Sidescan Sonar Segmentation using texture descriptors and active contours

Maria Lianantonakis

Yvan R. Petillot

School of Engineering and Physical Sciences

Heriot Watt University

Edinburgh, Scotland, UK

email: M.Lianantonakis@hw.ac.uk

Y.R.Petillot@hw.ac.uk

Abstract

This paper is concerned with the application of active contour methods to unsupervised binary segmentation of high resolution sonar images. First texture features are extracted from a side scan image containing two distinct regions. A region based active contour model of Chan and Vese [5] is then applied to the vector valued images extracted from the original data. The set of features considered is the Haralick feature set based on the co-occurence matix. To improve computational efficiency the extraction of the Haralick feature set is implemented by using sum and difference histograms as proposed by Unser [23]. Our implementation includes an automatic feature selection step used to readjust the weights attached to each feature in the curve evolution equation that drives the segmentation. Results are shown on simulated and real data. The influence of the algorithm paprameters and contour initialisation are analysed.

I. Introduction

The role of underwater seabed analysis is becoming increasingly important for many applications including marine science (habitat mapping, environmental monitoring), off-shore industry

and military applications. There are two types of existing approaches to the problem: the first is based on the analysis of acoustic and geoacoustic data of the sea bottom(see [2], [22] and references therein). The second approach is image based and makes use of, primarily, sidescan and bathymetric sonar data [3], [13], [19], [25].

Sidescan sonar systems are used to generate 2-D, high resolution images that represent large areas of the seabed. They are characterised by an acoustic signal which is emitted from the sides of the sonar system in a direction perpendicular to the direction of travel. The beam produced from each pulse is narrow in the horizontal direction and wide in the vertical direction, its purpose being to insonify a narrow, long strip of the seabed on either side of the platform and perpendicular to the navigation track. The intensity of the backscattered signal along a strip is displayed as a function of time corresponding to a row of data in the sidescan image, each time instant referring to a point in the seabed. The sidescan image is then produced by putting together a sequence of these signals along the navigation track to create a 2-D representation of large areas of seabed.

High-resolution sidescan images are characterised by visually distinct areas corresponding to: objects on the seabed, visualised as high intensity areas caused by the reflection of the acoustic wave, shadows, visualised as low intensity, textured areas caused by the lack of acoustic reverberation from areas neigbbouring the objects, and background, visualised as distinct areas with strong texture characteristics. The latter are the result of the backscatter caused by the seabed, with different type of seabed producing a distinct backscatter response.

Sidescan sonar image analysis is used in a number of applications, from object localisation and identification to seabed classification and 3-D reconstruction. There are two main approaches in sidescan image analysis: classification using texture features and supervised learning [3], [25] and unsupervised segmentation/classification based on Bayesian clustering methods using grey levels or texture features [13], [19]. The first of these approaches has produced useful results for seabed classification but existing methods are very sensitive both to training and to the viewpoint in the sonar data [1]. Existing usupervised segmentation algorithms for sidescan data are largely concerned with identifying objects on the seabed by segmenting the image in shadow, non-shadow areas ([13] and references therein) or into three regions corresponding to shadow, echo and sea-bottom reverberation ([19]). The approach used in [13], [19] is based on a novel hierarchical Markov tree model with prior knowledge being incorporated in the model. This

has produced some impressive results but feature selection and parameter estimation remain an issue, the latter being usually addressed using training or adhoc settings.

In this paper we concentrate on a different type of segmentation task which has received very little treatment in the literature, namely that of segmenting a sidescan sonar image in different types of seabed in an unsupervised manner. In particular, such an algorithm can be used to segment the areas of sonar images corresponding to sea-botom reverberation. This task is viewed here as an unsupervised texture segmentation problem. In addition, the extra assumption that the underlying image contains two dinstict regions is made.

There is a large litterature on unsupervised texture segmentation and a number of different approaches to the problem. In this paper a variational approach using region based active contours is adopted. Active contour methods have been extensively used over the past decade for boundary detection and image segmentation [4], [6], [7], [9]–[12], [18], [24], [26]. Originating from the classical parametric active contours (or snakes) [10], geometric active contours are represented implicitly as level sets of functions of two variables. Like the snakes they are usually derived by minimising a suitable data dependent energy functional, the central idea being that a minimum is attained when the contour optimally segments the underlying image. Unlike the snakes however they are able to make use of the level set techniques introduced by Osher and Sethian in [17] to handle topological changes and convergence problems. Energy functionals may depend on the data in a variety of ways but can be broadly divided into boundary and region based models depending on the type of information used to evolve the curve towards the boundary of the distict regions in the image. Region based models aim to evolve a curve by dynamically calculating some homogeneity measure over the entire region to be segmented. They are thus more suitable for images where the different regions are not defined through strong gradients as, for example, in blurry, noisy or textured images.

The active contour model used in this paper was first introduced by Chan and Vese in [6]. It is obtained as a special case of the Mumford Shah functional [14] and uses the mean intensity as a region homogeneity measure. In [5] the model has been extended to include vector-valued images and it is this form of the model that will be used here for texture segmentation of sidescan images. As expected from the results obtained in [6], this model turns out to be robust to the noise present in sonar data. In addition it holds good regularisation properties, similar to those of a Markov random field, as a result of the velocity dependence on the global region statistics

and the curvature of the contour. An additional feature of this model is its capacity to perform automatic feature selection when used for texture segmentation. This is implemented here by automatically readjusting the weights of the feature images to ensure that the contour evolution is driven by those that are the most discriminant.

The outline of the paper is as follows: In Section 2 we give some necessary background on level sets and briefly describe a general framework for region based geometric active contours. The Chan-Vese active contour models (scalar and vector-valued) are then described. In Section 3 the co-occurence matrices for texture extraction and their relationship with sum and difference histograms is reviewed. In Section 4 an algorithm for extracting the Haralick features from a sonar image is described. Finally, in Section 5 the implementation of the segmentation algorithm is described and several numerical results on simulated and real sonar images are presented.

II. The Chan-Vese active contour model for binary image segmentation.

A. Background

Let I be an image defined on a domain $\Omega \subseteq \mathbf{R}^2$. It will be assumed throughout this paper that I consists of two homogeneous regions Ω_1 and Ω_2 with Γ denoting their common boundary. The idea behind active contour segmentation methods is to evolve a curve $C(s,t)$ in Ω using a partial differential equation of the form:

$$\frac{\partial C}{\partial t} = F\vec{\mathbf{N}} \tag{1}$$

subject to an initial condition $C_0 = C(.,0)$ in such a way that the solution of (1) converges to the boundary Γ. Here $\vec{\mathbf{N}}$ is the inward normal vector of $C(s,t)$ and F is the speed with which C evolves in the direction of the vector $\vec{\mathbf{N}}$. The speed F may depend on many factors (e.g. local or global properties of the front, image data) but it is assumed that it is independent of the curve parametrisation. It has to be chosen in such a way that the evolving curve $C(s,t)$ is attracted by the boundary Γ of the two regions and becomes stationary at Γ.

The strength of this approach lies in its ability to make use of the level set methods introduced by Osher and Sethian in [17] and further developed by several authors over the past decade [15], [16], [21]. The level set formulation of equation (1) is given by:

$$\frac{\partial \phi}{\partial t} = F\|\nabla \phi\| \tag{2}$$

with initial condition $\phi(.,.,0) : \mathbf{R}^2 \rightarrow \mathbf{R}$ where the initial surface $\phi(.,.,0)$ is chosen so that its zero level set is given by the initial curve C_0 in (1), that is

$$\{(x,y) : \phi(x,y,0) = 0\} = C_0 \tag{3}$$

The family of curves $C(.,t)$, $t > 0$ satisfying (1) will then be given by the zero level sets of the surfaces $\phi(.,.,t)$, $t > 0$ that satisfy equation (2). In this way any topological changes in the evolving curve $C(.,t)$, as splitting or merging, can be handled naturally and powerful numerical schemes able to approximate the correct viscosity (weak) solution can be employed. [15], [21]

The velocity function F is usually derived through minimising a suitable image dependent energy although it is also possible to synthesize F directly from the image data [12], [20]. There are two types of approaches when choosing a velocity F: boundary-based and region-based. The former rely on the boundary Γ being described as the points in the image where $\|\nabla I\|$ is maximised and therefore tend to depend only on local information [4], [12], [20]. Region-based methods on the other hand aim to segment the two regions by considering various measures of homogeneity of each region. In this way global image information can be incorporated in the velocity function F [6], [9], [11], [18], [26]. A review of region-based methods can be found in [9].

