Robust Acquisition of Relocatable Targets

Hartmut Schimpf
FGAN Forschungsinstitut für Hochfrequenzphysik und Radartechnik (FHR)
D-53343 Wachtberg – Werthhoven, Neuenahrer Straße 20
GERMANY

ABSTRACT

Three target types, namely T72, ZSU 23-4 and BMP-2 were measured in a tower/turntable configuration in several articulations each. A set of geometric, statistical, structural and polarimetric features is used to study the robustness of classification. Based on the Kolmogoroff-Smirnov distance between histograms a metric is defined that at the same time allows to quantify intra-class robustness and inter-class separability for an individual feature. For sets of several features, a simple classification approach in connection with a reference confusion matrix allows to assess the robustness of classification. It is demonstrated, that averaging the feature reference over all available target articulations improves the classification performance as compared to a reference that is based on one articulation only. Also, it is shown that in most cases, the classification is the better the more precisely the target aspect angle can be estimated independently. -- The paper reports work that is done in the framework of the NATO RTO/SET-069 working group.

1 INTRODUCTION

Features are a means of statistical pattern recognition that ATR algorithms use to discriminate ground targets from the surrounding clutter background and, subsequently, to sort potential targets into one of several target classes (including the non-target case). Problems for ATR arise from the specular nature of radar imagery because small changes to the configuration of targets can result in significant changes to the resulting target signature [3][4]. This adds to the challenge of constructing a classifier that is both robust to changes in target configuration and target aspect, and which is capable of generalizing to previously unseen targets.

ATR features have to provide at the same time good inter-class separability and good intra-class stability. The reference vectors usually are obtained from former measurements of the respective target either on a turntable or by means of SAR and are stored in look-up tables. The test vectors are obtained on-line while the seeker is passing over the target area. In order for the ATR to provide reliable results both the test vectors and the reference vectors have to show **robustness** against target modifications, preferably including camouflage, different target realizations or articulations, slight changes in depression angle, aspect angle changes that occur during the time-on-target, and many more. Robustness has to be understood in the sense that the statistics of the test and reference vectors either remain unchanged, or that their changes are taken into account appropriately, and that the estimates that are obtained of these vectors are representative for this statistics. As a consequence, the classification performance should not be degraded. In order to obtain the desired robustness it is of great importance to eliminate those target variations that can be handled beforehand, the most crucial one being the aspect angle dependence. The analysis of tower/turntable measurements on typical targets shows that feature values as a function of aspect angle do not only fluctuate around a stable mean, but that their statistics themselves, i.e. mean and standard deviation, are a function of aspect angle. It has been demonstrated before [1][2] how important an independent determination of the target aspect angle is. Among the methods most commonly used are the Hough transform or a process of pattern matching [1] [5].

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.

Three different military vehicles were measured in a tower/turntable configuration at 35GHz with 800MHz bandwidth using the fully polarimetric MEMPHIS radar. Each vehicle was measured in several different articulations (hatches open or closed, turret turned to different positions) while its positioning on the turntable remained unchanged. All data underwent an identical polarimetric calibration to warrant optimal comparability.

As a means to assess feature robustness several metrics were developed to quantify the results. Two approaches are compared:

- the separability between feature histograms (using the Kolmogoroff-Smirnow distance as a distance measure)
- analysis of confusion matrices based on a generic classification scheme

Typical features of various types (geometric, statistical, polarimetric, scatterer power, structural etc.) are used, each one depending on one or two parameters that allow optimization.

The paper is organised in five sections. The first one gives a short description of the measurement setup. In the second, the features used for classification are described in some detail. Next, the relationship between robustness and inter-class separability is analysed. Section four shows how confusion matrices can be used to characterize robustness. Finally, some thoughts on the aspect angle behaviour of the features are presented.

2 MEASUREMENT SETUP

For the measurements that are analysed here, the FGAN operated fully polarimetric MEMPHIS radar [8] was located on top of a tower at a height of 47 meters. The three targets (T72, ZSU 23-4 and BMP) were positioned on a turntable at a distance of about 154m, giving rise to a slant range of 161m and a depression angle of 17°.

The MEMPHIS 35 GHz radar transmitted linear V polarisation, and received H and V simultaneously thus providing orthogonal VV and VH channels. The basic waveform is a linear chirp with 200 MHz bandwidth. In order to achieve higher range resolution, this chirp is combined with a stepped frequency mode with 8 steps of 100 MHz increment [9]. The resulting maximum processing bandwidth thus is 800 MHz. However, as this requires a 320-point DFT (2.5 MHz frequency sampling step), here only a reduced bandwidth of 640 MHz was processed allowing the use of a 256-point FFT. The resulting range resolution is about 0.24m which is sufficient for this kind of ATR analysis [10].

A full revolution of the turntable took place in 130 seconds, the effective PRF was $2300s^{-1}/8$ such that a 128-point Doppler FFT results in a cross-range resolution of 0.2m, sufficiently close to the desired square-pixel case. The targets were measured in the following configurations: the turret of the T-72 was positioned 20° to the left, and in 30° intervals from 0° (forward) to 180° (backward). In the case of the ZSU 4 different and of the BMP 5 different combinations of shut/closed driver's, commander's and turret hatches were realized, cf.[6].

3 CALCULATION OF THE CLASSIFICATION FEATURES

All feature values were computed on the basis of 2-D ISAR images with 0.24m (range) by 0.2m (cross-range) pixels. They were taken from a list prepared by the NATO SET-TG14 working group [7]. For geometrical, statistical, and structural features, the total power map ($|VV|^2 + |VH|^2$) was used, for the polarimetric features the VV and VH power map were used in parallel.

