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 [1] is then applied to the vector valued image extracted from the original data. Our implementation includes a new 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 parameters and contour initialisation are also 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 geo-acoustic data of the sea bottom(see [2], [3] and

references therein). The second approach is image based and makes use of, primarily, side-scan and bathymetric sonar data [4]–[7].

Side-scan 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 side-scan image, each time instant referring to a point in the seabed. The side-scan 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 side-scan 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 neighbouring 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 types of seabed producing a distinct backscatter response.

Side-scan sonar image analysis is used in a number of applications, from object localisation and identification to seabed classification and 3-D reconstruction. The analysis falls into two categories: classification using texture features and supervised learning [4], [5] and unsupervised segmentation/classification based on Bayesian clustering methods using grey levels or texture features [6], [7]. 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 [8]. Existing unsupervised segmentation algorithms for side-scan data are largely concerned with identifying objects on the seabed by segmenting the image in shadow, non-shadow areas ([6] and references therein) or into three regions corresponding to shadow, echo and sea-bottom reverberation ([7]). The approach used in [6], [7] is based on a novel hierarchical Markov tree model with prior knowledge being incorporated in the model.

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 side-scan 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-bottom reverberation. This task is viewed here as an unsupervised texture segmentation problem. In addition, the extra assumption that the underlying image contains two distinct regions is made.

There is a large literature on unsupervised texture segmentation and a number of different approaches to the problem. In this paper a variational approach using region based active contours based on level sets is adopted.

Active contour methods have been extensively used over the past decade for boundary detection and image segmentation [9]–[18]. Originating from the classical parametric active contours (or snakes) [9], 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 [19] 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 toward the boundary of the distinct 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.

One way of using this type of method for texture segmentation is to evolve a curve simultaneously over a set of feature images extracted from the original image. Several works based on this approach exist in the literature mostly using Gabor filters and wavelets as the texture feature extraction technique([20]–[24], [25]). Of these works only [24], [25] present unsupervised segmentation methods and are closely related to the techniques used here. In particular the active contour model and feature selection method are very similar to those in [25]. The primary aim of the work in this paper, however, is to apply these techniques referred to sonar data, and in particular to side-scan sonar images, while taking advantage of the narrower domain of application in order to improve and refine existing methods.

The active contour model used in this paper was first introduced by Chan and Vese in [14]. It is obtained as a special case of the Mumford Shah functional [26] and uses the mean intensity

as a region homogeneity measure. One of the attractive features of this model is that it is free of any *a priori* assumptions or statistical modelling of the image. The latter is desirable as textures present in side-scan images cannot be modelled by one type of distribution. Moreover this model is robust to noise and 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. In [1] 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 side-scan images. More specifically the Chan-Vese vector valued model is combined with the Haralick texture features calculated from the grey level co-occurrence matrices. The Chan-Vese model is used in a novel way making use of its ability to perform automatic feature selection. This is implemented by automatically readjusting the weights of the feature images to ensure that the contour evolution is driven by those features 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-occurrence matrices for texture extraction and their relationship with sum and difference histograms is reviewed and an algorithm for extracting the Haralick features from a sonar image is described. Finally, in Section 4 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. Let $\mathbf{C_o}(s) : S^1 \to \Omega$ be a closed, simple curve with curve parameter s. The idea behind active contour segmentation methods is to deform the initial curve $\mathbf{C}(s,0) = \mathbf{C}_0(s)$ in time using a suitable, data dependent partial differential equation of the form:

$$\frac{\partial \mathbf{C}(s,t)}{\partial t} = F(s,t)\vec{\mathbf{N}}(s,t) \tag{1}$$

so that the family of curves $\{\mathbf{C}(s,t) : s \in S^1, t \in (0,+\infty)\}$ satisfying (1) converges to the boundary Γ as $t \to +\infty$. Here s, t are the curve and time parameters respectively, $\vec{\mathbf{N}}$ stands for

the inward normal vector of $\mathbf{C}(s,t)$ and F is the speed with which $\mathbf{C}$ 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 $\mathbf{C}(s,t)$ 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 [19] and further developed by several authors over the past decade [27]–[29]. 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 $\mathbf{C}_0$ in (1), that is

$$\{(x,y) : \phi(x,y,0) = 0\} = \mathbf{C}_0 \tag{3}$$

The family of curves $\mathbf{C}(.,t)$, $t > 0$ 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 $\mathbf{C}(.,t)$, as splitting or merging, can be handled naturally and powerful numerical schemes able to approximate the correct viscosity (weak) solution can be employed. [27], [29]

The velocity function F is usually derived through minimising a suitable image dependent energy although it is also possible to synthesise F directly from the image data [10], [30]. 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 [10], [12], [30]. 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 [11], [14], [16]–[18]. A review of region-based methods can be found in [18].

A general framework that can unify many of the region-based approaches in the two region case was presented in [18]. This framework considers energies of the general form

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

where $\Omega_{out}, \Omega_{in}$ denote the outside and inside of the curve $\mathbf{C}$. The kernels $k^{(out)}, k^{(in)}$ play the role of "region descriptors" and are modelled as a combination of features globally attached to the evolving regions $\Omega_{out}, \Omega_{in}$. The main result in [18] is the computation of the speed F that will make a contour obeying equation (1) evolves toward a minimum of the energy $E(\mathbf{C})$ in equation (4).

B. The Chan-Vese model for scalar images.

The approach used in this paper was first introduced in [14] and generalised for vector-valued images in [1]. It can be seen as a restricted form of the Mumford-Shah functional for segmentation first proposed in [26]. Alternatively it can be viewed as a special case of the framework in [18].

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

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

where $u : \bar{\Omega} \to \mathbf{R}$ 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 $\Omega_{i i \in I}$ 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 $(u, \mathbf{C})$ of $E(u, \mathbf{C})$ will be an "optimal" piecewise-smooth approximation u of the initial, possibly noisy, image I while $\mathbf{C}$ will approximate the edges of I.

In the case where I consists of two distinct regions and when restricting the approximations u of I to piecewise constant functions (functions taking only two values, c_1, c_2 in the region inside $\mathbf{C}$ and the region outside $\mathbf{C}$ respectively), the Mumford Shah functional becomes

$$\begin{aligned} E(c_1, c_2, \mathbf{C}) = \mu \text{Length}(\mathbf{C}) \quad &+ \quad \lambda_1 \int_{int(\mathbf{C})} |I(x,y) - c_1|^2 dxdy \\ &+ \quad \lambda_2 \int_{ext(\mathbf{C})} |I(x,y) - c_2|^2 dxdy \end{aligned} \qquad (6)$$

In this case the constants c_1, c_2 are given by

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

This leads to the minimisation problem:

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

for $\mu \geq 0$, $\lambda_1, \lambda_2 > 0$. Now energy (7) is of the form of equation (4) considered in [18]. Following [18] the evolution equation derived from the minimisation of (7) is given by

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

where κ is the local curvature of the curve $\mathbf{C}$.

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

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

and can therefore be extended naturally outside the curve $\mathbf{C}$. As demonstrated in [14], 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 one of them. 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.

The approach used in this paper to extend the model of the previous section to vector-valued images is presented follows [1].