A general framework that can unify many of the region-based approaches in the two region case was presented in [9]. This framework considers energies of the general form

$$E(C) = \int_{\Omega_{out}} k^{(out)}(x,y,\Omega_{out})dxdy + \int_{\Omega_{in}} k^{(in)}(x,y,\Omega_{in})dxdy + \int_C k^{(b)}(x,y)ds \tag{4}$$

where $\Omega_{out}, \Omega_{in}$ denote the outside and inside of the curve C. The kernels $k^{(out)}, k^{(in)}$ play the role of "region descriptors" and are modeled as a combination of features globally attached to the evolving regions $\Omega_{out}, \Omega_{in}$. The main result in [9] is the computation of the speed F that will make a contour obeying equation (1) evolve towards a mimimum of energy (4).

B. The Chan-Vese model for scalar images.

The approach used in this paper was first introduced in [6] and generalised for vector-valued images in [5]. It can be seen as a restricted form of the Mumford-Shah functional for segmentation first proposed in [14]. Alternatively it can be viewed as a special case of the framework in [9].

The general form of the Mumford-Shah functional is given by

$$E(u,C) = \mu\text{Length}(C) + \lambda \int_{\Omega} |I(x,y) - u(x,y)|^2 dxdy + \int_{\Omega \setminus C} |\nabla u(x,y)|^2 dxdy \qquad (5)$$

where $u : \bar{\Omega} \rightarrow \mathbf{R}$ is continuous and piecewise smooth, μ and λ are positive parameters. The segmentation problem that the minimisation of (5) is designed to solve can be stated as follows: Given I find a decomposition Ω_i of Ω and an optimal approximation u of I such that u varies smoothly within each Ω_i and rapidly or discontinuously across the boundaries of Ω_i. A minimiser (u, C) of $E(u, C)$ will be an "optimal" piecewise-smooth approximation u of the initial, possibly noisy, image I while C will approximate the edges of I.

In the case where I consists of two dinstinct regions and when restricting the approximations u of I to functions taking only two values, c_1, c_2 in the region inside C and the region outside C respectively, the Mumford Shah functional becomes

$$\begin{aligned} E(c_1, c_2, C) = \mu\text{Length}(C) \quad &+ \quad \lambda \int_{int(C)} |I(x,y) - c_1|^2 dxdy \\ &+ \quad \lambda \int_{ext(C)} |I(x,y) - c_2|^2 dxdy \end{aligned} \qquad (6)$$

In this case the constants c_1, c_2 are given by

$$\begin{aligned} c_1 &= m_{in}(C) = \text{average}(I) \text{ inside } C \\ c_2 &= m_{out}(C) = \text{average}(I) \text{ outside } C \end{aligned}$$

This leads to the minimisation problem:

$$\begin{aligned} \inf_C E(C) = \inf_C \{\mu\text{Length}(C) \quad &+ \quad \lambda_1 \int_{int(C)} |I(x,y) - m_{in}(C)|^2 dxdy \\ &+ \quad \lambda_2 \int_{ext(C)} |I(x,y) - m_{out}(C)|^2 dxdy\} \end{aligned} \qquad (7)$$

for $\mu \geq 0$, $\lambda_1, \lambda_2 > 0$. Now energy (7) is of the form of equation (4) considered in [9]. Following [9] the evolution equation derived from the minimisation of (7) is given by

$$\frac{\partial C}{\partial t} = \{\lambda_1 (I - m_{in}(C))^2 - \lambda_2 (I - m_{out}(C))^2 + \mu\kappa\}\vec{\mathbf{N}} \qquad (8)$$

where κ is the local curvature of the curve C.

Equation (8) can be implemented by using a level set method as explained above. Note that in this case the speed function F is given by

$$F = \lambda_1 (I - m_{in}(C))^2 - \lambda_2 (I - m_{out}(C))^2 + \mu\kappa \qquad (9)$$

and can therefore be extended naturally outside the curve C. As demonstrated in [6], this model has several advantages: ability to detect boundaries with very smooth or blurred boundaries (boundaries without gradient), automatic change of topology and automatic detection of interior contours, scale adaptivity (through the parameter μ) and robustness to noise.

The main limitation of the model comes from the fact that it can only discriminate regions which have different mean intensities. In particular it is, in general, unable to segment images with strong textures. One way to overcome this is to extract features $I_1, I_2, \ldots, I_n$ from the initial image I and apply the above algorithm directly to those. An even better segmentation result can be achieved by combining the information in $I_1, I_2, \ldots, I_n$. This is explained below.

C. Extending the model to vector valued images.

Let I be as before and let $I_1, I_2, \ldots, I_n$ be n images on the domain Ω that have been extracted from I. For example $I_i, i = 1, \ldots, n$ may be the output of a filter bank applied on I. The idea is to evolve a curve C in Ω as before but making use of the information contained in all of the images $I_i, i = 1, \ldots, n$. One way to achieve this is to evolve C under equation (1) where the velocity function F is given as a weighted average of terms over all images:

$$F = \frac{1}{n}\sum_{i=1}^{n}\{\lambda_i^{(in)}[I_i - m_i^{(in)}]^2 - \lambda_i^{(out)}[I_i - m_i^{(out)}]^2\} + \mu\kappa \tag{10}$$

where $m_i^{(in)}, m_i^{(out)}$ are the mean values of images I_i inside and outside C. This approach was first implemented in [5] where it is also shown that the speed F will evolve C under equation (1) towards the minimum of the energy

$$\begin{aligned} E(C) &= \frac{1}{n}\sum_{i=1}^{n}\lambda_i^{(in)}\int_{\Omega_{in}}|I_i(x,y) - m_i^{in}(C)|^2 dxdy \\ &+ \frac{1}{n}\sum_{i=1}^{n}\lambda_i^{(out)}\int_{\Omega_{out}}|I_i(x,y) - m_i^{out}(C)|^2 dxdy + \mu\int_C ds \end{aligned} \tag{11}$$

As in the one-dimensional case this approach also fits into the framework in [9].

The coefficients $\lambda_i^{(in)}, \lambda_i^{(out)}$ can be used as weights attached to each image depending on the amount of information that it contains. In our implementation the weights $\lambda_i^{(.)}$ are initially set equal to 1 and readjusted automatically as the curve evolve depending on the magnitude of the quantities $|m_i^{(in)} - m_i^{(out)}|$. In this way the active contour also performs a feature selection.

III. Co-occurence matrices and sum and difference histograms

The feature set of Haralick et al [8] is probably one of the most famous methods for texture analysis. It is based on the calculation of the co-occurence matrix, a second order statistics of the gray levels in an image window.

Let W be a $N \times M$ grey level image containing L quantised grey levels. In most applications W will be either an image consisting of a single texture/pattern or a window contained in an image I of which the local texture we are interested in analysing. It is a standard assumption in statistical texture modelling that W is a realisation of a stationary and ergodic random process $\{X(n,m)\}$, $n = 0, 1, \ldots, N-1$, $m = 0, 1, \ldots, M-1$. Let $\vec{d} = (d_1, d_2)$ denote a (small) displacemt vector and let $X_{\vec{d}}$ denote the process associated with the image $W(n+d_1, m+d_2)$. The probability of observing the gray levels $i, j \in \{0, 1, \ldots, L-1\}$ in relative position $\vec{d}$ is given by the joint pdf $p(i,j)$ of the processes $X, X_{\vec{d}}$.

The co-occurence matrix $C_{\vec{d}}(i,j)$ of W with parameters (d_1, d_2) is defined [8] to be the number of pixel pairs $(n,m), (n',m')$ in W that have intesity values i and j respectively. The normalised quantity

$$\hat{p}(i,j) = C_{\vec{d}}(i,j)/(N \times M) \tag{12}$$

is thus an estimate of the joint pdf $p(i,j)$.

In [8] a set of 14 textural features is proposed known as the Haralick Feature Set. All of these features are extracted directly from the normalised co-occurence matrices $\hat{p}(i,j)$. The 7 most commonly used of these features are:

1) *Energy*:

$$f_1 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \hat{p}(i,j)^2$$

2) *Contrast* or *Inertia*:

$$f_2 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i-j)^2 \hat{p}(i,j)$$

3) *Correlation*:

$$f_3 = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i-\mu)(j-\mu)\hat{p}(i,j)$$

where μ denotes the estimated mean of the process, that is

$$\mu = \sum_{i=0}^{L-1} i \hat{p_W}(i)$$

where $\hat{p_W}$ is the histogram of W.

4) *Entropy*:

$$f_4 = -\sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\hat{p}(i,j)\log(\hat{p}(i,j))$$

5) *Homogeneity*:

$$f_5 = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}\frac{1}{1+(i-j)^2}\hat{p}(i,j)$$

6) *Cluster shade*:

$$f_6 = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}(i+j-2\mu)^3\hat{p}(i,j)$$

7) *Cluster prominence*:

$$f_7 = \sum_{i=0}^{L-1}\sum_{j=0}^{L-1}(i+j-2\mu)^4\hat{p}(i,j)$$

Experiments have shown (see [23] and references therein) that these features are significant in terms of their capacity to measure visually perceivable qualities of textures. On an intuitive level, for example, *energy* can be thought of as a measure of homogeneity of a texture image, *contrast* measures the amount of variation in gray tones present in an image, *correlation* is a measure of gray tone linear dependencies and *entropy* is a measure of complexity of an image.

