

Combining Features to Improve Oil Spill Classification in SAR Images

Darby F. de A. Lopes, Geraldo L.B. Ramalho, Fátima N.S. de Medeiros,
Rodrigo C.S. Costa, and Regia T.S. Araújo

Image Processing Research Group, Universidade Federal do Ceara
60455-760 - Fortaleza, CE, Brazil

{darby, fsombra, rodcosta, regia}@deti.ufc.br, glbramalho@gmail.com
<http://www.gpi.deti.ufc.br/>

Abstract. As radar backscatter values for oil slicks are very similar to backscatter values for very calm sea areas and other ocean phenomena, dark areas in Synthetic Aperture Radar (SAR) imagery tend to be misinterpreted. In this paper three feature sets are used to identify the oil slicks in SAR images. These images are submitted to different MLP architectures to verify the separability performance over each feature set. This analysis is very suitable for remote sensing of environment applications concerning marine oil pollution. The estimated resulting performance points out which feature set is the best suitable for the suggested application.

1 Introduction

Since the last decade Synthetic Aperture Radar (SAR) systems have played an important role in remote sensing of environmental disasters. These systems provide oil spills detection and monitoring, that seriously affect the marine ecosystem, providing a more rigorous and effective environment monitoring. Furthermore, SAR images have considerably contributed to understand atmospheric phenomena, land use mapping and monitoring, deforestation assessment, geographic evolution, urban growing rates assessment, agricultural crops monitoring and so on. The potential damage for the environment and economy of the area at stake requires that agencies be prepared to rapidly detect, monitor, and clean up any large spill [1]. Remote sensing of dark spots in the sea is a complex process, due to the simultaneous movement of radar and spots. The presence of an oil film on the sea surface damps out the small waves and reduces the rough surface due to the increased viscosity of the top layer and drastically reduces the measured backscattering energy, resulting in darker areas in SAR imagery [2]. The interest in appraising texture features in this work becomes from the different rough degrees presented in SAR images. Oil spill images are characterized by being less rough when compared to the similar slicks. Moreover, the procedures to extract texture features are independent of segmentation methods. The diffusion of the electromagnetic waves in the surface of the sea depends, mainly, on the rough surface which is influenced by the presence of winds, currents, waves

and parameters of the radar, such as incidence angle, frequency, polarization and resolution. The sea behaves as a specular surface when there are not waves and winds. However, dark areas might not be oil slicks but merely local wind effects or natural oil films due to low winds [3].

Automatic identification of oil spills in SAR images is a very complex task because similar images of oil spills frequently occur, particularly in low-wind conditions [4] requiring a careful interpretation. In general, the human interpreter determines if a dark object is an oil spill or a look-alike one. The contrast between oil spectral and water radiance around the oil determines which might be oil slicks. Studies have been carried out to improve methods to detect oil spills in satellite images. Liu *et al.* [5] proposed algorithms to detect and track mesoscale oceanic features employing multiscale wavelet analysis using the 2-D Gaussian wavelet transform to track oil slicks, eddies, fronts, whirlwinds and icebergs. The authors concluded that the wavelet analysis can provide a more cost-effective monitoring program that would keep track of changes in important elements of the coastal watch system. In [4] it was proposed a semi-automatic algorithm for spots detection which identifies objects in the scene with larger probability of being oil spills. A neural network approach for oil spills detection in European Remote Sensing Satellite-Synthetic Aperture Radar (ERS-SAR) imagery has been explored as an alternative tool in [2]. Del Frate *et al.* [2] proposed an algorithm to classify spots based on a set of geometric features extracted from real oil spots and look-alike ones. The input of the network consisted of a set of features regarding an oil spill candidate and the output concerns the probability for the candidate to be a real oil spill. The authors reported that the introduction of physical characteristics related to atmospheric conditions such as wind speed and water temperature could improve the algorithm results.

Concerning evaluation of feature selection issue, Jain and Zongker [6] applied feature selection algorithms to SAR images in order to classify land use combining features of four different texture models. The researchers also evaluated the potential difficulties of performing feature selection in small sample size situations due to the curse of dimensionality.

