlled by the scalar parameter λ) between original and generated training points

First tests were conducted in [9] using a three-class problem with MSTAR data that present a wide spectrum of RVMGs (dependent on λ). They vary from a machine (original RVM) 15 times faster and 10% lower quality than the SVM to a machine a little faster than the SVM and even better. In this paper we present experimental set-ups with up to ten classes.

3.3 Heuristics for Multi-Class Problems

Decision heuristics qualify the two-class kernel machines for multi-class problems. The approaches used have been assessed for suitability concerning problems with many classes. Different approaches were feasible.

The first one is the 1-to-rest-heuristic. It discriminates each class against the union of all other classes. Each machine gives +1 or –1 relative to the decision surface. A class is determined if an unambiguous decision exists. The test sample is rejected if any conflict occurs.

A slight modification of the 1-to-rest heuristic is possible if each machine gives the distance to the decision boundary. Then the class with maximum distance is chosen. A reject is not possible.

The third heuristic is the 1-to-1-heuristic. It uses classifiers for each pair of classes. A two stage majority decision follows: The test sample is rejected if all classes get 70% or lower of the possible votes. The winner class is determined by direct comparison of the three best classes.

All three heuristics were tested using ten classes of vehicles provided by the MSTAR public target dataset. Hereby we found that for all tested kernel machines the classification results obtained using the 1-to-1-heuristic are the most favorable. The 1-to-restMx (with maximum decision) is sometimes better but it contradicts the idea of a necessary reject class.

3.4 Preprocessing & Feature Extraction

Kernel machines map input data into a high dimensional feature space. This is sometimes taken to mean that feature extraction is not necessary for such classifiers. But nevertheless preprocessing and feature extraction are able to enhance the generalization and to speed-up the whole classification.

It is well known, e.g. see [3], that in many cases a simple data-independent low pass filter not only reduces the computational effort but also generally improves the quality of the classification. This behavior is confirmed by our tests presented below.

Other investigations on Fourier-coefficient-based feature extraction have been done. This is an adaptive method for pre-classifying. The 2D-SAR data normalized with respect to energy is transformed via 2D discrete Fourier transformation. Each single Fourier coefficient defines a weak classifier for the training dataset (two-class problem), i.e. a linear discrimination is defined. Then the best of these weak classifiers are chosen as feature extractors.

3.5 Reject Criterion and ROC Curves

If the classifier is only tested against trained classes the tests only deliver a proposition concerning a closed world performance. More reliable and robust results can be achieved by tests against a set of unknown signatures, so-called confusers. For this, a reject criterion parametrized by a scalar d_{min} is introduced. Therefore ROC (Receiver Operator Characteristics) curves can be computed, presenting the interrelationship of FAR (False Alarm Rate) and classification quality.

The reject criterion is sketched in Figure 6. It is given in the high dimensional feature space of the kernel machine that is defined implicitly by the kernel – here we use the RBF kernel only. Classification in this feature space is done by simple linear discrimination. Therefore using the 1-to-1 decision heuristic each class has been trained against each other, i.e. hyperplanes c_{ij} define the classes' boundaries. An acceptance region is defined by the minimal distance d_{min} to all related hyperplanes, i.e. we use the 1-norm.

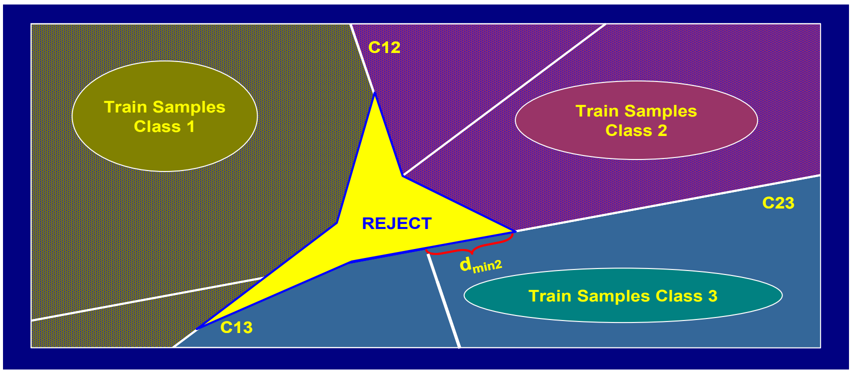


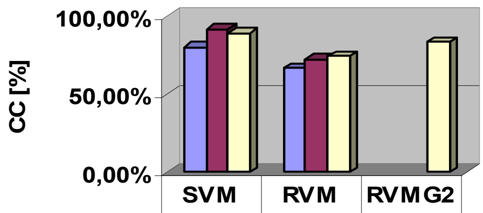
Figure 6: Parametrized reject criterion – An acceptance region is defined for each class.