In [23] Unser proposes an alternative, significantly more efficient way of computing the Haralick Feature Set. His approach relies on the following argument:

Let $X, X_{\vec{d}}$ be the random variables defined above. It can be shown that the random variables $S = X + X_{\vec{d}}$ and $D = X - X_{\vec{d}}$ are uncorrelated, that is

$$covar(S, D) = 0.$$

When S and D are independent their joint pdf p can be computed from

$$p(i,j) = P(S = i+j, D = i-j) = P(S = i+j)P(D = i-j)$$

This expression is always correct for Gaussian random variables. For arbitrary uncorrelated random variables the last equality is not always satisfied. However the joint pdf p can be approximated by the product on the right hand side, that is

$$p(i,j) \approx p'(i,j) = c_0 P(S = i+j)P(D = i-j) \tag{13}$$

where c_0 is a normalisation constant chosen so that

$$\sum_i \sum_j p'(i,j) = 1$$

The approximation error in (13) is given by the relative entropy $I(p,p')$ which is a measure of interdependence of S and D. It can be seen that

$$I(p,'p) = H_S + H_D - H_X - \log c_0 \tag{14}$$

where H_S, H_D, H_X denote the entropy of the variables S, D and X.

Returning to the problem of computing the Haralick Feature Set for W, equation (13) can be translated as follows:

The co-occurence matrix $\hat{p}$ of W with parameters $\vec{d} = (d_1, d_2)$ can be estimated by the histograms of the sum and difference images W_S and W_D

$$W_S(k,l) = W(k,l) + W(k+d_1, l+d_2)$$

$$W_D(k,l) = W(k,l) - W(k+d_1, l+d_2).$$

(Note that the images W_S, Y_D will have twice the dynamic range of the original image W and lower spatial dimensions depending on $\vec{d} = (d_1, d_2)$.)

The histograms of W_S and W_D can, in turn, be estimated directly from W. In other words

$$\hat{p}(i,j) \approx \hat{p}_S(i+j)\hat{p}_D(i-j) \tag{15}$$

where $\hat{p}_S$, $\hat{p}_D$ denote the histograms of W_S and W_D.

This leads to the following expressions of the features $f_1 \ldots f_7$ in terms of the sum and

difference histograms $\hat{p_S}$, $\hat{p_D}$:

$$
\begin{aligned}
f_1 &\approx \sum_i \hat{p_S}(i)^2 \sum_j \hat{p_D}(j)^2 \\
f_2 &= \sum_j j^2 \hat{p_D}(j) \\
f_3 &= \frac{1}{2}\left\{ \sum_i (i-2\mu)^2 \hat{p_S}(i) - \sum_j j^2 \hat{p_D}(j) \right\} \\
f_4 &\approx -\sum_i \hat{p_S}(i) \log \hat{p_S}(i) - \sum_j \hat{p_D}(j) \log \hat{p_D}(j) \\
f_5 &= \sum_j \frac{1}{1+j^2} \hat{p_D}(j) \\
f_6 &= \sum_i (i-2\mu)^3 \hat{p_S}(i) \\
f_7 &= \sum_i (i-2\mu)^4 \hat{p_S}(i)
\end{aligned}
\tag{16}
$$

The symbol $\approx$ indicates that the left hand side can be approximated by the expression on the right with a quantifiable error.

It is now clear why using the sum and difference histograms to compute the features $f_1 \ldots f_7$ results in significant gain in computation time: a double summation is replaced by a single one and memory requirement reduces by a factor $L/4$.

IV. Extracting the Haralick features

In this section we briefly outline our algorithm for feature extraction of the Haralick feature set.

Let I be an $N_x \times N_y$ image to be segmented. The main idea as in many extraction algorithms for image segmentation is as follows: consider a sequence of windows in I, extract the features given by (1) for each window and thus create a vector valued image defined on the centre of each window. The problem with this approach is that computing the sum and difference for each window separately is computationally expensive and would lead to a very slow implementation. Instead we use the following variation of the above:

1) Set up a lattice of pixels $\mathbf{L} = \{(n, m)\}$. Fix the density of $\mathbf{L}$ via the horizontal and vertical steps s_x, s_y. Fix the window size for feature extraction to $w_x \times w_y$. This should reflect the geometry of the image and the scale of the information to be extracted.

2) Fix a set of displacement parameters $\{\vec{d_i}\}$. This can be represented by a distance d and a set of directions θ_i. Typically θ takes the values $-\pi/2, -\pi/4, 0, \pi/4$.
3) Locally equalise image I using the pixels in $\mathbf{L}$ and window size $w_x \times w_y$. The aim here is to minimise the discriminative role of the first order statistics.
4) Compute the sum and difference I_S and I_D of the entire image I.
5) Set up sequences of windows $W_S(n,m)$, $W_D(n,m)$ centred at (n,m) in I_S and I_D respectively.
6) For each one of the pixels (n,n) use the histograms of W_s, W_D to compute features $(f_1,\ldots,f_7)$ as in (1). Realise features as a vector $(I_1,\ldots,I_7)$ of images extracted from the original image I.

The main benefit drawn from this approach is that the sum and difference of an image is computed only once over the entire image I.

We close this section by giving an example in figure 1 of a sonar image and the extracted features $f_1 - f_7$ with $s_x = s_y = 2$ and $\theta = -\pi/2$. The extraction time for an image of this size (400×360) was $45sec$. It's worth noting however that the local equalisation step takes up about half of this time.($\approx 21sec$).

V. Implementation and experimental results.

A. Implementation of the segmentation algorithm.

Let I be a two-region greyscale image and let $\vec{I} = (I_1,\ldots,I_n)$ denote a set of feature images extracted from I and defined on a common domain Ω. The following partial differential equation is implemented numerically by making use of a level set method as explained in Section 2:

$$\frac{\partial C}{\partial t} = \{\frac{1}{n}\sum_{i=1}^{n}\{\lambda_i^{(in)}[I_i - m_i^{(in)}]^2 - \lambda_i^{(out)}[I_i - m_i^{(out)}]^2\} + \mu\kappa\}\vec{\mathbf{N}} \tag{17}$$

with initial condition $C = C_0$ and where $m_i^{(in)}, m_i^{(out)}$ are the mean values of images I_i inside and outside C. More precisely the PDE

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= (F_0 + F_1)|\nabla\phi| \qquad (18)\\ F_0 &= \frac{1}{n}\sum_{i=1}^{n}\{\lambda_i^{(in)}[I_i - m_i^{(in)}]^2 - \lambda_i^{(out)}[I_i - m_i^{(out)}]^2\} \\ F_1 &= \mu\kappa = \mu\nabla\cdot\left(\frac{\nabla\phi}{|\nabla\phi|}\right) \end{aligned}$$

with initial condition $\phi(x, y, 0) = sgnd(x, y, C_0)$ is discretised. Here $sgnd(., ., C_0)$ denotes the signed distance function from the initial curve C_0 and is defined by

$$sgnd(x, y, C_0) = \begin{cases} -dist(x, y, C_0) \text{ if } (x, y) \text{ lies inside } C_0 \\ dist(x, y, C_0) \text{ if } (x, y) \text{ lies outside } C_0 \end{cases} \tag{19}$$

We use a first order monotone scheme to approximate the term $F_0|\nabla\phi|$ and a first order central difference approximation to the the curvature term $F_1|\nabla\phi|$. For computational efficiency the values of ϕ are updated only in a narrow band around the zero level set of ϕ. To ensure that the evolving curve remains well within the narrow band domain it is necessary to reinitialise ϕ when the zero level set of ϕ gets close to the boundary of the band. This is done by resetting ϕ to be equal to the signed distance from its zero level set using (19). We have found that in many cases it is necessary to reinitialise more often as the gradient of ϕ tends to get very large affecting the convergence of the numerical solution. In the case of sonar images the "roughnes" of the image data itself seems to affect the convergence of the numerical approximation. It is possible that a higher-order space accurate monotone scheme would improve the convergence and accuracy of the solution.

Another source of difficulties is the non-convexity of the energy associated with the speed F_1. Such an energy has more than one local minimum and this makes the algorithm susceptible to becoming stuck near local minima with the solution depending on the initialisation. We found that in general initialising with a sequence of uniformly distributed circles over the entire image gives better segmentation results and convergence is faster than when initialising with a single curve. For certain types of images however this is not true; for example images where the object or region in the central part of the image is small relative to the background region segment better when initialising with a single curve that intersects with the central region.

