Classification of objects in SAR images using scaling features

Uma S. Ranjan
DaimlerChrysler Research Centre India
uma.s.ranjan@daimlerchrysler.com

Akash Narayana
DaimlerChrysler Research Centre India
akash.narayana@daimlerchrysler.com

Abstract

Synthetic Aperture Radar (SAR) is an important technique used for imaging of objects. Its strength lies in the fact that it is the only successful all-weather imaging system. However, SAR images suffer from clutter and speckle, and much research has been devoted to developing a pre-processor which can eliminate these. In this paper, we show that classification based on scaling information is naturally invariant to speckle and clutter. The methodology makes use of two kinds of scaling information in images - Hölder exponents and wavelet transform. It has been shown that these two features correspond to two different multiscale formalisms and essentially capture different kinds of behaviour. When used in conjunction with each other, they yield accurate classification on the MSTAR public domain database images.

1. Introduction

Synthetic Aperture Radar (SAR) is an air-borne or space-borne radar which obtains a "photograph" of the ground below using transmission and reception of electromagnetic energy. The resolution in the x-direction is obtained by controlling the transmitted pulsewidth. The resolution in the y-direction is obtained from the incremental doppler shift of adjacent positions on the ground. A SAR system produces a two-dimensional image of the electromagnetic scatterers within a scene wherein each pixel in the image has corresponding range and cross-range values (coordinates). SAR imaging has received a tremendous amount of research attention since it is unaffected by seasonal variations and weather conditions and is the only successful all-weather imaging system[15]. One of the main applications of SAR imaging is in military surveillance for the recognition of military targets. While this can be considered to be a classical pattern recognition problem, there are some special factors which set apart SAR image processing: Firstly, due to the inherent imaging mechanism, SAR images are corrupted by a multiplicative noise known as speckle. Secondly, reflections from the background, referred to as clutter, also affect the quality of the image. Most features extracted from optical images are not invariant to multiplicative noise, and hence, specially designed features are required for SAR image classification. If classical features such as edges or moments are considered, a pre-filtering operation is required [2, 7] to remove the speckle and clutter.

Current approaches to SAR object recognition consist of a segmentation step followed by a recognition step. Apart from the additional computation involved in segmentation, the accuracy of the overall classifier is affected by the accuracy of the segmentation algorithm. Hence, there is a need to eliminate the segmentation pre-processing step from the feature-extraction step. The scaling features presented in this paper have been shown to represent the target characteristics well and can be used directly in a classifier without an initial segmentation. Further, these features are invariant to the relative size of an object and also to rotations and translations of the target. This is of particular importance in SAR imagery, where the target is imaged at different ranges and also from various angles.

The idea of making features invariant to object sizes is captured through the multiresolution representation. Features are extracted at each level of resolution. An advantage of this method is that certain features are more obvious at certain scales than in the original image. Further, since features are present in an image at multiple scales, one can obtain a reliable and complete characterization of all the image attributes. However, a disadvantage with the multiscale representation *per se* is that each level of representation gives rise to a separate set of feature vectors. These have to be combined in a meaningful manner to obtain a good classification. Some algorithms make use of multi-level classification and combining the results[12]. However, such a proposition is computationally very expensive. Further, common attributes which may be present at different scales of resolution are largely ignored. The purpose of this paper is to use multi-scale information in such a manner that a single feature vector is obtained and one round of classification is sufficient. Thus, the algorithm proposed here not only eliminates a pre-processing routine, but also the need for multiple classifications.

Feature vectors are sometimes found to perform well with only a particular kind of classifier. In order to ensure that the feature vector captures the distinguishing characteristics of the data, it has to perform well on any classifier.

In this paper, we have implemented two standard classifiers - the Nearest-Neighbour Classifier (NNC) and the Support Vector Machine (SVM). Both classifiers were used with a minor modification wherein a separate reject criterion was incorporated. This was done to enable the classifier to reject a particular sample in case it could not be assigned with reasonable accuracy to any of the specified classes.

This paper is organized as follows: Section 2 describes scaling information in images and their importance in object classification. Section 3 describes the classification algorithm used for SAR target identification. Results and conclusions are presented in Section 4 and Section 5 respectively.