- ft1 = range extent of 20 strongest scatterers
- ft2 = cross-range extent of 20 strongest scatterers
- ft3 = ft1*ft2 (= area of the "minimum bounding rectangle" (MBR))
- ft4 = mean/std.dev.(total power|MBR)
- ft5 = (powersum 10 strongest scatterers) /powersum(MBR)
- ft6 = log10(pmax(1)/pmax(5)) (ratio between strongest and 5^{th} strongest scatterer within the MBR)
- ft7 = log10(pmax(1)/pmin)|MBR (ratio between strongest and weakest scatterer within the MBR)
- ft8 = $\max(pvv/pvh)|_{dB}$ - $\min(pvv/pvh)|_{dB}$ (span of parallel/cross channel separation)
- ft9 = slope(pmax vs.dif)$|_{dB}$
- ft10 = shift(pmax vs.dif)$|_{dB}$

(in ft9 and ft10 "pmax" stands for the 10 strongest scatters within the MBR, sorted in descending order, "dif" contains the related channel differences pvv/pvh, shift and slope refer to a least squares line fit that is applied to these 10 pairs of values).

The rationale for the selection of this set of features is not that they constitute a "best" set. Rather they are considered to be a "generic" set with representatives from several feature types, namely geometric, statistical, scatterer power related or structural, and polarimetric. Of course, some of these features are more or less correlated with one another. This can be assessed either by determining all the mutual cross correlation coefficients, or by a principal component analysis (PCA, [11]). Therefore, only certain subsets out of these 10 features will form meaningful sets of ATR features.

4 INTRA-CLASS ROBUSTNESS VS. INTER-CLASS SEPARABILTY

The basic test of robustness is to analyse how strongly the feature statistics is changed when the respective target is modified. It is clear that each feature as a function of aspect angle will reflect any target modifications under those angles where they become effective. But in the ideal case, the overall statistics as measured over a certain aspect angle interval, should only slightly be affected, i.e. the probability density function (pdf) should be more or less the same.

There is a duality between intra-class robustness and inter-class separabilty. The more tolerant a feature or set of features or a classification scheme is towards different articulations of a certain target type, the less likely it is to precisely discriminate between a large number of different target types.

A convenient way to compare two probability density functions (pdf's) or histograms is by determining the Kolmogoroff-Smirnov distance (KSD) which is defined as the maximum difference between the two pertinent cumulative distributions:

Let $p_1(f)$ and $p_2(f)$ be two probability density functions (pdf's) of a certain feature "f" obtained from two different vehicles. The pertinent distribution functions then are

$$P_i(f) = \int_{-\infty}^{f} p_i(f')df' \quad (i=1,2)$$

and hence

$$KS(p_1, p_2) = \max_f |P_1(f) - P_2(f)|.$$

By definition, the KSD can vary between 0 and 1, where “0” means identity, and “1” means complete separation without any overlap.

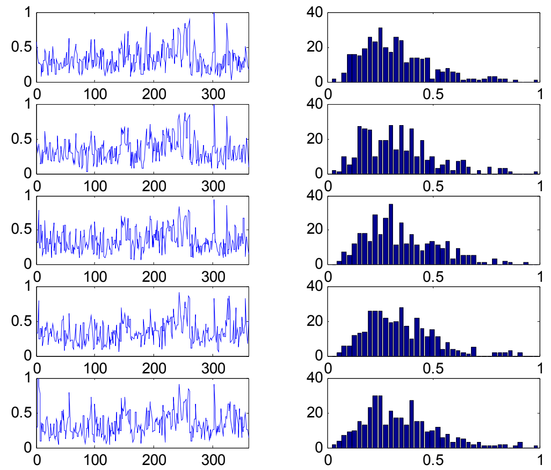

Figure1 plot of feature #6 vs. aspect angle, and related histograms for5 articulations of the BMP

Fig. 1 shows as an example the power feature #6 for the five different articulations of the BMP. As one sees, the polar curves in some cases may differ considerably e.g. in the interval 80° to 120° or near 330°. However, the histograms look rather similar. Thus, at first sight this seems to be a candidate for a robust feature.

On the other hand, robustness is only one criterion that an ATR feature has to fulfill. Good discrimination w.r.t. other targets is another important one, and both properties have to be examined and eventually a trade-off has to be made.

Let us now quantify the similarity between the pdfs of different target articulations by means of the KSD. This is best done using a table that lists all possible combinations of pairs of targets for a selected feature. Let us again look at feature #6 (Table 1). The KSD between pairs of different T72 are fairly low, mostly less than 0.1 with some outliers up to 0.144, which suggests a few major differences.

For pairs of ZSU or BMP, values are below 0.086 and hence close to zero as required. In the areas that are dark grey-shaded we have pairs of different target types. Here, in the ideal case we would expect values close to 1, i.e. complete separation. Of course, this is not the case, rather the values are around 0.2, hardly above 0.23. This is certainly not satisfactory, but one has to keep in mind that the classification will not be done with only one feature but that one will go to higher dimensions of the feature space where less overlap is expected.

The dark shaded areas of the triangular matrix K is where KSD values close to “1” are expected, all others should be close to zero. If we define a reference matrix R which contains only the desired values 0 or 1 in the appropriate positions, then the quality of a feature can be judged by computing the distance between the actual matrix and the reference matrix.