The parameters $\lambda_i^{(in)}, \lambda_i^{(out)}$ determine the degree of contribution of image I_i in the final result. In [5] the λ parameters are used to filter high frequency noise from different channels. In the case where the I_is are the output of a filter bank (e.g. Haralick features) the level of noise is the same in all channels and the primary role of the parameters is one of feature selection. In our implementation the coefficients $\lambda_i^{(in)}, \lambda_i^{(out)}$ are initially set to be equal to 1 for the first 300 iterations. They are subsequently reset at certain predetermined times to:

$$\lambda_i^{(in)} = \lambda_i^{(out)} = \frac{|m_i^{(in)} - m_i^{(out)}|}{M} \tag{20}$$

where

$$M = max_{1 \leq i \leq n} |m_i^{(in)} - m_i^{(out)}|.$$

This choice ensures that the features with maximum discriminatory capacity drive the curve evolution. As examples in the next section will demonstrate this type of feature selection can greatly improve the segmentation result.

B. Experimental results

In the last section of this paper we present numerical results of the proposed method on synthetic and real sonar images. One of the problems when segmenting sonar images is the noise or other artefacts resulting from imperfect data acquisition methods. It is thus desirable to first validate the proposed method on synthetic data. We then present some examples of images which segment well and for which the segmentation is robust under different initialisation. Finally examples of images which are difficult to segment with the proposed method are also given.

In the experiments described below the algorithm parameters were chosen as follows: In the feature extraction algorithm the window sizes w_x, w_y were set to $w_x = w_y = 21$, the step sizes s_x, s_y were set to $s_x = s_y = 2$, the grey levels were quantised to 32 and the set of displacement vectors $\{\vec{d}_i\}$ for the calculation of the co-occurence matrices is determined by: $|\vec{d}_i| = 2$, $\theta_i \in \{-\pi/2, -\pi/4, 0, \pi/4\}$. In the curve evolution scheme parameters $\lambda_i^{in}, \lambda_i^{out}$ were chosen as in (20) at iterations $100, 200, \ldots, 500$ but where we have also set $\lambda_i^{(in)} = \lambda_i^{(out)} = 0$ for those indices i where $\frac{|m_i^{(in)} - m_i^{(out)}|}{M} < 0.5$. Parameter μ controls the smoothness in the contour and thus the sharpness in the boundaries of the segmented regions. The value $\mu = 5 \cdot 10^{-4} \cdot (255)^2$ was found to be a good compromise between the level of accuracy in boundary detection and the level of noise or scale of detail within each separate region that needs be left undetected.

Figures 2, 3 and 4 show the segmentation results for three simulated side scan sonar images. The images in figures 2 and 3 contain two types of seabed whereas figure 4 consists of three types of seabed with the middle and top regions having a higher degree of similarity than any of the other two pairs. All three segmentation results are robust to different initialisations. Initialising with an array of circles uniformly distributed over the entire image speeds up convergence by a factor of two compared to initialising with a curve whose interior consists of a single connected component (e.g a single circle or rectangle). In figure 4 the two visually more similar regions (top and middle) have been given the same label.

Figures 5, 7, 8, 9 and 10 contain segmentation results on real raw sonar data. As can be seen the data is noisy and contains various artefacts. These results are robust to initialisation and to feature selection: running the algorithm with $\lambda_i = 1$ for all i gives the same segmentation as setting the coefficients as explained above. The speed of convergence however is greatly reduced when the feature selection is activated as only a few of the feature images contribute to the calculation of the curve flow. Figure 6 displays the six most discriminant features extracted from the image in figure 5 and selected automatically. These correspond to the features f_2 (contrast) in (1) in four different directions ($\theta = -\pi/2, -\pi/4, 0, \pi/4$) and f_5 (homogeneity) in the directions $\theta = -\pi/4$ and $\theta = 0$. Execution time per iteration for an image of 400×300 was $0.2sec$.

Figure 11 shows examples of images for which some feature selection is necessary for correct segmentation. The images in the middle column were obtained by setting the coefficients λ_i as described in the beginning of this section. The images in the right column were obtained by setting all of the λ_i equal to 1.

Figure 12 contains examples of images that are harder to segment using the proposed algorithm. The image in the top row can be succesfully segmented when initialising with a small circle (about a third of the image radius) placed in the middle of the image. This result is shown in the middle column. The result on the right is the one obtained when the initial curve has a large intersection with the surrounding region. The image in the middle row exhibits the opposite behaviour: initialising with too small a circle does not allow for region growing in order to capture the top half of the rippled seabed (middle). Initialising with a 3×4 array of circles however gives a good segmentation result (right). The difficulties in the last image come from the vertical line across the middle and the similarity of the two types of seabed present. As in the top image, initialising with a circle having a small intersection with the coarser texture gives a good result. Problems of this type can be addressed by using a multiresolution approach. For example evolving a curve on features extracted at a lower resolution can be used for initialisation. Alternatively a different pixel based segmentation algorithm can be used after the feature extraction as an initial step –e.g.K-means or fuzzy K-means– followed by regularisation with the proposed algorithm.

VI. Conclusions

In this paper an unsupervised binary segmentation algorithm is proposed and applied in particular to side scan sonar images. It combines the Haralick features for texture and an active

contour model for vector valued images proposed in [5]. The implementation makes use of the sum and difference histograms for computing the co-occurence matrix and level set methods for the curve evolution. A suitable set of parameters is identified and used throughout the experiments described here. The resulting algorithm is validated on several simulated and real sonar data and is robust to the noise naturally present in sonar data. As most of the examples in the last section demonstrate the proposed algorithm is also robust to other artefacts resulting from the data acquisition process. Most of the images in our data set were succesfully segmented even in cases where the two regions have similar textures and/or poorly defined boundaries. The main difficulty comes from the initialisation in the curve evolution a problem which can be adressed by a multiresolution approach.

The proposed method can be readily extended to other contexts. Different type of features may be used. By taking a slightly different viewpoint the method may also be used for sensor fusion (i.e. bathymetry, video). Finally it is possible to extend the method to segment sonar images in up to four classes by considering a coupled evolution of two contours [7]. Overall the results obtained in this paper suggest that curve evolution is a valid tool in sonar analysis.

REFERENCES

[1] J. M. Bell, M. J. Chantler, and T. Wittig. Sidescan sonar: a directional filter or seabed texture? *IEE Proc. Radar, Sonar Navig.*, 146(1):65–72, 1999.

[2] L. Bjorno and et al. Identification of seabed data from acoustic reflections: theory and experiment. *Acta Acoustica*, 2:359–374, 1994.

[3] D. R. Carmichael and et al. Seabed classification through multifractal analysis. *IEE Proc. Radar, Sonar Navig.*, 143(3):140–148, 1996.

[4] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. *Int. J. Comput. Vis.*, 22(1):61–79, 1997.

[5] T. Chan, B. Sanberg, and L.Vese. Active contours without edges for vector-valued images. *J.Visual Comm. Image Rep*, 11:130–141, 2000.

[6] T. Chan and L. Vese. Active contours without edges. *IEEE Transactions on Image processing*, 10(2):266–277, 2001.

[7] T. Chan and L. Vese. A level set algorithm for minimising the mumford- shah functional in image processing. In *Comput. Soc. Proc. of the 1st IEEE Workshop on Variational and Level Set Methods in Computer Vision*, pages 161–168, 2001.

[8] R. M. Haralick and et al. Textural features for image classification. *IEEE Trans. Syst. Man, Cybern,*, pages 610–621, 1973.

[9] S. Jehan-Besson, M. Barlaud, and G. Aubert. Defornable regions driven by an eulerian accurate minimisation method for image and video segmentation. *Int. J. Comp. Vis.*, 53(1):45–70, 2003.

[10] M. Kass, A. Witkin, and D. Terzopoulos. Snakes:active contour models. *Int. J. Comput. Vis.*, 1:321–331, 1988.

[11] J. Kim, J.W. Fisher, and et al. Non-parametric methods for image segmentation using information theory and curve evolution. In *Proceedings IEEE Int. Conf. Image Proc.*, 2002.

[12] R. Malladi, J.A. Sethian, and B. C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Trans. pattern Anal. machine Intell.*, 17:158–175, 1995.

[13] M. Mignotte, C. Collet, P. Perez, and P. Bouthemy. Sonar image segmentation using an unsupervised hierarchical mrf model. *IEEE Trans. Image Processing*, 9(7):1216–1231, 2000.

[14] D. Mumford and J. Shah. Optimal approximations by piecwise smooth functions and associated variational problems. *Comm.Pure Applied Math.*, pages 577–685, 1989.

[15] S. Osher and R. Fedkiw. *Level Set Methods and Dynamic Implicit Surfaces*. Springer, 2003.

[16] S. Osher and N. Paragios. *Geometric Level Set Methods in Imaging, Vision and Graphics*. Springer, 2003.

[17] S. Osher and J. Sethian. Fronts propagating with curvature-dependent speed:algorithms based on hamilton-jacobi formulation. *J.Comput.Phys.*, 79:12–49, 1988.