The already used 1-to-1 decision heuristic is extended by the reject criterion. Thus there are two reasons to reject samples: The sample is rejected by the parametrized reject criterion or due to a class conflict inclusive ambiguous class voting.

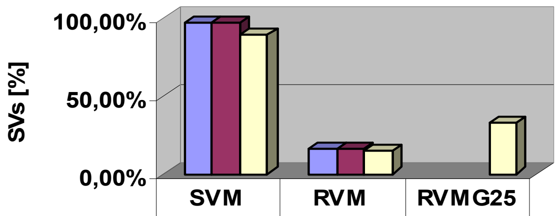
4.0 MSTAR DATA TESTS

The first investigations in classifying high resolution SAR data have been done with the MSTAR public dataset. The chosen MSTAR data consists of 3671 training and 3203 test chips organized in ten classes: BMP2, BTR70, T72, BTR60, 2S1, BRDM_2, D7, T62, ZIL131, and ZSU_23_4. The data has been taken under a depression angle of 17° for the training samples and of 15° for the test samples. Only the magnitude data have been used for the tests.

4.1 Heuristics and Kernel Machines for Robust Classification



	SVM	RVM	RVMG2
1-to-rest	80,08%	66,56%	
1-to-restMx	91,73%	72,40%	
1-to-1	88,95%	74,24%	83,39%



	SVM	RVM	RVMG25
1-to-rest	96,62%	15,88%	
1-to-restMx	96,62%	15,88%	
1-to-1	89,24%	14,87%	32,72%

Figure 7: Results of different combinations of heuristics and machines applied to the MSTAR ten classes problem. Left: Correct classified samples, Right: Percentage of used SVs or RVs

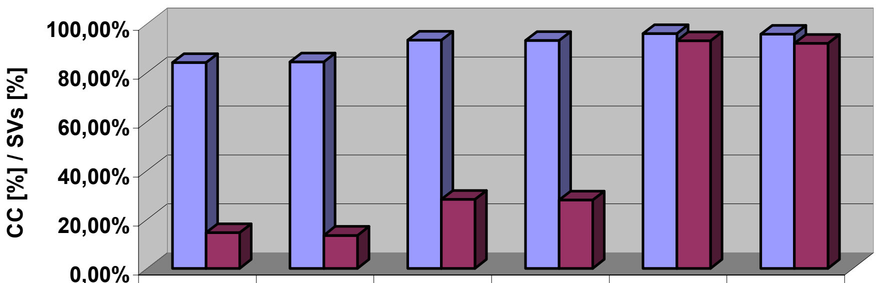
The decision heuristics and kernel classifiers introduced above are tested with the ten-class MSTAR problem. The results given in Figure 7 demonstrate the different behavior of the kernel machines.

For the SVM and RVM machines the results of the 1-to-1 heuristic are even better than these of 1-to-rest heuristic due to the stability of voting strategy. The 1-to-restMx with maximum decision yields good results but is off the discussion because of the necessity of a reject class. The RVMG (λ = 0.25) has been tested for the 1-to-1 heuristic only, because of the higher stability of the 1-to-1 heuristic – take a look at the number of SVs / RVs – and the huge computational effort for the training. The better generalization (low number of RVs) of the RVMG with respect to the SVM (factor 2.7) results in a 5.56% lower classification quality.

The nearest neighbor classifier produces a similar result with 88.17% (and 100% SVs). Here it should be remarked that we count the overall used number of SVs (RVs) and do not sum up the SVs used in all two-class sub-problems.

4.2 Real-Time-Capability by Fourier-Coefficient-Based Feature Extraction

Fourier-coefficient-based feature extraction is an adaptive method for pre-classifying and was tested for different kernel machines. The results are given in Figure 8. For each the best 500 or 1000 coefficients are chosen.