2. Scaling information in images

It has been widely recognized [9] that features "reside" at several scales in an image. As one zooms into the image, more detail is visible and the constituent length scale decreases. Often, the optimal scale at which features should be extracted is not known *a priori*. Hence, a complete characterization of an image involves representation at all its constituent scales. Images also display some common characteristics across several scales. Such information can often be used [5, 13] to study the generic nature of systems.

In a typical SAR image, the constituent scales (in decreasing order of length) are those corresponding to background, target, details of target/background and speckle noise. If it is required to distinguish between two different targets (as in the case of object recognition), both the highest and lowest scales (corresponding to background and speckle noise) do not play a role and need to be neglected. The comparison then needs to be performed on features in the intermediate scale range.

There are several ways of obtaining a multiscale representation such as Gaussian filtering, morphological transform, wavelet transform, Partial Differential Equations *etc.* Most of these methodologies represent an image at different resolutions by essentially performing a spatial averaging. The size of the image throughout the evolution remains the same. However, since coarser representations of the image carry lesser information content, they may be subsampled without loss of information. Two formalisms have exploited this idea to study systems at different scales:

- Wavelet Transforms
- Multifractals

2.1 The Wavelet transform

Mallat[11] proposed a Multiresolution Approximation (MRA) based on the wavelet transform wherein any function $f \in L^2(R)$ can be decomposed into two parts through a projection onto a space V and onto its orthogonal complement W. The space V is spanned by dilations and translations of the scaling function $\phi(x)$ and a projection onto this space is a low-pass filtering operation which retains only the low frequency components of the signal. The orthogonal complement of V is spanned by a basis generated through dilations and translations of the wavelet function $\Psi(x)$ and contains the high-frequency components of the function. The next level of approximation is obtained by decomposing the space V again into its two orthogonal complements. Thus a multiscale description of the signal can be obtained by repeated application of the wavelet decomposition. Projections onto the space V and W can both be subsampled by a factor of two without any loss of information. The energy of the wavelet transform at various scales yields local texture measurements over neighbourhoods of varying sizes. Thus, wavelets are particularly well-suited to analyse and extract local features in signals [17] which accounts for their popular use in signal and image processing.

2.2 Multifractal Formalism

The wavelet transform can capture the frequency information of a regular signal at various scales. However, singularities and irregular structures often carry essential information in an image. In order to characterize local regularities and singularities through the wavelet transform, one has to study the decay of the wavelet transform (or its modulus maxima) across different scales. Except for a small class of signals, there are no good algorithms to compute the rate of decay reliably. An equivalent method of obtaining the scaling exponents is through the notion of multifractals [1] which uses the idea of studying behaviour common to several scales and is also numerically efficient. Multifractals were first proposed [3, 4] to study systems with more than one scaling behaviour. The central concept in the notion of multifractal formalism is that of Hölder exponents $\alpha(x)$ (also called local Lipschitz exponent) which measure the local regularity of a function

$$\alpha(x) = \lim_{\epsilon \to 0} \frac{\log \mu(B_\epsilon(x))}{\log \epsilon} \tag{1}$$

where $\mu(\cdot)$ is obtaining by integrating the function over the ϵ-neighbourhood $B_\epsilon(x)$ of pixel x.

An advantage of Hölder exponents is that they can be computed for any kind of measure. Singularities of the measure can be detected as pixels whose exponents lie in a certain range [10]. Thus, they characterize both regular as well as singular behaviour. For instance, the Hölder exponent of any pixel in a constant gray-level region is 2, the topological dimension of the image. Edges, on the other hand, correspond to a Hölder exponent close to 1 [10]. This is also intuitively convincing, since edges are essentially

