

Unsupervised multiscale oil slick segmentation from SAR images using a vector HMC model

Stéphane Derrode^{a,*}, Grégoire Mercier^b

^aEGIM, GSM Team, Institut Fresnel (CNRS UMR 6133), Université Paul Cézanne d'Aix-Marseille III, Dom. Univ. de St Jérôme, F-13013 Marseille Cedex 20, France

^bGET/ENST Bretagne, TIME Team, TAMCIC (CNRS UMR 2658), Technopôle Brest-Iroise, CS 83818, F-29238 Brest Cedex, France

Received 22 September 2005; received in revised form 17 March 2006; accepted 27 April 2006

Abstract

This study focuses on the segmentation and characterization of oil slicks on the sea surface from synthetic aperture radar (SAR) observations. In fact, an increase in viscosity due to oil notably reduces the roughness of the sea surface which plays a major part in the electromagnetic backscattering. So, an oil spill is characterized by low-backscattered energy and appears as a dark patch in a SAR image. This is the reason why most detection algorithms are based on histogram thresholding, but they do not appear to be satisfactory since the number of false alarms is generally high.

By considering that a film has a specific impact on the ocean wave spectrum and by taking into account the specificity of SAR images, a vector hidden Markov chain (HMC) model adapted to a multiscale description of the original image is developed. It yields an unsupervised segmentation method that takes into account the different states of the sea surface through its wave spectrum. Thanks to mixture estimation, it is possible to characterize the detected areas and thus avoid most false alarms.

Results of segmentation are shown in two types of scenarios. The first one concerns an oil spill in the Mediterranean sea detected by the ERS SAR sensor at a resolution of 25 m. The second scenario is related to the wreck of the Prestige acquired by the Envisat ASAR sensor in a wide swath mode at a resolution of 150 m.

© 2006 Pattern Recognition Society. Published by Elsevier Ltd. All rights reserved.

Keywords: Oil slick detection; Synthetic aperture radar; Multiscale wavelet analysis; Hidden Markov chain; Unsupervised segmentation

1. Introduction

This study focuses on the segmentation and characterization of oil slicks in the marine environment, from data acquired by a synthetic aperture radar (SAR) sensor. On the contrary of optical sensors, SAR sensors have all-weather and all-day capabilities, which is more suitable for an operational framework (ship accidents or illegal discharges). The following introduces the oceanographical and physical motivations for the proposed multiscale strategy to detect oil slicks [1,2].

1.1. Oceanophysical background

The oceanic sea surface is complex and often governed by non-linear dynamic systems. Surface waves, that are found in the ocean, range from the millimeter scale to hundreds of meters. And, one can divide the sea surface wave spectrum into three domains [3]:

- (1) long waves (100 m in length) such as the swell;
- (2) intermediate waves (tens of meters in length) i.e., gravity waves;
- (3) short waves (less than 1 m), i.e., shortest gravity waves, gravity–capillary (few centimeters in length) and capillary (less than 1 cm in length).

* Corresponding author. Tel.: +33 491 28 28 49; fax: +33 491 28 88 13.
E-mail address: stephane.derrode@fresnel.fr (S. Derrode).

By considering an infinite sea surface, the wind induces capillary waves by friction. Capillary waves cannot propagate

far away and they vanish with the wind. But they transfer their energy to waves with a longer wavelength until they reach an equilibrium that depends on the wind. In addition, gravity waves transfer their energy to gravity–capillary waves. Several models have been proposed to characterize the sea surface spectrum with an accuracy that depends on the wavelength bandwidth. Those models integrate the wind, but also the current, atmospheric pressure and so on. An interesting comparison of some models may be found in Ref. [4] in the context of SAR imagery. On the one hand, capillary waves are generated by friction, and more specifically by friction stress related to wind speed and surface properties, and die down when the friction stress decreases. On the other hand, gravity waves are generated indirectly by sea spectrum energy spreading and propagating over long distances far from their origins. The sea surface may be roughly modeled through its wave spectrum that is governed by the following dispersion equations [3]:

$$\begin{cases} \omega^2 = \frac{\tau}{\rho} \kappa^3 & \text{for capillary waves,} \\ \omega^2 = g\kappa & \text{for gravity waves,} \\ \omega^2 = g\kappa + \frac{\tau}{\rho} \kappa^3 & \text{for gravity–capillary waves.} \end{cases} \quad (1)$$

$\omega = 2\pi/\lambda$ and λ is the wavelength, κ the surface wavenumber, g is the gravity acceleration, τ is the surface tension and ρ the water density. It is obvious from Eq. (1) that long waves are governed by gravity forces while the smaller are induced by surface tension. The next section will show that backscattering phenomena are directed by gravity–capillary waves (due to the wavelength of electromagnetic waves) while typical SAR systems are sensitive to gravity waves (due to their resolution).

Marangoni explained the process of wave damping in presence of a film on the sea surface in 1872 [5], which is detailed in several works by Alpers and co-authors [6–8]. When gravity–capillary waves propagate, successive compressions and extensions increase dissipation. When a film covers the sea surface, the wind has less effect on capillary waves. Peak-to-peak wave amplitude decreases, implying a surface stress gradient that induces an opposite force to this alternating motion. So-called Marangoni waves are created when this gradient appears, that is, with a visco-elastic film on the sea surface. The Marangoni waves are partly longitudinal in the upper layer, where dissipation is attenuated and partly tangential, linked with the stress gradient associated with visco-elastic film properties [7]. Small wave damping is associated with a decrease in surface tension. The latter being linked to the film’s elasticity. It stabilizes the sea surface and stops short wave generation by the wind. Non-linear interactions are efficient enough to transfer energy from gravity-waves to wavelengths where Marangoni damping occurs to counterbalance viscous dissipation. This transfer acts as a diffusion process through the wave spectrum. The more Marangoni damping occurs, the more non-linear energy transfer [8]. So, the overall wave spectrum

is modified by oil spills: either by energy damping (short waves vanishing), or by spreading the energy uniformly.

1.2. Sea surface observation with SAR data

Radar electromagnetic waves are backscattered by the sea surface. They are sensitive to surface roughness which is linked to gravity–capillarity waves and damped by slicks. Detecting slicks from SAR data depends on [9,10] (i) radar configuration (wavelength, polarization, incidence angle [11]), (ii) slick nature (natural or oil-based) and properties (viscosity, elasticity, thickness, temperature [8]) and, (iii) meteorological and oceanic conditions (wind and current strengths and directions). From numerous experimental studies (e.g. [12,8,11]), the wind appears to be the most important feature when characterizing a slick from SAR data. On one hand, at low wind speed the surface roughness is not uniform and the calmer areas look like oil spills since capillarity waves are not created and the contrast becomes very weak. On the other hand, at high wind speed, backscattering remains high in the slick and then contrast decreases. For the C band, which is generally preferred to X and Ku bands, a wind speed range from 2/3 to 10/14 m s⁻¹ allows an efficient detection [12].

For a 20–70° incidence angle θ , the Bragg resonance effect that links the sea surface and electromagnetic wavelengths is generally adopted to describe the backscattering process [3,13]. According to this model, the sea surface waves of wavenumber $\kappa_{\text{sea}} = 2\kappa_{\text{SAR}} \sin \theta$, with κ_{SAR} the wavenumber of the electromagnetic wave, contribute to the scattering process. By considering incidence angles from 15° to 45° according to the sensors and their acquisition modes, gravity–capillary waves of wavelengths from 4 to 14 cm intervene in the scattering process for the C band (i.e. a wavelength of about 5.6 cm). Nevertheless, SAR image resolution is much larger than the wavelength and the sensor is sensitive to large scale oceanographic phenomena. In fact, the Bragg mechanism considers waves of one wavelength only [14] while it is necessary to consider a bandwidth or, even better, the overall wave spectrum. Moreover, wave motion induces specific modulation on the wave spectrum acquired by the SAR sensor. Hence, surface roughness induced by short waves is modulated by longer waves allowing SAR imagery to characterize “indirectly” oceanophysical phenomena such as swell, internal waves, coastal bathymetry or oil slicks.

1.3. Oil spill detection strategies

A film-covered area is seen smoother than a clean sea surface since small wave generation and propagation is stopped by the viscosity of a spill. From the SAR sensor point of view, a spill is characterized by a lack of backscattered energy and then restituted through a dark area in comparison with the surrounding regions in a radar image [15].

According to the Bragg scattering theory, the backscattering process is mostly due to surface roughness. This typical effect is mostly found in the Mediterranean sea. This is due to oil characteristics and the sea surface spectrum shape where swell is less significant than in an ocean. That is why many spill detection algorithms are based on a thresholding technique [16–18]. Nevertheless, this process is not efficient when:

- the wind is increasing (basically over 10 m s^{-1});
- the sea state is different (more precisely its wave spectrum shape);
- visco-elastic oil properties are different (different kinds of oil and degradation);
- possible bad-detection with look-alikes (e.g. no wind area, up-welling, phytoplankton, algae, etc.).