	T72b	T72c	T72d	T72e	T72f	T72g	T72h	ZSUa	ZSUb	ZSUc	ZSUd	BMPa	BMPb	BMPc	BMPd	BMPe
T72a	0.052	0.108	0.041	0.069	0.061	0.044	0.061	0.136	0.088	0.111	0.113	0.180	0.144	0.138	0.136	0.127
T72b	0	0.091	0.047	0.066	0.047	0.077	0.069	0.119	0.075	0.091	0.116	0.227	0.180	0.161	0.169	0.161
T72c	0	0	0.113	0.122	0.108	0.122	0.144	0.097	0.116	0.1	0.072	0.247	0.208	0.186	0.211	0.205
T72d	0	0	0	0.055	0.052	0.075	0.072	0.136	0.1	0.108	0.113	0.2	0.175	0.158	0.144	0.144
T72e	0	0	0	0	0.077	0.083	0.083	0.15	0.102	0.125	0.144	0.216	0.175	0.169	0.163	0.172
T72f	0	0	0	0	0	0.077	0.077	0.108	0.069	0.094	0.097	0.205	0.186	0.147	0.166	0.166
T72g	0	0	0	0	0	0	0.05	0.15	0.113	0.119	0.136	0.177	0.133	0.116	0.122	0.102
T72h	0	0	0	0	0	0	0	0.166	0.127	0.144	0.147	0.194	0.161	0.136	0.141	0.133
ZSUa	0	0	0	0	0	0	0	0	0.072	0.041	0.036	0.227	0.225	0.202	0.233	0.211
ZSUb	0	0	0	0	0	0	0	0	0	0.05	0.063	0.222	0.202	0.172	0.197	0.175
ZSUc	0	0	0	0	0	0	0	0	0	0	0.069	0.222	0.213	0.177	0.205	0.186
ZSUd	0	0	0	0	0	0	0	0	0	0	0	0.222	0.216	0.172	0.211	0.188
BMPa	0	0	0	0	0	0	0	0	0	0	0	0	0.061	0.086	0.080	0.080
BMPb	0	0	0	0	0	0	0	0	0	0	0	0	0	0.086	0.077	0.063
BMPc	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.055	0.063
BMPd	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0.038

Table 1 KSD between pdf's of all 17 targets for feature #6

$$d = \sqrt{\frac{2}{17 \cdot 16} \sum_{i<j} \| K - R \|^2} \qquad (1$$

Table 2 shows an example. The smaller this value is, the closer the measured matrix is to the reference. Taking this metric, the range extent (feature 1) performs best.

However, this single value does no longer allow to differentiate between robustness and separability. Robustness is the better the closer the intraclass KSD are to zero, and separability is the better the closer the interclass KSD are to 1. One can therefore average all intraclass KSD (resulting in K_0) and all interclass KSD (resulting in K_1) and plot the results in K_0-K_1-coordinates (fig.3). The closer a feature is located to the point (0,1) the better its performance will be [6].

Feature #	d
1	.55
2	.68
3	.583
4	.647
5	.635
6	.695
7	.638
8	.758
9	.68
10	.68

Table 2 mean difference "d" between matrices K and R

As one sees from the definitions of the classification features, they all depend at least on one free parameter. First of all this is the number of scatterers (N_{sc}) that is is used to create the MBR on which all subsequent computations are performed. N_{sc}=20 was used for most of the analyses presented here. In addition, feature #5 uses a subset of scatterers within the MBR (here 10 out of 20), and feature #6 is the power ratio between the strongest and 5th-strongest scatterer within the MBR, thus introducing a second parameter for each of them. One can think of using these free parameters in connection with either the reference matrix described above, or the point (0,1) in the K_0-K_1-plane to somehow optimize the features. The values that were used in the original definitions may seem somewhat arbitrary and need a justification. Let us begin with the reference matrix and the related distance measure (1). If we vary the common parameter N_{sc}, we can consider d=d(N_{sc}) and try to find a minimum in a certain reasonable interval. Fig.2 shows d(N_{sc}) for all 10 features where N_{sc} was varied between 10 and 30. As one sees, there is almost no dependence on N_{sc}, features 1,3,4,5,7,9,10 show a very small decrease of d(N_{sc}) with increasing N_{sc}, features 2 and 8 show an increase of d(N_{sc}), and #6 is not influenced at all. From this one cannot get a clear recommendation towards a certain value of N_{sc}.

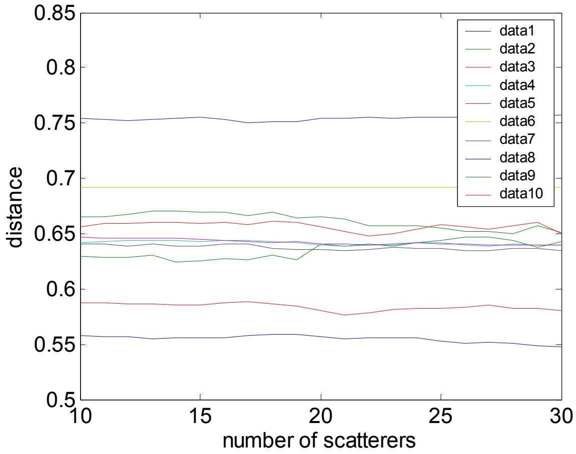

Figure 2 distance between K and R for varying N_{sc}

Because d(N_{sc}) cannot distinguish between robustness and separability, we now want to look to K_0 and K_1 which describe these two properties individually. Fig.3 shows how each feature moves in the K_0-K_1-plane when N_{sc} is varied between 10 and 30. The starting point (N_{sc}=10) is marked by 'o' while the final point is marked by '*'.