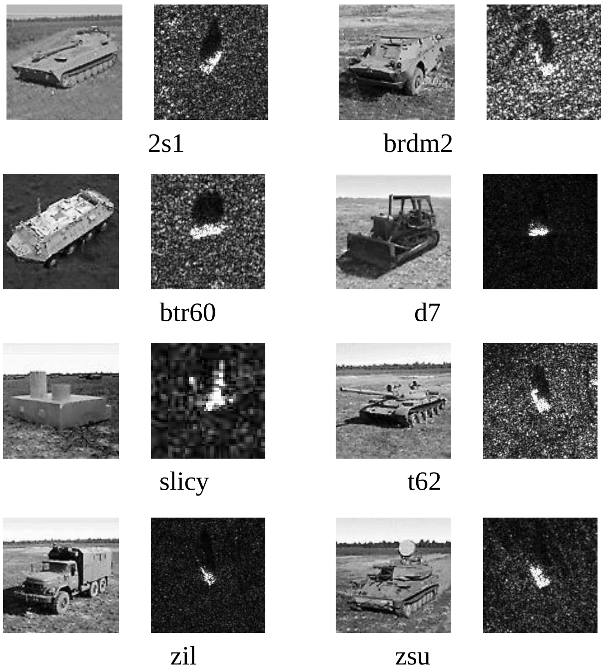

Figure 1: Sample targets and their SAR images

one-dimensional structures and carry a topological dimension 1. Hölder exponents can also be obtained through an information-theoretic characterization of an image [16, 19].

3. Classification of objects in SAR image

We now apply the principles of scaling theory to the detection of objects in SAR images. The images were taken from the public domain MSTAR library [8] and consist of 8 classes of targets : 2s1,brdm2, btr60,d7,slicy,t62,zil31,zsu[1] (sample images in Figure 1). These correspond to images of battle tanks. These targets are stationary and are imaged by a radar mounted on an aircraft at various depression and azimuth angles.

SAR image classification techniques almost invariably include a pre-filtering step [7]. However, while this reduces speckle, it also results in a smoothing of the object interiors and thus destroys any scaling information. Hence, in this work, we have not attempted any kind of segmentation. We have, however, performed a cropping of the image which encloses the target within a box which includes as little of the background as possible. The following features were considered for object detection:

Wavelet feature The image was decomposed into 5 levels, thus yielding one approximate image and 15 detail images. The average root mean square (rms) value of each image was taken to form a 16-dimensional feature vector (Figure 3). The rms value was chosen because it gives an indication of the total average strength of the singularities in an image. The Haar wavelet was chosen as the basis function. Apart from being one of the simplest wavelets, it is best suited for analysis of regular geometric objects such as military tanks. Military targets contain large regions of flat surfaces which can be represented by fewer number of coefficients.

Multifractal feature The Hölder exponent was computed at each pixel by considering the total measure inside neighbourhoods of different sizes ϵ. This was then plotted against ϵ on a double-logarithmic plot. The slope of the straight line fitted to the data yields the local Hölder exponent. The histogram of the Hölder exponents was used as the feature vector (Figure 2). Such a feature vector has the advantage of capturing local scaling information as well as global distribution of the scaling exponents.

On examining the histogram, it was found that some of the histogram values (those corresponding to the three lowest and highest Hölder exponents) were found to be identical across classes. This is also consistent with our earlier surmise that the lowest and highest scales correspond to areas outside the object and need to be neglected. Thus, the multifractal exponent captures only target scales. Scales corresponding to speckle as well as clutter are present in roughly the same magnitude in all targets and are averaged out in the computation of the Hölder exponent.

3.1 Classification

The Nearest-neighbour classifier was modified to incorporate a criterion for rejecting input patterns in case they could not be assigned to any of the classes with reasonable accuracy. The distance metric used was the Mahalanobis distance [18] $d_{\mathbf{x}i}$ between vector $\mathbf{x}$ and class i given by

$$d_{\mathbf{x}i} = (\mathbf{x} - \mu_i)\Sigma^{-1}(\mathbf{x} - \mu_i)$$

where μ_i is the mean vector of class i and Σ is the covariance matrix of class i.

The advantage in using the Mahalanobis distance instead of the Euclidean metric is that surfaces in R^d where the distance is constant are hyper-ellipsoids instead of hyper-spheres. Thus, if the cluster is non-spherical, the Mahalanobis distance follows the contour of the cluster better than the Euclidean distance. The unknown feature vector is said to belong to the class to which its Mahalanobis distance is the minimum. If this distance is greater than a threshold, the input data is assigned to the *reject* class. The threshold

[1] These are referred to as classes 1 through 8 in the results

for each class is defined to be the maximum Mahalanobis distance from any vector in a class to its class mean in the training data.