Moreover, a radiometric point of view shows some limitations since wave slopes that are not oriented to the sensor are restituted with a reduced radar backscatter that may be confused with small spills. That is why most spill detection strategies include a post-processing stage to remove small slicks through geometrical, morphological or contextual criteria [19–21]. A priori, the proposed multiscale strategy that is implemented to detect oil spills may be justified by several concepts mentioned above:

- the SAR sensor is only sensitive to surface roughness (at a centimeter scale) which is modulated by larger scale phenomena that induce shades of texture;
- the increase of viscosity, due to the presence of an visco-elastic film on the sea surface, affects the sea surface wave spectrum shape;
- oil films induce dark areas on the SAR images under certain conditions, such as a limited wind (i.e. under $10/14 \text{ m s}^{-1}$).

The paper is organized as follows: Section 2 presents the multiscale transformation and explains how the initial sea surface observation is represented as coarse radiometric information associated with a set of coefficients linked to the surface roughness at several scales. Section 3 presents the statistical segmentation procedure. Since the vector extension of the hidden Markov chain (HMC) model is straightforward, we focus on the way to deal with multidimensional non-Gaussian probability density functions (pdf) arising in the statistical model. Section 4 shows some results of oil spill detection from ERS and ENVISAT SAR images. Section 5 concludes and gives some perspectives.

2. Multiscale representation

The wavelet transform is an appropriate tool for local analysis of the sea wave spectrum. It is based on a mother

wavelet ψ and is defined as

$$\mathcal{W}_f(b, a) = \int_{-\infty}^{\infty} f(t) \frac{1}{\sqrt{a}} \psi^* \left(\frac{t-b}{a} \right) dt.$$

Many shapes may be found in mother wavelets (but they are subject to admissibility conditions [22]) and the choice of the *best* wavelet remains an open question, depending on the application. In this study, we focus on the multiscale decomposition with wavelet transform which is shift invariant. Hence, details and texture have the same signature whatever their location.

2.1. Multiscale wavelet transform

The decomposition is implemented in order to analyze the local shape of the sea surface wave spectrum. Then, a multiscale differential operator is used to characterize the sharp structures of a signal f . The wavelet operator is then defined (in one-dimensional) as

$$\mathcal{W}_\ell\{f\}(b, a) = a^\ell \frac{\partial^\ell}{\partial b^\ell} (f * \overline{\theta}_a)(b),$$

where ℓ is the decomposition level ($0 \leq \ell < L$) and $\overline{\theta}_a(x) = (1/\sqrt{a})\theta(-x/a)$. Convolution $f * \overline{\theta}_a$ acts as a smoothing of f over a domain a .

When considering the first derivatives, the modulus of $\mathcal{W}_1\{f\}(b, a) = a(\partial/\partial b)(f * \overline{\theta}_a)(b)$ is maximum where $(f * \overline{\theta}_a)(b)$ fluctuates. When considering the second derivative, the modulus maxima of $\mathcal{W}_2\{f\}(b, a) = a^2(\partial^2/\partial b^2)(f * \overline{\theta}_a)(b)$ correspond to the points of maximum curvature of f . More generally, wavelet coefficients are proportional to the components of the gradient of f smoothed by θ_a . Then, with such a multiscale differential operator each singularity of f is detected by following the modulus maxima of the wavelet coefficients towards finer scales. Moreover, it has been shown that if f adopts a band-limited Fourier transform and if the wavelet ($\psi = -\partial\theta/\partial x$) has compact support, then the wavelet modulus maxima define a complete and stable signal representation [23].

In order to analyze the sea surface wave spectrum shape yielded by a SAR observation, a wavelet-based multiscale edge detector is defined as a two-dimensional separable dyadic wavelet. Its Fourier transforms (in two-dimensional) may be written as

$$\begin{aligned} \mathcal{F}_{\psi^{\text{hori}}}(\omega_{\text{hori}}, \omega_{\text{vert}}) &= \mathcal{F}_g \left(\frac{\omega_{\text{hori}}}{2} \right) \mathcal{F}_\phi \left(\frac{\omega_{\text{hori}}}{2} \right) \mathcal{F}_\phi \left(\frac{\omega_{\text{vert}}}{2} \right), \\ \mathcal{F}_{\psi^{\text{vert}}}(\omega_{\text{hori}}, \omega_{\text{vert}}) &= \mathcal{F}_g \left(\frac{\omega_{\text{vert}}}{2} \right) \mathcal{F}_\phi \left(\frac{\omega_{\text{hori}}}{2} \right) \mathcal{F}_\phi \left(\frac{\omega_{\text{vert}}}{2} \right) \end{aligned}$$

with

$$\mathcal{F}_\phi(\omega) = \left(\frac{\sin \frac{\omega}{2}}{\omega/2} \right)^{m+1} e^{-j\varepsilon \frac{\omega}{2}}, \quad \varepsilon = \begin{cases} 1 & m \text{ even,} \\ 0 & m \text{ odd.} \end{cases}$$

This yields low-pass filter coefficients $\overline{\mathcal{F}}_h(\omega) = \sqrt{2}(\cos \omega/2)^{m+1} e^{-j\varepsilon\omega/2}$ that correspond to a box-spline

of degree m . $g(\cdot)$ is a discrete approximation of the derivative operator, of Fourier transform: $\mathcal{F}_g(\omega) = -j\sqrt{2}(\sin \omega/2)^{m+1} e^{-j\omega/2}$.

In order to analyze the observations as locally as possible, a scale function $\theta(\cdot)$ was chosen as short as possible and its Fourier transform is a cubic spline. The multiscale decomposition yields a set of images for which the wavelet coefficients represent the details at a coarser and coarser scale ℓ ($0 \leq \ell < L$), associated with a smooth image which is a coarse radiometric approximation (at scale L) of the original image (see Fig. 1). Then, an observation f may be represented with its details at different scales by

$$f \rightarrow \begin{cases} \Theta_L = f * \theta(2^{-(L-1)}x, 2^{-(L-1)}y), \\ \Psi_\ell^{\text{hori}} = f * \psi^{\text{hori}}(2^{-\ell}x, 2^{-\ell}y), & 0 \leq \ell < L, \\ \Psi_\ell^{\text{vert}} = f * \psi^{\text{vert}}(2^{-\ell}x, 2^{-\ell}y), & 0 \leq \ell < L, \end{cases} \quad (2)$$

which corresponds to cutting the Fourier spectrum into a dyadic sequence: $\omega_{\max}/2^{\ell+1}$, $0 \leq \ell < L$.

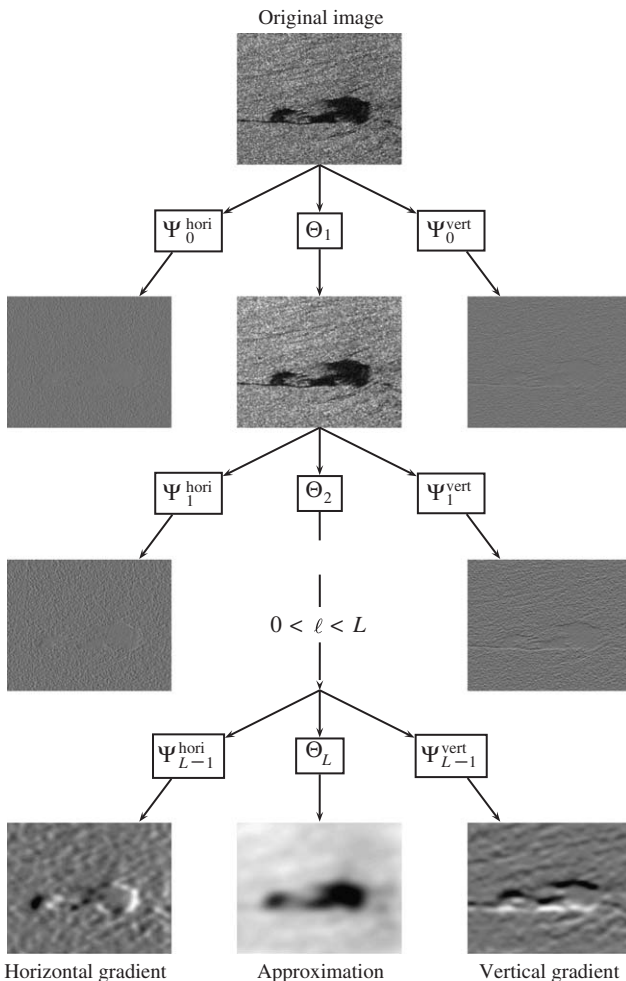


Fig. 1. Multiscale decomposition of a SAR image. Lossless characterization of the initial information is performed through images of wavelet coefficients (Ψ_ℓ coming from the convolution with functions $\psi(2^{-\ell}x, 2^{-\ell}y)$) and the coarse image at level L (coming from the convolution with $\theta(x/2^{L-1}, y/2^{L-1})$).