This paper proposes an analysis of the discrimination power of three different feature sets, comparing the performance of a classifier based on neural networks applied to each set: the physical-geometrical feature set generated by statistical measures on geometric characteristics [7], the texture feature set obtained as described in Bevk *et al* [8], and the third one as a composition of the previous sets. To minimize the computational effort of the classifier, principal component analysis (PCA) is used to reduce the dimensionality of each feature set and the reduced sets are compared with the original ones. The overall performance of the classifier is evaluated for different feature sets based on geometrics and texture attributes aiming at optimizing oil spills detection in SAR images. The proposed method can be used to support environmental remote monitoring.

This paper is organized as follows. The next section describes the methodology, the feature extraction process and the approach used to detect oil spills. Section 3 presents the simulation results and the last section concludes the paper.

2 Methodology

The feature data sets are generated using SAR images collected from different sources. After extracting the features from the spots, the data is divided into three feature sets and two different analysis are made: a) a classifier processes the original feature sets and b) a classifier processes the reduced sets using principal component analysis. The classifier estimated performance states the discrimination power of each set. Fig. 1 exhibits the block diagram of the previously described proposed methodology.

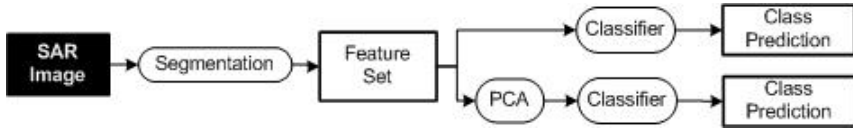


Fig. 1. General steps for feature sets evaluation of dark spots in SAR images

2.1 Feature Extraction

Texture analysis is able to provide an automatic classification of features presented in SAR images [1]. In general, texture characteristics are important for surface or object identification from aerial, satellites or biomedical images and for other applications such as industrial monitoring or product quality, remote sensing of natural resources, and medical diagnosis with tomography [9]. Despite its importance and ubiquity in image data, a formal approach or precise definition of texture does not exist [10]. The term is used to point to intrinsic properties of surfaces, especially those that do not vary smoothly in intensity. Texture includes intuitive properties like roughness, granulation and regularity. More formally, it can be defined as the set of local neighborhood properties of image grey levels [11].

Statistical information of texture characteristics is based on the representation of texture using properties governing the distribution and relationships of grey level values in the image [12]. The spatial grey level dependence matrix proposed in [13] is used to extract features, i.e., energy, contrast or entropy. In this paper the first and second order statistics of the segmented images are extracted to provide the textural features of oil spills.

The first-order probability distribution of the amplitude of a quantized image may be defined as:

$$H(g) = \frac{n_g}{N}; \quad g = 0, 1, \dots, G - 1 \tag{1}$$

where N represents the total number of pixels in the image, G denotes the number of grey levels and n_i denotes the number of pixels of grey value i in a given image. The histogram is a probability function of pixel values, therefore we can characterize its properties with a set of statistical parameters (also called first-order statistics). Many parameters may be derived from the histogram such as its mean, variance and percentiles. The following parameters are also computed:

mean (S_M), standard deviation or image contrast (S_D), skewness (S_S), kurtosis (S_K), entropy (S_{Ent}) and energy (S_E) [8].

Second-order statistics operate on the probability function ($P(i, j|d, \theta)$), that measures the probability of observing a pair of pixel values that are some vector \vec{d} apart in the image [8].

The grey level cooccurrence can be specified in a matrix of relative frequencies $P_{i,j}$ with which two neighboring pixels separated by distance d in a given direction, occur on the image, one with grey level i and the other with grey level j . Generally, the cooccurrence matrix is computed for a finite number of pixel orientations, formally for angles in intervals of 45° . The cooccurrence matrices are symmetric.

The results of the grey level cooccurrence are averaged for each angle with its transposed matrix as follows:

$$S(i, j) = \sum_{\theta=0,45,90,135^\circ} \frac{P(i, j|\theta, d) + P(i, j|\theta, d)^t}{8} \quad (2)$$

The second order statistics are extracted from the matrix shown in equation 2. Based on this matrix the following texture measures are computed: auto-correlation (A), cluster proeminence (CP), cluster shade (CS), contrast (C), correlation ($Corr$), covariance (Cov), energy (E), entropy (Ent), local homogeneity (H) and maximum probability (MAX). More detailed definitions of these features can be found in [13].

