Automatic screening of Synthetic Aperture Radar imagery for detection of oil pollution in the marine environment

Camilla Brekke

Forsvarets forskningsinstitutt/Norwegian Defence Research Establishment (FFI)

12 December 2007

FFI-rapport 2007/02391

1104

ISBN 978-82-464-1286-6

Keywords

Syntetisk Apertur Radar (SAR)

Oljesøl

Fjernmåling

Automatisk bildeanalyse

Approved by

Richard Olsen	Project manager
Jonny Bardal	Director

Populærvitenskaplig sammendrag

Utslipp av oljesøl i havet fra skip i internasjonale farvann er et alvorlig miljøproblem. Bevist dumping av olje skyldes ofte at skipene ikke tar seg råd og tid til å levere restavfallet i havn. Arbeidet presentert i avhandlingen kan føre til bedre overvåking av havområdene med hensyn på oljeforurensing.

Satellitter med Synthetic Aperturer Radar (SAR) sensorer blir brukt til observasjon av jorden, såkalt fjernmåling. Det kan gjøres SAR opptak uavhengig av vær- og lysforhold. Oljesøl er ofte synlig på havoverflata i SAR bilder. SAR egner seg spesielt godt som et kostnadseffektivt verktøy for overvåking av oljeforurensing over store havområder.

I tillegg til at havområdene som skal overvåkes er store, kommer det stadig flere SAR satellitter i bane rundt jorda. SAR bildene dekker ofte flere tusen kvadrat kilometer. Dette gir gode dekningsmuligheter, men også store mengder data som skal analyseres. Ressurskrevende manuelle analysemetoder av SAR bildene blir ofte benyttet. En ønsker å automatisere analysen av disse bildene. Et operativt system vil kunne redusere analysetiden samt gi en indikasjon på hvilke posisjoner som bør prioriteres med hensyn på videre rapportering.

En utfordring er å få automatiske system til å skille mellom faktiske oljesøl og andre fenomener som kan forveksles med olje i SAR bildene. Brekke har i sitt doktorgradsarbeid evaluert eksisterende metoder, samt utviklet nye metoder for automatisk å finne oljesøl i SAR bilder. Avhandlingen konkluderer med at automatiske metoder vil være et godt alternativ eller supplement til eksisterende manuelle metoder.

Forord

Denne rapporten inneholder en Dr. gradsavhandling som vil bli forsvart for graden *Philosophiae Doctor (Ph.D)* ved Institutt for informatikk ved det Matematiske og naturvitenskaplige fakultet ved Universitetet i Oslo (UIO) 18. januar 2008.

Doktorgradsarbeidet ble utført i perioden 2003 til 2007 ved gruppe for Digital Signalbehandling og Bildeanalyse ved Institutt for informatikk ved UIO og ved Forsvarets forskningsinstitutt på Kjeller. Mine veiledere har vært 1. amanuensis Anne Solberg (1. veileder) og professor Fritz Albregtsen ved Institutt for informatikk og professor Geir Storvik ved Institutt for matematikk (alle ved UIO). Avhandlingen ligger innenfor fagområdene fjernmåling og bildeanalyse.

Avhandlingen er en artikkelsamling. Arbeidet består av en innledning samt seks vitenskaplige publikasjoner (fire tidskriftartikler, en konferanseartikkel og ett bokkapittel).

Jeg ble tildelt et 3-årig doktorgradsstipend fra Norsk forskningsråd våren 2003 for å gjennomføre prosjektet *Algorithms for Automatic Detection of Oil Spills in SAR Images (ADOS)*. I tillegg har FFI bidratt med midler for at dette arbeidet skulle la seg gjennomføre.

Avhandlingen vil også bli gitt ut av Unipub forlag.

Preface

This thesis is submitted to the Faculty of Mathematics and Natural Science, University of Oslo (UIO) in partial fulfillment of the requirements for the degree *Philosophiae Doctor* (Ph.D).

This thesis consists of an introduction and 6 articles, Paper I - Paper VI. Paper IV has some overlap with Paper I and Paper II. The content of the articles appear in this thesis like they appear in the published and submitted versions, but the layout of the articles have been adapted to the general layout of this thesis. The articles cover topics related to oil spill detection based on remote sensing.

The study was started during the spring of 2003. In the beginning my working hours were shared between the work on this thesis and my work at Norwegian Defence Research Establishment (FFI). In 2005 my daughter Linnea was born and I had a break lasting for 8 months taking care of her.

Acknowledgements

First of all, I will like to express my gratitude to my supervisors, associate professor Anne Solberg (Department of Informatics, UIO), professor Geir Storvik (Department of Mathematics, UIO) and professor Fritz Albregtsen (Department of Informatics, UIO) for their guidance and support. A special thank to Anne, my first supervisor, for her encouragement and many highly appreciated advises. Kjell Viken at FFI is acknowledged for helping me with administrative matters making it possible to do this research in collaboration with both UIO and FFI. He also commented on several of my manuscripts. Thanks to Dan Weydahl and Richard Olsen at FFI for comments and suggestions on the introduction of this thesis. I will also like to thank my other colleagues at both the Digital Signal Processing and Image Analysis Group at the Department of Informatics and at the FFI, as many of them have helped me one way or the other during the work with this thesis. I greatly appreciate the financial support from the Research Council of Norway (the PROOF program: *Langtidsvirkninger av utslipp til sjø fra petroleumsvirksomhet*) and the FFI.

I will also like to thank my family for all their support and encouragement, specially my husband and daughter for being so patient with me.

Camilla Brekke,
Oslo, June 2007

Abstract

This thesis is focused on marine oil spill monitoring by Synthetic Aperture Radar (SAR). A survey of the methodologies and technology applied in the field of oil spill remote sensing is given. Various space-borne and air-borne sensors are discussed, and results from manual and automatic analysis approaches of SAR images are compared to results from aircraft surveillance. One of the papers included in this thesis emphasises on the experience from the Northern European seas with respect to oil pollution monitoring. Several of the other papers cover studies of automatic detection and classification algorithms for ENVISAT and RADARSAT-1 SAR imagery. The main objective of this work was to develop an automatic algorithm for detection and classification of deliberate operational oil spills.

Applying a supervised statistical classification approach, the algorithm consists of three main steps: detection of dark spots appearing on the sea surface in SAR images, extraction of a number of features from each dark spot, then the features make up a feature vector which is finally input to the classifier. In this thesis, it is shown that basing the classification decision only on the outcome of the statistical classifier produces too many false alarms, and therefore an additional step has been suggested. In this additional step, a confidence level for each dark spot classified as oil spill is estimated. This could be helpful when prioritising the alarms to be inspected by aerial surveillance.

The dark spot detector is based on adaptive thresholding of the SAR images. Improvements to the thresholding algorithm are suggested for better detection of thin, piecewise linear dark spots. Spot features are extracted from both the dark spots themselves and their surroundings. A modified feature vector, including both new and improved features, is suggested for better description of the dark spots. Taking into consideration the variations in the extracted features, the feature space is divided into subclasses. Within each subclass, a classification of the dark spots into either the *oil spill* class or the *look-alike* class is performed. It is shown that a statistical classifier with regularized covariance matrices adapted to each subclass outperforms the simpler solution of a statistical classifier with diagonal covariance matrices and C-SVM classifiers. It is also shown that when there is enough data in the training set and a regularized classifier is applied, the performance of the algorithm is not very sensitive to using subclasses to divide the feature space.

Publications included in this thesis

I. **Journal article (published):** C. Brekke and A. H. S. Solberg, *Oil spill detection by satellite remote sensing*, Remote Sensing of Environment, March, 2005, vol. 95, no. 1, pp. 1-13.

II. **Journal article (published):** A. H. S. Solberg, C. Brekke and P. O. Husøy, *Oil Spill Detection in Radarsat and Envisat SAR Images*, IEEE Transactions on Geoscience and Remote Sensing, March, 2007, vol. 45, no. 3, pp. 746-755.

III. **Journal article (submitted):** C. Brekke and A. H. S. Solberg, *Segmentation and Feature Extraction for Oil Spill Detection in ENVISAT ASAR Images*, International Journal of Remote Sensing, 2006.

IV. **Book chapter (to appear):** A. H. S. Solberg and C. Brekke, *Oil spill detection in nothern European waters: approaches and algorithms*, Remote Sensing of the European Seas (Vittorio Barale and Martin Gade (Eds.)), Springer Science and Business Media B.V., 2007. Invited contribution. Accepted based on review of full paper.

V. **Journal article (to appear):** C. Brekke and A. H. S. Solberg, *Classifiers and Confidence Estimation for Oil Spill Detection in ENVISAT ASAR Images*, IEEE Geoscience and Remote Sensing Letters, vol. 5, no. 1, Januray 2008.

VI. **Conference paper (published):** C. Brekke, A. Solberg and G. Storvik, *Classifying Oil Spills and Look-alikes in ENVISAT ASAR Images*, In proceedings of: *ENVISAT symposium*, Montreux, Switzerland, 23-27 April (ESA SP-636, July 2007). Accepted based on review of extended abstract.

Related publications

The following publications are written in the context of the PhD study, but are not included in full text in this thesis.

VII. **Conference paper (published):** A. H. S. Solberg and C. Brekke and R. Solberg and P. O. Husøy, *Algorithms for oil spill detection in Radarsat and ENVISAT SAR images*, Proc. IGARSS 20-24 September 2004, Anchorage, Alaska, vol. 7, pp. 4909-4912.

VIII. **Conference paper (published):** C. Brekke and A. H. S. Solberg, *Feature extraction for oil spill detection based on SAR images*, Proc. SCIA 19-20 June 2005, Joensuu, Finland,H. Kalviainen et al. (Eds.): *Lecture Notes in Computer Science*, vol. 3540, pp. 75-84

IX. **Conference paper (published):** C. Brekke and A. H. S. Solberg, *Classification Methods for Oil Spill Detection in ENVISAT ASAR Images*, Proc. SPIE, Image and Signal Processing for Remote Sensing XII, Lorenzo Bruzzone, Chair/Editor, (Sep. 13-14, 2006), Stockholm, Sweden, vol. 6365, 636512, 11 pages.

X. **FFI report (published):** C. Brekke, AUTOMATIC DETECTION OF OIL SPILLS BY SAR IMAGES - Dark Spot Detection and Feature Extraction, FFI/RAPPORT-2005/00893.

Contents

Chapter 1

Introduction

The focus of this thesis is on marine oil spill monitoring by space-borne Synthetic Aperture Radar (SAR).

Examples of major pollution accidents with oil tankers are the Prestige (2002) off the northwest coast of Spain, the Exxon Valdez (1989) in the north-eastern portion of Prince William Sound in Alaska and the Sea Empress (1996) in the entrance to Milford Haven, South West Wales. However, the focus of this thesis is not on large tanker accidents, but rather minor operational discharges, which according to the European Space Agency (ESA) [14] contribute with 45% of the oil pollution on a global basis (see figure 1.1). The most

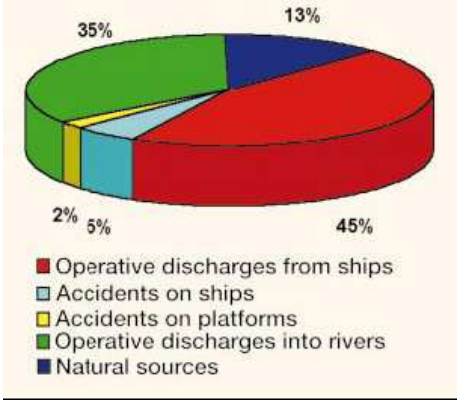


Figure 1.1: Sources of oil pollution into the seas. Source: ESA 1998 [14].

common pollution incidents occur during terminal operations when oil is being loaded or discharged, but a much greater quantity of oil enters the sea from normal tanker operations, usually associated with the cleaning of cargo residues. Other causes are tank cleaning in connection with dry docking and non-tanker accidents [49; 50].

Space-borne SAR images can be used to screen large ocean areas looking for possible oil pollution, while aircraft are needed to verify the spill and identify the polluter. The main

limitation for space-borne optical sensors is the need for daylight and cloud-free scenes. SAR imaging has the advantage of being independent of solar illumination and is generally unaffected by cloud cover. However, wind is essential in establishing the surface roughness necessary for capturing an oceanographic feature like a man-made oil spill.

SEASAT was the first civilian SAR satellite launched in 1978. Since then several nations have placed SAR satellites in orbit around the earth. The ongoing missions of RADARSAT-1 and ENVISAT, along with currently planned SAR missions (e.g., SENTINEL-1 and RADARSAT-2), mean that SAR data will be available also for years to come.

The aim of this thesis is to develop algorithms for automatic detection of possible oil spills observable on the sea surface in SAR images. The main objectives are:

- to develop a segmentation algorithm for dark spots appearing on the sea surface,
- to investigate the classification power of features and select suitable ones for the differentiation between possible oil spills and other dark structures,
- to develop a classification algorithm with respect to reducing the number of false alarms while keeping a high detection rate, and investigate the possibilities of and implement a procedure for confidence estimation,
- to test and analyse the performance of the various steps of the system with respect to expert analysis of SAR images and "ground truth" collected by aircraft.

1.1 Thesis Outline

This thesis consists of two parts:

- **Part I chapter 2-9:**
  An introduction to the main principles of SAR is given in chapter 2. Chapter 3 covers scattering mechanisms and chapter 4 covers imaging of oil pollution and its look-alikes with SAR. Chapter 5 presents the study area and the data material applied in this thesis. Chapter 6 gives a short literature survey with respect to automatic pattern recognition of oil spills. A short summary of the papers included in the second part is given in chapter 7 and the main contributions are pointed out in chapter 8. Some comments on the research done and thoughts about future research possibilities are discussed in chapter 9.
- **Part II chapter 10-15:**
  The second part includes the publications that contain the research contributions of this thesis.

Chapter 2

Imaging with Synthetic Aperture Radar

Each SAR system has its own configuration in terms of frequency, polarization, resolution, swath width etc., but the underlying operating concept for each is the same. A detailed description of the theory for the SAR is beyond the scope of this thesis, but a short introduction to the main principles of SAR is given in the following (kept on a "need to know" basis when reading this thesis).

2.1 SAR Principles

SAR is a side-looking imaging radar operating from a moving platform. A typically SAR flown on a satellite has a quite large rectangular antenna of about 10 m $\times$ 1 m (e.g. according to Curlander and McDonough [8], SEASAT had an antenna size of 10.7 m $\times$ 2.2 m, ERS-1 10 m $\times$ 1.0 m and RADARSAT-1 15 m $\times$ 1.6 m). The longest side is aligned with the orbit track and the radar beam is sent out to the side of the satellite. SAR produces two-dimensional (2-D) images. One dimension is called the *range* or across-track, the other dimension is called the *azimuth* (or along-track) and is perpendicular to the range (see figure 2.1).

2.1.1 Imaging Geometry of the SAR

Figure 2.1 shows the viewing geometry of the side-looking SAR moving in azimuth direction. The nadir is directly beneath the platform. The microwave beam is transmitted obliquely with respect to the direction of flight and illuminates a swath which is offset from nadir. Radar backscatter values are collected from a footprint area and later processed to form the SAR image. At all ranges the radar antenna measures the radial line of sight distance (slant range) between the radar and each target on the surface. The ground range distance is the true horizontal distance on the ground corresponding to each point measured in slant range.

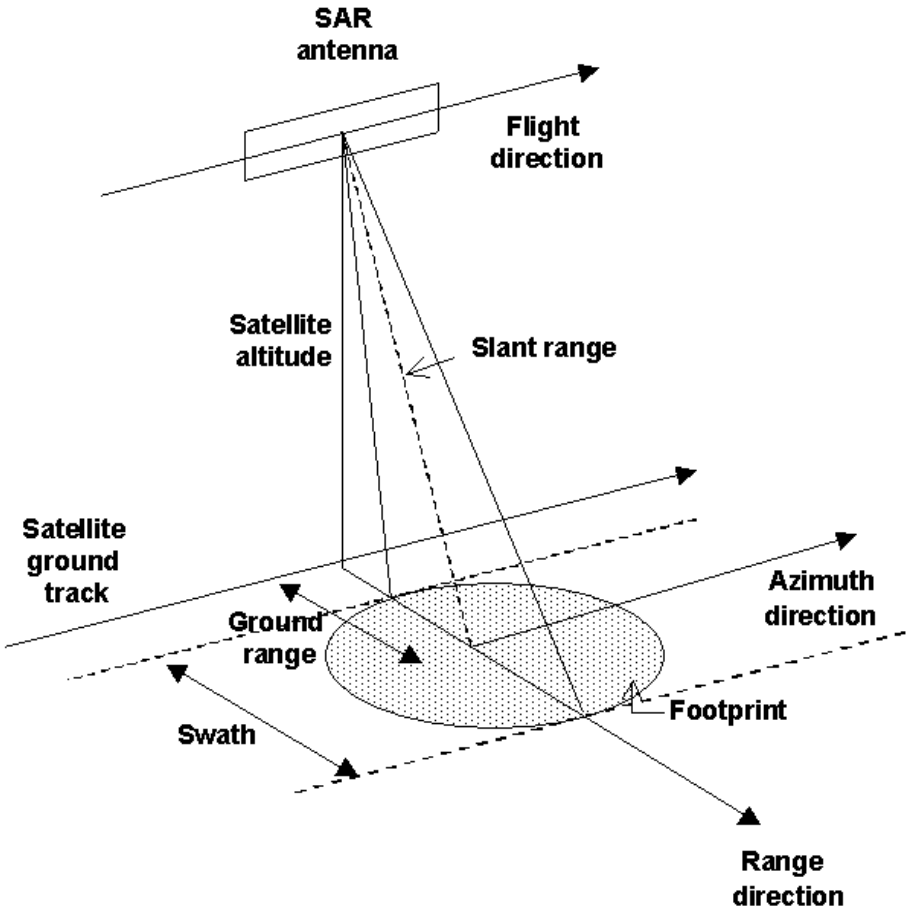


Figure 2.1: The side-looking SAR moving in azimuth direction.

2.1.2 Range Resolution

The resolution of the radar in (ground) range is defined as the minimum range separation of two objects that can be distinguished as separate by the system. If the arrival time of the pulse echo from the more distant point is later than the arrival time of the echo from the nearer point, each point can be distinguished in the time history of the radar echo. Range is determined by precisely measuring the time from transmission of a pulse to receiving the echo from a target. Objects that are located at the same distance from the SAR sensor, for a given azimuth value, will therefore be located at the same position in the SAR image. Because of this, certain geometrical effects can appear in SAR images. This is of particular importance for land applications, but of less relevance for ocean feature applications.

The ground range resolution is defined as

$$R_{ground\_range} = \frac{c}{2Bsin\theta} \tag{2.1}$$

where c is the speed of light, $B = \frac{1}{\tau}$ is the pulse bandwidth, τ is the pulse duration and θ is the incidence angle (see figure 3.3). Finer ground range resolution can be achieved by using a shorter pulse length. However, this can only be done within certain engineering design restrictions. Therefore, the radar system range resolution relies instead on the type of pulse coding and the way in which the return from each pulse is processed [8].

All radar systems like e.g. SLAR or SAR resolve targets in the range dimension in the same way, but it is the ability of SAR to produce relatively fine azimuth resolution (in the dimension parallel to the line of flight) that differentiates it from other radars.

2.1.3 Azimuth Resolution

The beam width defines the azimuth resolution. As the beam fans out with increasing distance from the radar the spatial resolution decreases. In addition to the range, the beam width depends on the antenna length. To obtain fine azimuth resolution, a physically long antenna in the along-track dimension is needed to focus the transmitted and received energy into a sharp beam. Antenna lengths of several hundred meters are often required. However, the key principle satellite SAR is to utilize the forward motion of the platform to synthesize a long antenna. As the SAR moves forward, a series of pulses is transmitted and received such that any given target on the surface is illuminated many times. The space-borne SAR then collects the data while flying and processes the data as if it came from a physically long antenna. This means that as the sensor moves along the satellite track, echoes are recorded coherently (the radar signal is recorded as a function of time) and combined in a processor to synthesize a much longer antenna (or aperture) than the physical one present. The distance the spacecraft flies while it records the reflected radar pulses from the target is known as the *synthetic aperture*. This is illustrated in figure 2.2. A target at far range will be illuminated for a longer period (due to the wider beam) of time than a target at near range. The expanding beamwidth, combined with the increased time the target is within the beam as ground range increases, balance each other. Therefore, the resolution remains constant across the entire swath.

A narrow synthetic beam width results from the relatively long synthetic aperture, which yields finer resolution than is possible from a smaller physical antenna. The resolution in the azimuth dimension is generally limited by:

$$L_a/2 \tag{2.2}$$

This simply states that the best possible azimuth resolution for a SAR system that can be achieved with a physical antenna of length L_a (azimuth dimension) is half the antenna length. This also states that improved resolution comes from smaller antennas [8].

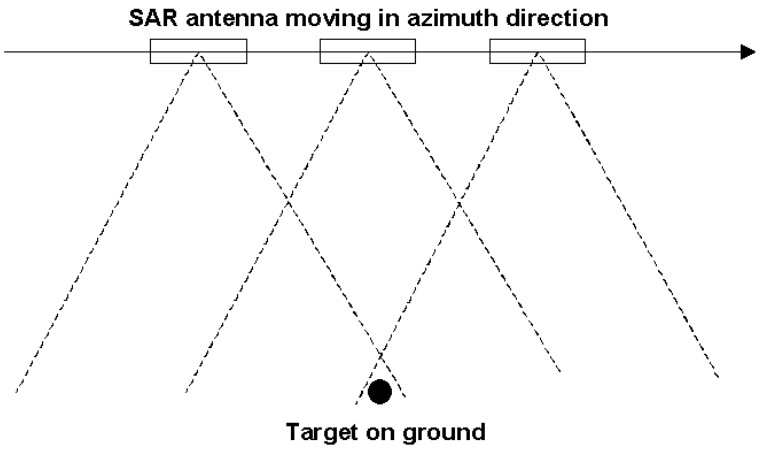


Figure 2.2: As a target first enters the radar beam, the backscattered echoes from each transmitted pulse begin to be recorded. As the platform continues to move forward, all echoes from the target for each pulse are recorded during the entire time that the target is within the beam. The point at which the target leaves the view of the radar beam determines the length of the synthesized antenna.

2.1.4 Microwaves and Polarization

The two primary factors influencing the transmission characteristics of the signals from any given radar system are the wavelength and the polarization of the energy pulse used [25].

The SAR transmits pulses of electromagnetic (EM) energy in the microwave range (wavelength: 1mm-1m) of the EM spectrum. Table 2.1 lists some of the wavelength bands. RADARSAT-1 and ENVISAT ASAR are examples of C-band SAR. According to Lillesand

Table 2.1: Radar bands. (Adapted from [25]).

Band	**Wavelength λ (cm)**	**Frequency (GHz)**
X	2.4-3.75	12.5-8.0
C	3.75-7.5	8.0-4.0
L	15-30	2.0-1.0
P	30-100	1.0-0.3

et al. [25], the wavelength of a radar signal determines the extent to which it is attenuated and/or dispersed by the atmosphere. Serious atmospheric effects on radar signals are restricted to the shorter wavelengths (less than about 4 cm). Even at these wavelengths, under most operating conditions the atmosphere only slightly attenuates the signal.

Polarization refers to the orientation of the electric field. SAR is an active sensor, and in

contrast to passive sensors it transmits a signal and measures the reflected wave. Most SARs are designed to transmit microwave radiation either horizontally polarized (H) or vertically polarized (V). Similarly, the antenna receives either the horizontally or vertically polarized backscattered energy, and some radars can receive both. Thus, we have four polarization combinations: HH - (like-polarized) for horizontal transmit and horizontal receive, VV - (like-polarized) for vertical transmit and vertical receive, HV - (cross-polarized) for horizontal transmit and vertical receive and VH - (cross-polarized) for vertical transmit and horizontal receive. Since various objects modify the polarization of the energy they reflect to varying degrees, the mode of signal polarization influences how the objects look in the resulting imagery.

SAR wavelengths and polarization combinations with respect to oil spill detection on the sea surface will be discussed in chapter 3 and 4.

2.2 Orbits and Coverage

Remote sensing satellites are often placed in *polar sun synchronous orbits*. A satellite in a polar orbit passes above or nearly above both poles of the planet on each revolution. It therefore has an inclination of (or very close to) 90° to the equator. A satellite in a polar sun synchronous orbit will pass over a given latitude at the same time every day (different for ascending and descending passes). A sun synchronous orbit also makes it possible to operate at a constant angle between the satellite solar panels and the sun. This is why many radar satellites are put in sun synchronous orbits.

RADARSAT-1 is operating in an orbit 798 km above the Earth, circling from pole to pole in a sun-synchronous orbit with an inclination of 98.6°. One orbit takes 100.7 minutes and the satellite has a 24 day repeat cycle [47]. ENVISAT carries 9 instruments, including the ASAR. ENVISAT also has a sun-synchronous orbit at an altitude of 800 km with an inclination of 98°. One orbit takes 101 minutes and ENVISAT has a 35 day repeat cycle [48]. Figure 2.3 and figure 2.4 show the number of images available for the North Sea in July 2004 for ENVISAT and RADARSAT-1.

Since orbit track spacing varies with latitude, the density of observations and the revisit rate are significantly greater at high latitude than at the equator. Coverage is also affected by the different swath widths. Current SAR systems are able to operate in different modes with different coverage and spatial resolution. Generally, high resolution modes cover smaller areas.

One method for increasing swath width is to use so-called ScanSAR imaging, where a swath widening (in range) can be achieved by the use of an antenna beam, which is electronically steerable in elevation. (The scene extension in azimuth is in principal only determined by the length of the observation period). Radar images are created by sequentially synthesizing images from the different beam positions. The area imaged by the different beams form sub-swaths. The principle of ScanSAR is to share radar operational time between two

or more separate sub swaths in such a way that full image coverage is obtained for each.

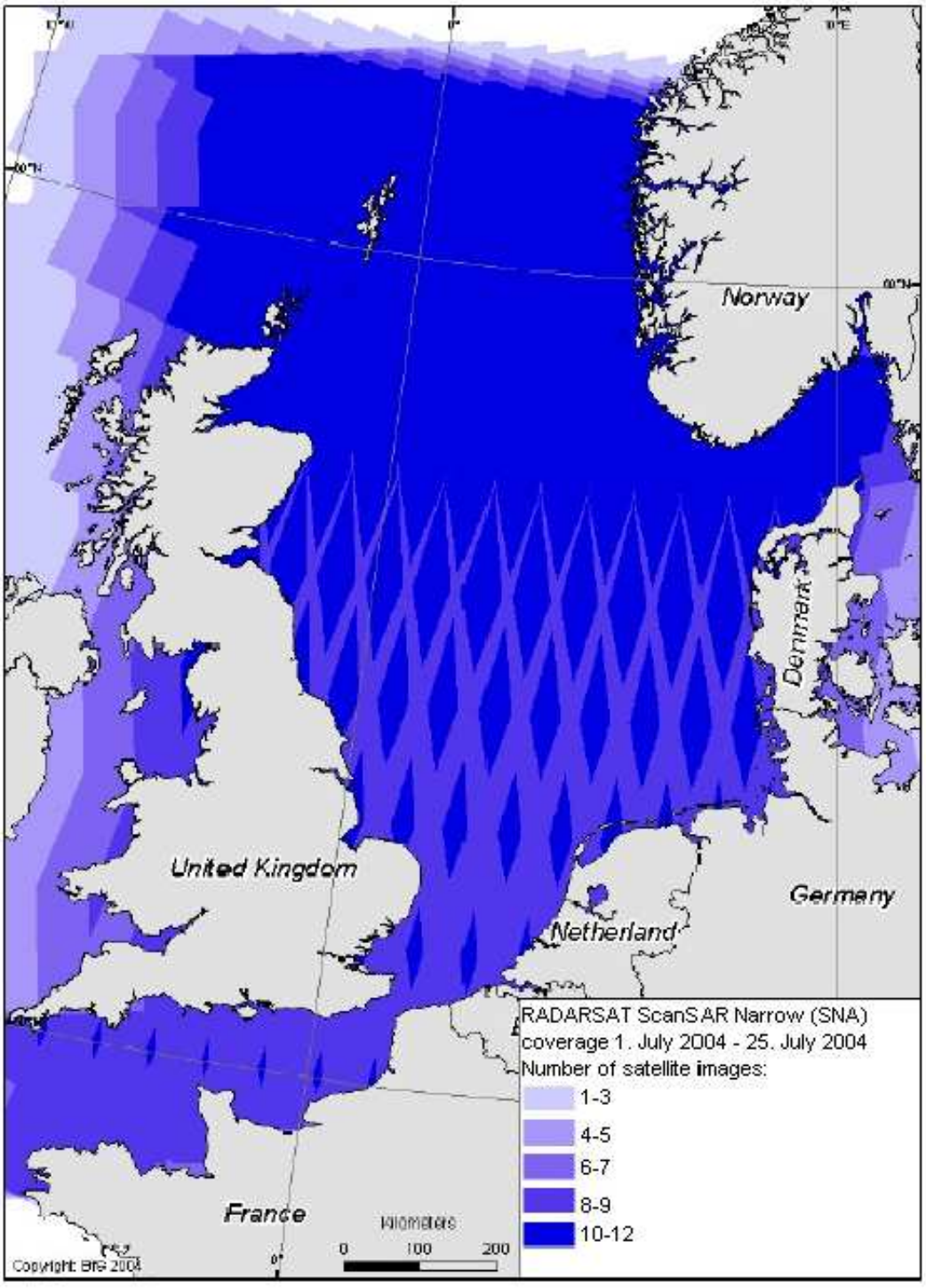


Figure 2.3: Coverage for RADARSAT-1 ScanSAR Narrow (300 km wide footprint) for the North Sea in July 2004 for one repeat cycle (1 July-24 July). It is assumed that all possible images are available (which is not usually the case in practice). Source: [43].

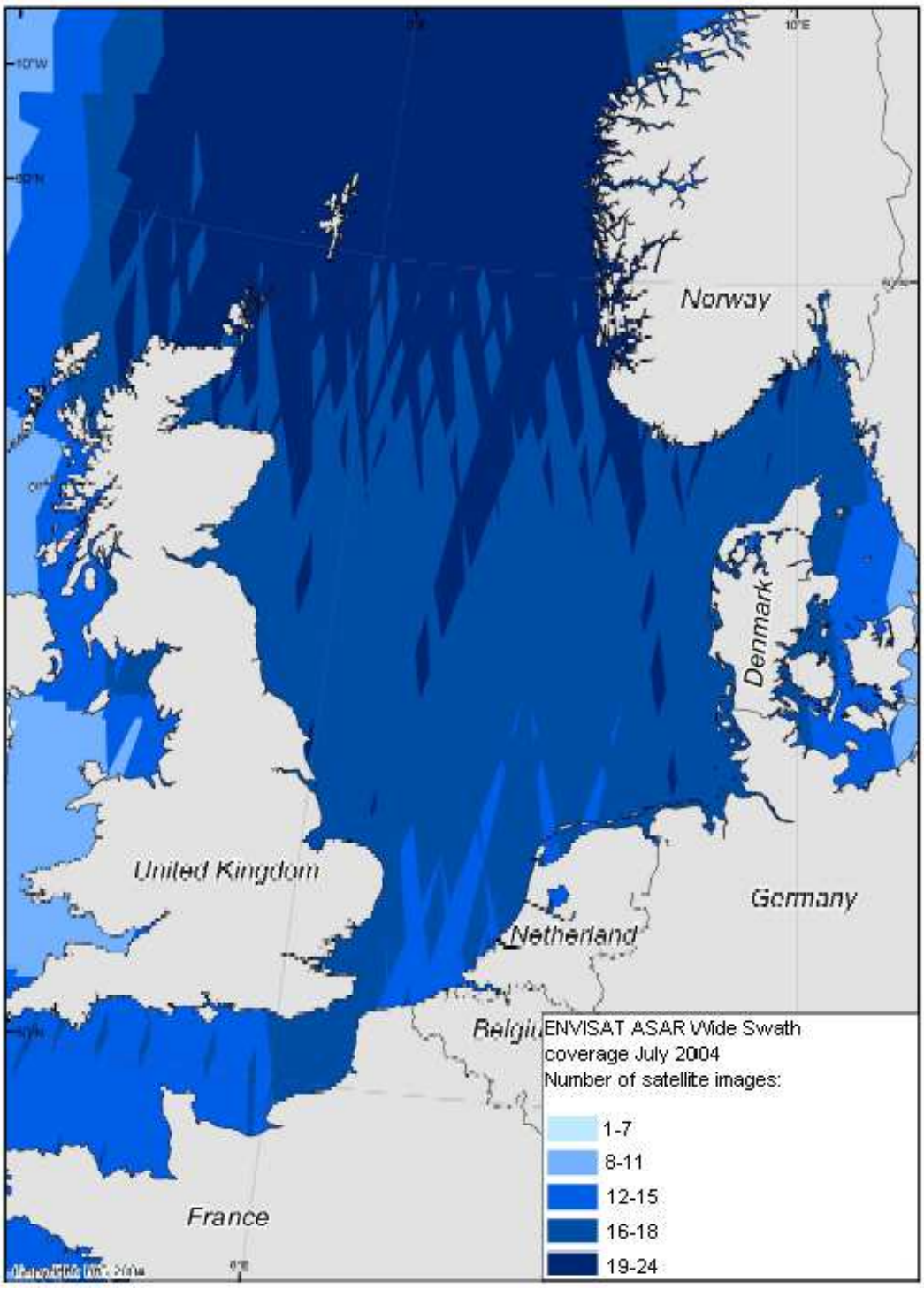


Figure 2.4: The number of images available for the North Sea in July 2004 for ENVISAT ASAR Wide Swath Mode (WSM) (400 km wide footprint). It is assumed that all possible images are available (which is not usually the case in practice). Source: [43].

Chapter 3

Scattering Mechanisms

In SAR imaging, there are several important factors that decide how strong a signal is reflected back from the target area. These factors can be divided into satellite system factors:

- the radar beam incidence angle
- the radar wavelength
- the polarization of the radar

and ground surface factors:

- the roughness of the surface
- the geometrical structure of the surface
- the dielectric properties of the surface
- the wind speed
- the angle between the radar beam and the wind

The intention of this chapter is to provide an overview of target scattering mechanisms as a fundament for the discussion of oil spill detection in the following chapters of this thesis.

3.1 Surface Scattering

For flat terrain, the local reflection angle is the same as the incidence angle as shown in figure 3.1 a). Most of the incident energy will be reflected away from the sensor, resulting in a very low return signal. Rough surfaces will scatter incidence energy in all directions and

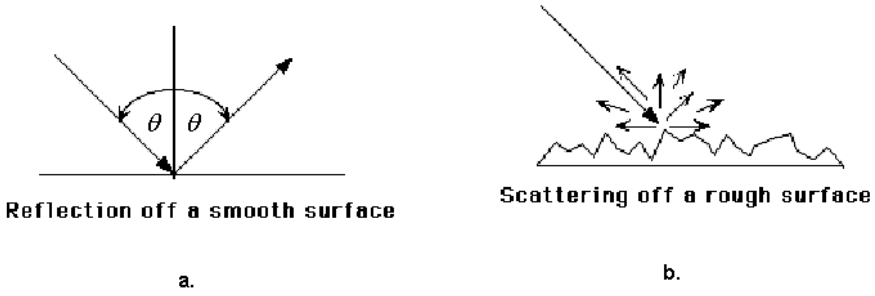


Figure 3.1: Scattering mechanisms. a) Reflection off a smooth surface. b) Scattering off a rough surface. (Adapted from [3]).

return a significant portion of the incident energy back to the antenna. This is illustrated in figure 3.1 b).

On the ocean surface it is the waves that make the surface rough. Whether the surface is perceived rough or not, depends on the wavelength of the SAR.

3.1.1 Bragg Resonance Model

The ocean surface is known to contain a spectrum of waves from short ripples of a few millimetres to waves hundreds of meters long. However, it is generally accepted that the dominating mechanism at work to support the backscattering is a type of Bragg resonance. The particular application of the Bragg resonance model to the ocean surface, which is a complex summation of a wide spectrum of different wavelengths, requires the assumption that the Bragg mechanism is able to select just those waves that are in resonance. In terms of the ocean wavelength, λ_w, this means that:

$$\lambda_w = \frac{n\lambda_r}{2sin\theta}, n = 1, 2, ... \tag{3.1}$$

defines the wavelength of the Bragg-selected waves. θ is the incidence angle and λ_r is the radar wavelength. (The dominant return will be for the wavelength where $n = 1$ [8]). Note that to be selected by the resonance, the Bragg waves need to propagate toward or away from the look direction of the radar antenna. Equation 3.1 implies that the surface waves which influences the radar backscatter are those of comparable wavelength to the microwaves. It is the short gravity and capillary-gravity waves to which the radar responds directly. The Bragg condition also implies, for a given SAR, that the resonant surface waves will be shorter at more oblique incidence angles. This also relates to the general observation that the backscatter for a given sea state decreases with increasing incidence angle [36] (the backscattered radar power is proportional to the spectral energy density of the Bragg waves and the spectral distribution decays at shorter wavelength), as will be discussed in section 3.3.

For RADARSAT-1 and ENVISAT ASAR with C-band frequency, a radar wavelength of 5.7 cm and incidence angles in the range of $20° - 50°$ will this model give Bragg resonant sea wavelengths λ_w in the range of 8.3-3.7cm.

In Equation 3.1, the Bragg resonant wave has its crest at right angles to the range direction. For surface waves with crests at an angle ϕ to the radar look direction we get:

$$\lambda_w^{'} = \lambda_w sin\phi \tag{3.2}$$

where $\lambda_w^{'}$ is the wavelength of the surface waves propagating at angle ϕ to the radar look direction. An illustration is given in figure 3.2. The resonant surface wavelengths will increase when ϕ increases.

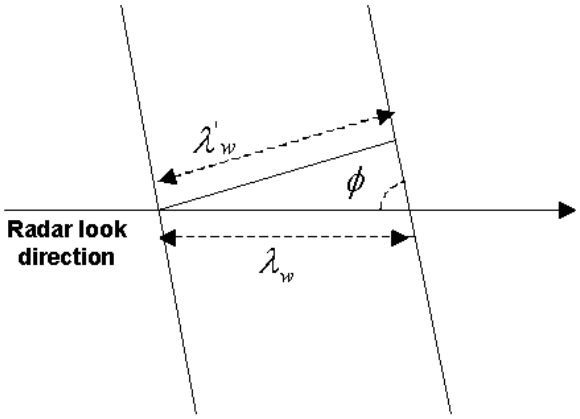


Figure 3.2: Crests at an angle ϕ to the look direction of the SAR.

3.2 Volume Scattering

Bragg models are most frequently used for describing scattering from the sea surface. Due to the large dielectric constant of water ($\epsilon = 80$ [12]), the scattering mechanism is exclusively surface scattering. Target areas that can be characterized by Bragg scatter are essentially special examples of the general scattering problem, which is significantly more complex. Thus, scattering from natural terrain and vegetation is generally a combination of surface scattering and volume scattering. Volume scattering results from dielectric discontinuities within the media [8]. Volume scattering will not be treated further in this thesis.

3.3 Range Brightness Variation

The radar incidence angle is defined relative to the vertical plane, and is thus smaller at near range compared to far range. Figure 3.3 gives an illustration. SAR images tend to

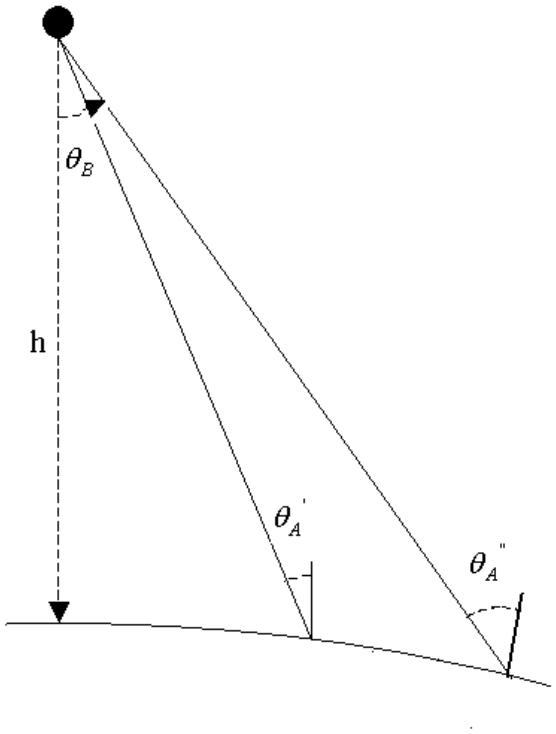


Figure 3.3: The viewing geometry of the SAR. θ_A = incidence angle (angle between the radar beam and the normal to the Earth's surface at a particular point of interest) and θ_B = look angle (angle at which the radar "looks" at the surface, relative to the vertical).

become darker with increasing range. Backscatter is related to the local incident angle (i.e. as the local incident angle increases, backscatter decreases), which is in turn related to the distance in the range direction. Backscatter is also related to wind speed. This is illustrated in figure 3.4, where a plot based on the CMOD-model [26] is presented. To some degree, mathematical models can be used to compensate for this effect [25].

Figure 3.5 is an ENVISAT ASAR image (no correction with respect to the incidence angle variations is done) with a difference in incidence angle from near to far range of about $16.9° - 41.6°$.

3.4 Speckle

The resolution cell of any SAR is large with respect to a wavelength of the radar system [8]. *Many* individual scatter will contribute to the radar echo coming from a particular resolution cell. Coherent processing of the scattered signals generates SAR images, and this gives the scenes a speckled appearance. Speckle produces a seemingly random pattern of brighter and

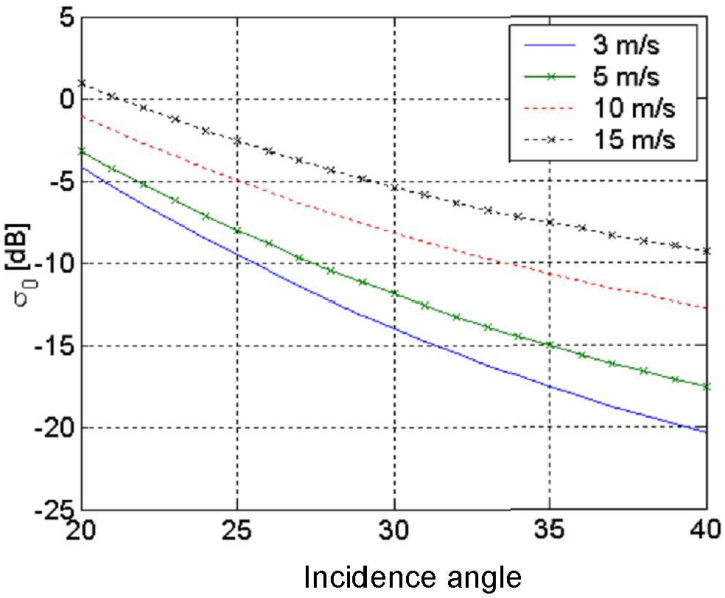


Figure 3.4: Backscattering, σ^0, from the sea estimated for different wind speeds, upwind, C-band and VV polarization. Source: FFI.

darker pixels in SAR images, giving a grainy appearance or speckle [25]. Speckle noise is multiplicative in nature, which means that the noise level increases with the average intensity. This is also apparent in the plot in figure 3.5

There are several ways to reduce the amount of speckle and thereby make the SAR image more interpretable. One solution is to subdivide the synthetic aperture and process independent images - or looks, from the SAR raw data. By averaging these uncorrelated looks (*look averaging*), the standard deviation will be reduced and hence smooth the image appearance. The signal-to-noise (S/N) ratio will also increase, but at the cost of poorer spatial resolution in the azimuth direction. Multi-look processing is normally done at the SAR processing facility.

Another way of smoothing the speckle noise is to do some kind of post-processing on the SAR image itself. A smoothing technique often applied in image processing is convolution (e.g., simple mean filters). More sophisticated filters are adaptive in the sense that they take into account the speckle statistics by using the mean and variance of the SAR image pixel values within a local region. The advantage with adaptive filters is that they can smooth a lot in large homogeneous regions, while inhomogeneous areas and edges are preserved.

A detailed discussion about the difference between various speckle filters is beyond the scope of this thesis. Generally the choice of speckle filter type is governed by the nature of the application, requirements on the processing time and the SAR image product involved [46].

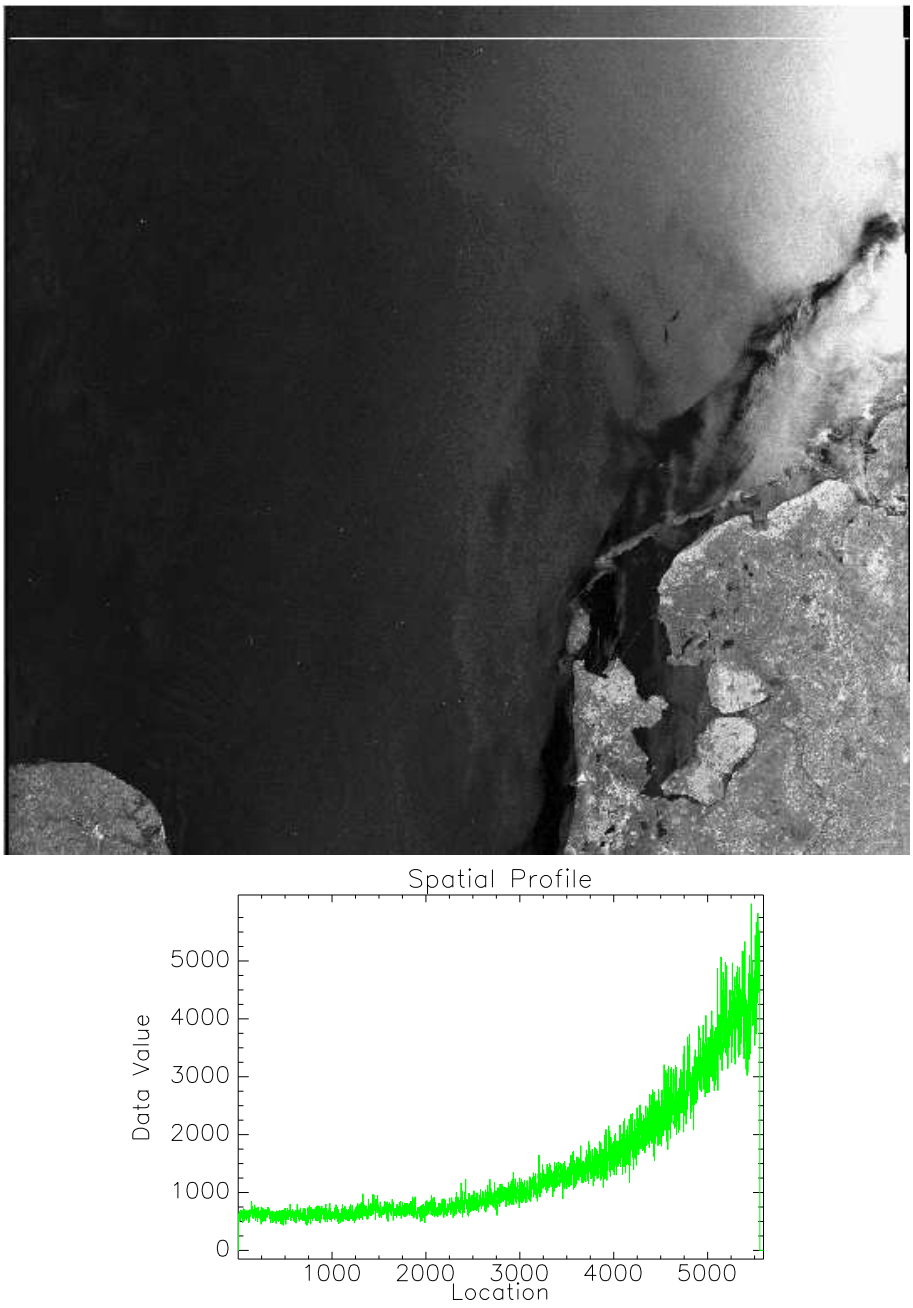


Figure 3.5: Top: ENVISAT ASAR WSM image from 28th of January 2005. Original data ©ESA/KSAT. From right to left: near to far range. Dimensions: 5596×5359. Bottom: Range profile at the white line (row 200).

Chapter 4

Oil Spill Imaging with Synthetic Aperture Radar

A wide range of ocean surface phenomena have been imaged with SAR. In addition to oil spills, several phenomena may dampen out the Bragg waves sensed by the SAR. When these waves are dampened, very little of the emitted signal will return to the SAR. A dark area will therefore appear in the SAR image.

4.1 Imaging of Low-backscattering Ocean Features

Both atmospheric processes that affect surface wind conditions (and thus the generation and modulation of Bragg waves (discussed in section 3.1.1)) and oceanic processes that directly modulate the Bragg wave spectrum, produce signatures imaged by SAR [6].

In general, lower wind speeds generate fewer Bragg waves. This produces a smoother ocean surface that appears in the SAR imagery as a dark area. Below a low wind speed threshold, little of the radar energy will be scattered back toward the SAR, and features depending on the modulation of Bragg waves to be imaged will not be visible in the image.

As discussed in section 3.3, the sea surface backscatter decreases rapidly with increasing radar incidence angle. Success in detecting low-backscatter ocean features may well depend on where the features lie within a swath. The wider the swath, the larger the contrast between near and far range.

The wind direction relative to the plane of the incident radar wave also affects the backscattering in a scene. A crosswind (wind blowing perpendicular to the range direction) produces lower backscattering than an upwind or downwind (wind blowing along the range direction) (see also section 3.1.1).

4.2 Man-made Pollutants

Man-made biogenic slicks are produced by the discharge of organic matter resulting from human activities such as fish processing, while *man-made mineral slicks* are typically caused by accidental spills or the illegal dumping of petroleum products.

According to Espedal [13], man-made pollutants could be:

I. *Oil spills.* Oil spills from ships may be due to accidents, and operational cleaning of tanks often occurs. Oil spills from oil installations may occur as a result of technical failure or operational errors.

II. *Produced water from oilrigs.* This is water coming from the reservoir together with the oil. It is cleaned before being released into the sea, but it will still contain some traces of chemicals and oil.

III. *Water-based or oil-based drilling mud.* This is used when drilling wells. After being reused several times, the oil-based mud is brought ashore for treatment, but the water based mud is released into the sea and may give a slick signature in the SAR image.

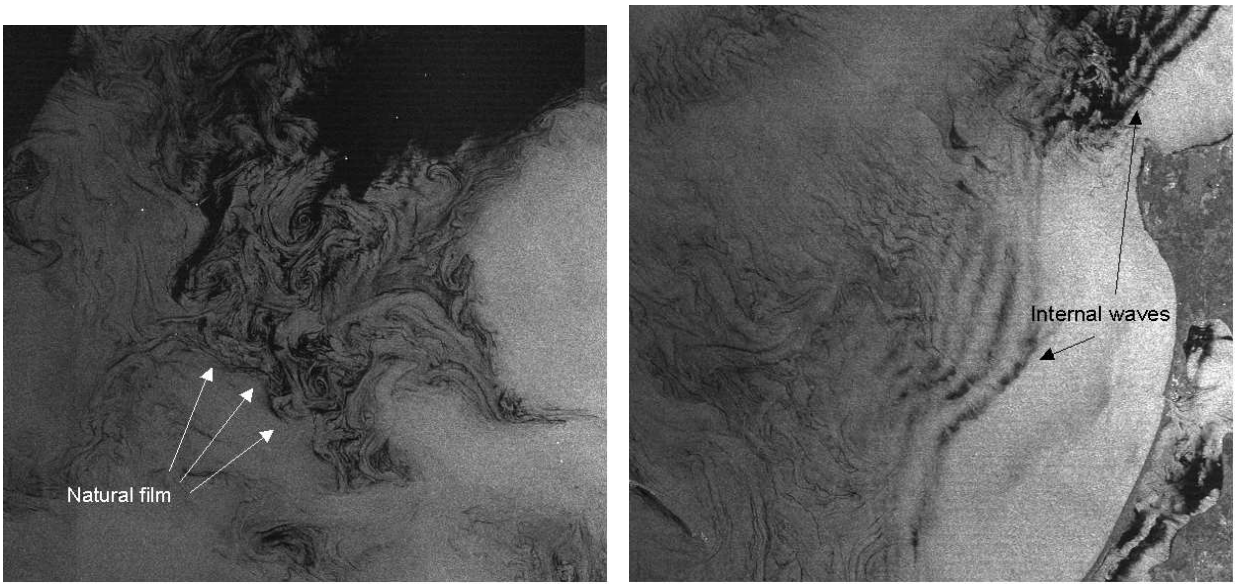
IV. *Drain water.* The platform drain water consists among other things of rainwater. It is collected through an open drain system and most of it is separated from possible oil contamination.

According to Robinson [36], an important consideration to acknowledge is that within the constraint of present technology and knowledge, there is no way of determining from SAR data alone what is the cause of a surface slick observed in a SAR image (see however a discussion of full polarimetry in chapter 9). However, as this thesis shows, based on expert knowledge of SAR image analysis and statistical methods applied to the problem, we can estimate the confidence of a slick in a SAR image being an actual oil spill or not.