Remark. The vector $(\Psi_\ell^{\text{hori}}, \Psi_\ell^{\text{vert}})^t$ is proportional to the gradient of f smoothed by θ_{2^ℓ} . In this kind of decomposition, there is no diagonal decomposition as is the case in most dyadic multiresolution decompositions. Moreover, this decomposition acts as a multiscale edge detector and there is no consideration to orthogonality between the subspaces generated by $\psi_{2^\ell}^{\text{hori}}$, $\psi_{2^\ell}^{\text{vert}}$ and θ_{2^ℓ} . Hence, it is necessary to consider a relationship between wavelet coefficients of different scales through the correlation. When applied to SAR images, it is not necessary to use more than three or four levels since $\Theta_{L=3}$ or $\Theta_{L=4}$ contains details at the swell scale, which is enough for our application.

2.2. Statistical characterization of the wavelet coefficients

The distribution of each band of the multiscale decomposition is characterized by parametric families of distributions. On the one hand, the Pearson system of distributions is used to describe low-pass coefficients Θ_L of Eq. (2), in accordance with laws that have to be taken into account for SAR images [24]. On the other hand, the generalized Gaussian family was selected for its capacity to better represent the shape of high-pass coefficients Ψ_ℓ^{hori} and Ψ_ℓ^{vert} .

2.2.1. Characterization of low-pass coefficients through the Pearson system

The low-pass image, Θ_L , obtained by convolution with a low-pass filter θ_{2^L} , shows a smooth histogram with a shape similar to the histogram of the original image.

It is now well accepted that the statistical model of radar data follows a Gaussian circular complex law in homogeneous regions (where physical characteristics of backscattering are stationary) [25]. In fact, in each radar resolution cell, a large number of N_d elementary backscatters is summed in a coherent way (by a contribution in magnitude and phase of the electromagnetic wave, and not only in energy). The contribution of the N_d elementary backscattered fields remains random since the number of backscatters and their orientation are unknown. In most radar images, it is well accepted that the intensity is distributed according to a Gamma or a generalized Gamma law, with statistical moments depending on the reflectivity of the homogeneous area. The multiplicative model of the speckle was introduced since the mean of the intensity of the pixel is proportional to its variance. However, for an image of the sea surface, usual assumptions of the fully developed speckle are not necessary acceptable and may be wrong since:

- the number of scatters may not be significant enough to make the law of large numbers valid and to consider the speckle as fully developed;
- the elementary backscatters are not independent. The description of the wave spectrum of the sea surface does not allow us to consider that the wave spectrum energy is

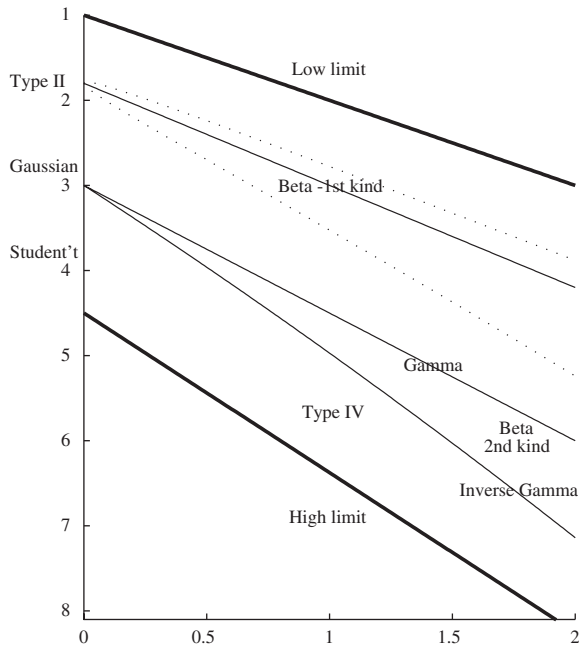


Fig. 2. Eight families of the Pearson system in the (β_1, β_2) -diagram. Note that the β_2 axis is reversed: (a) wavelet coefficient pdf; (b) generalized Gaussian distributions.

Gaussian for each wavelength and independent from the wavelength;

- the elementary backscatters are moving.

This is the reason why we decided not to filter the speckle before processing but, conversely, to use statistical laws induced by this phenomenon to segment the surface roughness.

Then, the Pearson system [26] was used since it appears adapted to the segmentation of SAR images [27,28]. This system is made up of mainly eight families of distributions (including Gaussian, Gamma and Beta) and offers a large variety of shapes (symmetrical or not, with finite or semi-finite support, etc.). Each law can be uniquely defined by its mean (μ_1) and its first three centered moments (μ_2, μ_3, μ_4). All of them can be represented in the so-called Pearson diagram (as shown in Fig. 2) in which axes β_1 and β_2 are given by

- Skewness $\sqrt{\beta_1}$ with $\beta_1 = \mu_3^2/\mu_2^3$,
- Kurtosis $\beta_2 = \mu_4/\mu_2^2$.

Gaussians are located at $(\beta_1 = 0, \beta_2 = 3)$, Gamma distributions (III) on the straight line $\beta_2 = 1.5 \beta_1 + 3$ and inverse Gamma distributions on the curve with the equation

$$\beta_2 = \frac{3}{\beta_1 - 32}(-13\beta_1 - 16 - 2(\beta_1 + 4)^{3/2}) \quad \text{with } \beta_1 < \frac{96}{25}.$$

First kind Beta distributions are located between the lower limit and the Gamma line, second kind Beta distributions

are located between the Gamma and the inverse Gamma distributions, and Type IV distributions are located between the inverse Gamma distributions and the upper limit. Then, it is possible to estimate empirical moments of a distribution from a sample and to assess the family of distributions from coordinates (β_1, β_2) and the parameters that precisely characterize the pdf within its family.

2.2.2. Characterization of high-pass coefficients through generalized Gaussians

We observed [29,30] that the pdf of wavelet coefficients, from most mother wavelets, can be globally modeled by the generalized Gaussian family, given by

$$p(x; \mu, \alpha, \beta) = \frac{\beta}{2\alpha \Gamma(1/\beta)} e^{-(|x-\mu|/\alpha)^\beta},$$

where $\Gamma(\cdot)$ is the Gamma function, μ is the mean, α is a scale parameter which characterizes the width of the pdf (i.e. the variance), and where β is a shape parameter. For example, the Laplace law comes from $\beta = 1$, and the Gaussian pdf of variance $\alpha/2$ from $\beta = 2$.

Several parameter estimation methods have been proposed [31], especially by means of the moment method like for the Pearson system. However, the maximum likelihood estimation is more robust for heavy-tailed distributions and when the number of samples is low. From an independent and identically distributed sample $\mathbf{x} = \{x_i; i = 1, 2, \dots, N\}$, the log-likelihood \mathcal{L} to be maximized is given by

$$\mathcal{L}(\mathbf{x}; \mu, \alpha, \beta) = \log \prod_{i=1}^N p(x_i; \mu, \alpha, \beta). \quad (3)$$

This optimization problem is similar to minimizing the difference of entropy between a generalized Gaussian with parameters (μ, α, β) and the pdf of sample \mathbf{x} . In general, this minimization gives an unique root in the case of generalized Gaussians. We obtain this system of equations:

$$\begin{cases} \min_{\mu} \sum_{i=1}^N |x_i - \mu|^\beta, \\ \frac{\partial \mathcal{L}}{\partial \alpha} = -\frac{N}{\alpha} + \frac{\beta}{\alpha^{\beta+1}} \sum_{i=1}^N |x_i - \mu|^\beta = 0, \\ \frac{\partial \mathcal{L}}{\partial \beta} = \frac{N}{\beta} + \frac{N \Psi(1/\beta)}{\beta^2} - \sum_{i=1}^N \left(\frac{|x_i - \mu|}{\alpha} \right)^\beta \log \left(\frac{|x_i - \mu|}{\alpha} \right) = 0, \end{cases}$$

where $\Psi(\cdot) = \Gamma'(\cdot)/\Gamma(\cdot)$ is the Digamma function. It is important to note that even if, globally (in the whole image), coefficients are centered, this not necessarily true for a local neighborhood of a pixel. Then, the estimation of the three parameters is necessary.