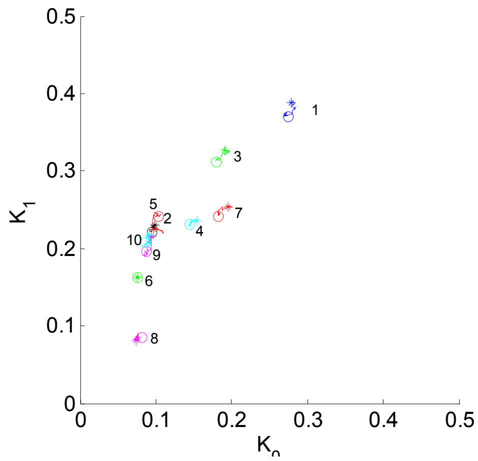

Figure 3 K_1 vs. K_0 for varying N_{sc}, features 1 - 10

What we hope to find is a tendency to approach K_0=0 and K_1=1. However, none of the 10 features follows this pattern. Rather, K_0 and K_1 both tend to increase with increasing N_{sc} (features 1,3,4,7,9) or are almost constant (features 6 and 8). Obviously, it is not possible to push K_0 beneath a value of ≈0.08. As one recognizes, none of the 10 features comes close to the desired locus (0,1) in the K_0-K_1-plane for values of N_{sc} between 10 and 30. One must conclude that N_{sc} offers only a very limited optimization potential for the 10 features under consideration. It seems as if the 10 features as defined above do not reflect sufficiently the geometrical and scatterer structure of the targets and therefore are too insensitive w.r.t. N_{sc}.

5 CONFUSION MATRICES

Another means to assess feature robustness is to apply a generic classification scheme to the available data and determine the probabilities of classification (P_c) which can either be probabilities of correct classification (P_{cc}) in the case of the like target class, or probabilities of false alarm (P_{fa}), for the other target classes. This classification is performed for certain sets of features. For this purpose one has to create reference feature vectors (or training data) for all available targets or target types. Then a target under test is chosen, a test feature vector determined, and the Euclidian distance

$$d_{eukl}(\alpha) = \sqrt{\sum_{i=1}^{N} \frac{(f_i(\alpha) - F_i^{ref}(\alpha))^2}{\sigma_i^2}}$$

in feature space computed between this test vector and all reference vectors. The target under test for this special feature vector then gets the label of the reference target to which its distance is minimum. This is repeated for a large number of test vectors of the respective target under test (either from a limited aspect angle interval or - as in our examples - from all aspect angles between 0° and 360°), the scores being summed up for all reference categories. The P_c values finally are determined as the ratio between the individual scores and the total number of test vectors. In this simple implementaion, we can talk of a "forced decision classifier" because the non-target case is not taken into account.

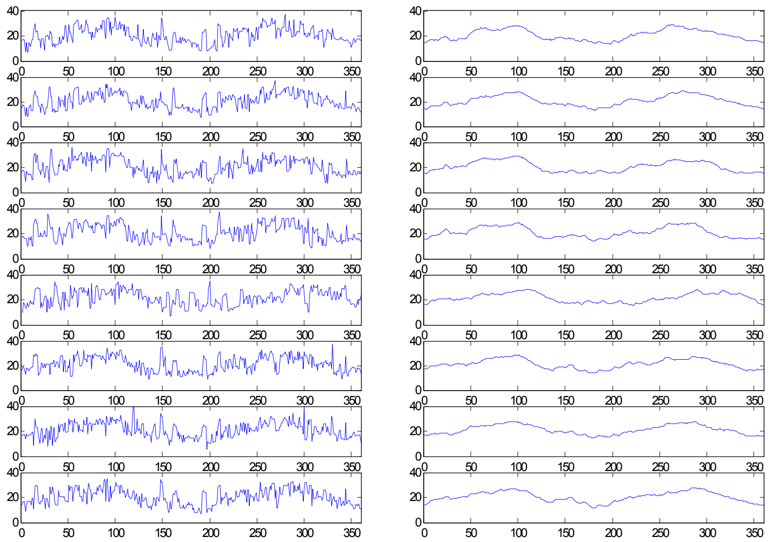

Figure 4: Feature 2 (cross-range extent for 8 T72, test values (left) and reference values (right)

Out of the 10 features analysed here, subsets of only a few of them were formed for classification purposes. The main requirement for feature selection is that they carry independent information, i.e. are statistically independent. There are several ways to achieve this goal. A common one is the principal component analysis (PCA, [11]) where, dependent on the eigenvalues of the covariance matrix, only the most "meaningful" features or linear combinations of features are retained. Another, simpler way is to determine the cross-correlation coefficients for all possible pairs of features, and then select only those sets that are essentially uncorrelated. The feature sets analysed in the following are the result of this latter procedure.

How can one create reference feature vectors? For this purpose we refer to results from former analyses [1][12] that demonstrated the importance of an independent determination of the orientation of the target under test. Thus, comparison has only to be done to reference feature vectors out of a limited aspect angle interval instead of [0°, 360°] which considerably increases the classification performance. An achievable value for the precision of pose estimation is 10° to 20° or even better. Therefore, for the present analysis, a sliding window averaging was applied to the original features over a ±10° interval with respect to each aspect angle thus creating the pertinent reference value. Fig.4 shows the effect of this averaging. The dependence of the classification result on the width of this interval, i.e. the precision with which the aspect angle can be determined, is analysed in [13].