Let I be as before and let $I_1, I_2, \ldots, I_n$ be n feature maps on the domain Ω that have been extracted from I. For example $I_{ii=1,\ldots,n}$ may be the output of a filter bank applied on I. The idea is to evolve a curve $\mathbf{C}$ 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 $\mathbf{C}$ 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 $\mathbf{C}$. This approach was first implemented in [1] where it is also shown that the speed F will evolve $\mathbf{C}$ under equation (1) toward the minimum of the energy

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

As in the one-dimensional case this approach also fits into the framework in [18].

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. Feature Extraction

Side-scan sonar images are often characterised by textured regions corresponding to different types of seabed. Many approaches have been proposed to model texture (filter banks, morphology, run length encoding, Markov random fields, fractals, co-occurrence) [4]–[7], [21], [31]. The appropriate choice of feature sets is a complex and largely unresolved problem. Co-occurrences matrices belong to the statistical representation of textures (fractals, Markov random fields) which view texture as a stationary random stochastic process. A stochastic model is well adapted to sonar imagery due to the amount of noise present in the images. Co-occurrences matrices have been chosen here as they are simple, well-established and fast to calculate. There is also evidence that they are well adapted to sonar imagery [31]–[33]. The main emphasis of this paper is not the selection of the best texture feature set but the demonstration that texture features can be integrated into a curve evolution framework to efficiently segment side-scan sonar images. Other

feature sets could be readily integrated in the existing scheme if they were better adapted to the specific nature of the images to be analysed, be they side-scan or other.

The feature set used here is the well known Haralick set [34]. Only a subset of the 14 Haralick features (*Energy, Contract, Correlation, Entropy, Homogeneity, Cluster Shade and Cluster Prominence*) representing the most commonly chosen ones is used here. They are referenced throughout the text as f_1 to f_7. For computational reasons, these features are not extracted for each pixel of the original image I but on a sub-sampled lattice determined by horizontal and vertical sub-sampling steps (s_x, s_y). For each point of the lattice, the features are extracted using an estimation window of size $w_x \times w_y$. The distance between the couples of pixels for which the co-occurrence is estimated is denoted d and the orientation of the couples is denoted θ. More details on the extraction of Haralick features can be found in [34]. A fast implementation of an estimate of the Haralick set using sum and differences of histograms introduced by Unser [35] is used here. Our implementation is based on non-optimised C code running on a Pentium 1.8GHz, 512MB RAM computer with Linux operating system. In figure VI an example sonar image and the extracted features $f_1 - f_7$ with $s_x = s_y = 2$ and $\theta = -\pi/2$ is shown. The extraction time for this image (400×360 pixels) was $45s$. It is worth noting however that a local equalisation step used to make the textures more stationary, performed as a pre-processing stage, takes up about half of this time($\approx 21s$).

IV. Implementation and experimental results.

A. Implementation of the segmentation algorithm.

Let I be a two-region grey-scale 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 partial differential equation 12 is implemented numerically by making use of a level set method as explained in Section 2:

$$\frac{\partial \mathbf{C}}{\partial t} = \{\frac{1}{n}\sum_{i=1}^{n}\{\lambda_i^{(in)}[I_i - m_i^{(in)}]^2 - \lambda_i^{(out)}[I_i - m_i^{(out)}]^2\} + \mu\kappa\}\vec{\mathbf{N}} \tag{12}$$

with initial condition $\mathbf{C} = \mathbf{C}_0$ and where $m_i^{(in)}, m_i^{(out)}$ are the mean values of images I_i inside and outside $\mathbf{C}$. More precisely the PDE

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= (F_0 + F_1)|\nabla \phi| \\ F_0 &= \frac{1}{n}\sum_{i=1}^{n}\{\lambda_i^{(in)}[I_i - m_i^{(in)}]^2 - \lambda_i^{(out)}[I_i - m_i^{(out)}]^2\} \\ F_1 &= \mu\kappa = \mu\nabla \cdot \left(\frac{\nabla \phi}{|\nabla \phi|}\right) \end{aligned} \tag{13}$$

with initial condition $\phi(x, y, 0) = sgnd(x, y, \mathbf{C}_0)$ is discretized. Here $sgnd(., ., \mathbf{C}_0)$ denotes the signed distance function from the initial curve $\mathbf{C}_0$ and is defined by

$$sgnd(x, y, \mathbf{C}_0) = \begin{cases} -dist(x, y, \mathbf{C}_0) \text{ if } (x, y) \text{ lies inside } \mathbf{C}_0 \\ dist(x, y, \mathbf{C}_0) \text{ if } (x, y) \text{ lies outside } \mathbf{C}_0 \end{cases} \tag{14}$$

where $dist(x, y, \mathbf{C})$ denotes the Euclidean distance between a point of coordinates (x, y) and a curve $\mathbf{C}$.

We have used 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 (14). We have found that in some 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.

The stopping criterion for the algorithm is formulated in terms of the stationarity of the zero level set of the solution ϕ: If the zero level set of ϕ has remained unchanged over a given number of iterations then the algorithm is declared to have converged. In our implementation, the algorithms is said to have converged after 50 stationary iterations of the zero level set of ϕ. Convergence of the scheme can also be monitored by calculating the energy (11) and determining whether it is decreasing over time but has not been used here.

It is important to note that convergence to the global minimum of the energy functional (11) associated with the speed F_1 cannot however be guaranteed. First the proposed method is a 'gradient descent' technique and will only reach the closest local minimum. Second, such an

energy has in general more than one local minimum. This makes the algorithm susceptible to initialisation. Two initialisation schemes have been used here: initialisation with a sequence of uniformly distributed circles over the entire image and initialisation with a single central circle (see figures 10. Generally, uniformly distributed circles give better segmentation results and convergence is faster than when initialising with a single curve. However, for certain types of images where the object or region in the central part of the image is small relative to the background region, the final segmentation is better when initialising with a single curve that intersects with the central region. Generally, it was found that the algorithm was not overly sensitive to initialisation.

The choice of features and the setting of the coefficients $\lambda_i^{(in)}, \lambda_i^{(out)}$ is another issue that needs to be addressed. The parameters $\lambda_i^{(in)}, \lambda_i^{(out)}$ determine the degree of contribution of image I_i in the final result. In [1] the λ parameters are used to filter high frequency noise from different channels. In the case where the I_is are the output of feature extraction (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. Here, as in [25], the automatic feature selection is based on maximising the difference between the means of a given feature inside and outside the curve. In [25] the number of features to be selected is fixed *a priori* and the coefficients $\lambda_i^{(in)}, \lambda_i^{(out)}$ are set manually. In our implementation the coefficients $\lambda_i^{(in)}, \lambda_i^{(out)}$ are set automatically and the number of selected features can be indirectly determined via a threshold $0 < \alpha \leq 1$ (in our implementation, α is set at 0.5). The parameters $\lambda_i^{(in)}, \lambda_i^{(out)}$ are initially set to be equal to 1 for the first 300 iterations. They are subsequently reset at regularly following equation:

$$\lambda_i^{(in)} = \lambda_i^{(out)} = \frac{|m_i^{(in)} - m_i^{(out)}|}{M} \tag{15}$$

where

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

Setting the coefficients in this manner ensures that the features with maximum discriminatory capacity drive the curve evolution. In addition the number of features contributing in the algorithm can be controlled using the criterion $\lambda_i \geq \alpha$. As examples in the next section illustrate this type of feature selection can greatly improve the segmentation result.