4.3 Look-alikes: Dark Structures Resembling Oil Spills

The following are examples of natural phenomena that can create oil spill *look-alikes* appearing in SAR imagery [6; 13]:

- **Natural biogenic surfactants/natural film:** Natural biogenic slicks are produced by plankton and fish substances normally released into the environment. Surfactants accumulate in convergent zones by internal waves and current/eddy fields, but are mixed into the upper ocean and rapidly disperse and disappear under windy conditions. Fresh-water run-off containing biogenic material can also cause natural slicks. Figure 4.1(a) shows an example of natural film.



(a) Occurrences of natural film.

(b) Internal waves.

Figure 4.1: Parts of an ENVISAT WSM Image from 22nd of September 2005. Original data ©ESA/KSAT.

- **Natural mineral surfactants:** Natural mineral slicks are the result of ocean-bottom oil seeps. (This phenomena should only be considered a look-alike if we are strictly looking for man-made oil pollution).

- **Grease ice:** Sea ice can also dampen ocean surface waves. In particular, grease ice (composed of small crystals that form when seawater begins to freeze) dampens Bragg waves and produces areas of extremely low backscattering. As it accumulates on the sea surface, grease ice forms slick patterns similar to those produced by mineral or biogenic surfactants. Figure 4.2 shows an example of grease ice, and how polarimetry can be used to reveal that the dark spots are look-alikes.

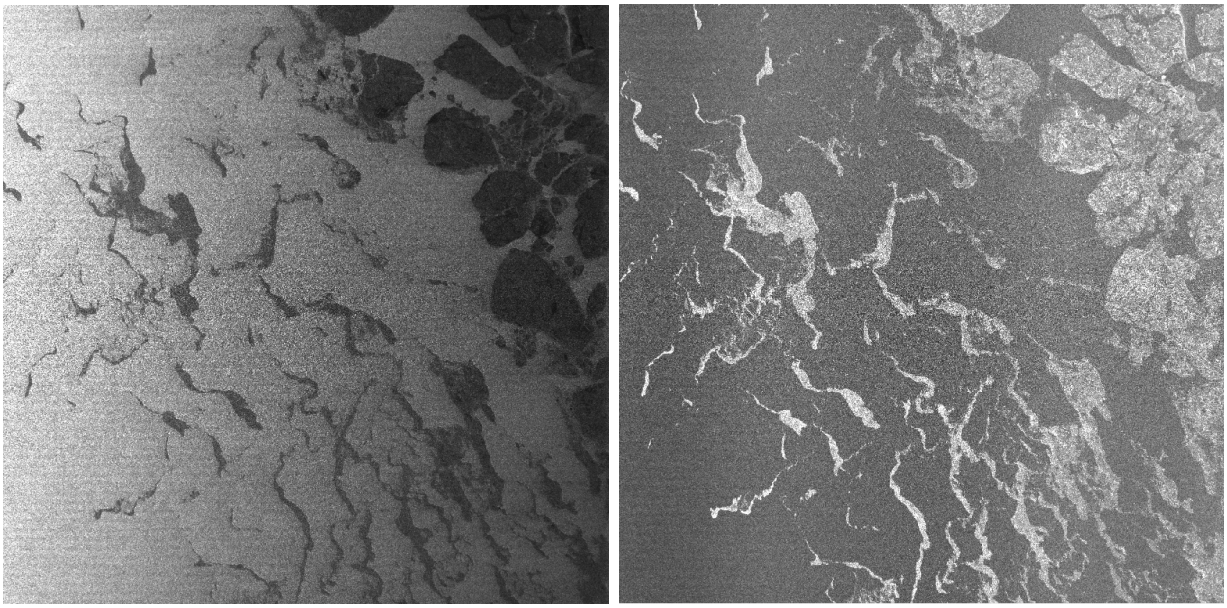


Figure 4.2: ENVISAT Alt Pol Mode Precision Image. 9th of November 2004. The dark patches in the HH-polarized image to the left are thin ice. On the HV polarized image to the right it is clear that these patches are ice and not mineral or biogenic slicks because of their bright reflection. Original data ©ESA/KSAT.

- **Low surface winds:** As the sea surface roughness is dependent on the wind conditions, an often seen feature imaged by SAR over the ocean is the wind speed variability itself. Dark areas appear with wind speeds below the threshold wind speed of about 3 m/s (which is the threshold for generation of Bragg waves). Areas of wind shadowing by coastal topography are also commonly observed in SAR imagery. The islands and high mountains shelter the water surface from the wind, and the Bragg wave growth is reduced on the lee side. An example is shown in figure 4.3.

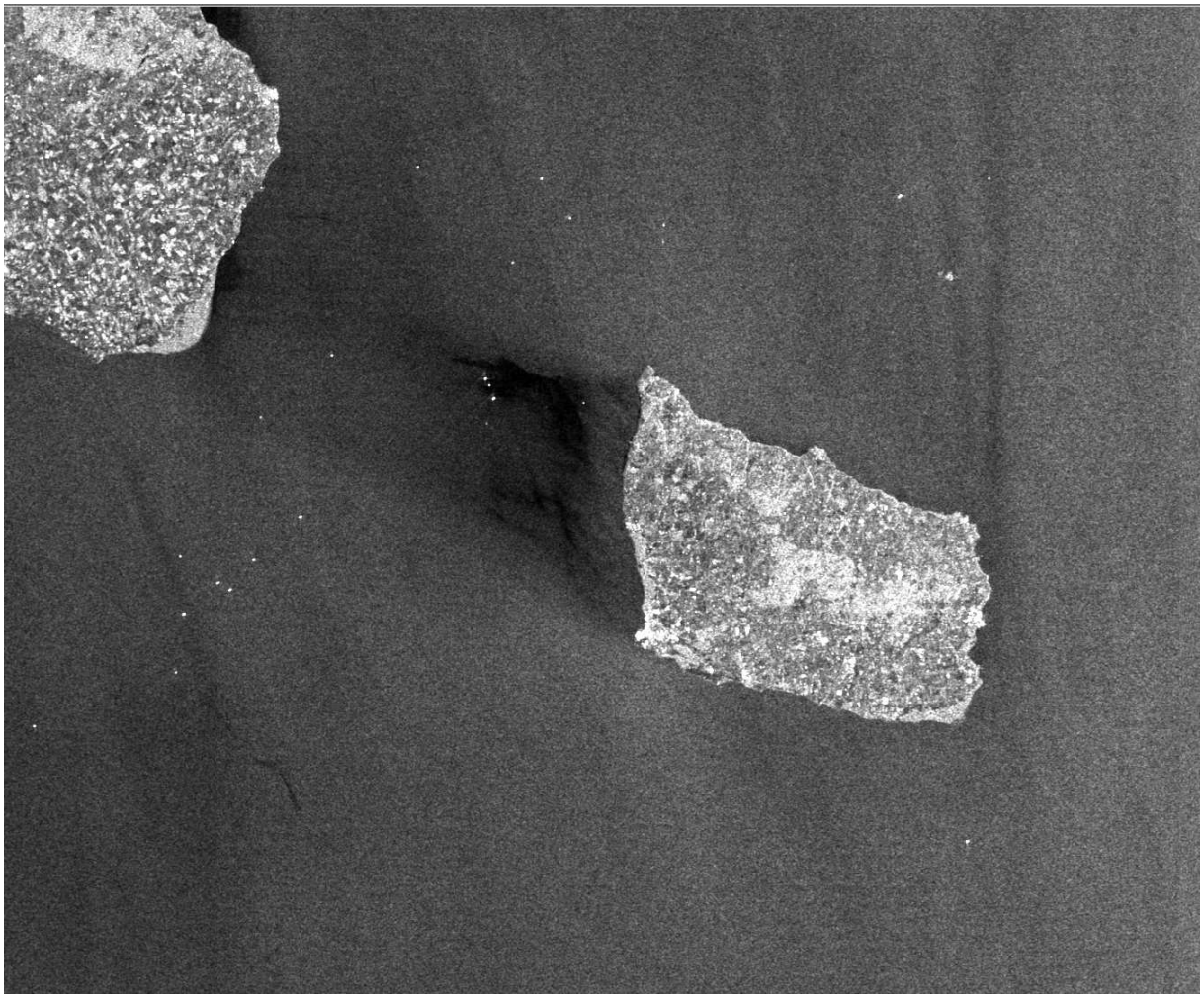


Figure 4.3: Part of an ENVISAT ASAR WSM image from the 25th of September 2005. Example of wind shadowing by the island of Bornholm in the Baltic Sea. Original data ©ESA/KSAT.

- **Rain cells:** There are two processes involved when low-backscatter signatures are caused by rain in SAR imagery. First, atmospheric attenuation due to volume scattering will tend to decrease the backscattering toward the SAR over an area under a rain system. Second, depending on the wind speed and Bragg wave scale, the raindrop impact on the sea surface may tend to dampen the Bragg waves. C-band is affected more by rain volume scattering, while L-band is more sensitive to Bragg wave dampening by the rain.

- **Shear zones:** Shear zones appear as narrow, bright or dark curving signatures in SAR images. Shear zones occure in areas of strong currents.

- **Internal waves:** Internal gravity waves in the ocean can affect the local sea surface velocities and thus the Bragg wave spectrum. This modulation allows imaging internal waves by SAR. The radar image of internal waves consists of adjacent bright and dark bands. Internal waves can also accumulate surfactants, in which case the internal waves are imaged as parallell dark bands. Internal waves appear in shallow water, and the wavelength is typically several kilometers. Figure 4.1(b) shows an example of what could be internal waves.

These ocean features reflect either meteorological or oceanographic conditions. There are also some low-backscattering phenomena caused by large oil installations or ships. An example is turbulent *ship wakes* that decrease the surface roughness when wind waves are present and they are often observed in SAR images.

Table 4.1 shows weather limitations and damping characteristics of some of the low-backscattering features described in this chapter.

Table 4.1: Weather limitations and damping of some low-backscattering features. (Adapted from [13]).

Phenomenon	**Weather limitations**	**Damping [dB]**
Oil spill	Wind speed <= 15 m/s	0.6 - 13.0
Natural film	Wind speed <= 7 m/s	0.8 - 11.3
Grease ice	Winter season and cold nights close to ice edge.	14.0 - 19.0
Threshold wind speed area	Wind speed <= 3 m/s	9.6 - 18.5
Shear zones	Wind speed <= 10-12 m/s	1.4 - 6.2
Internal waves	Wind speed <= 8 m/s	0.8 - 6.0

4.4 SAR Configuration for Oil Spill Detection Services

VV polarization gives higher radar backscatter from the sea surface than HH polarization because of the large dielectric constant of the ocean surface [12]. VV polarized SAR should therefore provide more contrast (better signal-to-noise ratio) when oil is floating on the sea surface, and is therefore the preferred polarization for detecting oil pollution at sea [1]. To illustrate the difference between VV and HH, figure 4.4 shows a possible oil spill imaged by the ENVISAT ASAR Alternating Polarization (AP) mode (co-pol), and the horizontal profiles of both bands (VV and HH). VH and HV reflections occur from more complex reflection mechanisms, and are not particular suitable for oil spill detection because the reflections from the sea surface will be below the noise floor of the SAR in moderate to strong wind conditions (the ocean backscatter should be at least 3 dB above the noise floor

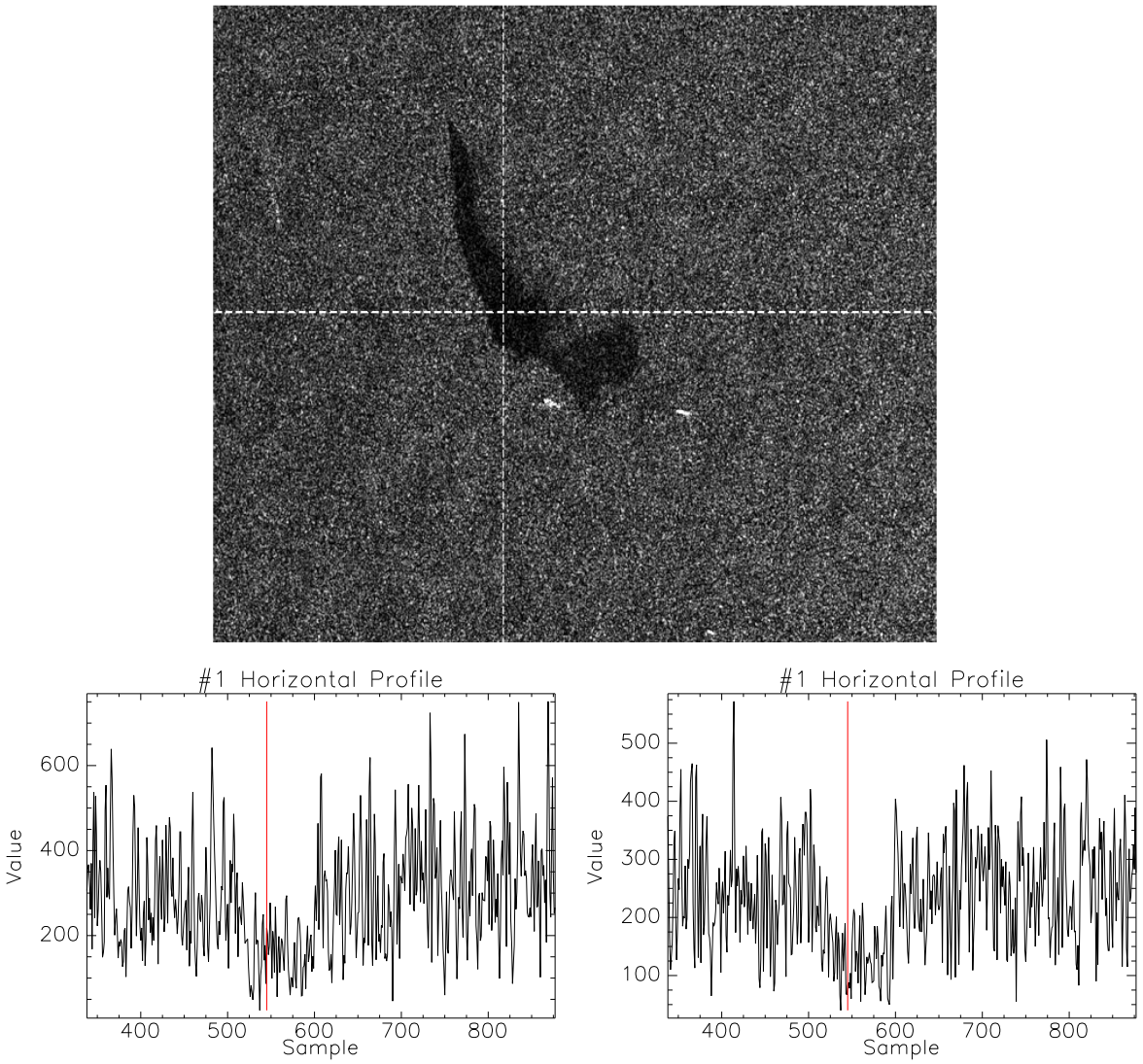


Figure 4.4: Top: Part of an ENVISAT ASAR AP image (VV band), 28th of March 2003. Original data ©ESA. Bottom left: Horizontal profile of the VV band. Bottom right: Horizontal profile of the HH band.

in order to support oil slick detection, and for ENVISAT WSM the signal levels are expected to be above the noise floor for most conditions [35]).

X- and C-bands are more efficient for oil spill detection than L- and P-bands [1] (see definition of X, C, L and P band in Table 2.1). The C-band frequency seems to be the most suitable for oil spill detection, allowing oil spills to be visible in SAR images up to a wind speed of about 10 to 14 m/s (depending on the type of oil and the thickness of the slick).

A bright point target (i.e. possible ship/oilrig) connected to or near by a detected dark spot can increase the confidence that a spot is an oil spill. Therefore, the distance to the closest bright spot is an important feature in oil spill detection. Estimated detectability of ships with respect to a selection of wind speeds and incidence angles for ENVISAT WSM was presented in Arnesen and Olsen [2], showing that HH polarization gives somewhat smaller (better) numbers than VV polarization. This is due to the fact that the backscatter from the sea for HH is lower than for VV, giving a slightly higher contrast to the targets. This indicates that HH polarized images should be preferred compared to VV for ship detection, which is in contradiction to the recommended VV mode for oil spill detection. According to Attema [4], ship detection has largest success using HH polarised data for incidence angles > 45 degrees while cross-polarised data should be used for incidence angles < 45 degrees. VH polarization has shown promise in detecting ships, and the VV/VH mode of RADARSAT-2 might therefore be a suitable acquisition mode in the future [9].

Within the frame of this thesis, only a simple ship detector was applied to detect potential targets in the SAR image data sets (the SAR image sets applied in this work are presented in Chapter 5). Details on the ship detection algorithm can be found in Paper II.

A wider range of sensors for oil spill detection is discussed in Paper I.

Chapter 5

Data Material and Area of Study

The experiments carried out within the framework of this thesis are based on data from the ENVISAT Advanced Synthetic Aperture Radar (ASAR) instrument and the RADARSAT-1 SAR instrument. RADARSAT-1 is a Canadian satellite and it was launched 4th of November 1995, while ENVISAT is an ESA satellite launched 1st of March 2002 (from Kourou in French Guiana).

All satellite images applied in the work with this thesis were processed and delivered by Kongsberg Satellite Services (KSAT) (former Tromsø Satellite Station) in Tromsø, Norway. KSAT has provided a service utilising satellite SAR images for detection of oil spills since 1994. The analysis relies upon human interpretation supported by meteorological and geographical information. The KSAT oil spill service is one of the few operational satellite services having operational customers [40]. Another example is the Canadian Integrated Satellite Tracking of Pollution (ISTOP) program that was operationalized in 2006. This service applies RADARSAT-1 data to monitor marine coastal areas for oil spill pollution [9].

5.1 Selected SAR Modes

Both ENVISAT ASAR and RADARSAT-1 have several image modes (see figure 5.1). In addition, ENVISAT ASAR has the possibility of several polarization options. The ENVISAT ASAR WSM and RADARSAT-1 ScanSAR Narrow (SCN) mode were selected for this study. The VV-polarization option for the ENVISAT WSM mode was chosen, while the RADARSAT-1 SCN images are HH-polarized. The main reason for this choice is that we got an agreement of sharing data acquired by the European Commission (EC) project Oceanides [51] (a large number of WSM and SCN images with aircraft verifications of possible oil spills, see section 5.2). ENVISAT ASAR WSM is aimed primarily at oceanographic applications (like oil spill detection), where there is a special interest in obtaining a wide area view with high temporal frequency. RADARSAT-1 and ENVISAT are separated in orbit by approximately 4 hours, and combined use of these two satellites can improve the temporal

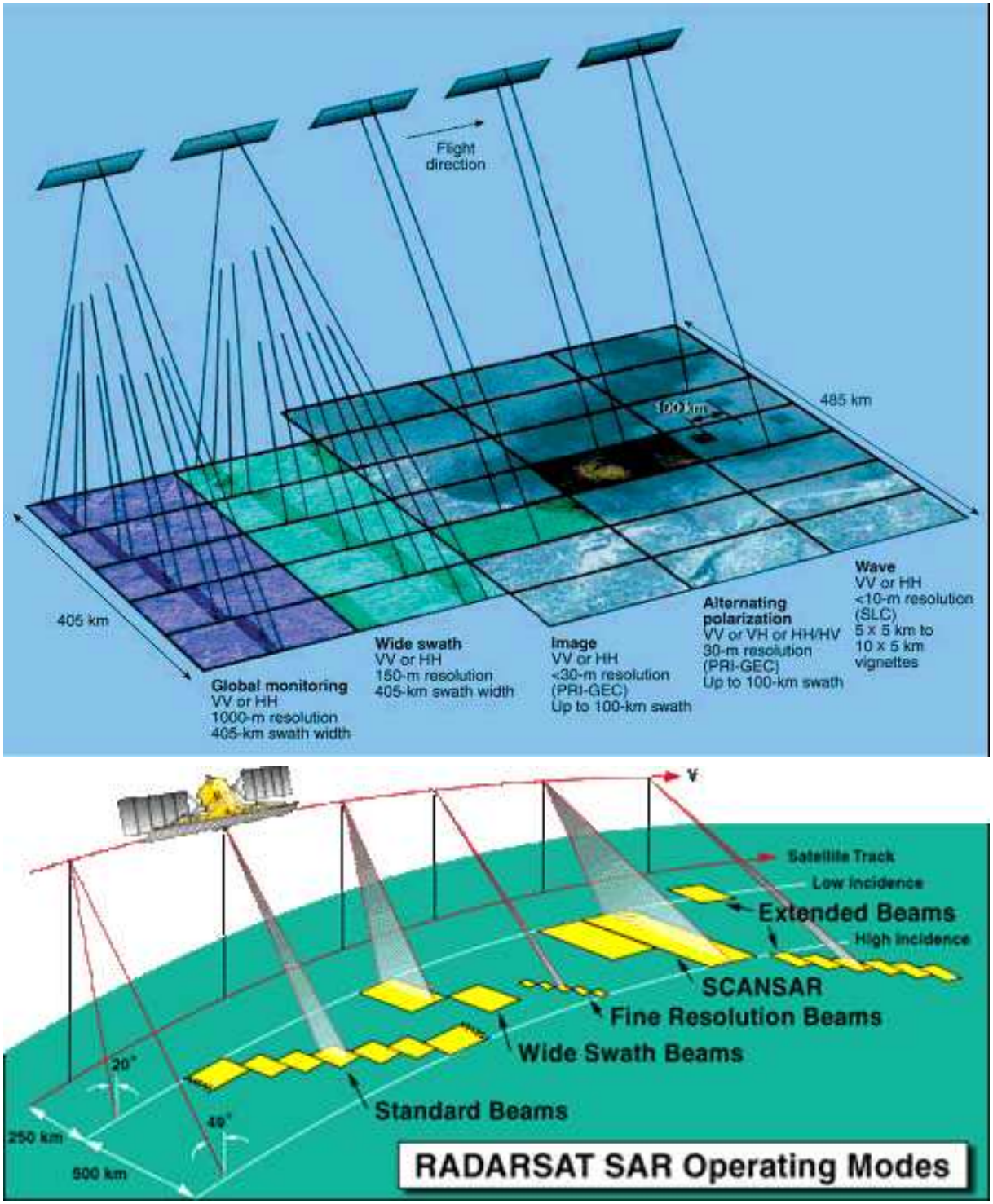


Figure 5.1: Top: the operational mode swaths of ENVISAT ASAR ©ESA. Bottom: the operational mode swaths of RADARSAT-1 ©CCRS.

and spatial coverage [40].

Both ENVISAT ASAR and RADARSAT-1 can operate according to the ScanSAR principle (see section 2.2). ASAR uses five predefined overlapping antenna beams that cover the wide swath [3] (see figure 5.2).

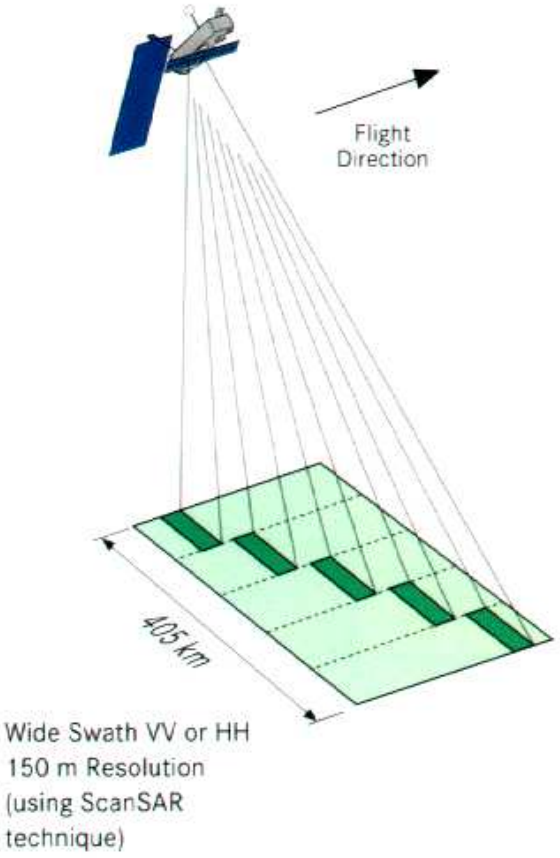


Figure 5.2: ENVISAT ASAR WSM. ©ESA.

5.2 Area of Study and Ground Truth

As part of the Oceanides project, a joint satellite-airborne campaign was performed during 2003 to establish a data set consisting of SAR images with associated aircraft verifications by the German and Finnish pollution control authorities. The campaign covered the Finnish and German sectors of the Baltic sea, in addition to the German sector of the North Sea. A total of 59 RADARSAT-1 and ENVISAT images were acquired between July and December 2003. This campaign was organized in such a manner that KSAT downloaded the satellite images,

analysed them in near-real time and reported possible oil spills to the Finnish and German pollution control authorities, which could check the reported dark spots and determine if they were oil or look-alike. Not all of the possible oil spills in these images were checked due to e.g. flight cancellations or slicks occurring outside the Finnish and German maritime territories. Figure 5.3 shows the three test sites from the joint aircraft and satellite campaign. The images from this campaign have been used for performance testing.

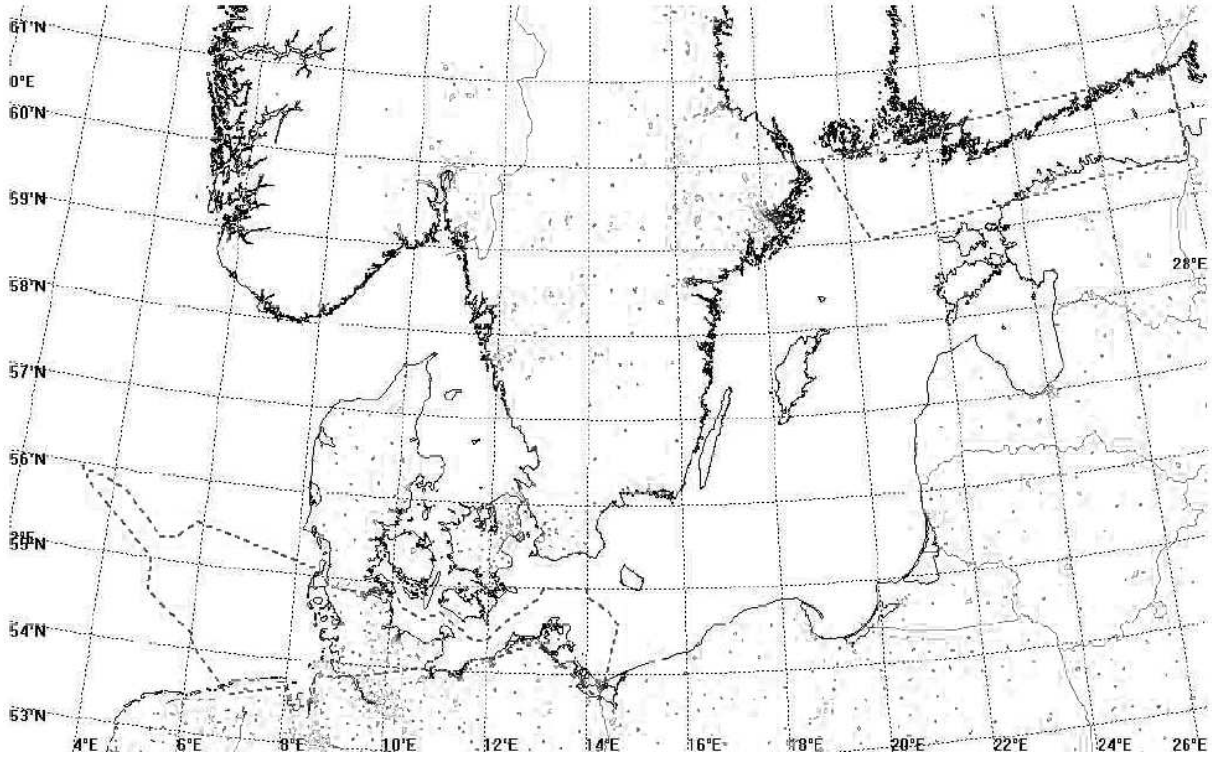


Figure 5.3: Test sites for the 2003 Oceanides campaign. The stippled lines indicates the maritime territories where the SAR images were acquired. Source: [38].

In addition to the images from the joint satellite-aircraft campaign, 71 RADARSAT-1 and 76 ENVISAT images were applied in this work. These images were used for training and parameter estimation. For ENVISAT, 56 of the training images were collected during March to December 2003 and January to April 2004, while 20 of the training images were collected later on between January and October 2005. See Table 5.1 for an overview of the data sets.

Oil spill service reports (including e.g. geographical location, extent, wind information, and confidence estimates) from operators at KSATs manual oil spill detection service were available to us for many of the SAR scenes in the data set. When available for the training images, these were used as support during analysis.

Table 5.1: The number of images from RADARSAT-1 and ENVISAT used in this thesis.

Satellite Image Mode	Training Set	Test/Benchmark Set	Total
RADARSAT-1 SCN	71	32	103
ENVISAT ASAR WSM	56+20	27	103

5.3 Pre-processing

Some pre-processing of the SAR images were done, consisting of:

- **Geo-referencing.**
- **Land masking.** Since the radar backscatter over land areas is of less interest for oil spill detection, a land mask [45] was created in the original geometry of the SAR image.
- **Range (incidence angle) dependent scaling and conversion to logarithmic scale.** This was done due to the fact that the radar backscatter from the ocean depends on the incidence angle, as discussed in section 3.3. A large variation in the backscatter intensity makes both visual inspection and automatic analysis of the large scenes more difficult.

The pre-processing of the data sets was performed by Per Ove Husøy (former employee at Norwegian Computing Center), and some details can be found in [20].

5.4 Image Masking

Collecting SAR image data containing both examples of oil spills and look-alikes was crucial to be able to succeed with the research objectives of this thesis. Finding an oil spill means detecting a rare but important event among a large number of other dark structures observed in SAR images. To maximize the number of positive examples, most of the SAR images applied in this work contain possible oil spills. Most look-alikes appearing in the images are natural phenomena and independent of the oil spill occurrences. Therefore we expect the system to perform well also on new unknown images mostly containing no oil.

The process of using data to determine the parameters of the classifier is referred to as *training*, and the most effective methods for developing classifiers involve learning from example patterns [11]. If a large amount of example patterns are available for training, it is more likely that the algorithm will perform well on new data samples. When designing a pattern recognition system, an important part is also to evaluate how the classifier performs on new samples. A large amount of data in the test set will give a good estimate of the real error rate. In the work with this thesis, the test set was used for performance testing.

A considerable amount of working hours was put into analysing and preparing the images by marking various phenomena visible on the ocean surface. To label the training set, we created a training mask for each SAR image. By manual inspection, we systematically scanned through all images tagging off dark spots as either *oil spill* or *doubt*. The doubt category was only used for a limited number of cases, as in some cases a trained operator cannot discriminate perfectly between oil slicks and look-alikes based on a SAR image alone. Doubt cases were left out of the training process. In principle, all suspicious dark spots likely to be oil pollution were tagged off during analysis. All slicks not marked as oil or doubt were used to train the *look-alike* class. Input from manual inspection by KSAT operators was used when available.

A similar procedure was performed for the test set. All slicks we knew were oil pollution from the aircraft verifications, and all additional slicks we were convinced by manual inspection to be oil spills and therefore should be detected by the algorithm, were tagged off in the mask files of the test set.

The ideal situation would have been to have "ground truth" for all dark spots appearing in the SAR images. This is however not realistic with the technology of today. The time delay (ca. 1-3 hours) between the satellite overpass and the aircraft surveillance made it difficult in some cases to verify the slicks. Another difficulty arose when a slick appeared outside the maritime territory of the pollution control authorities.

Chapter 6

Automatic Analysis

When constructing a system for automatic oil spill detection in SAR imagery there is a number of issues to consider [24]. One important issue is that construction and selection of appropriate features (region descriptors) are required. Another issue is that look-alikes occur much more frequent than oil spills, and the training data set is likely to be imbalanced. Furthermore, when developing a fully automatic oil spill detection algorithm, one of the first problems that arise is how to perform the *segmentation*, i.e. to distinguish the dark spots occurring on the sea surface from the background.

A very brief bibliography of other automatic oil spill detection approaches is given below.

6.1 Detection of Dark Spots

Oil spills appear as dark spots compared to a brighter background in SAR images. The images we are dealing with are gray-scale intensity images (i.e. represented by pixel-wise intensity values) and a popular approach for dark spot segmentation is threshold techniques [22; 23; 34; 39]. Thresholding is applied to each individual image pixel or group of pixels on a local (or global) level. The pixel value is compared to the mean value of all pixels in the local window lowered by a threshold value.

As oil spills damps the gravity-capillary waves, applying wavelets for detecting local variations of the wave spectra has been suggested [28; 29].

Other proposed segmentation methods are based on fuzzy clustering (to allow dealing with mixed surface-cover classes and unsharp boundaries among regions) [5], and mathematical morphology [18].

Segmentation is covered in Paper II and III.

6.2 Feature Extraction

Discrimination between oil spills and look-alikes is often based on a number of features computed for each suspicious dark spot on the sea surface [10; 15; 23; 33; 39; 41]. Good, discriminative features are very important for the success of the following classification step. The selection of features are often based on expert domain knowledge. Most of the features applied in the literature are covered by the following types:

- The geometry of the dark spot.
- Physical characteristics of the backscatter level of the dark spot and its surroundings.
- Dark spot contextual features.
- Texture features of the dark spot and the surroundings.

The feature vector is usually fed into a classifier.

The topic of feature extraction is covered in Paper III.

6.3 Classification Methods

In some of the published papers covering oil spill detection, classification techniques are applied where the decision boundaries are constructed by optimising an error criterion, like Neural-networks (NN) [10; 42] and Support Vector Machines (SVM) [29; 30]. In multi-layer perceptron NNs the separating hyperplane is iteratively updated as a function of the distance of the misclassified patterns from the hyperplane. SVMs handles non-separable training data, and maps the input vector to a high dimensional feature space (through a kernel function) to be able to find a linear separable hyper plane in that space. NNs and SVMs do not apply any a priori assumptions concerning the particular probability distribution of the data set.

Another category of classifiers applied is based on a probabilistic approach [15; 34; 39]. In statistical decision approaches, the decision boundaries are determined by the probability distributions of the patterns belonging to each class, which is normally estimated from training data [21].

Keramitsoglou et al. [23] apply a classifier based on fuzzy logic defined by human experts on the set of features extracted from the SAR images.

All detection algorithms suffer from false alarms because the SAR sensor cannot in all situations separate between a look-alike (e.g. natural slick) and an oil spill. SAR imaging of oil spills can be limited by atmospheric and oceanographic conditions and external data should be taken into account [17].

Classification methodology is the research topic of Paper V and VI.

Chapter 7

Summary of Publications

A summary of the publications included in this thesis follows.

7.1 Paper I: Oil Spill Detection by Satellite Remote Sensing

C. Brekke and A. H. S. Solberg
Published in: *Remote Sensing of Environment*
March 2005, vol. 95, no. 1, pp. 1-13

Developing a machine vision system for oil spill detection requires knowledge about how to analyse satellite images and which sensor and mode to select. It is also of importance to look into what has been done by others in the field; what the new trends are and the technology that exists or is about to emerge. This study presents results of a literature review. The main objective of this paper is to give an extensive overview of the present methodologies and technology applied in the field of oil spill remote sensing. The findings in this study were also used to identify research areas that later were looked into during the work with this thesis.

SAR is the most applicable space-borne sensor for operational oil spill monitoring of large ocean areas. However, SAR imagery has some limitations when it comes to classifying the type of pollution, estimating the thickness of the spill and in some cases to distinguish oil spills from other natural phenomena (*look-alikes*). There are other remote sensing devices that have some potential as a supplement. Optical sensors can be used to some extent to distinguish between oil spills and areas with high level of chlorophyll due to algal bloom (both giving a dark signature in SAR images). A combination of infrared/ultra violet (IR/UV) sensors can be used for oil spill detection and to estimate the thickness of the spill. Microwave radiometer (MWR) can be a useful tool for measuring the thickness and estimating the volume of the oil spill, and laser-fluoro-sensor (LFS) can be used for oil type classification.

Automatic or semi-automatic oil spill algorithms typically consist of three main steps: dark spot detection, dark spot feature extraction and dark spot classification. The value of an extensive comparison of classifiers based on the same data set of features is identified in this paper. This is addressed further in Paper V, where, based on the same data set two variations of a statistical classifier are compared to Support Vector Machines (SVM).

Research has shown that manual operators show some variance in detecting oil spills, particularly in assigning an oil spill confidence estimate (more results from this research are presented in Paper II and Paper IV). The value of developing an automatic algorithm with a reliable and objective oil spill confidence estimator is also identified in this paper. Research on how this could be done and an attempt to develop an automatic confidence estimator were therefore conducted, and the result is presented in Paper V.

Presently, aircraft equipped with additional sensors like side-looking airborne radar (SLAR), IR/UV, MWR and LFS are needed for collecting evidence (to verify the oil spill and to identify the source) to prosecute the polluters. In this paper, we recommended that a future oil spill system should be an integrated system (applying information from several sensors and databases), including automatic SAR algorithms followed by a manual step checking the alarms before notifying the aircraft, a database of "hotspots" (e.g. oilrigs, sunken ships and seepages), ship lane information, algae information, and more extensive use of wind information acquired from e.g. meteorological prognoses.

7.2 Paper II: Oil Spill Detection in Radarsat and Envisat SAR Images

A. H. S. Solberg, C. Brekke and P. O. Husøy
Published in: *IEEE Transactions on Geoscience and Remote Sensing*
March 2007, vol. 45, no. 3, pp. 746-755

In this paper, an early version of our oil spill detection and classification algorithm is described. This version was based on the Norwegian Computing Center (NR) algorithm developed for ERS satellite images by Solberg et al. [39]. Here, the algorithm has been adapted to analyse RADARSAT-1 ScanSAR and ENVISAT ASAR WSM scenes. This version of the algorithm is used as a reference for developments and improvements discussed in some of the following papers (Paper III, V and VI).

One of the objectives of the work conducted in the European Commission (EC) project Oceanides, was to compare the performance of different automatic or semi-automatic algorithms and manual approaches for SAR image analysis with respect to oil spill detection. Another objective was to compare the analysis results based on satellite imagery with results from aircraft surveillance. In this study, our algorithm was benchmarked against manual oil spill detection (KSAT's oil spill service) and a semi-automatic approach (QinetiQ's oil spill

algorithm). In addition to presenting the initial version of the oil spill algorithm, the primary contribution of this paper is presenting and discussing the outcome of the Oceanides benchmark with respect to the algorithm.

In the algorithm, a supervised Bayesian classification scheme is applied to obtain the posterior probability for a detected dark spot being an oil spill. Because of a very unbalanced training set with a limited number of oil spill observations, a classifier with common diagonal covariance matrices is applied in this paper. To overcome a high false alarm ratio produced, a large set of rules adjusting the posterior probabilities in the statistical classifier is introduced (the rules are based on feature values from misclassified dark spots in the training data set). However, establishing these rules was a tedious process and the approach was discarded at a later stage when automatic confidence estimation was introduced (this is covered in Paper V and VI).

After dark spot detection, sometimes thin, piecewise linear slicks got fragmented into several shorter segments or were lost completely in the segmentation process. The need for an improved dark spot detector is identified in this paper (this is covered in Paper III).

The benchmark experiments were performed on both RADARSAT-1 and ENVISAT images. The performance of our automatic system was quite comparable to manual inspection, both in terms of performance in detecting verified oil spills, and in terms of the number of false alarms. However, aircraft are still needed for oil spill verification and to identify the polluter. The dark spots reported by satellite but verified as not oil by the aircraft were studied. In the SAR scenes collected in the Baltic Sea, some cases of verified algae occurrences caused false alarms. Additional information about algal blooms, or the likelihood of observing algae in a given geographical location should be incorporated into the system to eliminate this problem.

When comparing the inter-operator variance between trained human experts with respect to the number of detections and the capabilities of detecting verified oil spills, quite a large difference was found. The automatic oil spill detection and classification algorithm was found to be a good (and objective) alternative to manual inspection when large ocean areas are to be inspected. Running the automatic algorithm as part of the oil spill detection service and including a manual step prior to sending out the aircraft could probably improve the existing operational fully manual service run in Norway.

Some more results from the benchmark study concerning manual confidence estimation are discussed in Paper IV.

7.3 Paper III: Segmentation and Feature Extraction for Oil Spill Detection in ENVISAT ASAR Images

C. Brekke and A. H. S. Solberg
Submitted to: *International Journal of Remote Sensing, 2006*

Many of the papers in the literature (e.g. [10; 23]) describing oil spill detection algorithms list the features (i.e. region descriptors) applied to describe the dark spots, but give very few arguments on why a particular feature or a combination of features are chosen. This paper covers the topics of dark spot detection and feature extraction. The novelties of this paper include a new segmentation algorithm for better detection of thin, piecewise linear slicks and three new features. The importance of an adapted segmentation process to this particular problem domain is demonstrated. Combined with an improved feature vector, an increase in the classification accuracy for the oil spill class is shown.

Preliminary results and the initial ideas for this paper are published in two short papers, Paper VII and Paper VIII.

First, building upon the general thresholding approach presented in Paper II, an improved thresholding algorithm for better detection of thin, piecewise linear slicks is described, where small fragments are merged to continuous segments. Based on the improved segmentation, the true shape of these slicks can better be reflected in shape features. By including the new segmentation approach into the algorithm the classification accuracy for oil spills increased from 78% to 86%, while the classification rate for look-alikes was unchanged at 99%.

Second, three new region descriptors are proposed including a feature measuring the curvature of the slick border, an improved slick border detector and a feature measuring the variance of the slick pixel values. By introducing the new region descriptors the classification accuracy for oil spills increased again from 86% to 89%, but at the same time the classification ratio for look-alikes was reduced from 99% to 98%. However, we consider misclassifying oil spills as look-alikes more serious than misclassifying look-alikes as oil spills, but in an operational context the false alarm ratio should be reduced as much as possible. In this paper, the focus is on improving the detection of oil spills through better segmentation and a better set of features. Paper V focuses on classification and the trade off between a high detection ratio for oil spills and a low false alarm ratio.

It is important to point out that the work presented in this paper was performed more or less in parallel with the benchmark study presented in Paper II. Therefore, not all the aircraft verifications were available at the time the test data set was labelled *oil spill* or *look-alike*. When all the aircraft verifications were available, we re-analysed the whole test data set to include all the aircraft verifications.

7.4 Paper IV: Oil Spill Detection in Northern European Waters: Approaches and Algorithms

A. H. S. Solberg and C. Brekke
To appear as a book chapter in: *Remote Sensing of the European Seas (Vittorio Barale and Martin Gade (Eds.))*

Springer Science and Business Media B.V., 2007

The aim of this paper is to give the reader an overview of the field of oil pollution monitoring with respect to the experience from Northern European seas. The paper was written as an invited contribution to the book *Remote Sensing of the European Seas*. The scope of the book determined that the paper should be written also for non image analysis specialists.

In this paper, the advantages and limitations of combined use of satellite imaging of oil spills and aircraft monitoring are discussed. In the Baltic Sea and the North Sea, many national pollution control authorities already have experience with combined satellite and airborne routine monitoring. Aircraft are used to verify potential spills first spotted by the satellite. The use of satellite surveillance allows the users to better target the aircraft used for oil spill surveillance and to cover larger areas. According to Tufte et al. [43], the best approach for a specific national authority depends on the size and shape of the area and other resources available. International cooperation with neighbouring countries on planning satellite acquisitions is important to reduce costs. Knowledge about oil spill statistics is also important for optimisation of the oil spill surveillance efforts. However, care must be taken in using oil spill statistics in terms of the number of observed spills if they are not normalized in terms of how frequently the areas are monitored. Good agreement is found between aircraft detections and satellite-based detections when the time offset between the acquisitions is low [51]. Therefore, a combined used of satellite-based SAR and surveillance aircraft carrying additional sensors is recommended as a cost-effective way of monitoring.

In Paper II, automatic detection of oil spills was found to be an interesting complement to manual detection, which currently is the primary approach applied by the operational oil spill services (e.g. [9; 40]). A main contribution of this paper is a study of the inter-operator variance between trained human experts with respect to assigned confidence levels. A large inter-operator variance in confidence estimation was found, indicating that there still is some subjectivity involved. Automatically estimated confidence levels can be helpful as an objective source of support when prioritising the alarms (automatic confidence estimation is covered in Paper V).

7.5 Paper V: Classifiers and Confidence Estimation for Oil Spill Detection in ENVISAT ASAR Images

C. Brekke and A. H. S. Solberg
Submitted to: *IEEE Transactions of Geoscience and Remote Sensing, 2007*

According to the results presented in Paper II, automatic algorithms could speed up the analysis process of the SAR images but the false alarm ratio produced by such algorithms

should be reduced. According to the findings in Paper IV, it would also be desirable that the automatic system not only reports which slicks are classified as oil spills, but also the confidence associated with a certain slick being classified as oil. The estimated confidence levels of the slicks could then be used as support when directing the surveillance aircraft to the targets.

The aim of this study was 1) to improve the classifier with respect to the number of false alarms, 2) to reduce the number of parameters to be tuned manually during training and 3) to develop an automatic confidence estimator to give the operator a tool for prioritising the alarms.

Preliminary results and the initial ideas for this paper are published in a short paper, Paper IX.

This study is based on the ENVISAT ASAR WSM image set. In this paper, the statistical classifier with diagonal common covariance matrices (Paper II) is compared to a statistical classifier with regularized covariance matrices and a C-Support Vector Classifier (C-SVC). A main contribution of this study is that we show that it is possible to largely reduce the false alarm ratio of the statistical classifier by introducing regularization of the covariance matrix estimates (due to the small number of oil spills observed). The regularized statistical classifier was also compared to the C-SVC, showing worse performance on the problem. More experiments with the regularized statistical classifier (the classifier found to perform best on this particular application) were done later, and the results are presented in Paper VI.

In this paper, it is also shown that it is feasible to develop and train an algorithm for automatic assignment of confidence levels with satisfactorily performance. After applying the regularized classifier, each dark spot with a higher posterior probability of being an oil spill than a look-alike is automatically assigned a confidence level. The automatic confidence estimator is based on conditions on the feature values extracted from the dark spots. The conditions are tuned on oil spill observations from the training set.

The two-step classification approach suggested in this paper (classification followed by confidence estimation) is easier to train and the number of parameters to be adjusted is largely reduced compared to the early version of the algorithm (presented in Paper II). In the early version of the algorithms, a large rule-base was applied to reduce the false alarm ratio involving a large number of parameters to be tuned. This rule-base now becomes redundant.

7.6 Paper VI: Classifying Oil Spills and Look-alikes in ENVISAT ASAR Images

C. Brekke, A. Solberg and G. Storvik
To appear in the proceedings of: *ENVISAT symposium*
Montreux, Switzerland, 23-27 April, 2007

In the statistical classifier, we model with Gaussian densities based on the 8-dimensional feature vector that was selected in Paper III. In Paper II, III and V, the feature space was divided into 10 subclasses based on two features: an estimate of the roughness of the sea surrounding the dark spots and a shape descriptor of the dark spots. By splitting the problem and assuming different densities for each subclass, it was expected to be appropriate to assume the densities to be Gaussian within each subclass. The subclasses were selected from experiments on the training set based on a manual approach. A classifier was applied within each subclass, before finally estimating the confidence levels of the dark spots classified as oil spills.

In this study, the optimal number of subclasses and the optimal boundaries between the subclasses were searched for. Optimal boundaries were estimated for different numbers of subclasses. The search for the boundaries between the subclasses was based on an optimiser minimizing the sum of the negative log-likelihood for each subclass. In this paper, the classification accuracies on the ENVISAT ASAR WSM test set are compared after applying the various subclass configurations.

The main contribution of this paper is the estimate of the optimal number and location of each subclass. The results show that when there is enough data in the training set, both a division into 10 subclasses and *no* subclasses could be applied in combination with a regularized classifier.

Chapter 8

Main Contributions

Figure 8.1 illustrates the information chain in a combined satellite-airborne oil spill service. In this context, the contributions of Paper II, III, V and VI belong to step B) in the figure, while Paper I and IV discusses issues related to B) in the larger context presented in the figure.

The main contributions of this thesis are:

- An extensive overview of the present methodologies and technology applied in the field of oil spill remote sensing is given (Paper I).
- Based on a benchmark study (where automatic, semi-automatic and manual SAR image analysis approaches are compared with "ground truth" from aircraft surveillance), the potential of automatic algorithms for oil spill detection is documented (Paper II and Paper IV).
- An improved adaptive thresholding algorithm for better detection of thin, piecewise linear dark spots is developed (Paper III).
- The classification power of various features are investigated and new features are introduced. Specially a shape feature, a contrast feature and a texture feature are included for better description of the dark spots (Paper III).
- With respect to reducing the number of false alarms while keeping a high detection rate, it is shown that a statistical classifier with regularized covariance matrices outperforms the simpler solution of a statistical classifier with common diagonal covariance matrices and Support Vector Machines (SVM) (Paper V).
- To give the operator a tool for prioritising the alarms, confidence estimation is suggested as a second step of the classification approach following classification with the regularized statistical classifier. A confidence estimator is developed, letting the user tune the system with respect to the trade-off between the number of true positives and false positives (Paper V).

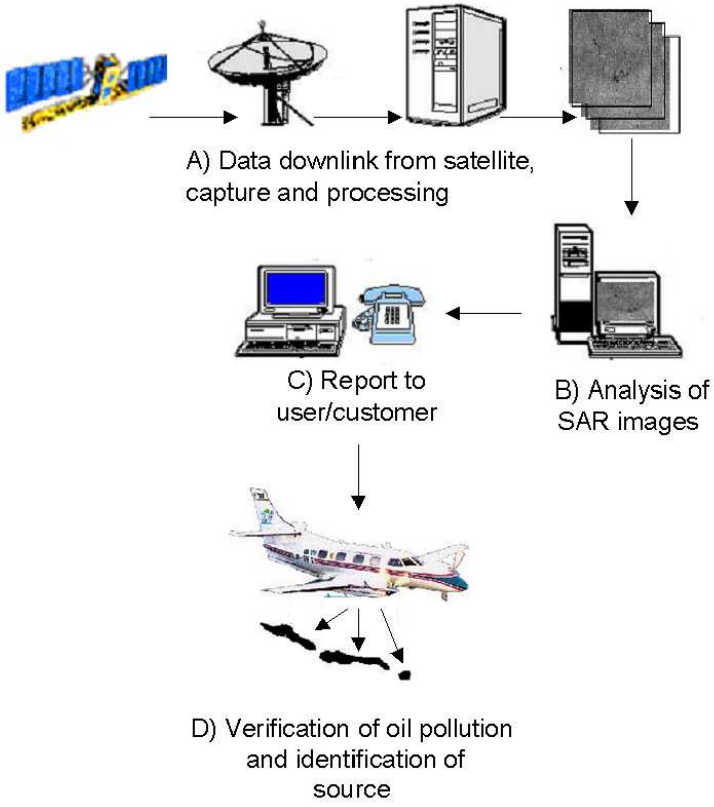


Figure 8.1: The elements and flow of information in an operational oil spill service.

- The behaviour of the features will change under different wind conditions. To handle multimodal conditional densities, experiments with a division of the feature space into subclasses are done. It is shown that, when there is enough data in the training set, a division of the feature space into 10 subclasses and no division at all gives comparable results when the two-step classification approach is applied. Therefore, a division of the feature space into subclasses might not be as important as first assumed (Paper VI).

Chapter 9

Discussion and Directions for Further Research

In this chapter, various aspects of this thesis are discussed. Details that were not satisfactory solved are highlighted and ideas for improvements are suggested. Thoughts about future research possibilities that probably should be looked into and ideas for improvements of the system are discussed as well.

Granularity

Classification can be done on the pixel level, region level (pixels may be grouped into segments) or on image level. The coarsest granularity works with the whole image and it simply states that a given image contains an oil spill. This means that the operator has to manually analyse the scene reported containing one or more oil spills by the system. On a region level, the system reports the coordinates of dark regions that are detected as oil spills. The finest granularity works with pixels, letting the user know if a particular pixel is part of an oil spill. However, if pixels are classified individually there is no guarantee that the *oil spill* pixels will form coherent regions in an image. We chose to let the system detect and classify regions. The output from the automated analysis is oil spill candidates and their geographical coordinates.

Region Descriptors

In Paper III, the feature SUM_EXTERNAL_ANGLES is defined as the sum of all local curvature measurements along the boundary of a dark spot. The implementation is based on ideas presented by Lobregt and Viergever [27]. However, in Paper III the boundary was traced and vertices were inserted with a three-pixel spacing. This might not be an optimal solution, as it will be uncertain if the true vertices of the contour are located. To improve

this, one possibility could be to apply the algorithm for polygonal approximation proposed by Wall and Danielsson [44], where the algorithm outputs a new line segment when the area deviation (the difference between the areas above and below the current segment enclosed by the curve) divided by the length of the current segment exceeds a pre-specified value.

Hu's seven combinations of the second and third order normalized central moments [19] are rotation, scale and position invariant and can be used for invariant pattern recognition. In Paper III, Hu's 1st invariant planar moment is found to separate well between dark spots with thin, piecewise elongated shapes and spots with other arbitrary shapes. Adding more moment features (e.g. higher order moments) can perhaps improve the description of spots with arbitrary shapes. However, if we could describe the shape of the spots perfectly, there would be no guarantee that this would improve the classification performance. The reason is that it is very hard to tell the difference between oil spills and look-alikes only based on the shape. Due to the problem of *the curse of dimensionality* in machine vision, the number of features should also be kept low relative to the number of training samples [21]. Therefore, it is likely to be more appropriate to base the classification on a combination of feature types, like features describing the slick surroundings, texture and contrast features in addition to shape features.