Feature #1	Reference target class		
Test target ↓	T72	ZSU 23-4	BMP2
T72	39.7±12.6	35.8±14.0	24.4±10.2
ZSU 23-4	27.4±13.6	55.5±16.9	17.2±9.5
BMP2	26.0±15.0	20.1±11.3	53.8±19.6

Feature #2	Reference target class		
Test target ↓	T72	ZSU 23-4	BMP2
T72	47.3±5.3	26.9±5.0	25.8±4.4
ZSU 23-4	31.7±5.0	39.8±7.2	28.6±3.5
BMP2	31.2±3.7	24.3±4.4	44.6±4.1

Table 3 Means and standard deviations of P_c values (%) for range extent (#1) and cross-range extent (#2)

We have now at our disposition a total of 17 test targets (8 T72, 4 ZSU, 5 BMP) and 17 references, accordingly. If we want to construct a confusion matrix, we have to choose a triplet of test targets out of the three target classes, as well as a triplet of references. For each triplet we have 8·4·5=160 ways to choose from the given data, hence 160·160 ways to obtain a confusion matrix. It is not reasonable to perform all possible combinations, though. Instead, for a total of 500 randomly generated combinations, the confusion matrices were computed and the means and standard deviations of all matrix elements determined. By definition, the main diagonale of the confusion matrix contains the P_{cc} values whereas the off-diagonale values represent P_{fa}. All lines sum up to 100%. The P_c standard deviations can be used as another measure of robustness. The smaller they are the less sensitive the classification process is to different articulations of a target, and to the influence of selecting test and reference targets.

As an example (table 3) let us consider the geometrical features #1 (range extent) and #2 (cross-range extent). Ft.1 performs quite well in the case of the ZSU and the BMP, but almost fails for the T72. Ft.2, on the other hand, performs well for the T72 and the BMP, but less well for the ZSU. What is striking, however, are the standard deviations, which are in the range of 4%-7% for ft.2, but between 11% and almost 20% for ft.1! Obviously, the range extent is less stable and reliable and consequently less robust than the cross-range extent. This is due to self-masking (shadowing) effects that cause scatterers at the rear parts of the target to be invisible under many aspect angles. This confirms results from former analysis [14].

Ft. set #1	Reference target class		
Test target ↓	T72	ZSU 23-4	BMP2
T72	50.0±8.0	25.5±6.0	24.5±5.2
ZSU 23-4	28.9±9.4	49.8±14.7	21.3±6.2
BMP2	22.1±4.2	20.1±3.9	57.8±5.0

Ft. set #2	Reference target class		
Test target ↓	T72	ZSU 23-4	BMP2
T72	48.3±12.5	28.5±12.2	23.2±8.4
ZSU 23-4	24.3±14.9	59.3±20.4	16.4±8.4
BMP2	26.9±15.0	15.4±9.6	57.6±19.4

Table 4 Means and standard deviations of P_c values (%) for feature set #1 (fts.2,6,8,9) and set #2 (fts.1,5,7,8,9)

The second example (table 4) deals with sets of features instead of individual features. Set #1 consists of features 2,6,8,9, set #2 consists of features 1,5,7,8,9. Both show comparable performance for the T72 and the BMP, set #2 yields 10% higher P_c for the ZSU. However, set #2 has much higher standard deviations, especially for the P_{cc}(ZSU) the value is 20% which reflects a lack of reliability of the result. It may be excellent, but it may as well be insufficient, depending on the random choice of the test and reference sets. This striking difference may be due to the use of feature 1 in set #2, and of feature 2 in set #1.

Another approach to assess classification robustness is suggested by a former result [6] which states that the most stable reference (which is not necessarily the one with the best performance) is obtained by averaging the references from all available articulations of a certain target class. We will analyse this result a little more in-depth. Averaging the 8 references of T72, 4 of ZSU and 5 of BMP results in one overall reference triplet (designated by ORT). Without averaging, we can define a total of 8·4·5=160 different reference triplets as pointed out earlier. Thus we can obtain a total of 161 confusion matrices of size 17 x 3 if we test all 17 targets against all possible reference triplets. We then want to compare the 160 cases of individual reference triplets (IRT) with the ORT case.