Another set of features used to describe a dark spot is extracted after the segmentation step. These measures are the physical-geometrical characteristics. Del Frate *et al* [7] state that some of these characteristics take into account the geometry and the shape of the dark spot, other part contains information about the backscattering intensity (calculated in dB) gradient along the border of the analyzed dark spot and others focus on the backscattering in the dark spot and/or in the background. The following measures, corresponding to the physical-geometrical set, are computed: area (Ar), average backscattering inside the area ($ABIA$), standard deviation of the backscattering inside the area ($SDBIA$), average backscattering outside the area ($ABOA$) and standard deviation of the backscattering outside the area ($SDBOA$). From the previous ones the following parameters are calculated: ratio between area and perimeter (AP), ratio between average backscattering inside and outside the area ($RBIO$), ratio between average backscattering and its standard deviation inside the area ($RBSDI$), ratio between average backscattering and its standard deviation outside the area ($RBSDO$), ratio between backscattering standard deviation inside and outside the area ($RSDIO$) and ratio between $SBSDI$ and $RBSDO$ ($RBSDIO$).

2.2 Principal Component Analysis

The use of more features extracted from patterns may lead to a better characterization and thus a better classification with a lower error rate, but in practice,

the opposite is observed. For a given problem the error rate initially drops with an increasing number of features, but at a certain point the error rate saturates or rises if additional features are included. This phenomenon is called curse of dimensionality. The origin of this phenomenon is the fact that classifier design relies on the inference of statistical properties from the data such as the estimation of the likelihoods or the estimation of the parameters of a distribution [14].

The problem of feature selection is defined as follows: given a set of candidate features, select a subset that performs the best under some classification system. This procedure can reduce not only the cost of recognition by reducing the number of features that need to be collected, but in some cases it can also provide a better classification accuracy due to finite sample size effects [5]. The term feature selection is taken to refer to algorithms that output a subset of the input feature set [6]. Principal components analysis (PCA) is a multivariate procedure which rotates the data such that maximum variabilities are projected onto the axes, mapping the image data into a new, uncorrelated co-ordinated system or vector space [15]. It produces a space in which the data has the most variance along its first axis, the next largest variance along a second mutually orthogonal axis, and so on. The later principal components would be expected, in general, to show little variance. These could be considered therefore to contribute little to separability and could be ignored, thereby reducing the essential dimensionality of the classification space and thus improving the classification speed. It is useful to know that due the nonlinearity of some data sets, the PCA space transformation not always leads to an optimal feature subspace. In this case further analysis using another space transformation methods are necessary to achieve better results.

3 Simulation Results

The experiments were obtained by using a set of 20 real dark spot images, where half of them are oil spill images and the other half consist of look-alike images. Figure 2a and Figure 2b are SAR image examples of a typical oil slick and a natural film, respectively. The first two sets are physical-geometrical features ($S1$) and texture features ($S2$). The third one is formed by the union ($S3 = S1 \cup S2$) of the both cited. The sets $S1$, $S2$ and $S3$ are respectively 8, 15 and 23-dimensional.

The classifiers performance assessment is shown in Figures 3 and 4. The results were obtained by running the classifier algorithm 100 times using a hold-out method varying the training size from 10% to 90% of the whole sample set. As the performance for the compound set degrades due to its higher dimensionality, we also tested different MLP architectures. Using the same N inputs and M outputs, where N is the size of the input vector and M is the number of different classes, we changed the number of neurons in the hidden layer from 2 until 20. Indeed, we experimented individual higher classification rates as the classifier fitted more the data and noise. This can be observed in Fig. 3(a) for the feature set $S3$. We decided to use the 5 hidden neurons MLP architecture, beside its