In Paper II, the Sobel operator was used in the computation of the mean of the magnitude of the border gradients of the dark spot border area. The Sobel operator can be used as a simple detector of horizontality and verticality of edges when the following two 3×3 kernels are convolved with the original image:

$$h_1 = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}, \; h_2 = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}$$

At each point in the image, the resulting gradient approximations can be combined to give the gradient magnitude (if the h_1 response is x and h_2 response is y, $\sqrt{x^2 + y^2}$). The problem is that we get inaccurate results for thin, linear slicks, particularly lines that are 1 pixel wide, which is sometimes the case for oil spills imaged by the WSM mode of ENVISAT. In Paper III, a combination of the Sobel operator and a line detector (consisting of four additional masks) was suggested. This is an ad hoc solution, and a better approach might be to sum the square of the responses from the following modified Sobel masks (replacing h_1):

$$h_3 = \begin{bmatrix} 0 & -1 & 1 \\ 0 & -2 & 2 \\ 0 & -1 & 1 \end{bmatrix}, \; h_4 = \begin{bmatrix} -1 & 1 & 0 \\ -2 & 2 & 0 \\ -1 & 1 & 0 \end{bmatrix}$$

and (replacing h_2):

$$h_5 = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 2 & 1 \\ -1 & -2 & -1 \end{bmatrix}, \; h_6 = \begin{bmatrix} 1 & 2 & 1 \\ -1 & -2 & -1 \\ 0 & 0 & 0 \end{bmatrix}$$

This gives an unchanged estimate for dark lines that are two or three pixels wide (compared to applying h_1 and h_2), but a single top for lines that are one pixel wide.

Estimates of Wind Speed

Information about wind speed and wind direction derived from models by meteorological institutions could be incorporated into the system, as it could probably improve the confidence estimation of a detected dark spot being an oil spill. Wind prognoses are available from e.g. Meteorologisk institutt (met.no) with a resolution of 10 km (HIRLAM). Another possibility is to used automatic methods for wind estimation, where the wind speed is derived directly from the SAR image. SAR wind retrieval is discussed in Christiansen's thesis [7]. According to Christiansen, generally, the standard deviation of SAR wind speed retrievals is below ± 2 ms^{-1}. See also Salvatori et al. [37] who applied an inverted CMOD4 model in their oil spill detection system.

Reducing False Alarms

Many of the SAR scenes in the data set are from the same geographical area. During analysis, we discovered that some look-alikes appeared at the exact same location in several images. As suggested in Paper I, to build a database of "hotspots" for look-alikes and include this into the algorithm could reduce the number of false alarms.

Analysis of the characteristics of the verified false alarms in Solberg et al. [38] did not reveal any particular characteristics. It might be the case that false alarms do not have any unique characteristics appearing in the SAR images to distinguish them from oil spills. The data sets applied in this thesis were partly collected from the Baltic Sea. During the summer months, algae blooming in the Baltic Sea could lead to many false detections. To include algae information from external sources could also help reduce the number of false alarms. Web-services covering information about algal blooms exists for some areas. An example is the algae service provided by Danmarks Meteorologiske Institut (DMI). Figure 9.1 shows the average of ENVISAT MERIS data collected during a week. By averaging the data products, the problems with cloud cover are reduced. This information could be used as prior knowledge and incorporated into the classifier.

The Spaceborne Imaging Radar- C/X-Band Synthetic Aperture Radar (SIR-C/X-SAR) [1] sensor was designed and built to make eight different measurements at the same time: L-band and C-band with four different polarization combinations, including HH, HV, VH and VV. Experiments done on the C-band polarimetric SAR [16; 31; 32] show that polarimetric SAR data can improve/assist oil spill detection. It was found that it is possible to distinguish

[1] SIR-C/X-SAR operated on board the Endeavour Shuttle on mission STS-59 and STS-68 in 1994. It was the first mission where a multi-frequency, multi-polarization imaging radar system was flown in space.

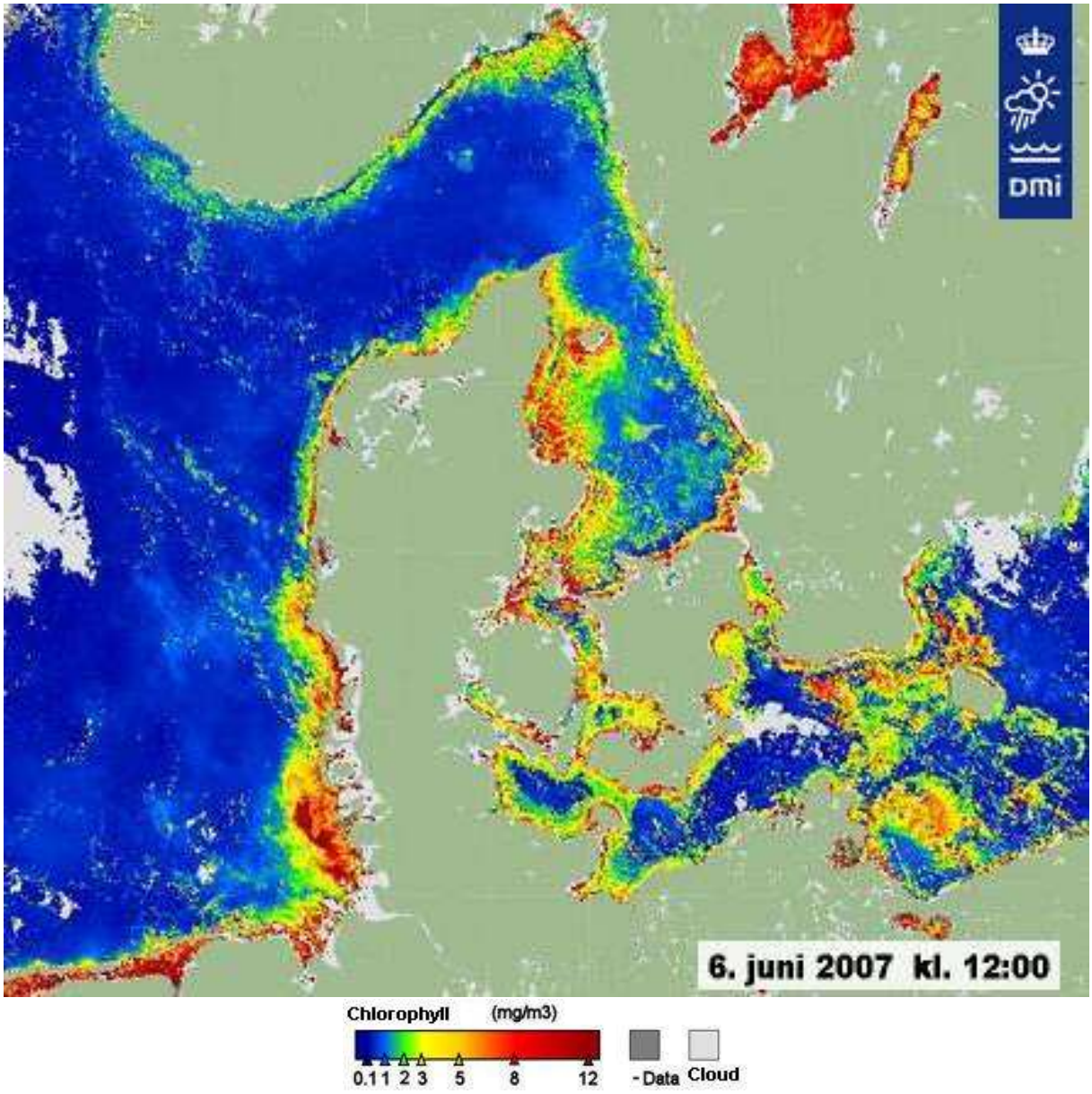


Figure 9.1: Algae consentration. The image is an average of (4-5) ENVISAT MERIS images collected over one week.
Available at: http://www.dmi.dk/dmi/index/danmark/algekort.htm

oil-free from oil-covered areas and in some cases to distinguish among oil spills and biogenic look-alikes. The technique only needs HH and VV data [16]. According to Migliaccio et al. [32], there is now general consensus that radar polarimetry is able to provide additional information to environmental remote sensing.

Noise Reduction

Noise filters for oil spill segmentation in SEASAT images were studied in a short paper by Barni et al. [5]. A Sigma filter (7×7 kernel size) was found to give the best segmentation result and was the fastest among the filters compared. It would be natural to also evaluate the effect of different noise removal filters, applying different window sizes, on the segmentation of slicks with various shapes and contrasts in ENVISAT ASAR WSM images. The ENVISAT images in the present study were pre-processed with a mean filter, in addition a Sigma filter was applied in the segmentation approach.

Parameter Tuning

In Paper VI, we formalized the selection of subclass boundaries based on optimising the negative of the *log*-likelihood function. The number of subclasses was selected based on the classification performance of different subclass configurations. In Paper V, an attempt to automatically set the confidence condition limits was made, however some manual adjustments were still found necessary. In the algorithm, there are still quite a few variables that are manually tuned based on a trial-and-error approach. Ideally, the algorithm should be able to adapt rapidly to changes in the training data set. The possibility of automating a larger part of the parameter tuning should be looked into, even if the complex nature of the SAR scenes could make this difficult.

Selection of Classifier

According to Jain et al. [21], the four best known approaches for pattern recognition are: 1) template matching, 2) statistical classification, 3) syntactic or structural matching, and 4) neural networks. In this work, mainly statistical classification and SVM have been explored. Neural networks is a popular classification technique, and has been applied by others for oil spill detection based on SAR images [10; 41]. It would be interesting to compare the classification performance of our regularized statistical classifier to neural networks based on the same feature data set.

Here, a supervised, density-based approach for classification is applied where the parameters for Gaussian densities are estimated. In Paper VI, it is shown that some of the features appear to have a skewed distribution. The application of other distributions should be explored.

Context Sensitive Learning

All our training images are collected from the same geographical area (the Baltic Sea and the North Sea). An open research issue is to look into how well the algorithm generalizes to other geographical areas, and how sensitive the performance is to seasonal climate variations.

Adapting the algorithm to detect other ocean features

It would also be interesting to look at the possibility to train and adapt the algorithm to detect other ocean features, e.g. fronts, shear zones, internal waves and eddies, which are frequently observed in SAR images.

Coming/Planned Sensors Suitable for Oil Spill Detection

Avoiding a gap with respect to the ERS and ENVISAT missions, SENTINEL-1 is a planned ESA two-satellite system to be operated as a constellation for maximized coverage/repeat cycle. The first satellite will be launched in the 2011-2012 timeframe and the second some 12-15 months later. SENTINEL-1 data should be available for at least 10 years. The SAR system will be C-band [4].

RADARSAT-2 is another SAR mission. RADARSAT-2 is a Canadian satellite planned to be launched in the summer of 2007. The SAR instrument will be C-band like RADARSAT-1, but there will be a flexibility in the selection of polarizations (HH, HV, VH, VV) [52]. The ScanSAR Wide mode with a nominal swath width of 500 km, an approximate resolution of 100 m $\times$ 100 m, approximate incidence angles of 20° - 49° and VV/VH polarization should be suitable for oil spill detection (including potential source detection).

The German X-band satellite TerraSAR-X was launched the summer of 2007 [53]. The oil spill detection capabilities of this sensor should be looked into. The ScanSAR mode has 100 km swath width and 16 m resolution. As discussed in section 2.1.4, with the short wavelengths of X-band, it might be that the SAR imaging will be less independent of atmospheric conditions.

References

[1] Alpers, W. and Espedal, H., 2004, Oils and Surfactants. Chapter 11 in *Synthetic Aperture Radar Marine User's Manual*, National Oceanic and Atmospheric Administration (NOAA). Jackson, C.R. and Apel, J.R.(Eds.), 263-275.

[2] Arnesen, T. N., Olsen, R. B., 2004, VURDERING AV ENVISAT ASAR FOR SKIPSDETEKSJON, *FFI/RAPPORT-2004/02121.*

[3] ASAR Product Handbook, Issue 21, 28 March 2006, *European Space Agency.*

[4] Attema, E., 2005, Mission Requirements Document for the European Radar Observatory Sentinel-1, Requirement Specification, European Space Agency, ES-RS-ESA-SY-0007, issue 1, revision 4.

[5] Barni, M., Betti, M., Mecocci, A., 1995. A Fuzzy Approach to Oil Spill Detection on SAR Images. In *Proc. IGARSS 1995*, 1, 157–159.

[6] Clemente-Colon, P. and Yan, X.-H., 2000, Low-Backscatter Ocean Features in Synthetic Aperture Radar Imagery. *Johns Hopkins Apl Technical Digest,* 21 (1), 116–121.

[7] Christiansen, M. B., 2006, Wind Energy Applications of Synthetic Aperture Radar, Ph.D. thesis, University of Cophenhagen, Denmark.

[8] Curlander, J. C. and McDonough, R. N., 1991, Synthetic Aperture Radar: Systems and Signal Processing. *New York: Wiley.*

[9] De Abreu, R., Gauthier, M.-F. and Van Wychen, W., 2006, SAR-Based Oil Pollution Surveillance in Canada: Operational Implementation and Research Priorities. In *Proc. of OceanSAR 2006 - Third Workshop on Coastal and Marine Applications of SAR, St. John's, NL, Canada, October 2006.*

[10] Del Frate, F, Petrocchi, A., Lichtenegger, J., Calabresi, G., 2000. Neural Networks for Oil Spill Detection using ERS-SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 38 (5), 2282–2287.

[11] Duda, R. O, Hart, P. E., Stork, D. G., 2001. *Pattern Classification*, John Wiley & Sons, Inc.

[12] Elachi, C. and van Zyl, J., 2006. *Intorduction to the Physics and Techniques of Remote Sensing*, John Wiley & Sons, Inc.

[13] Espedal, H. A., 1998, Detection of oil spill and natural film in the marine environment by spaceborne Synthetic Aperture Radar. Ph.D. thesis, Department of Physics University of Bergen and Nansen Environmental and Remote Sensing Center, Norway.

[14] European Space Agency, 1998, Oil pollution monitoring. *ESA brochure: ERS and its Applications - Marine, BR-128, 1.*

[15] Fiscella, B., Giancaspro, A., Nirchio, F., Pavese, P. and Trivero, P. Oil spill detection using marine SAR images. *International Journal of Remote Sensing*, 21(18), 3561–3566, 2000.

[16] Gambardella, A., Nunziata, F. and Migliaccio, M., 2007, Oil spill observation by means of co-polar phase difference, *POLinSAR 2007*, European Space Agency.

[17] Girard-Ardhuin, F., Mercier, G., Collard, F. and Garello, R., Operational Oil-Slick Characterization by SAR Imagery and Synergistic Data. *IEEE Journal of Oceanic Engineering.* 30(3), 487–495, 2005.

[18] Gasull, A., Fabregas, X., J.Jimenez, Marques, F., V.Moreno, Herrero, M., 2002. Oil spills detection in SAR images using mathematical morphology. In *Proc. EUSIPCO'2002, Toulouse, France, September 2002* 1, 25–28.

[19] Hu, M.-K., 1962. Visual pattern recognition by moment invariants. *IRE Trans. Inform. Theory*, 8, 179–187.

[20] Husøy, P. O., Norwegian Computing Center, Short note, *Pre-processing of ASAR and RADARSAT scansar data for oil detection.*

[21] Jain, A. K., Duin, R. P. W., and Mao, J., 2000, Statistical Pattern Recognition: A Review. *IEEE Trans. on Pattern Analysis and Machine Intelligence,* 22(1), 4–37.

[22] Kanaa, T. F. N., Tonye, E., Mercier, G., Onana, V., Ngono, J., Frison, P., Rudant, J. and Garello, R., 2003. Detection of oil slick signatures in SAR images by fusion of hysteresis thresholding responses. In *Proc. IGARSS 2003*, 4, 2750–2752.

[23] Keramitsoglou, I., Cartalis, C. and Kiranoudis, C. T., 2006, Automatic identification of oil spills on satellite images. *Environmental Modelling & Software,* 21, 640–652.

[24] Kubat, M., Holte, R. C. and Matwin, S., 1998, Machine learning for the detection of oil spills in satellite radar images. *Machine Learning*, 30, 195–215.

[25] Lillesand, T. M., Kiefer, R. W. and Chipman, J. W., 2004, Remote Sensing and Image Interpretation, *Wiley*.

[26] Lecomte, P., 1993, CMOD4 Model Description, Tech. Report, ESRIN DPE/OM, Doc. No.: ER-TN-ESA-GP-1120, Issue: 1.2.

[27] Lobregt, S. and Viergever, M. A., 1995, A discrete dynamic contour model. *IEEE Transactions on Medical Imaging,* 14(1), 12–24.

[28] Mercier, G., Derrode, S., Pieczynski, W., Caillec, J.-M. L., Garello, R., 2003. Multiscale oil slick segmentation with Markov Chain Model. In *Proc. IGARSS 2003*, 6, 3501–3503.

[29] Mercier, G. and Girard-Ardhuin, F. Oil slick detection by SAR imagery using Support Vector Machines. In *Oceans 2005 - Europe*, 1, 90–95, 2005.

[30] Mercier, G. and Girard-Ardhuin, F. Partially Supervised Oil-Slick Detection by SAR Imagery Using Kernel Expansion. IEEE Trans. on Geoscience and Remote Sensing, 44(10), 2839-2846, 2006.

[31] Migliaccio, M. and Tranfaglia, M., 2005, A study on the Use of SAR Polarimetric Data to Observe Oil Spills, *Ocean-Europe*, 196–200.

[32] Migliaccio, M., Gambardella, A. and Tranfaglia, M., 2007, SAR Polarimetry to Observe Oil Spills, *IEEE Transactions on Geoscience and Remote Sensing*, 45 (2), 506–511.

[33] Montali, A., Giacinto, G., Migliaccio, M., Gambardella, A., 2006, Supervised Pattern Classification Techniques for Oil Spill Classification in SAR Images: Preliminary Results, In *Proc. SeaSAR 2006, workshop 23-26 January 2006, ESA, ESRIN, Frascati, Italy.*

[34] Nirchio, F., Sorgente, M., Giancaspro, A., Biamino, W., Parisato, E., Ravera, R. and Trivero, P., Automatic detection of oil spills from SAR images. *International Journal of Remote Sensing*, 26(6), 1157–1174, 2005.

[35] Olsen, R. B., Wahl, T. and Engen, G., Expected Performance of the ENVISAT ASAR for Near Real-Time Maritime Applications, *Proc. IGARSS 1999*, 1, 962–964, 1999.

[36] Robinson, I. S., 2004, *Measuring the Oceans from Space: The principles and methods of satellite oceanography*, Springer.

[37] Salvatori, L., Bouchaib, S., Frate, F. D., Lichtenneger, J., Samara, Y., 2003. Estimating the wind vector from radar SAR images when applied to the detection of oil spill pollution. *Fifth International Symposium on GIS and Computer Catography for Costal Zone Management, CoastGIS'03.*

[38] Solberg, A., Indregard, M., Clayton, P., 2004, D2-report on benchmarking oil spill recognition approaches and best practice. Tech. rep., Oceanides project, European Commission, Archive No. 04-10225-A-Doc, Contract No: EVK2-CT-2003-00177.

[39] Solberg, A. H. S., Storvik, G., Solberg, R., Volden, E., 1999, Automatic detection of oil spills in ERS SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 37 (4), 1916–1924.

[40] Steinbakk, L., Pedersen, J. P., Indregard, M., 2005, ESA/EOMD of Envisat Based Marine Surveillance Services SAR Based Oil Spill and Ship Detection Services. In *Proc. of the 2004 Envisat & ERS Symposium, Salzburg, Austria, 6-10 September 2004.*

[41] Stathakis, D., Topouzelis, K. and Karathanassi, V., 2006, Large-scale feature selection using evelved neural networks, In *Proc. SPIE 2006, Stockholm, Sweden*, 6365, 636513, 9 pages.

[42] Topouzelis, K., Karathanassi, V., Pavlakis, P. and Rokos, D., 2006, Dark formation detection using recurrent neural networks and SAR data, In *Proc. SPIE 2006, Stockholm, Sweden*, 6365, 638511, 7 pages.

[43] Tufte, L., Indregard, M., Solberg, A. and Huseby, R. B., 2005, D11-oil spill monitoring guidelines and methodology to combine satellite oil spill recognition with routine airborne surveillance. Technical report, Oceanides project, European Commission, Archive No. 04-10225-A-Doc, Contract No: EVK2-CT-2003-00177.

[44] Wall, K. and Danielsson, P.-E., 1984, A Fast Sequential Method for Polygonal Approximation of Digitized Curves. *Computer Vision, Graphics, and Image Processing*, 28, 220–227.

[45] Wessel, P. and Smith, W. H. F., 1996, A global self-consistent, hierarchical, high-resolution shoreline database. *Journal of Geophysical Research*, 101 (1), 8741–8743.

[46] Weydahl, D. J., 1998, Analysis of satellite SAR images for change detection over land areas, Dr. Scient thesis, Department of Geography, University of Oslo, Norway.

[47] Web-site of the Canadian Space Agency, *http://www.space.gc.ca/asc/eng/default.asp*

[48] Web-site of the European Space Agency, *http://www.esa.int*

[49] Web-site of the International Maritime Organization, *http://www.imo.org*

[50] Web-site of the International Tanker Owners Pollution Federation Limited, *http://www.itopf.com*

[51] Web-site of the Oceanides project, *http://oceanides.jrc.it*

[52] Web-site of RADARSAT-2, *http://www.radarsat2.info*

[53] Web-site of TerraSAR-X, *http://www.terrasar.de/*

Chapter 10

Paper I: Oil Spill Detection by Satellite Remote Sensing

Camilla Brekke and Anne H. S. Solberg

Published in: *Remote Sensing of Environment*,
15 March 2005, vol. 95, no. 1, pp. 1-13.

Oil Spill Detection by Satellite Remote Sensing

Camilla Brekke,
Norwegian Defence Research Establishment,
Postboks 25, 2027 Kjeller, Norway.
Department of Informatics, University of Oslo,
Postboks 1080 Blindern, 0316 Oslo, Norway.
E-mail: Camilla.Brekke@ffi.no.
and
Anne H. S. Solberg,
Department of Informatics, University of Oslo,
Postboks 1080 Blindern, 0316 Oslo, Norway
E-mail: anne@ifi.uio.no

Abstract

This paper presents the state of the art for oil spill detection in the world oceans. We discuss different satellite sensors and oil spill detectability under varying conditions. In particular, we concentrate on the use of manual and automatic approaches to discriminate between oil slicks and look-alikes based on pattern recognition. We conclude with a discussion of suggestions for further research with respect to oil spill detection systems.

Key words: Synthetic Aperture Radar, oil spill, detectability, manual detection, automatic algorithms, dark spot detection, feature extraction, classification

1 Introduction

Oil spills on the sea surface are seen relatively often. Observed oil spills correlate very well with the major shipping routes (e.g. in the Southeast Asian Waters [52; 51], and in the Yellow and East China Sea [40]) and commonly appear in connection with offshore installations (e.g. in the North Sea [18]). Annually, 48% of the oil pollution in the oceans are fuels and 29% are crude oil. Tanker accidents contribute with only 5% of all pollution entering into the sea [22]. After analysing 190 ERS-1[1] SAR images of the Mediterranean Sea, Pavlakis et al. [58] found that "deliberate" oil spills appear with considerably higher frequency than oil spills corresponding to reported ship accidents. According to the European Space Agency [20], 45% of the oil pollution comes from operative discharges from ships. When taking into account how frequent such spillages occur, controlled regular oil spills can be a much greater threat to the marine environment and the ecosystem than larger oil spill accidents like the Prestige tanker (carrying $> 77,000$ ton of fuel oil [57]) accident at Galice, northwest coast of Spain in 2002. The impact of not monitoring oil spills is presently unknown, but the main environmental impact is assumed to be sea-birds mistakenly landing on them and the damage to the costal ecology as spills hit the beach [65]. Simecek-Beatty and Clemente-Colón [68] describes how oiled birds lead to the use of SAR for locating a sunken vessel leaking oil.

Active microwave sensors like Synthetic Aperture Radar (SAR) capture 2-dimensional images. The image brightness is a reflection of the microwave backscattering properties of the surface. SAR deployed on satellites is today an important tool in oil spill monitoring due to its wide area coverage and day and night all-weather capabilities.

Satellite-based oil pollution monitoring capabilities in the Norwegian waters were demonstrated in the early 1990's by using images from the ERS-1 satellite (e.g. Bern et al. [6]; Skøelv and Wahl [69]; Wahl et al. [80]). A demonstrator system based on ERS for the Spanish coast were presented by Martinez and Moreno [55]. Today, RADARSAT-1 and ENVISAT are the two main providers of satellite SAR images for oil spill monitoring.

Access to an increased amount of SAR images means a growing workload on the operators at analysis centres. In addition, recent research shows that even if the operators go through extensive training to learn manual oil spill detection they can detect different slicks and give them different confidence levels [39]. Algorithms for automatic detection that can help in screening the images and prioritising the alarms will be of great benefit. Research on this field has been ongoing for more than a decade, and this paper reviews various methods for satellite-based oil spill detection in the marine environment.

As SAR is just one of many remote sensing sensors available an evaluation of the applicability of other satellite sensors for oil spill monitoring is included as well. Most studies done on airborne remote sensing techniques are excluded. For a review of airborne sensor technology for oil spill observation see Goodman [33]. The detectability of oil spills in SAR images are discussed, in terms of wind conditions, sensor characteristics and ambiguities

[1] ERS-1: the First European Remote Sensing Satellite.

caused by other phenomena than oil spills. Finally, our emphasis is on methodology and algorithms for oil spill detection in spaceborne SAR imagery.

2 Satellite Sensors for Oil Spill Detection

Microwaves are commonly used for ocean pollution monitoring by remote sensing. They are often preferred to optical sensors due to the all-weather and all day capabilities, and examples of SAR-equipped satellites are presented in Table 1.1. Mainly spaceborne instruments

Table 1.1: Some satellites carrying SAR instruments.

Satellite (sensor)	**Operative**	**Owner**	**Character-istics**
SEASAT	1978-off same year	NASA	L-band, HH-pol
ALMAZ-1	1991-1992	Russian Space Agency	S-band, HH-pol
ERS-1	1991-1996	ESA	C-band, VV-pol
ERS-2	1995-operating	ESA	C-band, VV-pol
RADARSAT-1	1995-operating	CSA	C-band, HH-pol
ENVISAT (ASAR)	2002-operating	ESA	C-band, HH and VV, alt. pol. and cross pol. modes

(C-band 4-8 Ghz, λ 3.75 - 7.5 cm, L-band 1-2 Ghz, λ 15 - 30 cm and S-band 2-4 Ghz, λ 7.5-15 cm).

are covered here, but airborne Side-Looking Airborne Radar (SLAR) is another possibility. SLAR is an older but less expensive technology than SAR, but SAR has greater range and resolution [23]. Airborne surveillance is limited by the high costs and is less efficient for wide area surveillance due to its limited coverage. While spaceborne SAR can be used for a first warning, aircrafts are more suitable to be brought into action to identify the polluter, the extent, and the type of spill. An example is the German aerial surveillance, which locates oil discharges by SLAR, infrared/ultra violet (IR/UV) scanning is used to quantify the extent of the film, a microwave radiometer (MWR) is used to quantify the thickness and a laser-fluoro-sensor (LFS) is used for oil type classification [75].

In addition to SAR, there are other spaceborne remote sensing devices that have some potential for oil spill monitoring. Friedman et al. [27] compare a RADARSAT-1 SAR image with a corresponding Sea-viewing Wide Field-of-view Sensor (SeaWiFS, visible sensor) image. SeaWiFS measures high levels of chlorophyll for areas with algal bloom, while the SAR images has low backscatter levels in these regions. It is concluded that multiple data sets can be used to discriminate between for example algal blooms and man made slicks. Indregard et al. [39] points out that additional information (in addition to SAR) about algal bloom is desirable, particularly in the Baltic Sea. This could be taken from optical imagery,

from algae maps or other related information. Figure 1.1 shows two examples of algal bloom imaged by SAR.

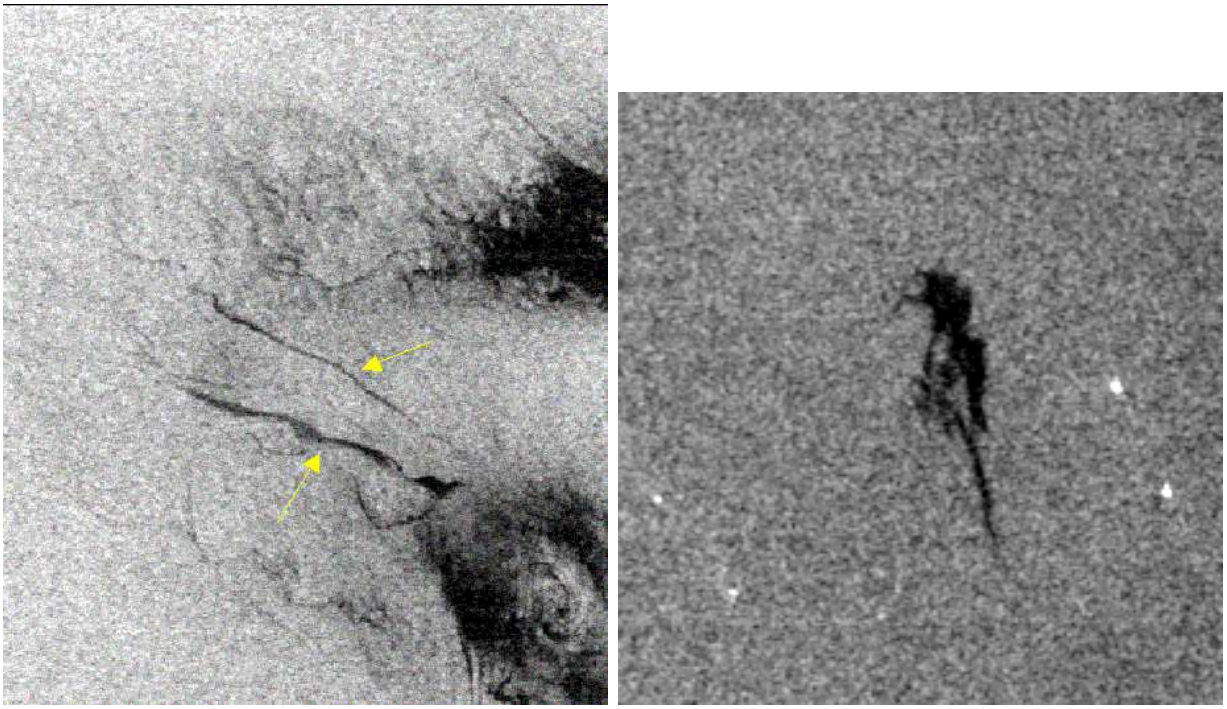


Figure 1.1: Two subscenes of an ENVISAT ASAR Wide Swath Mode (WSM) image from the German Baltic Sea (7th of August 2003). Left: The two slicks, indicated by yellow arrows, was classified as algae by a German surveillance aircraft. Several other phenomena (possible low wind) causing low backscattering are visible in the scene. The size of the subscene is 346×405 pixels. Right: The slick was classified as algae by a German surveillance aircraft. The size of the subscene is 236×241 pixels. ©ESA/KSAT 2004.

A drawback of the SeaWiFS sensor is its coarse spatial resolution of $\sim$ 1km. Hu et al. [37] demonstrates the possibility of oil spill monitoring by the Moderate-Resolution Imaging Spectroradiometer (MODIS) instrument, carried onboard the NASA satellites Terra and Aqua, by an example from Lake Maracaibo, Venezuela. The MODIS instrument has moderate resolution bands of 250 m and 500 m and a wide spectral range. This allows the MODIS instrument to provide images of daylight-reflected solar radiation and day/night time thermal emissions. The MODIS instrument was originally designed for land imaging, and with medium resolution it also shows potential for daily monitoring of the coastal zones looking for oil spills. Cloud cover and the lack of sun light limits the use of optical sensors. On the other hand, multiple wavelengths can give additional information to distinguish slicks produced by algal blooms from oil spills. Yet it is difficult to establish automated feature recogni-

tion systems that identifies oil spills as external knowledge about the region, environment and current events are needed in the visual analysis of the satellite images to discriminate between various events.

Hyperspectral sensors used for oil spill monitoring have a potential for detailed identification of materials and better estimation of their abundance. With more that 200 wavelengths provided by a hyperspectral sensor the spectral signature of oil can be exploited and used to distinguish between different oil types (crude or light oil). This can also eliminate the false alarm rate of ocean features that have the same colour and appearance as oil. Salem and Kafatos [63] found that a signature matching method based on airborne hyperspectral imaging (looking at chemical composition) is more accurate than the conventional techniques, where analysis is based on visual interpretation of the oils colour and its appearance in the satellite image. There is currently no commercial spaceborne hyperspectral sensor in orbit. The NASA EO-1 Hyperion hyperspectral sensor is an example of a spaceborne technology demonstrator that was launched in 2000. However, its major drawback is its small swath width of only 7.5×100 km.

Oil absorbs solar radiation and re-emits a portion of this energy as thermal energy. IR sensors observe thick oil slicks as hot, intermediate thickness of oil as cool, while thin oil is not possible to detect [23]. At night a thick spill can appear cooler than the water since it releases heat quicker than its surrounding water [76]. Tseng and Chiu examined the use and capability of the visible and IR sensors of NOAA Advanced Very High Resolution Radiometer (AVHRR) for early detection and monitoring of oil spills. Oil spills from the 1991 Persian Gulf war were studied. Thick and thin oil layers and the boundary between water and oil were possible to detect by the IR channel, but the oil spills may not have a significant different temperature signature from the surrounding water at night. Oil spills could be detected in the visible images only under highly favourable lighting and sea conditions.

UV technology can be used to detect oil spills as the spill displays high reflectivity of UV radiation even at thin layers. The UV instrument is not usable at night, and wind slicks, sun glints and biogenic material can cause false alarms in the UV data. These interferences are often different from those for IR, and a combination of IR and UV can provide a more reliable indication of oil and can be used for estimating oil thickness [23].

A MWR is another passive sensor. The instrument looks at the microwave radiation in the wavelength cm to mm range that the ocean emits, and therefore is almost weather-independent [75]. Oil slicks emit stronger microwave radiation than the water and appear as bright objects on a darker sea. According to Robinson [62], oil slicks can have strong surface-emissivity signatures, but as a spatial resolution of tens to hundreds of meters is desirable for the determination of oil slicks, this type of sensors for oil spill thickness monitoring is most appropriately pursued by aircraft sensor. Zhifu et al. [83] did some experiments using airborne (AMR-OS) and a ship borne (K-band) MWRs looking at various oil types and thickness. They found that MWRs are useful tools for measuring the thickness and estimating the volume of the spills, but the resolution is not fine enough to give accurate results. Fingas and Brown [23] summarize studies done on this field and conclude that the

potential of radiometers as a reliable device for measuring slick thickness is uncertain.

RADARSAT-1 and Landsat-5/Thematic Mapper (TM) were used to capture images of the Guanabara Bay, Brazil, after an oil spill emergency where a pipeline ruptured [4]. Bentz and de Miranda found that RADARSAT-1 provided suitable temporal coverage, while cloud cover, haze and the 8-day revisit schedule (using both Landsat-5 and -7) prevents Landsat from being used systematically for oil spill monitoring. However, many sea surface temperature and ocean colour sensors have a large swath width and can provide daily coverage of the earth surface (e.g. the revisit time for MODIS is 2 days and for SeaWiFS 1 day).

In summary, SAR is still the most efficient and superior satellite sensor for oil spills detection, though it does not have capabilities for oil spill thickness estimation and oil type recognition. SAR is useful particularly for searching large areas and observing oceans at night and at cloudy weather conditions. Usually even small volumes of oil cover large areas (several hundred meters) and thus the need for very high spatial resolution in SAR images is not crucial. Wahl et al. [79]; Bern et al. [5]; Wahl et al. [81] found Low Resolution ERS-1 SAR images with a spatial resolution of 100 m sufficient for oil spill detection. The original ERS images were filtered using a 5×5 mean filter, which gave better noise characteristics than the full resolution images, and they were therefore easier to analyse. SAR also has some limitations, as a number of natural phenomena can give false oil spill detections. In addition, SAR is only applicable for oil spill monitoring in a certain range of wind speeds. The usefulness of SAR in terms of responding to oil spills at various conditions is covered in more detail in the next section.

3 Detectability of Oil Spills in SAR Images

Oil slicks dampen the Bragg waves (wavelength of a few cm) on the ocean surface and reduce the radar backscatter coefficient [2]. This results in dark regions or spots in a satellite SAR images. Fig. 1.2 gives two examples.

Kotova et al. [44] emphasises the importance of *weathering processes* [3], as they influence an oil spills physicochemical properties and detectability in SAR images. The processes that play the most important role for oil spill detection is evaporation, emulsification and dispersion. Lighter components of the oil will evaporate to the atmosphere. The rate of evaporation is dependent on oil type, thickness of the spill, wind speed and sea temperature. Emulsification is estimated based on water uptake as a function of the wind exposure of the actual oil type. Dispersion is an important factor in deciding the lifetime of an oil spill and

[2]Normalized Radar Cross Section (NRCS) (σ_0): The normalised measure of the radar return from e.g. the ocean is called the radar backscatter coefficient, or sigma nought. It is defined as the reflective strength of the radar target, σ, per unit area on the ground, A: $\sigma_0 = \frac{\sigma}{A}$ [12]. (How to derive the σ_0 value from the RADARSAT-1, ERS-1 and -2 SAR products are described in Laur et al. [46]; Shepherd [67]).

[3]Weathering processes include: spreading, drift, evaporation, dissolution, dispersion, emulsification, sedimentation, biodegradation and photooxidation [44].

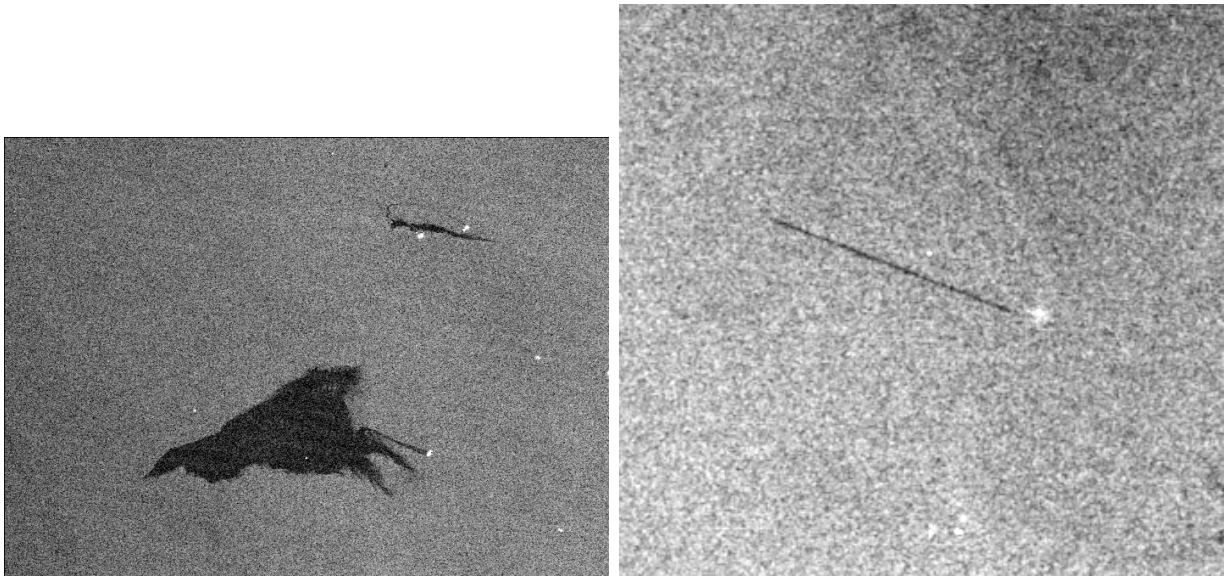


Figure 1.2: Left: A subscene of a RADARSAT-1 ScanSAR Narrow (SCN) near range mode image (30th of July 2002) containing two oil spill examples (classified as oil spill by manual inspection) on a homogeneous background. The size of the subscene is 1432×1032 pixels. Right: A subscene of an ENVISAT ASAR WSM image (24th of July 2003) containing a linear oil spill (classified as oil spill by manual inspection). The size of the subscene is 337×320 pixels. ©CSA/ESA/KSAT 2004.

it is strongly dependent on the sea state.

3.1 Discrimination Between Oil Spills and Look-alikes

A part of the oil spill detection problem is to distinguishing oil slicks from other natural phenomena that dampen the short waves and create dark patches on the surface. Natural dark patches are termed oil slick *look-alikes*. Oil slicks may include all oil related surface films caused by oil spills from oilrigs, leaking pipelines, passing vessels as well as bottom seepages, while look-alikes do include natural films/slicks[4], grease ice, threshold wind speed areas (wind speed < 3 m/s), wind sheltering by land, rain cells, shear zones, internal waves, etc. [17]. Oil spills in a narrow sense are only man-made slicks associated with crude petroleum and its products, heavy and light fuel. Fig. 1.3 shows an example of a dark spot detected as oil spill by three different satellite based analysis systems while verified as look-alike by aircraft.

The fact that the radar cross section values for oil spills are not unique poses a problem

[4]Natural film: microlayer of organic substances secreted by fish and several planctonic species [36].

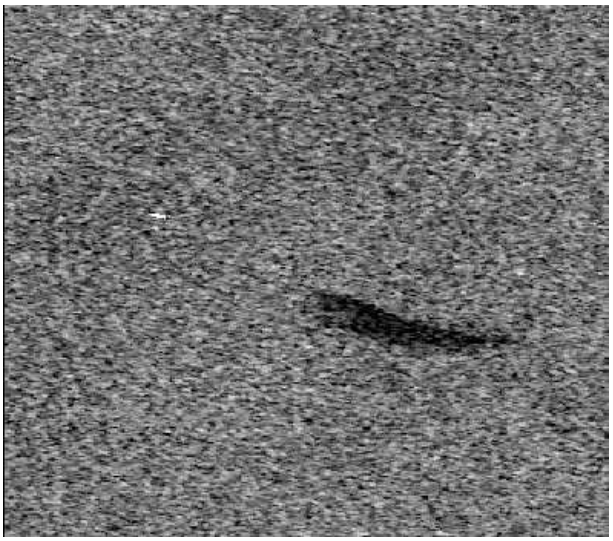


Figure 1.3: RADARSAT-1 SCN near range mode subscene (19th of July 2003) containing an oil slick look-alike (at 55,47N, 12,28E) detected as oil spill by Kongsberg Satellite Services AS (KSAT), QinetiQ and Norwegian Computing Centre (NR). The size of the subscene is 394×359 pixels. ©CSA/KSAT 2004.

in the development of an oil spill detection and monitoring system [35]. Natural films can be very difficult to distinguish from oil spills. Examples are presented in Hovland and Johannessen [35]. Due to a higher viscosity than natural films, oil spills tend to remain more concentrated and in turn provide larger dampening. In Espedal [17] some trends in slick properties were found for oil spills and natural films, but no one-to-one relationships were discovered. Oil spills are reported to give a dampening to the surrounding sea in the range of 0.6 dB to 13.0 dB, and natural films in the range of 0.8 dB to 11.3 dB. A preliminary version of a conceptual model for distinguishing oil spills from other slicks is proposed in Hovland et al. [36], and later on in Espedal [17, 16].

Even though we here focus on single frequency and single polarization SAR images, it is worth mentioning the possibility of a discrimination algorithm based on differences in multi-frequency and multi-polarization signatures. Maio et al. [53] proposes such an algorithm for discrimination between oil spills and false alarm templates. However, Gade et al. [28] did some experiments to investigate whether spaceborne L, C and X-band multi-polarization SARs are capable of discriminating between films of different chemical properties, and found that discrimination is only possible at low to moderate wind speeds. Differences between polarization signatures (film-covered as well as film-free surfaces) were found only for low wind speeds and small incidence angles.

Results from testing the performance of a semi-automatic oil spill detection algorithm, on 59 ERS-1 SAR images, show that oil spills that are misclassified as look-alikes fall into

three main categories [71]:

- Thin, piecewise-linear slicks. (These slicks might be caused by moving ships changing direction, or by changes in currents or wind directions affecting oil releases from stationary objects.)
- Low-contrast slicks in homogeneous sea.
- Slicks on a very heterogeneous background.

In addition to look-alikes, different kinds of pollution can cause slicks that are detectable in SAR images. Wahl et al. [80] gives some examples of ERS-1 images of fish oil and diesel, run-off water from an open depository, a controlled chemical spill and drilling fluid from an oil rig. The SAR sensor is currently not capable of distinguishing between the different pollutants. A single SAR-frequency is possibly not enough to estimate the thickness of the oil spill, but Jones [41] found for the large *Sea Empress* oil spill a good correlation between the largest reduction in backscatter and the thickest oil as determined by visual observations for a limited range of wind speeds (ca. 5-6 m/s).

These experiences require some attention when choosing features for discrimination between oil spills and look-alikes. Important features like wind speed, physical, geometrical and geographical parameters must be used to discriminate between oil spills and look-alikes.

3.2 The Wind Vector

SAR instruments have the advantage over optical sensors that they can acquire images of the oceans and coastal areas day and night and despite any weather conditions. However, the wind level influences the backscatter level and the visibility of slicks on the sea surface. Oil slicks are visible only for a limited range of wind speeds. Table 1.2 gives an overview for ERS SAR images. Gade et al. [30] studied oil spill pollution in the Baltic Sea, the North

Table 1.2: Visibility of slicks in SAR images (Bern et al. [5]; Perez-Marrodan [60]).

Wind speed	**Slick signatures**
0 m/s	No backscatter from the sea surface, hence no signature of oil slicks.
3 m/s	No impact from the wind on oil slicks.
	A high probability of oil slick look-alikes due to local wind variations.
3 to 7-10 m/s	Fewer false alarms from local low-wind areas.
	Oil slicks still visible and more homogeneous background.
>7-10 m/s	Only thick oil visible.
	Thinner oil slicks will be invisible due to a combination of oil spill dispersion.
	Thick oil can be visible with wind stronger than 10 m/s.

Sea and the north-western Mediterranean (700 ERS images). They found that with high

wind speed (> 10 m/s) few oil spills were detected in the SAR images, which is in agreement with Table 2. They also compared data from airborne surveillance with ERS SAR data. For both sensors they found that the maximum number of detected pollutions was found during summer time, April to September. A reason for this could possibly be that the average wind speed is higher at wintertime at all test sights (e.g. mean wind speed in the North Sea is above 10 m/s). A wind speed between 12 m/s to 14 m/s should possibly be considered as the upper limit for all spaceborne SAR imagery of oil spills [49], but the maximum wind speed for slick detection depends on oil type and the age (i.e. time since release) of the spill [5]. Thus, an estimate of the wind speed is valuable information for oil spill detection.

For the oil spill detection algorithm described by Solberg et al. [72], the wind level is set manually based on inspecting the image visually and it is used as input to a threshold procedure. Other possibilities are to incorporate wind information delivered by an external source (e.g. a forecasting centre) or to use automatic methods. With automatic methods for wind estimation, the wind speed can be derived directly from the SAR image. Salvatori et al. [64] estimates the wind speed from the SAR image by applying an inverted CMOD4[5] model. The wind vector appeared useful in knowing the evolution of the spill and to obtain correct classification.

3.3 Satellite Configurations for Oil Spill Detection

It is well understood that the radar image is a representation of the backscatter return and mainly proportional to the surface roughness at the scale of the radar wavelength (a phenomenon known as "Bragg scattering") [14]. The radar backscatter coefficient is also a function of the viewing geometry of the SAR, and the backscatter coefficient decreases with increasing incidence angle. The scattering properties of a material depend on the polarization of the incoming radar signal, and according to Girard-Ardhuin et al. [32] is the most suitable SAR configuration for slick detection C-band single-polarized VV SAR at incidence angles in the range of 20° to 45°. Dokken [13] writes that oil spill detection is possible under suitable meteorological conditions when $\sigma_{ocean} > noise\_floor + 4dB$. Oil spill detection is possible for incidence angles in the range of 20° (lower limit) to the intersection between the σ_{ocean} and the noise floor + 4 dB (upper limit). Because little multiple reflection of the signal occurs over the ocean there is little effect on the polarization. It is therefore not much to gain by using cross-polarization (HV or VH) [62]. Fortuny-Guasch [25] discusses the potential of polarimetric SAR for improved oil spill detection and classification. Its use might extend the validity ranges of wind speed and incidence angles.

For RADARSAT-1, Vachon et al. [78] recommend the SCN near range mode, see Table 1.3, for slick detection. ScanSAR Wide (SCW) can cause limitations because of the poor resolution of 100 m. RADARSAT-1 is not particularly suited for oil spill detection because of its HH polarization, as predictions show that oil spill detection will not be possible for

[5]CMOD4: developed by ESA for wind vector evaluation of C-band scatterometer.

the largest incident angles especially at low wind speeds [79]. However, it is successfully used in operating oil spill monitoring. No significant difference in practical performance between the detection capabilities of RADARSAT-1 versus ENVISAT has yet been reported for operational use, however the experience might still be limited (see also Shepherd et al. [66]). ENVISAT's ASAR WSM covers a much wider swath than its predecessors ERS-1 and -2, but the resolution is significantly lower. From Table 1.3 we can see that there is a trade-

Table 1.3: Examples of satellite modes.

SAR sensor	**mode**	**resolution (m)**	**pixel spacing (m)**	**swath width (km)**	**incidence angle (deg)**
ERS-2	PRI	30×26.3	12.5×12.5	100	20-26
ENVISAT	Image Mode (Precision Image)	30×30	12.5×12.5	100	15-45 (7 swaths)
RADARSAT-1	SCN	50×50	25×25	300	20-46
RADARSAT-1	SCW	100×100	50×50	450-500	20-49
ENVISAT	WSM	150×150	75×75	400	16-44

Approximate values are given. Product overviews: RADARSAT International [61]; European Space Agency [21]; ERS-2 Web-site [15]. Precision Image Mode (PRI).

off between image resolution and swath coverage. Generally, for efficient oil spill monitoring larger swath widths should be chosen on the expense of somewhat lower resolution.

Since the SAR satellites usually have polar orbits, the coverage depends on the latitude. Coverage is good in the polar regions and decreases with the distance from the poles. For the Mediterranean Sea the number of satellites passes per day is 0.04 for ERS, 0.27 for RADARSAT-1 (SCN near range) and 0.36 for ENVISAT (Wide Swath Mode) [60]. As the visibility of oil spills reduces with time (e.g. the rate of natural dispersion at moderate wind speed conditions is about 0.5-2% of the oil volume/hour [44]) and an early warning is wanted, a high number of passes per day is favourable. With steerable antennas, like the one planned for the RADARSAT-2 SAR, we get more flexible observation possibilities and thus less dependent e.g. on the repeat cycles of the satellites.

These are all important factors in estimating the usefulness and efficiency of the SAR, and in designing future satellite missions for pollution monitoring.

3.4 Speckle Noise

Speckle arises because the resolution of the sensor is not sufficient to resolve individual scatters within a resolution cell. Speckle is a large problem in SAR images since even a homogeneous area has a statistical distribution with large standard deviation. Incoherent

averaging multiple looks may reduce speckle [6] [47]. Another possibility is to smooth the speckles after the image has been formed. Barni et al. [2] tested two types of filters, general noise-reducing filters that do not assume any a priori speckle model and adaptive filters assuming a multiplicative speckle model (speckle noise by a multiplicative model is described by Lee [48]). Better results were reported for the latter type in their oil spill detection algorithm.

For oil spill applications, a filter should suppress speckle noise, but still preserve small and thin oil spills.

4 Methodology for Oil Spill Detection in SAR Images

We distinguish between manual approaches and automatic algorithms for oil spill detection. Detection of oil spills can be divided in [39]:

- Detection of suspected slicks.
- Manual verification of the slicks (oil/look-alike) and assignment of confidence levels.

This section addresses issues regarding the design of oil spill detection systems.

4.1 Manual Inspection

Since 1994 KSAT in Norway has provided a manual oil spill detection service. Here operators are trained to analyse SAR images for the detection of oil pollution. The KSAT approach is described by Indregard et al. [39]. External information about wind speed and direction, location of oilrigs and pipelines, national territory borders and coastlines are used as support during the analysis. The operator uses an image viewer that can calculate some spot attributes, but he/she still has to go through the whole image manually. This is time consuming. Possible oil spills found are assigned either high, medium or low confidence levels. The assignment is based on the following features: the contrast level to the surroundings, homogeneity of the surroundings, wind speed, nearby oilrigs and ships, natural slicks near by, and edge and shape characteristics of the spot. The determination of a confidence level is not exact science and there will always be an uncertainty connected to the results from manual inspection.

4.2 Manual Detection Compared to Automatic Detection

During manual inspection, contextual information is an important factor in classifying oil spills and look-alikes. A challenge is to somehow incorporate the "expert knowledge" into

[6] The bandwidth of the SAR system is divided into N discrete contiguous segments, each segment generating a single-look lower resolution image. The intensities of the N single-look images are incoherently averaged to form a multiple-look image [77].

the automatic algorithm. In Solberg and Volden [73] a set of rules and knowledge about external conditions (e.g. wind speed) are used to adjust prior probabilities of oil slicks in the scene. This information is incorporated into a classifier based on a multivariate probability distribution function.