The first equation is not derivable and depends on two parameters. So we developed a numerical method that involves in minimizing $\sum_{i=1}^N |x_i - \mu|^\beta$ from successive estimations

of parameters β and μ according to

$$\begin{cases} \hat{\beta} \text{ such that } \frac{\Gamma(2/\hat{\beta})}{\sqrt{\Gamma(1/\hat{\beta})\Gamma(3/\hat{\beta})}} = \frac{1/N \sum_{i=1}^N |x_i - \hat{\mu}|}{\sqrt{1/N \sum_{i=1}^N (x_i - \hat{\mu})^2}}, \\ \hat{\mu} \text{ such that } \min_{\hat{\mu}} \sum_{i=1}^N |x_i - \hat{\mu}|^{\hat{\beta}}. \end{cases}$$

The first estimation of μ is obtained by the empirical mean of the sample: $\hat{\mu} = (1/N) \sum_{i=1}^N x_i$. It is possible to refine the estimation of β , once $\hat{\mu}$ is fixed, by looking for the root of

$$1 + \frac{\Psi(1/\hat{\beta})}{\hat{\beta}} - \frac{\sum_{i=1}^N |x_i - \hat{\mu}|^{\hat{\beta}} \log |x_i - \hat{\mu}|}{\sum_{i=1}^N |x_i - \hat{\mu}|^{\hat{\beta}}} + \frac{\log(\hat{\beta}/N \sum_{i=1}^N |x_i - \hat{\mu}|^{\hat{\beta}})}{\hat{\beta}}.$$

Finally, the scale parameter α is computed according to

$$\hat{\alpha} = \left(\frac{\hat{\beta}}{N} \sum_{i=1}^N |x_i - \hat{\mu}|^{\hat{\beta}} \right)^{1/\hat{\beta}}.$$

In order to illustrate the method, Fig. 3 shows the histograms of the wavelet coefficients at three different levels from the

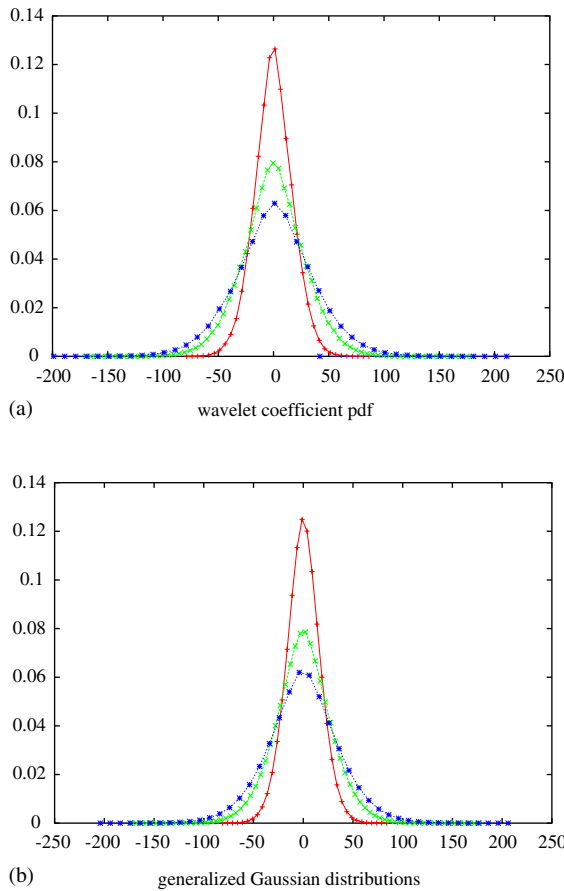


Fig. 3. Histograms of wavelet coefficients from the three band decompositions (a) and the three corresponding generalized Gaussians estimated from the maximum likelihood (b).

multiscale decomposition of Fig. 1 and the estimated generalized Gaussians.

3. Segmentation using a vector HMC model

From the multiscale decomposition, the initial observation is characterized by a vector of data with $M = 2L + 1$ components: $(\Theta_L, \Psi_{L-1}^{\text{hori}}, \Psi_{L-1}^{\text{vert}}, \dots, \Psi_\ell^{\text{hori}}, \Psi_\ell^{\text{vert}}, \dots, \Psi_0^{\text{hori}}, \Psi_0^{\text{vert}})^t$. The segmentation of the multiscale representation of the original image requires the implementation of a vector-based classification algorithm. We propose to adapt the HMC model to the vector context. This method is quite general and has been applied in different situations (multi-spectral [32], multitemporal [33]). The specifications of this algorithm in the multiscale context come from the choice of the pdf shapes that characterize observations.

3.1. Hidden Markov Chain

The interest of Markovian models is mainly due to the fact that when the unobservable signal process X can be modeled by a finite Markov chain and when the noise is not too complex, then X can be recovered from the observed process Y by using Bayesian classification techniques such as a maximum a posteriori (MAP) or a marginal posterior mode (MPM). For unsupervised classification, Markovian parameters have to be estimated from the observed data only. Well-known iterative methods like estimation–maximization (EM) or stochastic EM (SEM) can be used. In this work, we considered a third strategy called iterative conditional estimation (ICE) [34], whose interest in image segmentation has been illustrated several times (in remote sensing [35], in SAR [27,36], in sonar [37], or in medical [38] imaging) and for various structures like field, chain or tree.

Recently, it has been shown that, in some situations, the HMC model can compete with hidden Markov random field (HMRF) based methods in terms of classification accuracy, while being much faster, even though the latter provide a finer and more intuitive modeling of the spatial relationships [36]. In a vector context, the first step is to transform the M images into M one-dimensional chains by using a Hilbert–Peano scan [39]. The size of each chain is equal to the number of pixels (N) in the image. Then, M series of N data have to be considered. Let $\mathbf{y}^m = (y_1^m, \dots, y_N^m)$ be the data sequence for each component m , $1 \leq m \leq M$. For each index n , let $\bar{\mathbf{y}}_n = (y_n^1, \dots, y_n^M)^t$ be the M -component vector, and $\bar{\mathbf{y}} = (\mathbf{y}^1, \dots, \mathbf{y}^M)^t = (\bar{\mathbf{y}}_1, \dots, \bar{\mathbf{y}}_N)$ the set of all sequences (see Fig. 4).

3.2. Segmentation algorithm

The statistical approach involves considering that each observation $\bar{\mathbf{y}}_n \in \mathbb{R}^M$ is a realization of a random vector $\bar{\mathbf{Y}}_n$ that has to be classified among a set of K classes $\Omega = \{\omega_1, \dots, \omega_K\}$. The segmentation result is a realization

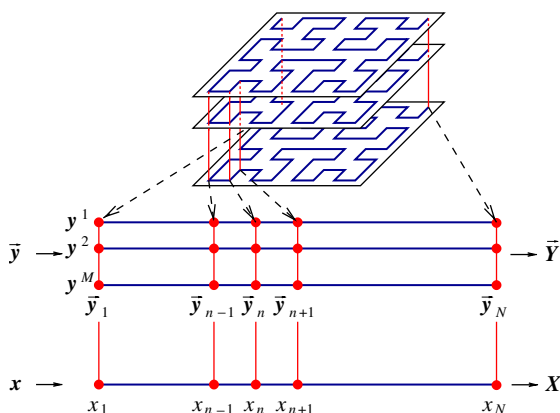


Fig. 4. Hilbert–Peano scan for a 8×8 multicomponent image (M : number of bands in the image, N : number of pixels in each band).

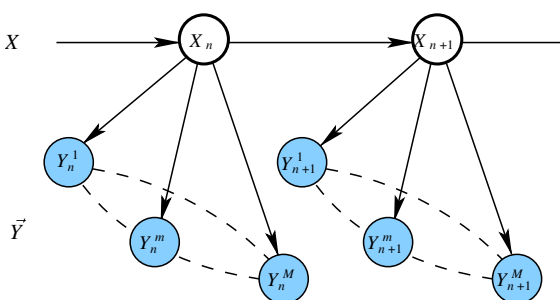


Fig. 5. Hidden Markov chain dependence graph. The M components of the observations (y_n^1, \dots, y_n^M) are linked (dotted line) by correlation.

\mathbf{x} of the process $\mathbf{X} = (X_1, \dots, X_n, \dots, X_N)$, with $X_n \in \Omega$, assumed to be Markovian and stationary. Given the two classical assumptions:

- random vectors $(\vec{Y}_n)_{1 \leq n \leq N}$ are independent conditionally to \mathbf{X} :

$$p(\vec{Y} = \vec{y} | \mathbf{X} = \mathbf{x}) = \prod_{n=1}^N p(\vec{Y}_n = \vec{y}_n | \mathbf{X} = \mathbf{x}),$$

- the distribution of \vec{Y}_n conditionally to \mathbf{X} is identical to its distribution conditionally to X_n :

$$p(\vec{Y}_n = \vec{y}_n | \mathbf{X} = \mathbf{x}) = p(\vec{Y}_n = \vec{y}_n | X_n = x_n),$$

we obtain the dependence structure of Fig. 5. It should be emphasized that no assumption has been made about the independence of components Y_n^m conditionally to X_n . Thus, it is possible to take into account the statistical links between the bands of the multiscale decomposition (links are represented by dotted curves)—see Section 3.3.

In unsupervised classification, the distribution $p(\vec{Y} = \vec{y} | \mathbf{X} = \mathbf{x})$ is unknown and must first be estimated in order to apply

a Bayesian classification technique (MAP or MPM). Therefore, the following sets of parameters need to be estimated:

γ set: The K^2 parameters $c(\omega_\ell, \omega_k) = p(X_{n-1} = \omega_\ell, X_n = \omega_k)$ have to be estimated in order to calculate a priori probabilities $p(\omega_k)$ and the transition matrix of components $p(X_n = \omega_k | X_{n-1} = \omega_\ell)$. All these probabilities are independent of n since \mathbf{X} is assumed to be stationary.

δ set: The parameters of the K M -dimensional pdfs $f_{\omega_k}(\vec{y}_n) = p(\vec{Y}_n = \vec{y}_n | X_n = \omega_k)$ have to be estimated. For a Gaussian pdf, $M \times (M + 1)$ parameters have to be considered: M -component mean and an $M \times M$ covariance matrix.