[18] N. Paragios and R. Deriche. Geodesic active regions: A new framework to deal with frame partition problems in computer vision. *J. Vis. Comm. Im. Repr.*, 13:249–268, 2002.

[19] S. Reed, Y. Petillot, and J. Bell. An automatic approach to the detection and extraction of mine features in sidescan sonar. *IEEE J. Ocean Eng.*, 28(1):90–105, 2003.

[20] G. Sapiro. *Geometric Partial Differential Equations and Image Analysis*. Cambridge University Press, 2001.

[21] J.A. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge University Press, 1999.

[22] D.D. Sternlicht and C.P. de Moustier. Time-dependent seafbor acoustic backscatter (10-100khz). *J. Acoust. Soc. Am.*, 114:2709–2725l, 2003.

[23] M. Unser. Sum and difference histograms for texture classifi cation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(7):717–727, 1989.

[24] A. Yezzi, A. Tsai, and A.S. Willsky. A statistical approach to snakes for bimodal and trimodal imagery. In *International Conference on Computer Vision-volume 2*, 1999.

[25] B. Zerr, E. Maillard, and D. Gueriot. Sea-fbor classifi cation by neural hybrid system. *IEEE Proc. OCEANS '94*, pages 239–243, 1994.

[26] S.C. Zhu, T.S. Lee, and A. Yuille. Region competition: unifying snakes, region growing and bayes/mdl for multi-band image segmentation. *IEEE Trans. Patt. Anal. Machine Intell.*, 18(9):884–900, 1996.

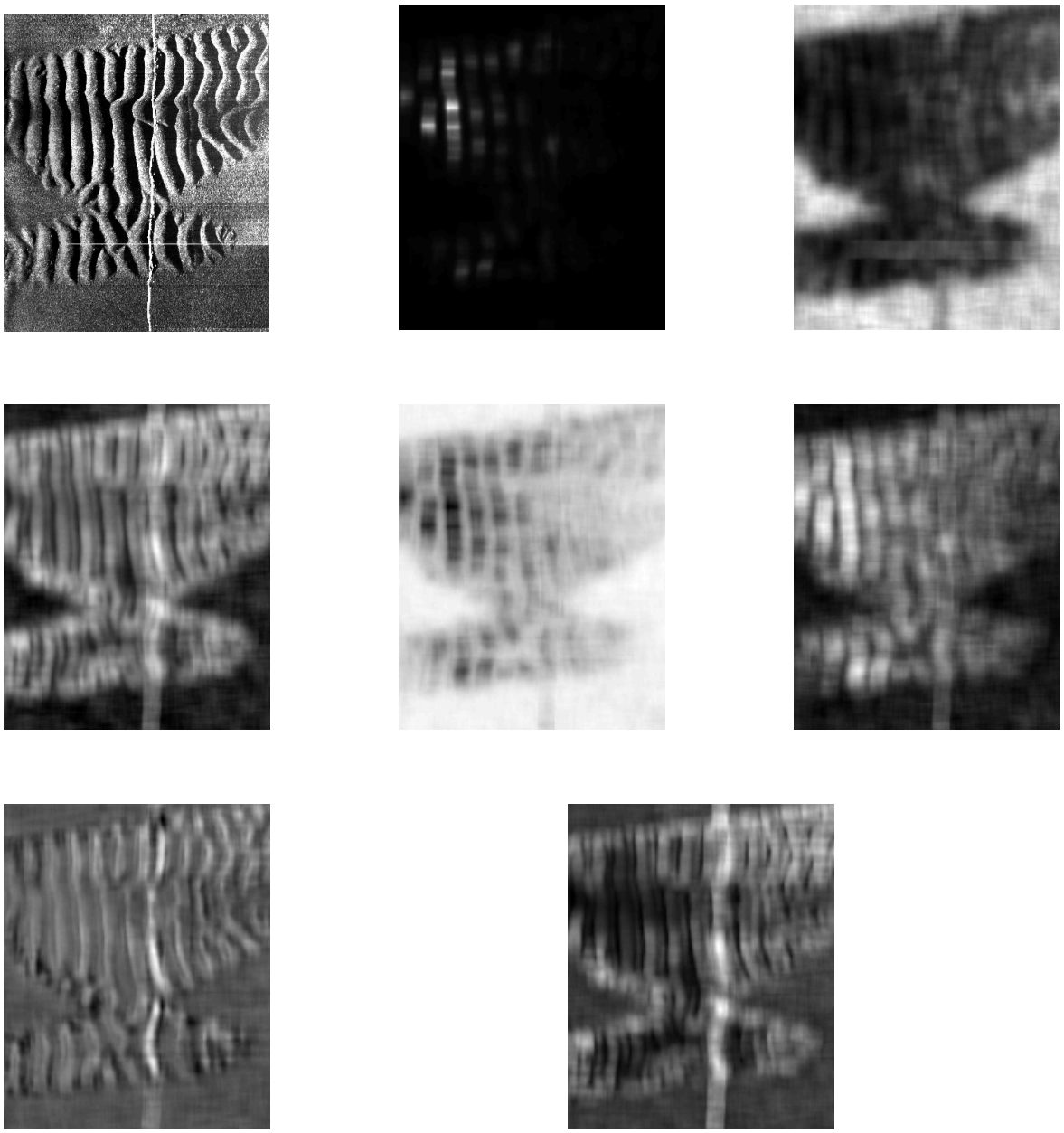

Fig. 1. From left to right: Sonar image and extracted features (energy, contrast, correlation, entropy, homogeneity, cluster shade and cluster prominence).

Fig. 2. Simulated image and segmentation result

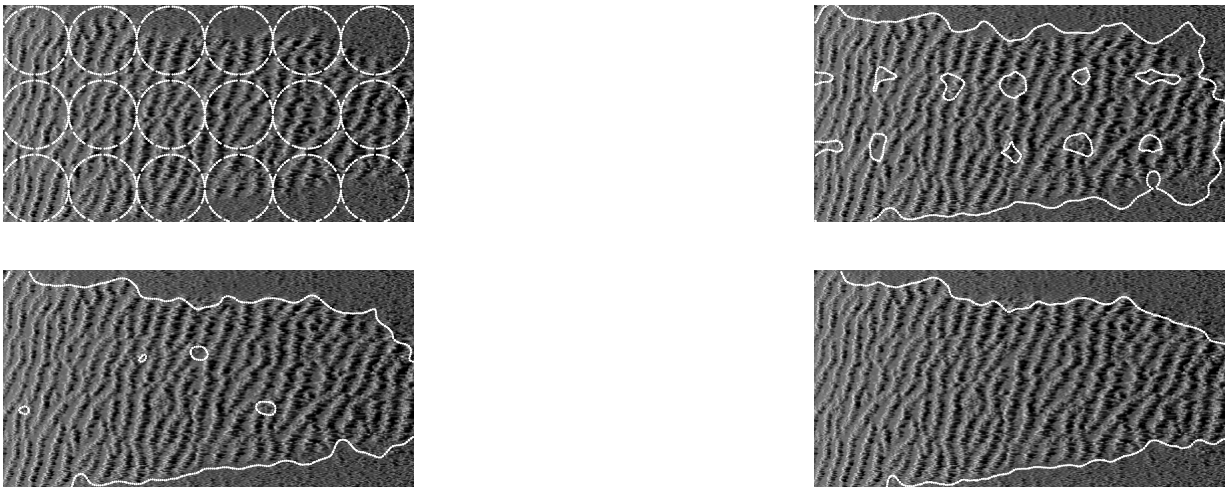

Fig. 3. Curve evolution on simulated image at iterations 0, 20, 40 and 100.

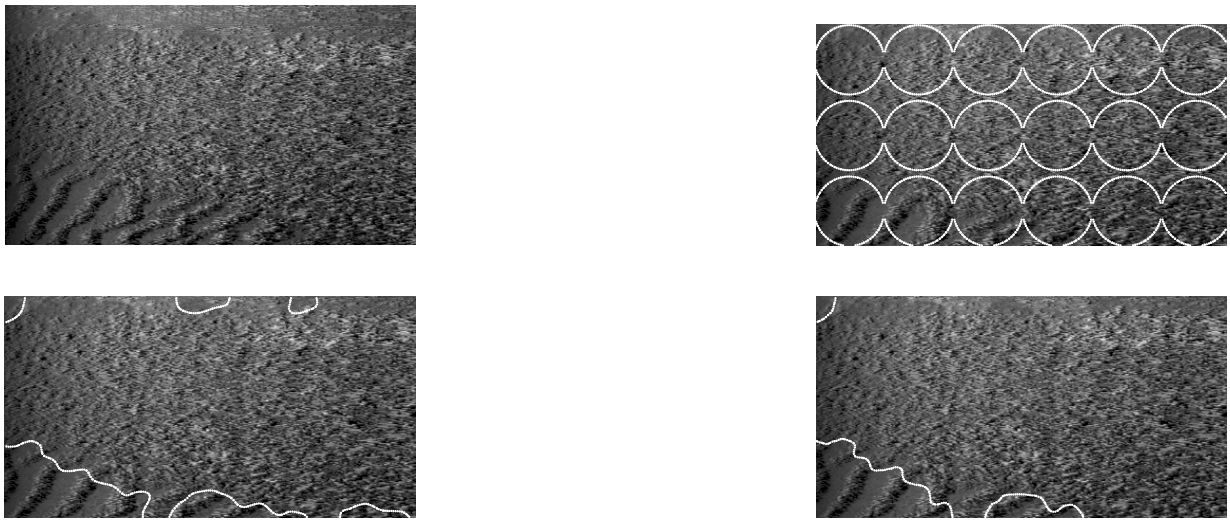

Fig. 4. Curve evolution on simulated image containing three types of seabed at iterations 0, 200 and 1000.