The parameter μ controls the smoothness in the contour and thus the sharpness of the boundaries of the segmented regions. The setting of μ is somehow empirical as it weights the contri-

bution of the smoothness term against the contribution of the data driven term, the latter being difficult to bound and depends on the image size. However the results do not seem to be very sensitive to the value of μ and it only needs to be adjusted within the correct operational range for a given type of images.

B. Experimental results

In the last section of this paper we present numerical results of the proposed method on synthetic and real sonar images. These were obtained using a Pentium 1.8GHz processor with 512MB of RAM. The code is written in non-optimised C++ compiled with g++.

One of the problems when segmenting sonar images is the noise or other artifacts 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.

The choice of the window size used to extract the texture features is ultimately driven by the scale of the textures to be segmented. Large window sizes will provide good texture features but blur the edges between texture regions while smaller windows will preserve the edges at the expense of feature stability. There is also a computational cost to large windows. A compromise must be found between those parameters. A multi-resolution implementation is always possible to obtain a fix analysis window size irrespective of the scale of the texture but this has a high implementation cost and the scale at which the analysis has to be performed needs to be determine. For those reasons, we have used a fixed window size representing a good compromise for the textures under analysis in the underwater images of our database.

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 $w_x = w_y = 21$. Finally 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-occurrence 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 (15) at iterations $100, 200, \ldots, 500$ but where we have also set $\lambda_i^{(in)} = \lambda_i^{(out)} = 0$ for those indices i where $\frac{|m_i^{(in)} - m_i^{(out)}|}{M} < \alpha (\alpha = 0.5)$, effectively removing 'bad' features from the curve evolution process.

For sonar images, the value $\mu = 5 \cdot 10^{-4} \cdot (255)^2$, where images were of size $256 \times 256 pixels$, 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. The setting of μ follows suggestions form [14].

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. The image in 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. This is an example where the present method is used to segment a homogeneous region within an image consisting of more than one regions. In this case a segmentation in three regions could be achieved by re-applying the method on both regions separately. In general however this is not a robust approach for multi-region segmentation, at least when the curve evolution is used to minimise the particular energy functional used here. Initial experiments on real sonar data confirm this.

Figures 5, 7, 8 contain segmentation results on real raw sonar data. As can be seen the data is noisy and contains various artifacts. 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 (VI) 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 $0.2 sec.$

Figure 9 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 10 contains examples of images that are harder to segment using the proposed algo-

rithm. The image in the top row can be successfully 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. In all, for such "harder images" the difficulty seems to come from the lack of a good initialisation rather than the choice of features. Problems of this type can be addressed by using a multi-resolution 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.

V. Conclusions and future work

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 [1].

The results obtained strongly suggest that the Haralick set of features is suitable for texture discrimination in side-scan sonar data. In particular, using the automatic feature selection step in the algorithm, the features contrast and homogeneity turn out to be consistently the most significant. The algorithm is shown to have good regularisation properties and to be robust to noise and other artifacts resulting from the data acquisition process.

As a result of the non-convexity of the energy functional one of the limitations of the algorithm is its sensitivity to initialisation. One way to address this problem is by pre-segmenting the image at a coarser scale using a pixel-based segmentation method. A multi-resolution approach would also be beneficial as the value of parameters for the co-occurrence matrices depends on the scale of the textures in the sonar image. This may be achieved by enlarging the feature set using different parameters corresponding to different scales and using the feature selection step as a

means of identifying the most suitable values.

The results obtained so far and presented in this paper suggest that curve evolution is a useful tool in sonar analysis.

The proposed method can be readily extended to other contexts. Different types of features may be used. In particular it is possible to "fuse" different feature sets to allow information not captured by the Haralick features to be considered. By taking a slightly different viewpoint the method may also be used for sensor fusion (for example bathymetry or video). Finally it is possible to extend the method to segment sonar images to n classes, $n > 2$, by considering coupled evolutions of several contours. The main difficulty in achieving this is the representation of different regions through the evolving curves. A number of different approaches exist, such as [36]–[39]. All of these methods can, in principle, be used in the context of this paper with suitable adjustments. This is a direction of research that we will pursue in the future.

VI. Acknowledgements

We would like to thank the NATO Undersea Research Centre for providing the side-scan sonar data used in this paper. We would also like to thank the editor and reviewers of the paper for their helpful comments which have improved the exposition considerably.

References

[1] T. Chan, B. Sanberg, and L.Vese, "Active contours without edges for vector-valued images," *Journal of Visual Communication and Image Representation*, vol. 11, pp. 130–141, 2000.

[2] L. Bjorno, J. Papadakis, P. Papadakis, J. Sageloli, J. Sessarego, S. Sun, and M. Taroudakis, "Identification of seabed data from acoustic reflections: theory and experiment," *Acta Acoustica*, vol. 2, pp. 359–374, 1994.

[3] D. Sternlicht and C. de Moustier, "Time-dependent seafloor acoustic backscatter (10-100khz)," *Journal of the Acoustical Society of America*, vol. 114, pp. 2709–2725l, 2003.

[4] B. Zerr, E. Maillard, and D. Gueriot, "Sea-floor classification by neural hybrid system," *Proceedings of the IEEE Oceans Conference*, pp. 239–243, 1994.

[5] D. R. Carmichael, L. M. Linnet, S. J. Clarke, and B. R. Calder, "Seabed classification through multi-fractal analysis," *IEE Proceedings Radar, Sonar and Navigation*, vol. 143, no. 3, pp. 140–148, 1996.

[6] M. Mignotte, C. Collet, P. Perez, and P. Bouthemy, "Sonar image segmentation using an unsupervised hierarchical mrf model," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1216–1231, 2000.

[7] S. Reed, Y. Petillot, and J. Bell, "An automatic approach to the detection and extraction of mine features in side-scan sonar," *IEEE Journal of Oceanic Engineering*, vol. 28, no. 1, pp. 90–105, 2003.

[8] J. M. Bell, M. J. Chantler, and T. Wittig, "Side-scan sonar: a directional filter or seabed texture?" *IEE Proceedings Radar, Sonar and Navigation*, vol. 146, no. 1, pp. 65–72, 1999.

[9] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes:active contour models," *International Journal of Computer Vision*, vol. 1, pp. 321–331, 1988.

[10] R. Malladi, J. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: A level set approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 158–175, 1995.

[11] S. Zhu, T. Lee, and A. Yuille, "Region competition: unifying snakes, region growing and bayes/mdl for multi-band image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, no. 9, pp. 884–900, 1996.

[12] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International Journal of Computer Vision*, vol. 22, no. 1, pp. 61–79, 1997.

[13] A. Yezzi, A. Tsai, and A. Willsky, "A statistical approach to snakes for bimodal and trimodal imagery," in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2, 1999.

[14] T. Chan and L. Vese, "Active contours without edges," *IEEE Transactions on Image processing*, vol. 10, no. 2, pp. 266–277, 2001.

[15] ——, "A level set algorithm for minimising the mumford- shah functional in image processing," in *Computer Society Proceedings of the 1st IEEE Workshop on Variational and Level Set Methods in Computer Vision*, 2001, pp. 161–168.

[16] J. Kim, J. Fisher, A. Yezzi, M. Cetin, and A. Willsky, "Non-parametric methods for image segmentation using information theory and curve evolution," in *International Conference on Image Processing*, 2002.