The extension of the ICE algorithm to the vector case is almost straightforward; further details of the method can be found in the appendix. Nevertheless, at each ICE iteration and for each class, an estimation of an M -dimensional density—which is not necessarily Gaussian—has to be performed.

3.3. Multidimensional pdf estimation

At an ICE iteration p and for a given class ω_k , let \vec{z} be the subset of \vec{y} that belongs to class ω_k . The parameters that characterize f_{ω_k} have to be estimated. In a non-Gaussian context, multidimensional estimation is quite difficult. The multivariate analysis involves decomposing the estimation of the M -dimensional pdf into M estimations of one-dimensional pdfs. Several strategies are available, depending on the assumptions made on the statistical links between the data.

- If data are considered independent [40], then f_{ω_k} is the product of M densities $f_{\omega_k}^m$:

$$f_{\omega_k}(\vec{z}_n) = \prod_{m=1}^M f_{\omega_k}^m(z_n^m). \tag{4}$$

However, this assumption cannot be justified in our context (as shown in Section 2).

- Independent component analysis (ICA) allows us to project random vectors \vec{z}_n in a new space where components are statistically independent [41]. The Central Limit Theorem says that the sum of ℓ random variables tends to a law towards a Gaussian distribution when ℓ tends to infinity. The corollary of this theorem says that the sum of two independent random variables has a more Gaussian distribution than any of the two initial random variables. Then, the estimation of independent components involves finding a linear mixture of the observation so that the yielded vector has the least Gaussian component [42,43]. Unfortunately, most of the time, the maximization of a non-Gaussianity criterion leads to multimodal distributions.
- A sub-optimal solution involves using a PCA. The principle is to decorrelate the original data in order to use Eq. (4) as in Ref. [44]. If there exists a matrix \mathbf{A} such that data $\vec{t}_n = \mathbf{A}\vec{z}_n$ are not correlated, then the covariance

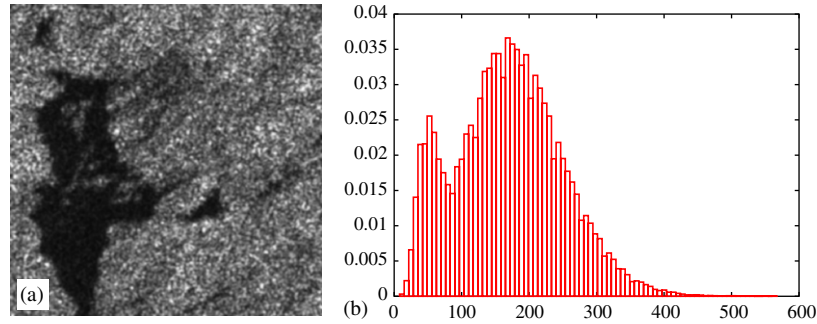


Fig. 6. ERS SAR image of an oil spill (a) and its histogram (b).

matrix of data \vec{t}_n is diagonal (unary matrix for example). Hence, it is possible to write $A\Gamma_{\vec{z}}A^t = Id$, with $\Gamma_{\vec{z}}$ the covariance matrix of data \vec{z} . By using the Choleski decomposition, a solution A of this equation is given by

$$A^t = (\text{Choleski}(\Gamma_{\vec{z}}))^{-1}.$$

It can be easily verified that variances of the projected data are unitary and their covariances are zero. For example, for $M = 2$ components, we get

$$A = \begin{pmatrix} \frac{1}{\sigma_1} & 0 \\ \frac{-\rho}{\sigma_1\sqrt{1-\rho^2}} & \frac{1}{\sigma_2\sqrt{1-\rho^2}} \end{pmatrix},$$

where σ_1^2 and σ_2^2 denote variances of \vec{z}^1 and \vec{z}^2 , respectively, and ρ the intercorrelation. If distributions $g^m(\cdot)$, $1 \leq m \leq M$, denote the one-dimensional pdf estimated from the decorrelated data \vec{t}_n , thus

$$f_{\omega_k}(\vec{z}_n) = |\det A| \prod_{m=1}^M g_{\omega_k}^m(t_n^m). \quad (5)$$

No assumption about the shape of densities in principal components is necessary. The method can be applied in a Gaussian context (where decorrelation stands for independence) as well as in a generalized mixture context where components belong to different families of distributions. This is the case for the multiscale data used here to detect and characterize oil spills.

4. Detection of oil spills

The multiscale segmentation method is the following. The multiscale decomposition described in Section 2 is applied to a sea surface SAR observation y . Then, the observation is characterized by $\vec{y} = (\Theta_L, \Psi_{L-1}^{\text{hori}}, \Psi_{L-1}^{\text{vert}}, \dots, \Psi_{\ell}^{\text{hori}}, \Psi_{\ell}^{\text{vert}}, \dots, \Psi_0^{\text{hori}}, \Psi_0^{\text{vert}})^t$ as in Eq. (2), without loss of information. This multicomponent information is segmented via the vector HMC model presented in Section 3. The pdf of the low-pass coefficients Θ_L has been modeled by families from

the Pearson system of distributions, and the pdf of the high-pass coefficients $\Psi_{\ell}^{\text{hori}}$ and $\Psi_{\ell}^{\text{vert}}$ ($0 \leq \ell < L$) by the generalized Gaussian family.

The number of classes used to segment the images is set manually. This value is typically equal to two if one wants to detect the presence or absence of oil slicks. Nevertheless, one or two more classes may be necessary to take into account a meteorological phenomenon like a wind front, a squall or a particular state of oil such as an emulsion of oil and sea water. We tested the algorithm on two types of SAR images showing oil spills. Below, local results from a high-resolution image (decametric resolution for a satellite sensor) with ERS data are shown first, then more global results with a lower resolution image (Envisat ASAR, Wide Swath mode, 150 m resolution).

4.1. Results on ERS data

We used an ERS SAR image showing an oil slick in the Mediterranean sea (cf. Fig. 6). The dark part of the image shows a compact spill, whereas the free sea appears lighter due to a lower viscosity and the presence of waves. Mixed areas are observable on the right-hand part of the image. Since the size of these zones is small, these cannot be assigned to local variations in the wind speed. But they could be an emulsion or, on the contrary, volatile oil in which the damping of waves is not so great. This scene is a typical spill in the Mediterranean sea. The difficulty in the location of oil spills is not in detecting large slicks but in determining precisely frontiers and detecting more ambiguous areas like the dark zones in the right-hand part of the image. It is known for example that a marginal spill does not have the same behavior windward or leeward.

Fig. 7 groups all the segmentation results in order to make some comparisons. The segmentation of the image was performed with different methods and two or three classes (first and second columns). The thresholding value was set manually, according to the histogram shape. The thresholding and the blind ICE algorithms illustrate the limits of methods based on radiometry only. Even if the slick is well localized, Fig. 7(a) and (d) show a large number of false alarms in clean