Fiscella et al. [24] found that a human image interpreter and a classification algorithm have similar ability to discriminate oil spills from look-alikes, but the image set used contained only 21 oil spill candidates.

A study of best practise, based on a comparison of KSAT's manual approach, NR's automatic algorithm (described in Bjerde et al. [8]; Solberg and Solberg [71]; Solberg and Volden [73] and lately in Solberg et al. [72, 70]) and QinetiQ's semi-automatic oil spill detection approach, has been performed by the ongoing Oceanides project [39]. QinetiQ's semi-automatic approach covers only the first step of an automatic algorithm, dark spot detection, and therefore the output targets must be classified visually by an operator. In this study the three satellitebased approaches were compared to airborne verifications in a satellite-airborne campaign. The study was done without the operators or the algorithms knowing of the aircraft verifications. (The benchmark set consisted of 32 RADARSAT-1 images). This data set contained 17 verified oil spills. KSAT detected 15 of these slicks, NR's algorithm detected 14, and QinetiQ detected 12. The results show that a challenge is to have all operators pick out the same spots and assign the same confidence levels. NR's algorithm is objective (with one exception of manual wind level assignment, see section 3.2) and produces the same result repeatedly. Good agreement was found for high-contrast slicks among the various methods, but there were some differences on low-contrast slicks. The operators at KSAT use 3-25 minutes to analyse a scene (on average 9 min.), the NR's algorithm used about 3 minutes and QinetiQ's algorithm used 20 minutes per scene in average. This shows that automatic approaches are more feasible as the volume of SAR data grows.

4.3 Design Issues for Automatic Detection Systems

Kubat et al. [45] identified a number of issues during their development of a machine learning component for an oil spill detection system. The first issue is the *scarcity of data*, as most satellite SAR images acquired contain no oil spills. The second is the *imbalanced training set*. Oil spill detection is an application where the classifier should detect a rare but important event (look-alikes appear much more frequent). Oil spills happen to appear in *batches*, where examples drawn from the same image constitute a single batch. A similarity within batches and dissimilarity between batches can be present, which can influence the classifier training. And finally, the *performance* of the classifier relates to the users possibility to decrease the number of alarms at the expense of missing genuine oil spills (specificity). Indregard et al. [39] also points out that automatic algorithms should be tuned in order to be certain that no suspicious oil spills that would be detected by an operator are missed.

The significant differences in mode characteristics for different SAR sensors, as presented by Table 1.3, suggests a development of sensor specific modules.

All of these are considerations to be taken into account when designing an automatic system for oil spill detection.

5 Automatic Techniques for Oil Spill Detection in SAR Images

Several of the published papers on oil spill algorithms for SAR images (e.g. Fiscella et al. [24]; Frate et al. [26]; Solberg et al. [72]) describe a structure comparable with the one in Fig. 1.4.

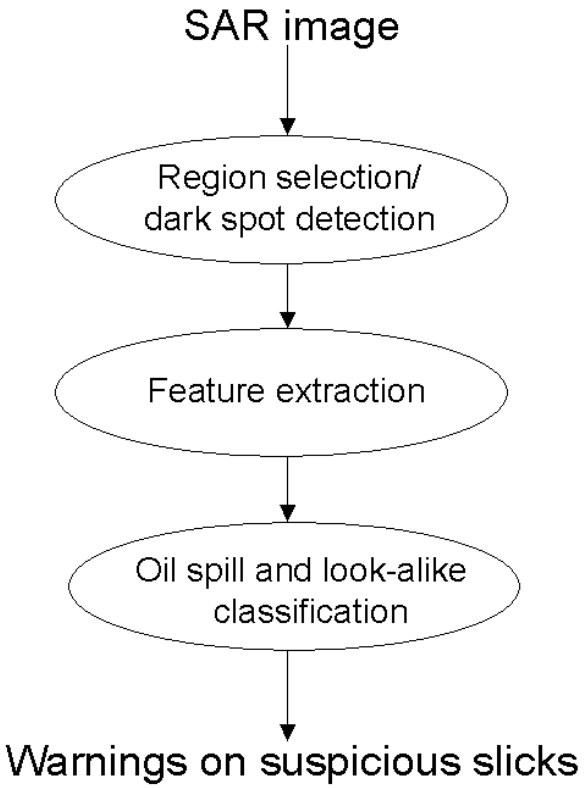


Figure 1.4: A framework for oil spill detection algorithms.

The importance of the wind vector was emphasised in section 3.2, and Salvatori et al. [64] include two additional steps of wind direction estimation and wind speed calculation. Manual wind estimation was included by Solberg et al. [72]. SAR image calibration, land masking, speckle reduction and class signature databases belong in this framework as well, but Fig. 1.4 shows the core modules of an oil spill detection and classification algorithm.

This section covers algorithms and techniques related to each module of Fig. 1.4.

5.1 Segmentation Techniques

As oil spills are characterised by low backscattering levels this suggest the use of thresholding for dark spot segmentation. An early attempt on segmentation of ERS-1 SAR images is described by Skøelv and Wahl [69]. The algorithm simply looks for bimodal histograms in widows of size $N \times N$ pixels (N was sat to 25 pixels). This is reported as a good method for detection of oil spills provided that the spill is not to thin. A similar approach is briefly described in Vachon et al. [78]; Manore et al. [54]. This algorithm, which is developed for RADARSAT-1 SAR data, spatially averages the image before a user defined adaptive threshold is applied. As both these algorithms lack a classification step look-alikes will be detected as well.

Solberg et al. [72, 70] apply an adaptive algorithm were the threshold is sat k dB below the mean value estimated in a moving window. The thresholding is combined with a multiscale pyramid approach and a clustering step to better separate the spill from its surroundings. Noise reduction by a mean filter smoothes the edges.

The use of hysteresis thresholding was introduced by Canny [9] and is applied by Kanaa et al. [42] for detecting oil spills in ERS amplitude images. A search is done in the 8 neighbourhood directions followed by a merging step of the responses. Linear features are reported accentuated by this method.

An oil slick detection approach, based on the Laplace of Gausian (LoG) and Difference of Gaussian (DoG) operators, is described in Change et al. [10] and Chen et al. [11]. The LoG operator is applied on the coarsest layer of a 2×2 pixel reduced pyramid with three layers. The concave areas of the grey level surface are selected. The DoG is used to locate those areas with more than half of the slick boundary pixels greater than $\mu + 1.75\sigma$ (as selected for ERS-1, where μ and σ are mean and standard deviation over all image pixels). To improve the result the finer layers of the pyramid are processed.

The use of wavelets in ocean feature detection (including oil spills) is described by Liu et al. [50] and Wu and Liu [82]. In the general linear feature detection scheme the analysing wavelet is defined as the LoG. Regions with multiple histogram peaks are selected for the wavelet transform. The wavelet is applied as an edge detector as the contours of the zero crossing indicates the feature edges (see Canny [9]).

As oil spills dampens the capillary waves, Mercier et al. [56] suggests a segmentation method based on detecting local variations of the wave spectra. First a multi-resolution analysis is achieved by a wavelet packet transform then a Hidden Markov Chain (HMC) model is applied to the wavelet coefficients. The technique is tested on an ERS PRI image.

QinetiQ's dark spot algorithm uses a Constant False Alarm Rate (CFAR) algorithm to locate dark regions. The dark spots are merged according to a clustering radius and a threshold, and the Hough transform is used specially to identify linear targets [39].

To allow dealing with mixed surface-cover classes and unsharp boundaries among regions Barni et al. [2] proposes an algorithm based on fuzzy clustering. A membership function $u_A(x)$ is assigned to each pixel x, which measures how much the pixel belongs to a set A. The Fuzzy C-means (FCM) algorithm is applied, and a pyramid structure is used in finding the membership values. Uncertain pixels are tested in the lower pyramid level. Neighbouring regions are identified, and a Sobel operator is used to enhance the main edges of the original filtered image. Regions, whose common border does not have a high enough percentage of large gradient points, are merged together. One difficulty with fuzzy clustering is to find the optimum number of clusters.

A method using mathematical morphology for oil spill segmentation is presented by Gasull et al. [31]. Combinations of opening and closing [7] operations are used for oil spill filtering and thresholding. The algorithm aims at detecting spills from sailing tankers, and some features used are the elongatedness and dampening of the spill.

Even though a variety of methods are applied, the common goal is to detect all suspicious slicks and to preserve the slick shapes. The latter is of most importance for the success of discriminating oil spills from look-alikes in the following steps.

5.2 Slick Feature Extraction

From the thresholded dark spot image, feature extraction is used to compute features for each slick. Table 1.4 summarizes the features used in three different algorithms. In a feature vector that is input to the classifier, the individual features are typically covered by the following classes:

- The geometry and shape of the segmented region.

  Geometric and shape features are applied by all methods in Table 1.4. To detect pollution from sailing tankers cleaning their tanks, an important feature is elongatedness which can be expressed as a ratio between the width and length of the slick [31]. Another possible feature useful in identifying these spots is the first invariant statistical moment of Hu [38].

- Physical characteristics of the backscatter level of the spot and its surroundings.

  Frate et al. [26] found that features containing most valuable information for classification by neural networks were features covering information on the gradient of the backscattering value when passing from background to spill (#13, #14 and #15). In addition, the background standard deviation (#10) was found important which is a parameter highly affected by the wind level and is generally high for natural sea slicks. Similarly in Fiscella et al. [24], features connected to the background surrounding the

[7] Opening: erosion followed by dilation. Closing: dilation followed by erosion (see Sonka et al. [74] for an introduction to mathematical morphology).

Table 1.4: Features applied by various algorithms.

#	**Feature**	**1**	**2**	**3**
1	Slick area (A)	x	x	x
2	Slick perimeter (P)	x		x
3	P/A			x
4	Slick complexity	x	x	
5	Spreading (low for long thin slicks, high for circular shape)	x		
6	Slick width		x	
7	First invariant planar moment [38]		x	
8	Dispersion of slick pixels from longitudal axis			x
9	Object/dark area standard deviation	x		x
10	Background/outside dark area standard deviation	x		x
11	Max contrast (between object and background)	x		
12	Mean contrast (between object and background)	x		
13	Max border gradient	x		
14	Mean border gradient	x	x	
15	Gradient standard deviation	x		
16	Local area contrast ratio		x	
17	Power-to-mean ratio of the slick		x	
18	Homogeneity of surroundings		x	
19	Average NRCS inside dark area			x
20	Average NRCS in limited area outside dark area			x
21	Gradient of the NRCS across the dark area perimeter			x
22	Ratio #9 to #10			x
23	Ratio #19 to #9			x
24	Ratio #20 to #10			x
25	Ratio #23 to #24			x
26	Ratio #19 to #20			x
27	Distance to a point source		x	
28	Number of detected spots in the scene		x	
29	Number of neighbouring spots		x	

1: Frate et al. [26] , 2: Solberg et al. [72] and 3: Fiscella et al. [24]. x indicates that the parameter is used in the feature vector of the particular algorithm.

slick were found to be important due to the wind speed dependence of oil spill observations (these features could also be classified as contextual features).

- Spot contextual features.

  Examples are slick location relative to the shore and distance to ships and oilrigs. In the contextual analysis of the supervised discrimination algorithm described by Espedal [16] is a "hot-spot" pollution source database used. Improved classification results were found by Solberg and Volden [73] when the dark spots are classified in the context of their surroundings and weather information is incorporated. Espedal and Wahl [19] suggest using wind history information for slick classification and slick age estimation.

Wind history can also be looked at as an indirect spot feature.

- Texture.

  In contradiction to the pixel intensity itself, texture provides information about the spatial correlation among neighbouring pixels. Assilzadeh and Mansor [1] describe an early warning system where texture features based on grey level co-occurrence matrixes (GLCM)[8] are used. Homogeneity and Angular Second Moment were found effective in separation of oil spills from other objects. Power-to-mean ratio of the slick and the surroundings is used by Solberg et al. [72] as a measure of homogeneity.

Even though the different methods in Table 1.4 does not apply the exact same features are several of the features different measures of the same characteristic.

Fractal texture description can be used to describe natural surfaces [59]. The use of fractal dimension[9] as a feature for classifying observed ocean radar signatures is suggested in Gade and Redondo [29]. A box-counting algorithm (the method is described in Keller et al. [43]) is used to find the fractal dimension D. A difference in D of oil spills compared to other oceanic phenomena is reported found. In another paper on fractal dimension by Benelli and Garzelli [3], was a steady fractal dimension value of $D = 2.45$ found for the sea surface, while an average value of $D = 2.15$ was found for oil spills. A smaller D indicates less roughness.

Good features are important, but the lack of good guidelines on how to acquire them has been pointed out by Kubat et al. [45].

5.3 Classification Methods

As a number of phenomena can create dark patches in a SAR image, the purpose of the classifier is to distinguish oil spills from the other cases.

Based on results from selecting the dark regions with a NRCS lower than one half of the average NRCS of the sea area in the image, Fiscella et al. [24] apply a Mahalanobis classifier to estimate the probability p of a dark spot being an oil spill. For ERS images, $p > 2/3$ and not looking at uncertain cases, 93% of the oil spills were correct classified (*a priori*) and 82% of the oil spills in the test set were correct classified. This was compared with a compound probability classifier. Here $p > 2/3$ gave a 85% correct classification rate (*a priori*) of oil spills and 91% of the oil spills in the test set were correct classified. A training set of 80 oil spills and 43 look-alikes and a test set of 11 oil spills, 4 uncertain and 6 look-alikes were used.

In Solberg et al. [72] is a probability assigned to a spot from a multivariate Gaussian density function. The unknown parameters are derived from a signature database. (Fiscella et al.'s method is in a similar way based on previous measurements.) This is combined

[8] GLCM: approximates the grey level joint probability distribution [34].

[9] A surfaces fractal dimension corresponds closely to our intuitive notion of roughness [59].

with a prior model for the number of look-alikes, a model for the presence of a slick in the vicinity of a bright object and a rule-based modification of the probability density to take into account feature combinations that are indications of certain scene conditions. The leave-one-out method with 84 scenes gave a correct classification of 94% for oil spills.

An artificial neural network (NN) is a mathematical model composed of many neurons[10] operating in parallel. An approach for oil spill detection, tested on ERS images, based on a multilayer perceptron (MLP) neural network with two hidden layers is described by Frate et al. [26]. Selection of a dark object and image fragmentation is performed with manual interaction, while feature extraction and classification into oil spill or look-alike is performed automatically. The input (to the net) is also here a feature vector. The neuron is the elemental building block of each layer in the net, and it is mainly characterized by an activation function (in this case a non-linear sigmoid function). The neuron computes the sum of its inputs, adds a bias term, and drives the result through the activation function. A single output is produced from each neuron. Using the leave-one-out method with 139 sample oil spills, 18% were reported misclassified as look-alikes.

The proposed automatic methods for oil spill detection discussed here reports accuracies from 82% to 94% correct classification of oil spills. These studies are however performed on different data sets, and the approaches differ both in terms of segmentation approach, feature extraction methods, and classification methodology. The most important steps in this process are segmentation and feature extraction. If a slick is not detected during segmentation, it cannot be classified correctly. If the features have good discriminatory power, the classification problem will be easier and several classifiers will work. We believe that the variance in shape, contrast, and surroundings of oil slicks and look-alikes is so large that it is necessary to subdivide the problem into subclasses, and to guide the classifier by using as much prior information about the problem as possible. This was done in the approach presented in Solberg et al. [72], which is the approach with the highest accuracy reported so far.

6 Conclusion and Suggestions for Further Work

Synthetic aperture radar is the most applicable spaceborne sensor for operational oil spill detection, mostly because of its all weather/all day detection capabilities and wide coverage. It can operate from light wind to wind speeds up to 12-14 m/s, but the maximum wind speed for oil slick detection depends on oil type and age. Sensors operating in wide swath mode with a spatial resolution of 50-150 m are found to be sufficient and allows covering large ocean areas efficiently.

The largest challenge in detection of oil spills in SAR images is accurate discrimination between oil spills and look-alikes. Most low wind situations can be handled by analysing the

[10] A neuron is a non-linear computational element connected to other neurons by links characterized by different weights (see also Bishop [7]).

surroundings of a slick, but natural films cannot always be properly distinguished from oil spills based on a SAR image alone. Additional information about algal blooms is desired, particularly for the Baltic Sea, where algae is common during the summer. Such information can be derived from optical sensors. Future oil spill systems should incorporate algae information either from multisensor studies, or by using prior knowledge about the likelihood of observing alga in a given area at a certain time of the year.

For operational purposes there is a need for coordination between satellite overpasses and aerial surveillance flights. Presently, aerial surveillance is needed for collecting evidence to prosecute the polluters. The combination of coordinated satellite image acquisitions and aerial surveillance flights is presently used operationally by many countries in Northern Europe. Many of these services use the KSAT manual approach to identify oil spills from the satellite images.

Future SAR missions are crucial for sustainable operational oil spill detection services. There is a number of commercial SAR missions planned; the Japanese Advanced Land-Observing Satellite (ALOS) and the European TerraSAR-L are both satellites with L-band SAR, which means an increased wavelength (24 cm) compared to e.g. ERS and ENVISAT. TerraSAR-X and COSMO/SkyMed (dual use) are German and Italian X-band satellites. TerraSAR-X will have a best resolution of 1 m, but will also provide a ScanSAR mode with 16 m resolution and 100 km swath width. Concerning oil spill detection, the most promising of the coming missions is RADARSAT-2 with its C-band SAR. RADARSAT-2 is planned launched in 2005, and will provide improved resolution and flexibility in selection of polarization options (both single and full polarization modes will be available). A future spaceborne C-band SAR to replace ENVISAT ASAR would be needed to sustain the quality of the oil spill detection services. This is acknowledge by ESA, as many other ocean applications use C-band SAR images regularly.

A study of best practice of manual versus automatic oil spill systems showed that operators show some variance in detecting spills, particularly in assigning an oil spill confidence estimate. An automatic algorithm with a reliable and objective oil spill confidence estimate would be highly desirable. The need for automatic algorithms depend on the number of images to be analysed, but for monitoring large ocean areas it is a cost-effective alternative to manual inspection.

More work on the direct comparison of the performance of manual versus automatic methods for oil spill detection is needed. Up to now, the automatic systems have been tested off-line, thus, additional spills reported by the automatic systems cannot be verified. We still believe that the slicks classified as oil by automatic algorithms should go through a manual inspection prior to sending out aircrafts. In that case, inspection of a couple of slicks per scene would be much more efficient than inspection of the complete scene as currently done. As part of the ESA project Northern View, NR's automatic algorithm will be deployed in KSAT's operational environment.

Automatic oil spill detection algorithms are normally divided into three steps, dark spot detection, dark spot feature extraction, and dark spot classification. Few papers are pub-

lished on automatic algorithms for classification of oil spills and its look-alikes as most authors focus on the detection step. Large-scale classification studies with acceptable classification performance are reported based on statistical classification [24; 72], and neural nets [26]. An extensive comparison of the classifiers used by the different approaches, based on the same data set of features, would be desirable.

To increase the performance further, incorporation of more knowledge is needed. We believe that the future oil spill system should be an integrated system, including automatic algorithms, a database of "hotspots" (e.g. oilrigs, sunken ships and seepages), ship lanes, algae information, and more extensive use of wind information.

Acknowledgements

This work is performed as a part of a PhD study funded by the Norwegian Research Council and the Norwegian Defence Research Establishment. The authors would like to thank the Oceanides project, in particular KSAT, for the SAR scenes used for illustration purposes.

References

[1] Assilzadeh, H., Mansor, S. B., 2001. Early warning system for oil spill using SAR images. *Proc. ACRS 2001 - 22nd Asian Conference on Remote Sensing, 5-9 November 2001, Singapore* 1, 460–465.

[2] Barni, M., Betti, M., Mecocci, A., 1995. A fuzzy approach to oil spill detection on SAR images. *Proc. IGARSS '95*, 1, 157–159.

[3] Benelli, G., Garzelli, A., 1999. Oil-spills detection in SAR images by fractal dimension estimation. *Proc. IGARSS'99*, 1, 218–220.

[4] Bentz, C., de Miranda, F. P., 2001. Application of remote sensing data for oil spill monitoring in the Guanabara Bay, Rio de Janeiro, Brazil. *Proc. IGARSS'01*, 1, 333–335.

[5] Bern, T.-I., Moen, S., Wahl, T., Anderssen, T., Olsen, R., Johannessen, J. A., 1992. Oil spill detection using satellite based SAR. Completion report for Phase 0 and 1. Tech. rep., OCEANOR report no. OCN-R92071, Trondheim.

[6] Bern, T.-I., Wahl, T., Anderssen, T., Olsen, R., 1992. Oil spill detection using satellite based SAR: Experience from a field experiment. Proc. 1st ERS-1 Symposium, Cannes, France, 4-6 November 1992, 829–834.

[7] Bishop, C. M., 1995. *Neural Networks for Pattern Recognition*. Oxford University Press.

[8] Bjerde, K. W., Solberg, A. H. S., Solberg, R., 1993. Oil spill detection in SAR imagery. *Proc. IGARSS'93*, 3, 943–945.

[9] Canny, J., 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8 (6), 679–698.

[10] Change, L. Y., Chen, K., Chen, C., Chen, A., 1996. A multiplayer-multiresolution approach to detection of oil slicks using ERS SAR image. *Proc. ACRS 1996 - 17th Asian Conference of Remote Sensing, Sri Lanka*.

[11] Chen, C. F., Chen, K. S., Chang, L. Y., Chen, A. J., 1997. The use of satellite imagery for monitoring coastal environment in Taiwan. *Proc. IGARSS'97*, 3, 1424–1426.

[12] Curlander, J. C., McDonough, R. N., 1991. *Synthetic Aperture Radar, Systems and Signal Processing*. Wiley series in remote sensing, John Wiley & Sons, Inc.

[13] Dokken, S. T., 1995. Optimal bruk av avanserte radarsatellitter. Master's thesis, Institutt for Tekniske Fag, Norges Landbrukshøgskole.

[14] Elachi, C., 1987. *Introduction to the Physics and Techniques of Remote Sensing*. John Wiley & Sons Inc.

[15] ERS-2 Web-site, 2004. http://earth.esa.int/services/pg/pgerssarpri.xml, accessed 20 October 2004.

[16] Espedal, H., 1999. Detection of oil spill and natural film in the marine environment by spaceborne SAR. *Proc. IGARSS'99*, 3, 1478–1480.

[17] Espedal, H. A., 1998. Detection of oil spill and natural film in the marine environment by spaceborne synthetic aperture radar. Ph.D. thesis, Department of Physics University of Bergen and Nansen Environment and Remote Sensing Center, Norway.

[18] Espedal, H. A., Johannessen, O. M., 2000. Detection of oil spills near offshore installations using synthetic aperture radar (SAR). *Int. J. Remote Sensing*, 21 (11), 2141–2144.

[19] Espedal, H. A., Wahl, T., 1999. Satellite SAR oil spill detection using wind history information. *Int. J. Remote Sensing*, 20 (1), 49–65.

[20] European Space Agency, 1998. Oil pollution monitoring. *ESA brochure: ERS and its Applications - Marine, BR-128* 1.

[21] European Space Agency, 2002. ASAR product handbook. Tech. rep., European Space Agancy - ENVISAT Product Handbook, Issue 1.1, 1 December 2002.

[22] Fingas, M., 2001. *The Basics of Oil Spill Cleanup*. Lewis Publishers.

[23] Fingas, M. F., Brown, C. E., 1997. Review of oil spill remote sensing. *Spill Science & Technology Bulletin*, 4, 199–208.

[24] Fiscella, B., Giancaspro, A., Nirchio, F., Pavese, P., Trivero, P., 2000. Oil spill detection using marine sar images. *Int. J. Remote Sensing*, 21 (18), 3561–3566.

[25] Fortuny-Guasch, J., 2003. Improved oil spill detection and classification with polarimetric SAR. *Proc. workshop on Application of SAR Polarimery and Polarimetric Interferometry, ESA-ESRIN Frascati, Italy, January 14-16 2004*, available at http://earth.esa.int/polinsar/pr.html, accessed 1 September 2004.

[26] Frate, F. D., Petrocchi, A., Lichtenegger, J., Calabresi, G., 2000. Neural networks for oil spill detection using ERS-SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, 38 (5), 2282–2287.

[27] Friedman, K. S., Pichel, W. G., Clemente-Colón, P., Li, X., 2002. GoMEx - an experimental GIS system for the Gulf of Mexico Region using SAR and additional satellite and ancillary data. *Proc. IGARSS'02*, 6, 3343–3346.

[28] Gade, M., Alpers, W., Bao, M., 1996. Measurements of the radar backscattering over different oceanic surface films during the SIR-C/X-SAR campaigns. *Proc. IGARSS'96*,, 860–862.

[29] Gade, M., Redondo, J., 1999. Marine pollution in European coastal waters monitored by the ERS-2 SAR: a comprehensive statistical analysis. *OCEANS '99 MTS/IEEE. Riding the Crest into the 21st Century*, 3, 1239–1243.

[30] Gade, M., Scholz, J., von Viebahn, C., 2000. On the detectability of marine oil pollution in European marginal waters by means of ERS SAR imagery. *Proc. IGARSS 2000*, 6, 2510–2512.

[31] Gasull, A., Fabregas, X., J.Jimenez, Marques, F., V.Moreno, Herrero, M., 2002. Oil spills detection in SAR images using mathematical morphology. *Proc. EUSIPCO'2002, Toulouse, France, September 2002* 1, 25–28.

[32] Girard-Ardhuin, F., Mercier, G., Garello, R., 2003. Oil slick detection by SAR imagery: potential and limitation. *Proc. OCEANS 2003*, 1, 164–169.

[33] Goodman, R., 1994. Overview and future trends in oil spill remote sensing. *Spill Science & Technology Bulletin*, 1 (1), 11–21.

[34] Haralick, R. M., 1979. Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67, 786–804.

[35] Hovland, H. A., Johannessen, J. A., 1994. Norwegian surface slick report, Final report to Norwegian Defence Research Establishment and Norwegian Space Centre. Tech. rep., Nansen Environmental and Remote Sensing Center.

[36] Hovland, H. A., Johannessen, J. A., Digranes, G., 1994. Slick detection in SAR images. *Proc. IGARSS'94*, 4, 2038–2040.

[37] Hu, C., Müller-Krager, F. E., Taylor, C. J., Myhre, D., Murch, B., Odriozola, A. L., Godoy, G., 2003. MODIS detects oil spills in Lake Maracaibo, Venezuela. *EOS, Transactions, American Geophysical Union*, 84 (33).

[38] Hu, M.-K., 1962. Visual pattern recognition by moment invariants. *IEEE Trans. Inform. Theory*, 8, 179–187.

[39] Indregard, M., Solberg, A., Clayton, P., 2004. D2-report on benchmarking oil spill recognition approaches and best practice. Tech. rep., Oceanides project, European Commission, Archive No. 04-10225-A-Doc, Contract No: EVK2-CT-2003-00177.

[40] Ivanov, A., He, M.-X., Fang, M.-Q., 2002. Oil spill detection with the RADARSAT SAR in the waters of the Yellow and East China Sea: A case study. *Proc. ACRS 2002 - 23rd Asian Conference on Remote Sensing, November 25-29, 2002, Kathmandu, Nepal*.

[41] Jones, B., 2001. A comparison of visual observations of surface oil with synthetic aperture radar imagery of the Sea Empress oil spill. *Int. J. Remote Sensing*, 22 (9), 1619–1638.

[42] Kanaa, T. F. N., Tonye, E., Mercier, G., Onana, V., Ngono, J., Frison, P., Rudant, J., Garello, R., 2003. Detection of oil slick signatures in SAR images by fusion of hysteresis thresholding responses. *Proc. IGARSS'03*, 4, 2750–2752.

[43] Keller, J. M., Chen, S., Crownover, R. M., 1989. Texture description and segmentation through fractal geometry. *Computer Vision, Graphics, and Image Processing* 45, 150–166.

[44] Kotova, L. A., Espedal, H. A., Johannessen, O. M., 1998. Oil spill detection using spaceborne SAR; a brief review. Proc. 27th ISRSE, Tromsø, Norway.

[45] Kubat, M., Holte, R. C., Matwin, S., 1998. Machine learning for the detection of oil spills in satellite radar images. *Machine Learning*, 30, 195–215.

[46] Laur, H., Bally, P., Meadows, P., Sanchez, J., Schaettler, B., Lopinto, E., Esteban, D., 2002. Derivation of the backscattering coefficient σ^0 in ESA ERS SAR PRI products. Tech. rep., Document No.: ES-TN-RS-PM-HL09, Issue 2, Rev. 5d, 17. September, European Space Agency (ESA).

[47] Leberl, F. W., 1990. *Radargrammetric image processing*. Artech House.

[48] Lee, J.-S., 1981. Speckle analysis and smoothing of synthetic aperture radar images. *Computer Graphics and Image Processing*, 17, 24–32.

[49] Litovchenko, K., Ivanov, A., Ermakov, S., 1999. Detection of oil slicks parameters from ALMAZ-1 and ERS-1 SAR imagery. *Proc. IGARSS'99*, 3, 1484–1486.

[50] Liu, A. K., Peng, C. Y., Chang, S. Y.-S., 1997. Wavelet analysis of satellite images for costal watch. *IEEE Journal of Oceanic Engineering*, 22 (1), 9–17.

[51] Lu, J., 2003. Marine oil spill detection, statistics and mapping with ERS SAR imagery in south-east Asia. *Int. J. Remote Sensing*, 24 (15), 3013–3032.

[52] Lu, J., Lim, H., Liew, S. C., Bao, M., Kwoh, L. K., 1999. Ocean oil pollution mapping with ERS synthetic aperture radar imagery. *Proc. IGARSS'99*, 1, 212–214.

[53] Maio, A. D., Ricci, G., Tesauro, M., 2001. On CFAR detection of oil slicks on the ocean surface by multifrequency and/or multipolarization SAR. *In Radar Conference, 2001. Proceedings of the 2001 IEEE*, 351–356.

[54] Manore, M. J., Vachon, P. W., Bjerkelund, C., Edel, H. R., Ramsay, B., 1998. Operational use of RADARSAT SAR in the costal zone: The Canadian experience. *27th International Symposium on Remote Sensing of the Environment, Tromsø, Norway, June 8-12*, 115–118.

[55] Martinez, A., Moreno, V., 1996. An oil spill monitoring system based on SAR images. *Spill Science & Technology Bulletin*, 3 (1/2), 65–71.

[56] Mercier, G., Derrode, S., Pieczynski, W., Caillec, J.-M. L., Garello, R., 2003. Multiscale oil slick segmentation with Markov Chain Model. *Proc. IGARSS'03*, 6, 3501–3503.

[57] Oceanides Web-site, 2004. http://oceanides.jrc.cec.eu.int/prestige.html, accessed 13 August 2004.

[58] Pavlakis, P., A.Sieber, Alexandry, S., 1996. Monitoring oil-spill pollution in the Mediterranean with ERS SAR. *ESA Earth Observation Quarterly, June 1996* (52).

[59] Pentland, A. P., 1984. Fractal-based description of natural scenes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6 (6), 661–674.

[60] Perez-Marrodan, M., 1998. ENVISYS - environmental monitoring warning and emergency management system. *Proc. of the AFCEA Kiev Seminar, 28-29 May*, 122–132.

[61] RADARSAT International, 2000. D4, RADARSAT, data products specification. Tech. rep., Doc. No: RSI-GS-026, RADARSAT International.

[62] Robinson, I. S., 1994. *Satellite oceanography. An Introduction for Oceanographers and Remote-Sensing Scientists*. Wiley-Praxis series in remote sensing.

[63] Salem, F., Kafatos, P. M., 2001. Hyperspectral image analysis for oil spill mitigation. *Proc. ACRS 2001 - 22nd Asian Conference on Remote Sensing, 5-9 November 2001, Singapore* 1, 748–753.

[64] Salvatori, L., Bouchaib, S., Frate, F. D., Lichtenneger, J., Samara, Y., 2003. Estimating the wind vector from radar SAR images when applied to the detection of oil spill pollution. *Fifth International Symposium on GIS and Computer Catography for Costal Zone Management, CoastGIS'03*.

[65] Shepherd, I., 2004. Developing an operational oil-spill service in GMES. *Version 2, OCEANIDES workshop, 25 May, 2004, EEA, Copenhagen*, available at: http://intelligence.jrc.it/marine/oceanides/oceanides.html, accessed 23 August 2004.

[66] Shepherd, I., Bauna, T., Chesworth, J., Kourti, N., Lemoine, G., Indregard, M., 2004. Use of ENVISAT at JRC for marine monitoring in 2003. *JRC technical note*, available at:
http://pta.jrc.cec.eu.int/entity.gx/_entity.object/SEC—
000000000002CB39/_entity.name/envisatreport-7.pdf, accessed 23 August 2004.

[67] Shepherd, N., 2000. Extraction of beta nought and sigma nought from RADARSAT CDPF products. Tech. rep., Report No.: AS97-5001, Rev. 4-28 April, Altrix Systems.

[68] Simecek-Beatty, D., Clemente-Colón, P., 2004. Locating a sunken vessel using SAR imagery: detection of oil spilled from the SS Jacob Luckenbach. *Int. J. Remote Sensing*, 25 (11), 2233–2241.

[69] Skøelv, Å., Wahl, T., 1993. Oil spill detection using satellite based SAR, Phase 1B completition report. Tech. rep., Norwegian Defence Research Establishment.

[70] Solberg, A. H. S., Dokken, S. T., Solberg, R., 2003. Automatic detection of oil spills in Envisat, Radarsat and ERS SAR images. *Proc. IGARSS'03*, 4, 2747–2749.

[71] Solberg, A. H. S., Solberg, R., 1996. A large-scale evaluation of features for automatic detection of oil spills in ERS SAR images. *Proc. IGARSS'96*, 3, 1484–1486.

[72] Solberg, A. H. S., Storvik, G., Solberg, R., Volden, E., 1999. Automatic detection of oil spills in ERS SAR images. *IEEE Transactions on Geoscience and Remote Sensing*, 37 (4), 1916–1924.

[73] Solberg, A. H. S., Volden, E., 1997. Incorporation of prior knowledge in automatic classification of oil spills in ERS SAR images. *Proc. IGARSS'97*, 1, 157–159.

[74] Sonka, M., Hlavac, V., Boyle, R., 1999. *Image Processing, Analysis, and Machine Vision*. Brooks/Cole Publishing Company.

[75] Trieschmann, O., Hunsänger, T., Tufte, L., Barjenbruch, U., 2003. Data assimilation of an airborne multiple remote sensor system and of satellite images for the North- and Baltic sea. *Proceedings of the SPIE 10th Int. Symposium on Remote Sensing, conference "Remote Sensing of the Ocean and Sea Ice 2003"*, 51–60.

[76] Tseng, W. Y., Chiu, L. S., 1994. AVHRR observations of Persian Gulf oil spills. *Proc. IGARSS'94*, 2, 779–782.

[77] Ulaby, F. T., Moore, R. K., Fung, A. K., 1986. *Microwave Remote Sensing: Active and Passive. Vol. 3: From Theory to Applications*. Artech house, Inc.

[78] Vachon, P. W., Thomas, S. J., Cranton, J. A., Bjerkelund, C., F.W.Dobson, R.B.Olsen, 1998. Monitoring the costal zone with the RADARSAT satellite. *Oceanology International 98, UK, March 10-13*, 10 pages.

[79] Wahl, T., Anderssen, T., Skøelv, Å., 1994. Oil spill detection using satellite based SAR, Pilot Operation Phase, final report. Tech. rep., Norwegian Defence Research Establishment.

[80] Wahl, T., Skøelv, Å., Andersen, J. H. S., 1994. Practical use of ERS-1 SAR images in pollution monitoring. *Proc. IGARSS'94*, 4, 1954–1956.

[81] Wahl, T., Skøelv, Å., Pedersen, J. P., Seljelv, L.-G., Andersen, J. H., Follum, O. A., Anderssen, T., Strøm, G. D., Bern, T.-I., Espedal, H. H., Hamnes, H., Solberg, R., 1996. Radar satellites: A new tool for pollution monitoring in costal waters. *Costal Management*, 24, 61–71.

[82] Wu, S. Y., Liu, A. K., 2003. Towards an automated ocean feature detection, extraction and classification scheme for SAR imagery. *Int. J. Remote Sensing*, 24 (5), 935–951.

[83] Zhifu, S., Kai, Z., Baojiang, L., Futao, L., 2002. Oil-spill monitoring using microwave radiometer. *Proc. IGARSS'02*, 5, 2980–2982.

Chapter 11

Paper II: Oil Spill Detection in Radarsat and Envisat SAR images

Anne H. S. Solberg, Camilla Brekke and Per Ove Husøy

Published in: *IEEE Transactions on Geoscience and Remote Sensing*,
March 2007, vol. 45, no. 3, pp.746-755.

Oil Spill Detection in Radarsat and Envisat SAR Images

Anne H. S. Solberg
is with Department of Informatics, University of Oslo, 0314 Oslo, Norway,
and with the Norwegian Computing Center, 0314 Oslo, Norway
(e-mail: anne@ifi.uio.no)
and
Camilla Brekke
is with the Norwegian Defence Research Establishment
and with the Department of Informatics, University of Oslo
and
Per Ove Husøy
was with the Norwegian Computing Center

Abstract

We present algorithms for automatic detection of oil spills in SAR images. The algorithms consist of three main parts; (i) detection of dark spots; (ii) feature extraction from the dark spot candidates; and (iii) classification of dark spots as oil spills or look-alikes. The algorithms have been trained on a large number of Radarsat and Envisat ASAR images. The performance of the algorithm is compared to manual and semi-automatic approaches in a benchmark study using 59 Radarsat and Envisat images. The algorithms can be considered to be a good alternative to manual inspection when large ocean areas are to be inspected.

Index Terms– SAR, oil spill detection, feature extraction, classification.

This work was supported by the EC-project Oceanides and the Norwegian Research Council. The authors would like to thank Marte Indregard, Peter Clayton and Lars Tufte for valuable input on the benchmark comparisons.

1 Introduction

Marine pollution arising from illegal oily discharges from tank cleaning or bilge pumping represent a serious threat to the marine environment. Discharges from ballast water or tank cleaning concerns mainly tankers, while engine room effluent discharges concern all types of ships. Due to such operations, large amounts of oil are deliberately pumped into the marine environment every day. Oil spills correlate well with the major shipping routes, and they also often appear in connection with offshore installations.

The combined use of satellite-based SAR images and aircraft surveillance flights is a cost-effective way to monitor large areas and catch the polluters. Radarsat and Envisat ASAR images enables covering large areas, but aircraft observations are needed to prosecute the polluter, and in certain cases to verify the oil spill.

Oil spills appear as dark areas in the SAR images because the oil dampens the capillary waves of the sea surface. A major part of the oil spill detection problem is to distinguish oil spills from other natural phenomena (look-alikes) that dampen the short waves and create dark patches or on the surface. An oil spill may result in several oil slicks. By the term "oil slick" we will denote a region on the sea surface covered by oil. We distingush between the terms "oil spill" and "oil slick" to be able to describe each dark region in the image that the oil spill consists of.

A review of algorithms for automatic detection of oil spills can be found in [1]. Several of the papers [2; 3; 4] describe a methodology consisting of dark spot detection followed by feature extraction and classification. Solberg et al. [4] apply an adaptive thresholding algorithm for dark spot segmentation. Kanaa et al. [5] use hysteresis thresholding. An edge detection approach based on the Laplace of Gaussians or Difference of Gaussians is presented in [6; 7]. The use of wavelets for ocean feature detection is described by [8; 9]. QinetiQ's dark spot algorithm uses a Constant False Alarm Rate (CFAR) algorithm followed by clustering and Hough transform to identify linear targets [10]. Barni et al. [11] propose an algorithm based on fuzzy clustering. A method using mathematical morphology for oil spill segmentation is presented by Gasull et al. [12]. Even though a variety of methods are applied, the common approach is to detect all suspicious slicks and to preserve slick shapes.

In [1] a discussion of previous approaches to oil spill feature extraction is included. Most of the features used for slick feature extraction in the literature are typically covered by the following types: the geometry and shape of the segmented region; physical characteristics of the backscatter level of the spot and its surroundings; contextual features describing the slick in relation to its surroundings; and textural features.

Various classifiers have been applied to classify a slick as oil or look-alike. Fiscella et al. [2] applied a Mahalanobis classifier. Solberg et al. [4] combined a statistical classifier using a Gaussian model and several subclasses with a rule-based modification of prior probabilities to take into account feature combinations that are indications of certain scene conditions. They tested their method on 11 oil spills and 6 look-alikes. A neural-network approach is described by Del Frate et al. [3]. Their data set consisted of 71 oil slicks and 68 look-alikes,

and they used leave-on-out for error estimation. A recent study based on fuzzy logic is given by [13], who used 9 images for training and 26 images for testing. A comparison of these different oil spill detection approaches in terms of classification accuracy is difficult. They use different data sets and different classifiers, so the reported classification accuracy cannot be directly compared. For an automatic algorithm that should operate under all wind conditions, including low wind where a high number of look-alikes can be expected, the false alarm ratio becomes very important. By using e.g. loss functions, the number of oil spills classified as look-alikes can be lowered, but at the cost of a higher false alarm ratio. In this paper we test the performance of the algorithm on a test set of many thousand look-alikes.

Comparing the results of the automatic algorithm to manual inspection by a trained operator is very interesting. A significant contribution of this paper is a benchmark study involving 59 Radarsat and ENVISAT SAR images that compares the algorithm to the manual detection done as part of the operational oil spill detection service at Kongsberg Satellite Station in Tromsø, Norway, and aircraft verification.

In this paper, we present a methodology for automatic oil spill detection in Radarsat and Envisat SAR images. The algorithms are trained on 70-80 images from each sensor, and they are benchmarked against aircraft verifications and manual detections.

2 SAR Imaging of Oil Spills

A SAR sensor is the most efficient satellite sensor for oil spill detection, though it does not have capabilities for oil spill thickness estimation and oil type recognition. A discussion of other sensors used for oil spill imaging is given in [1]. Oil spills dampen the Bragg waves on the ocean surface and reduce the radar backscatter coefficient, thus creating dark spots in the SAR image. Oil spill look-alikes are phenomena that result in dark areas in the SAR image. Oil spills may include all oil-related surface films caused by oil spills from oil rigs, leaking pipelines, passing vessels as well as bottom seepages, while look-alikes include natural films/slicks, algae, grease ice, threshold wind speed areas, wind sheltering by land, rain cells, shear zones, internal waves etc. [14]. A trained human interpreter is able to discriminate between oil spills and look-alikes based on experience and prior information concerning location, differences in shape and contrast, and weather information. Knowledge about ship lanes and oil rig positions can be extracted from a database, but ship and oil rig positions can also be derived from the SAR image (ships of a certain minimum size). Information about the wind level is very important to consider for oil spill classification. The wind level influences the backscatter level and the visibility of oil spills and look-alikes. The visibility of slicks in SAR images is discussed in [15; 16]. In wind speeds less than 3-5 m/s the probability of observing look-alikes is high. Wind speeds from 5-10 m/s produce fewer dark spots from local low-wind areas. As the wind increases, the contrast between the oil and the surrounding sea will decrease. In high wind, only thick oil will be visible, and look-alikes are rare. The upper limit for observing oil in the SAR image is not known

exactly. In an operational oil spill detection service at Kongsberg Satellite Station in Tromsø, an upper limit of 15 m/s is used.

3 The Oil Spill Detection Approach

An overview of the oil spill detection approach is given in Figure 1.1. Pre-processing consists of land masking the images by converting a land mask to the SAR image grid to avoid re-sampling the speckle pattern, followed by range-dependent scaling and conversion to radar brightness values (beta naught values). The land mask is obtained from a shoreline database [17].

The radar backscatter from the ocean depends strongly on the incidence angle, yielding a much stronger signal at low incidence angles. The incidence dependency depends on the wavelength, polarization, and wind and weather conditions on the ocean surface. Instead of trying to estimate and compensate for the weather dependent effects, we have opened for the simple solution of using a standard incidence angle compensation for all weather conditions. Two different correction functions were identified in a small experimental study, one for Radarsat and one for ASAR. For ASAR, a $\tan^4(\alpha)$ term is used, where α is the incidence angle, while for Radarsat a $\tan^2(\alpha)$ is sufficient to yield a visually satisfactory range intensity for the relevant scenes.

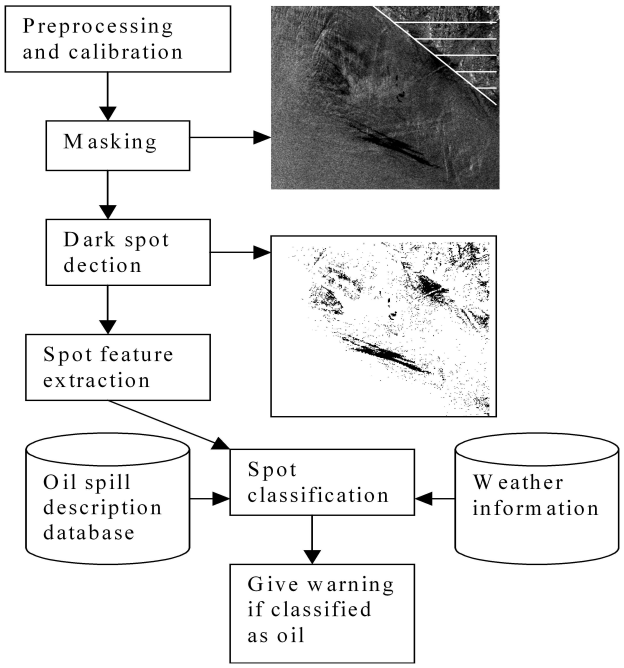


Figure 1.1: Overview of the oil spill detection approach.

3.1 Dark Spot Detection and Segmentation

The first step in the oil detection algorithm is to segment out oil spill candidates. We have chosen to segment out all dark regions in the image, and later discriminate between oil spills and look-alikes based on features computed for each region. This approach can result in a large number of dark areas being segmented, particularly for low-wind conditions where look-alikes are frequently observed.

Sensor-specific modules for dark spot detection have been developed for Radarsat ScanSAR and Envisat Wide Swath ASAR images. Sensor-specific modules were necessary because of different spatial resolution and contrast. Radarsat ScanSAR Narrow full resolution images (HH polarization) have a resolution of approximately 50 m and a pixel spacing of 25 m. Envisat Wide Swath images (VV polarization) have a resolution of approximately 150 m and a pixel spacing of 75 m.

Initial segmentation experiments with the algorithm developed for ERS images [4] showed that a single-scale approach did not perform well in segmenting both large and small regions. Thus, a multiscale approach was developed.

First, an image pyramid is created by averaging pixels in the original image. From the original image, the next level in the pyramid is created with half the pixel size of the original image, and so on. Adaptive thresholding is then applied to each level in the pyramid.

The following adaptive thresholding algorithm is used to segment each level in the pyramid:

I. For each pixel i, compute the mean value (μ) and the power-to-mean (PMR) value (σ/μ) in a local window of size W (where σ is the standard deviation).

II. Compute *homogeneity category* k from PMR as described in Table 1.1.

III. Get threshold value ΔdB_k (in dB) given *homogeneity category* from Table 1.1.

IV. Set $T_i = \mu - \Delta dB_k$.

V. Threshold pixel i with the computed value of T_i.

The threshold is thus set adaptively based on estimates of the roughness of the surrounding sea. The motivation for this is the following: In low wind with many look-alikes, the Power-to-mean ratio (PMR) will be high, and a high contrast between slicks and their surroundings can be expected. As the wind increases, the PMR value will decrease, and the expected contrast between the oil and the surrounding sea will also decrease.

The parameter values in Table 1.1 were found by experiments on the training data set (see Section 3.3).

It is possible to replace the homogeneity category with wind predictions or estimated wind levels. For the data set we worked with, no wind predictions were available, and we did not have access to an algorithm for wind speed estimation based on the SAR image.

Table 1.1: Parameters values for thresholding

PMR intervals	Homogeneity category (k)	Thresholds ΔdB_k in dB (Envisat)	Thresholds in dB (Radarsat)
PMR > 0.15	1	4.0	4.0
PMR $\in [0.05, 0.15]$	2	2.7	3.0
PMR $\in [0.04, 0.05]$	3	2.4	2.0
PMR $\in [0.03, 0.04]$	4	1.5	1.8
PMR $\in [0.015, 0.03]$	5	1.3	1.5
PMR < 0.015	6	1.0	1.0

If μ and the PMR value is recomputed for every single pixel, this procedure is fairly slow for a 400×400 km^2 image (Envisat WS). The segmentation step is the most computationally intensive step of the algorithm. To speed up the processing, two skip factors are introduced. Skip factor s_1 is used for thresholding, the adaptive window is moved s_1 pixels for each threshold computation, while skip factor s_2 is used when computing the mean and PMR value inside the window (only every s' pixel is included in the mean and variance computations). The following parameter values are used: $W = 121$, $s_1 = 5$, and $s_2 = 11$.

After segmenting each level in the pyramid, the different levels are merged using a very simple procedure. All segmented pixels in the highest resolution level are kept, and the lower resolution levels are used to fill holes in the segmentation result using a logical operator. The Radarsat segmentation module uses three pyramid levels, while the Envisat module uses two levels.

The segmented image will contain all oil spills, but also a high number of look-alikes. The segmentation step works well in general, but in some cases thin linear slicks can be fragmented. Fig. 1.2 shows an example of an image and the corresponding segmentation result.

3.2 Dark Spot Feature Extraction

After segmentation, "'region objects"' are formed for all objects in the segmented image. Small objects (< 20 pixels) are not further processed. The subsequent processing considers only region objects. For each region, a set of features is computed. These features are later used to classify each slick as oil or look-alike.

The feature set consists of a mix of features specially developed for oil spill detection and traditional descriptors from the image analysis literature.

The features can be grouped into descriptors for *shape*, *contrast*, *slick surroundings*, and *slick homogeneity*. We previously performed a study regarding oil spill features in ERS

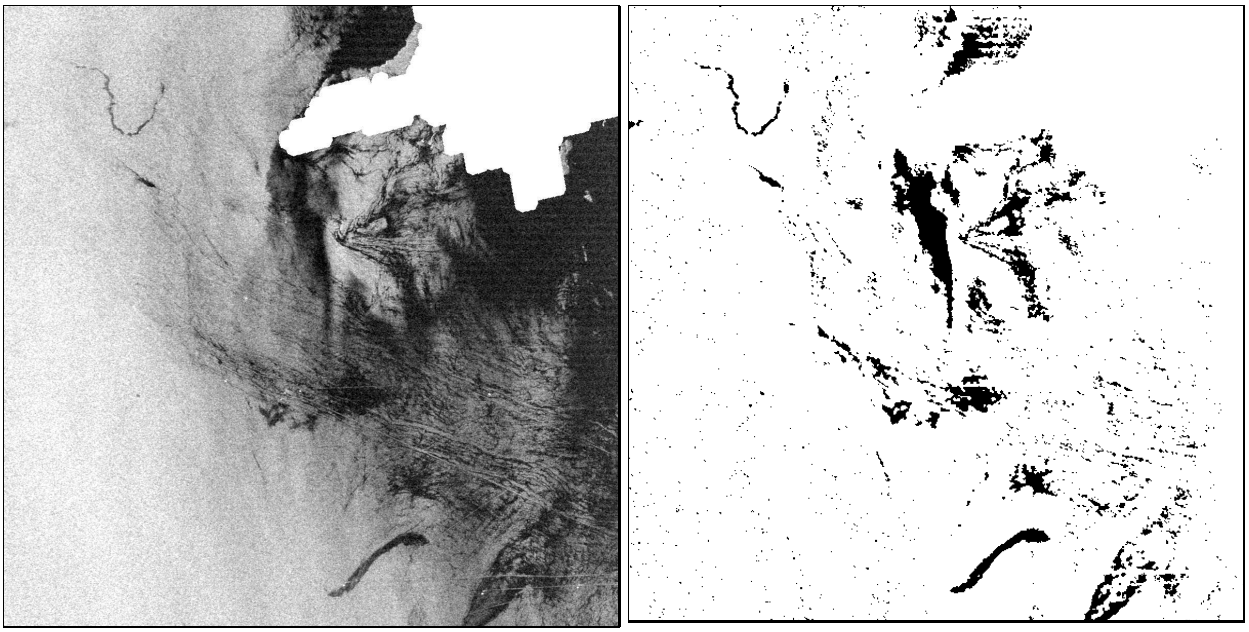


Figure 1.2: Part of the Envisat image from July 21, 2003, and the corresponding segmentation result. ©ESA/KSAT/NR

images [4] and the features used here are mainly from [4], but some of them have been modified.

The following features are used:

Shape features

- *Slick complexity* is defined by $C = P^2/A$, where P is the perimeter and A is area of the region.
- *Slick width* is the ratio between the area of the region and the width of the branches of the skeleton of the region.
- *Slick area* is the size (in number of pixels) of the region.
- *Slick moment* is the first invariant planar moment [18], defined as $\phi_1 = \eta_{20} + \eta_{02}$, where η_{pq} are normalised central moments: $\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}}$, and μ_{pq} are the non-normalised central moments.

Contrast features

- *Slick local contrast* is the difference between the mean value of the slick and the mean value of a larger window surrounding the slick.
- *Border gradient* is the mean of the magnitude of gradient values of the region border area. The Sobel operator is used to compute the gradient.