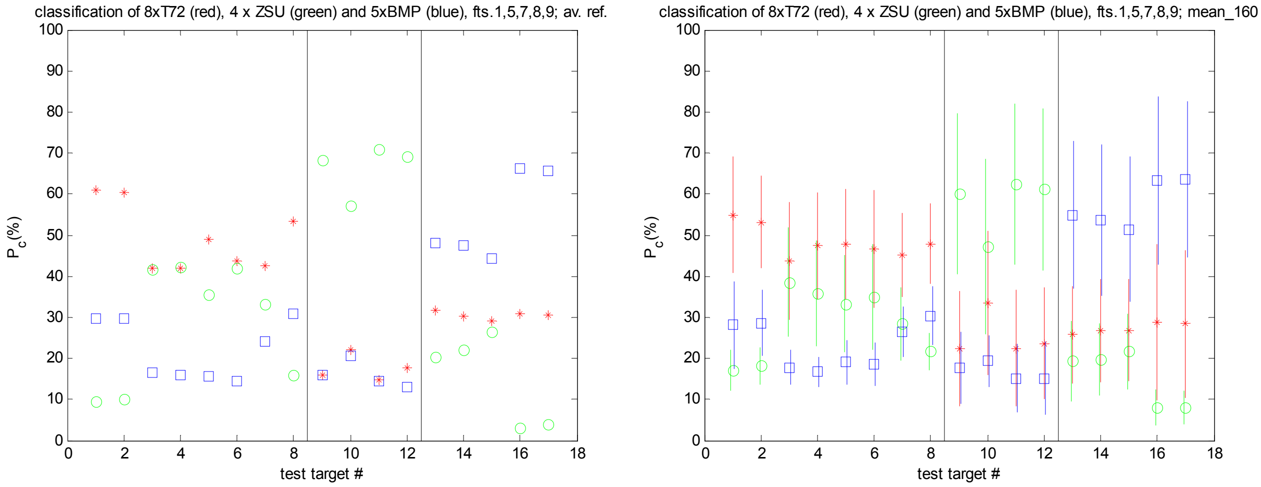

Figure 5 ORT (left) vs. <IRT> case for set of features 1,5,7,8,9

As an example we show the case of a feature set consisting of features 1,5,7,8,9. Fig.5 shows on the left side the ORT case, on the right side the <IRT> case, i.e. mean±std.dev. from all 160 combinations. On the abscissa we have the 17 test targets, subdivided by vertical lines into three classes. The ordinate represents P_c(%). We see that all 17 test targets get the highest score in their respective class, although in two cases (ORT, T72 #3 and #4) the results are very tight w.r.t a possible misclassification as ZSU. For the ZSU, the ORT results are better than the IRT average, but the large std.dev. indicates that there are IRT combinations that outperform the ORT case. For the BMP there is no clear tendency: the "weak" cases (#13,14,15) get weaker with ORT, the "strong" cases (#16,17) get stronger. For the T72, only #1,2 and 8 show an advantage for ORT, for the remaining articulations, the IRT shows a larger P_c-difference to a possible ZSZ misclassification.

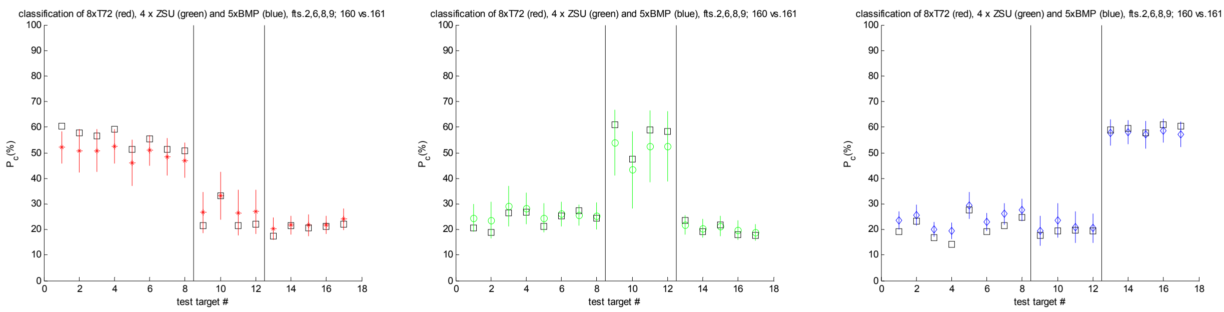

Figure 6 <IRT> results for T72 (left), ZSU(center) and BMP(right), set of features 2,6,8,9

An easier comparison is shown in fig.6, this time for a feature set consisting of fts.2,6,8,9. The diagrams show the results of testing against the T72 reference (left), the ZSU reference (center) and the BMP reference (right). The <IRT> results are indicated by their respective symbols and color, including ±1σ-error bars. The pertinent ORT results are marked by black squares. In all 17 cases, the ORT results are better than the <IRT> results, although only slightly for the BMP. – After looking in many more examples of individual features and sets of several features one may summarize that in the "robust" cases (characterized by small standard deviations) the ORT is to be preferred to the IRT, whereas in less robust cases IRT may be better, although a clear tendency is often missing. Also, one can state that the classification results in the ORT case become more homogeneous, i.e. each articulation is recognized with essentially the same probability, as was already found in [6].

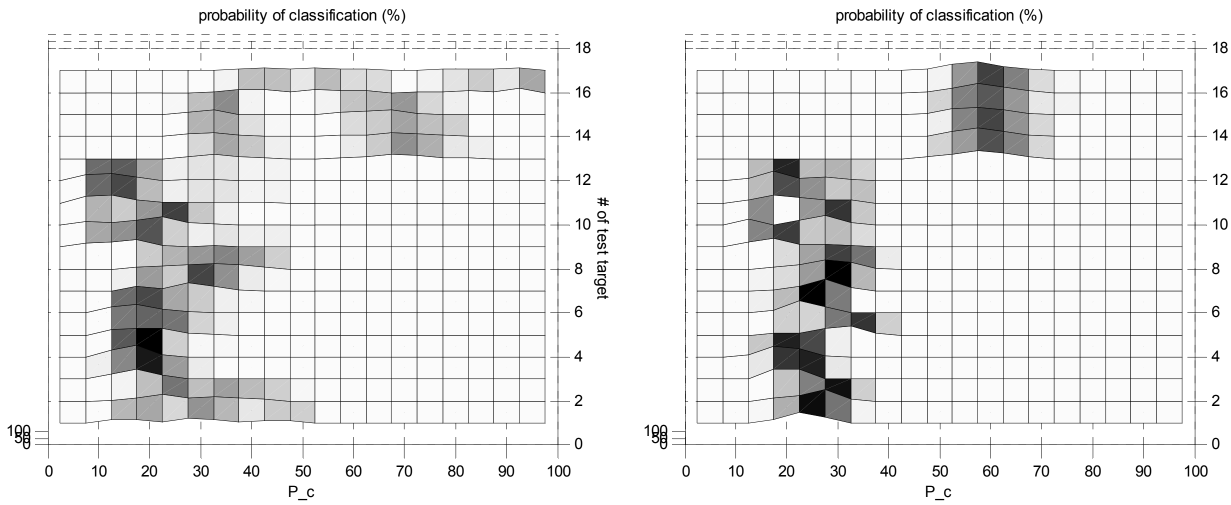

Figure 7 greyshade histograms of classification probabilities for two feature sets