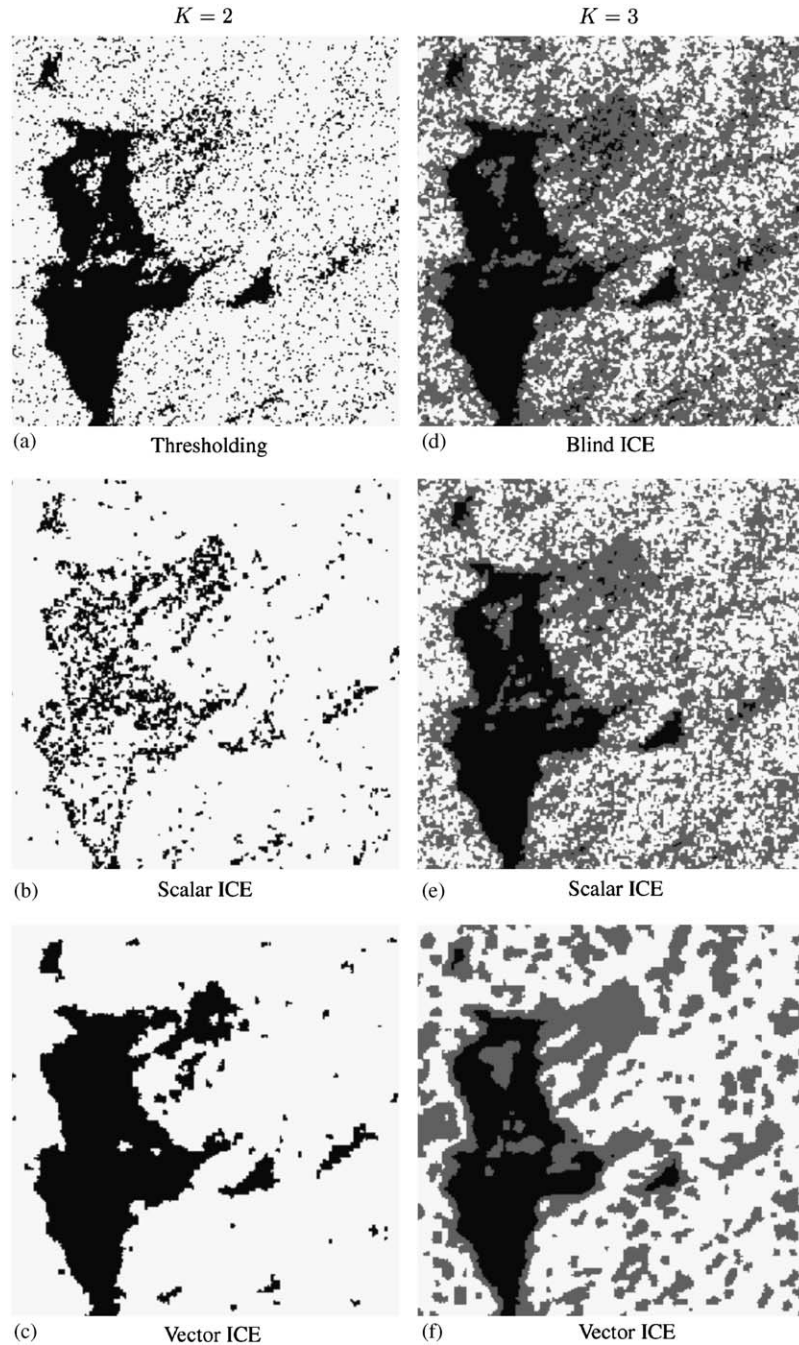


Fig. 7. Segmentation results of the image in Fig. 6 for $K = 2$ and $K = 3$ classes. The first row corresponds to the thresholding algorithm (2 classes) and blind ICE method (3 classes). The second row corresponds to the scalar HMC applied to the original image directly. The third row corresponds to the vector HMC algorithm applied on a multiscale decomposition.

water. The algorithms cannot take into account the texture of the sea surface, and the third class introduced in Fig. 7(d) split up clean water. Regarding the two Markovian models, the error rate in clean water was reduced in the case of two classes—Fig. 7(b)—but detection of the spill is not satisfactory. Nevertheless, the image Fig. 7(e) shows the precise location of the spill. Radiometric and spectral information from the contextual Markovian model allows the quality of

detection to be improved significantly. But the clean water inside the spill has been absorbed by the class representing the oil spill.

The specific interest of the vector HMC model in multiscale decomposition is illustrated by the image Fig. 7(f). In addition, the algorithm achieves the segmentation of the local texture variations and gives much more satisfactory results. This shows that the quality of segmentation is not

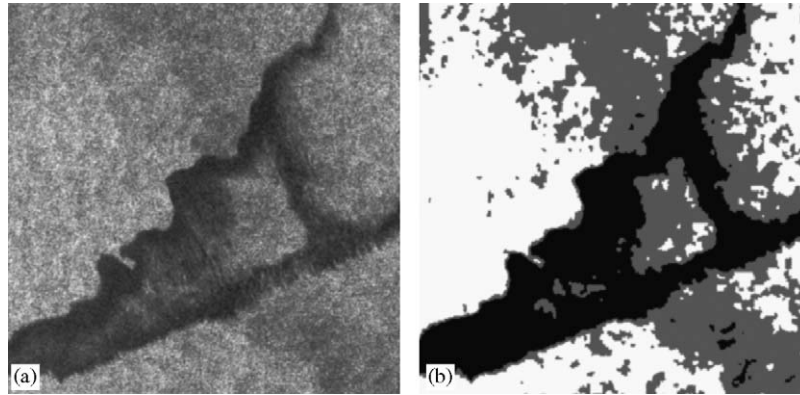


Fig. 8. (a) ASAR image from Envisat acquired during the wreck of the Prestige ©ESA. (b) Segmentation of oil by using the vector HMC model on the multiscale decomposition of the image.

only due to the markovianity but also to the way the scene is represented. The meaning of the multiscale representation is twofold: the coarse information Φ_L acts as a regularization of the initial observation, while the set of wavelet coefficients $\Psi_\ell^{\text{hori or vert}}$ characterizes all the shades of local texture. From a thematic point of view, the slick is perfectly detected. In a similar way, the clean water and the oil spill are characterized without any false alarms. The intermediate class can be found in two areas: first, at the interface between the sea and the spill and, second, in ambiguous zones where ground measurements could show fragmented spills or volatile oil. In fact, when oil is in contact with the sea, it loads up water and on disintegration makes an emulsion whose viscosity is very different from that of the initial oil. On the contrary, light oil diffuses on the water surface.

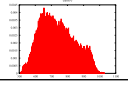
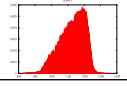
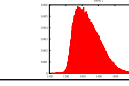
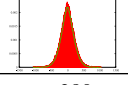
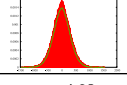
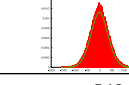
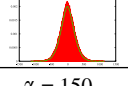
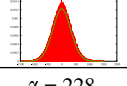
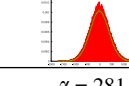
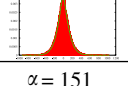
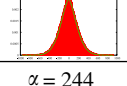
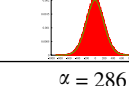
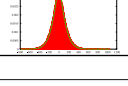
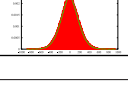
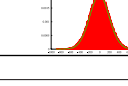
4.2. Results on envisat data

On November 19, 2002, the wreck of the Prestige confirms the interest of radar remote sensing for the location of oil spills in an operational context. In adverse weather conditions with a cloudy sky or at night, SAR sensors yield images of identical quality, while optical sensors are useless. However, the presence of dark areas in images is not sufficient to claim the presence of oil slick on the sea surface. Such areas can be generated by a lack of wind, by really strong wind, by oceanographic phenomena like up-wellings or the presence of phytoplankton, etc.

The additional interest of the proposed approach is the simultaneous estimation of a mixture of laws. So, once the segmentation is achieved, an analysis of the type and shape of laws in the mixture can be used to confirm or not the presence of oil in a dark area. The image in Fig. 8 (a) shows an oil trail just before the Prestige wreck, on November 17, 2002. The oil spill has divided into two trails due to the sea state, the wind and the viscosity of oils. So, the characterization of the oil between these two trails becomes difficult. Furthermore, atmospheric conditions induce a particular pressure on the sea surface and a front appears with a

Table 1

Histograms and pdfs of the three classes obtained by the vector HMC method during the segmentation process of the multiscale decomposition of the original image

Band	Class		
	oil	ambiguous	sea free of pollution
$\Theta_{L=2}$	$\beta_1 = 0.1$ $\beta_2 = 2.25$ 	$\beta_1 = 0.25$ $\beta_2 = 2.53$ 	$\beta_1 = 0.4$ $\beta_2 = 3.2$ 
$\Psi_{\ell=1}^{\text{hori}}$	$\alpha = 255$ $\beta = 1.62$ 	$\alpha = 406$ $\beta = 1.79$ 	$\alpha = 499$ $\beta = 1.87$ 
$\Psi_{\ell=1}^{\text{vert}}$	$\alpha = 280$ $\beta = 1.69$ 	$\alpha = 460$ $\beta = 1.92$ 	$\alpha = 549$ $\beta = 1.94$ 
$\Psi_{\ell=0}^{\text{hori}}$	$\alpha = 150$ $\beta = 1.29$ 	$\alpha = 228$ $\beta = 1.60$ 	$\alpha = 281$ $\beta = 1.91$ 
$\Psi_{\ell=0}^{\text{vert}}$	$\alpha = 151$ $\beta = 1.36$ 	$\alpha = 244$ $\beta = 1.74$ 	$\alpha = 286$ $\beta = 1.94$ 

lower radiometry without revealing an oil spill. The image in Fig. 8(b) presents the result of segmentation with three classes. A first class is characteristic of the oil spill, a second one of the ambiguous zone where the sea surface appears darker, and a third one of the sea free of pollution.

If the presence of an ambiguous class makes the segmentation result ambiguous, the analysis of statistical laws (Table 1) can be used to decide the presence of oil with more

safety for each class. It is interesting to note that wavelet coefficients from oil spills are characterized by pdfs with a shape closer to a Laplace pdf than laws of wavelet coefficients from clean water, which are more similar to Gaussian densities. Hence, this ambiguous class can be assigned a posteriori to an area free of pollution, and surely represents an atmospheric phenomenon, like a squall. Finally, we can conclude that this kind of classification segments the local wave spectrum observed from a SAR sensor.

Remark. From an operational point of view, radar images represent a huge amount of data (8000×8000 for ERS images) that it is necessary to take into account during the process. However, the Markov chain model is not suited to dealing with large images due to the amount of memory needed for computation. To mitigate this problem, it is possible to apply the algorithm on a partition of the image into blocks (of size 1000×1000 for example). This approach is coherent for the following two reasons:

- (1) The multiscale description is local since it is built by convolution with finite impulse response filters. Furthermore, tiling (overlapping of blocks) is possible, to reduce border effects in the segmentation of the whole image.
- (2) The Markovian model assumes that process X is stationary, so that blocks are also assumed stationary. This assumption is not always true for images with oil slicks. However, the influence of the neighborhood of each pixel is spatially limited since correlation reduces exponentially with distance in a Markovian model.

Finally, it is possible to join all the blocks together to yield a global segmentation map for the overall image.