- *Smoothness contrast* ($\frac{N_r/G_r}{N_b/G_b}$) is defined as the ratio between the number of region pixels N_r and the sum of the region pixel gradient values G_r, and the ratio of the number of background pixels N_b and the sum of the background gradient values G_b .

Homogeneity features

- *Surrounding PMR* is computed as the PMR ratio (σ/μ) in a large window (containing slick surroundings).
- *Slick PMR* is computed as the PMR ratio of the slick.

Slick surroundings

- *Number of detected spots in the scene.*
- *Number of neighboring spots in a large window centered at the region.*
- *Number of neighboring spots in a small window centered at the region.*
- *Distance to ship/rig* A simple point detector to detect ships or oil rigs has been implemented. To detect point targets, pixel cliques (a clique consists of two or more neighboring pixels) within a center window of size 3×3 is compared to the mean value of a larger window (size 15×15). If the difference between the clique with the largest mean value and the mean of the large window is larger than 7 dB, the center location is said to be a point target. From every region from spot detection, the distance to the closest point target is computed.

3.3 Dark Spot Classification

In designing a classifier to discriminate between oil spills and look-alikes, several factors are important to consider. One factor is the expected number of oil spills compared to the expected number of look-alikes. The number of look-alikes depends on the wind speed (and also location). With the given segmentation approach we sometimes observe several hundred look-alikes in a single scene during low-wind conditions. In high wind, only a few look-alikes are expected. The expected number of oil spills does not depend strongly on wind speed. In low wind, even very thin oil films are visible in the SAR image, whereas in high wind, only thicker oil films are visible. The expected number of oil spills in a scene might also depend on the location. Oil spills are commonly seen near oil rigs or major shipping lanes. Fig. 1.3 shows the number of reported oil spills in the Baltic sea during 1990-2002. Note that this information is not normalized in terms on monitoring frequency, so it cannot be directly used to calculate the yearly expected number of oil spills at a given location. Studying the spatial distribution and temporal trends for oil spills is beyond the scope of this paper, for further information see [19].

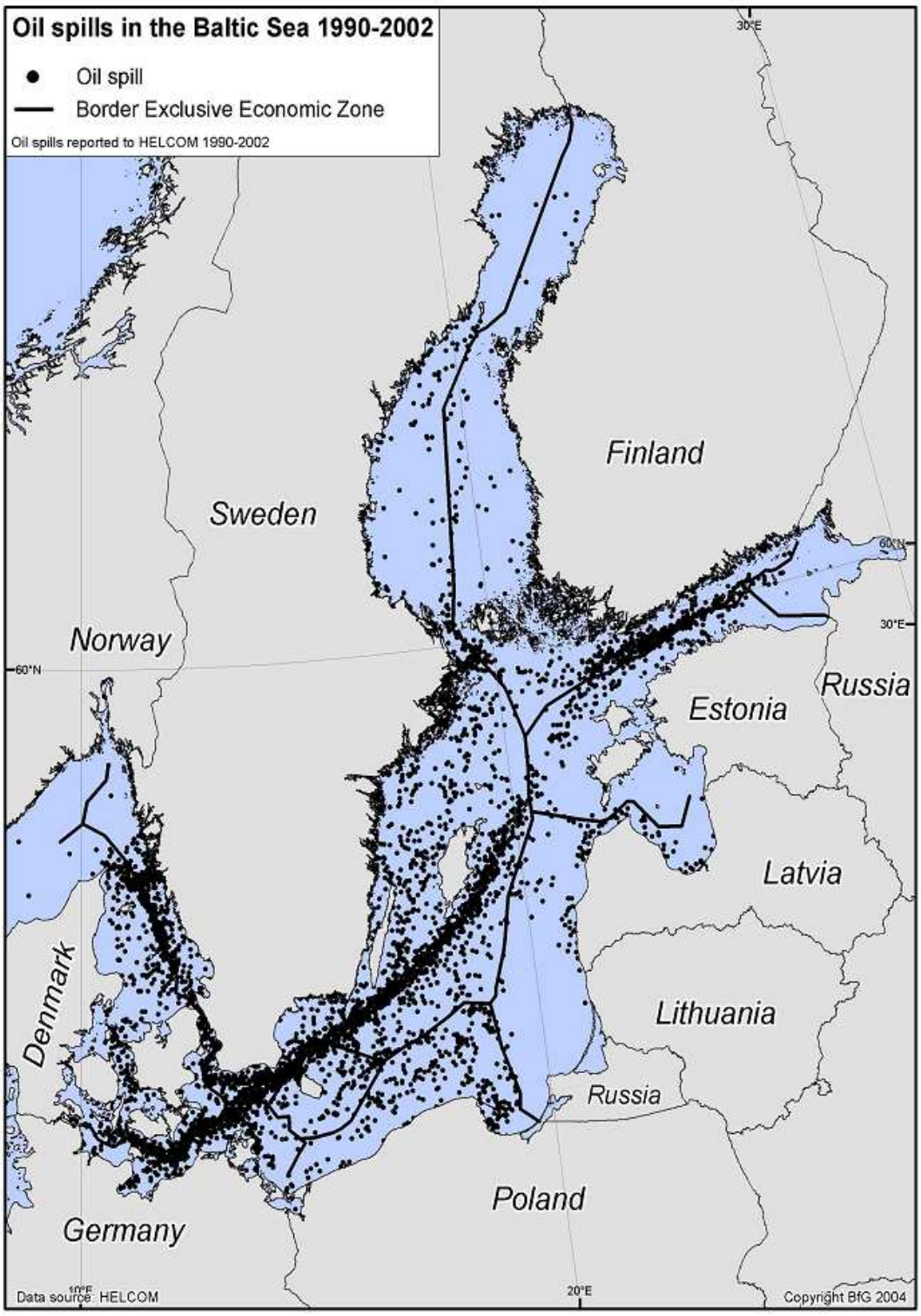


Figure 1.3: Oil spills In the Baltic Sea 1990-2002 reported to HELCOM. (Copyright BfG 2004)

The difference between the expected number of oil spills and the number of look-alikes will influence the estimates of the corresponding class-conditional probability distributions. A training set consisting of e.g 100 images is likely to contain thousands of look-alikes and about 100 oil spills. We need to consider this when estimating the variance of the distributions.

Basic Classifier

In the following we describe our basic classifier for oil spill classification. It is given in [4], but it is included here for easier reference. For each of the detected dark spots, a number of features are computed and collected in the feature vector $\mathbf{x}_i$. These features are constructed such that they typically will be different depending on if the dark spot is an oil spill or a look-alike. However, the behavior of these features will change with different wind levels. Denote the wind level by w.

A prior distribution and a probability density for the features are combined through Bayes theorem to obtain the *posterior* probability for a detected spot being an oil spill. Let c be the unknown class membership of a detected spot. Then

$$\begin{aligned}\Pr(c = o|\mathbf{x}_i, w) =& \frac{\pi_o(w) f_{o,w}(\mathbf{x}_i)}{\pi_o(w) f_{o,w}(\mathbf{x}_i) + (1 - \pi_o(w)) f_{l,w}(\mathbf{x}_i)} \\ =& \frac{\pi_o(w) \frac{f_{o,w}(\mathbf{x}_i)}{f_{l,w}(\mathbf{x}_i)}}{\pi_o(w) \frac{f_{o,w}(\mathbf{x}_i)}{f_{l,w}(\mathbf{x}_i)} + (1 - \pi_o(w))}. \end{aligned} \tag{1.1}$$

$\pi_o(w)$ and $\pi_l(w)$ are the prior model for the probability that a detected spot is oil or look-alike given wind level w. $f_{o,w}(\mathbf{x}_i)$ and $f_{l,w}(\mathbf{x}_i)$ are the probability densities for the observed features $\mathbf{x}_i$ in classes $o =$ oil spills and $l =$ look-alikes, respectively.

In the following paragraphs, we describe how the class-conditional proability densities are computed, the use of prior models for the expected number of oil spills and look-alikes in an image, the use of loss functions to model false positives and false alarms, and the use of rule-based corrections of the prior probabilities.

Specification of Class-conditional Probability Densities The simplest choice of densities is to use multivariate Gaussian ones. Even within each wind level both the oil spills and the look-alikes may vary quite a lot in shape and other features. Describing the feature density by a unimodal density such as the Gaussian is therefore not appropriate. Instead we have assumed different densities depending on the value of a shape descriptor s (we have used the feature *slick moment*). The sample space of s was divided into four subgroups ranging from wide slicks with regular shape to wide slicks with a complex boundary to thin, linear spots. The density for class c and wind level w is then given by $f_{c,w}(\mathbf{x}_i) = f_{c,w,g}(\mathbf{x}_i)$ if s is in subgroup g. This hierarchical structure is illustrated in Fig. 1.4. Finally, the densities

within each subgroup are assumed Gaussian:

$$f_{c,w,g}(\mathbf{x}_i) = \frac{1}{(2\pi)^{\frac{d}{2}} \mid \Sigma_{w,g} \mid^{\frac{1}{2}}} \times \exp\{-\frac{1}{2}(\mathbf{x}_i - \mu_{c,w,g})^T \Sigma_{w,g}^{-1}(\mathbf{x}_i - \mu_{c,w,g})\},$$

where $c \in \{o, l\}$, d is the number of features, $\mu_{c,w,g}$ is the mean vector under wind level w and shape descriptor in subgroup g for class c and $\Sigma_{w,g}$ is a diagonal covariance matrix, common for both classes given (w, g). A diagonal covariance matrix is used because the number of oil spills in each subclass in the training data set is small, typically 10-20. Equal variance for a given feature for a given oil spill subclass and its corresponding look-alike subclass is used because the look-alike classes have many times more observations than the oil spill classes, and experiments showed that using class-conditional covariance estimates resulted in classifying all slicks as look-alikes.

Only the features *slick complexity*, *slick local contrast*, *border gradient*, *smoothness contrast*, *slick PMR*, *slick width* and *number of neighboring spots* are used to compute the probability densities. These features were selected in previous studies [4].

Prior Model for the Number of Slicks Prior distributions for the presence of oil spills can in its simplest form be specified trough the number of oil slicks in the scene. In particular, we will assume that the probability of k oil slicks in the scene is given by

$$\pi(k) = \alpha^k(1 - \alpha) \tag{1.2}$$

where $1/\alpha$ is a parameter describing the expected number of oil slicks in the scene.

Denote by M the number of detected spots. We assume that all the present oil slicks will be detected as dark spots. Given M_o, the number of oil slicks, the stochastic variable is $M_l = M - M_o$, the number of look-alikes that will appear in the scene. We will assume that

$$\Pr(M_l = m | M_o = k, w) = \alpha_w^m(1 - \alpha_w),$$

that is a geometric distribution with the parameter depending on the wind level (given by the dependence on w). In particular, we have assumed that the wind level is divided into 4 categories. These categories are given in Table 1.2 together with their corresponding values of $\frac{1}{1-\alpha_w}$. We use the following as the prior $\pi_o(w)$ for an object being an oil slick:

$$\Pr(c = o | M, w) = \frac{\frac{\alpha}{\alpha_w}}{\frac{\alpha}{\alpha_w} + 1}.$$

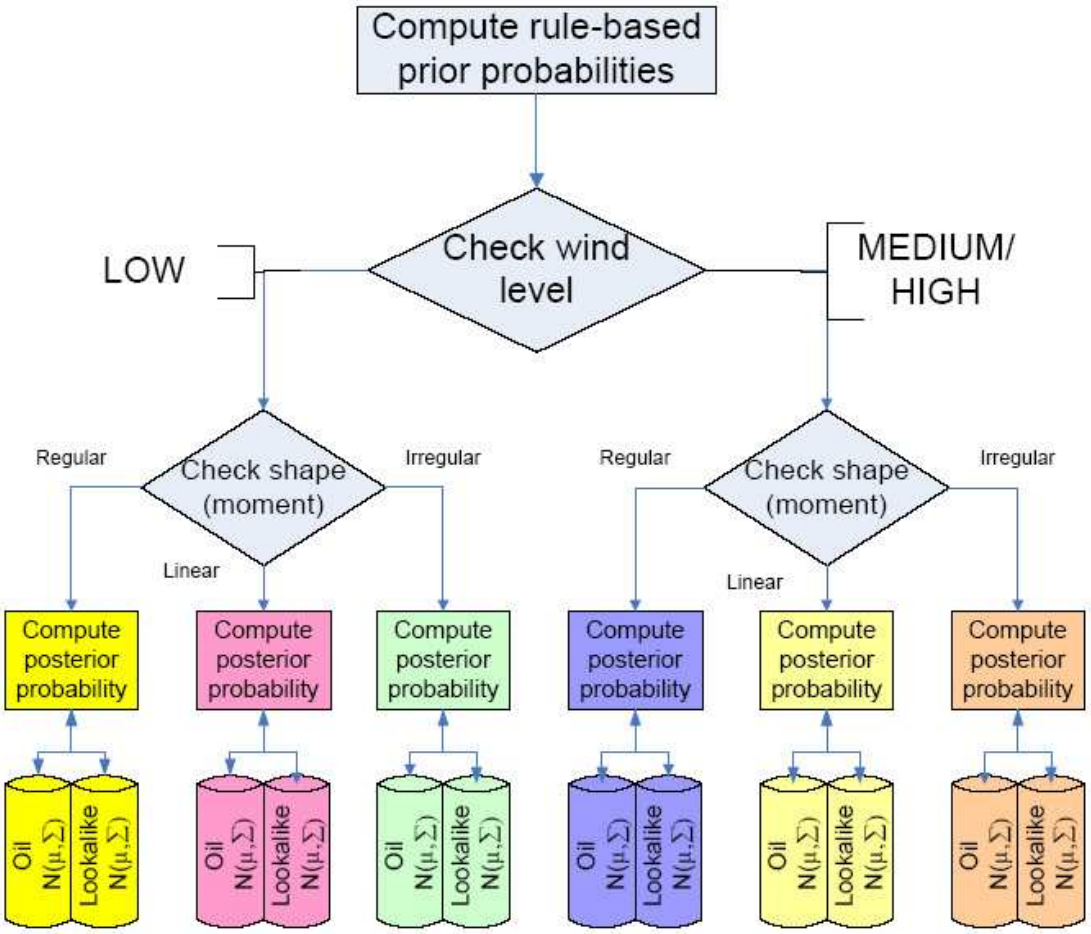


Figure 1.4: Subclass structure based on wind and shape.

Table 1.2: *The expected number of look-alikes is given by* $1/(1-\alpha_w)$ *where* α_w *is the corresponding parameter in the geometric distribution for the number of look-alikes.*

Wind level	Expected no. look-alikes
Low	10000
Low-to-moderate	100
Moderate	20
High	7

Loss Functions We want the classifier to find all spots with a certain probability of being oil. All spots classified as oil are then inspected by an operator. Misclassifying oil as look-

alikes is considered more serious than misclassifying look-alikes as oil. This can be modeled in terms of a simple loss function. Let l_1 be the loss associated with misclassifying a true oil slick as a look-alike and l_2 the loss associated with misclassifying a true look-alike as an oil slick. Then the optimal classification for spot i is given by

$$\hat{c} = \begin{cases} o & \text{if } \Pr(c = o|\mathbf{x}_i, w) > \frac{l_2}{l_1+l_2}; \\ l & \text{otherwise.} \end{cases}$$

Extended Classifier: Rule-based Corrections

The classification performance using the basic classifier was not satisfactory, mainly because a high number of look-alikes were classified as oil. A set of rules adjusting the prior probabilities based on a certain combination of features was developed.

The basic classifier resulted in a high false alarm ratio, which is unacceptable because a scene in low-wind conditions can contain hundreds of look-alikes. To overcome this, we inspected the misclassified slicks (in the training data set) to understand why they were misclassified. What we found was that a human operator in many cases could identify many such slicks as look-alikes based on a combination of features. We then studied the feature values for misclassified slicks in relation to why an expert would label them look-alike. Then a rule for these circumstances was established. The rule specifies a multiplicative factor adjusting the Gaussian densities (or actually the ratios between the densities of oil slick and look-alikes). Establishing these rules was a tedious process for the first images, but after performing it for 10-20 images the procedure tended to converge and the number of false alarms was greatly reduced.

The procedure resulted in 37 rules for Radarsat imagery. Examples of the rules used are given in Table 1.3. Three of these were used to set prior probabilities for regions close to point sources (using the DIST feature). Most of the rules (28) were used to reduce false alarms, and 8 rules were used for false negatives. For Envisat, the rule set was extended with 13 rules tailored to Envisat statistitics. Generally, we observed more look-alikes in Envisat images than Radarsat images (as could be expected due to the higher sensitivity of Envisat's VV polarization), making additional rules necessary.

4 Experimental Results

The oil spill detection system is trained and tested on Envisat ASAR and Radarsat SAR images. A set of 56 Envisat ASAR Wide Swath images and 71 Radarsat SAR images from 2003 and 2004 from European waters (mainly the Baltic and the North Sea) was used to train the algorithm. All the SAR images were processed by Kongsberg Satellite Services (KSAT).

For some of the images, the reports made by KSAT as part of their oil spill detection service were available. To label all of the images in the training set, we created a training

Table 1.3: *Examples of rules used to compute the rule-based modifications of the prior probabilities. The feature abbreviations are: DIST = Distance to ship/rig, NLN = Number of neighboring spots in a large window, AREA = Slick Area, COMPLEXITY = Slick Complexity, PMRSURR = Surrounding PMR, LCONT = Slick local contrast, NSN = Number of neighboring spots in a small window, MOM = Slick Moment*

Increase prior probability for oil spills if
$DIST < 20$ AND $NLN = 0$ AND $AREA > 20$
$DIST < 10$ AND $NLN = 0$ AND $COMPLEXITY < 20$
Increase prior probability for look-alikes if
$AREA < 20$ AND $PMRSURR > 0.05$
$NSN > 5$ AND $PMRSURR > 0.05$ AND $LCONT < 3.0$
$WIND = LOW$ AND $NLN > 5$ AND $MOM < 1$

mask for each SAR image. KSAT reports were used when available. We scanned through all images labeling all possible oil spill candidates as either *oil spill* or *doubt*. Doubt cases were used because in many cases a trained operator cannot discriminate perfectly between oil slicks and look-alikes based on the SAR image alone. All slicks not marked as oil or doubt were used to train the *look-alike* class (thus the slicks marked *doubt* were left out of the look-alike class).

We first studied the performance of the oil spill detection algorithm compared to the test masks made by manual inspection. The Envisat test set of 27 images contained 37 slicks that were manually labeled as oil spills, in addition to 12110 look-alikes. This set of images will also be used for benchmarking against other approaches in the following subsection. Table 1.4 summarizes the classification accuracy for the Envisat test data set. To study the effect of the rule-based adjustments of the prior probabilities, the performance is reported with and without the rules. Without the rule-based corrections, the average classification accuracy is 89% for oil spills and 71% for look-alikes. However, since the total number of look-alikes is 12110, 3504 false alarms occur. This is not acceptable. With the rule-based corrections, the average classification accuracy is 78% for oil spills and 99.4% for look-alikes. This result in an acceptable false alarm ratio.

Benchmarking Oil Spill Detection Systems

As part of the EC project Oceanides, a joint satellite-airborne campaign was performed during 2003 to establish a data set consisting of SAR images with associated aircraft verifications by the German and Finnish pollution control authorities. The campaign covered the Finnish and German sectors of the Baltic Sea, in addition to the German sector of the North Sea. A total of 59 Radarsat and Envisat images were acquired between July and December 2003.

Table 1.4: Classification accuracies on the Envisat test set with and without the rule-based confidence levels.

Basic classifier, no rules		
	Classified as oil	Classified as look-alike
Marked as oil	33 (89.2)%	4 (10.8%)
Marked as look-alike	3504 (28.9%)	8613 (71.1%)
Advanced classifier with rules		
	Classified as oil	Classified as look-alike
Marked as oil	29 (78.4%)	8 (21.6%)
Marked as look-alike	77 (0.7%)	12033 (99.4%)

This campaign was organized in such a manner that KSAT downloaded the satellite images, analyzed them in near-real time, and reported possible oil spills to the Finnish and German pollution control authorities, which would check the reported locations and determine if the reported slicks were oil or look-alikes.

The Radarsat benchmark data set consisted of 32 Radarsat images, and the Envisat benchmark data set of 27 Envisat images. The details of this benchmarking can be found in [10]. The images were analyzed by KSAT, by a semi-automatic algorithm developed at QinetiQ, and by our automatic algorithm. For the benchmark comparison, KSAT let another operator analyze the images without knowing the results of the previously reported oil slicks. The results from this operator are termed KSAT-B and are used in the benchmark tables.

The benchmarking was done without any of the persons/algorithms knowing the aircraft detections, so that none of the approaches could be tuned to the aircraft results. The joint aircraft-satellite campaign was organized in such a manner that KSAT downloaded the satellite images, analyzed them in near-real time, and reported possible oil spills to the Finnish and German pollution control authorities, which would check the reported locations. Aircraft verification of all slicks was not possible, as many of the slicks were outside the Finnish or German territories (meaning that the aircraft was not allowed to fly to the location), or because flight plans or satellite acquisitions sometimes had to be changed.

The performance in detecting verified oil slicks is given in Table 1.5. The Radarsat data set contained 18 oil slicks verified by the aircraft. 15 of these slicks were found by the KSAT-B operator, the algorithm described in this paper found 14, while the QinetiQ approach found 12. The Envisat data set contained 11 oil slicks verified by the aircraft. 8 of these slicks were found by the KSAT operator, 8 by the algorithm, and Qinetiq's oil spill system found 5 of these. Fig. 1.5 shows examples of verified oil slicks which were correctly classified.

A very interesting scene is from July 21, 2003. For this date, the same area in the Finnish sector of the Baltic sea was covered by both Envisat and Radarsat images. Parts of this scene is shown in Figure 1.6. The Radarsat image was taken at 16.05 and the Envisat image

Table 1.5: Performance in detecting verified oil slicks

	Envisat data set	Radarsat data set
Slicks verified by aircraft	11	18
Detected by operator KSAT-B	8 (72%)	15 (83%)
Detected by algorithm	8 (72%)	14 (77%)
Detected by QinetiQ	5 (45%)	12 (66%)

at 19.45. In the Radarsat image, the leftmost slick was reported as oil by KSAT, QinetiQ and the algorithm, while the rightmost slick was not reported. Both these slicks have been verified as oil. In the Envisat image, the rightmost slick was reported by KSAT, the algorithm and QinetiQ, while the leftmost slick was not reported. The leftmost slick appears larger and clearer in the Radarsat image. It is possible that the oil has partly evaporated during this time. The rightmost slick appears with good contrast and separated from the surrounding low-wind pattern in the Envisat image, but in the Radarsat image the contrast is lower, and it can be mistaken to be part of the low-wind pattern. Note also that in the Envisat image a ship or oil rig can be seen adjacent to the slicks. However, this is probably not the source of the slick as the slick was also visible on the Radarsat image taken 3.5 hours earlier.

In addition to the capabilities of detecting the verified oil slicks, we can also study the slicks reported by satellite, but verified as not oil by the aircraft. For ENVISAT, 9 slicks were detected by the algorithm, the location checked by the aircraft, but found not to contain any oil. 6 of these 9 slicks (all from the Baltic Sea) were verified as algae. Algae is a problem in the Baltic Sea because the type of algae observed here will dampen the capillary waves and result in a dark spot in the SAR image. Additional information about algal activity is desirable in the Baltic during the summer. For 2 of the 9 false alarms, nothing was visible when the aircraft checked the location a few hours later. In this case, the slicks were small, and the oil might have disappeared in the time between satellite image aquisition and aircraft inspection. An example of a slick verified as algae is given in Figure 1.7.

We can also try to compare the number of alarms for the automatic algorithm compared to the operator. For the Radarsat data set, the operator produced 75 alarms, while the algorithm produced 71 alarms. Thus, they are expected to be comparable in terms of false alarm ratio. (Note here that all additional slicks reported by the algorithm could not be verified because the images were not processed in real-time.) For the Envisat data set, the operator reported 69 slicks, while the algorithm reported 112 slicks.

In judging the performance of an automatic algorithm compared to manual detection, it is also intersting to study the variance in detections done by different operators at KSAT. In the benchmark study, two different operators (called KSAT-1 and KSAT-2) inspected the same SAR images (without knowing the results of the other inspections). For Radarsat, the two operators at KSAT detected 75 and 68 oil slick candidates, while the algorithm

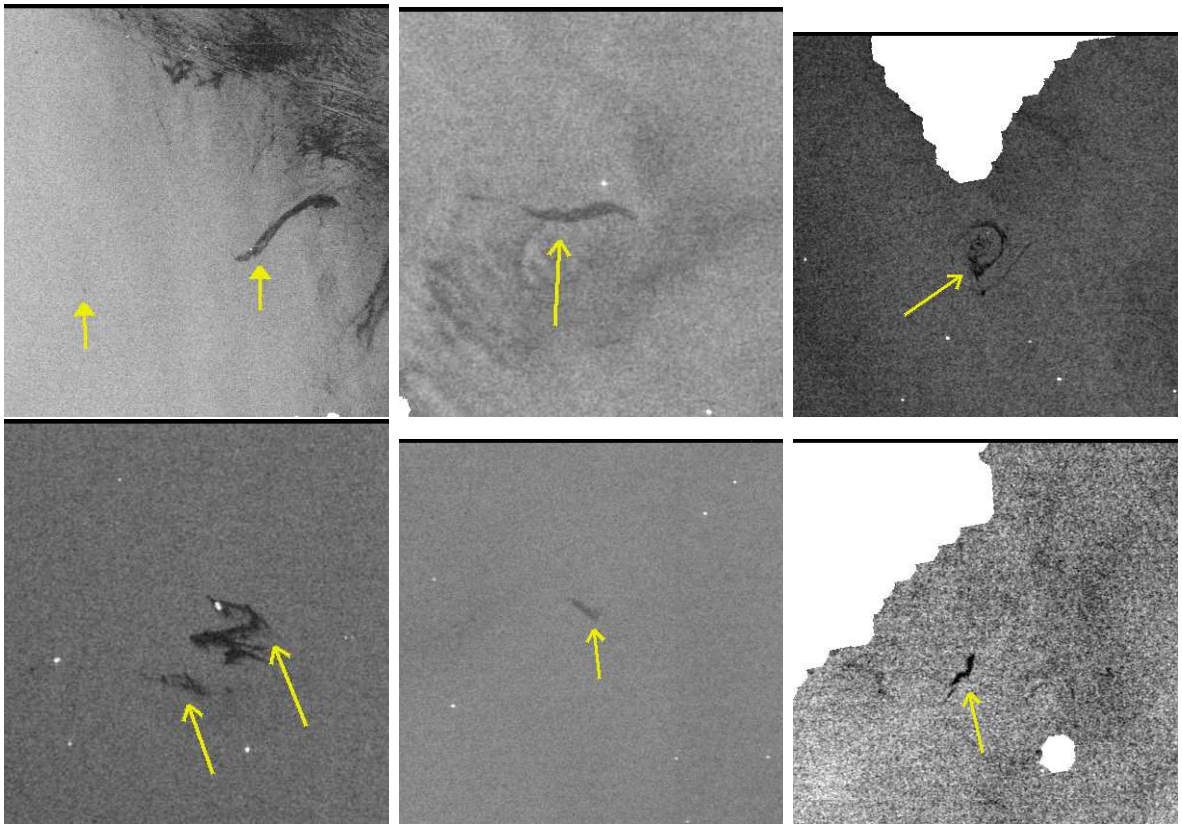


Figure 1.5: Envisat images with examples of verified oil slicks correctly classified as oil. ©ESA/KSAT/NR

detected 71 slick candidates. Note that not all these slicks could be verified by aircraft, as only KSAT-1 inspected the images in real time. Of the 18 slicks detected by KSAT-1 and verified, KSAT-2 detected 15 and the algorithm 14. For ENVISAT, KSAT-1 detected 11 slicks that were verified by aircraft. KSAT-2 detected 8 of these slicks, and the algorithm also detected 8 slicks.

A question that arises is why some of the verified oil slicks were not detected by any of the approaches. Fig. 1.8 shows verified oil slicks from Radarsat images that were not detected. The two slicks in the upper row are low-contrast slicks in a locally homogeneous background, while the scene does contain many low-wind areas with many look-alikes. We expect that the algorithm did not classify these two slicks as oil because the wind estimate was too low because the surroundings (in a larger local window) contained low-wind areas. Improving the local wind estimate will help this. The slicks in the second row of Fig. 1.8 are slicks in very heterogeneous surroundings with complex look-alike patterns. It is very difficult to discriminate these slicks from the surroundings. Low contrast slicks on a locally homogeneous background was found to be difficult to detect on Radarsat images, but this was

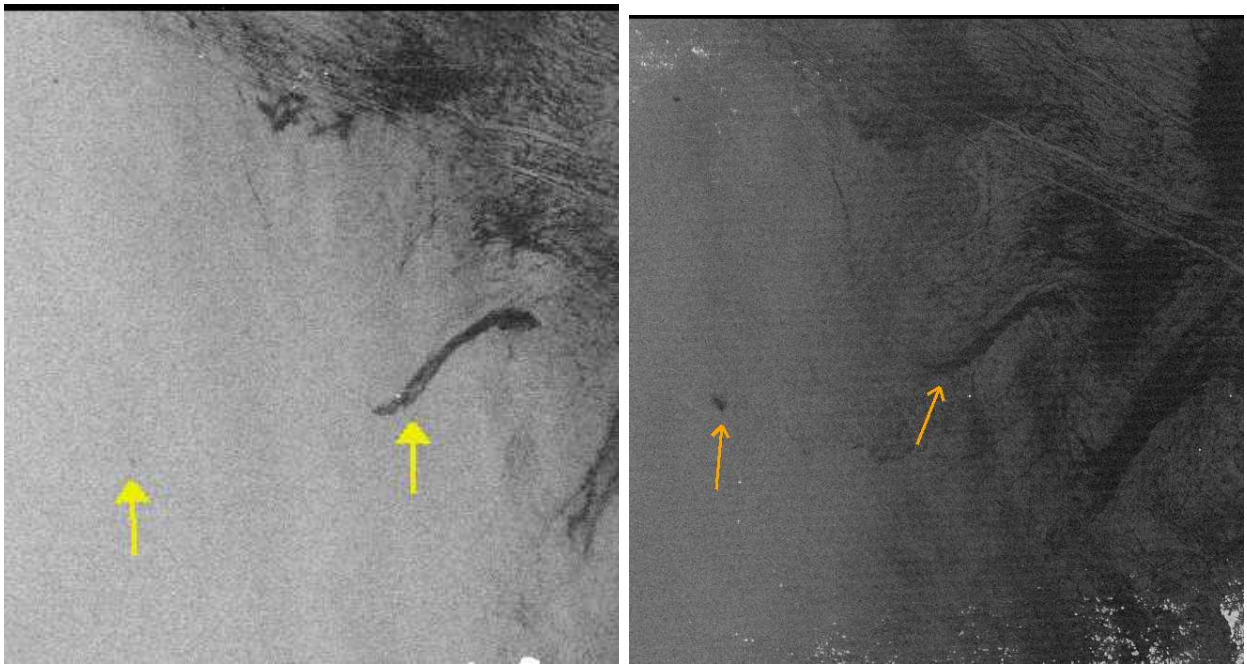


Figure 1.6: Envisat (left) and Radarsat (right) image of the same scene, taken on the same date. On the Envisat image, the leftmost spill was missed, while the rightmost spill was correctly classified as oil. In the Radarsat image, the result was the opposite. ©ESA/Canadian Space Center/KSAT/NR

not a problem for ENVISAT images (perhaps due to the more favourable VV polarization). The slicks not detected on ENVISAT images were slicks that overlapped with dark areas due to low wind (see Fig 1.9 for examples).

For Radarsat images, the processing time for manual inspection at KSAT was between 3 to 25 minutes with an average inspection time of 9 minutes. QinetiQ's semi-automatic approach took 7 to 32 minutes per scene with an average of 20 minutes. The automatic algorithm had an average processing time of 3 minutes per Radarsat image. The processing time for Envisat images for KSAT was 10 minutes per scene in average, and varied between 5 and 23 minutes. QinetiQ used 6-29 minutes with an average of 18 minutes, while the automatic algorithm had an average processing time of 1.45 min on a 2 GHz Linux computer. Note that even if the scene size for Radarsat SCN is smaller that for Envisat (300×300 km compared to 400×400 km), the resolution is higher for Radarsat, thus the processing time is longer.

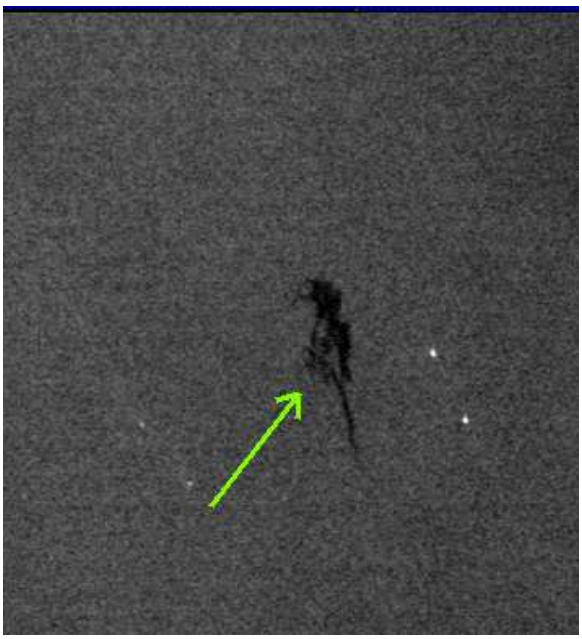


Figure 1.7: Slick from August 7, 2003, verified as algae by the aircraft. ©ESA/KSAT/NR

5 Discussion and Conclusions

In this paper, we have presented a system for automatic detection of oil spills in Radarsat and Envisat images. A major part of the oil spill detection problem is to distinguish oil slicks from other natural phenomena that create dark patches in the SAR image. The oil spill approach can be divided into detection of dark areas, extracting features for all dark areas, and then classifying all spots as oil spills or look-alikes. A multi-scale approach for dark spot segmentation was presented. Sensor-specific modules for Radarsat and Envisat images have been developed. Slick feature extraction was based on features describing slick shape, contrast, homogeneity, and surroundings. The classification approach is a mix between a statistical classifier with subclasses based on wind and shape, and a rule-based approach. The rule-based approach was introduced to reduce the number of false alarms.

In some cases, algae can be misclassified as oil because certain algae types have a similar effect to oil in dampening the radar backscatter signal. Additional information about algal blooms, or the likelihood of observing algae in a given geographical location, should be utilized in the classification step. We also consider to include oil spill hot spot information in terms of a database with previously reported oil slicks and their location.

The dark spot segmentation step works well in most situations, but it can sometimes fragment a thin, linear slick into several shorter segments. We are currently working on improving the segmentation for such slicks. Future work on this topic will also include a thorough evaluation of the use of different classifiers for the dark spot classification step.

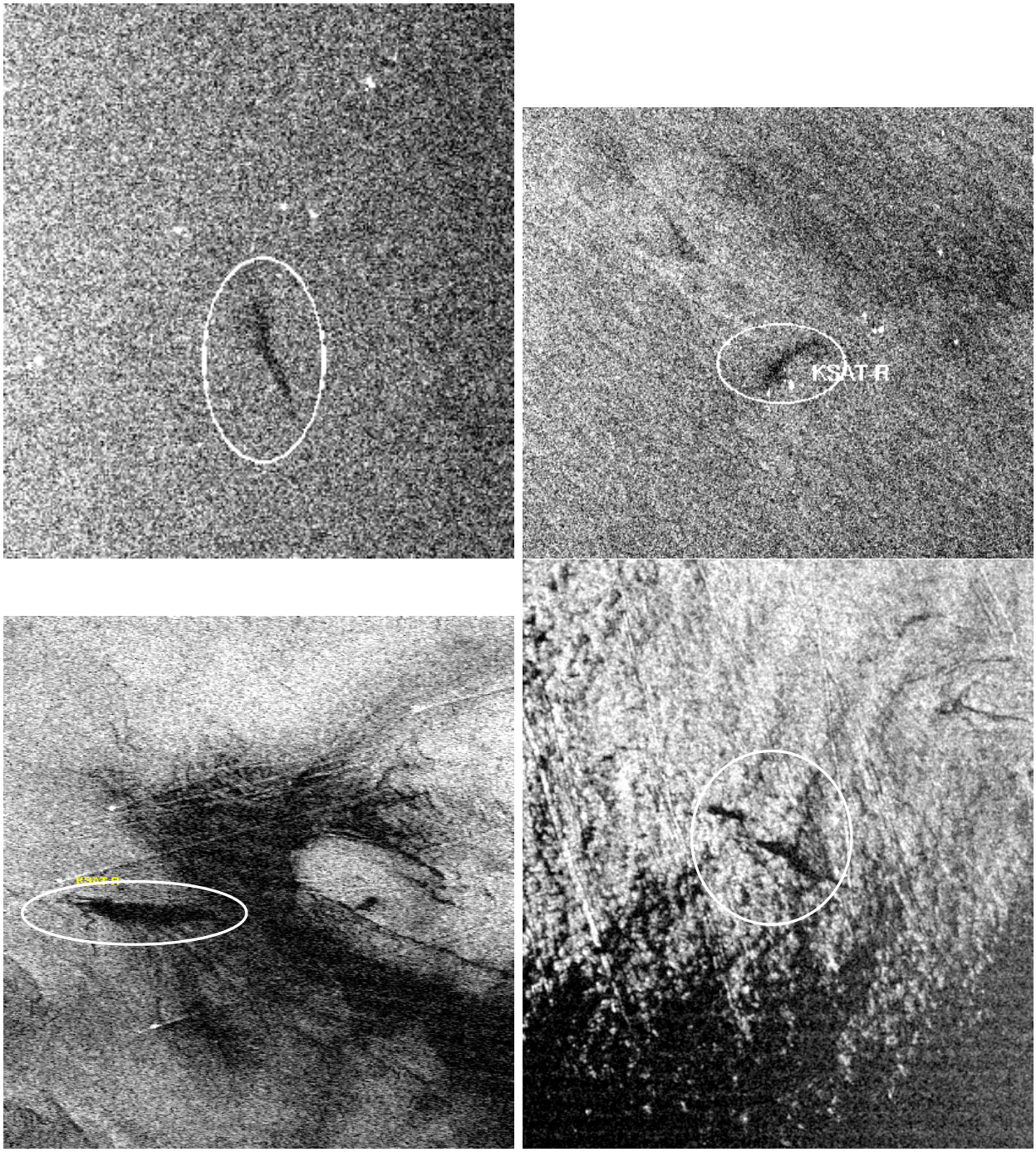


Figure 1.8: Examples of oil slicks verified by the aircraft, but not detected by the operator or the algorithm. ©ESA/KSAT/NR

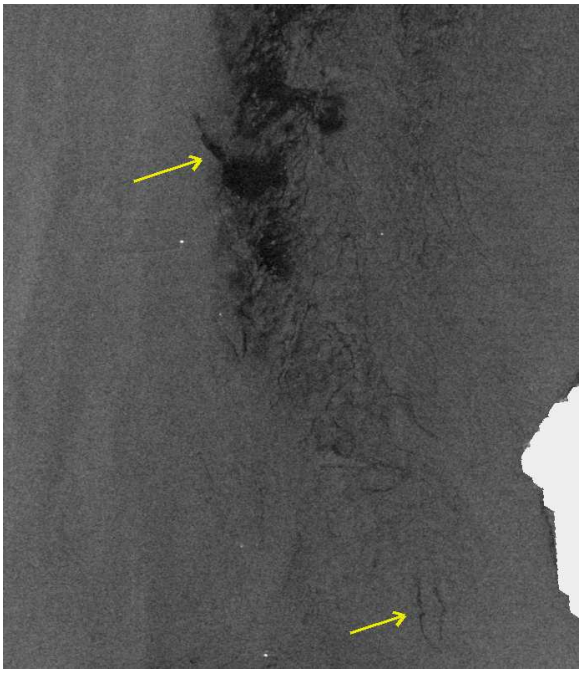


Figure 1.9: Examples of oil slicks from Sept 2., 2003 (ENVISAT), verified by the aircraft, but not detected by the operator or the algorithm. ©ESA/KSAT/NR

The algorithm has been benchmarked against manual oil spill detection and a semi-automatic approach on 59 Radarsat and Envisat images. The benchmark showed that two manual operators do not identify exactly the same oil slicks. The performance of the automatic system is quite comparable to manual inspection, both in terms of performance in detecting verified oil spills, and in terms of the number of alarms. The autmatic algorithm is faster than manual inspection, with an average processing time of 1.45 min compared to 10 minutes for manual inspection.

In determining the best procedure for operational monitoring of oil spills, we find that the algorithm can be a very intesting supplement to manual inspection. Aircrafts are still needed for oil spill verification. We propose to run the algorithm as part of an oil spill detection service by including a manual step prior to sending out the aircraft. This would improve the existing fully manual service, as the benchmark revealed that there is some variance between different operators. Currently, a prototype version of this algorithm has been implemented at KSAT.

References

[1] C. Brekke and A. Solberg, "Oil spill detection by satellite remote sensing," Remote Sensing of Environment, vol. 95, no. 1, pp. –13, 2005.

[2] B. Fiscella, A. Giancaspro, F. Nirchio, P. Pavese, and P. Trivero, "Oil spill detection using marine SAR images," Int. J. Remote Sensing, vol. 21, no. 18, pp. 3561–3566, 2000.

[3] F. Del Frate, A. Petrocchi, J. Lichtenegger, and G. Calabresi, "Neural networks for oil spill detection using ERS-SAR data," IEEE Transactions on Geoscience and Remote Sensing, vol. 38, no. 5, pp. 2282–2287, 2000.

[4] A. H. S. Solberg, G. Storvik, R. Solberg, and E. Volden, "Automatic detection of oil spills in ERS SAR images," IEEE Transactions on Geoscience and Remote Sensing, vol. 37, no. 4, pp. 1916–1924, 1999.

[5] T. F. N. Kanaa, E. Tonye, G. Mercier, V. Onana, J. Ngono, P. Frison, J. Rudant, and R. Garello, "Detection of oil slick signatures in SAR images by fusion of hysteresis thresholding responses," Proc. IGARSS'03, vol. 4, pp. 2750–2752, 2003.

[6] L. Y. Change, K. Chen, C. Chen, and A. Chen, "A multiplayer-multiresolution approach to detection of oil slicks using ERS SAR image," Proc. ACRS 1996 - 17th Asian Conference of Remote Sensing, Sri Lanka, 1996.

[7] C. F. Chen, K. S. Chen, L. Y. Chang, and A. J. Chen, "The use of satellite imagery for monitoring coastal environment in Taiwan," Proc. IGARSS'97, vol. 3, pp. 1424–1426, 1997.

[8] A. K. Liu, C. Y. Peng, and S. Y.-S. Chang, "Wavelet analysis of satellite images for costal watch," IEEE Journal of Oceanic Engineering, vol. 22, no. 1, pp. 9–17, 1997.

[9] S. Y. Wu and A. K. Liu, "Towards an automated ocean feature detection, extraction and classification scheme for SAR imagery," Int. J. Remote Sensing, vol. 24, no. 5, pp. 935–951, 2003.

[10] M. Indregard, A. Solberg, and P. Clayton, "D2-report on benchmarking oil spill recognition approaches and best practice," Oceanides project, European Commission, Archive No. 04-10225-A-Doc, Contract No: EVK2-CT-2003-00177, Tech. Rep., 2004.

[11] M. Barni, M. Betti, and A. Mecocci, "A fuzzy approach to oil spill detection on SAR images," Proc. IGARSS '95, vol. 1, pp. 157–159, 1995.

[12] A. Gasull, X. Fabregas, J. Jimenez, F. Marques, V. Moreno, and M. Herrero, "Oil spills detection in SAR images using mathematical morphology," Proc. EUSIPCO'2002, Toulouse, France, September 2002, vol. 1, pp. 25–28, 2002.

[13] I. Keramitsoglou, C. Cartalis, and C. Kiranoudis, "Automatic identification of oil spills on satellite images," Environmental Modelling & Software, vol. 21, no. 5, pp. 640–652, 2006.

[14] H. A. Espedal and O. M. Johannessen, "Detection of oil spills near offshore installations using synthetic aperture radar (SAR)," Int. J. Remote Sensing, vol. 21, no. 11, pp. 2141–2144, 2000.

[15] T.-I. Bern, S. Moen, T. Wahl, T. Anderssen, R. Olsen, and J. A. Johannessen, "Oil spill detection using satellite based SAR. Completion report for Phase 0 and 1," OCEANOR report no. OCN-R92071, Trondheim, Tech. Rep., 1992.

[16] M. Perez-Marrodan, "ENVISYS - environmental monitoring warning and emergency management system," Proc. of the AFCEA Kiev Seminar, 28-29 May, pp. 122–132, 1998.

[17] P. Wessel and W. H. F. Smith, "A global self-sonsistent, hierarcical, high-resolution shoreline database," J. Geophys. Res,, vol. 101, no. 1, pp. 8741–8743, 1996.

[18] M.-K. Hu, "Visual pattern recognition by moment invariants," IEEE Trans. Inform. Theory, vol. 8, pp. 179–187, 1962.

[19] L. Tufte, "D6(a)-report on oil spill data standardisation, oil spill pattern analysis and hot spots, as well as pollution conditions in test sites including oil spill maps and time series." Oceanides project, European Commission, Archive No. 04-10225-A-Doc, Contract No: EVK2-CT-2003-00177, Tech. Rep., 2004.

Chapter 12

Paper III: Segmentation and Feature Extraction for Oil Spill Detection in ENVISAT ASAR Images

Camilla Brekke and Anne H. S. Solberg

Submitted: *International Journal of Remote Sensing*, 2007.

Segmentation and Feature Extraction for Oil Spill Detection in ENVISAT ASAR Images

Camilla Brekke,
Norwegian Defence Research Establishment,
Postboks 25, 2027 Kjeller, Norway.
Department of Informatics, University of Oslo,
Postboks 1080 Blindern, 0316 Oslo, Norway.
E-mail: Camilla.Brekke@ffi.no.
and
Anne H. S. Solberg,
Department of Informatics, University of Oslo,
Postboks 1080 Blindern, 0316 Oslo, Norway.
E-mail: anne@ifi.uio.no

Abstract

An automatic algorithm processing SAR images for the purpose of detecting illegal oil spill pollution in the marine environment has earlier been developed [28; 31]. The framework consists of modules for dark spot detection, dark spot feature extraction, and a classifier that discriminates between oil slicks and other oceanographic phenomena which resemble oil slicks (*look-alikes*). Based on this framework, we have in this study aimed at identifying new features that lead to significant improvements in classification performance for ENVISAT ASAR Wide Swath Mode images. Both traditional region descriptors, features tailored to oil spill detection, and techniques originally associated with other applications are evaluated. Ahead of feature extraction, a successful segmentation of dark spots on the sea surface is crucial. Improvements done to the segmentation process are discussed as well, and the importance of an adapted segmentation process to this particular problem domain is demonstrated. The novelties of this paper include a new segmentation algorithm for better detection of thin, linear slicks and three new features. The algorithm is evaluated on a set of 83 satellite images, and experimental results are presented. The combination of the presented improvements gave a significant increase in classification accuracy from 78% to 89% in the number

of detected oil spills, and as many as 98% of the look-alikes were correctly classified by our method.

1 Introduction

Oil spills correlate very well with the major shipping routes, and do often appear in connection with offshore installations. When taking into account how frequent illegal discharges from ships appear, controlled regular oil spills can be a much greater threat to the marine environment and the ecosystem than larger oil spill accidents like the Prestige tanker accident in 2002.

Spaceborne Synthetic Aperture Radar (SAR) has proven to be the most efficient satellite sensor for oil spill monitoring of the worlds oceans. Oil spills appear as dark areas in the SAR images because the oil dampens the capillary and short gravity waves on the sea surface. A part of the oil spill detection problem is to distinguish oil slicks from other natural phenomena (*look-alikes*) that dampen the short waves and create dark patches in a similar way. Oil slicks may include all oil related surface films caused by oil spills from oilrigs, leaking pipelines, passing vessels as well as bottom seepages, while look-alikes do include natural films/slicks, grease ice, threshold wind speed areas (wind speed < 3 m/s), wind sheltering by land (like in the lee of an island), rain cells, shear zones, internal waves, etc. [10]. These ambiguities put a challenge on the selection of suitable features for oil spill detection (see figure 1.1).

A review of algorithms for automatic detection of oil spills can be found in Brekke and Solberg [5]. Several papers [9; 11; 28; 31] describe a methodology consisting of dark spot detection followed by feature extraction and classification. Solberg *et al.* [28, 31] apply an adaptive thresholding algorithm for dark spot segmentation. Kanaa *et al.* [16] use hysteresis thresholding. An edge detection approach based on the Laplace of Gaussians or Difference of Gaussians is presented in Change *et al.* [7]; Chen *et al.* [8]. The use of wavelets for ocean feature detection is described by Liu *et al.* [20]; Wu and Liu [38]. QinetiQ's dark spot algorithm uses a Constant False Alarm Rate (CFAR) algorithm followed by clustering and Hough transform to identify linear targets [30]. Edge detection by a CFAR filter is also applied in Migliaccio *et al.* [22], where experiments were done on airborne X-band SAR images (single-look). Barni *et al.* [3] propose an algorithm based on fuzzy clustering. A method using mathematical morphology for oil spill segmentation is presented by Gasull *et al.* [14]. Even though a variety of methods are applied, the common approach is to detect all suspicious slicks and to preserve slick shapes. At the feature extraction step, parameters that can be used to discriminate oil spills from other phenomena appearing on the sea surface are extracted. The features extracted vary between approaches. Most of the features used for slick feature extraction in the literature can be sorted into these classes: the geometry and shape of the segmented region, physical characteristics of the backscatter level of the spot and its surroundings, contextual features describing the slick in relation to its surroundings, and textural features. Little attention seems to have been given to the feature extraction

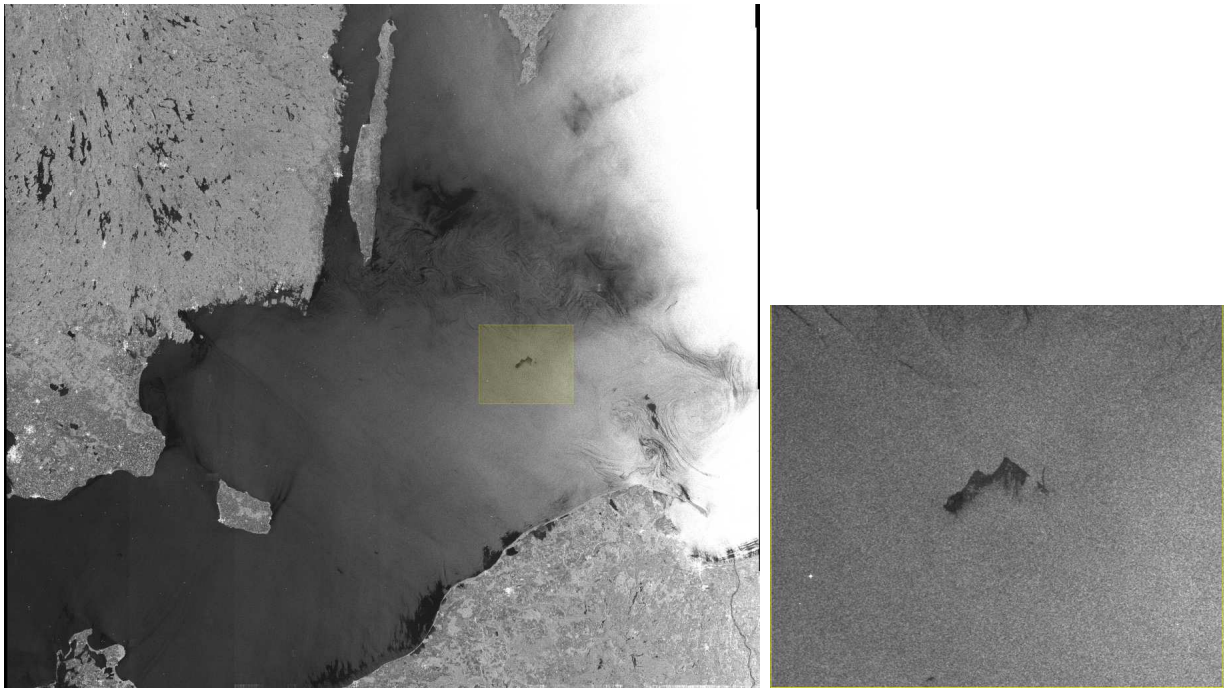


Figure 1.1: Left: ENIVSAT Advanced Synthetic Aperture Radar (ASAR) Wide Swath Mode (WSM) image from 6th of September 2005. Dimentions: 5596×5051. ©ESA/KSAT 2005. The image shows an example of a possible oil spill (within the shaded rectangle). A wind speed of 4 m/s was reported in the area, and low wind look-alikes are present. Right: zoom-in on the shaded rectangle. A dark spot with high contrast is visible. Kongsberg Satellite Services (KSAT) reports medium confidence that this slick is an oil spill.

step. An early study on feature extraction for oil spill detection based on ERS images is described by Solberg and Solberg [27], and an evaluation of the discrimination efficiency of typically used features can be found in Topouzelis *et al.* [33]. Various classifiers have been applied to classify a slick as oil or look-alike. Fiscella *et al.* [11] applied a Mahalanobis classifier and a compound probability classifier, and the probabilistic approach was improved by Nirchio *et al.* [23]. Solberg *et al.* [28, 31] combined a statistical classifier using a Gaussian model and several subclasses with a rule-based modification of prior probabilities to take into account feature combinations that are indications of certain scene conditions. A neural-network approach is described by Del Frate *et al.* [9], and Keramitsoglou *et al.* [17] developed a classifier based on fuzzy logic.