[17] N. Paragios and R. Deriche, "Geodesic active regions: A new framework to deal with frame partition problems in computer vision," *Journal of Visual Communication and Image Representation*, vol. 13, pp. 249–268, 2002.

[18] S. Jehan-Besson, M. Barlaud, and G. Aubert, "Deformable regions driven by an eulerian accurate minimisation method for image and video segmentation," *International Journal of Computer Vision*, vol. 53, no. 1, pp. 45–70, 2003.

[19] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed:algorithms based on hamilton-jacobi formulation," *Journal of Computational Physics*, vol. 79, pp. 12–49, 1988.

[20] C. Sagiv, N. A. Sochen, and Y. Y. Zeevi, "Integrated active contours for texture segmentation," *IEEE Transactions on Image Processing*, vol. 1, no. 1, pp. 1–17, 2004.

[21] S. Liapis, E. Sifakis, and G. Tziritas, "Colour and texture segmentation using wavelet frame analysis, deterministic relaxation, and fast marching algorithms," *Journal of Visual Communication and Image Representation*, 2003.

[22] N. Paragios and R. Deriche, "Geodesic active regions and level set methods for supervised texture segmentation," *International Journal of Computer Vision*, vol. 46, no. 3, pp. 223–247, 2002.

[23] J.-F. Aujol, G. Aubert, and L. Blanc-Féraud, "Supervised classification for textured images," INRIA Report, Tech. Rep. 4640, 2002.

[24] M. Rousson, T. Brox, and R. Deriche, "Active unsupervised texture segmentation on a diffusion based feature space," *International Conference on Computer Vision and Pattern Recognition*, vol. 2, pp. 699–706, 2003.

[25] B. Sandberg, T. Chan, and L. Vese, "A level-set and gabor-based active contour algorithm for segmenting textured images," UCLA CAM Report 02-39, Tech. Rep., 2002.

[26] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communication on Pure Applied Mathematics*, pp. 577–685, 1989.

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

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

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

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

[31] I. Karoui, J.-M. Boucher, R. Fablet, and J.-M. Agustin, "Multi-scale segmentation of texture sonar images using occurrence statistics," in *International Conference on Image Processing*, 2004.

[32] N. Pican, E. Trucco, M. Ross, D. M. Lane, Y. Petillot, and I. T. Ruiz, "Texture analysis for seabed classification: Co-occurrence matrices versus self-organising maps," in *Proceeding of the IEEE Oceans Conference*, Nice, France, Sept. 1998.

[33] J. M. Preston, A. C. Christney, S. F. Bloomer, and I. L. Beaudet, "Seabed classification of multi-beam sonar images," in *Proceedings of the IEEE Oceans Conference*, 2001.

[34] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man and Cybernetics*, pp. 610–621, 1973.

[35] M. Unser, "Sum and difference histograms for texture classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, no. 7, pp. 717–727, 1989.

[36] L.Vese and T. Chan, "A multiple level set framework for image segmentation using the mumford and shah model," *International Journal of Computer Vision*, vol. 50, no. 3, pp. 271–293, 2002.

[37] J.-F. Aujol, G. Aubert, and L. Blanc-Féraud, "Wavelet-based level set evolution for classification of textured images," *IEEE Transactions on Image processing*, vol. 12, no. 12, pp. 1634–1641, 2003.

[38] I. Ayed, A. Mitiche, and Z. Belhadj, "Multi-region level set partitioning of synthetic aperture radar images," *IEEE Transactions on Image processing*, vol. 27, no. 5, pp. 793–800, 2005.

[39] C. Vazquez, A. Mitiche, and I. B. Ayed, "Image segmentation as regularized clustering: A fully global curve evolution method," in *Proceedings of the IEEE International Conference on Image processing*, Oct 2001, pp. 3467–3470.

[40] J. M. Bell and L. M. Linnett, "Simulation and analysis of synthetic side-scan sonar images," *IEE Proceedings Radar, Sonar and Navigation*, vol. 144, no. 4, pp. 219–226, 1997.

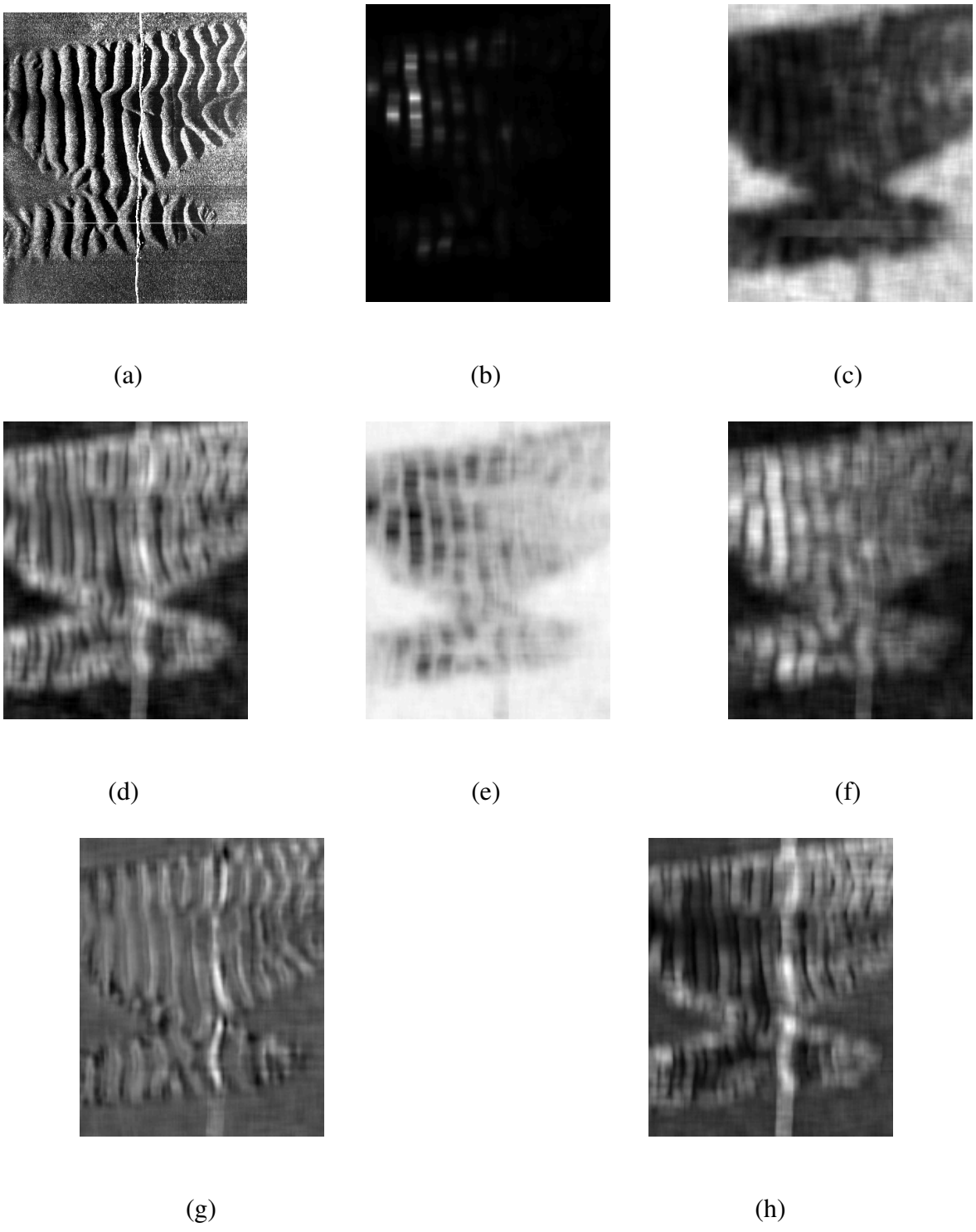


Fig. 1. From left to right: example of sonar image (a) and extracted Haralick features [energy (b), contrast (c), correlation (d), entropy (e), homogeneity (f), cluster shade (g) and cluster prominence (h)]. The figure clearly shows that some features are better suited than other for the segmentation of this example image. Therefore automatic feature selection will be important.