5. Conclusion

In this work, an unsupervised method for the segmentation of oil spills based on a multiscale decomposition and a Markovian model have been presented. More precisely, the multiscale decomposition was implemented with a wavelet transform which acts as a multiscale differential operator. The segmentation is performed thanks to a vector extension of the HMC model. The estimation of the multidimensional pdf arising in the algorithm was achieved by PCA in which each component comes from the Pearson system of distributions (for the low-pass coefficients) or from the family of generalized Gaussians (for the high-pass coefficients). This method performs the segmentation of the sea surface thanks to a local characterization of the wave spectrum observed by a SAR sensor. This segmentation strategy may detect different phenomena that have an impact on the sea surface wave spectrum and seems to be appropriate for the detection of oil spills on the sea surface.

However, the stationarity assumption of the Markov chain can be a limitation for the analysis of full size radar images in an operational context. The lack of validity of this

assumption induces a wrong segmentation where the oceanographic phenomena are mixed together in order to yield a Markov chain where parameters are stationary. It becomes difficult to make the distinction between an oil spill with low viscosity and a slow wind area. Hence, a natural evolution of the method will involve taking into account the non-stationary characteristic of a sea surface in presence of oil. In addition, an interesting point to consider for further work is to measure the ability of the method to discriminate between oil spills and look-alikes such as natural films or other pollutants. One idea is to use the shape of the distributions estimated during the segmentation step, compared to other features based on the shape or the backscatter of the segmented regions [45]. Finally, ships induce strong scatters with a high level of radiometry on SAR images. These anomalies are highly penalizing for the statistical characterization of wavelet coefficients and have a strong impact on the segmentation result. It is thus necessary to suppress the ship signatures and two solutions are possible: by detecting ships in a pre-processing step or by using a non-linear decomposition with median filters. This latter idea constitutes the main perspective for further work.

Appendix A. ICE description

The skeleton of the entire segmentation algorithm is divided into three steps, see Fig. 9, denoting $\theta = (\gamma, \delta)$:

Step 1—Initialization. The objective is to make a preliminary estimation of model parameters. The low-pass image (Θ_L) is segmented with a blind ICE algorithm, in a Gaussian context, yielding x_0 . The estimation is performed according to the complete data (x_0, \vec{y}) and gives θ_0 ;

Step 2—ICE parameter estimation. This step is performed according to ICE. Each iteration gives an estimation θ_p . Since the number P of iterations cannot be set a priori, the algorithm stops when parameters do not vary under a limit or when the maximum number of iterations is reached.

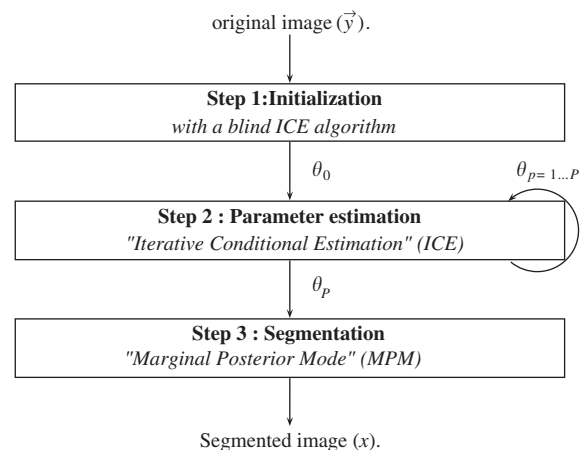


Fig. 9. Overall skeleton of the segmentation algorithm. Step 2 is detailed in Fig. 10.

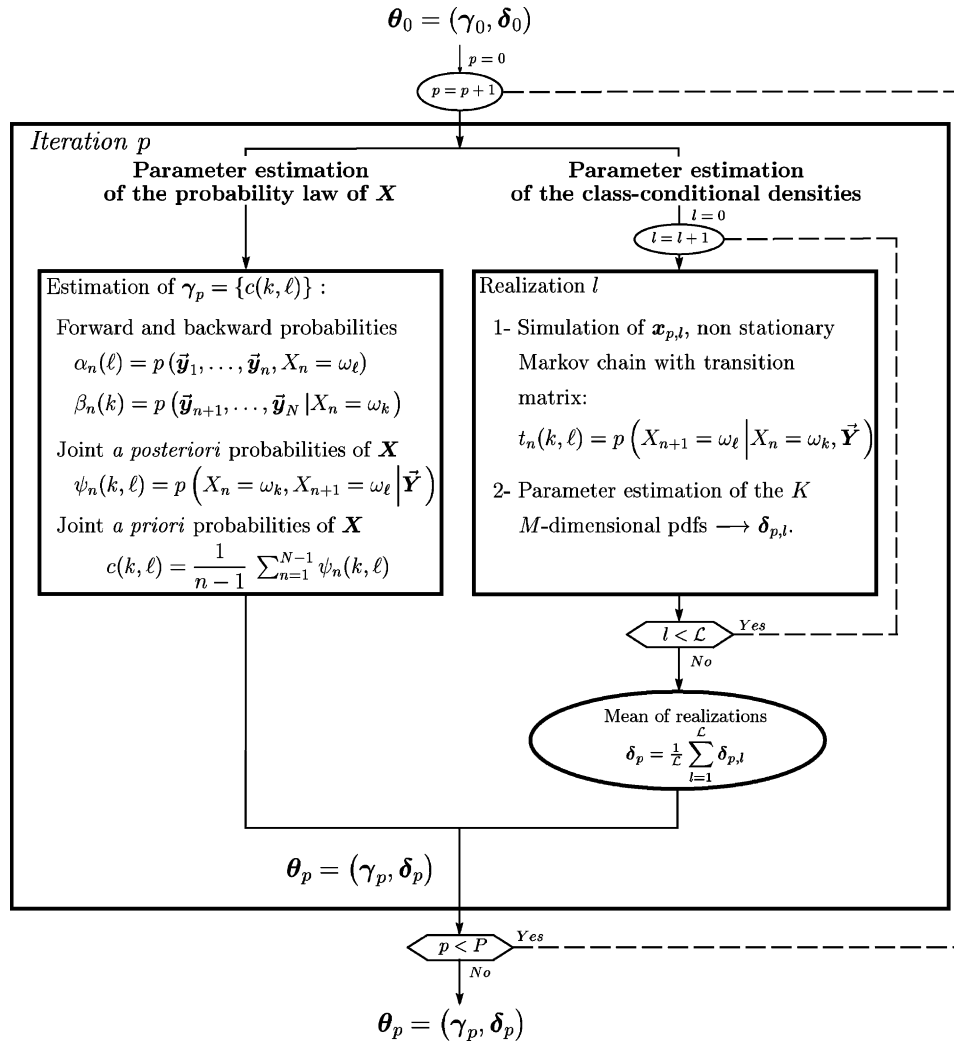


Fig. 10. Skeleton of the ICE parameter estimation procedure (stage 2).

Step 3—Bayesian segmentation This last step involves segmenting the image from estimated parameters (θ_p) by using the MPM criterion.

The ICE principle is based on the conditional expectation of some estimators from the complete data (x, \vec{y}) [34]. ICE is an iterative procedure which gives a series of estimations θ_p of parameter θ in the following way:

- (1) Initialization of θ_0 .
- (2) Calculation of $\theta_{p+1} = E_{\theta_p}[\hat{\theta}(X, \vec{Y}) | \vec{Y} = \vec{y}]$, where $\hat{\theta}(X, \vec{Y})$ is an estimator of θ from the complete data (X, \vec{Y}) .
- (3) Stop at iteration P if $\theta_p \simeq \theta_{p+1}$.

In the context considered here, this procedure gives two different situations:

γ set: The expectation of estimators of parameters $\hat{\gamma}$ can be calculated analytically, by using the Baum–Welch algorithm (“Forward” and “Backward” probabilities) [46].

δ set: The expectation of estimators of parameters cannot be calculated analytically. However, it can be estimated by computing the mean of several estimations according to $\theta_{p+1} = (1/\mathcal{L}) \sum_{l=1}^{\mathcal{L}} \hat{\theta}(x^l, \vec{y})$, with x^l an a posteriori realization of X conditionally to \vec{Y} , simulated by using θ_p . Indeed, it can be shown that the distribution $p(X = x | \vec{Y} = \vec{y})$ is a non-homogeneous Markov chain whose parameters can be computed with the Baum–Welch algorithm, which allows its simulation.

The skeleton of the ICE parameter estimation procedure is given in Fig. 10 and corresponds to a zoom on stage 2 of the general skeleton of Fig. 9.

References

- [1] G. Mercier, S. Derrode, W. Pieczynski, J.-M. Le Caillec, R. Garello, Multiscale oil slick segmentation with Markov chain model, in: Proceedings of the IEEE IGARSS’03, Toulouse, Fr., 21–25 July, 2003.