	RVM 500	RVM 1000	RVMG25 500	RVMG25 1000	SVM 500	SVM 1000
CC %	84,11%	84,33%	93,23%	93,16%	95,88%	95,69%
SVs %	14,55%	13,48%	28,22%	27,89%	92,97%	91,94%

Figure 8: Real time ability by Fourier coefficients (500 or 1000 coefficients have been selected) Classification quality (CC) and percentage of SVs or RVs for MSTAR ten classes are given.

RVMG25 with 500 coefficients is real-time-capable because it is 116 times faster (more than 100 samples per second without I/O) than the nearest neighbor classifier (CC_{NN}=88.17%) without feature extraction. This speed increase is due to the reduction in dimensions from 16384 (original data) to 500 (Fourier coefficients).

5.0 FULLY POLARIMETRIC QINETIQ DATA TESTS

It is a very interesting question how the classifier performance can be improved by fully polarimetric data. We investigated this problem using the QinetiQ dataset. The QinetiQ dataset contains 4006 images subdivided into 9 target classes A to I. Each of these classes is further subdivided into a training set consisting of 335 different target aspects and into a test set consisting of 110 test samples. Each of the complex-valued images of size 150x100 is depicting one single target. However, the target positions are varying from image to image. As a fixed target position is a prerequisite for achieving good results using SVM, 64x64 image sub-windows were selected in a pre-processing step. In the new images the centre of gravity of the binary object mask of the target coincides with the centre of the image window. Because the investigations were carried out with magnitude based features, only the magnitudes are stored (QinetiQ64). In order to be able to assess the influence of the image resolution, a further dataset (QinetiQ32) was generated degrading the 64x64 images to 32x32 using a 2x2 window.

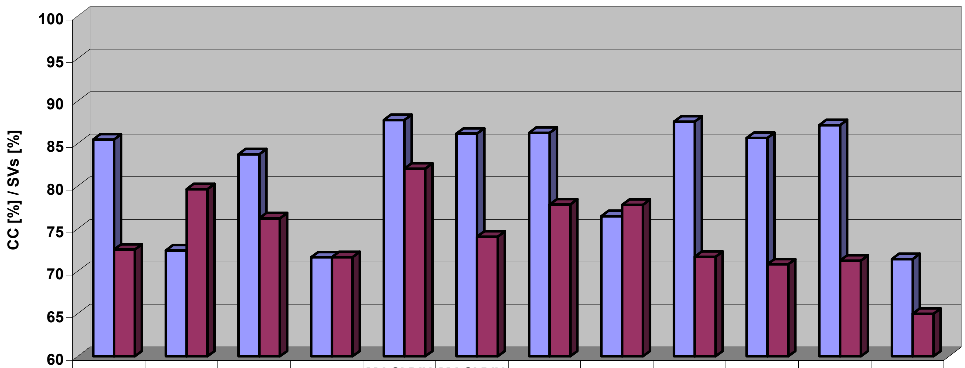
5.1 Polarimetric Feature Selection

For both QinetiQ32 and QinetiQ64 different polarimetric features were investigated using the SVM. The classification results obtained using complex-valued features, especially the Pauli decomposition, were lower than 50%. Therefore in the following only magnitude based features were considered.

The tested polarimetric features based on the magnitudes are:

- a single coplanar channel (VV, HH)
- the average of the two coplanar channels (MAGVVHH)
- the concatenation of the two coplanar channels (VVHH)
- the average of all four channels (MAG)
- the concatenation of all four channels (ALL)

The investigations were done using QinetiQ64 as well as QinetiQ32 applying the SVM with Gaussian filtering (σ = 0.75 for QinetiQ32 , σ = 1.5 for QinetiQ64). A survey of the results is shown in Figure 9. The classification quality is represented in blue whereas the computational effort is depicted in purple.



