International Journal of Applied Earth Observation and Geoinformation 30 (2014) 146–157

Contents lists available at ScienceDirect

International Journal of Applied Earth Observation and Geoinformation

journal homepage: www.elsevier.com/locate/jag

Oil spill detection using synthetic aperture radar images and feature selection in shape space

Yue Guo*, Heng Zhen Zhang

College of Information Engineering, Shanghai Maritime University, Shanghai, 201306 PR China

ARTICLE INFO

Article history:
Received 8 October 2013
Accepted 15 January 2014
Available online 7 March 2014

Keywords:
SAR
Oil-spill
Lookalikes
Feature selection
Shape space

ABSTRACT

The major goal of the present study is to describe a method by which synthetic aperture radar (SAR) images of oil spills can be discriminated from other phenomena of similar appearance. The optimal features of these dark formations are here identified. Because different materials have different physical properties, they form different shapes. In this case, oil films and lookalike materials have different fluid properties. In this paper, 9 shape features with a total of 95 eigenvalues were selected. Using differential evolution feature selection (DEFS), similar eigenvalues were extracted from total space of oil spills and lookalike phenomena. This process assumes that these similar eigenvalues impair classification. These similar eigenvalues are removed from the total space, and the important eigenvalues (IEs), those useful to the discrimination of the targets, are identified. At least 30 eigenvalues were found to be inappropriate for classification of our shape spaces. The proposed method was found to be capable of facilitating the selection of the top 50 IEs. This allows more accurate classification. Here, accuracy reached 94%. The results of the experiment show that this novel method performs well. It could also be made available to teams across the world very easily.

1. Introduction

An oil spill is the release of a liquid petroleum hydrocarbon into the environment, especially marine areas, due to human activity, and is a form of pollution. Oil spills are a major threat to marine ecosystems. Space-borne SAR can be used to gather information about maritime oil spills in any type of weather. Previous works have focused on the dark areas in SAR images, but oil spills are not the only things that can cause them. Darker areas may also be caused by other factors, such as currents, eddies, up- and down-welling, ice, wind fronts, sheltering provided by land, rain cells, and internal waves (Brekke and Solberg, 2005a,b; Fiscella et al., 2000; Nirchio et al., 2005; Ferraro et al., 2007, 2009; Karathanassi et al., 2006). These objects (dark areas) resembling oil spills are called “lookalikes” (Brekke and Solberg, 2005a,b; Topouzelis et al., 2008).

Oil films and lookalikes all have texture features, gray features, and frequency domain characteristics. Several researchers have used multi-features to detect oil spill from lookalikes. Fiscella et al. (2000) used 14 geometrical characteristics for both oil spill and natural features. Solberg and Theophilopoulos (1997) selected 11 features including surroundings and shape. Del Frate et al. (2000) considered 11 geometrical features in terms of its extension and its shape. A general description of the features calculated was published by Espedal and Johannessen (2000). They were the first to introduce texture features. Keramitsoglou et al. (2005) described 14 features. Karathanassi et al. (2006) described 13 features. Both these teams addressed physical, geometrical, and textural behavior. Several schemes have been designed for the purpose of unifying features with similar characteristics (Brekke and Solberg, 2005a,b; Migliaccio and Trangaglia, 2004; Montali et al., 2006). Most of the published papers that describe attempts to detect oil spills discuss features such as shape, physical characteristics, and texture. However, none of the methods described in these studies addressed the issue of which features are useful and which are not.

Because oil films and lookalikes involve different materials, they have different physical properties and may take different shapes under certain wave and current conditions. Alessandro et al. (2007) simulate shift of oil-films shape. It indicates that shape features of oil films tend to be regular. In the present work, 9 space features were selected: marking ratio, solidity, rectangular saturation, circularity, narrowness, and edge density, interior angles based on bounding polygons (IABP), Hu moment invariance, and elliptical Fourier descriptors. These features were used to detect the oil spills from lookalikes. The results suggest that not all selected features have an effect on the classification process. The purpose of the present work was to identify the most important features (IEs) of the oil spills using differential evolution and a statistical repair mechanism (Rami et al., 2007, 2011). Then test samples were re-tested with IEs.

* Corresponding author. Tel.: +86 20 38284485; fax: +86 21 38284485.
E-mail address: yueguo@shmtu.edu.cn (Y. Guo).