(a) (b)

Fig. 2. Simulated image (a) and segmentation result (b). The simulated image was realised using a sonar simulator developed in [40]. It is based on ray-tracing and provides realistic ground-truthed sonar images enabling the validation of the method proposed.

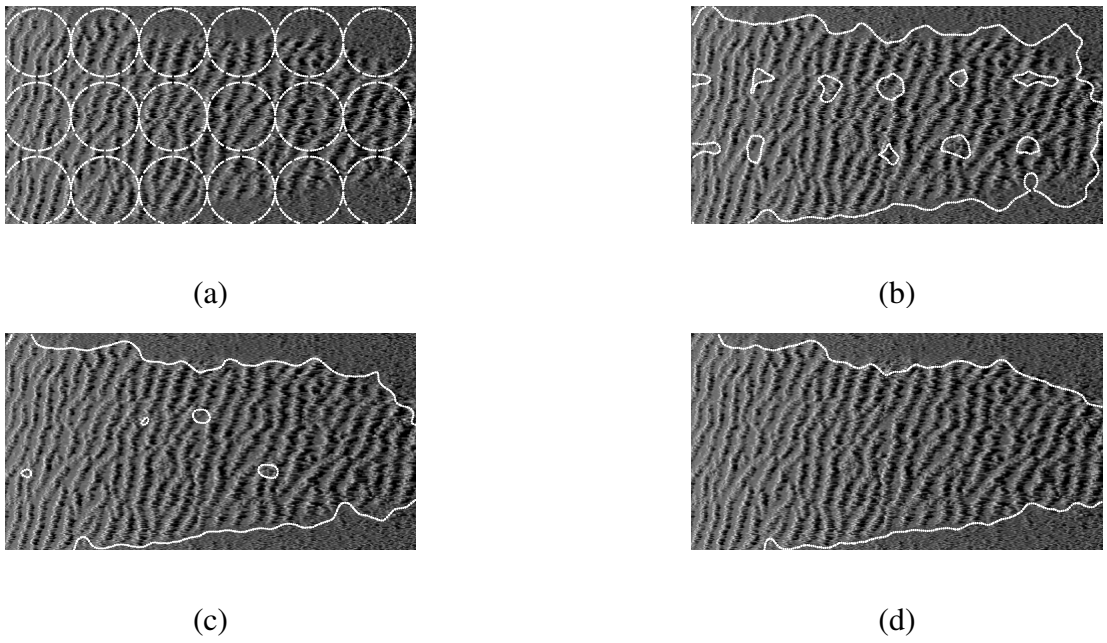


Fig. 3. Curve evolution on simulated image at iterations 0 (a), 20 (b), 40 (c) and 100 (d). Image (a) corresponds to the original image. Image (b) to the initial curve. The other correspond to the zero level set of ϕ at different stages of the process. The image here is composed of two distinct textures, one is sand ripples and the other flat seabed. The final curve very precisely follows the boundary between the two regions without being perturbed by the vertical artifact which locally interferes with the feature extraction. This example demonstrates the global region-based nature of the technique

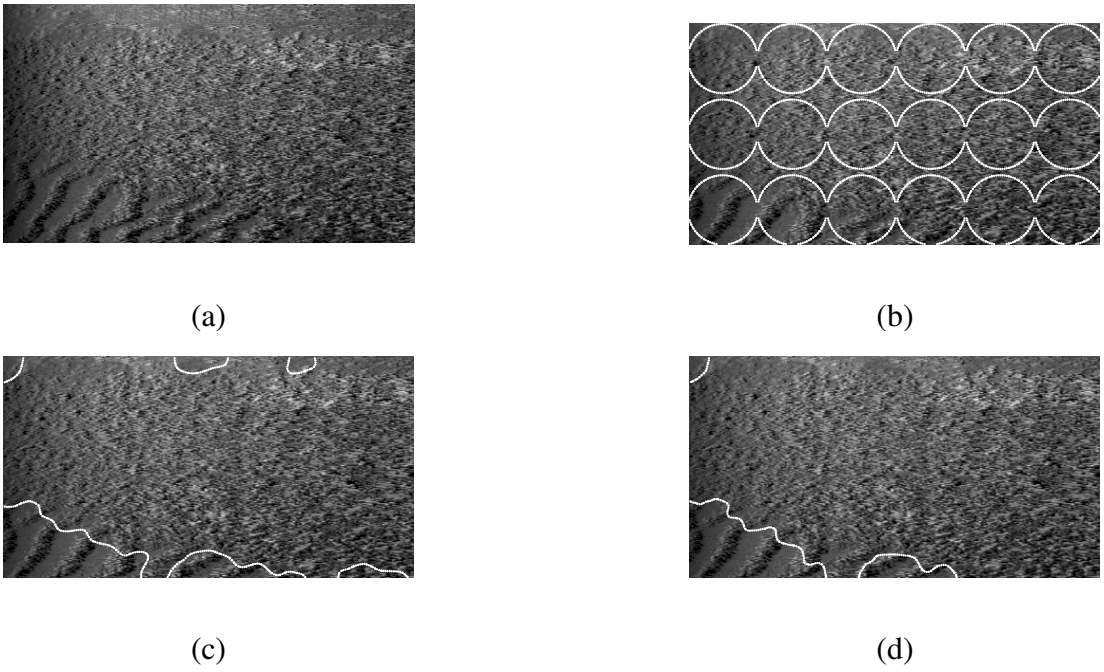
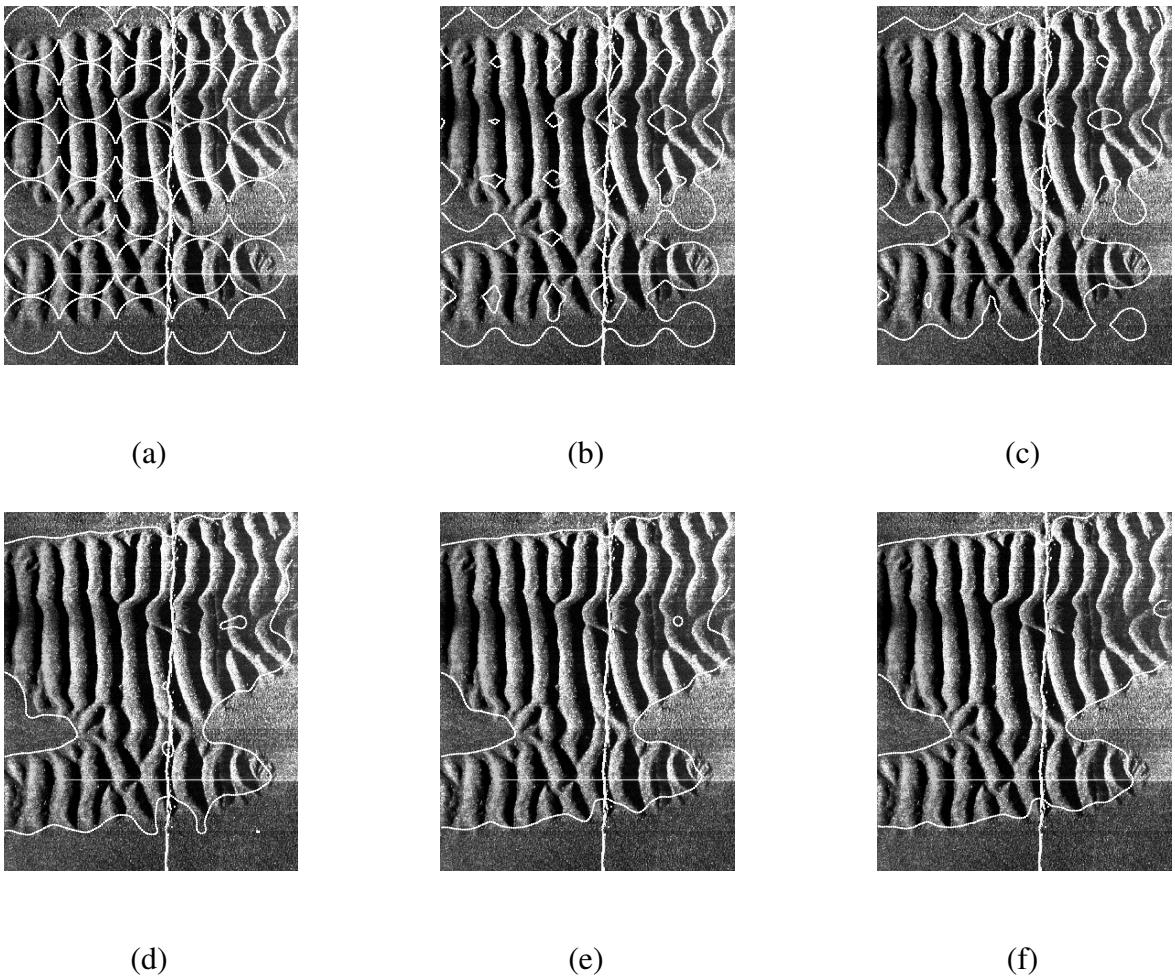