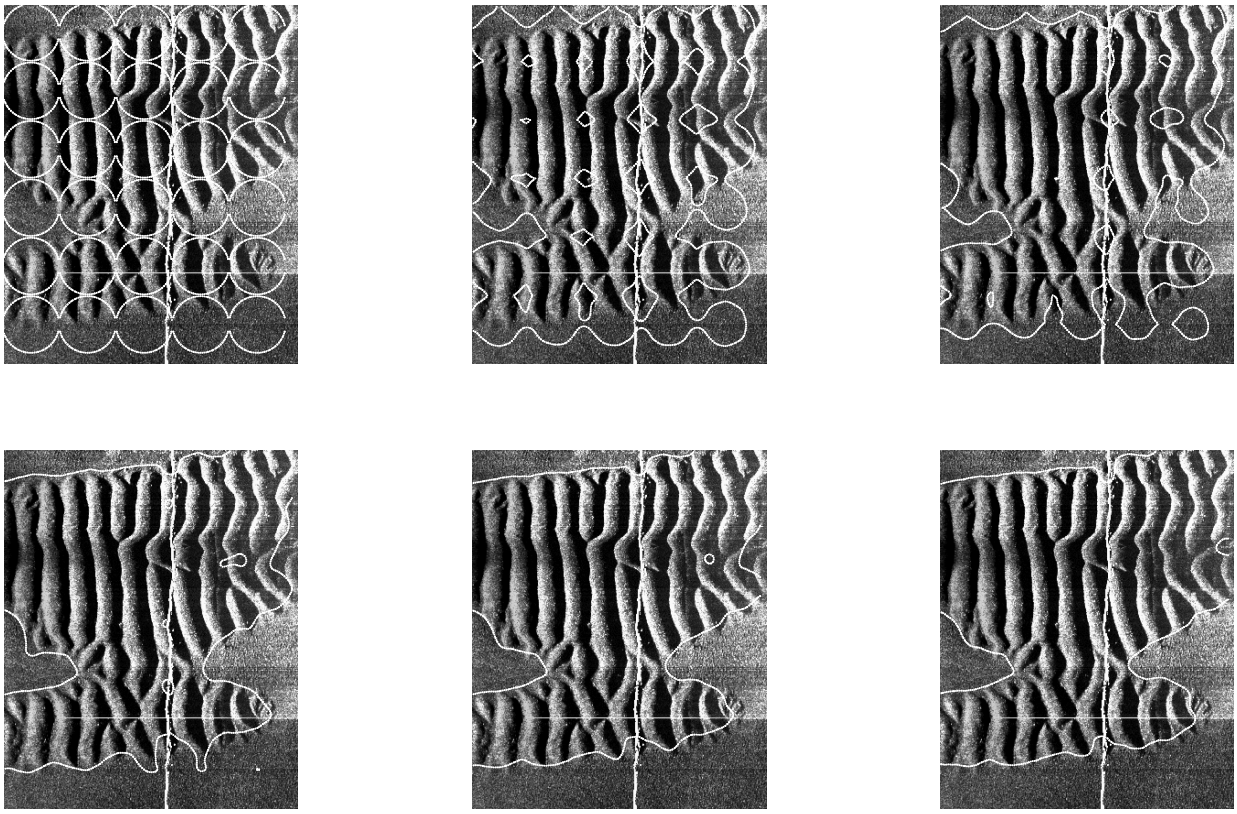

Fig. 5. Contour evolution at $0, 10, 20, 40, 60$ and 200 iterations.

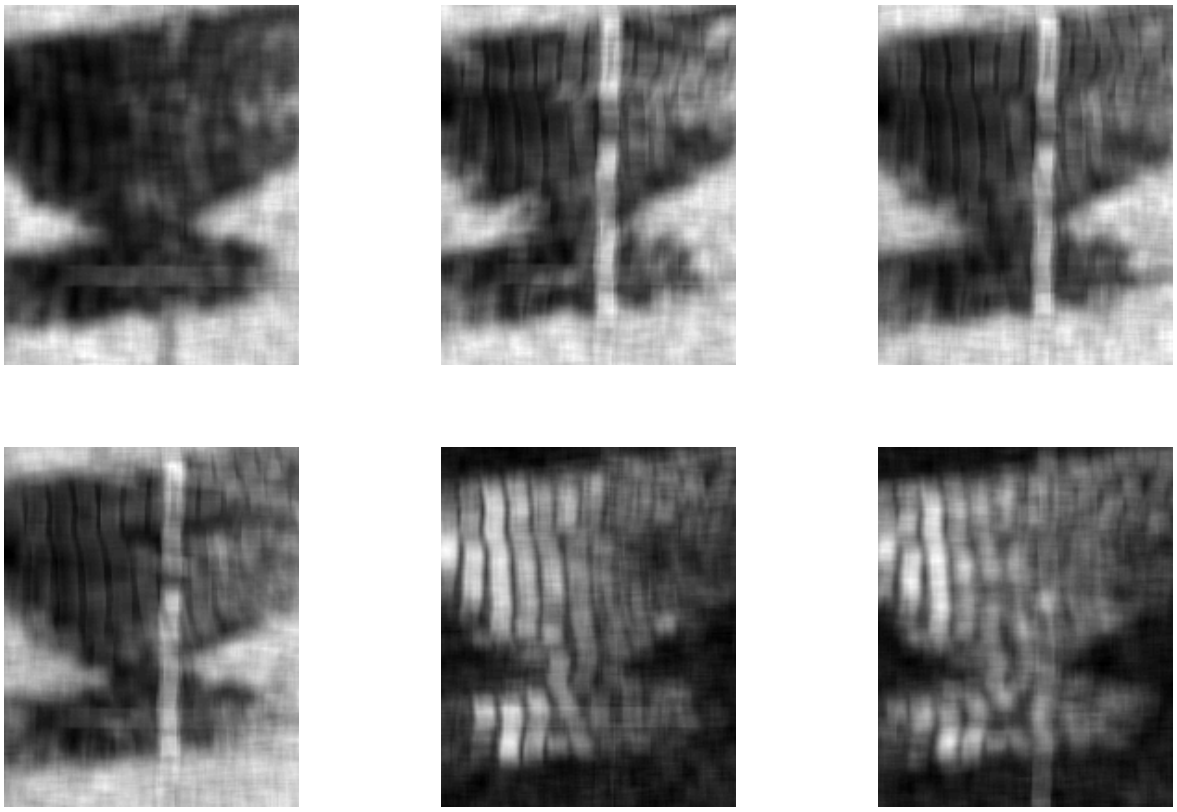

Fig. 6. The six most discriminant features selected for segmentation of the image in figure 5.