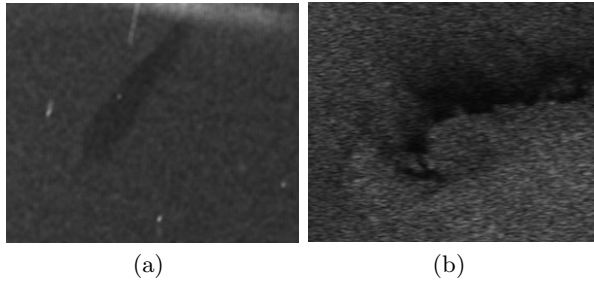


Fig. 2. SAR image examples of (a) an oil slick and (b) a natural film

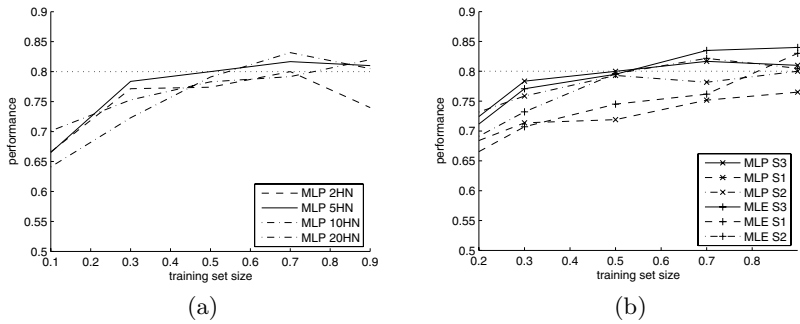


Fig. 3. MLP performance comparison (a) using different number of hidden neurons over the S3 feature space and (b) with a reference model MLE over all three feature sets

higher computational training cost, because of the generalization loss caused by *overfitting* when using the MLPs with more hidden neurons.

In Fig. 3(b) we provide a comparison between a three layer MLP with 5 hidden neurons and a *maximum-likelihood* estimator (MLE) [16] used as a reference model. The *maximum-likelihood* estimator tries to fit one gaussian probability function to each class centered on their means using unitary covariances and based on assumption of data independence. The maximum class probability is taken to assign a class label to the sample. The error probability is computed according the bayesian decision rule: $P_e = p_1P(e|C_1)+p_2P(e|C_2)$, where $P(e|C_n)$ is the conditional error probability for the input vector classified as belonging to class C_n and p_n is the *a priori* probability for the classes.

The use of PCA to reduce the dimensionality has achieved a better classification performance. Fig. 4 shows the PCA transformed data set presents a slightly better separability. Unfortunately this varies as the linearity changes from one data set to another. Thus, for S3 it is a good solution, but to the rest of the data sets the classifier performance is worse than working on the original space or quite the same.

Table 1 shows a rounded average *confusion matrix* computed from 100 classification rounds. The feature set S3 was applied to a MLP classifier with 5

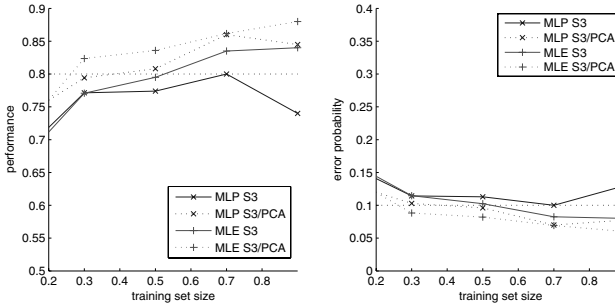


Fig. 4. MLP x Naïve Bayes performance and error probability comparison between the original space and the PCA transformed space

Table 1. Confusion matrix for *S3* feature set applied to MLP classifier with 5 hidden neurons

Predicted Class	True Class	
	C_1	C_2
C_1	9	2
C_2	1	8

Table 2. Variance comparison between original and PCA transformed feature spaces

Classifier	Original space			PCA space		
	S1	S2	S3	S1	S2	S3
MLP	0.0201	0.0255	0.0135	0.0215	0.0208	0.0172
BAYES	0.1010	0.0570	0.0712	0.0928	0.0339	0.0681

hidden neurons using 70% training size. Oil Spill samples are represented by class C_1 and the look-alike ones by the class C_2 . It is worthy of notice that this low false-alarm rate was achieved using only 20 image samples.

The classifier variances obtained in the experiments are shown in Table 2. The variances for the original feature space and for the PCA transformed ones are very similar. The result obtained by adding the texture features (feature set *S3*) has shown that a better classification performance can be reached without loss of generality. Although the Naïve Bayes classifier has achieved higher correct classification rates, as expected, the MLP has provided better generalization.