Fig. 4. Curve evolution on simulated image (a) containing three types of seabed at iterations 0 (b), 200 (c) and 1000 (d). Image (a) corresponds to the initial curve. The other correspond to the zero level set of ϕ at different stages of the process. The image here is composed of three distinct textures. The algorithm is currently limited to two classes and amalgamates two classes together. Note that they are perceptively the closest as well.



(a) (b) (c)

(d) (e) (f)

Fig. 5. Contour evolution at $0(a), 10(b), 20(c), 40(d), 60(e)$ and $200(f)$ iterations. Image (a) corresponds to the initial curve. The other correspond to the zero level set of ϕ at different stages of the process. The image here is composed of two distinct textures, one is sand ripples and the other flat seabed. The final curve very precisely follows the boundary between the two regions.

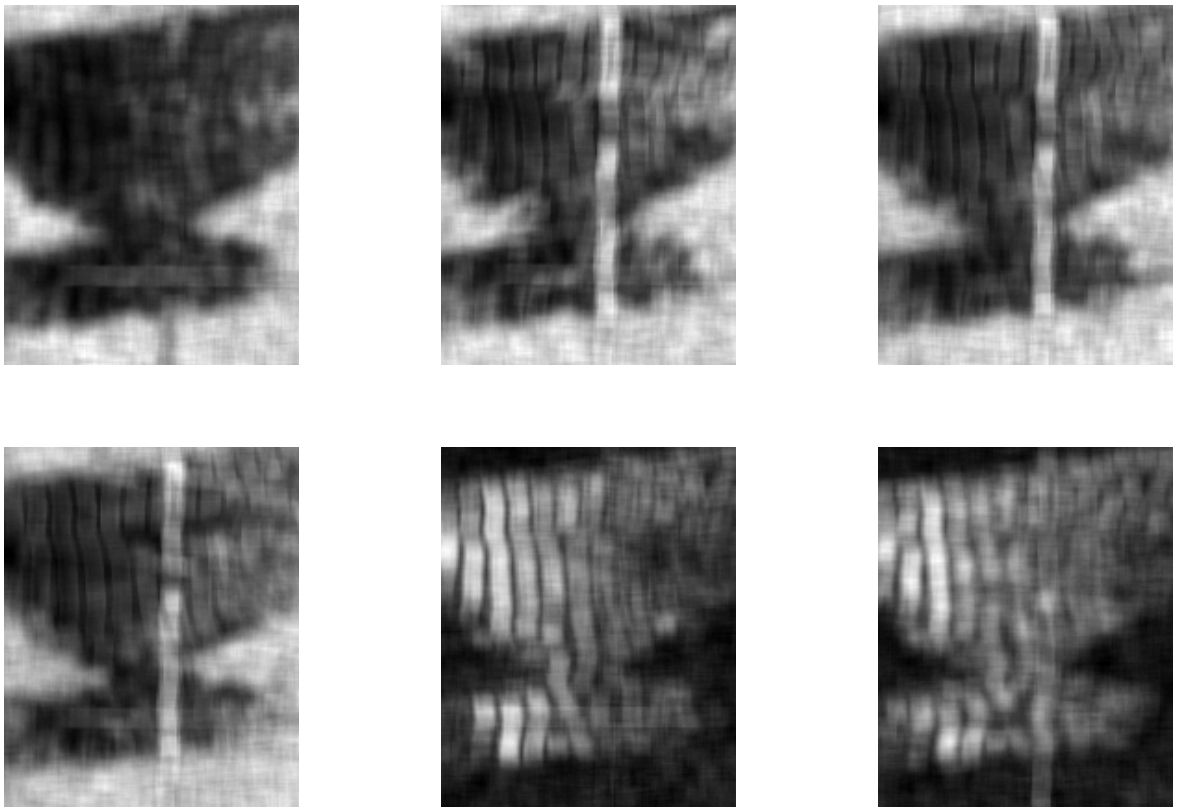


Fig. 6. The six most discriminant features selected for segmentation of the image in figure 5. This clearly shows that the feature selection is effective at identifying the most discriminant features in the feature set.