All detection algorithms suffer from false alarms, and slicks classified as oil spills may be confused with look-alikes. SAR imaging of oil slicks can be limited by atmospheric and oceanographic conditions. Girard-Ardhuin *et al.* [13] suggest a four step approach including the three steps 1) detection, 2) characterization, 3) classification and a fourth additional step taking external data into account. To improve classification and assess the slick nature,

characteristics of the automatically detected slicks from the SAR images and meteorological and oceanic data are combined through a multisensor approach (including e.g. surface wind measurements (from SeaWinds on QuikSCAT), sea-surface temperature (from MODIS on Terra and Aqua) information about atmospheric fronts and clouds (from AVHRR, Meteosat) and chlorophyll (from SeaWiFS)). A multisensor approach is also applied by Migliaccio *et al.* [22] where SAR data and wind fields from the SeaWinds scatterometer are combined.

Segmentation of dark spots and feature extraction are crucial parts of the algorithms for oil spill detection. If a slick is not detected during segmentation, it will not enter the classification procedure and cannot be classified correctly. Turning to classification, if the features have good discriminatory power, the classification problem will be easier and several classifiers can work. Keeping these issues in mind, we hereby present results from a study aiming at identifying suitable features that lead to significant improvements in classification performance for ENVISAT ASAR WSM images (some early results were presented in Brekke and Solberg [6]). In addition, we discuss improvements done to the detection of dark spots, leading to a segmentation approach contributing to the improved performance (the ideas were first outlined in Solberg *et al.* [29]). The novelties of this paper include a new segmentation algorithm for better detection of thin, linear or piecewise slicks and three new features. The new features include a feature measuring the curvature of the slick shape, an improved slick border detector and a feature measuring the variance of the slick pixel values.

This paper is organized in the following manner. Section 2 describes the automatic algorithm that constitutes the fundament of our work. The details of our new and latest contributions to the algorithm design are as well discussed here. Section 3 outlines the experimental design and presents the experimental results. Finally, the conclusions can be found in section 4.

2 The Automatic Oil Spill Detection Algorithm

Our goal is to develop an automatic system for oil spill detection, in which objects with a high probability of being oil slicks are automatically identified. These slicks should then be presented to an operator for manual inspection/verification before the information is forwarded to a surveillance aircraft. Processing techniques are considered *semi automatic* when a human expert selects the SAR image area to be processed and assist interpretation. However, our algorithm is considered *fully automatic* (in the sense that there is no need for manual interaction) up to the point when the operator is notified of a slick assigned high confidence of being an oil spill.

The framework of this study is a fully automatic advanced oil spill detection algorithm (see figure 1.2). The algorithm was originally intended to process ERS-1/-2 SAR images, but has now been adapted to work for RADARSAT-1 SAR and ENVISAT ASAR images. The algorithm includes sensor specific modules for dark spot detection, spot feature extraction and a classifier that discriminates between oil spills and look-alikes. Sensor-specific modules

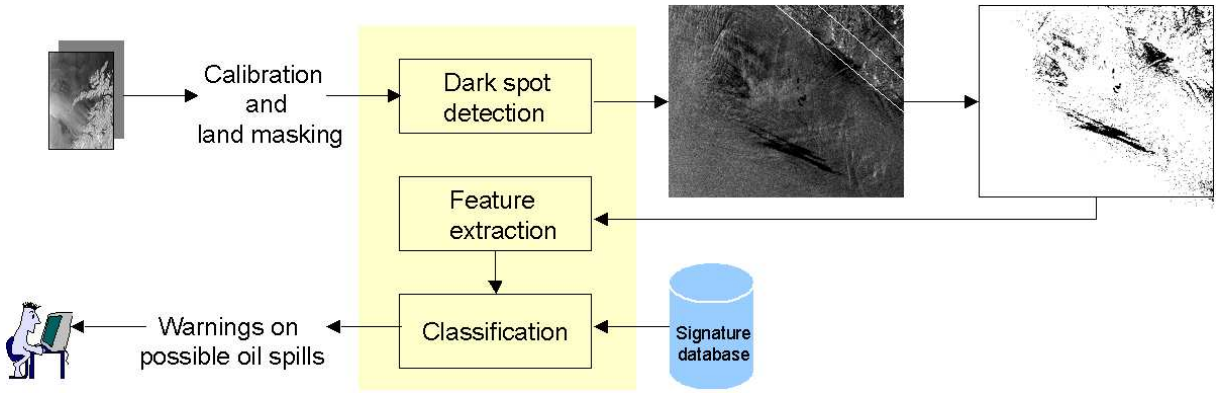


Figure 1.2: The oil spill detection algorithm and its context. Arrows indicate data flow.

are necessary because of different spatial resolution and contrast. RADARSAT-1 ScanSAR Narrow full resolution images (HH polarization) have a resolution of 50 m $\times$ 50 m and a pixel spacing of 25 m $\times$ 25 m. ENVISAT ASAR WSM images (VV polarization) have a resolution of 150 m $\times$ 150 m and a pixel spacing of 75 m $\times$ 75 m.

Pre-processing, consisting of converting a land mask to the image grid and a normalization of the backscatter with respect to incidence angles, is performed ahead of the segmentation step for ENVISAT ASAR images. The radar backscatter over land areas is not of importance for successful oil spill detection in the marine environment, therefore a land mask is created. To avoid re-sampling the speckle pattern, the land mask is created in the original geometry of the SAR image. The geo-referencing information available in the SAR product is used, and the land mask is obtained from Wessel and Smith [37]. The incidence angle range varies between 16° - 44° for ENVISAT ASAR WSM (see also table 1.8 in Appendix A). The radar backscatter from the sea depends strongly on the incidence angle, yielding a much stronger signal at low incidence angles. The dependency is not constant, however it depends on both wavelength and polarization of the SAR sensor and wind and weather conditions on the ocean surface. We have applied a standard incidence angle compensation for all weather conditions to reduce the impact of incidence angle variations on the statistical parameters extracted from the image (this is a however somewhat simplified solution, not estimating and compensating for the weather dependent effects).

2.1 Dark Spot Detection

The dark spot detector applies an adaptive threshold where the threshold is set k dB below the estimated local mean in a large window. The window is moved across the image in small steps to threshold all pixels in the scene. The goal of this step is to segment out all possible oil spill candidates (i.e. to create a binary image where the dark spots are represented by one pixel value and the background by another). A high number of look-alikes will also be

segmented out, but these will hopefully be classified as look-alikes during the classification step. From the segmented image, "region objects" are formed for all objects representing a dark spot (applying 8-pixels connectivity). In the following feature extraction step (described in section 2.2), only region objects are considered.

As earlier experiments showed that a single-scale segmentation approach did not perform satisfying on both large and small regions, a two-level multiresolution approach has been developed for ENVISAT ASAR images [31]. The resolution levels are created by averaging pixels in the original image. The first resolution level has the same dimensions as the original image, while the next resolution level is created with half the row and column dimensions of the original image.

Each resolution level is segmented using the adaptive threshold after applying a mean filter with a small window consisting of 3×3 pixels. The threshold is computed given the mean (μ) and the power-to-mean (PMR) ratio (σ/μ) (computed in a local window consisting of 121×121 pixels), and a *homogeneity* category. The homogeneity category is set according to the local PMR level. A look-up table for the parameter values can be found in Solberg *et al.* [31]. The threshold is thus set adaptively based on estimates from the SAR image of the roughness of the surrounding sea. The motivation is: in low wind with many look-alikes, the PMR value will be high, and also a high contrast between slicks and their surrounding can be expected. As the wind increases, the PMR value will decrease, and the expected contrast between the oil and the surrounding sea will also decrease. After segmenting both resolution levels, the results are merged.

The details of the initial segmentation algorithm can be found in Solberg *et al.* [31]. Figure 1.3 illustrates the result from the segmentation process.

Detection of Thin, Linear Slicks The procedure described above works very well in general, but in some cases it does not work for thin, linear slicks. These slicks are often piecewise linear due to moving ships changing directions, or wind or current altering the shape of the slick. Here, we present a new approach developed for ENVISAT ASAR WSM images for detecting these cases. The main idea is to grow elongated segments in the direction of their orientation, if certain criteria are fulfilled. After a proper segmentation of these slicks, their suspicious shapes can more easily be reflected through geometry and shape features.

The following algorithm is used to detect thin, linear slicks:

I. Threshold the SAR image with the approach described above, but fix k at a small dB value. $k = 1.3\ dB$ is here selected, which is based on experiments on the training data set and reported among the parameter values for thresholding in Solberg *et al.* [31]. (This gives an image with too many object pixels only intended to be used to locate fragments that can be part of a linear oil spill).

II. Compute Hu's first invariant planar moment (definition given in table 1.1) for every region object in the segmented image in I. Select the fragments with a moment value >

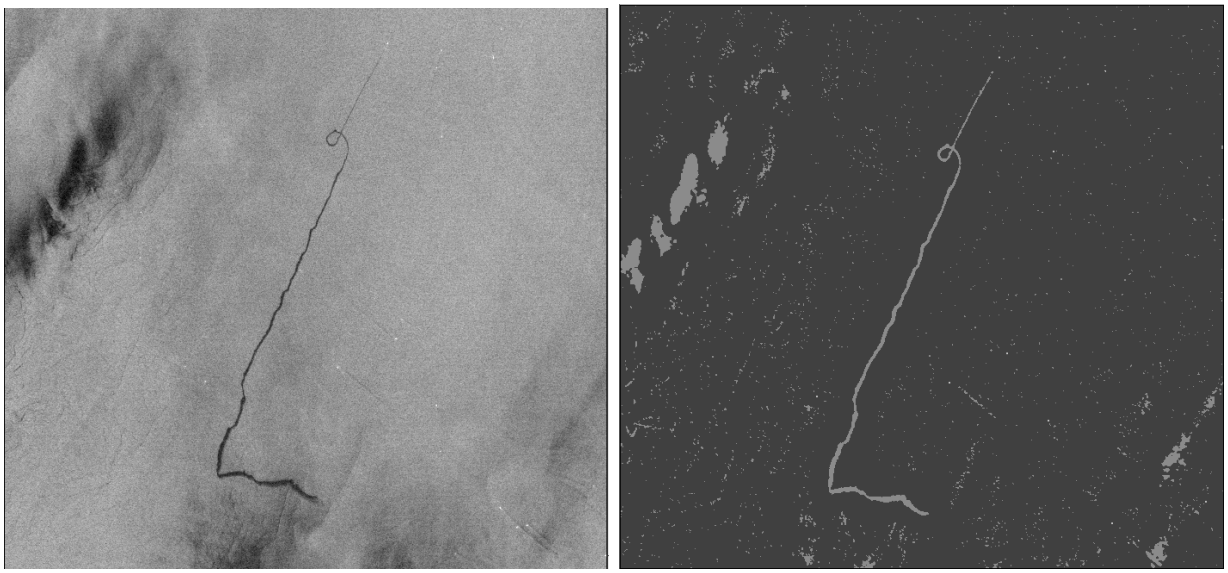


Figure 1.3: Left: Section of an ENVISAT ASAR WSM scene (16th September 2003). ©ESA/KSAT 2003. A long oil spill with a loop is clearly visible. A likely source can be seen at the upper end of the spill. Right: segmentation result. The oil spill and the ship, in addition to several other dark phenomena, have been separated from the background.

0.5 for further processing. The threshold is empirically derived from elongated regions in the training data set. (Hu's first invariant planar moment is a good indicator of elongatedness. See also figure 11).

III. For every selected fragment, find the *object-oriented bounding box* B_1, which is the smallest rectangle around the fragment having one side parallel to the orientation of the fragment, where the fragments orientation is defined as the angle (relative to the x-axis) of the axis through the center of mass that gives the lowest moment of inertia. See also Sonka *et al.* [32].

IV. For denoising slightly the ASAR image, apply a sigma filter [1] on the fragment and its nearby surroundings.

V. Set $T = \mu_{B_2} - \Delta dB$, where μ_{B_2} is the mean value inside an object-oriented box B_2, which is found by expanding the width and the length of B_1 (to cover a larger part of the background pixels). ΔdB is a small dB value found by experiments on the training data. $\Delta dB = 0.3$ is here selected.

[1] The window size used is 7×7 pixels. The center pixel is replaced with the average of all pixels in the window whose value is within ± 1 standard deviation, σ, from the center pixel. The σ is estimated in a larger window of 41×41 pixels surrounding the center pixel of the smaller window, giving us a local estimate of the σ for every window position (see Lee [19] for the original definition of the sigma filter).

VI. Extend the object-oriented bounding box B_1 by about 15 pixels in both directions of the fragments orientation, and reduce the width of B_1 to a couple of pixels if the ratio of the fragments width and length is larger than 0.2 (to narrow the area to be searched for broad and complex shaped fragments).

VII. Apply T as a threshold on every pixel i inside the extended parts of B_1. If the backscatter value of i is $< T$, accept i only if i or one of its 8-neighbour pixels represents an edge pixel, using the Sobel operator (an idea discussed in Trier and Jain [34]).

VIII. Add the segmented slick to the original segmented image.

See figure 1.4 for an illustration of the process.

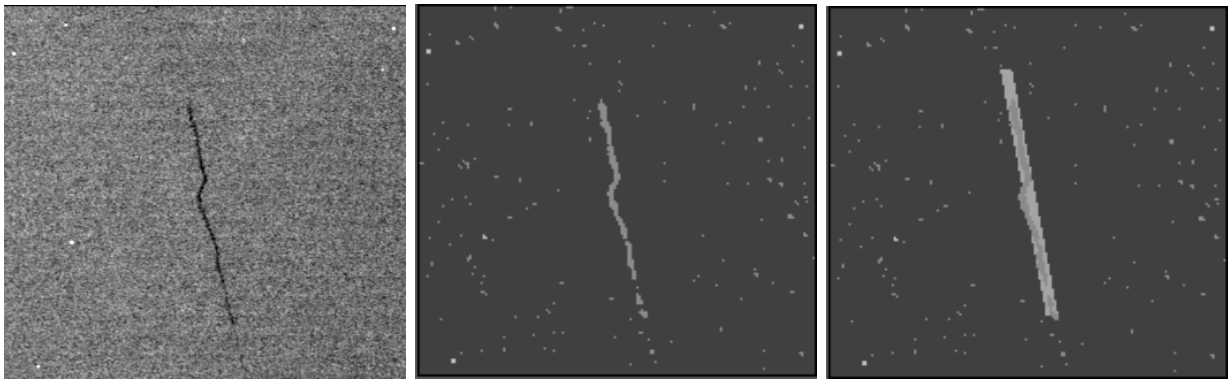


Figure 1.4: Left: section of an ENVISAT ASAR image (5th of December 2003). ©ESA/KSAT 2003. Middle: original segmentation result. Right: an extended object-oriented box (light grey) and the final segmentation result (darker grey).

The improved segmentation process works generally better than the basic, adaptive thresholding approach for thin, linear slicks. Figure 1.5 shows that fragments are merged and that the shape of the thin, linear slick is preserved by this approach. Figure 1.6 gives another examples of a slick affected by the approach.

After the segmentation process, a set of features are computed and extracted from every region above a certain minimum size. (This step eliminates the small regions found in figure 1.5 and figure 1.6).

2.2 Slick Feature Extraction

First, we briefly report the feature set derived from previous studies, then the new features studied in this paper are described.

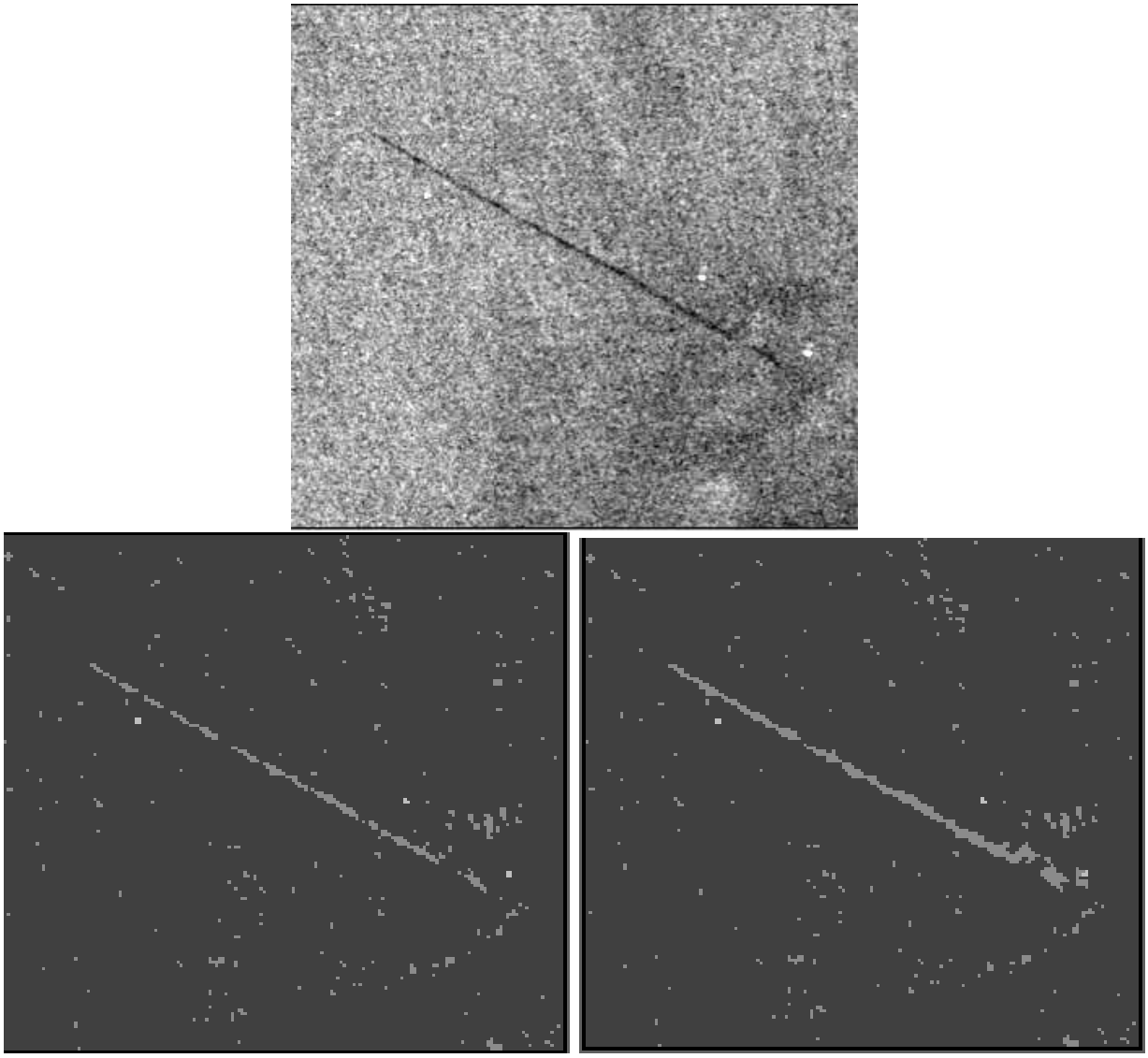


Figure 1.5: Top: section of an ENVISAT ASAR WSM image (19th of September 2003). ©ESA/KSAT 2003. Bottom: the original segmented image and the improved result after including the object-oriented box approach. Note that small regions are filtered out at a later stage.

Existing Set of Features

A basic set of features is described in table 1.1. The features are a mix of standard region descriptors and features tailored to oil spill detection. The features were presented in Solberg *et al.* [28, 31].

As no wind predictions were available here, and we did not have access to an algorithm

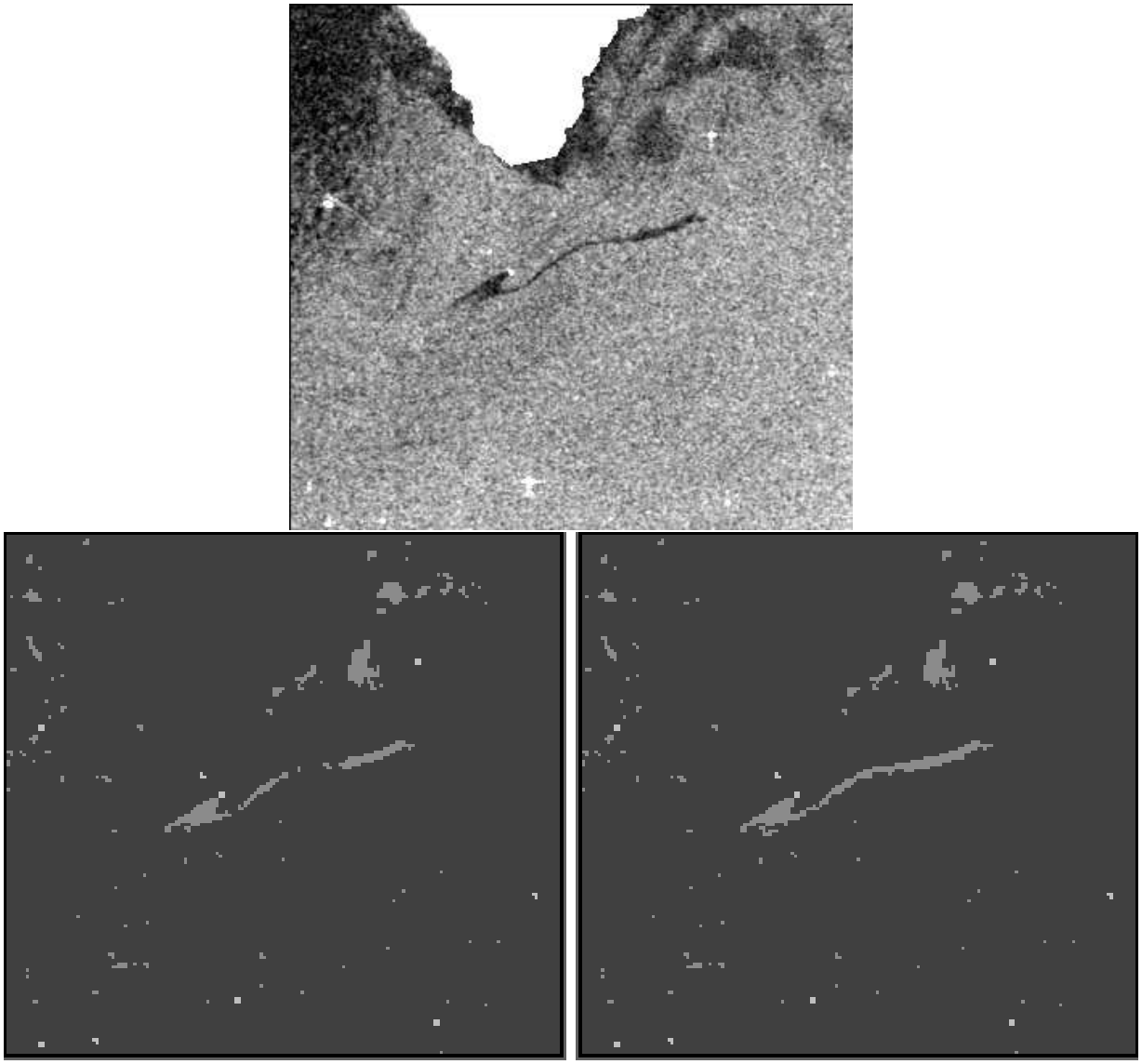


Figure 1.6: Top: section of an ENVISAT ASAR WSM image (17th of August 2003). ©ESA/KSAT 2003. The white area at the top of the image is part of a land mask. Bottom: the original segmented image and the improved result after including the object-oriented box approach. Note that small regions are filtered out at a later stage.

for wind speed estimation based on the SAR image, homogeneity categories were used to represent the WIND feature (#1 in Table 1.1). (As explained in section 2.1, a look-up table for the homogeneity categories based on the PMR feature can be found in Solberg *et al.* [31]). The categories are from LOW to HIGH.

Improvements to the feature vector are discussed in the next section. A goal of this paper

Table 1.1: Basic feature set. The features are based on geometrical characteristics, texture, the backscatter/contrast level, and contextual information.

#	Feature	Description
1	WIND (Wind level)	The local wind level in the scene.
2	MOM (Slick Moment)	Hu's first invariant planar moment of the region, defined as $\phi_1 = \eta_{20} + \eta_{02}$, where η_{pq} are normalised central moments: $\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\gamma}, \gamma = \frac{p+q}{2} + 1, \forall p+q >= 2$, μ_{qp} are central moments of order $p+q$ and μ_{00} are zero-order moments [15; 24].
3	COMPL (Slick complexity)	$C = P^2/A$, where P is the perimeter and A is the area of the region.
4	PMR (Power-to-mean ratio)	Homogeneity of the regions surroundings, defined as the ratio of standard deviation, σ_b, and mean, μ_b, of near-by background pixels surrounding the region. Window size used: 150×150 pixels.
5	LCONT (Slick local contrast)	Local area contrast, defined as $\mu_b - \mu_r$, where μ_b is the background pixel mean and μ_r is the region pixel mean (estimated in a window surrounding the region containing a larger amount of background than region pixels).
6	THICK (Slick thickness)	Defined as the ratio between the area of the region and the width of the branches of the region skeleton.
7	NSN (Number of regions in small neighbourhood)	The number of neighbouring regions completely within a window, consisting of 150×150 pixels, centered at the slick.
8	BGRAD (Slick border gradient)	The mean of the magnitude of the region border gradient. Sobel is used to compute the gradients.
9	SMC (Slick smoothness contrast)	Defined as the ratio between the ratio of the number of region pixels and the sum of the region gradient values, and the ratio of the number of background pixels and the sum of the background gradient values (window used contains a larger amount of background than region pixels).
10	AREA (Slick area)	The number of pixels in the region.
11	DIST (Distance)	The distance from the region to the closest bright spot (ship/oilrig).
12	NLN (Number of regions in large neighbourhood)	The number of neighbouring regions touching within a window, consisting of 150×150 pixels, centered at the slick.
13	NREG (Number of regions)	The total number of detected regions in the scene.

is to discuss new features and compare their performance to the existing.

New Features

Kubat *et al.* [18] recognize that good features are important for oil spill classification, but there are few guidelines on how to estimate them. In our search for new features that could add some discrimination information and improve the performance, we emphasised on:

- weaknesses in the existing set of features from table 1.1

- features reported valuable for the purpose of oil spill detection by other algorithms described in the literature
- other general features used in image analysis that we expected, based on our domain knowledge, could add some discrimination power.

Table 1.2 presents the features from table 1.1 grouped into feature categories. Shape,

Table 1.2: The features from table 1.1 grouped into categories.

Feature class	Feature
Shape	MOM, COMPL, THICK, and AREA
Texture	PMR
Contrast	LCONT, BGRAD, and SMC
Contextual	WIND, NSN, DIST, NLN, and NREG

texture, and contrast features will be discussed here. New features are introduced in the following, and the basic feature set has been extended with some of these. Selected feature combinations are evaluated through performance testing in section 3.2.

Contrast Features LCONT, BGRAD, and SMC defined in table 1.1 are features based on the backscatter level contrast between oil slicks and the sea. We have here focused on improving the BGRAD feature. It works generally well, but it seems to give inaccurate results for thin, linear regions. The Sobel operator is an edge detector that has been suggested used for oil slick border gradient estimation, see e.g. Girard-Ardhuin *et al.* [12, 13]; Solberg *et al.* [28]. In table 1.1, the mean value of the Sobel gradient magnitude is applied in the BGRAD feature. The main problem is that the edge response does not match the real borders of the region (e.g. a one pixel broad line gets no response for the object pixels but a one pixel top on both sides of the line). The top row to the right of figure 1.7 illustrates the response of the Sobel operator on the oil spill to the left in the same figure. As we can see, the largest gradient magnitude appears outside the true region border. The following four additional convolution masks are suggested as a line detector for thin oil spill regions:

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & -4 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & -4 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \textit{and} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -4 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

The response from the mask returning the largest magnitude is selected. The bottom row of figure 1.7 illustrates the response to these masks. The response from the additional masks decreases as the slick gets thicker. If the Sobel operator gives stronger magnitude response to any of the border pixels that value is kept, otherwise the response from the additional

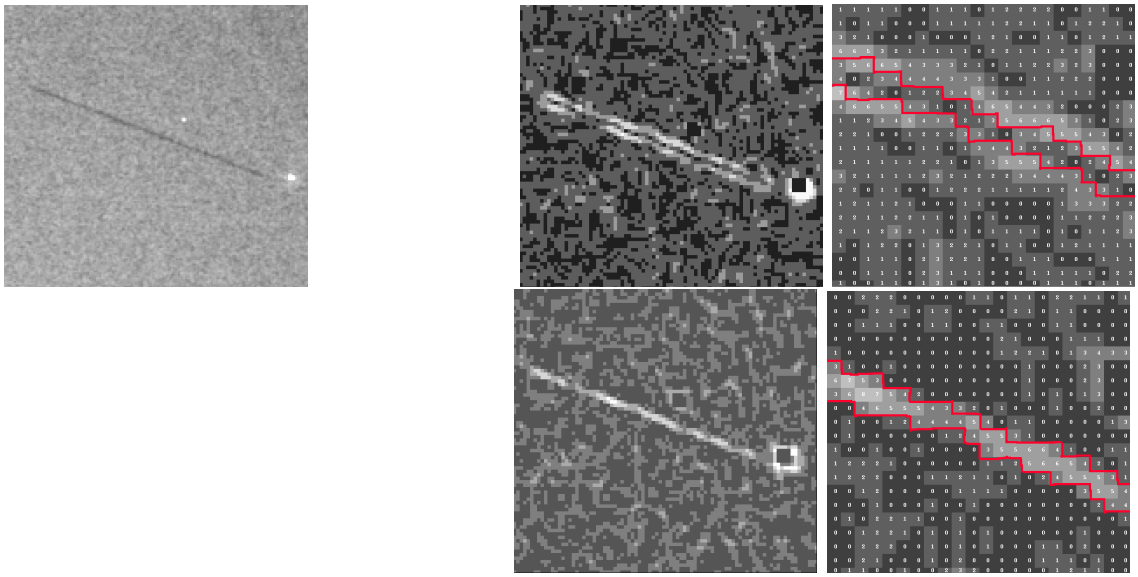


Figure 1.7: Left: section of an ENVISAT ASAR WSM image (24th of July 2003). ©ESA/KSAT 2003. A possibly thin, linear oil spill with a likely source visible. Top row to the right: the Sobel operator yields the largest border gradient magnitude response outside the slick region. The detected edge is misplaced with respect to the real border of the region. (The red lines shows the outline of the region border). Bottom row to the right: response from improved border estimation.

masks is used. This gives us an ad hoc solution where a gradient detector and a line detector are combined, and further refinement of this method is likely possible. However, the mean of this border detector gives us an improved indication of the contrast to the surrounding background for thin regions, and it is used in the feature SLICK_BORDER to replace the BGRAD feature in table 1.1.

The standard deviation of the slick border gradient values were found significant for the success of classification in Del Frate *et al.* [9]. To be able to evaluate this feature, we implemented it (SLICK_BORDER_STDEV) based on our improved slick border detector described above. The result from the evaluation is presented in section 3.2.

Texture Texture refers to the properties that represent the surface or structure of an object, but there is no precise mathematical definition of texture due to its wide variability. In table 1.1 there is no feature representing the texture of the slick it self, only of its surroundings. Solberg *et al.* [28] have earlier suggested the PMR of the slick, defined as σ_r/μ_r where σ_r is the standard deviation and μ_r is the mean value of the pixels belonging to the slick. Del Frate *et al.* [9] have simply used the standard deviation of the slick as a texture measure. However, the standard deviation of the intensity values of the pixels belonging to a slick

is highly correlated with the area/size of the region. This is due to the inherent nature of speckle statistics. Speckle is a large problem in SAR images since even a homogeneous area has a statistical distribution with large standard deviation. As the region grows larger the variance in intensity values will increase as well. A better choice would be to look at the ratio σ_r^2/A, where σ_r is the standard deviation and A is the area of the slick. This is implemented as the new feature SLICK_VARIANCE. After normalization by area, the measured feature values of larger oil spills are comparable to smaller samples.

Geometrical/Shape Properties MOM, COMPL, THICK, and AREA defined in table 1.1 are all shape features. Generally, look-alikes are much more complex in shape with less smooth borders than man made oil spills. Oil spills are often elongated or have wide, regular shapes caused by stationary objects releasing a large amount of oil. We have here aimed at identifying features that can represent curvature and elongatedness, as this seems like two important properties of shape that can help to distinguish oil spills from natural slicks.

Features based on the ratio between the perimeter and the area, aiming at describing the shape complexity of regions, have been used in several algorithms described in the literature [11; 28; 9; 33]. Del Frate *et al.* [9] applies the formula $C = P/2\sqrt{\pi A}$, where P is the perimeter and A is the area of the region. According to Del Frate *et al.*, this feature is generally expected to get a small numerical value for regions with simple geometry, while a larger value for more complex regions. In contradiction to common intuition, the thin, linear oil spill to the right of figure 1.8 gets a larger complexity value than both the others when using COMPL in table 1.1. Del Frate *et al.*'s [9] formula gives very similar but differently scaled results. This indicates that the ratio between perimeter and area is not a good complexity measure as it is not possible to separate complex shaped slicks from linear slicks. This weakness is also pointed out by Nixon and Aguado [24], and Topouzelis *et al.* [33] found that the feature gave little contribution to oil spill detection.

To resolve this ambiguity we could introduce additional shape measures, or replace this measure with a more robust one. A possibility is to look at the number of branching points [2] in the skeleton of each region (see figure 1.9). Because we only look at the number of branching points, the information level is decreased so much that again it is often not possible to distinguish simple regions from more complex ones (e.g. a straight line would get the same feature value as an "S" shaped region). It is clear that it is important to preserve more shape information in features expressing geometrical complexity.

Instead, we propose a new measure of complexity. Contour or snake models are commonly applied in medical image segmentation, e.g. ultrasound images. In Lobregt and Viergever [21] a discrete dynamic model for defining contours in 2-D images is presented. Lobregt and Viergever gives a definition of a contours local curvature as a part of their model description. As the definition is general, it can be adapted as a feature extracted from the contour of dark structures appearing in SAR images. The edge segment leaving from vertex V_i is represented

[2]We define *the number of branching points* as a point with three lines or more connected to it.

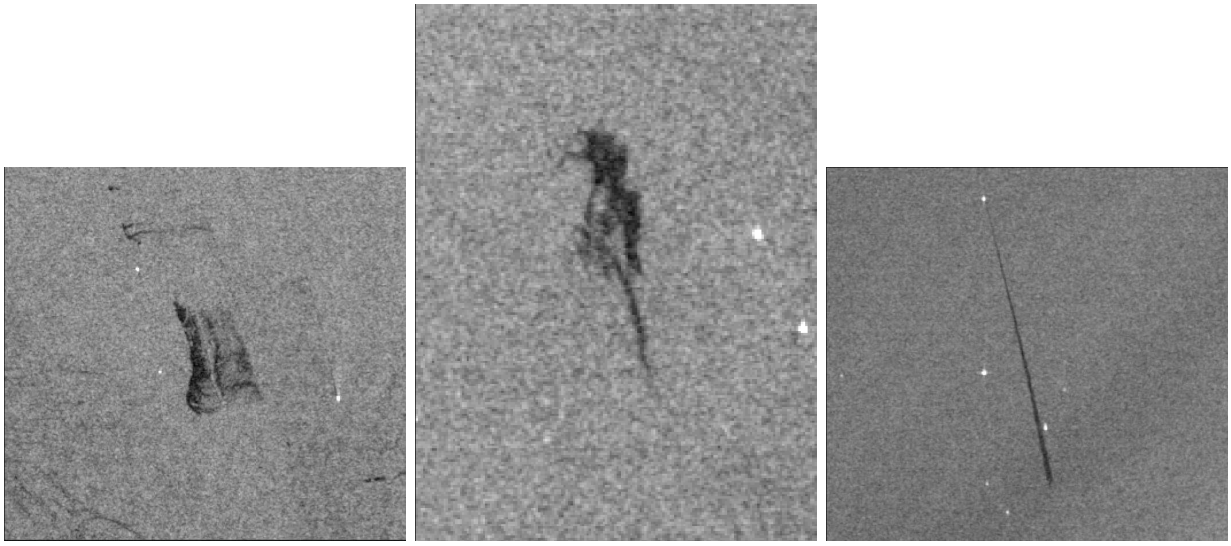


Figure 1.8: Sections of ENVISAT ASAR WSM images. Left: 3rd of August 2003. Middle: 7th of August 2003. Slicks with wide, complex, irregular shapes. Right: 2nd of December 2003. Thin, linear oil spill with a regular shape. ©ESA/KSAT 2003.

by a vector $\mathbf{d}_i$; its direction is described by the unit vector $\hat{\mathbf{d}}_i$ (see figure 1.10). The local curvature $\mathbf{c}_i$ is defined as the difference between the directions of two edge segments that join at a vertex: $\mathbf{c}_i = \hat{\mathbf{d}}_i - \hat{\mathbf{d}}_{i-1}$. The local curvature has length and direction. This provides a measure of the angle between two joining edge segments. The length of the curvature vector depends only on this angle and is not influenced by the lengths of the two edge segments. In our implementation of the curvature feature, we have traced the boundary of every region and inserted vertexes with a three-pixel spacing [3]. This gives us an approximation of the contour given by vertexes which are connected by straight line segments or edges. The angle between two edge segments is calculated as described above, and the final SUM_EXTERNAL_ANGLES feature is the sum of all local curvature measures along the boundary. Each external angle in a regular n-sided polygon will be $2\pi/n$ and the sum of the angles will be 2π. More complex regions get a higher curvature measure.

Concerning elongatedness, Topouzelis *et al.* [33] report the standard deviation of the asymmetries of sub-objects as one of the most effective features of those that were evaluated in their study. Asymmetry were expressed as the ratio of the lengths of the minor and major axes of the object approximated ellipse [4]. Topouzelis et al. emphasise that the lengthier an object is the more asymmetric it is. We implemented and calculated the asymmetry feature

[3]The distance between the vertexes represents the resolution of the approximation. If it is large, the model will not be able to follow variations of small scale along the contour. However, this regression might not be optimal as it is uncertain if the true vertexes of the contour are located.

[4]The *object ellipse* is defined as the ellipse whose least and greatest moments of inertia equal those of the object. The length of the semimajor and semiminor axes can be expressed as [39]:

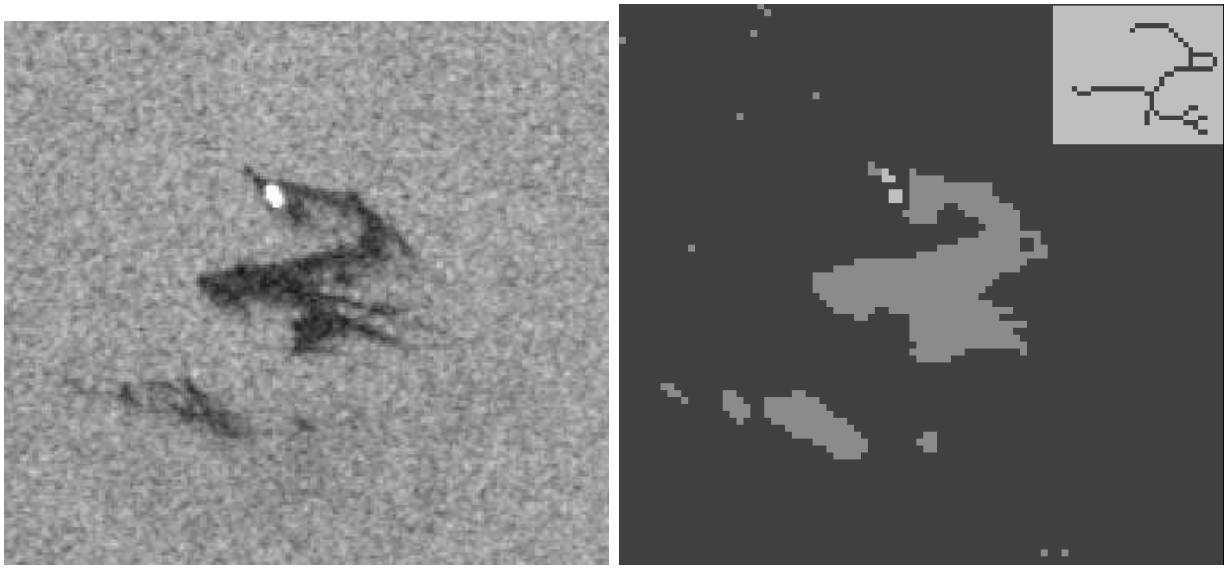


Figure 1.9: Left: section of an ENVISAT ASAR WSM image (3rd of August 2003). ©ESA/KSAT 2003. Right: example of a segmentation result of the oil spill and its skeleton (presented in the small frame in the upper right corner).

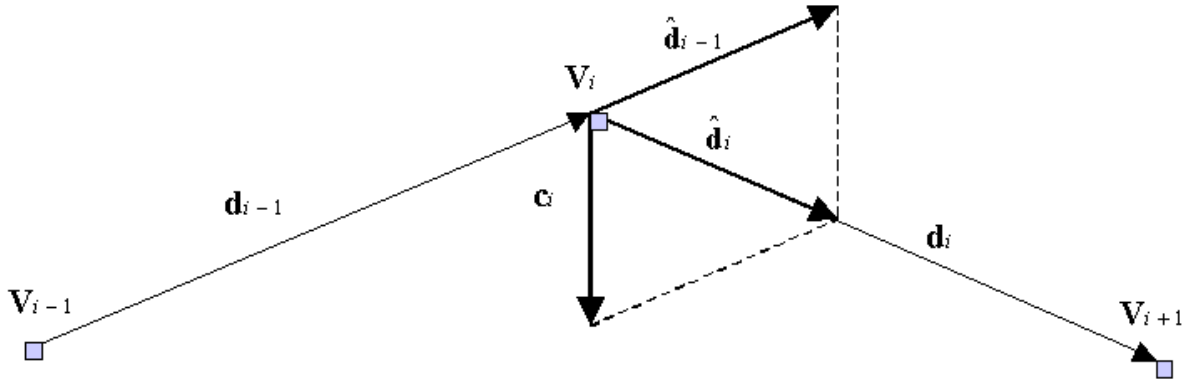


Figure 1.10: Local curvature $\mathbf{c}_i$. $\hat{\mathbf{d}}_{i-1}$ and $\hat{\mathbf{d}}_i$ are the directions (unit vectors) of the edge segments $\mathbf{d}_{i-1}$ and $\mathbf{d}_i$ meeting at vertex V_i.

for a number of dark regions. Figure 1.11 compares the asymmetry feature with MOM, and as we can see the asymmetry feature is comparable (by an opposite correlation) to Hu's first invariant planar moment. However, figure 1.11 also shows that the latter more significantly reflects the elongated regions.

$(\alpha, \beta) = \sqrt{\frac{2[\mu_{20}+\mu_{02} \pm \sqrt{(\mu_{20}-\mu_{02})^2+4\mu_{11}^2}]}{\mu_{00}}}$

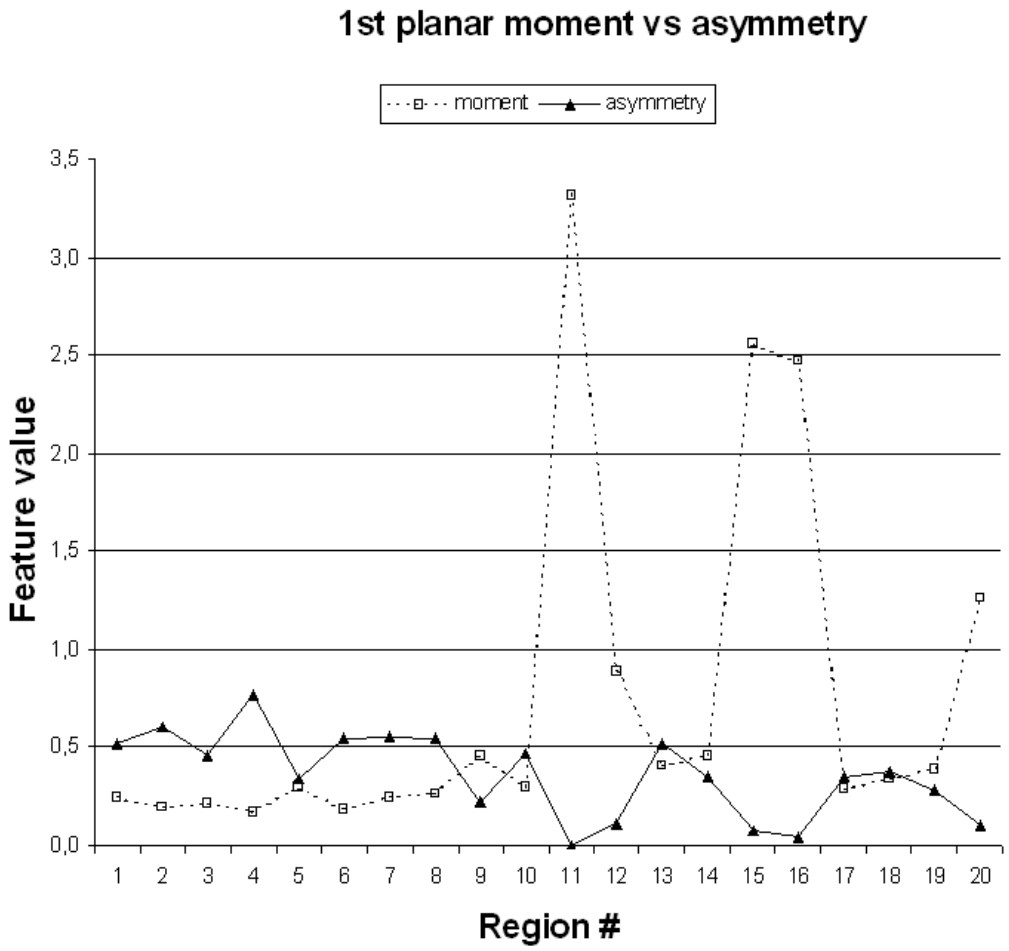


Figure 1.11: Comparison of Hu's first invariant planar moment and asymmetry. 20 regions (out of thousands) are selected as examples from the segmented images. Region #11, #12, #15, #16, and #20 represents elongated regions.

2.3 Statistical Classification

After a set of M dark spots has been detected, we want to classify them as either oil spills or look-alikes. For this purpose, a classification algorithm has been developed (see Solberg *et al.* [28, 31]) combining a statistical model for oil spills of different shapes and seen under different wind conditions, a prior model for the probability of observing oil and look-alikes, and a rule-based approach which takes care of certain expert knowledge related to oil spill detection.

The WIND feature is first used to divide the objects in two different subclasses and

then these are each further divided into five subclasses based on the shape descriptor MOM. This is due to the fact that the variance in shape, contrast, and surroundings of oil slicks and look-alikes depends on the wind level. The densities inside each subgroup are assumed Gaussian. Common covariance matrices for both the oil slick and the look-alike class (given the wind level and shape) are used. This is because look-alikes occur much more frequently than oil slicks and their features will vary more than the features for oil slicks. Because of limited training data for oil slicks, a further reduction in the number of unknown parameters to be estimated is obtained by assuming the covariance matrices to be diagonal. This choice is an important one, because correlations between features are ignored by assuming diagonal covariance matrices.

A set of rule-based corrections of the prior probabilities in certain situations were applied in Solberg *et al.* [31]. Without these rules, the classifier resulted in a high number of false alarms, because a high number of look-alikes were classified as oil. These rules were determined by inspecting the misclassified slicks to understand why they were misclassified. By inspecting the feature values for these slicks, a set of rules for these circumstances was established.

Combinations of features are known to give strong indications either toward an oil slick or toward a look-alike [28]. Because we use diagonal covariance matrices in the Gaussian densities, knowledge of such combinations are instead utilized through the rule-based corrections of the Gaussian densities (or actually the ratios between the densities of oil slicks and look-alikes). These corrections are specifically important in reducing the number of false alarms (as will be illustrated in section 3).

The basic feature vector, feature #3 - #9 in table 1.1, was used to compute the probability densities in Solberg *et al.* [28, 31]. Here, we have introduced new features and therefore we need to alter the feature vector. Which feature vector that is used to compute the probability densities are specified together with every classification result presented in the next section. However, our final result is based on a feature vector consisting of feature #3 - #7, SLICK_BORDER, #9, and SLICK_VARIANCE. The SUM_EXTERNAL_ANGLES feature is included together with feature #2, #4 - #7, SLICK_BORDER, and #9 - #13 in the rule-based corrections of the probability densities.

The classifier is trained (that is, the means and the covariance matrices in the Gaussian densities are estimated) on a large set of labelled samples. The training process is further described in section 3.

3 Performance Testing

According to Alpers and Espedal [2], C-band single-polarized VV SAR has so far proven to be the most efficient SAR configuration for oil spill detection. No significant difference in practical performance between the detection capabilities of the HH-polarized RADARSAT-1 ScanSAR mode versus the VV-polarized ENVISAT ASAR WSM mode has yet been reported

for operational use, however the experience might still be limited (Shepherd *et al.* [26]; Solberg *et al.* [30]). Usually even small volumes of oil cover large areas and thus the need for very high spatial resolution in SAR images is not crucial. Bern *et al.* [4]; Wahl *et al.* [36, 35] found Low Resolution ERS-1 SAR images with a spatial resolution of 100 m sufficient for oil spill detection. ENVISAT's ASAR WSM covers a much wider swath than its predecessors ERS-1 and -2, but the resolution is lower. Generally, for efficient oil spill monitoring larger swath widths should be chosen on the expense of somewhat lower resolution. Because of ENVISAT's good oceanography and oil spill detection properties, we have chosen its WSM mode (VV polarization) for our experimental design. Olsen and Wahl [25] argue that if we base our evaluation on a contrast requirement of -2 to -3 dB for labelling features as potential oil slicks, this means that ocean backscatter should be at least 3 dB above the noise floor to support slick discrimination. The estimated noise floor is about 4 dB below the expected low-wind backscatter for VV-polarized data at the outer edge of the swath, while HH-polarized data will be around the expected noise floor. Therefore, VV-polarized wide swath data would be the best option for monitoring slicks. (See also Brekke and Solberg [5] for a discussion on satellite sensors and modes for oil spill detection). Table 1.8 in Appendix A gives a description of some key features of the SAR sensor and mode selected for this study.

This section covers the experimental design and the experimental results produced during performance testing of the improved segmentation process and the new features discussed in the previous sections.

3.1 Experimental Design

Our results are based on a set of 83 ENVISAT ASAR WSM images. We have aircraft verifications collected by the European Commission (EC) project Oceanides for 27 of the scenes. This is done in collaboration with Kongsberg Satellite Services (KSAT), QinetiQ, Norwegian Computing Center (NR), German pollution control authorities (MLZ) and Finnish pollution control authorities (SYKE) (see Solberg *et al.* [30]).

For performance testing, the set of SAR scenes is split into two parts. 56 of the SAR scenes are used for training and adjusting the model parameters, and the 27 benchmark scenes are used as a validation/test set to estimate the generalization error. The training set is collected from the German and Finnish Baltic Sea, the North Sea and some along the Norwegian coastline during March to December 2003 and January to April 2004. The benchmark set is collected mainly from the German and Finnish Baltic Sea and the German North Sea between July and December 2003.

To label the training set, we created a training mask for each SAR image. By manual inspection, we systematically scanned through all images tagging off dark spots as either *oil spill* or *doubt*. The doubt category was only used for a limited number of cases that could either be classified as oil spills or look-alikes. Doubt cases were left out of the training process. All slicks not marked as oil or doubt were used to train the *look-alike* class. A

similar procedure was performed for the test/benchmark set. All together, 37 dark spots were tagged off as oil spills and thousands of dark segments as look-alikes. Not all of the possible oil spills were checked due to e.g. flight cancellations or oil spills occurring outside the Finnish and German maritime territories. The classification performance was evaluated on the total set of dark spots, and not on the image level. Doubt cases are left out of all the classification results presented in the following.

3.2 Intermediate Classification Results - Improved Feature Extraction

Table 1.3 presents the results from classifying the complete benchmark set of 27 scenes by applying feature #3 - #9 in table 1.1 as a basic feature vector. This classification was done

Table 1.3: Classification results based on the basic feature vector.

Segmentation	Basic feature vector	Rule-based corrections	Correctly classified oil spills	Correctly classified look-alikes
Basic, adaptive thresholding extended with the object-oriented box approach	#3 - #9	No	89%	90%

without the rule-based corrections of the probability densities described in Solberg *et al.* [28, 31] and in section 2.3. (The rule-based corrections are based on the observed values of the features on the training set. When replacing some of the features, the rules have to be modified [5]. To adjust the rule-based corrections according to new features is a time consuming process, and it is not necessary for the purpose of demonstrating the efficiency of the new features. The rule-based corrections are left out of all performance results presented in this section, but are however included in the final results presented in the next section). All other parameter settings, except for the feature vector, were fixed in the algorithm during feature evaluation. The results in table 1.3 can for this reason be used as a reference for table 1.4 and 1.6.