	HH32	HH64	VV32	VV64	MAGVVH H32	MAGVVH H64	VVHH32	VVHH64	MAG32	MAG64	ALL32	ALL64
CC[%]	85,47	72,45	83,75	71,64	87,79	86,18	86,28	76,49	87,59	85,67	87,18	71,44
SV [%]	72,54	79,64	76,22	71,67	82,06	74,06	77,84	77,81	71,71	70,81	71,24	64,98

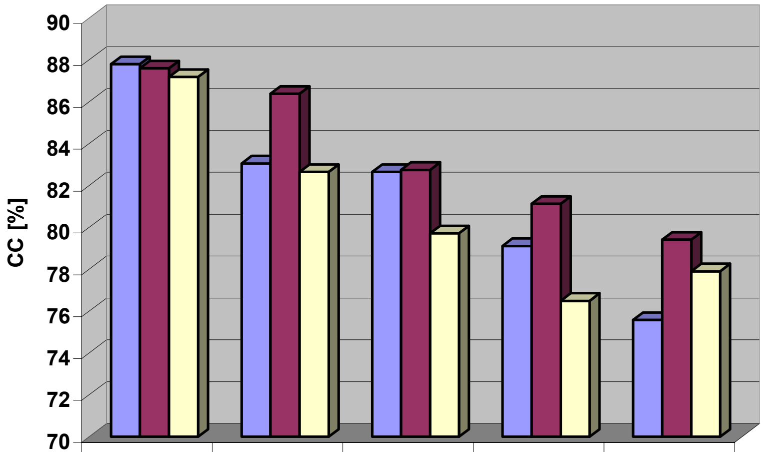
Figure 9: Efficiency and quality achieved using SVM

The results pointed out that it is more favorable to use QinetiQ32 with the reduced geometrical but improved signal-to-noise-ratio than to use QinetiQ64 possessing the fully geometrical information. However, there is a clear distinction between the features based on averaging and those based on concatenation. In case of averaging the increase of the classification rate is small, less than 2%. In case of concatenation the increase is essentially larger, more than 10%.

Using the fully polarimetric information is also of value. For QinetiQ32 an increase, concerning the classification rate, of about 2% was noticeable. For the QinetiQ64 an increase of about 5% was detected. In both cases the best classification results were obtained using both coplanar channels. The main advantage of using all four channels is the reduction of the number of support vectors. Assessing the computational effort it is necessary to consider that concatenation enlarges the dimension of the support vectors by two or even four times. Therefore a reduction of the computational effort can be achieved only for averaging based features. Thus the most promising feature is MAG32 because of its good classification results and its low computational effort. Considering only the classification rate MAGVVHH32 and ALL32 are also well suited features. Finally HH32 is also a good choice as the classification performance is diminished only by 2% and the support vector quota is similar to the one of MAG32.

5.2 Real-Time-Capability by Advanced Methods

For usage in an operational ATR system the classification task must be solved in real time. In the case of high-resolution SAR data this is actually not state of the art. To achieve this objective we use the RVM and RVMG which minimize the number of support vectors, as well as adaptive Fourier coefficients and the combination of these using the most promising features MAG32, MAGVVHH32 and ALL32.



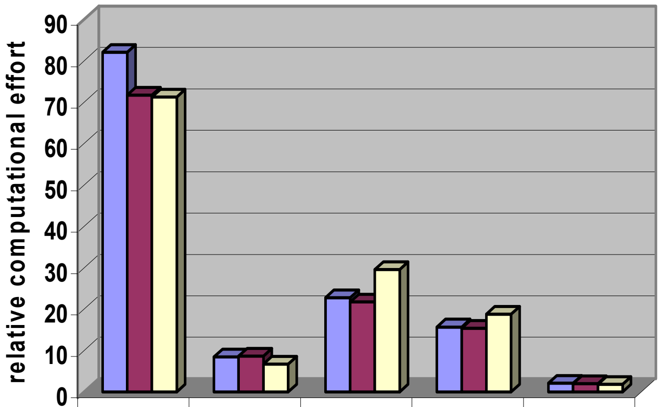
	SVM	SVMF	RVMG25	RVMG10	RVMG10F
MAGVVHH32	87,79	83,05	82,64	79,11	75,58
MAG32	87,59	86,38	82,74	81,13	79,41
ALL32	87,18	82,64	79,72	76,49	77,90

Figure 10: Classification quality of the most favorable features

The classification results obtained using these new methods with Gaussian filtering are depicted in Figure 10. The results for the SVM are given in the first column. Results of the SVM with image data transformed into 100 Fourier coefficients (SVMF) are indicated in the second column. The third and fourth column present the results of the new RVMG when using the control value $\lambda = 0.1$ (RVMG10) respectively $\lambda = 0.25$ (RVMG25). Finally, the fifth column indicates the results of the RVMG with image data transformed to 100 Fourier coefficients (RVMG10F).