- [2] G. Mercier, S. Derrode, W. Pieczynski, Segmentation multi-échelle de nappes d'hydrocarbure, *Traitement du Signal* 21 (4) (2004) 329–346.
- [3] F.T. Ulaby, R.K. Moore, A.K. Fung, *Microwave Remote Sensing: Active and Passive*, Addison-Wesley, Reading, MA, 1982.
- [4] T. Elfouhaily, B. Chapron, A comparison of wind wave spectra used in ocean remote sensing modeling, in: *Proceedings of the IEEE IGARSS'96*, vol. 1, Lincoln, NE, USA, 27–31 May, 1996, pp. 606–608.
- [5] C. Marangoni, Sul principio della viscosità superficiale dei liquidi stabili, *Nuovo Cimento* 2 (5/6) (1872) 239–273.
- [6] H. Hühnerfuss, W. Alpers, W.D. Garrett, P.A. Lange, S. Stolte, Attenuation of capillarity and gravity waves at sea by monomolecular organic surface films, *J. Geophys. Res.* 88 (C14) (1983) 9809–9816.
- [7] W. Alpers, H. Hühnerfuss, Radar signatures of oil films floating on the sea surface and the Marangoni effect, *J. Geophys. Res.* 93 (C4) (1988) 3642–3648.
- [8] W. Alpers, H. Hühnerfuss, The damping of ocean waves by surface films: a new look at an old problem, *J. Geophys. Res.* 94 (C5) (1989) 6251–6265.
- [9] B. Fiscella, A. Giancaspro, F. Nirchio, P. Pavese, P. Trivero, Oil spill detection using marine SAR images, *Int. J. Remote Sensing* 21 (18) (2000) 3561–3566.
- [10] F. Girard-Ardhuin, G. Mercier, F. Collard, R. Garello, Operational oil-slick characterization by SAR imagery and synergistic data, *IEEE J. Oceanic Eng.* 30 (3) (2005) 487–495.
- [11] V. Wismann, M. Gade, W. Alpers, H. Hühnerfuss, Radar signatures of marine mineral oil spills measured by an airborne multi-frequency radar, *Int. J. Remote Sensing* 19 (18) (1998) 3607–3623.
- [12] K.P. Singh, A.L. Gray, R.K. Hawkins, R.A. O'Neil, The influence of surface oil on c- and ku-band ocean backscatter, *IEEE Trans. Geosci. Remote Sensing* 24 (5) (1986) 738–744.
- [13] I.S. Robinson, *Satellite Oceanography*, Wiley, New York, 1995.
- [14] J.W. Wright, A new model for sea clutter, *IEEE Trans. Antennas Propag.* 16 (1968) 217–223.
- [15] M. Gade, W. Alpers, H. Hühnerfuss, V. Wismann, P. Lange, On the reduction of the radar backscatter by oceanic surface films: scattometer measurements and their theoretical interpretation, *Remote Sensing Environ.* 66 (1998) 52–70.
- [16] M. Gade, J. Scholz, C. von Viebahn, On the detectability of marine oil pollution in European marginal waters by means of ERS SAR imagery, in: *Proceedings of the IEEE IGARSS'00*, vol. 6, Honolulu, HI, USA, 24–28 July, 2000, pp. 2510–2512.
- [17] P. Lombardo, D. Conte, A. Morelli, Comparison of optimised processors for the detection and segmentation of oil slicks with polarimetric SAR images, in: *Proceedings of the IEEE IGARSS'00*, vol. 7, Honolulu, HI, USA, 24–28 July, 2000, pp. 2963–2965.
- [18] K.W. Bjerde, A.H.S. Solberg, G. Storvik, Oil spill detection in SAR imagery, in: *Proceedings of the IEEE IGARSS'93*, vol. 3, Tokyo, Japan, 18–21 August, 1993, pp. 943–945.
- [19] A.H.S. Solberg, G. Storvik, R. Solberg, E. Volden, Automatic detection of oil spills in ERS SAR images, *IEEE Trans. Geosci. Remote Sensing* 37 (1999) 1916–1924.
- [20] H. Espedal, Detection of oil spill and natural film in the marine environment by spaceborne SAR, in: *Proceedings of the IEEE IGARSS'99*, vol. 3, Hamburg, Germany, 28 June, 2 July, 1999, pp. 1478–1480.
- [21] M. Bertacca, F. Berizzi, E. Dalle Mese, A. Capria, A FARIMA-based analysis for wind falls and oil slicks discrimination in sea SAR imagery, in: *Proceedings of the IEEE IGARSS'04*, Anchorage, AK, USA, 20–24 September, 2004.
- [22] I. Daubechies, *Ten lectures on wavelets*, Society for Industrial and Applied Mathematics, Philadelphia, Pennsylvania, 1992.
- [23] C. Kiecy, C. Lennard, Unique reconstruction of band-limited signals by a Mallat–Zhong wavelet transform algorithm, *Fourier Anal. Appl.* 3 (1) (1997) 63–82.
- [24] F. Henderson, A. Lewis, *Principles and applications of imaging radar, Manual of Remote Sensing*, third ed., vol. 2, Wiley, New York, 1998.
- [25] P. Beckmann, A. Spizzichino, *The Scattering of Electromagnetic Waves from Rough Surfaces*, Macmillan, New York, 1963.
- [26] N.L. Johnson, S. Kotz, *Distribution in Statistics: Continuous Univariate Distributions*, vols. 1 and 2, Wiley, New York, 1994.
- [27] Y. Delignon, A. Marzouki, W. Pieczynski, Estimation of generalized mixture and its application in image segmentation, *IEEE Trans. Image Process.* 6 (10) (1997) 1364–1375.
- [28] S. Derrode, G. Mercier, J.-M. Le Caillec, R. Garello, Estimation of sea-ice SAR clutter statistics from Pearson's system of distributions, in: *Proceedings of the IEEE IGARSS'01*, Sydney, Australia, 9–13 July, 2001.
- [29] S. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. Pattern Anal. Machine Intell.* 11 (17) (1989) 674–693.
- [30] M. Do, M. Vetterli, Wavelet-based texture retrieval using generalized Gaussian density and Kullback–Leibler distance, *IEEE Trans. Image Process.* 11 (2) (2002) 146–158.
- [31] K. Sharifi, A. Leon-Garcia, Estimation of shape parameter for generalized Gaussian distributions in subband decompositions of video, *IEEE Trans. Circuits Syst. Video Technol.* 5 (1) (1995) 52–56.
- [32] G. Mercier, S. Derrode, M. Lennon, Hyperspectral image segmentation with Markov chain model, in: *Proceedings of the IEEE IGARSS'03*, Toulouse, Fr., 21–25 July, 2003.
- [33] S. Derrode, G. Mercier, W. Pieczynski, Unsupervised change detection in SAR images using a multicomponent HMC model, in: *Proceedings of the Second International Workshop on MultiTemp'03*, Ispra, It., 16–18 July, 2003.
- [34] W. Pieczynski, Statistical image segmentation, *Mach. Graphics Vision* 1 (2) (1992) 261–268.
- [35] J.-M. Boucher, P. Lena, Unsupervised bayesian classification, application to the forest of Paimpont (Brittany), *Photo Interpretation* 32 (1994/4, 1995/1–2) (1995) 79–81.
- [36] R. Fjørtoft, Y. Delignon, W. Pieczynski, M. Sigelle, F. Tupin, Unsupervised segmentation of radar images using hidden Markov chains and hidden Markov random fields, *IEEE Trans. Geosci. Remote Sensing* 41 (3) (2003) 675–686.
- [37] M. Mignotte, C. Collet, P. Pérez, P. Boutheymy, Three-class Markovian segmentation of high-resolution sonar images, *Comput. Vision Image Understanding* 76 (3) (1999) 191–204.
- [38] M. Mignotte, J. Meunier, Three-dimensional blind deconvolution of SPECT images, *IEEE Trans. Biomed. Eng.* 47 (2) (2000) 274–280.
- [39] W. Skarbeck, Generalized Hilbert scan in image printing, in: R. Klette, W.G. Kropfsh (Eds.), *Theoretical Foundation of Computer Vision*, Akademik Verlag, Berlin, 1992.
- [40] N. Giordana, W. Pieczynski, Estimation of generalized multisensor hidden Markov chains and unsupervised image segmentation, *IEEE Trans. Pattern Anal. Machine Intell.* 19 (5) (1997) 465–475.
- [41] J.-F. Cardoso, Blind signal separation: statistical principles, *Proc. IEEE* 9 (10) (1998) 2009–2025.
- [42] C. Jutten, J. Héroult, Blind separation of sources, part I: an adaptive algorithm based on neuromimetic architecture, *IEEE Trans. Signal Process.* 24 (1991) 1–10.
- [43] P. Common, Independent component analysis, a new concept?, *IEEE Trans. Signal Process.* 36 (1994) 287–314.
- [44] W. Pieczynski, J. Bouvrais, C. Michel, Estimation of generalized mixture in the case of correlated sensors, *IEEE Trans. Image Process.* 9 (2) (2000) 308–311.
- [45] C. Brekke, A.H.S. Solberg, Oil spill detection by satellite remote sensing, *Remote Sensing Environ.* 95 (2005) 1–13.
- [46] P.A. Devijver, Baum's forward–backward algorithm revisited, *Pattern Recognition Lett.* 3 (1985) 369–373.