The SLICK_BORDER_STDEV feature (see section 2.2) were added to the basic feature vector (feature #3 - #9 from table 1.1) and the classification results are presented in table 1.4. By comparing this result with table 1.3, it is clear that this feature does not add any significant value to the classification accuracy. For this reason this feature were not taken into further consideration.

Table 1.5 gives a definition of the new features we finally selected for our experiment, and Table 1.6 presents the intermediate classification results after introducing one at the

[5] An example of a rule is [31]: *"Increase prior probability for oil spills if* $DIST < 10$ *AND* $NLN = 0$ *AND* $COMPL < 20$". DIST, NLN and COMPL are features (see table 1.1). If we change the definition of these features we then might have to adjust the limit values as well.

Table 1.4: Evaluation of the slick border standard deviation feature.

Segmentation	Feature vector	Rule-based corrections	Correctly classified oil spills	Correctly classified look-alikes
Basic, adaptive thresholding extended with the object-oriented box approach	#3 - #9 and SLICK_BORDER_STDEV	No	89%	90%

Table 1.5: New set of features.

#	Feature	Description
14	SLICK_BORDER	Defined as a combination of the mean of the magnitude of the Sobel gradient values of the region border area and the four additional masks described in section 2.2.
15	SLICK_VARIANCE	Defined as the ratio σ_r^2/A, where σ_r is the standard deviation and A is the area of the slick.
16	SUM_EXTERNAL_ANGLES	Defined as the sum of all local curvature measures (changes of slope) along the boundary.

time of the new features into the basic feature vector. The BGRAD feature in table 1.1 is

Table 1.6: Intermediate classification results after introducing the new features.

Segmentation	Feature vector	Rule-based corrections	Correctly classified oil spills	Correctly classified look-alikes
Basic, adaptive thresholding extended with the object-oriented box approach	#3 - #7, #14, and #9	No	95%	88%
Basic, adaptive thresholding extended with the object-oriented box approach	#16 and #4 - #9	No	89%	91%
Basic, adaptive thresholding extended with the object-oriented box approach	#3 - #9, and #15	No	92%	89%

substituted with the improved border gradient detector SLICK_BORDER, COMPL with SUM_EXTERNAL_ANGLES, and SLICK_VARIANCE is added as an additional feature to the feature vector. As the new features either replace old features or measure a new property not represented in the old feature set, we avoid high correlation among the features and the assumption of diagonal covariance matrices is still valid. By introducing each of the new

features in table 1.6, either the number of correctly classified oil spills or the number of look-alikes correctly classified has increased compared to table 1.3. For SLICK_BORDER and SLICK_VARIANCE have the numbers of correctly classified oil spills increased, while there is a reduction in the numbers of look-alikes correctly classified compared to table 1.3. For SUM_EXTERNAL_ANGLES the number of correctly classified look-alikes has increased, while the number of correctly classified oil spills is unchanged.

3.3 Final Classification Results

In the previous section, we illustrated the importance of suitable features for oil spill detection. A good segmentation result can as well contribute significantly to the success of oil spill detection. Table 1.7 presents the classification results for the benchmark set before (first table row) and after (second table row) introducing the object-oriented box approach (described in section 2.1). The basic feature vector, feature #3 - #9 in table 1.1, is kept in both cases. The object-oriented box approach aims at detecting thin, linear oil spills often left

Table 1.7: Final classification results. Rule-based corrections of the probability densities were included in the algorithm.

Segmentation	Feature vector	Rule-based corrections	Correctly classified oil spills	Correctly classified look-alikes
Basic, adaptive thresholding	#3 - #9	yes	78%	99%
Basic, adaptive thresholding extended with the object-oriented box approach	#3 - #9	yes	86%	99%
Basic, adaptive thresholding extended with the object-oriented box approach	#3 - #7, #14, #9, and #15	yes	89%	98%

out or fragmented by the basic, adaptive thresholding approach (described in section 2.1). This class of oil slicks are often visible behind vessels in motion, discharging oil while sailing. Figure 1.12 presents two of the cases that were missed by the basic, adaptive thresholding, but classified correctly as oil spills after adding the improvements to the segmentation process. The large increase from 78% to 86% in the number of detected oil spills shows the importance of a tailored segmentation process to this particular application.

Finally, the last row of table 1.7 presents the classification results after also introducing the new features from table 1.5 and adjusting the rule-based corrections of the probability densities according to the new feature values of the training set. The SUM_EXTERNAL_ANGLES feature (#16 in table 1.5) is not included in the final feature vector (used to compute the probability densities), but instead applied directly in the rule-based corrections of the probability densities. SUM_EXTERNAL_ANGLES was shown in table 1.6 to decrease the

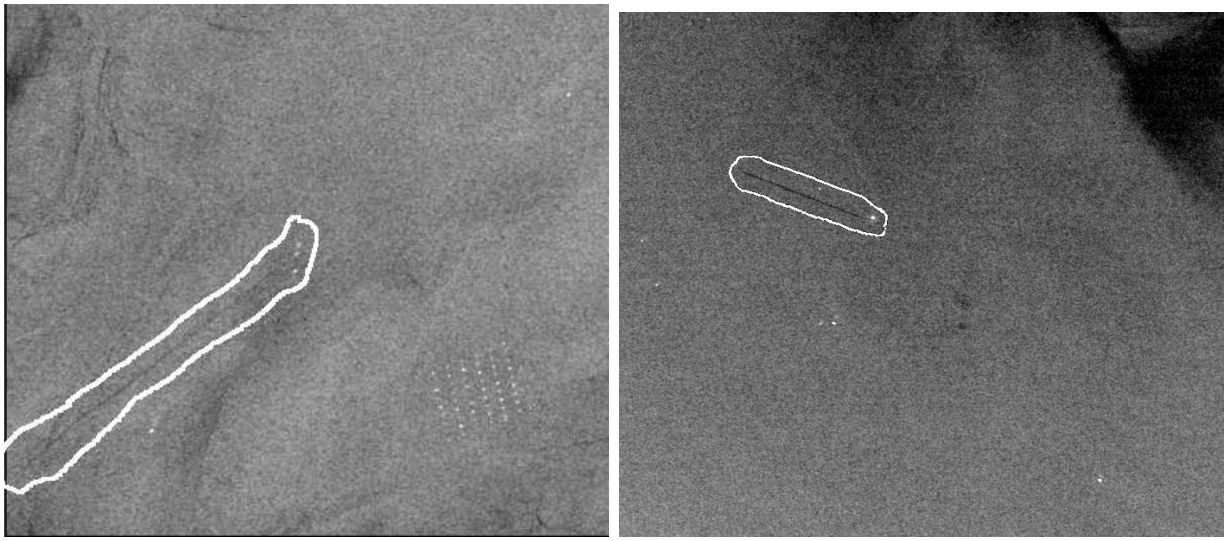


Figure 1.12: Sections of ENVISAT ASAR WSM images (13th of July 2003 and 24th of July 2003). ©ESA/KSAT 2003. Thin, linear slicks (marked off) that were correctly detected as oil spills after the improvements to the dark spot detector were added to the algorithm.

false alarm rate, and contributes with this property through the rule-based corrections. The results show that the combination of an improved segmentation process and an improved set of features lead to a significant better classification result for oil spills. It gives us an increase from 78% to 89% when comparing the first and third row of table 1.7. In contradiction to the intermediate results presented above (section 3.2), we have here in table 1.7 added rule-based corrections of the probability densities. This contributes strongly to the very high classification rate of 98% for look-alikes. We do only get 238 false alarms out of 12 247 regions tagged off as look-alikes.

Even though some of the desired alarms on possible oil spills are lost, this motivates the use of the rule-based corrections as a classification rate of 88-91% for look-alikes (table 1.6) is far too low considering the number of false alarms. However, there are room for improvement because of the large amount of time and effort necessary to produce such a rule-base when e.g. adapting the algorithm to a new sensor mode.

Detecting the dark spots (segmentation) is the most computationally intensive step of the algorithm. Adding the object-oriented box approach to this step might call for an optimization of the algorithm with respect to processing time.

Del Frate *et al.* [9], Solberg *et al.* [28], and Fiscella *et al.* [11] describe automatic oil spill detection algorithms and report accuracies from 82% to 94% in the number of correctly classified oil spills. These studies are however performed on different data sets, and the methods differ both in segmentation approach, features extracted, and classification methodology. For this reason, these results are not directly comparable with our results.

4 Conclusions

In this paper, we have developed new features for oil spill detection that lead to improvements in classification performance of SAR images. In addition, a new segmentation algorithm aiming at detecting thin, linear oil spills is presented.

We have studied properties of the border gradient and texture measures of the slicks. These features are based on the backscatter level characteristics of the slicks and their surroundings. In addition, we have evaluated several features measuring various aspects of the slicks shape. The use of curvature, as adopted from the well-known concepts of contour models (snakes), is suggested as a more robust complexity feature than those commonly applied in the oil spill remote sensing literature.

We have as well illustrated through examples and improved classification results the importance of an adapted segmentation process. A combination of an improved segmentation of thin, linear oil spills, an improved set of features, and adjusted rule-based corrections of the probability densities has lead to a significant increase from 78% to 89% in the number of correctly classified oil spills. 98% of the look-alikes are correctly classified. The algorithm has been evaluated on a set of 83 ENVISAT ASAR WSM images.

Using a statistical classifier, our experimental results show that the rule-based corrections of the probability densities are very important in keeping the false alarm rate down, but sometimes, unfortunately, at the expense of loosing some of the desired alarms on suspected slicks. However, such modifications of the feature densities are not common in statistical classification, but more often used in knowledge-based classification methods [28]. To improve the algorithm, incorporation of external knowledge about algae blooms (from e.g. numerical models, see e.g. Albretsen *et al.* [1], optical sensors, and geographical information related to seasonal algae-blooms) and a database of "hotspots" (e.g. oilrigs, sunken ships, seepages and other phenomena known to create dark patches on given locations) should be considered.

Some large-scale classification studies with acceptable classification performance, have been reported earlier using e.g. statistical classification and neural nets. As the performance results are based on different data sets and sometimes even different SAR sensors, a comparison of various classifiers based on the same data set would be desired.

Appendix A: Data Set Characteristics

Table 1.8 presents information concerning the data sets used in this study with respect to SAR product details.

Acknowledgment

The work of Camilla Brekke was performed as part of a Ph.D. study funded by the Norwegian Research Council and the Norwegian Defence Research Establishment. Thanks to Per Ove

Table 1.8: SAR product details.

Sensor	ENVISAT ASAR
Mode	ASA_WSM_1P: Wide Swath-Medium Resolution Image
Frequency (GHz)	5,331 (C-band)
Polarization	V/V
Incidence angles (°)	16 - 44
Approx. coverage (km) (range × azimuth)	400 × 400
Nominal resolution (m) (range × azimuth)	150 × 150
Nominal pixel spacing (m) (range × azimuth)	75 × 75
Processing station	KSAT

Husøy (NR) for pre-processing the SAR data.

References

[1] Albretsen, J. , Winther, N., Søiland, H. and Røed, L. P., 2004, Models in MONCOZE. Technical report, Norwegian Meteorological Institute, Report no. 14/2004.

[2] Alpers, W. and Espedal, H., 2004, Oils and Surfactants. *Chapter 11 in Synthetic Aperture Radar Marine User's Manual*, National Oceanic and Atmospheric Administration (NOAA). Jackson, C.R. and Apel, J.R.(Eds.), pp. 263-275.

[3] Barni, M. and Betti, M. and Mecocci, A., 1995, A Fuzzy Approach to Oil Spill Detection on SAR Images. *In Proc. of International Geoscience and Remote Sensing Symposium, 1995,* pp. 157–159.

[4] Bern, T.-I., Moen, S., Wahl, T., Anderssen, T., Olsen, R. and Johannessen, J. A., 1992, Oil Spill Detection Using Satellite Based SAR. Completion Report for Phase 0 and 1, OCEANOR report no. OCN-R92071, Trondheim.

[5] Brekke, C. and Solberg, A.H.S., March 2005, Oil spill detection by satellite remote sensing. *Remote Sensing of Environment,* **95 (1)**, pp. 1–13.

[6] Brekke, C. and Solberg, A.H.S., 2005, Feature extraction for oil spill detection based on SAR images. *Lecture Notes in Computer Science,* Kalviainen, H. et al. (Ed.), **3540**, pp. 75–84.

[7] Change, L. Y., Chen, K.S., Chen, C.F. and Chen, A.J., 1996, A Multiplayer-Multiresolution Approach to Detection of Oil Slicks Using ERS SAR Image. *In Proc. of ACRS 1996 - 17th Asian Conference of Remote Sensing, Sri Lanka.*

[8] Chen, C. F., Chen, K. S., Chang, L. Y. and Chen, A. J., 1997, The Use of Satellite Imagery for Monitoring Coastal Environment in Taiwan. *In Proc. of International Geoscience and Remote Sensing Symposium, 1997*, **3**, pp. 1424–1426.

[9] Del Frate, F., Petrocchi, A., Lichtenegger, J. and Calabresi, G., 2000, Neural Networks for Oil Spill Detection Using ERS-SAR Data. *IEEE Transactions on Geoscience and Remote Sensing,* **38 (5)**, pp. 2282–2287.

[10] Espedal, H.A., 1998, Detection of oil spill and natural film in the marine environment by spaceborne Synthetic Aperture Radar. Ph.D thesis, Department of Physics University of Bergen and Nansen Environment and Remote Sensing Center, Norway.

[11] Fiscella, B., Giancaspro, A., Nirchio, F., Pavese, P. and Trivero, P., 2000, Oil spill detection using marine SAR images. *International Journal of Remote Sensing,* **21 (18)**, pp. 3561–3566.

[12] Girard-Ardhuin, F., Mercier, G. and Garello, R., 2003, Oil slick detection by SAR imagery: potential and limitation. *In Proc. of OCEANS'03,* **1**, pp. 164–169.

[13] Girard-Ardhuin, F., Mercier, G., Collard, F., Garello , R., July 2005, Operational Oil-Slick Characterization by SAR Imagery and Synergistic Data. *IEEE Journal of Oceanic Engineering.* **30 (3)**, pp. 487–495.

[14] Gasull, A., Fabregas, X., Jimenez, J., Marques, F., Moreno, V. and Herrero, M.A., 2002, Oil Spills Detection in SAR Images Using Mathematical Morphology. *In Proc. of EUSIPCO'2002, Toulouse, France, September 2002,* **1**, pp. 25–28.

[15] Hu, M.-K., 1962, Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory,* **8**, pp. 179–187.

[16] Kanaa, T. F. N., Tonye, E., Mercier, G., Onana, V.P., Ngono, J.M., Frison, P.L., Rudant, J.P. and Garello, R., 2003, Detection of Oil Slick Signatures in SAR Images by Fusion of Hysteresis Thresholding Responses. *In Proc. of International Geoscience and Remote Sensing Symposium, 2003,* **4**, pp. 2750–2752.

[17] Keramitsoglou, I., Cartalis, C. and Kiranoudis, C. T., 2006, Automatic identification of oil spills on satellite images. *Environmental Modelling & Software,* **21**, pp. 640–652.

[18] Kubat, M., Holte, R. C. and Matwin, S., 1998, Machine Learning for the Detection of Oil Spills in Satellite Radar Images *Machine Learning,* **30**, pp. 195–215.

[19] Lee, J.S., 1983, Digital Image Smoothing and the Sigma Filter. *Computer Vision, Graphics, and Image Processing,* **24**, pp. 255–269.

[20] Liu, A. K., Peng, C. Y. and Chang, S. Y.-S., 1997, Wavelet Analysis of Satellite Images for Costal Watch. *IEEE Journal of Oceanic Engineering,* **22(1)**, pp. 9–17.

[21] Lobregt, S. and Viergever, M.A., 1995, A discrete dynamic contour model. *IEEE Transactions on Medical Imaging,* **14(1)**, pp. 12–24.

[22] Migliaccio, M., Tranfaglia, M. and Ermakov, A., July 2005, A Physical Approach for the Observation of Oil Spills in SAR Images. *IEEE Journal of Oceanic Engineering,* **30(3)**, pp. 496–507.

[23] Nirchio, F., Sorgente, M., Giancaspro, A., Biamino, W., Parisato, E., Ravera, R. and Trivero, P., 20 March 2005, Automatic detection of oil spills from SAR images. *International Journal of Remote Sensing,* **26(6)**, pp. 1157–1174.

[24] Nixon, M. and Aguado, A., 2002, *Feature Extraction & Image Processing,* Newnes.

[25] Olsen, R.B. and Wahl, T., 2000, The Role of Wide Swath SAR in High-Latitude Coastal Management. *John Hopkins APL Technical Digest,* **21 (1)**, pp. 136–140.

[26] Shepherd, I., Bauna, T., Chesworth, J., Kourti, N., Lemoine, G. and Indregard, M., (2004). Use of ENVISAT at JRC for Marine Monitoring in 2003. JRC technical note. Available Online at:
http://pta.jrc.cec.eu.int/entity.gx/_entity.object/SEC—
000000000002CB39/_entity.name/envisatreport-tn-2.pdf

[27] Solberg, A.H.S. and Solberg, R., 1996, A large-scale evaluation of features for automatic detection of oil spills in ERS SAR images. *In Proc. of Geoscience and Remote Sensing Symposium, 1996*, **3**, pp. 1484–1486.

[28] Solberg, A.H.S., Storvik, G., Solberg, R. and Volden, E., 1999, Automatic detection of oil spills in ERS SAR images. *IEEE Transactions on Geoscience and Remote Sensing,* **37 (4)**, pp. 1916–1924.

[29] Solberg, A.H.S., Brekke, C., Solberg, R. and Husøy, P.O., 2004, Algorithms for oil spill detection in Radarsat and ENVISAT SAR images. *In Proc. of Geoscience and Remote Sensing Symposium, 2004*, **7**, pp. 4909–4912.

[30] Solberg, A., Clayton, P. and Indregard, M., 2005, D2-Report on benchmarking oil spill recognition approaches and best practice. Technical report, Oceanides project, European Commission, Archive No. 04-10225-A-Doc, Contract No: EVK2-CT-2003-00177.

[31] Solberg, A.H.S., Brekke, C. and Husøy, P.O., 2007, Oil spill detection in Radarsat and Envisat SAR images, *IEEE Transactions on Geoscience and Remote Sensing*, **45**(3), 746–755.

[32] Sonka, M., Hlavac, V. and Boyle, R., 1999, *Image Processing, Analysis and Machine Vision,* Brooks/Cole Publishing Company.

[33] Topouzelis, K., Karathanassi, V., Pavlakis, P. and Rokos, D., 2003, Oil spill detection: SAR multi-scale segmentation & object features evaluation. *In Proc. of SPIE, Remote Sensing of the Ocean and Sea ice 2002, 23-27 September, Crete, Greece,* Bostater and Santoleri (Ed.), **4880**, pp. 77–87.

[34] Trier, Ø.D. and Jain, A.K., 1995, Goal-Directed Evaluation of Binarization Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **17 (12)**, pp. 1191–1201.

[35] Wahl, T., Anderssen, T., and Skøelv, Å, 1994, Oil spill detection using satellite based SAR, Pilot Operation Phase, Final Report, Norwegian Defence Research Establishment.

[36] Wahl, T., Skøelv, Å., Pedersen, J. P., Seljelv, L.-G., Andersen, J. H., Follum, O. A., Anderssen, T., Strøm, G. D., Bern, T.-I., Espedal, H. H., Hamnes, H. and Solberg, R., 1996, Radar Satellites: A New Tool for Pollution Monitoring in Costal Waters. *Costal Management,* **24**, pp. 61–71.

[37] Wessel, P. and Smith, W. H. F., 1996, A global self-consistent, hierarcical, high-resolution shoreline database. *Journal of Geophysical Research,* **101 (1)**, pp. 8741–8743.

[38] Wu, S. Y. and Liu, A. K., 2003, Towards an automated ocean feature detection, extraction and classification scheme for SAR imagery. *International Journal of Remote Sensing,* **24 (5)**, pp. 935–951.

[39] Yang, L. and Albregtsen, F., 1996, Fast and Exact Computation of Cartesian Geometric Moments Using Discrete Green's Theorem. *Pattern Recognition,* **29 (7)**, pp. 1061–1073.

Chapter 13

Paper IV: Oil Spill Detection in Northern European Waters: Approaches and Algorithms

Anne H. S. Solberg and Camilla Brekke

To appear as a chapter in: *Remote Sensing of the European Seas, (Vittorio Barale and Martin Gade (Eds.)), Springer Science and Business Media B.V., 2007.*

Oil Spill Detection in Northern European Waters: Approaches and Algorithms

Anne H. S. Solberg [1,2] and Camilla Brekke [1,3]
[1]Department of Informatics, University of Oslo. P.O. Box 1080 Blindern, 0316 Oslo
E-mail: anne@ifi.uio.no
[2]also at Norwegian Computing Center
[3]also at Norwegian Defence Research Establishment

Abstract

The combined use of satellite-based synthetic aperture radar (SAR) images and aircraft surveillance flights is a cost-effective way to monitor deliberate oil spills in large ocean areas and catch the polluters. SAR images enable covering large areas, but aircraft observations are needed to prosecute the polluter, and in certain cases to verify the oil spill. We discuss the limitations of satellite imaging of oil spills compared to aircraft monitoring. Automatic detection of oil spills has proven to be an interesting complement to manual detection. We present an overview of algorithms for automatic detection, and discuss their potential compared to manual inspection as part of an operational oil spill detection framework. Experimental results show that automatic algorithms can perform comparable to manual detection, both in terms of accuracy in detecting verified oil slicks, false alarm ratio, and they can also speed up the image analysis process compared to fully manual services.

1 Introduction

Marine pollution arising from illegal oily discharges from ships represents a serious threat to the marine environment. Oil pollution caused by large accidents like the Prestige event in 2002 capture many headlines, but the majority of the oil pollution cases are caused by operational discharges from tankers. Observed oil spills commonly appear in connection with off-shore installations and correlate well with major shipping routes. A combination of aircraft and satellite sensors are currently used to monitor large ocean areas to detect oil spills and catch the polluter. The inclusion of satellite surveillance allows the user to better target the aircrafts used for oil spill surveillance and to cover larger areas.

2 Remote Sensing Sensors for Oil Spill Detection

For routine monitoring of illegal oil discharges from ships and offshore installations both aircraft sensors and satellite sensors can be used. Satellite-based SAR images can be used to screen large ocean areas, while aircrafts are more suitable to be brought into action to identify the polluter, the extent, and the type of spill.

2.1 Aircraft Sensors

Most surveillance aircrafts used for oil pollution monitoring in Europe are equipped with a combination of sensors (Side-Looking Airborne Radar (SLAR), infrared/ultraviolet (IR/UV), Laser-Fluoro-Sensor (LFS) , Microwave Radiometer (MWR)). For an overview of aircraft sensors for oil spill detection, see [10; 19]. SLAR is the main sensor for long-range detection of oil pollution on the sea surface. The SLAR is used to locate possible spill locations. Then the spill is inspected more closely using additional sensors and/or visual inspection. The sensor configuration used on board surveillance aircrafts varies from country to country. An example is the German aerial surveillance, which locates the oil spills by SLAR, IR/UV scanning is used to quantify the extent of the film, a MWR is used to quantify the thickness, and a LFS is used for oil type classification. The SLAR, IR, and LFS can operate at night. A number of different aircraft types are used for oil spill aerial surveillance. They differ in terms of endurance, cruising speed and SLAR sensor equipment resulting in different SLAR area coverage during one flight hour. One hour of airborne remote sensing over the sea at a speed of 335 km/h covers an area of 13400 km^2 [20].

2.2 Satellite Sensors

SAR is the main spaceborne remote sensing instrument for oil spill imaging, with all-weather and all-day operation capabilities, although it is not capable of oil spill thickness estimation and oil type recognition. The main limitation for spaceborne optical sensors is the need for

daylight and cloud-free scenes, but they have a potential to discriminate between oil and algal blooms. A more detailed discussion of other satellite sensors for oil spill detection is given in [3].

SAR is particularly useful for searching large areas. Usually even small volumes of oil cover large areas and thus the need for very high spatial resolution in SAR images is not crucial. SAR has however some limitations, as a number of natural phenomena can produce similar dark objects in the SAR images (see Section 3).

Currently, RADARSAT-1 and ENVISAT ASAR are the two main SAR sensors used for oil spill detection. The best trade-off between spatial coverage and spatial resolution is achieved using RADARSAT-1 ScanSAR and ENVISAT ASAR Wide Swath image modes. Table 1.1 describes the coverage and resolution of these sensors. The costs of satellite images are much lower than the costs of covering the same area by aircraft. The actual time that the satellite passes over a given location will vary with latitude, but the overflight will be fixed in time.

Table 1.1: Coverage and resolution for selected RADARSAT-1 and ENVISAT ASAR products.

	RADARSAT-1 ScanSAR Narrow	ENVISAT ASAR Wide Swath
Spatial coverage per scene	300 km × 300 km	400 km × 400 km
Spatial resolution per scene	50 m × 50 m	150 m × 150 m

The number of available RADARSAT-1 or ENVISAT ASAR images of a given area on a given date depends on the geographic latitude and the observation period. Daily coverage is possible in Northern Europe, and a coverage a couple of times a week is possible for all European waters.

2.3 Satellite vs. Aircraft - Advantages/Limitations

The advantages and limitations of satellite-based vs. aircraft monitoring are summarized in Table 1.2. In order to cover the same area as a RADARSAT-1 ScanSAR Narrow scene or an ENVISAT ASAR Wide Swath scene with an aircraft, 6 or 12 flight hours, respectively, are needed. A limitation with the satellite monitoring is that the images are taken at fixed times of the day. The fate and persistence of oil in seawater are controlled by processes that vary considerably in space and time. The amount of oil spilled, its initial physical and chemical characteristics, and the prevailing climatic and sea conditions have great impact on the lifetime of an oil spill. A reasonable assumption might be that most illegal oil discharges are bilge oil (i.e. a mixture of several kinds of oils (fuel oil, hydraulic oil, etc.)). For instance, in the Finnish surveillance area it is estimated that 1-5 percent of the detected slicks are

thicker slicks that persist several days, and 95-99% are bilge oil that persists only some hours [20]. To cover oil spills occurring at all times of the day, aerial surveillance can be used to supplement the fixed coverage times of the SAR satellites.

Table 1.2: Advantages and drawbacks with SAR satellite and aerial surveillance. (Adapted from [20]).

Satellite SAR	**Aerial surveillance**
Advantages	**Advantages**
Large and well-defined spatial coverage.	Flexible monitoring.
Less expensive than airborne surveillance.	High accuracy of oil spill detection.
Can be used to cue aircraft to improve	Can be deployed at short notice.
aircraft operational efficiency.	Can identify polluter.
	Can identify additional oil parameters.
Limitations	**Limitations**
False targets can occur in analysis.	High cost.
Fixed monitoring schedule.	Smaller spatial coverage.
Limited to certain wind conditions.	Limited to certain weather conditions.

3 SAR Imaging of Oil Spills

Oil slicks dampen the Bragg waves (wavelength of a few cm) on the ocean surface and reduce the radar backscatter coefficient. This results in dark regions or dark spots in a satellite SAR image. A part of the oil spill detection problem is to distinguish oil slicks from other natural phenomena that dampen the short waves and create dark patches on the surface. Natural dark patches are termed oil slicks look-alikes. Oil slicks include all oil related surface films caused by oil spills from oilrigs, leaking pipelines, passing vessels as well as bottom seepages, while look-alikes include natural films/slicks, grease ice, threshold wind speed areas, wind sheltering by land, rain cells, shear zones, internal waves etc.

A service for oil spill detection based on SAR images must contain an oil spill detection step where the SAR images are analyzed, and dark regions that might be oil spills are identified. In this process, the factors that can be used to discriminate between an oil spill and a look-alike are important. Due to higher viscosity, oil spills tend to remain more concentrated and provide larger damping to the surrounding sea than natural films [11]. A newly released oil spill will have reasonably sharp borders to the surrounding sea. As the weathering effects the spill, the borders can become more fuzzy. The shape will be altered by wind and current. Depending on the source of the outlet, certain shapes of the oil spills can be expected. Oil slicks from moving ships are thin, linear or piecewise slicks, while oil slicks

from stationary sources can be wide and regular if a significant amount of oil is released in short time.

Other types of pollution can also cause slicks that are visible in the SAR image Algae can also create dark patches in the SAR image very similar to oil slicks. This is in particular a problem in the Baltic Sea, where a certain algae type that dampens the Bragg waves is common during the summer season. If additional information from e.g. ocean color sensors is available, it can be used to identify algae.

The wind speed is also important for imaging of oil spills. With very low wind, no backscatter from the sea surface will be seen. Look-alikes are very frequently observed in low to moderate wind conditions (approximately 3 to 7 m/s). As the wind speed increases, the expected number of look-alikes will be lower. For oil spills, the contrast between the spill and the surrounding sea will decrease with higher wind speeds. At high wind speeds (>10 m/s) only larger slicks with thicker oil will be visible. The upper limit for observing oil in the SAR image is not known exactly. In an operational oil spill detection service at Kongsberg Satellite Services (KSAT) in Tromsø, Norway, an upper limit of 15 m/s is used.

4 SAR Oil Spill Detection: Manual vs. Automatic

Oil spills in a SAR image can be identified by manual inspection, or the image can first be screened by an automatic algorithm for oil spill detection, followed by manual inspection of the suspect alarms only. Manual interpretation of e.g. a 400 km × 400 km ENVISAT ASAR image can be a complex and time consuming task because the image is so large that the operator can only view a small part of the scene at a time to be able to detect thin oil slicks. Recent benchmarks [15; 14] comparing automatic algorithms to manual detection shows that an automatic algorithm can be a valuable tool when a large number of images are to be inspected.

4.1 Manual Oil Spill Detection

A well-established operational service for oil spill detection is run at KSAT. Trained operators detect oil spills by inspecting the SAR images. In addition to the image, they can use external information about wind speed and direction, oil rig/pipeline location, national territory borders and coast lines. After a possible oil slick has been detected, it is assigned confidence level low, medium or high based on a certain set of rules [15]. The location of detected oil spills and their confidence level is then immediately sent to the surveillance aircrafts.

4.2 Automatic Approaches

A literature review of automatic techniques for oil spill detection in SAR images can be found in Brekke and Solberg [3]. In this section we give a short update on the state-of-the-art in

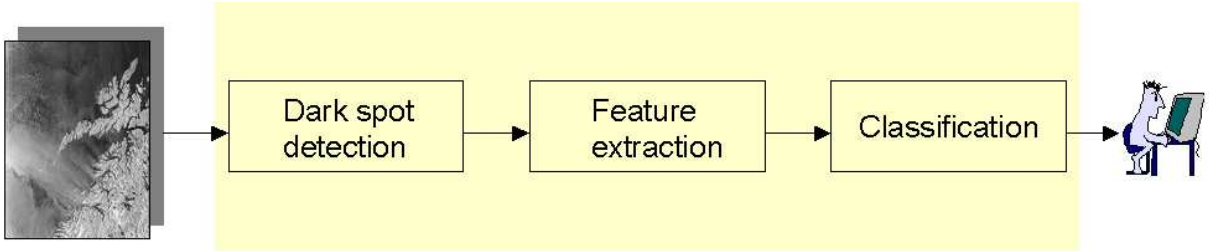


Figure 1.1: A framework for oil spill detection algorithms.

this field.

Several of the published papers on oil spill algorithms for SAR images (e.g. [8; 7; 17; 14]) describe a methodology consisting of dark spot detection followed by feature extraction and a classification step (see Fig. 1.1).

Segmentation Techniques

As oil spills are characterized by low backscattering levels, the use of thresholding for dark spot segmentation is commonly applied (see e.g. Keramitsoglou et al. [12]; Solberg et al. [17]; Brekke and Solberg [3]). As SAR images tend to become darker with increasing range and as local variations in the wind level and other meteorological and oceanic conditions occur, thresholding algorithms where the threshold is set adaptively based on local statistical estimates should be preferred.

Dark Spot Feature Extraction

Discrimination between oil spills and look-alikes are often based on a number of features computed for each suspicious dark spot on the sea surface (see e.g. [18; 17; 5; 7; 8]).

Good features are very important for the sucess of the following classification step. Most of the features applied in the literature are covered by the following types:

- The geometry and shape of the dark spot.
- Physical characteristics of the backscatter level of the dark spot and its surroundings.
- Dark spot contextual features.
- Texture features of both the dark spot and the surroundings.

Slick Classification

The purpose of the classification step is to distinguish oil slicks from look-alikes. How difficult the classification task is depends on the variability in the feature values for objects in the

oil spill class relative to the difference between feature values for objects in the look-alike class. Effective methods for developing classifiers involve learning from example patterns (*training*). In statistical approaches, the classification decision is based on the probability and the cost of a certain decision (e.g. [17; 6; 4]).

Various classifiers have been applied to the oil spill detection problem [8; 13; 7; 12]. All detection algorithms suffer from false alarms, and dark spots classified as oil spills may be confused with look-alikes (e.g. natural film and low wind areas). Applying external data to improve classification and assess the slick nature has been suggested. Girard-Ardhuin et al. [9] combines characteristics of the detected dark spots from the SAR images and meteorological and oceanic data through a multisensor approach (including information about surface wind measurements, sea-surface temperature, atmospheric fronts and clouds and chlorophyll).

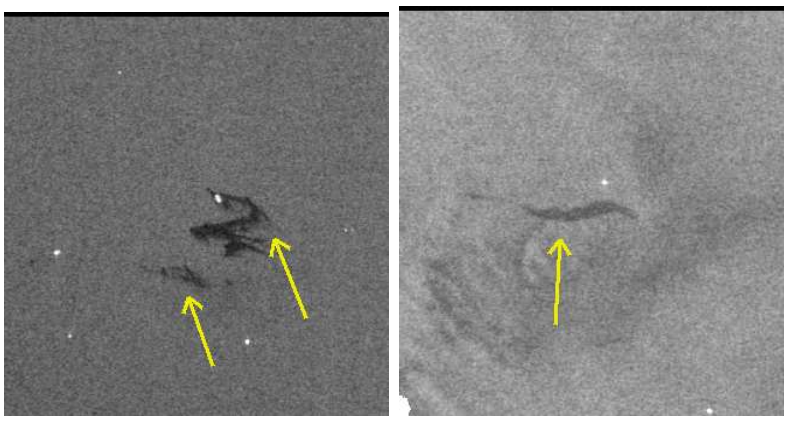
5 A Benchmark Study of Oil Spill Detection Approaches

As part of the EC project Oceanides, a benchmark study comparing oil spill recognition approaches was performed. Manual oil spill detection based on SAR images was compared to semi-automatic and automatic approaches. A joint satellite-airborne campaign was performed during 2003. The campaign covered the Finnish and German sectors of the Baltic Sea, in addition to the German sector of the North Sea. The campaign was organized in such a way that a trained operator at KSAT (KSAT1) analyzed the SAR images, and reported possible oil spills to the Finnish and German pollution control authorities. They would check the positions and verify the slicks, and report additional slicks found by the aircraft. This was done for both ENVISAT and RADARSAT-1 images.

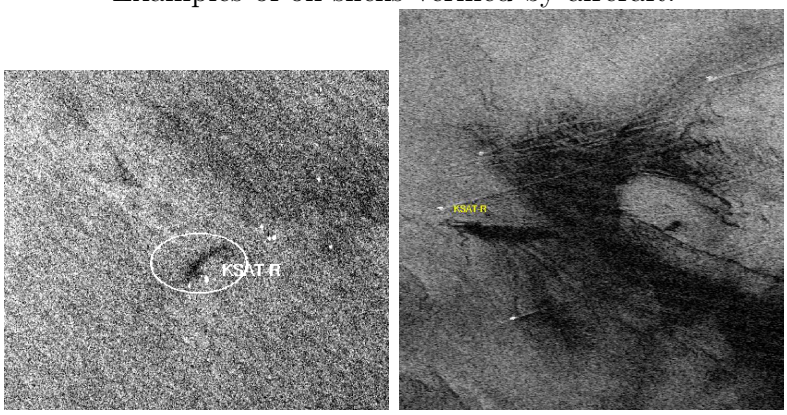
For benchmark comparisons, KSAT let another operator (KSAT2) inspect the same SAR images without knowing the aircraft detections or the result of the previous inspection to study the inter-operator variance. The automatic oil spill detection approach developed at Norwegian Computing Center (NR) [14] was used to analyze all images without knowing aircraft detections or KSAT results. Fig. 1.2 shows examples of both correctly classified oil slicks, false alarms, and slicks detected only by aircraft.

The benchmark data set contained 27 ENVISAT images and 32 RADARSAT-1 images. The real-time inspection of the ENVISAT images at KSAT (KSAT1) detected 11 oil slicks that were verified as oil slicks by the aircraft. The repeated inspection by another operator, KSAT2, detected 8 of these verified slicks, the automatic algorithm (NR) also detected 8 of the verified slicks. For RADARSAT-1 data, there was 18 verified oil slicks, KSAT2 found 15 of these, while the NR algorithm found 14.

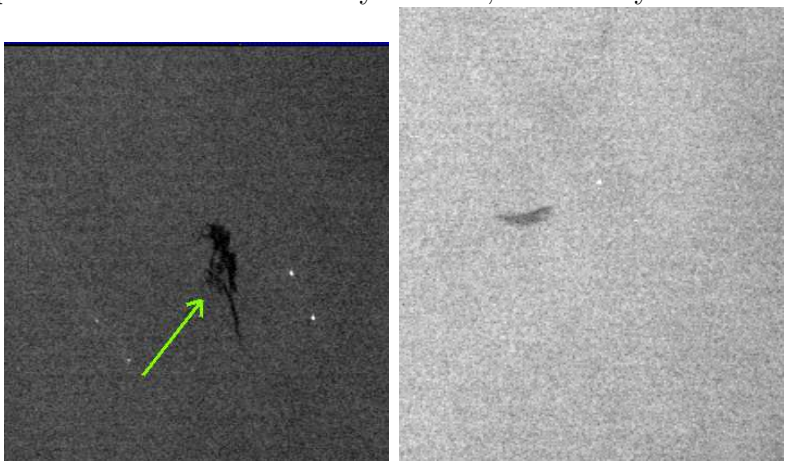
This demonstrates that the inter-operator variance was significant. KSAT has later taken measures to reduce this variability by increasing operator training and harmonising the interpretation process. The performance of the NR algorithm is also almost comparable to KSAT2, so it can be a valuable alternative or supplement to manual inspection.



Examples of oil slicks verified by aircraft.



Examples of oil slicks detected by aircraft, but not by KSAT and NR.



Examples of false alarms. The slick in the image on the left was verified as algae, while the clearly suspect dark spot in the right image could not be found when the aircraft inspected the position (possibly resolved).

Figure 1.2: ENVISAT images with examples of oil slicks and false alarms. ©ESA/KSAT/NR

It is also of interest to study additional slicks reported by the aircraft, but not detected from satellite image analysis. In general, these slicks involved a small amount of oil. For some of them, the time of aircraft pass was several hours after the satellite image acquisition, so the release could be new. For other cases, the satellite image was taken several hours after the air-craft pass, and a small amount of oil could very well be resolved during this time period.

The number of false alarms, the number of slicks reported by satellite, but verified as not oil by the aircraft, was also studied. All false alarms are discussed in detail in [15]. Some linear slicks with good contrast were detected in the satellite image but verified as algae.

The confidence levels that are associated with reported possible oil slicks from KSAT can be used by the surveillance aircraft crew to prioritize which oil spill positions they will inspect first. Slicks that were assigned confidence level High had very low false alarm ratio, slicks that were assigned confidence Medium had reasonably low false alarm ratio, while slicks that were assigned confidence Low had relatively high false alarm ratio (again the details can be found in [15]). However, the confidence assigned by two different operators at KSAT varied. The automatic algorithm can also be used to compute confidence levels using a set of rules that simulate the rules used by the operators at KSAT (see [16] for how this is done). By comparing the confidence levels assigned by two KSAT operators and NR's automatic algorithm on a set of 22 selected oil slicks, it was found that KSAT1 and KSAT2 assigned the same confidence for 7 out of 22 slicks, while the algorithm and KSAT1 agreed for 13 of 22 slicks, and had a confidence level difference of one (indicating e.g. High vs. Medium) for six additional slicks. What this indicates is that there is still some subjectivity involved in confidence assignment, and using an algorithm to get a "second opinion" to use in confidence assignment might be valuable.

The processing time for manual inspection at KSAT varied between 3 and 25 minutes for RADARSAT-1 images, with an average of 9 minutes. The automatic algorithm had an average processing time of 3 minutes. For ENVISAT images, the average time for manual inspection at KSAT was 10 minutes, while the automatic algorithm had an average processing time of 1.45 minutes.

6 Discussion and Conclusions

The combined use of satellite-based SAR images and aircraft surveillance flights is a cost-effective way to monitor large areas and catch the polluters. The coverage in terms of the number of weekly satellite passes of European waters is very good in Northern Europe and decent in the Mediterranean.

Oil spill detections from aircrafts and satellite images were compared in a benchmark study. In general, there were good agreement between aircraft detections and satellite-based detections when the time offset between the image acquisitions was low.

Some information about oil spill statistics and hot spots exist, see e.g. [2; 20]. Hot spots

coincide well with major shipping routes, pipelines, oil rigs etc. Tufte et al. [20] discuss sampling requirements as guidelines for operational oil spill monitoring in European waters. They also summarize current monitoring efforts for many countries in Northern Europe, and their experience in using a combination of aerial- and satellite monitoring. The best approach for a specific national authority depends on the size and shape of the area to be monitored, and other resources available. International cooperation with neighboring countries on planning satellite acquisitions and sharing costs is important.

Many different techniques are proposed for automatic detection of oil spills. They often consist of three main parts: segmentation, feature extraction and classification. However, there is still a challenge in reducing the number of false alarms, and automatically assigned confidence levels can be helpful in prioritizing the alarms. Interoperator variance in manual detection should be reduced by better training, or introducing the algorithm as a second information source. This applies to both detection and confidence assignment.

Presently, one of the biggest challenges for operational oil spill detection services is obtaining sustainability in terms of data availability. SENTINEL-1 is a planned two-satellite system (C-band) to be operated as a constellation for maximized coverage/repeat cycle. The first satellite will be launched in 2010 and the second some 12-15 months later [1]. RADARSAT-2 is a Canadian SAR satellite planned to be launched in the summer of 2007. The SAR instrument will be C-band like RADARSAT-1, but there will be more flexibility in the selection of polarizations.

Acknowledgments

The authors would like to thank the Oceanides project, in particular Marte Indregard, Peter Clayton and Lars Tufte for contributions to the benchmark study on oil spill approaches.

References

[1] E. Attema. Mission Requirements Document for the European Radar Observatory Sentinel-1, Requirement Specification. Technical Report ES-RS-ESA-SY-0007, issue 1, revision 4, European Space Agency, 2005.

[2] T. Bauna and P. Clayton. D3c-state of knowledge of potential oil spill hotspots in european waters. Technical report, Oceanides project, European Commission, Archive No. 04-10225-A-Doc, Contract No: EVK2-CT-2003-00177, 2004.

[3] C. Brekke and A. Solberg. Oil spill detection by satellite remote sensing. *Remote Sensing of Environment*, 95(1):–13, 2005.

[4] C. Brekke, A. Solberg, and G. Storvik. Classifying Oil Spills and Look-alikes in ENVISAT ASAR Images. In *(to appear) Proc. of ENVISAT symposium, Montreux, Switzerland*, 23-27 April 2007.

[5] C. Brekke and A. H. S. Solberg. Segmentation and feature extraction for oil spill detection in ENVISAT ASAR images. Submitted: *Int. Journal of Remote Sensing*, 2006.

[6] C. Brekke and A. H. S. Solberg. Classifiers and Confidence Estimation for Oil Spill Detection in ENVISAT ASAR Images. Submitted: *IEEE Trans. on Geoscience and Remote Sensing*, 2007.

[7] F. Del Frate, A. Petrocchi, J. Lichtenegger, and G. Calabresi. Neural networks for oil spill detection using ERS-SAR data. *IEEE Trans. on Geoscience and Remote Sensing*, 38(5):2282–2287, 2000.

[8] B. Fiscella, A. Giancaspro, F. Nirchio, P. Pavese, and P. Trivero. Oil spill detection using marine SAR images. *Int. J. Remote Sensing*, 21(18):3561–3566, 2000.

[9] F. Girard-Ardhuin, G. Mercier, F. Collard, and R. Garello. Operational Oil-Slick Characterization by SAR Imagery and Synergistic Data. *IEEE Journal of Oceanic Engineering*, 30(3), 2005.

[10] R. Goodman. Overview and future trends in oil spill remote sensing. Spill Science & Technology Bulletin, 1(1):11–21, 1994.

[11] H. A. Hovland, J. A. Johannessen, and G. Digranes. Slick detection in SAR images. *Proc. IGARSS'94*, 4:2038–2040, 1994.

[12] I. Keramitsoglou, C. Cartalis, and C. Kiranoudis. Automatic identification of oil spills on satellite images. *Environmental Modelling & Software*, 21(5):640–652, 2006.

[13] F. Nirchio, M. Sorgente, A. Giancaspro, W. Biamino, E. Parisato, R. Ravera, and P. Trivero. Automatic detection of oil spills from SAR images. *Int. Journal of Remote Sensing*, 26(6):1157–1174, 2005.

[14] A. Solberg, C. Brekke, and P. Husøy. Oil spill detection in Radarsat and Envisat SAR images. *IEEE Trans. on Geoscience and Remote Sensing*, 45(2):746–755, 2006.

[15] A. Solberg, P. Clayton, and M. Indregard. D2-report on benchmarking oil spill recognition approaches and best practice. Technical report, Oceanides project, European Commission, Archive No. 04-10225-A-Doc, Contract No: EVK2-CT-2003-00177, 2005.

[16] A. H. S. Solberg. Automatic oil spill detection and confidence estimation. In *Proc. International Symposium Remote Sensing of Environment*, pages 943–945, St. Petersburg, Russia, 2005.

[17] A. H. S. Solberg, G. Storvik, R. Solberg, and E. Volden. Automatic detection of oil spills in ERS SAR images. *IEEE Trans. on Geoscience and Remote Sensing*, 37(4): 1916–1924, 1999.

[18] K. Topouzelis, V. Karathanassi, P. Pavlakis, and D. Rokos. Oil spill detection: SAR multi-scale segmentation & object features evaluation. In *Proc. of SPIE, Remote Sensing of the Ocean and Sea Ice*, series 4880:77–87, 2002.

[19] O. Trieschmann, Th. Hunsänger, L. Tufte, and U. Barjenbruch. Data assimilation of an airborne multiple remote sensor system and of satellite images for the North- and Baltic sea. *Proceedings of the SPIE 10th Int. Symposium on Remote Sensing, conference "Remote Sensing of the Ocean and Sea Ice 2003"*, pages 51–60, 2003.

[20] L. Tufte, M. Indregard, A. Solberg, and R. B. Huseby. D11-oil spill monitoring guidelines and methodology to combine satellite oil spill recognition with routine airborne surveillance. Technical report, Oceanides project, European Commission, Archive No. 04-10225-A-Doc, Contract No: EVK2-CT-2003-00177, 2005.

Chapter 14

Paper V: Classifiers and Confidence Estimation for Oil Spill Detection in ENVISAT ASAR Images

Camilla Brekke and Anne H. S. Solberg

To appear: *IEEE Geoscience and Remote Sensing Letters*,
vol. 05, no. 1, January 2008.

Classifiers and Confidence Estimation for Oil Spill Detection in ENVISAT ASAR Images

Camilla Brekke
is with the Norwegian Defence Research Establishment, PO Box 25, NO-2027 Kjeller, Norway
and with the Department of Informatics, University of Oslo
and
Anne H. S. Solberg
is with Department of Informatics, University of Oslo
and with the Norwegian Computing Center

Abstract

An improved classification approach is proposed for automatic oil spill detection in SAR images. The performance of statistical classifiers and Support Vector Machines (SVM) are compared. Regularized statistical classifiers prove to perform the best on this problem. To allow the user to tune the system with respect to the trade-off between the number of true positive alarms and the number of false positives, an automatic confidence estimator has been developed. Combining the regularized classifier with confidence estimation leads to acceptable performance.

Index Terms– SAR, oil spill, classification.

1 Introduction

Oil spills appear as dark areas in SAR images because the oil damps the short gravity waves on the sea surface. A part of the oil spill detection problem is to distinguish oil spills from other low-backscatter ocean phenomena (*look-alikes*) creating dark structures in the images. The framework of our algorithm is a dark spot detector, a feature extractor and a dark spot classifier. Dark spots in the images are primarily detected by adaptive thresholding. For each of them a number of features are computed in order to classify the slick as either an oil spill or a look-alike. Segmentation and feature selection were discussed in [1; 12], here we focus on the classification step.

Various classifiers have been applied to the oil spill detection problem. A Mahalanobis classifier and a compound probability classifier was applied in [5]. The probabilistic approach was improved in [9] by using a different classification algorithm including a multi-regression approach where the regression coefficients represent the contribution of each feature on the prediction of a new sample. A neural-network approach is described in [6], and in [7] a classifier based on fuzzy logic is presented. In [8] a SVM technique for oil spills detection was applied.

Oil spill detection is an application where the classifier should detect a rare but important event (look-alikes appear much more frequent). Based on a statistical classifier for oil spill detection earlier described [12], we here suggest introducing regularization of the covariance matrices to decrease the number of false positives. SVMs has been used for a number of applications often showing good performance [2]. We compare SVM to the regularized statistical classifier. Confidence levels automatically assigned to the slicks can be useful for prioritising the alarms and helpful as a "second opinion" to manual analysis [11]. We present a novel algorithm for automatic assignment of confidence levels.

2 SAR Images

Our system is trained and tested on 103 complete ENVISAT ASAR Wide Swath Mode images. The image set is split into three sets: training (56 images), validation (20 images) and testing (27 images). The images are collected from the Baltic Sea and the North Sea from 2003-2005. Aircraft verifications, from a joint airborne-satellite campaign [10], are available for the test set. To label the training and validation sets, oil slick candidates were manually masked and associated categories "high-confidence" and "moderate-confidence" (this information was used in the training of the confidence estimator, section 3.3). For more details on the masking process see [12].

3 Classification Methodology

For each of the segmented dark spots i, a number of features (selected in previous studies [1; 13]) are computed: **Shape features:** "Sum of External Angles", "Slick Moment", "Slick Area", "Slick Complexity", "Slick Width", **Contrast features:** "Slick Local Contrast", "Slick Border", "Smoothness Contrast", **Texture:** "Power-to-mean Ratio", "Slick Variance" and **Surroundings:** "Regions in Small Neighbourhood", "Distance to Ship/Oilrig". In addition, a new feature "Low Wind Area" is developed. It is defined as a binary feature, assigned 1 if the region overlaps with a large region (≥ 4000 pixels, i.e. larger than the largest oil spill observation among the training samples) in a coarse segmented version of the image where large dark structures at the size of large low wind areas will be detected.

The simplest classifier to consider is to use a multivariate Gaussian classifier. However, the behaviour of the features will change with different wind levels. Describing the conditional density by a unimodal density such as a Gaussian is therefore not appropriate. If we split the problem and assume different densities depending on the wind level, we can train a classifier for each subclass. The data within each subdivision would then be reasonably homogeneous and it might be appropriate to assume the densities to be Gaussian.

The class-conditional multivariate normal density is completely specified by $d+d(d+1)/2$ parameters, namely the elements of the class-specific mean vector μ_k and the independent elements of the covariance matrix Σ_k. d is the dimensionality of the feature vector. As oil spills just occasionally occur, we need to consider robust estimation of Σ_k due to the low number of training samples. We suggest regularizing the estimated covariance matrix, by forming a combination of the fully estimated covariance matrix and the diagonal covariance matrix.

Oil spill detection based on SAR is today used in combination with surveillance aircraft in many European countries. In addition to the cost of sending out aircraft, the false alarm ratio acceptable by the customers could vary. It would be desirable that the automatic system not only reports which slicks were classified as oil, but also the probability associated with that decision. Using the computed posterior probability for this purpose would be the first option to consider. However, initial experiments indicated that this did not perform well. The posterior probability for a two-class Gaussian tends to be very close to either 1 or 0. Comparing probabilities across subclasses also turned out to be difficult because of the difference in variance between the subclasses. To reduce the number of false positives and to give the operator a tool for prioritising the alarms, we have developed an automatic confidence estimator for all spots labelled oil spill by the classifier. First, a classifier is applied to each feature vector $\mathbf{x}_i$, then for each slick with a higher posterior probability of being an oil spill than a look-alike ($Pr(oil\ spill|\mathbf{x}_i) > Pr(look-alike|\mathbf{x}_i)$) a confidence level is automatically estimated.