As expected the best results were obtained applying the SVM achieving a classification rate of more than 87%. The poorest results were obtained using RVMG10F with a classification rate varying for the different features between slightly more then 75% and slightly less than 80%. The best results were obtained using MAG32 with a loss of about 1% in case of SVMF and about 8% in case of RVMG10. With the exception of RVMG10F with the MAGVVHH32 feature, all classification rates decrease by less than 10%.

An important assessment criterion for the tested methods was the computational effort. Comparing the different methods we have to consider that in case of SVMGF and RVMGF the image data consisting of 1024 pixels was reduced to a vector of 100 Fourier coefficients which causes a nearly 90% reduction of the computational effort.



	SVM	SVMF	RVMG25	RVMG10	RVMG10F
MAGVVHH32	82,06	8,49	22,75	15,72	2,08
MAG32	71,71	8,65	21,79	15,36	2,01
ALL32	71,24	6,77	29,59	18,81	1,87

Figure 11: Computational effort of the most favorable features in percent of that of an SVM with 100% SVs without Fourier coefficients

In Figure 11 a survey of the corresponding adapted computational efforts is given. All new methods have an essentially lower effort. For the RVMG25 test the speed up factor is considerably higher than 2. The RVMG10F yields a speed up factor of nearly 40, with respect to the SVM.

5.3 Optimization of FAR against Quality

In the preceding sections the performance concerning the discrimination between trained classes was investigated. In case of ATR however not only objects of the trained classes but also objects of non-trained classes or even artifacts will occur. Therefore an important property of a classifier would be the capability to reject objects not belonging to one of the trained classes.

In order to test whether the implemented reject criterion fulfills this demand, the following experiment was carried out. Only the four classes A, B, D and G of the QinetiQ dataset were trained. Then the test samples of all nine classes were classified. The results obtained are depicted in Figure 12 in case of MAG32, in Figure 13 in case of MAGVVHH32 and in Figure 14 in case of ALL32. The results are controlled by two parameters. The curves family is parametrized by the reject criterion, i.e. the minimal distance (1-norm) d_{min} to all related hyperplanes. The curves were determined by varying the kernel factor σ of the RBF.

The results demonstrate that for distinct FAR intervals related optimal d_{min} values and optimal features exist. Generally using a larger d_{min} i.e. a stronger reject criterion will result in a poorer classifier performance. However, as MAG32 is indicating, this global trend may no longer be valid for small FARs – usually the most interesting part of the curves. In this specific example the reject criterion $d_{min} = 0.5$ is the most favorable one. This may be the reason that generally only small values of the normalization factor σ can guarantee small FARs. But these settings also cause the rejection of a large number of real targets. Reducing the false alarms by increasing the strength of the reject criterion permits the rise of the factor σ normalizing the RBF kernel. By this, the efficiency of the classifier will increase, too.

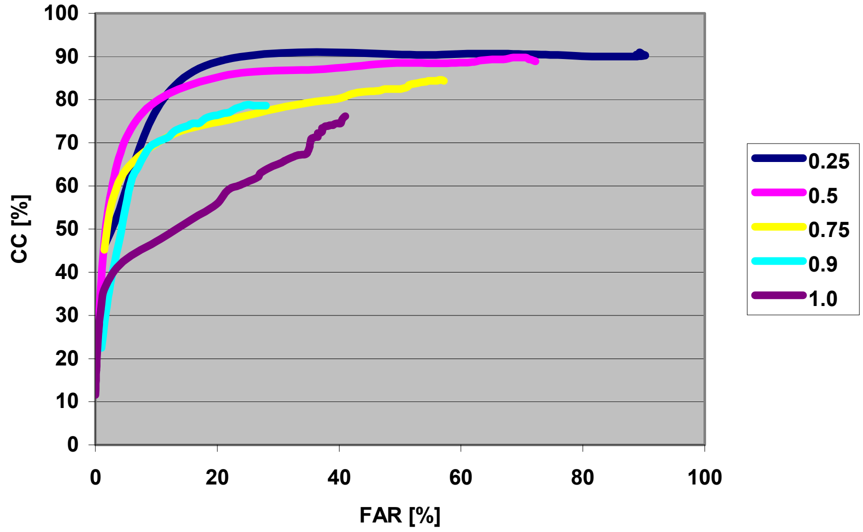


Figure 12: Result of the reject criterion for MAG32

Evaluating the tested features shows that no feature is always better than the two other ones. For a demand of a FAR below 15% using MAGVVHH32 is more favorable than using MAG32 or ALL32. However in case of a FAR above this level MAG32 is the best choice.