4 Conclusions

This paper presented a methodology to improve oil spill classification in SAR images. In this approach, a small set of images is described by a large number of features. Thus, for this purpose a non-parametrical classifier like MLP is more suitable than the statistical parameters based ones, like Fisher Discriminant

Analysis (FLDA) for example. This occurs because the higher-order moments, necessary to establish the discriminant, are poorly estimated which leads to errors. The maximum-likelihood estimator, used in this paper, can give only a good point of observation, which we use to compare the performances of the classifiers. The overall misclassification achieved with a MLP classifier is low enough but we have a lot of work to do in order to reduce false alarms to permit the use of this methodology in reliable marine surveillance applications. Further investigation is required to choose a more robust classifier in order to achieve a higher rate of correct classification and improve its reliability for environment surveillance applications.

The error probability is smaller as the number of training samples grows up. We believe that with a larger data set it is possible to develop a MLP architecture that can reach even higher performances. Finally, the feature sets tested on these experiments have shown that textural features provide important effect in the performance improvement for oil spill detection application. The results reported in this paper point out that the use of texture features can add significantly discrimination power for oil spill detection applications without loss of generality. This improvement is reached when using that set combined with physical-geometrical features. As the use of PCA transformation also accomplished a less complex classifier, the overall computational cost was maintained low. It is noteworthy that a very small data set was used, furthermore we concluded that any performance improvement can be a very hard task to perform with this set. Although we consider these results an advance for automatic oil spill detection systems, the misclassification rate is not lower enough. In future works, we will investigate improvements on this approach by using methods for automatic feature selection using classifier combination.

Acknowledgement

The authors would like to thank CNPq (#476177/2004-9) and FUNCAP for their financial support.

References

1. Marghany, M.: Radarsat automatic algorithms for detecting coastal oil spill pollution. *Asian Journal of Geoinformatics* **3** (2001) 191–196
2. Frate, F.D., Salvatori, L.: Oil spill detection by means of neural networks algorithms: a sensitivity analysis. *IEEE International Geoscience and Remote Sensing Symposium* **2** (2004) 1370–1373
3. Calabresi, G., Frate, F.D., Lichtenegger, J., Petrocchi, A.: Neural networks for the oil spill detection. *IEEE International Geoscience and Remote Sensing Symposium* **1** (1999) 215–217
4. Solberg, A.H.S., Dokken, S.T., Solberg, R.: Automatic detection of oil spills in envisat, radarsat and ers sar images. *IEEE International Geoscience and Remote Sensing Symposium* **4** (2003) 2747–2749

5. Liu, A., Peng, C., Chang, S.S.: Wavelet analysis of satellite image for coastal watch. *IEEE International Journal of Oceanic Engineering* **22** (1997) 9–17
6. Jain, A., Zongker, D.: Feature selection: Evaluation, application, and small sample performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19** (1997) 153–158
7. Nirchio, F., Sorgente, M., Giancaspro, A., Biaminos, W., Parisato, E., Ravera, R., Trivero, P.: Automatic detection of oil spills from sar images. *International Journal of Remote Sensing* **26** (2005) 1157–1174
8. Bevk, M., Kononenko, I.: A statistical approach to texture description of medical images: a preliminary study. *Proceedings of the 15th IEEE Symposium on Computer-Based Medical Systems* (2002) 239–244
9. Chang, T., Kuo, C.C.J.: A wavelet transform approach to texture analysis. *IEEE Transactions on Image Processing* **4** (1992) 429–441
10. Haralick, R.M., Shapiro, L.G.: *Computer and robot vision*. Addison-Wesley, New York (1992)
11. Livens, S.: *Image Analysis for Material Characterization*. PhD thesis, Universiteit Antwerpen (1998)
12. Castellano, G., Bonilha, L., Li, L.M., Cendes, F.: Texture analysis of medical images. *Clinical Radiology* **59** (2004) 1061–1069
13. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics* **3** (1973) 610–621
14. de Wouwer, G.V.: *Wavelets for Multiscale Texture Analysis*. PhD thesis, Universiteit Antwerpen (1998)
15. Richards, J., Jia, X.: *Remote Sensing Digital Image Analysis - An Introduction*. Springer (1999)
16. Webb, A.R.: *Statistical Pattern Recognition*. 2 edn. Wiley, England (2002)