0303-2434/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.jag.2014.01.011

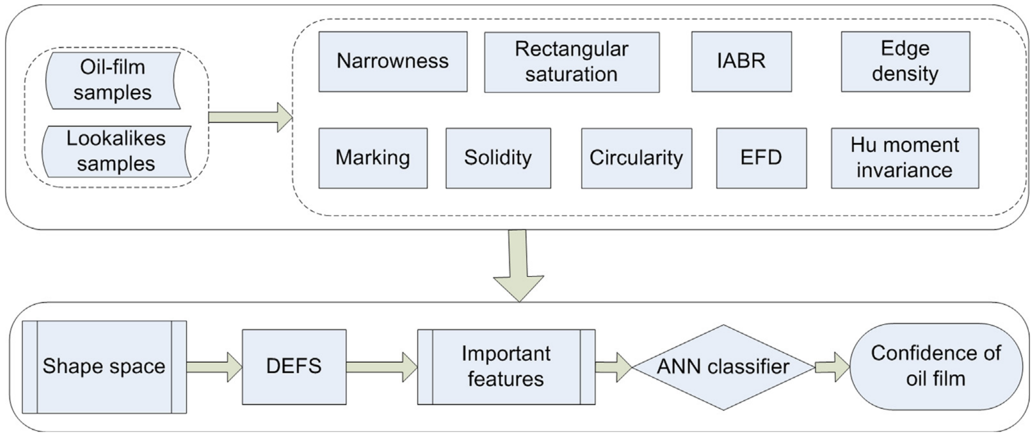

Fig. 1. Spill identification.

This paper is organized in the following manner. Section 2 introduces the method under discussion, and Section 3 shows the dataset. Results and the contributions to the detection of oil spill are discussed in Sections 4 and 5, respectively. The paper is concluded in Section 6.

2. Methods

Generally, oil spill detection systems include two parts: image pre-processing and oil spill identification. Pre-processing is mainly responsible for filtering and segmenting SAR images. It provides samples (oil-spills and lookalikes) for identification (Yue and XiaoFeng, 2011). This paper mainly addresses identification (Fig. 1). First, the total eigenvectors from 9 shape features were computed. Next, a differential evolution feature selection (DEFS) algorithm was used in the total shape feature space, and eigenvalues that were similar among oil spills and lookalikes were identified. It was here assumed that these similar eigenvalues indicated the similarity of each oil spill and lookalike and that they would negatively impact classification. Then the same eigenvalues were removed from the feature space and differential feature subsets were obtained. After comparing the importance of eigenvalues in the differential feature subsets the important eigenvalues (IEs) are made. The results of the experiments indicate that this method can discriminate oil spills from lookalikes very effectively. Compared to other methods, this method is effective, and the recognition accuracy is relatively high.

2.1. Extraction of shape features

The shape of an oil spill depends on the amount of oil, wind, and current. Oil spills and lookalikes may display different physical fluid characteristics under the same wind and current conditions. This suggests that they may also have different shape features. In the present work, 9 shape features taken from binary images are compared. A total of 95 eigenvalues were observed in the feature space.

2.1.1. Marking ratio (M)

Marking ratios describe the total area covered by all the closed areas (spots) in the image relative to that image's area. This character does not vary when translated, resized, flipped, or rotated. *Sp* is the number of pixels in each interested region; *A* is the total number of pixels in the image area. The ratio is computed as follows:

$$M = \frac{Sp}{A} \tag{1}$$

2.1.2. Solidity (S)

Solidity here refers to the degree of filling in the region contour. This feature does not vary when translated, resized, scaled, flipped, or rotated. In fact, it reflects the region's solid level. *A* is the area of the shape region; *Cp* specifies the area of polygonal approximation. Solidity is computed as follows:

$$S = \frac{A}{Cp} \tag{2}$$

2.1.3. Rectangular saturation (Rs)

Rectangular saturation is the ratio of the number pixels in the region to the number of pixels in the entire bounding box. This feature does not vary when translated, scaled, flipped, or rotated. It indicates the regional expansion of the scope. *A* is the area of the shape, and B_{area} is area of the minimum bounding box. It refers to the box with the smallest measure within all the points of dark area lie. *RS* is computed as follows:

$$Rs = \frac{A}{B_{area}} \tag{3}$$

2.1.4. Circularity (C)

Circularity is used to the complexity of the object's boundaries (van der Werff and van der Meer, 2008). The circularity of a star is much greater than that of a circle. *P* stands for the region circumference, *A* is the area. The minimum value of *C* is 1. For all shapes other than circles, $C > 1$. The more complex the shape, the greater the value of *C* is.

$$C = \frac{P^2}{4\pi \times A} \tag{5}$$

2.1.5. Narrowness

Narrowness is how narrow the shape is (Ai-Bin, 2009). Here, the narrowness of a circle is 0 and that of a straight line is 1. The greater of the narrowness, the more slender the object is. Here *a* is for the long axis of the object, and *b* is the short axis of the object.

$$e = 1 - \frac{b}{a} \tag{6}$$

2.1.6. Edge density (ED)

Edge density is average density of the edge pixels in a given region. It describes the distribution of information within an image. *P* is the pixels in the edge region, and *A* is the total area. It is computed as follows:

$$ED = \frac{P}{A} \tag{7}$$