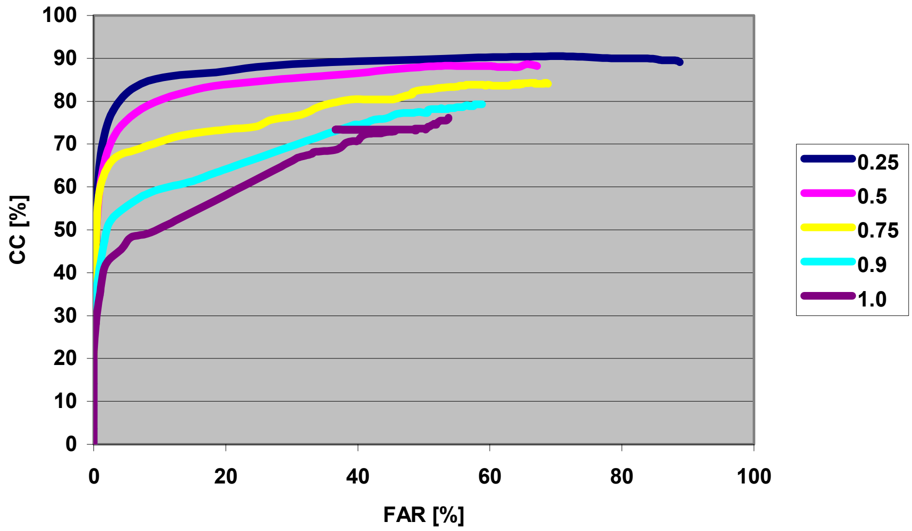


Figure 13: Result of the reject criterion for MAGVVHH32

The ROC curve of ALL32 is always positioned below the ROC curves of either MAG32 or MAGVVHH32 indicating a poorer performance of this feature. By this, the ranking of the features presented in the preceding section is confirmed.

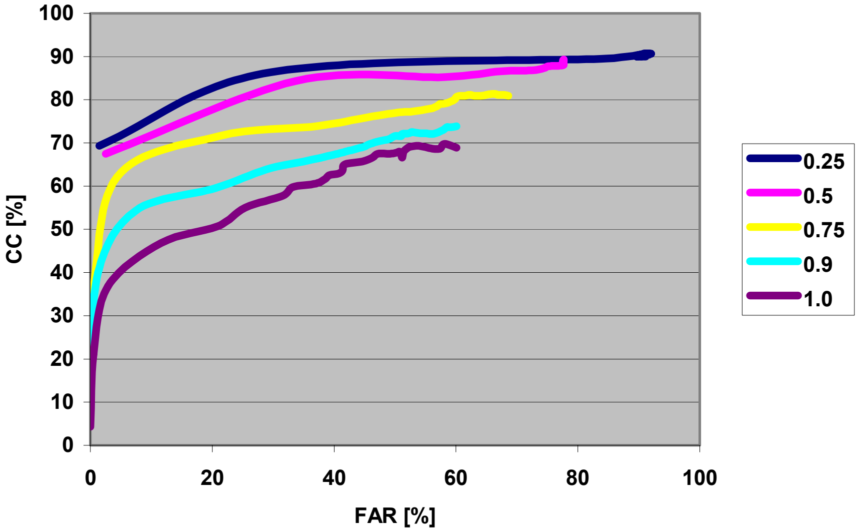


Figure 14: Result of the reject criterion for ALL32

6.0 CONCLUSION

In this paper we introduced classification methods for the processing chain of ATR in high-resolution SAR images. Kernel machines as robust classification methods are the basis of our approach. A novel kernel machine was presented that controls the trade-off between classification quality and computational effort, i.e. number of relevance vectors. The multi class classification capability is given by an efficient 1-to-1 decision heuristic. An adaptive feature extraction based on Fourier coefficients enables the module for real time execution.