In order to show the insensitivity of the feature vector to the classifier, we have also used a Support Vector Machine (SVM)[6] for classification. SVM solves a 2-class problem by an appropriate nonlinear mapping which transforms the given feature vector to a higher dimensional vector. In this higher dimension, two classes are linearly separable by a hyperplane. John Platt's algorithm of Sequential Minimal Optimisation(SMO)[14] was used for training the SVM with parameters $C = 1000$ and $\epsilon = 0.1$. The hyperplane was computed using a Radial-basis kernel of $\sigma = 0.01$. In order to use the SVM algorithm on the MSTAR database, the 8-class problem was transformed into eight 2-class problems where the positive samples were from one particular class and the rest of the classes formed the negative samples. Thus, 8 SVMs were trained. During testing, the SVM which gave the highest positive output was deemed to be the winner. If all outputs were negative, the sample was rejected.

4. Results

In this section, we present the results of two classification algorithms using scaling features on the MSTAR image database. The results from the two classifiers gave almost identical results. Each class contains approximately 200-300 images, of which roughly half were chosen for training and the remaining for validation. The data also corresponds to the following angles of depression: $15°$, $17°$, $30°$ and $45°$

Results with the Nearest-Neighbour classifier were as follows: for a depression angle of $17°$, multifractal features gave an average classification accuracy of 85% and average misclassification error[2] of 15% (Table 1) while wavelet features gave an average classification accuracy of 86% with an average misclassification error of 9% (Table 2). On combining the two features, the classification accuracy improved to 97.2%(Table 3) with a misclassification error of 0%. Classification accuracies for other depression angles are summarized in Table 4.

The results of the SVM classifier are detailed in Table 5 for data relating to $17°$ and in Table 6 for the other angles of depression. It can be seen that the results are almost similar to those obtained with the Nearest-Neighbour Classifier. We may therefore conclude that the scaling features used are capable of distinguishing the classes irrespective of the classifier used.

Effect of background A very common pitfall in SAR object classification is that the background plays a significant role in the classification and may aid the classification. To avoid such a situation, the image was cropped to include the entire target and as little of the background as possible. Another crop of the image was taken to include only the background and classification separately performed on the two sets to determine the relative contribution of the object and background. Since the images corresponding to pure background were very small in size (sometimes just 20×30), Hölder exponents could not be computed reliably (Since these exponents are computed as a straight-line fit, it is required that the box-sizes over which measurements are taken span at least one order of magnitude.) Hence, classification in this case was performed with only the wavelet features (described in the previous section). Classification based on the background alone yielded an accuracy of 37%. Moreover, the error was distributed uniformly among the various classes (Table 7) which indicates that the background did not play a significant role in the classification.

Effect of depression angle The effect of depression angle on the classification accuracy was found to be quite pronounced. For the combined feature set, the classification accuracy was good for angles close to that on which the classifier has been trained, but poor for

[2]obtained as percentage of images that were neither assigned to the correct class nor were rejected

	Classification								
	1	2	3	4	5	6	7	8	None
1	97	1	0	2	0	0	0	0	0
2	0	72	26	1	0	0	0	1	1
3	1	20	79	0	0	0	0	0	0
4	0	0	0	100	0	0	0	0	0
5	0	0	0	0	100	0	0	0	0
6	11	0	0	1	0	66	22	0	0
7	1	0	0	1	0	18	72	7	0
8	1	5	0	0	0	1	1	92	1

Table 1: Classification using NNC with Hölder exponents

	Classification								
	1	2	3	4	5	6	7	8	None
1	96	0	0	0	0	0	0	0	4
2	0	94	1	0	0	0	0	0	5
3	0	0	91	0	2	0	0	0	8
4	0	64	0	36	0	0	0	0	0
5	0	0	0	0	97	0	0	0	3
6	0	0	3	0	0	90	0	0	7
7	0	0	0	0	0	0	93	1	6
8	0	0	0	0	0	0	0	95	5

Table 2: Classification using NNC with wavelet coefficients

	Classification (percentage)								
	1	2	3	4	5	6	7	8	None
1	100	0	0	0	0	0	0	0	0
2	0	100	0	0	0	0	0	0	0
3	0	0	96	0	0	0	0	0	4
4	0	0	0	96	0	0	0	0	4
5	0	0	0	0	91	0	0	0	9
6	0	0	0	0	0	97	0	0	3
7	0	0	0	0	0	0	100	0	0
8	0	0	0	0	0	0	0	98	2

Table 3: Classification using NNC with combined features

Depression angle	Classification	Reject	Error
15	96.4	3.6	0
17	97.2	2.8	0
30	98.4	1.6	0
45	97.7	2.3	0

Table 4: Classification accuracy using NNC for combined features

angles far from that on which it was trained. For instance, if data from 15° was used to train the classifier, the accuracy of classification was good for data corresponding to 15° and 17° but deteriorated gradually for 30° and 45°. However, if data from all angles was used to train the classifier, classification accuracy remained good for all depression angles (Table 8).