3.1 Statistical Classifier

In [13; 12], a prior distribution and a probability density for the features are combined through Bayes theorem to obtain the *posterior* probability for a detected spot being an oil spill. Let c be the unknown class membership of a detected spot (we are dealing with a two class problem: oil spill or look-alike). Then

$$Pr(c = o|\mathbf{x}_i) = \frac{\pi_o f_o(\mathbf{x}_i)}{\pi_o f_o(\mathbf{x}_i) + (1 - \pi_o) f_l(\mathbf{x}_i)} \tag{1.1}$$

where π_o is the prior model for the probability that a detected spot is oil. $f_o(\mathbf{x}_i)$ and $f_l(\mathbf{x}_i)$ are the probability densities for the observed features $\mathbf{x}_i$ in classes o = oil spills and l = look-alikes, respectively. The densities are assumed Gaussian:

$$f_c(\mathbf{x}_i) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma|^{\frac{1}{2}}} \times exp\{-\frac{1}{2}(\mathbf{x}_i - \mu_c)^T \Sigma^{-1}(\mathbf{x}_i - \mu_c)\} \tag{1.2}$$

where $c \in \{o, l\}$, d is the number of features, μ_c is the mean vector for class c and Σ is the covariance matrix, common for both classes due to the imbalanced data set. Only "Slick Complexity", "Power-to-mean Ratio", "Slick Local Contrast", "Slick Width", "Regions in Small Neighbourhood", "Slick Border", "Smoothness Contrast" and "Slick Variance" are included in feature vector $\mathbf{x}_i$.

Division of Each Class Into a Set of Subclasses

Even within each wind level W, both the oil slicks and the look-alikes may vary quite a lot in shape, contrast and other features. The wind level W is first used to divide the samples in two different subclasses and then these are divided into five subclasses based on the shape descriptor "Slick Moment (MOM)" (see Table 1.1). The wind level is represented by a homogeneity category $\in \{1, ..., 6\}$ estimated based on the "Power-to-mean Ratio (PMR)" of the surroundings [12]. The division of the feature space might not be optimal, and is the subject of ongoing research.

Covariance Matrix Estimation

With the subclass division, the density for class c and wind level W is then given by $f_{c,W,g}(\mathbf{x}_i)$ if "Slick Moment" is in shape subgroup g. Different densities depending on the value of "Slick Moment" are assumed. Applying Gaussian densities, the simplest classifier consists of using common diagonal covariance matrices for each subclass: $\hat{\Sigma}_{W,g} = diag\{\Sigma_{W,g}\}$ where $\Sigma_{W,g}$ is the fully estimated common covariance matrix. Replacing $\hat{\Sigma}_{W,g}$ with a regularized covariance matrix leads to a more general family of covariances indexed by $\rho_{W,g}$:

$$\widetilde{\Sigma}_{W,g}(\rho_{W,g}) = \rho_{W,g}[diag\{\Sigma_{W,g}\}] + (1 - \rho_{W,g})\Sigma_{W,g} \tag{1.3}$$

Table 1.1: Subclasses and the number of observations in the training set.

Subclass	MOM	PMR	Oil Spills	Look-alikes
1	< 0.3	$>=0.04$	37	5267
2	$\in [0.3, 0.5)$	$>=0.04$	34	4202
3	$\in [0.5, 0.8)$	$>=0.04$	23	1681
4	$\in [0.8, 1.2)$	$>=0.04$	12	617
5	$>=1.2$	$>=0.04$	16	278
6	<0.3	< 0.04	23	786
7	$\in [0.3, 0.5)$	< 0.04	25	604
8	$\in [0.5, 0.8)$	< 0.04	14	152
9	$\in [0.8, 1.2)$	< 0.04	9	51
10	$>=1.2$	< 0.04	14	27

Here $\rho_{W,g} \in [0, 1]$ allows a continuum of models, and needs to be specified through experiments. If class-dependent covariance matrices are used, the variance of the look-alike class could be huge compared to the oil class, and the resulting class-conditional probabilities for the look-alikes would be of another magnitude than for oil resulting in a large "bias" in the classification. A way of handling the very unbalanced data set is to avoid using class-conditional covariance matrices. However, how imbalanced the training set is varies among the subclasses. If we use class-conditional covariance matrices within each subclass, regularization can be expressed as:

$$\widetilde{\Sigma}_{c,W,g}(\rho_{W,g}) = \rho_{W,g}[diag\{\Sigma_{c,W,g}\}] + (1 - \rho_{W,g})\Sigma_{c,W,g} \tag{1.4}$$

where $c \in \{o, l\}$ and $\Sigma_{c,W,g}$ is the fully estimated class-conditional covariance matrix.

Loss Functions and Prior Probabilities

We are considering misclassifying oil spills as look-alikes more serious than misclassifying look-alikes as oil spills. A spot is classified as oil if:

$$\frac{f_l(\mathbf{x}_i)}{f_o(\mathbf{x}_i)} < \frac{\pi_o l_1}{(1 - \pi_o) l_2} \tag{1.5}$$

Equal prior probabilities for the oil spill and the look-alike classes ($\pi_o = \pi_l = 0.5$) are applied, and $l_1 = \beta 0.6$ and $l_2 = 0.4$. $\pi_o = 0.5$ is far from realistic, as an oil spill is considered a rare event compared to look-alikes, however π_o is adjusted in eq. (1.5) through β which is based on a model presented in [13; 12] for the expected number of look-alikes and oil spills depending on the wind level. In total, this gives a reasonably realistic model for the imbalanced class problem.

Training the Regularized Classifier

A search for the optimal regularization parameter, ρ, and experiments to compare regularization with common vs. class-conditional covariance matrices have been done. As some of the subclasses have a quite limited number of training samples, we chose to use cross-validation with *leave-one-out* on the training set to identify optimal values for ρ within each subclass. The validation set was used to compare the performance of eq. (1.3) and (1.4). The result is presented in table 1.2. To study the effect of regularization on the classifier (section 4),

Table 1.2: The optimal selection $\widetilde{\Sigma}$ and ρ for each subclass.

Subclass	ρ	$\widetilde{\Sigma}$
1	0	common
2	0	common
3	0.1	common
4	0	class-conditional
5	0	class-conditional
6	0.1	common
7	0	common
8	0.1	common
9	0.8	common
10	1	common

two versions of the statistical classifier were trained: a classifier applying common, diagonal covariance matrices as in [12], and a regularized classifier. Our regularization technique adapts to the subclasses, and is likely to generalize better than diagonal covariance matrices. For the small subclasses 9 and 10 the optimal ρ was found to be close to 1 (giving diagonal covariance matrices). The results also showed that for subclass 1-8 (a majority of the subclasses) the optimal ρ is close to 0 which means that the classifier performs best using a non-diagonal covariance matrix. We found that class-conditional (compared to common) covariance matrices give a better result on the validation set only for subclasses 4 and 5 (which are the two least imbalanced subclasses among those with highest PMR).

3.2 Support Vector Machines

As the training set is imbalanced, we have chosen to use C-Support Vector Classification (C-SVC) for the special case where different penalty parameters are used for the oil spill and look-alike classes. We have applied an implementation by Chang and Lin [3]. In regular C-SVC [4], a common penalty parameter C is applied for both the classes.

Given a training set of feature vector - label pairs $(\mathbf{x}_i, y_i), i = 1, 2, ..., N$ where $\mathbf{x}_i \in R^n$ and $y_i \in \{1, -1\}$ (note that we here are dealing with a two class approach since we are considering a pair of classes, oil spill or look-alike), C-SVC requires training that involves the minimization of the error function:

$$\min_{\mathbf{w},b,\xi} \frac{1}{2}\mathbf{w}^T\mathbf{w} + C_1 \sum_{y_i=1} \xi_i + C_{-1} \sum_{y_i=-1} \xi_i \tag{1.6}$$

subject to:

$$y_i(\mathbf{w}^T\phi(\mathbf{x}_i) + b) >= 1 - \xi_i,\ \xi_i >= 0,\ i = 1, ..., N$$

where the ξ_i are slack variables which measure the degree of misclassification of $\mathbf{x}_i$ and $C_1 > 0$ is the oil spill and $C_{-1} > 0$ is the look-alike penalty parameters of the error terms, $\boldsymbol{w}$ is the vector of coefficients and b a constant. The feature vectors are mapped into a higher dimensional space by the kernel ϕ. A RBF kernel is used: $K(\mathbf{x}_i, \mathbf{x}_j) = exp(-\gamma||\mathbf{x}_i - \mathbf{x}_j||^2)$, where $\gamma > 0$ is the kernel parameter to be specified. The same feature vector as applied for the statistical classifiers was applied here.

Training the C-SVC

In early attempts on applying regular C-SVC almost all slicks were classified as look-alikes. Since we are dealing with an imbalanced data set, we added a cost model that makes errors on the oil spill examples more expensive. This improved the results. It is emphasized in eq. (1.6) how the class dependent penalty parameters C_1 and C_{-1} are applied in C-SVC. The penalty parameter for the oil spill class (C_1) is set equal to $w_1 * C$, where w_1 is a weight. The penalty parameter for the look-alike class (C_{-1}) is set equal to $w_{-1} * C$, where w_{-1} is another weight. We selected $w_1 = \frac{number\ of\ look-alikes\ in\ subclass}{number\ of\ oil\ spills\ in\ subclass}$ and $w_{-1} = 1$. A standard *grid-search* was done where pairs of (C, γ) were tried out and the one with the best performance accuracy (on the validation set) was picked. The exponentially growing sequences $C = 2^{-5}, 2^{-4}, ..., 2^{15}$ and $\gamma = 2^{-15}, 2^{-14}, ..., 2^3$ were used for a coarse grid search. Then, a region on the grid containing the best validation accuracy was selected for a grid search with finer resolution. The final selection of C, γ, w_1 and w_{-1} used in the training of the classifier is presented in table 1.3. Each attribute of the feature vectors were scaled to the range $[-1, 1]$, as commonly done for SVM.

3.3 Automatic Confidence Estimation

A study performed by Oceanides [10] showed that a confidence assigned manually by the operator would be of high value to the surveillance aircraft if it is reliable. However, experiments involving several operators and aircraft detections showed that reliable and consistent assignment of confidence levels is difficult and the manual procedure was subjective. We have thus developed an automatic (and objective) confidence estimator as a second step in

Table 1.3: Optimal selection of C, γ, w_1 and w_{-1} for C-SVC.

Subclass	$\gamma[2^x]$	$C[2^x]$	w_1	w_{-1}
1	-15.25	16.75	142	1
2	-8	15.5	123	1
3	-2.75	-4.25	73	1
4	-2	12	51	1
5	2.5	-3	17	1
6	-0.75	-5.75	33.5	1
7	-3.5	-6	23.5	1
8	-0.25	-4.25	10.5	1
9	-3.5	-2.25	5.2	1
10	-1.5	1	1.9	1

our two-step classification approach. A confidence level is estimated for all slicks classified as oil spills in the first step of the classification approach.

To determine the confidence level of a slick, the operators at KSAT use a set of guidelines [10]. Early experiments showed that it is not sufficient to base the design of an automatic procedure only on these guidelines [11]. In addition to translating several of the KSAT criteria into computed features, we included additional features that we found important for reliable confidence estimation. An example is the KSAT guidelines for *High* confidence: *"The slick has a large contrast to gray-level surroundings. The surroundings are homogenous, with a constant gray-level. The wind speed is moderate to high, i.e. approximately 6-10 m/s. Ship or platform directly connected to slick."* and the following rule established in our automatic confidence estimator: *IF "Slick Local Contrast"* ≥ 0.91 *AND "Distance to Ship/Oilrig"* ≤ 85.84 *AND "Power-to-mean Ratio"* ≤ 0.04 *AND "Regions in Small Neighbourhood"* ≤ 6 *AND "Slick Moment"* ≥ 0.25 *AND "Slick Area"* ≥ 20 *THEN* HIGH (the feature limits are automatically estimated).

We use four confidence levels, *High, Medium, Low* and *Very Low*, and the following features for ***High:*** "Slick Local Contrast", "Power-to-mean Ratio", "Regions in Small Neighbourhood", "Distance to Ship/Oilrig", "Slick Moment" and "Slick Area". ***Medium:*** "Slick Local Contrast", "Slick Border", "Sum of External Angles", "Power-to-mean Ratio", "Slick Area", "Regions in Small Neighbourhood" and "Low Wind Area". ***Low:*** "Low Wind Area", "Power-to-mean Ratio", "Regions in Small Neighbourhood", "Sum of External Angles", "Slick Border", "Distance to Ship/Oilrig", "Slick Local Contrast", "Slick Area", "Smoothness Contrast" and "Wind".

Both the training and the validation sets were used in the training of the confidence estimator. The conditions for level *High* and *Medium* confidence were semi-automatically trained, while level *Low* was manually trained. For *High*, the feature value limits were

estimated automatically from the computed feature vectors of the oil slicks. Depending on the feature, the limit was set at either the first or the last quartile of all the sorted feature values. We only applied those slick candidates in the training and the validation set that were marked "high-confidence". The limit for the feature "Power-to-mean Ratio" was however slightly adjusted, as the automatic estimate appeared to be too strict. Confidence level *Medium* was trained in a similar manner as *High*, but here we applied the first or the last quartile of *all* the feature vectors masked as oil spills in both the training and the validation sets for the initial estimation of the limits. Some of the limits were afterwards manually adjusted after inspection of their performance on the validation set. The manual training of *Low* was done basically by visually studying the slicks in the validation images and their computed feature values.

4 Comparing the Classifiers Performance

Table 1.4 presents the performance accuracy on the test set for both the previously published statistical classifier and the new regularized statistical classifier. If we compare the classifier

Table 1.4: Classification results on the test set.

Classifier with Common Diagonal Covariance Matrices		
	Classified as Oil	Classified as look-alike
Marked as Oil	36 (87.8%)	5 (12.2%)
Marked as Look-alike	1879 (15.3%)	10366 (84.7%)
Classifier with Regularized Covariance Matrices		
	Classified as Oil	Classified as look-alike
Marked as Oil	38 (92.7%)	3 (7.3%)
Marked as Look-alike	1256 (10.3%)	10989 (89.7%)
C-SVC		
	Classified as Oil	Classified as look-alike
Marked as Oil	34 (82.9%)	7 (17.1%)
Marked as Look-alike	2768 (22.6%)	9476 (77.4%)

with common diagonal covariance matrices with the regularized one, there are 623 less false alarms after having introduced the regularization technique. If we have a look at the oil spill class, we can see that we lose 5 out of 41 oil spills with the earlier published version of the classifier, while we only lose 3 out of 41 with the regularized one. Binary C-SVC for imbalanced data sets was evaluated against the regularized statistical classifier. It is clear from table 1.4 that C-SVC performs worse than the statistical classifiers for both the oil spill and the look-alike classes. Based on this, no further investigations with C-SVC were done.

5 Confidence Estimation

Based on the results presented in section 4, we selected the regularized statistical classifier for the first step of our classification approach. Each slick with a higher posterior probability of being an oil spill than a look-alike is then automatically assigned a confidence level. Table 1.6 summarizes the final classification accuracy. Here, the classifier was trained on 76 images

Table 1.5: Performance results for the two-step classification approach.

High		
	Classified as Oil	Classified as Look-alike
Marked as Oil	13 (31.7%)	28 (68.3%)
Marked as Look-alike	8 (0.1%)	12237 (99.9%)
Medium (includes all High and Medium slicks)		
	Classified as Oil	Classified as Look-alike
Marked as Oil	21 (51.2%)	20 (48.8%)
Marked as Look-alike	30 (0.2%)	12215 (99.8%)
Low (includes all High, Medium and Low slicks)		
	Classified as Oil	Classified as Look-alike
Marked as Oil	32 (78.0%)	9 (22.0%)
Marked as Look-alike	94 (0.8%)	12151 (99.2%)
Very Low (includes all High, Medium, Low and Very Low slicks)		
	Classified as Oil	Classified as Look-alike
Marked as Oil	38 (92.7%)	3 (7.3%)
Marked as Look-alike	1200 (9.8%)	11045 (90.2%)

(both the training and the validation set), compared to 56 in Table 1.4, which reduces the a number of false alarms to 1200 for *Very Low* confidence level.

The trend found is that the surroundings of the detected slicks get more and more inhomogeneous and the number of look-alikes present increases for lower confidence levels. Fig. 1.1 presents some examples from the test set.

Counting the actual number of alarms, there were twice as many true positive as false positive alarms for the confidence category *High*, about an equal amount of true positive alarms and false positive alarms for *Medium* and about half as many true positives as false positives for *Low*. The number of correctly recognized positive examples can thus be increased at the cost of an increased number of false alarms, or vice versa. We believe that for practical use the *Low* confidence level would probably be the most suitable in most cases. The system is thus designed to produce some false alarms, which should be sorted out by manual inspection.

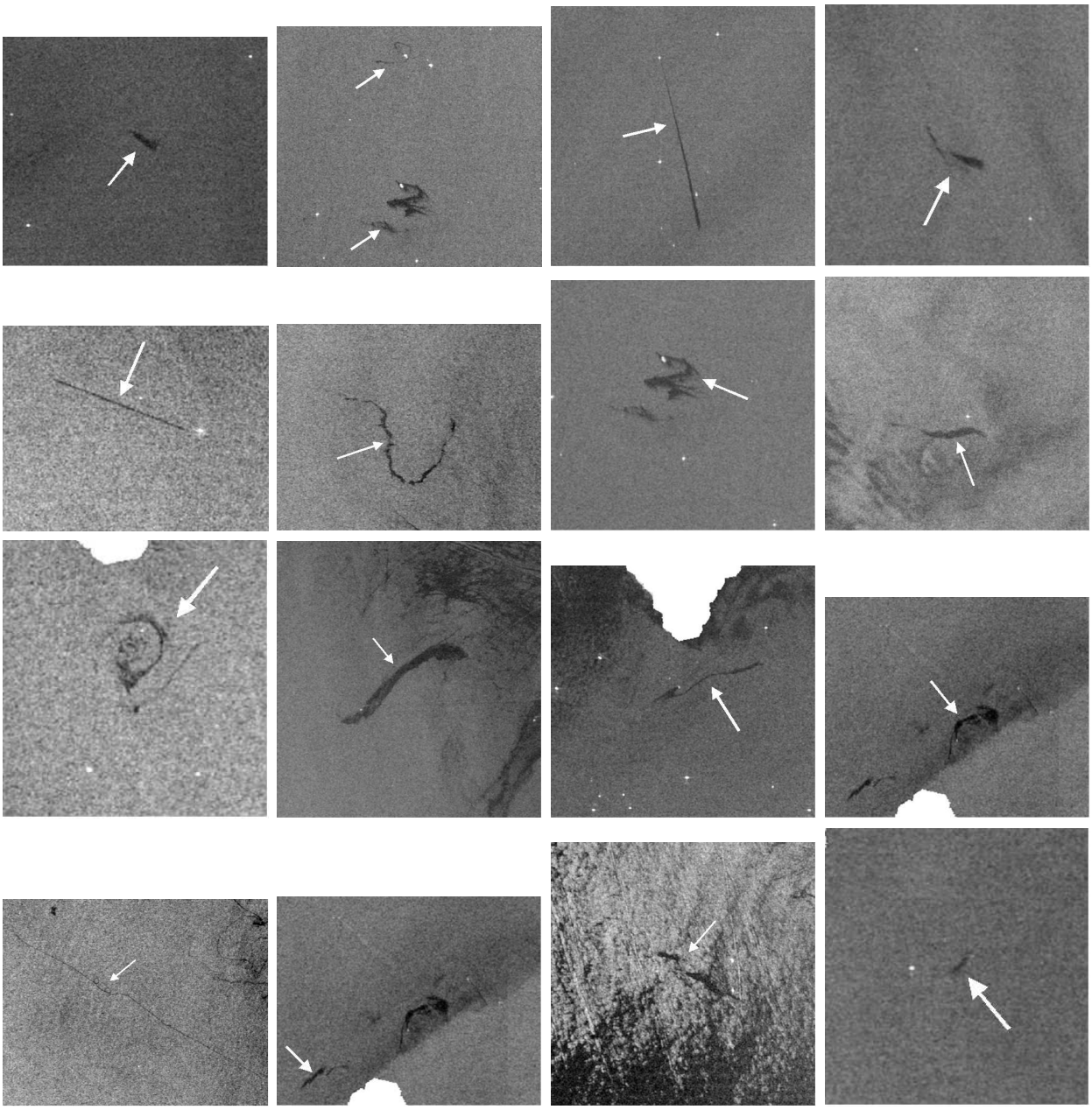


Figure 1.1: Slicks automatically assigned *High* (first row), *Medium* (second row), *Low* (third row) and *Very Low* (fourth row) confidence. Note that as the confidence decreases, the surroundings of the slicks get more and more heterogeneous.

6 Conclusion

We propose a two-step classification procedure for oil spill detection in SAR images, consisting of a regularized statistical classifier and automatic confidence estimation of the detected

slicks. This combination leads to an acceptable classification performance if the algorithm is run prior to a manual verification step before sending out aircraft/vessel. The performance is comparable to results presented in [12] where a large rule-base was trained to reduce the false alarm ratio. However, the approach proposed here is much easier to train as it is done semi-automatically and the number of parameters to be adjusted is largely reduced. Compared to the classifier earlier proposed, the results presented here show that it is possible to largely reduce the false alarm ratio by introducing regularization of the covariance matrices. By selecting a confidence level the user is also able to tune the system with respect to the frequency of true positives vs. false positives. The associated confidence level seems to be well correlated with the homogeneity of the surroundings. Information about algal blooms and wind is planned incorporated in future versions of the algorithm.

References

[1] C. Brekke and A. H. S. Solberg. Segmentation and feature extraction for oil spill detection in ENVISAT ASAR images. *Submitted: Int. J. of Remote Sens.*, 2006.

[2] C. J. C. Burges, A Tutorial on Support Vector Machines for Pattern Recognition, *Data Mining and Knowledge Discovery*, 2, 121-167, 1998.

[3] C.-C. Chang and C.-J. Lin. *LIBSVM: a library for support vector machines*, 2001. Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm.`

[4] C. Cortes and V. Vapnik. Support-Vector Networks. *Machine Learning*, 20, 273–297, 1995.

[5] B. Fiscella, et al. Oil spill detection using marine SAR images. *Int. J. of Remote Sens.*, 21(18), 3561–3566, 2000.

[6] F. Del Frate, et al. Neural Networks for Oil Spill Detection Using ERS-SAR Data. *IEEE Trans. on Geosci. and Remote Sens.*, 38(5), 2282–2287, 2000.

[7] I. Keramitsoglou, et al. Automatic identification of oil spills on satellite images. *Environmental Modelling & Software*, 21, 640–652, 2006.

[8] G. Mercier and F. Girard-Ardhuin. Partially Supervised Oil-Slick Detection by SAR Imagery Using Kernel Expansion. *IEEE Trans. on Geosci. and Remote Sens.*, 44(10), 2839–2846, 2006.

[9] F. Nirchio, et al. Automatic detection of oil spills from SAR images. *Int. J. of Remote Sens.*, 26(6), 1157–1174, 2005.

[10] A. Solberg, et al. D2-Report on benchmarking oil spill recognition approaches and best practice. Tech. report, Oceanides project, EC, Arch. No. 04-10225-A-Doc, Contr. No: EVK2-CT-2003-00177, 2005.

[11] A. H. S. Solberg. Automatic oil spill detection and confidence estimation. In *Proc. Int. Symp. Remote Sens. of Environment, St. Petersburg, Russia*, 943–945, 2005.

[12] A. H. S. Solberg, et al. Oil spill detection in Radarsat and Envisat SAR images. *IEEE Trans. on Geosci. and Remote Sens.*, 45(3), 746–755, 2007.

[13] A. H. S. Solberg, et al. Automatic Detection of Oil Spills in ERS SAR Images. *IEEE Trans. on Geosci. and Remote Sens.*, 37(4), 1916–1924, 1999.

Chapter 15

Paper VI: Classifying Oil Spills and Look-alikes in ENVISAT ASAR Images

Camilla Brekke, Anne Solberg and Geir Storvik

Published: *In the proceedings of ENVISAT symposium 2007, Montreux, Switzerland*, 23-27 April 2007 (ESA SP-636, July 2007).

Classifying Oil Spills and Look-alikes in ENVISAT ASAR Images

Camilla Brekke[1,2], Anne Solberg[2] and Geir Storvik[3]
[1] Norwegian Defence Research Establishment,
PO Box 25, NO-2027 Kjeller, Norway, E-mail: Camilla.Brekke@ffi.no
[2] Department of Informatics, University of Oslo,
PO Box 1080 Blindern, 0316 Oslo, Norway, E-mail: anne@ifi.uio.no
[3] Department of Mathematics, University of Oslo,
PO Box 1053, Blindern, 0316 Oslo, Norway, E-mail: geirs@math.uio.no

Abstract

We propose an improved classification approach for automatic oil spill detection in Synthetic Aperture Radar (SAR) images in the framework of a dark spot detector, a dark spot feature extractor and dark spot classification. New ideas for optimal subclass estimation are discussed. A regularized statistical classifier for oil spill and look-alike classification is applied within each subclass. To allow the user to tune the system with respect to the trade-off between the number of true positive alarms and the number of false positives, an automatic confidence estimator has been developed. The system is trained on 76 ENVISAT ASAR images and performance tested on 27 images.

1 Introduction

Oil spills and several other ocean features (*look-alikes*) dampen out the small scale waves on the sea surface. This reduces the backscattering back to the SAR antenna and dark slicks appear in the SAR images. A part of the oil spill detection problem is to distinguish oil slicks from the look-alikes. Our goal is to develop an automatic system for oil spill detection, in which objects with a high probability of being oil spills are identified.

The framework of our algorithm is a dark spot detector, a dark spot feature extractor and a dark spot classifier. Dark spots in the images are primarily detected by adaptive thresholding. For each of them a number of features are computed in order to classify the slick as either an oil spill or a look-alike. The classification scheme is based on statistical modelling.

As the behaviours of the features will change with different wind levels, we spilt the feature space into subclasses. A search for an optimal feature space division for the purpose of oil spill classification within each subclass has been performed based on Maximum Likelihood Estimation (MLE) and is the focus of this paper.

Various classifiers have been applied to the oil spill detection problem: A Mahalanobis classifier and a compound probability classifier were applied in [4]. The probabilistic approach was improved in [6]. A neural-network approach is described in [3], and a classifier based on fuzzy logic was developed in [5]. The work presented here is based on a statistical classifier [10; 8; 1; 2].

2 Algorithm Design

The algorithm consists of the following main steps: 1) dark spot detection based on segmentation of the SAR image, 2) feature extraction from the segmented image, 3) classification of the detected dark spots as oil spills or look-alikes and finally 4) estimation of confidence levels for each of the slicks classified as oil spills.

Pre-processing of the SAR image, consisting of converting a land mask to the image grid and a normalization of the backscatter with respect to incidence angles, is performed ahead of the segmentation step.

2.1 Segmentation of Dark Spots

In the segmentation step, dark spots are separated from the background. This is the most computationally intensive step of the algorithm.

First, an image pyramid is created by averaging pixels in the original image. From the original image, the next level in the pyramid is created with half the pixel size of the original image, and so on. Adaptive thresholding is then applied to each level in the pyramid. The threshold is set adaptively based on estimates from the SAR image of the roughness/texture

of the surrounding sea. After segmenting each level in the pyramid, the different levels are merged. More details are given in [8].

To detect thin linear slicks, elongated segments are first located in a coarse segmented version of the image. Then these segments are grown in the direction of their orientation if certain criteria are fulfilled (e.g. the backscatter value of a pixel has to be below a threshold value and it or one of its 8-neighbour pixels has to represents an edge pixel). See [1] for more details.

The 27 segmented test images contains 12245 look-alikes and 41 oil spills (64 segmented regions).

2.2 Dark Spot Feature Extraction

Discrimination between oil spills and look-alikes is based on a number of features computed for each of the segmented dark spots. The features are computed and collected in a feature vector $\mathbf{x}_i$. These features are constructed such that they typically will be different depending on if the dark spot is an oil spill or a look-alike. The following features were selected in previous experiments [1].

- **Shape features:** "Sum of External Angles", "Slick Moment (MOM)", "Slick Area", "Slick Complexity" and "Slick Width"
- **Contrast features:** "Slick Local Contrast", "Slick Border", "Smoothness Contrast"
- **Texture:** "Power-to-mean Ratio (PMR)", "Slick Variance"
- **Surroundings:** "Regions in Small Neighbourhood", "Distance to Ship/Oilrig", "Wind", "Low Wind Area"

2.3 Optimal Subclass Boundary Estimation

When selecting a classifier to solve this problem, a possibility is to use Bayesian estimation techniques to calculate the *posterior* probability for a detected spot being an oil spill. Then, a distribution needs to be selected to model the classes. The simplest choice would be multivariate Gaussian densities.

However, the behaviour of the features will change with different wind levels. In low wind, the backscatter difference between an oil spill and the surrounding sea will be large. As the wind level increases, the backscatter contrast will be lower and the feature values will change. A multimodal conditional density will likely occur. Describing the conditional density by a unimodal density (e.g. Gaussian) is therefore not appropriate. If we split the problem and assume different densities depending on the wind level, the data within each subdivision is more likely to be more homogeneous.

Even within each wind level, both the oil slicks and the look-alikes may vary quite a lot in shape, contrast and other features. In [2], the wind level was first used to divide the samples in two different subclasses and then these were divided into five subclasses based on the shape descriptor. We have applied MOM as a shape descriptor and the wind level is represented by PMR. In low wind with many look-alikes, the PMR value will be high, and high contrast between slicks and their surrounding can be expected. As the wind increases the PMR value will decrease, and the expected contrast between the oil and the surrounding sea will also decrease. Let $\Omega_1, ..., \Omega_K$ represent the subclasses, where K = the number of subclasses. Tab. 1.1 shows the initial configuration with $K = 10$ as applied in earlier studies [2]. The goal of the current study is to estimate these subclass borders. $\hat{\theta} = \{p, m_1, m_2, m_3, m_4\}$

Table 1.1: The initial subclass division.

	Slick	**Surrounding**
Subclass	**MOM**	**PMR**
Ω_1	MOM $< m_1$	PMR $\geq p$
Ω_2	MOM $\in [m_1, m_2 >$	PMR $\geq p$
Ω_3	MOM $\in [m_2, m_3 >$	PMR $\geq p$
Ω_4	MOM $\in [m_3, m_4 >$	PMR $\geq p$
Ω_5	MOM $\geq m_4$	PMR $\geq p$
Ω_6	MOM $< m_1$	PMR $< p$
Ω_7	MOM $\in [m_1, m_2 >$	PMR $< p$
Ω_8	MOM $\in [m_2, m_3 >$	PMR $< p$
Ω_9	MOM $\in [m_3, m_4 >$	PMR $< p$
Ω_{10}	MOM $\geq m_4$	PMR $< p$

are the parameters defining the subclasses and their values need to be optimised for a best possible performance. Tab. 1.2 presents the parameter values applied in [2]. These values are here used as the initial input to an optimiser searching for the optimal $\hat{\theta}$.

Table 1.2: θ: best guess for $\hat{\theta}$.

p	m_1	m_2	m_3	m_4
0.04	0.3	0.5	0.8	1.2

For the parameter search, the oil spill and the look-alike classes are merged in the training set, and the sum of the negative log-likelihood for each subclass is used as a criteria for optimisation. During the search, Gaussian densities are assumed within each subclass: $f_{\Omega_k} =$

$N(\eta_{\Omega_k}, \mathrm{T}_{\Omega_k})$, and the $\hat{\theta}$ minimizing the sum of all $l(\hat{\theta})_{\Omega_k}$ for $k = 1, ..., K$ is selected, where $l(\hat{\theta})_{\Omega_k}$ is defined as:

$$\begin{aligned} l(\hat{\theta})_{\Omega_k} = -\sum_{\mathbf{x}_i \in \Omega_k} [-\frac{p}{2}\log(2\pi) - \frac{1}{2}\log|\mathrm{T}_{\Omega_k}| \\ -\frac{1}{2}(\mathbf{x}_i - \eta_{\Omega_k})^T \mathrm{T}_{\Omega_k}^{-1}(\mathbf{x}_i - \eta_{\Omega_k})] \end{aligned} \tag{1.1}$$

where p is the number of features. Because η_{Ω_k} and T_{Ω_k} are unknown, these must be estimated as well. Using maximum likelihood estimates for these, Eq. 1.1 reduces to

$$l(\hat{\theta})_{\Omega_k} = -(-\frac{1}{2} n_{\Omega_k} \log|\widehat{\mathrm{T}}_{\Omega_k}|) \tag{1.2}$$

where n_{Ω_k} represents the number of samples within subclass Ω_k and $\widehat{\mathrm{T}}_{\Omega_k}$ is the maximum likelihood estimate for T_{Ω_k} based on these n_{Ω_k} samples.

2.4 *Log*-transform of the Data Set

If we look at the training data, some of the features appear to have a skewed distribution, therefore we did some experiments with a *Log*-transform ($x' = \log(x - min(x) + 1)$) on the data set. As an example, Fig. 1.1 shows histograms for the feature "Slick Complexity" before and after the *Log*-transform. A Linear Bayes Normal Classifier (LDC) in Matlab was trained for each subclass to evaluate the effect of the *Log*-transform on the test set. The initial configuration with 10 subclasses (Tab. 1.1) and θ (Tab. 1.2) was applied. The results presented in Tab. 1.3 shows that a normalization with the *Log*-transform does not improve the result. Therefore, we have not applied any transformation on the feature vectors before

Table 1.3: *Log*-transform. Classification results.

LDC		
	No transform	*Log*-transform
Oil spills (error rate)	9/64	12/64
Look-alikes (error rate)	1947/12245	1394/12245
Total error rate:	15.0%	15.1%

the classification step.

2.5 Classification

A prior distribution and a probability density for the features are combined through Bayes theorem to obtain the *posterior* probability for a detected spot being an oil spill. Let c be

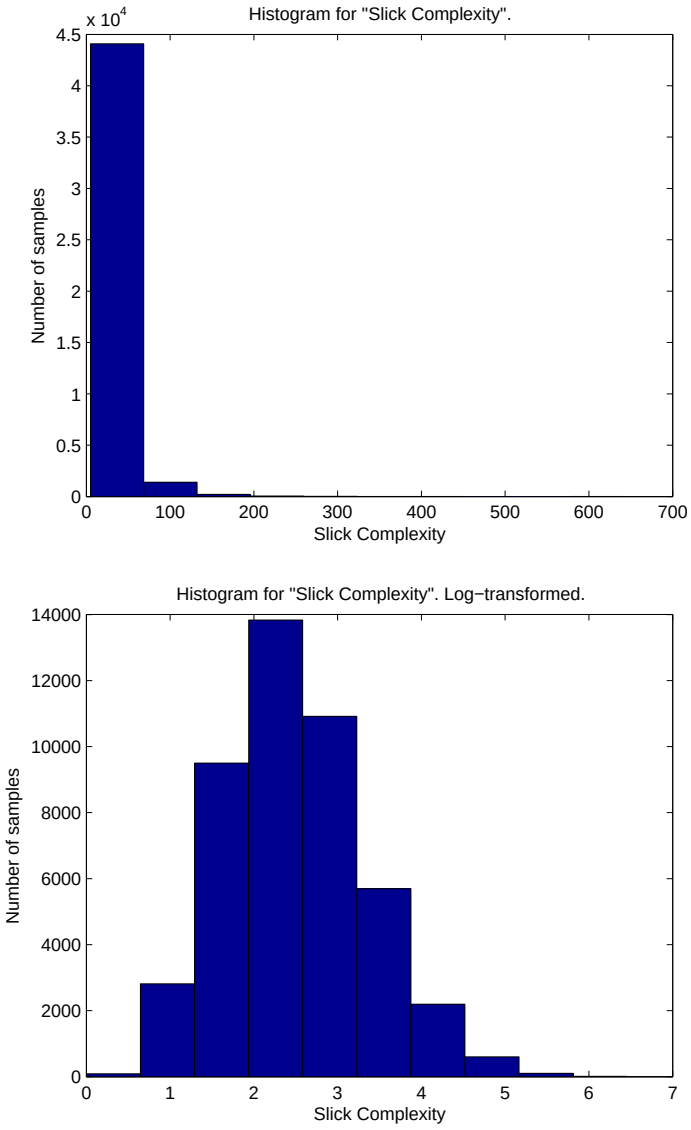


Figure 1.1: Histograms on the training set for the feature "Slick Complexity". Top: No transform. Bottom: *Log*-transformed data.

the unknown class membership of a detected spot (we are dealing with a two class problem: oil spill or look-alike). Then, for $\mathbf{x}_i \in \Omega_k$,

$$Pr(c = o|\mathbf{x}_i) = \frac{\pi_o f_{o,\Omega_k}(\mathbf{x}_i)}{\pi_o f_{o,\Omega_k}(\mathbf{x}_i) + (1 - \pi_o) f_{l,\Omega_k}(\mathbf{x}_i)} \tag{1.3}$$

where π_o is the prior model for the probability that a detected spot is oil. $f_{o,\Omega_k}(\mathbf{x}_i)$ and $f_{l,\Omega_k}(\mathbf{x}_i)$ are the probability densities for the observed features $\mathbf{x}_i$ in classes o = oil spills and l = look-alikes, respectively.

The densities are assumed Gaussian:

$$f_{c,\Omega_k}(\mathbf{x}_i) = \frac{1}{(2\pi)^{\frac{d}{2}}|\Sigma_{\Omega_k}|^{\frac{1}{2}}} \times \exp\{-\frac{1}{2}(\mathbf{x}_i - \mu_{c,\Omega_K})^T \Sigma_{\Omega_k}^{-1}(\mathbf{x}_i - \mu_{c,\Omega_k})\} \tag{1.4}$$

where $c \in \{o, l\}$, d is the number of features, μ_{c,Ω_k} is the mean vector for class c and Σ_{Ω_k} is the covariance matrix, common for both classes due to the imbalanced data set.

The features "Slick Complexity", "Power-to-mean Ratio", "Slick Local Contrast", "Slick Width", "Regions in Small Neighbourhood", "Slick Border", "Smoothness Contrast" and "Slick Variance" are included in feature vector $\mathbf{x}_i$.

Covariance Matrix Estimation

Regularized covariance matrices are used in the classifier. With the subclass division, the density for class c within a subclass Ω_k is then given by $f_{c,\Omega_k}(\mathbf{x}_i) = N(\mu_{c,\Omega_k}, \widetilde{\Sigma}_{\Omega_k})$.

Applying Gaussian densities and regularized covariance matrices $\widetilde{\Sigma}_{\Omega_k}(\rho_{\Omega_k})$, leads to a general family of covariances indexed by ρ_{Ω_k}. Regularization of the common covariance matrices can be expressed as follows:

$$\widetilde{\Sigma}_{\Omega_k}(\rho_{\Omega_k}) = \rho_{\Omega_k}[diag\{\Sigma_{\Omega_k}\}] + (1 - \rho_{\Omega_k})\Sigma_{\Omega_k} \tag{1.5}$$

where ρ_{Ω_k} is the regularization parameter and Σ_{Ω_k} is the fully estimated common covariance matrix. Here $\rho_{\Omega_k} \in [0, 1]$ allows a continuum of models, and needs to be estimated from experiments as described in [2].

2.6 Confidence Estimation

After first applying a regularized statistical classifier within each subclass, each slick with a higher posterior probability of being an oil spill than a look-alike is automatically assigned a confidence level. We have developed an automatic confidence estimator as a second step in our two-step classification approach (see Fig. 1.2). The confidence estimator will automatically assign a slick one of four confidence levels: *High, Medium, Low* or *Very Low*.

If the system is operating on level *Medium*, all slicks with *High* and *Medium* confidence are reported. If the system is operating on level *Low*, all slicks with *High, Medium* and *Low* confidence are reported. All slicks that are not detected as *High, Medium* or *Low* are given *Very Low* confidence. When operating on level *Very Low* all slicks detected as oil spills in the first step of the classification approach (see Fig. 1.2) are reported.

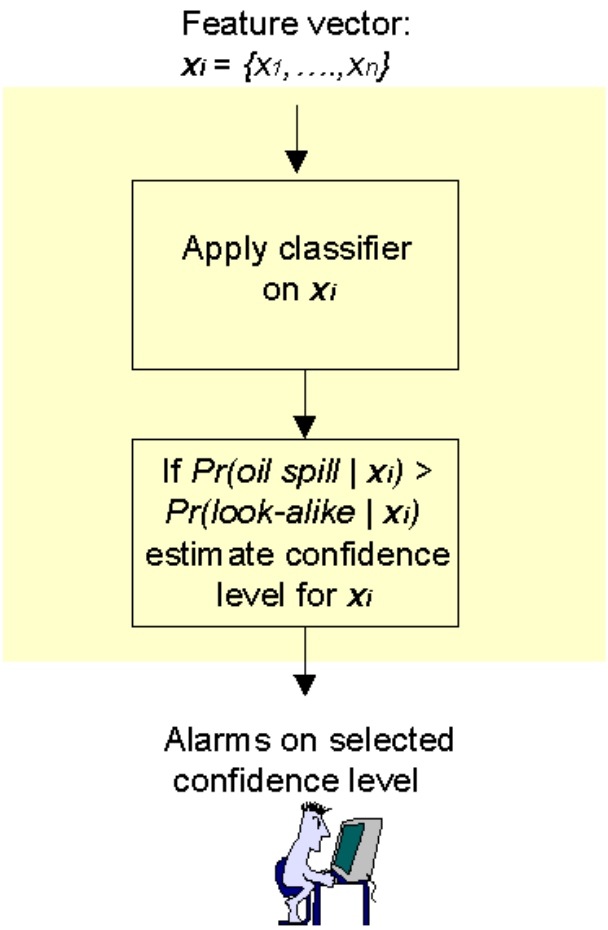


Figure 1.2: The two-step classification approach.

Kongsberg Satellite Services (KSAT) manual oil spill service chain is described in [9]. To determine the confidence level of a slick, the operators use a set of guidelines. We used these guidelines as a starting point when designing our automatic procedure for confidence estimation. However, experiments showed that it is not sufficient to base the design only on these guidelines [7]. In addition to translating several of the criteria into computed features, we included additional features that we found important for reliable confidence estimation. The development and training of the confidence estimator is described in more detail in [2].

3 Results

The results from estimating the subclass boundaries are presented in Sect. 3.2. The final classification results are presented in Sect. 3.3.

3.1 Data Set

The data set applied in this study consist of 103 ENVISAT ASAR WSM images. The training set contains 76 scenes, while the test set contains 27 scenes. The images are collected mainly from the German and Finnish Baltic Sea and the German North Sea from 2003 to 2005. Aircraft verifications, collected during a combined satellite and aircraft campaign, were available for the test set.

3.2 Results - Optimal Subclass Division

Tab. 1.4 shows the classification results from the first step of the classification approach (see Fig. 1.2) for no subclasses, 2, 4, 6 and 8 subclasses. For $K = 8$, the optimiser was given $\theta = [p = 0.04, m_1 = 0.3, m_2 = 0.5, m_3 = 0.8, m_4 = 1.2]$ as initial values and returned $\hat{\theta} = [p = 0.1285, m_1 = 0.2513, m_2 = 0.3210, m_3 = 0.4646, m_4 = 2.5844]$ which should have given 10 subclasses. However some of the subclasses were empty or had very few samples. After merging some of the subclasses, 8 subclasses were left. For $K = 6$, $\theta = [p = 0.04, m_1 = 0.5, m_2 = 0.8]$ was applied as an initial guess, while the optimal estimate was $\hat{\theta} = [p = 0.1510, m_1 = 0.2996, m_2 = 0.5171]$. For $K = 4$, $\theta = [p = 0.04, m_1 = 0.5]$ was applied as the initial guess, and $\hat{\theta} = [p = 0.1510, m_1 = 0.3971]$ was the optimal estimate. For $K = 2$, $\theta = [p = 0.04]$ was applied as the initial guess, and $\hat{\theta} = [p = 0.1510]$ was the optimal estimate.

Tab. 1.4 shows that $K = 1$ (no subclasses) gives the best classification result. Fig. 1.3 presents the two oil spills misclassified as look-alikes when applying $K = 1$. Tab. 1.5 presents the number of training samples within each of the two classes and the regularization parameter ρ_{Ω_1}. As $\rho_{\Omega_1} = 0$, this means that there is enough data to get a good estimate of the common covariance matrices (see eq. 1.5) when applying $K = 1$ (all ρ_{Ω_k} were also estimated to be 0 for 2, 4 and 6 subclasses).

3.3 Results - Confidence Estimation

Applying $K = 1$ (no subclasses), Tab. 1.6 shows the classification accuracies for the four confidence levels. The trend found in [2], with respect to the automatic confidence estimator, is that the surroundings of the detected slicks get more and more inhomogeneous and the number of look-alikes present increases for lower confidence levels.

Table 1.4: Classification accuracies on the test set applying no subclasses, 2, 4, 6 and 8 subclasses, and common regularized covariance matrices.

$K = 1$**: No subclasses.**		
	Classified as Oil spill	Classified as Look-alike
Marked as Oil Spill	39 (95.1%)	2 (4.9%)
Marked as Look-alike	1196 (9.8%)	11049 (90.2%)
$K = 2$**: 2 subclasses.**		
	Classified as Oil spill	Classified as Look-alike
Marked as Oil Spill	39 (95.1%)	2 (4.9%)
Marked as Look-alike	1494 (12.2%)	10751 (87.8%)
$K = 4$**: 4 subclasses.**		
	Classified as Oil spill	Classified as Look-alike
Marked as Oil Spill	39 (95.1%)	2 (4.9%)
Marked as Look-alike	1676 (13.7%)	10569 (86.3%)
$K = 6$**: 6 subclasses.**		
	Classified as Oil spill	Classified as Look-alike
Marked as Oil Spill	39 (95.1%)	2 (4.9%)
Marked as Look-alike	1606 (13.1%)	10639 (86.9%)
$K = 8$**: 8 subclasses.**		
	Classified as Oil spill	Classified as look-alike
Marked as Oil Spill	38 (92.7%)	3 (7.3%)
Marked as Look-alike	1607 (13.1%)	10638 (86.9%)

Table 1.5: The number of training samples within each class and the estimated ρ_{Ω_1} when applying $K = 1$ (no subclasses).

	Training Set	
ρ_{Ω_1}	**Oil Spills**	**Look-alikes**
0	285	18779

4 Conclusion

In earlier studies [10; 2], the feature space was divided into 10 subclasses. The division boundaries were selected from experiments on the training set based on a manual approach. Here, we have searched for the optimal division boundaries based on an optimiser minimizing the negative log-likelihood function.

Table 1.6: Classification accuracies on the test set.

High		
	Classified as Oil Spill	Classified as Look-alike
Marked as Oil Spill	13 (31.7%)	28 (68.3%)
Marked as Look-alike	12 (0.1%)	12233 (99.9%)
Medium		
	Classified as Oil Spill	Classified as Look-alike
Marked as Oil Spill	21 (51.2%)	20 (48.8%)
Marked as Look-alike	42 (0.3%)	12203 (99.7%)
Low		
	Classified as Oil Spill	Classified as Look-alike
Marked as Oil Spill	32 (78.0%)	9 (22.0%)
Marked as Look-alike	122 (1.0%)	12123 (99.0%)
Very Low		
	Classified as Oil Spill	Classified as look-alike
Marked as Oil Spill	39 (95.1%)	2 (4.9%)
Marked as Look-alike	1196 (9.8%)	11049 (90.2%)

We have compared the classification accuracies from applying no subclasses with dividing the feature space into 2, 4, 6 and 8 subclasses. The results show that no division of the feature space gives the best performance result on the test set.

As the final results from our two-step classification approach are comparable to earlier results where 10 subclasses were applied [2], this study shows that when there is enough data in the training set both a division into 10 subclasses and no subclasses could be applied in combination with a regularized classifier. A regularized classifier is preferred (compared to a classifier with diagonal covariance matrices as applied in early versions of the algorithm [10]) to avoid a low detection rate for oil spills and a larger amount of false alarms.

For *Low* confidence, 78.0% of the oil spills are here correctly classified while 99.0% of the look-alikes are correctly classified. This confidence level gives a good trade-off between detecting significant oil spills and having a low number of false alarms.

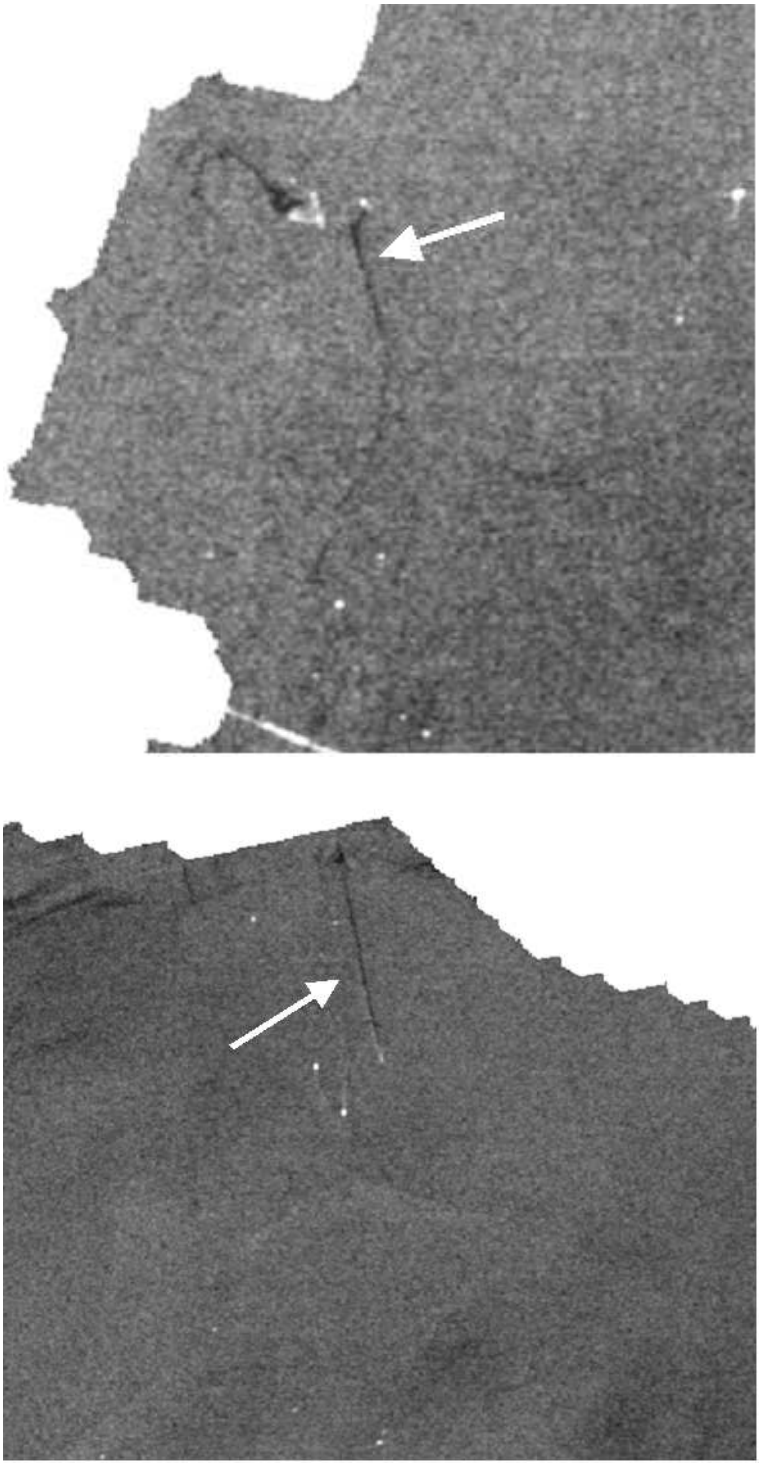


Figure 1.3: Oil spills classified as look-alikes when $K = 1$ (no subclasses) are applied.

References

[1] Brekke, C. & Solberg, A. H. S. (2006). Segmentation and feature extraction for oil spill detection in ENVISAT ASAR images. Submitted.

[2] Brekke, C. & Solberg, A. H. S. (2007). Classifiers and Confidence Estimation for Oil Spill Detection in ENVISAT ASAR Images. Submitted.

[3] Del Frate, F. Petrocchi, A., Lichtenegger, J. & Calabresi, G. (2000). Neural Networks for Oil Spill Detection Using ERS-SAR Data. *IEEE Trans. on Geosci. and Remote Sensing*, **38**(5), 2282–2287.

[4] Fiscella, B., Giancaspro, A., Nirchio, F., Pavese, P. & Trivero, P. (2000). Oil spill detection using marine SAR images. *Int. J. of Remote Sensing*, **21**(18), 3561–3566.

[5] Keramitsoglou, I., Cartalis, C. & Kiranoudis, C. T. (2006). Automatic identification of oil spills on satellite images. *Environmental Modelling & Software*, 21, 640–652.

[6] Nirchio, F., Sorgente, M., Giancaspro, A., Biamino, W., Parisato, E., Ravera, R. & Trivero, P. (2005). Automatic detection of oil spills from SAR images. *Int. J. of Remote Sensing*, **26**(6), 1157–1174.

[7] Solberg, A. H. S. (2005). Automatic oil spill detection and confidence estimation. In *Proc. Int. Symp. Remote Sensing of Environment, St. Petersburg, Russia*, 943–945.

[8] Solberg, A. H. S., Brekke, C. & Husøy, P. O. (2007). Oil spill detection in Radarsat and ENVISAT SAR images. *IEEE Trans. on Geosci. and Remote Sensing*, **45**(3), 746–755.

[9] Solberg, A., Clayton, P. & Indregard, M. (2005). D2-Report on benchmarking oil spill recognition approaches and best practice. Tech. report, Oceanides project, EC, Archive No. 04-10225-A-Doc, Contr. No: EVK2-CT-2003-00177.

[10] Solberg, A. H. S., Storvik, G., Solberg, R. & Volden, E. (1999). Automatic Detection of Oil Spills in ERS SAR Images. *IEEE Trans. on Geosci. and Remote Sensing*, **37**(4), 1916-1924.