The investigations have taken place for the MSTAR public database. Tests with a ten-class problem have been analyzed. For example the RVMG25 machine with a pre-classifier using 500 Fourier coefficients is 116 times faster than the original nearest neighbor. The classification quality has been improved from CC_{NN}=88.17% to CC_{RVMG25_500}=93.23%.

Further investigations have used the fully polarimetric QinetiQ dataset with nine classes. The tests carried out indicate that the most favorable polarimetric features for the depicted hard targets are the magnitude based ones. Especially the two coplanar polarizations embody the essential information for the class distinction, e.g. the SVM result of MAGVVHH32 with CC=87.79% is similar to MAG32 with CC=87.59%. But the crossplanar channels cause a better generalizability, e.g. the SVM gives SV=82.06% for MAGVVHH32 and SV=71.71% for MAG32. On the other side complex features like those based on the Pauli decomposition are off the discussion because of their weak performance.

An important property of a classifier used in the ATR framework is the capability to reject objects not belonging to one of the trained classes. Therefore the QinetiQ data have been divided into two class systems: the training and test classes and the confusion objects. The classification module with reject criterion is controlled by the reject parameter and the kernel parameter of the RBF. To determine ROC curves we have varied both parameters. For example the SVM with MAGVVHH32, reject parameter r=0.25, and RBF parameter σ=0.1 yields a very low FAR=2.9% and a classification quality of CC=77.73%.

By the way classification quality of CC_{NN}=91.83% was achieved for the nine class QinetiQ data by the nearest neighbor classifier with MAGVVHH32 and a simple pre-filtering. This single result also approves the well selected polarimetric feature for hard targets.

This paper has shown that polarimetric data is also useful for ATR tasks in high-resolution SAR data. Future work should deal with a more physically founded analysis of polarimetric features. Especially the whole complex information should be made available in a proper way for the classifiers. This could lead to a better generalizability of SAR image catalogs.

ACKNOWLEDGEMENTS

The authors wish to thank D. Blacknell at QinetiQ for the polarimetric data set. Moreover many thanks to A. Ebert at FGAN-FOM for his technical assistance in implementing the software system and testing.

REFERENCES

[1] V. Vapnik, A. Chervonenkis, “Theory of Pattern Recognition“ [in Russian], Nauka, Moscow, 1974, (German Translation: W. Wapnik, A. Tscherwonenkis, Theorie der Zeichenerkennung, Akademie-Verlag, Berlin, 1979)

[2] V. Vapnik, Statistical Learning Theory, Wiley, New York, 1998

[3] B. Schölkopf, “Support Vector Learning”, PhD Thesis, R. Oldenbourg Verlag, Munich, 1997

[4] N. Cristianini, J. Shawe-Taylor, “An Introduction to Support Vector Machines and other kernel-based learning methods”, Cambridge University Press, 2000

[5] M. E. Tipping, “The Relevance Vector Machine”, Advances in Neural Information Processing Systems, 12, S. A. Solla, T. K. Leen, K.-R. Muller (eds), MIT Press, Cambridge, MA, 2000.

[6] M. E. Tipping, “Sparse Bayesian Learning and the Relevance Vector Machine”, Journal of Machine Learning Research, 1, 211-244, 2001.

[7] H. Byun, S.-W. Lee, “A Survey on Pattern Recognition Applications of Support Vector Machines”, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 17, No. 3, 459-486, 2003

[8] W. Middelmann, U. Thoennessen, “Classification in High-Resolution SAR Data”, Proceedings of SPIE - Algorithms for Synthetic Aperture Radar Imagery X, E. G. Zelnio, F. D. Garber (eds.), 5095, pp. 325-335, 2003

[9] W. Middelmann, “Classification with Controlled Robustness in High-Resolution SAR Data”, Proceedings 25th DAGM Symposium, Magdeburg, Germany, September 2003, B. Michaelis, G. Krell (eds.), LNCS 2781, pp. 92-99, Springer, 2003.

[10] MSTAR, Air Force Research Lab., Model Based Vision Lab., Sensor Data Management System, http://www.mbvlab.wpafb.af.mil/public/sdms/datasets/mstar/overview.htm