5. Conclusion

A new methodology for classification of military targets from SAR images has been presented. This method is based on a multiresolution approach and incorporates features at each level of decomposition (via the wavelet transform) as well as features that remain invariant across multiple scales (via the multifractal formalism). A common attribute with both kinds of features is that they characterize the scaling information in an image. The overall scaling dimension is represented by the energy of the wavelet transform detail signal[11]. Details of different scaling behaviour in different regions of the image are given by the multifractal spectrum. A combination of the two features has been shown to provide accurate classification. This indicates that the features capture the object characteristics accurately. In order to reduce the misclassification further, the classifiers were modified to incorporate a reject criterion. The features were found to be sensitive to the depression angle, which is to be expected since changing the depression angle changes the aspect ratio of the image and hence, its scaling behaviour. However, this problem can be overcome by using data from all depression angles for training the classifier. Further, it has been shown that the background and object correspond to different patterns of scaling and do not interfere with each other for purposes of classification. Thus, although in the present work, we have performed a clipping of the image to include only the target area, such a procedure may not be required.

	Classification (percentage)								
	1	2	3	4	5	6	7	8	None
1	100	0	0	0	0	0	0	0	0
2	0	99	0	0	0	0	0	0	1
3	0	0	99	0	0	1	0	0	0
4	0	0	0	100	0	0	0	0	0
5	0	0	1	0	99	0	0	0	0
6	0	0	0	0	0	100	0	0	0
7	0	0	0	0	0	0	100	0	0
8	0	0	0	0	0	0	2	93	5

Table 5: Classification using SVM with combined features

Depression angle	Classification	Reject	Error
15	97.4	1.9	0.7
17	98.8	0.8	0.4
30	99.5	0.3	0.2
45	98.8	0.8	0.4

Table 6: Classification accuracy using SVM for combined features

References

[1] R. Benzi, G. Paladin, G. Parisi, and A. Vulpiani. On the multifractal nature of fully developed turbulence and chaotic systems. *Journal of Physics A*, 17:3521–3531, 1984.

[2] H. Derin, P. A. Kelly, G. Vezina, and S. G. Labitt. Modelling and segmentation of speckled images using complex data. *IEEE Trans. Geoscience and Remote Sensing*, 28:76–87, 1990.

[3] P. Grassberger and I. Procaccia. *Physica D*, 13, 1984.

[4] T. C. Halsey, R. V. Jensen, L. P. Kadanoff, I. Procaccia, and B. I. Shraiman. Fractal measures and their singularities: The characterization of strange sets. *Physical Review A*, 33, 1986.

[5] T. C. Halsey, P. Meakin, and I. Procaccia. *Physical Review Letters*, 56, 1986.

[6] S. Haykins. *Neural Networks*. Prentice Hall, 2nd edition, 1999.

[7] D. H. He and L. Wang. Texture features based on texture spectrum. *Pattern Recognition*, 24:391–399, 1991.