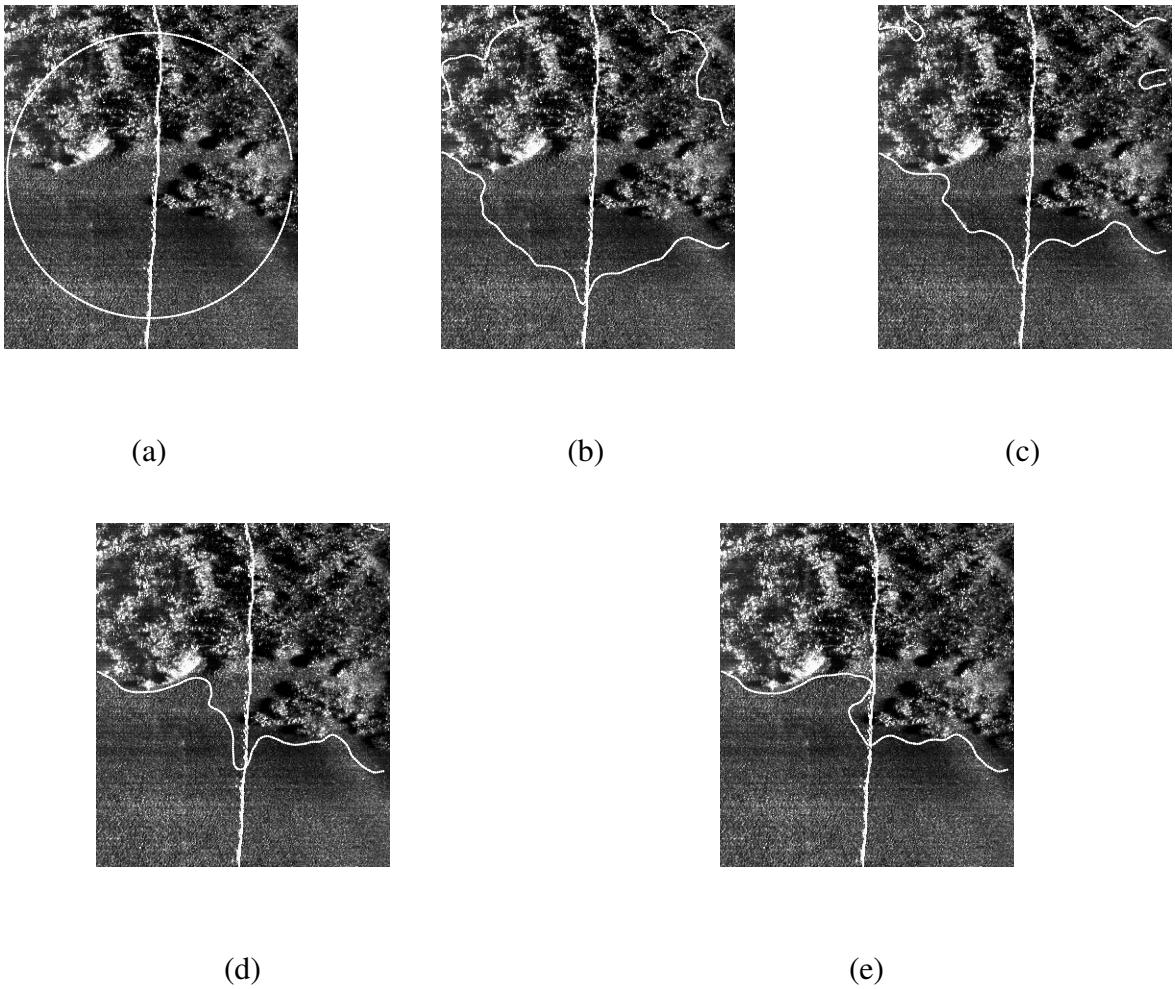
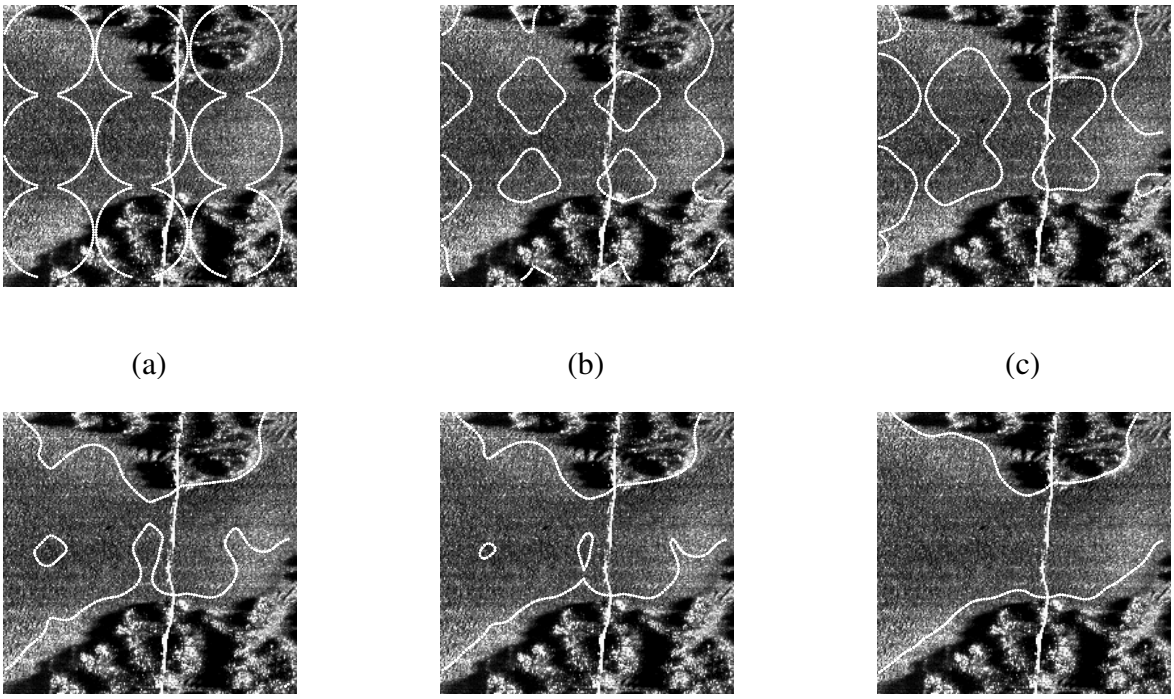


Fig. 7. Contour evolution at $0(a), 80(b), 140(c), 240(d)$ and $460(e)$ iterations. Image (a) corresponds to the initial curve. The other correspond to the zero level set of ϕ at different stages of the process. The initialisation in this case is a single large circle. Convergence is still demonstrated showing that the proposed algorithm is not very sensitive to initialisation.



(a) (b) (c)

(d) (e) (f)

Fig. 8. Contour evolution at $0(a), 20(b), 50(c), 70(d), 90(e)$ and $120(f)$ iterations. Image (a) corresponds to the initial curve. The other correspond to the zero level set of ϕ at different stages of the process. The final curve very precisely follows the boundary between the two regions.