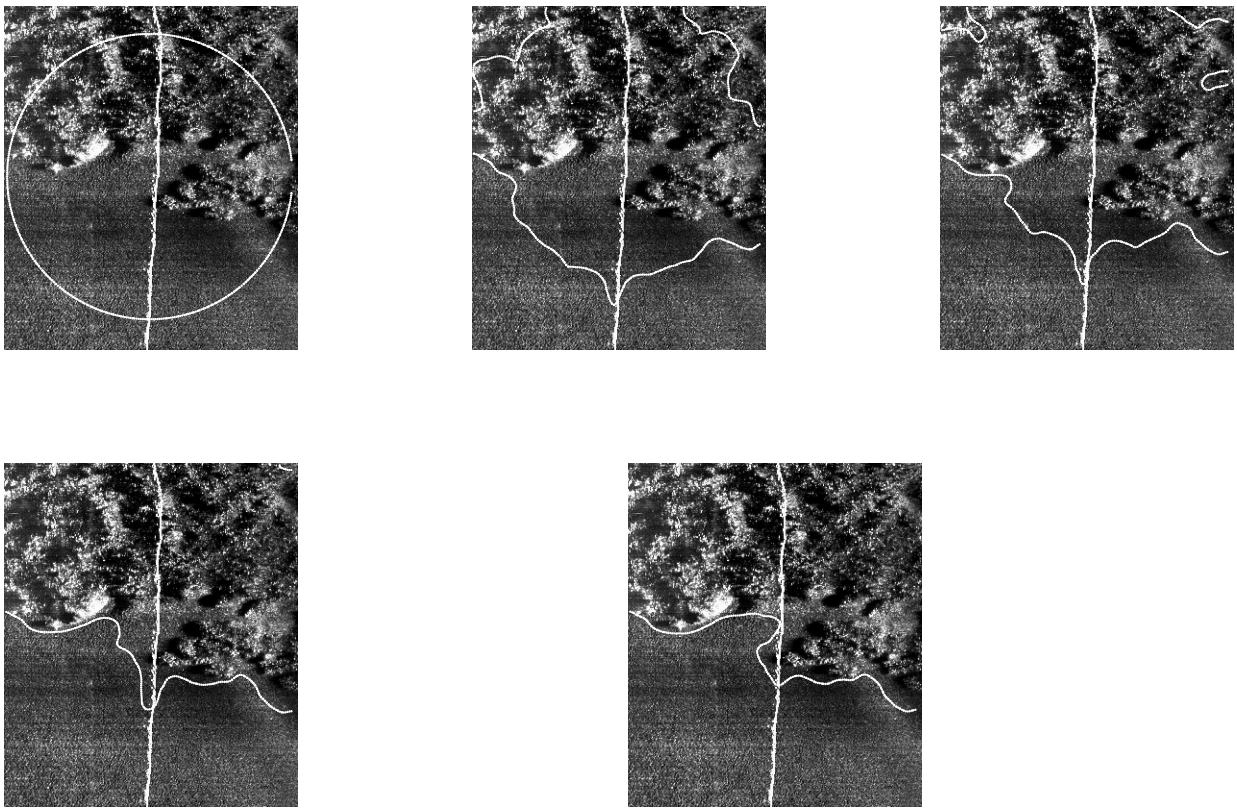

Fig. 7. Contour evolution at $0, 80, 140, 240$ and 460 iterations.

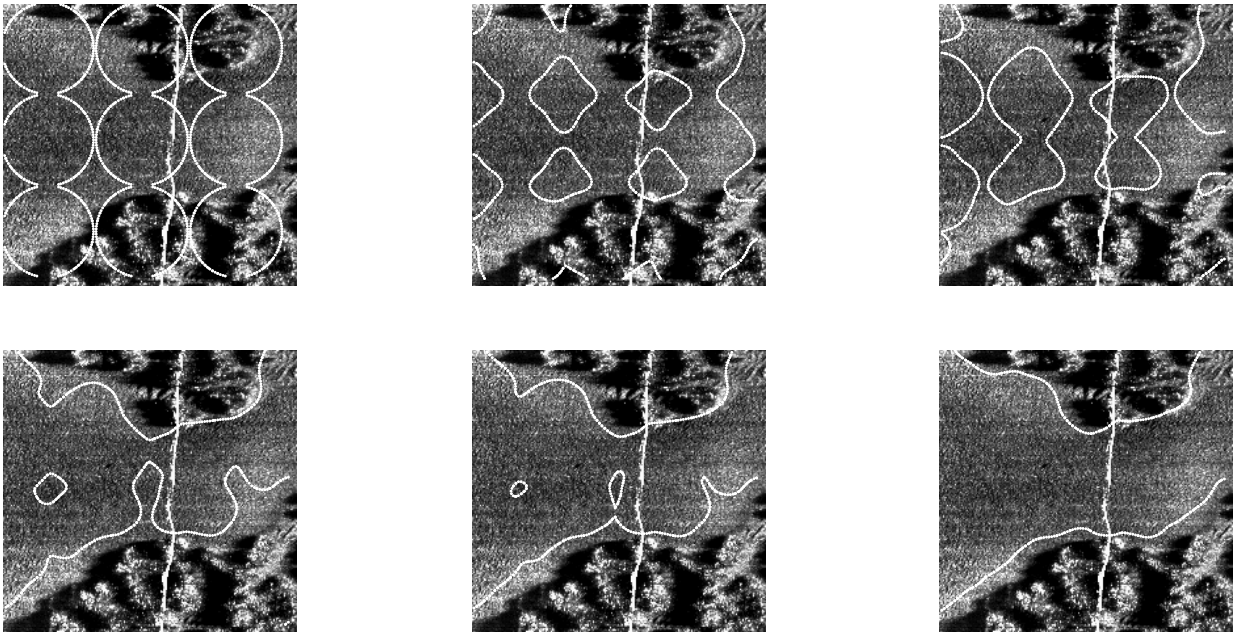

Fig. 8. Contour evolution at $0, 20, 50, 70, 90$ and 120 iterations.

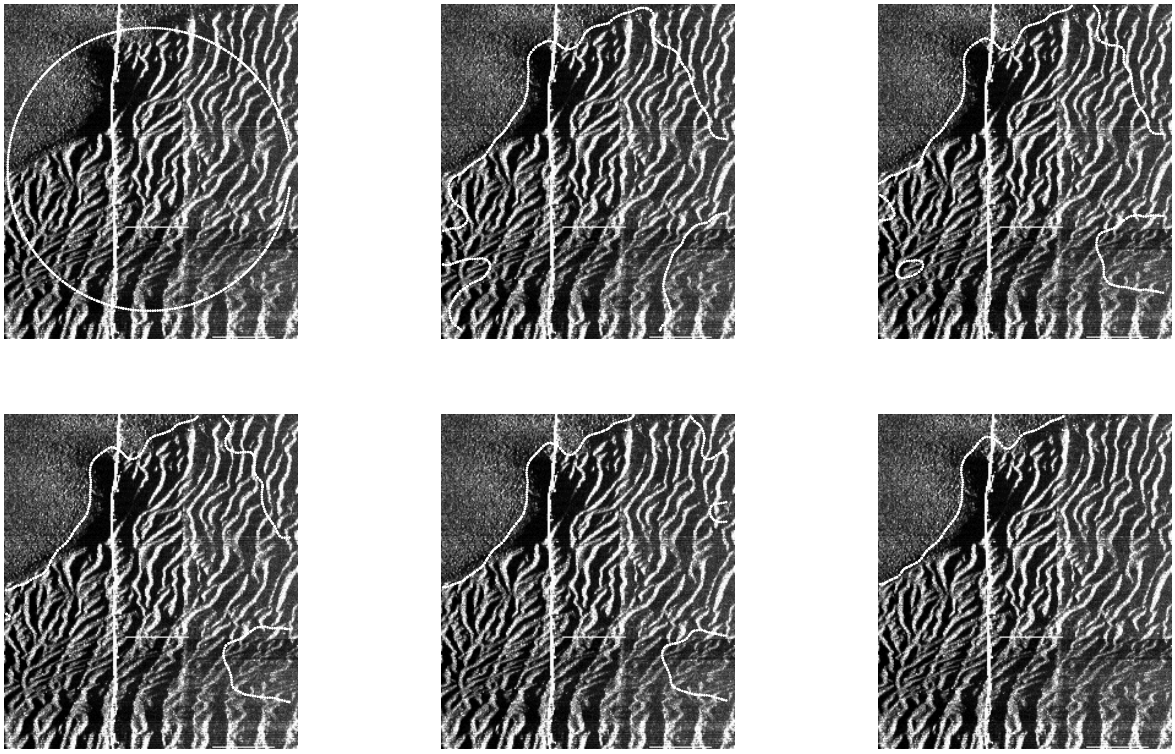

Fig. 9. Contour evolution at $0, 60, 180, 260, 400$ and 920 iterations.

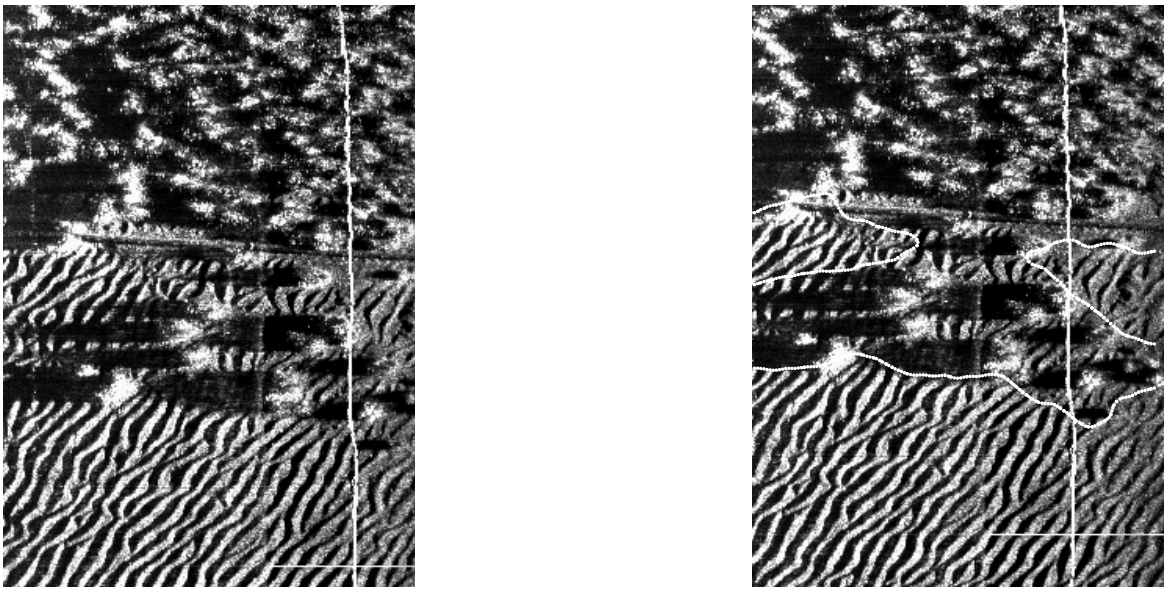

Fig. 10. From left to right: Original image and segmentation result.