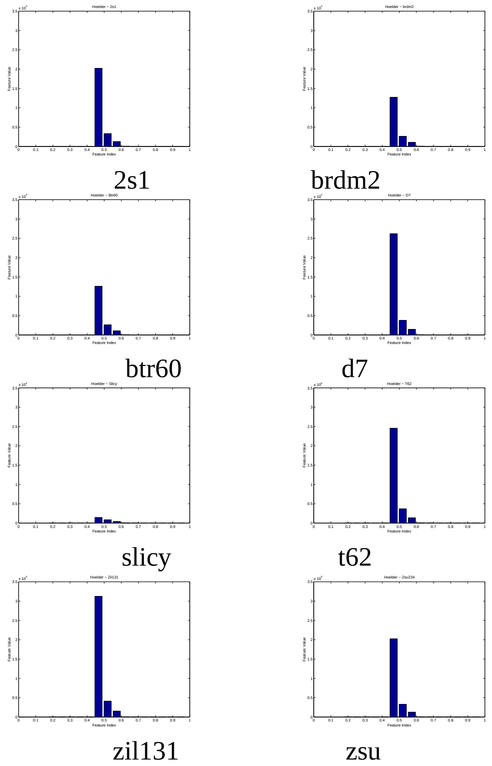

Figure 2: Sample Hoelder features

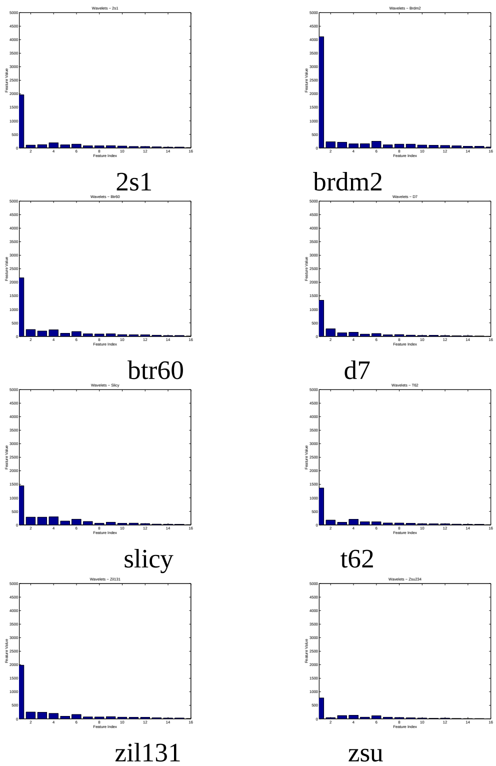

Figure 3: Sample Wavelet features

	Classification (percentage)								
	1	2	3	4	5	6	7	8	None
1	42	11	44	0	1	0	1	0	1
2	5	31	60	1	1	0	0	0	2
3	1	16	81	0	0	0	1	0	0
4	1	32	47	9	1	0	9	0	0
5	4	5	32	0	57	1	1	0	0
6	24	7	43	0	7	16	4	0	0
7	6	12	35	0	0	0	42	0	4
8	1	0	55	0	8	0	5	31	0

Table 7: Classification using NNC based on image background pixels

Depression angle	Classification	Reject	Error
15	98.9	0.3	0.8
17	99.0	0.1	0.9
30	99.2	0.3	0.5
45	98.8	0	1.2

Table 8: Classification using NNC trained with all depression angles

[8] https://www.mbvlab.wpafb.af.mil/public/datasets/mstar/MSTARmixedtgts.htm.

[9] Koenderick. The structure of images. *Biological Cybernetics*, 1984.

[10] J. Levy Vehel, P. Mignot, and J.-P. Berrior. Texture and multifractals: New tools for image analysis. Technical Report 1706, Institut National de Recherche en Informatique et en Automatique, 1992.

[11] S. Mallat. A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11:674–693, 1989.

[12] S. Mallat. *A Wavelet Tour of Signal Processing*. Academic Press, 1998.

[13] C. Meneveau and K. R. Srinivasan. *Physical Review Letters*, 59, 1987.

[14] J. Platt. Sequential minimal optimization: A fast algorithms for training support vector machines. Technical Report MSR-TR-98-14, Microsoft Research, 1998.

[15] R. Rao and P. R. Mahapatra. Synthetic aperture radar: A focus on current problems. *Defence Science Journal*, 11:517–536, 1987.

[16] A. Rényi. *Probability Theory*. North-Holland/American Elsevier, 1970.

[17] O. Rioul and M. Vetterli. Wavelets and signal processing. *IEEE Signal Processing magazine*, 8:14–38, 1991.

[18] R. Schalkoff. *Pattern recognition - Statistical, Structural and Neural Approaches*. John Wiley, 1992.

[19] K. Uma, K. R. Ramakrishnan, and G. Ananthakrishna. Image analysis using multifractals. In *Proc. IEEE Int. Conf. Acoustic, Speech and Signal Proc.*, 1996.