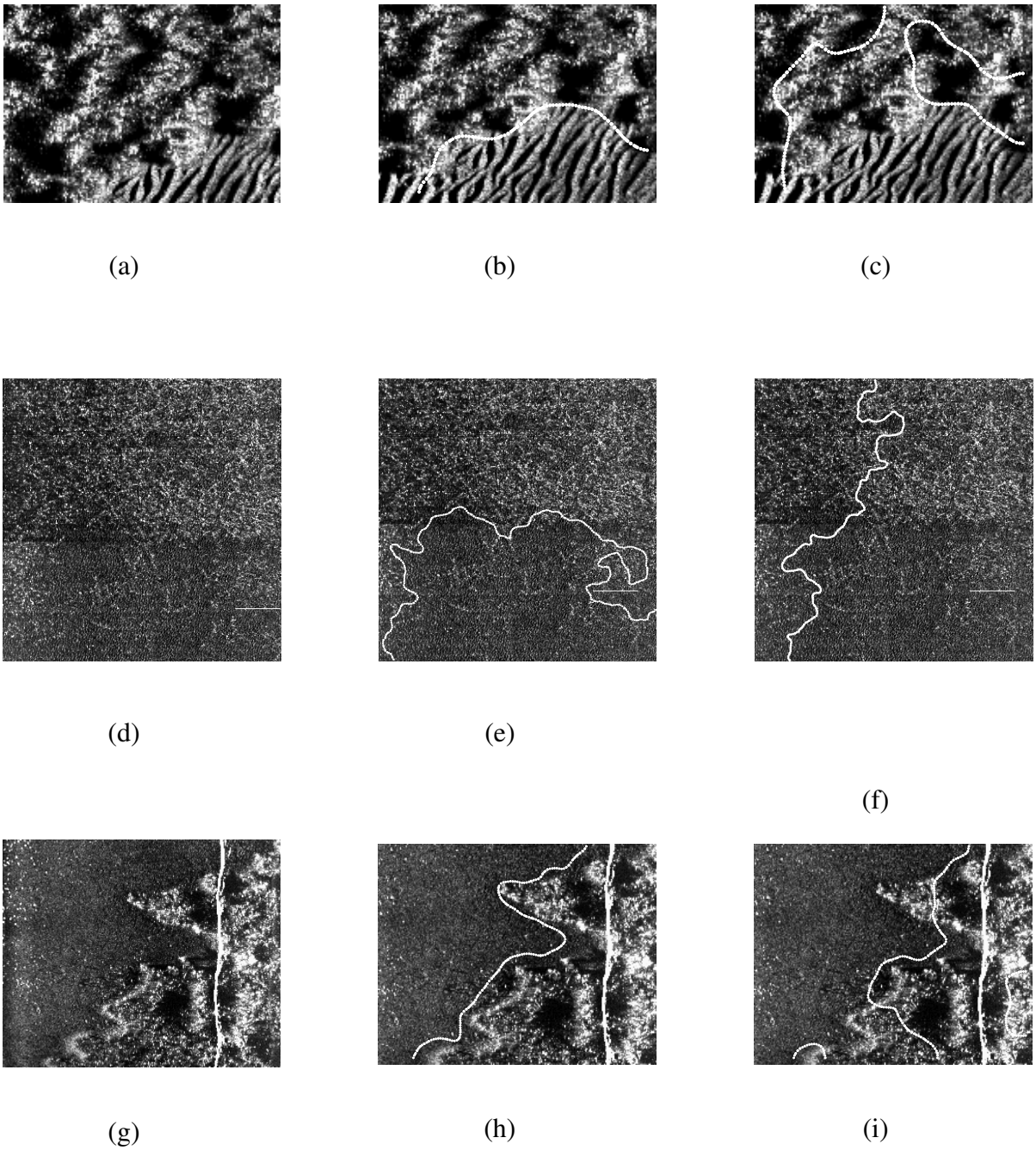


Fig. 9. Original images (a,d,g) and corresponding segmentation results with (b,e,h) and without (c,f,i) feature selection. The results show the importance of feature selection for convergence. Non discriminant features will create extra local minima in the energy functional.

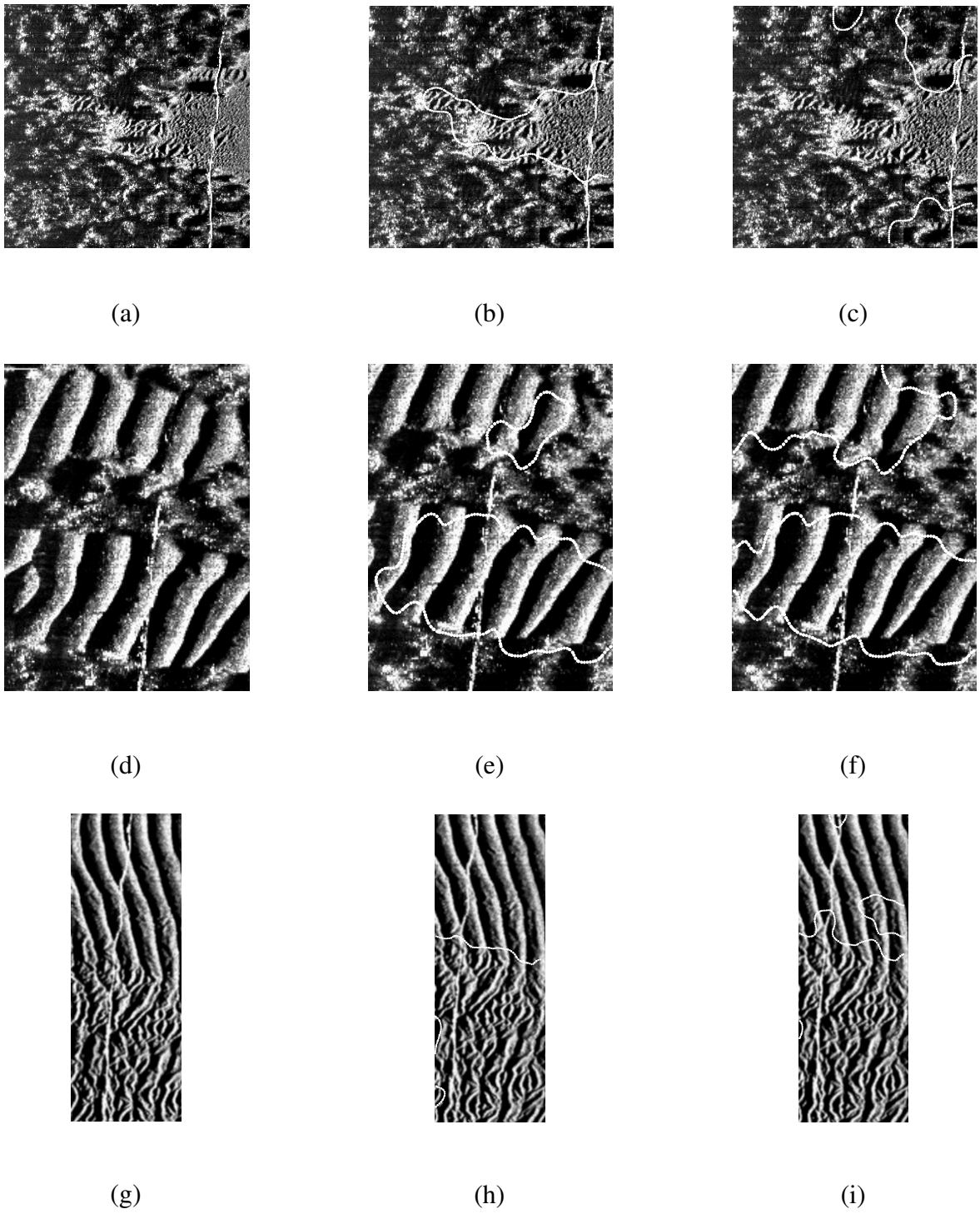


Fig. 10. Original images (a,d,g) and segmentation results when initialising with one circle (b,e,h) and an array circles (c,f,i). This figure demonstrates the importance of initialisation in specific cases. In the first image, initialisation with a big central circle outperforms the initialisation with an array of small circles while the situation is inverted for the last two images. There is no rule as yet to decide on a better initialisation scheme. Prior segmentation using more rudimentary techniques (k-means) could be used to initialised the curves and improve convergence.