It is interesting to look not only at the mean and standard deviation of the 160 IRT cases but also to study their complete histograms (fig.7). These are represented by grey-shading where 'black' means highest frequency of occurrence. We compare the two feature sets 1,5,7,8,9 (left) and 2,6,8,9 (right) where the reference in both cases is the BMP. The abscissa represents the P_c(%) values, the ordinate shows all 17 test targets. Set #1 shows some strange phenomena: first, the histograms for test targets #13-16 are bimodal, i.e.there is one group of cases with P_c up to 60-80%, another group, where P_c is only around 30% which means no classification at all. Second, the histogram in the case of test target #17 is smeared almost uniformly between 40% and more than 90%. Clearly, this set of features is not robust because it does not provide reliable and reproduceable results. Picking a random set of references can mean anything from success to failure. On the other hand, set #2 shows narrow and well defined histograms which indicate reproduceability and robustness.

6 CONSIDERATIONS OF ASPECT ANGLE DEPENDENCE

We want to conclude with some thoughts on the aspect angle dependence of classification features. For certain features, especially geometric ones like range and cross-range extent, it is clear that they will vary as a function of aspect angle. For others, like statistical or polarimetric features, it is not clear what behaviour to expect, although an aspect angle dependence should be anticipated in every case.

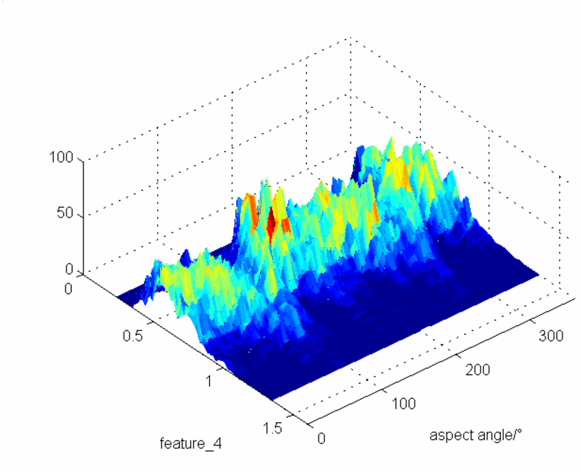

Fig.8 feature #4 histograms for sliding windows

An example is shown in fig.8 where a typical statistical feature (f4) is represented. F4 is defined as the ratio between mean and standard deviation of the 20 strongest scatterers belonging to the target, its area being declared to be the "minimum bounding rectangle" (MBR) within each 2-d ISAR image. These ISAR images are processed with angular increments of about 1/40 of a degree (as a cross-range resolution of 0.2m at 35GHz requires an angular increment of 1.2°, this means overlapping ISAR processing). Thus, an aspect angle interval of 12° which may be assumed to be a typical value for the precision with which the target orintation can be determined, gives rise to about 500 templates. The resulting feature values are transformed into a histogram which represents the f4 statistics at the respective aspect angle. Fig.8 shows the full series of histograms between 0° and 360°. As one sees, the statistics of f4 is by no means constant.

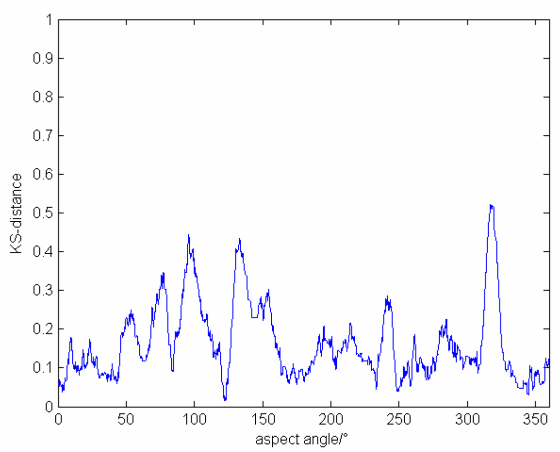

Fig.9 KS distance between global and local f4 pdf's

Fig.9 shows this KS distance between the overall pdf of f4 (out of 360°) and the "local" pdf's as a function of aspect angle. The deviation in this example can be as high as 0.5!

How does this aspect angle dependence influence the target recognition process? For each potential target, a reference vector has to be established in a multi-dimensional feature space. These reference vectors normally are derived from tower/turntable or spot SAR measurements if available. The target that has to be classified provides a test vector (or a series of test vectors during the time-on-target) that now is compared to the available reference vectors. A certain distance measure in feature space is defined, and the reference which is closest to the test vector determines the target class.

Now, if no information is available on the target orientation and the variability of the respective feature statistics, the reference vector has to be determined from the [0°, 360°] interval with no preferred aspect angle. Consequently, this reference can be unnecessarily far away from the true reference that would apply for the present target aspect. This could result in a misclassification and hence a degradation in performance.