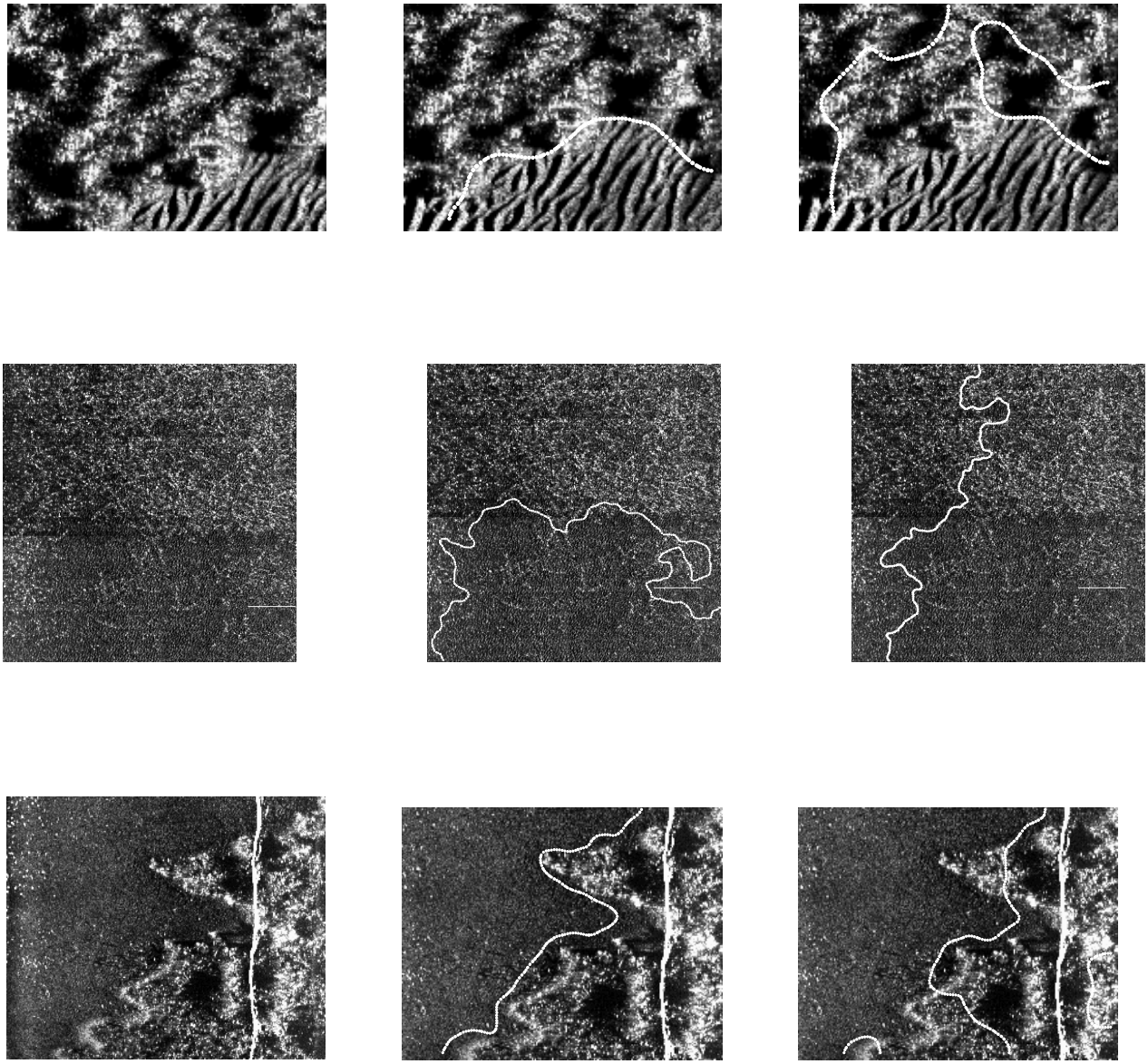

Fig. 11. From left to right: Original image and segmentation results with and without feature selection .

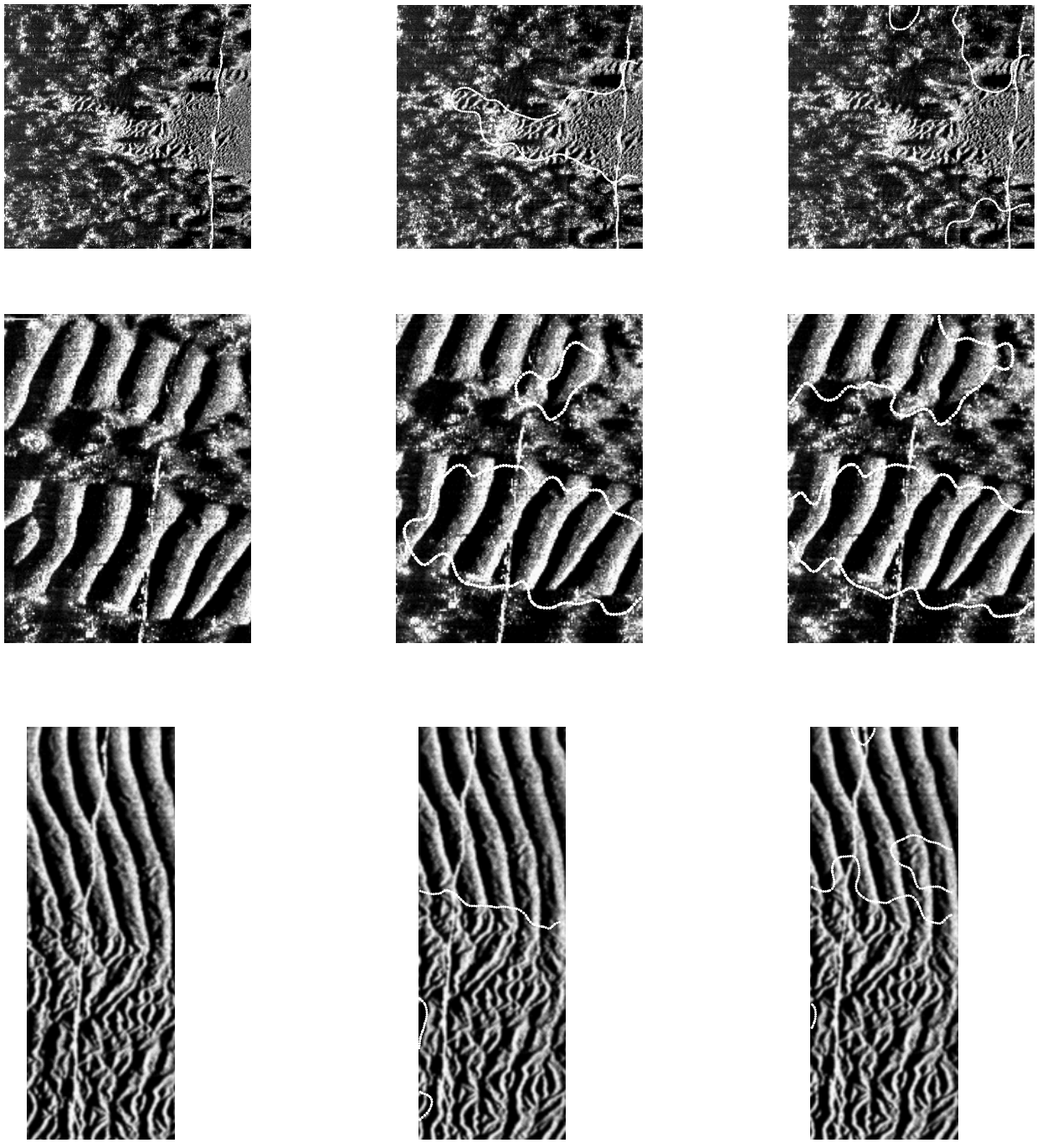

Fig. 12. From left to right: Original image and segmentation results when initialising with one circle (middle) and an array circles (right).