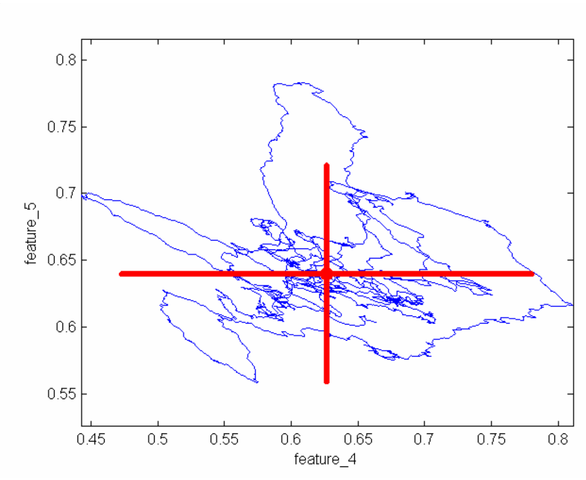

Fig.10 behaviour of features 4&5 references between 0° and 360°

This effect is demonstrated in fig.10. As a simple example, a 2-dimensional feature space is shown defined by f4 as above, and f5, where f5 is the fraction of backscatter energy that is contained in the 20 sttrongest scatterers as compared to the total energy within the MBR. The red cross marks the overall (i.e. averaged over 360°) reference vector with the pertinent standard deviations "σ" of features f4 and f5. The irregular line shows the behaviour of the "local" reference as a function of aspect angle, where each local reference is averaged over a sliding 12° interval. As one sees, the aspect angle dependent reference vectors can lie far outside the 1σ-ellipse around the overall reference vector so that using the latter clearly may lead to erroneous classification. All this underlines the importance of an independent determination of the target orientation to initialize the ATR process, as was already found in former analyses [1][2][5]. More details can be found in [13].

7 SUMMARY AND CONCLUSIONS

Three target types, namely T72, ZSU 23-4 and BMP-2 were measured in a tower/turntable configuration in 8, 4 and 5 articulations, respectively. Based on 2-D ISAR images in the VV and VH channel, a set of 10 geometric, statistical, structural and polarimetric features was calculated which was used to study the robustness of classification. The Kolmogoroff-Smirnov distance measure between histograms (pdf's) was used to define a metric that at the same time allows to quantify intra-class robustness and inter-class separability for an individual feature. For sets of several features, a simple classification approach in connection with a reference confusion matrix allows to assess the robustness of classification. At the same time this reference matrix can be used to maximize robustness by varying the free parameters of the feature definitions such that the difference of the measured confusion matrix with respect to the reference matrix is minimized. It was found that the number of scatterers N_{sc} does not offer a good potential for optimization.

As former analysis has shown the importance of an independent pose estimation of the target under test, reference feature vectors were computed as sliding window averages over +/-10° aspect intervals. It could be demonstrated further, that averaging this reference over all available target articulations improves the classification performance as compared to a reference that is based on one articulation only.

The set of 17 measurements was used to establish a statistics of the confusion matrices. The standard deviations of the P_c values vary widely depending on the type of feature or feature set. They lend themselves to be used as another metric to characterize robustness.

Finally, it was demonstrated that the feature statistics may be strongly dependent on the aspect angle of the target. As a consequence, the ATR performance has to be improved by independently determining the target orientation, e.g. by means of a Hough transform or pattern matching.

8 REFERENCES

[1] **A.C.van den Broek, R.J.Dekker, W.L.van Rossum, A.J.E.Smith, L.J.van Ewijk,** Feature Extraction for Automatic Target Recognition in High Resolution and Polarimetric SAR Imagery, TNO Report FEL-00-A236, Den Haag, Feb.2001

[2] **H.Schimpf**, Automatic Recognition of Military Targets using High Resolution Signatures at mmw frequencies, NATO RTO Symposium "High Resolution Radar Techniques", Granada, Proceedings MP-40, Nov.1999

[3] **Guy T.Maskall, Andrew R.Webb**, Nonlinear feature extraction for MMW image classification: an unsupervised approach, SPIE 2002

[4] **Adrian Britton, Keith D. Copsey, Guy T. Maskall, Andrew R. Webb and Karl West,** Nonlinear feature extraction and Bayesian mixture model approaches to target classification using MMW ISAR imagery: a preliminary study, SPIE Proc.**4033** #14, Orlando, April 2000

[5] **Albertus van den Broek, Rob Dekker, and Philippe Steeghs,** Robustness of Features for Automatic Target Discrimination in High Resolution Polarimetric SAR Data, SPIE Proc.**5094** #34, Orlando, April 2003

[6] **Hartmut Schimpf**, Millimeter Wave ATR - A Study on Feature Robustness, SPIE Proc.**5426** #28, Orlando, April 2004

[7] **NATO-RTO/SET/TG.14 and SET-069 Research and Study Groups**, List of features for Automatic Target Recognition, unpublished

[8] **H.Schimpf, H.Essen, S.Boehmsdorff, T.Brehm,** MEMPHIS – a Fully Polarimetric Experimental Radar, Proc.IGARSS 2002, Toronto, Canada, June 2002

[9] **H.Schimpf; A.Wahlen, H.Essen,** High range resolution by means of synthetic bandwidth generated by frequency-stepped chirps, El.Letters, **39**,18, pp.1346-48, Sept.2003

[10] **L.Novak,** Automatic Target Recognition using enhanced resolution SAR data, IEEE AES-35, pp.157-175 (1999)

[11] **R.O.Duda et.al.,**Pattern Classification, 2nd ed., , chapter 3, Wiley 2001

[12] **H.Schimpf,** Polarimetry and High Resolution as Key Features for Millimeter Wave Automatic Target Recognition, Proc.GRS 2000, p.283ff., Berlin, 11-12 Oct.2000

[13] **H.Schimpf, M.Hägelen,** The Influence of Target Aspect Angle Estimation on Robust Target Acquisition, Proc. "Radar 2004", Toulouse, Oct.2004

[14] **H.Schimpf,** The Estimation of Target Length from Radar High Range Resolution Profiles, Proc. Int'l Radar Symposium "IRS 2003", pp.663-668, Dresden, Sept.2003

Acknowledgement My thanks are due to my colleagues from FGAN for performing the measurements and creating an outstanding data set, and to my colleagues from NATO SET-069 Research and Study Group on "Robust Acquisition of Relocatable Targets using Millimeterwave Sensors" for numerous fruitful discussions and inspiring comments.