Sidescan sonar segmentation using active contours and level set methods

Maria Lianantonakis
School of Engineering and Physical Sciences
Heriot Watt University
Edinburgh, Scotland, UK
Email: M.Lianantonakis@hw.ac.uk

Yvan R. Petillot
School of Engineering and Physical Sciences
Heriot Watt University
Edinburgh, Scotland, UK
Email: 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 [3] 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 [16]. 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 parameters and contour initialisation are analysed.**

I. Introduction

Sidescan sonar image analysis is used in a number of underwater 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 [1], [17] and unsupervised classification and segmentation (typically in shadow/non-shadow or echo/shadow/sea-bottom reverberation areas) based on Bayesian clustering methods using grey levels or texture features [9], [13]. These have produced useful results for seabed classification and segmentation but are generally sensitive to training, feature selection and parameter estimation.

This paper concentrates on a different type of segmentation task namely that of segmenting a sidescan sonar image in different types of seabed in an unsupervised manner. This task is viewed as an unsupervised texture segmentation problem. An unsupervised binary segmentation algoritm for texture-based seabed discrimination is presented. It relies on geometric active contour models and level set methods. In addition a novel, automatic feature selection step is integrated within the algorithm.

In the first stage of the segmentation algorithm a set of texture features from the sidescan image is extracted. The texture descriptors used here are the Haralick features based on the estimation of co-occurence matrices within a window surrounding each pixel in a lattice in the sonar image. This feature set is well known to be a good descriptor of the textures present in sonar data. To improve computational efficiency the feature extraction algorithm implemented here uses the sum and difference histograms approximation method introduced by Unser [16].

The second phase of the segmentation process relies on the active contour model for vector-valued images introduced by Chan and Vese in [3], [4] which uses the mean intensity as a region homogeneity measure. The model is robust to noise and has 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. To implement the contour evolution a level set approach is used leading to a Hamilton-Jacobi equation. This is numerically approximated through a first order monotone scheme.

One of the problems when working with texture features is feature selection. Identifying a set of features able to discriminate between the textures present in all images of a data set is not always possible. One of the novel features of our implementation is an automatic feature selection step which has been integrated in the evolution scheme. This ensures that for a given image the features with maximum discriminatory capacity drive the contour evolution.

II. Background

Suppose that I is an image defined on a domain $\Omega \subseteq \mathbf{R}^2$ consisting 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 [12] and further developed by several authors over the past decade [10], [11], [15]. This is achieved by representing the curve C implicitly as the zero level set of a surface $\{(x, y, z) : z = \phi(x, y)\}$. The level set formulation of equation (1) is then 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. [10], [15]

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 [8], [14]. 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 [2], [8], [14]. 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. A review of region-based methods can be found in [7].

III. The Chan-Vese active contour model for binary segmentation.

The approach used in this paper was first introduced in [4] and generalised for vector-valued images in [3]. It can be seen as a restricted form of the Mumford-Shah functional for segmentation in which the original image I is approximated by a binary image.

The Chan-Vese model is based on the minimisation of the following energy functional:

$$\begin{aligned} \inf_C E(C) &= \inf_C \{\lambda_1 \int_{int(C)} |I(x,y) - m_{in}(C)|^2 dxdy \\ &+ \lambda_2 \int_{ext(C)} |I(x,y) - m_{out}(C)|^2 dxdy \\ &+ \mu \text{Length}(C)\} \end{aligned} \tag{4}$$

where

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

and $\mu \geq 0$, $\lambda_1, \lambda_2 > 0$. Following [7] the evolution equation derived from the minimisation of (4) is given by

$$\frac{\partial C}{\partial t} = \{\lambda_1 (I - m_{in}(C))^2 - \lambda_2 (I - m_{out}(C))^2 + \mu\kappa\}\vec{\mathbf{N}} \tag{5}$$

where κ is the local curvature of the curve C.

Equation (5) can be implemented by using a level set method as explained above. Note that in this case the speed function F is given by

$$F = \lambda_1 (I - m_{in}(C))^2 - \lambda_2 (I - m_{out}(C))^2 + \mu\kappa \tag{6}$$

and can therefore be extended naturally outside the curve C. As demonstrated in [4], 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 (e.g. the output of a filter bank applied on I) and evolve C over all images under equation (1) where the velocity function F is given as a weighted average of terms over all images:

$$F = \frac{1}{n} \sum_{i=1}^{n} \{\lambda_i^{(in)} [I_i - m_i^{(in)}]^2 - \lambda_i^{(out)} [I_i - m_i^{(out)}]^2\} + \mu\kappa \tag{7}$$

Here $m_i^{(in)}, m_i^{(out)}$ are the mean values of images I_i inside and outside C. This approach was first implemented in [3] where it is shown that the speed F will evolve C under equation (1) towards the minimum of the energy

$$\begin{aligned} E(C) &= \frac{1}{n} \sum_{i=1}^{n} \lambda_i^{(in)} \int_{\Omega_{in}} |I_i(x,y) - m_i^{in}(C)|^2 dxdy \\ &+ \frac{1}{n} \sum_{i=1}^{n} \lambda_i^{(out)} \int_{\Omega_{out}} |I_i(x,y) - m_i^{out}(C)|^2 dxdy \\ &+ \mu \int_C ds \end{aligned} \tag{8}$$

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.

IV. Co-occurence matrices and sum and difference histograms

The feature set of Haralick et al [6] 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. The co-occurence matrix $C_{\vec{d}}(i,j)$ of W with parameters (d_1, d_2) is defined [6] to be the number of pixel pairs $(n,m),(n',m')$ in W that have intesity values i and j respectively. The normalised quantity

$$\hat{p}(i,j) = C_{\vec{d}}(i,j)/(N \times M) \tag{9}$$

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

In [6] 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 [16] 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 [16] Unser proposes an alternative, significantly more efficient way of computing the Haralick Feature Set. His approach relies on approximating the co-occurence matrix $\hat{p}$ of W by the product of the histograms of the sum and difference images W_S, W_D defined as

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

This leads to the following expressions of the features $f_1 \ldots f_7$ in terms of the sum and difference histograms $\hat{p_S}$, $\hat{p_D}$:

$$\begin{aligned}
f_1 &\approx \sum_i \hat{p_S}(i)^2 \sum_j \hat{p_D}(j)^2 \\
f_2 &= \sum_j j^2 \hat{p_D}(j) \\
f_3 &= \frac{1}{2}\left\{\sum_i (i-2\mu)^2 \hat{p_S}(i) - \sum_j j^2 \hat{p_D}(j)\right\} \\
f_4 &\approx -\sum_i \hat{p_S}(i)\log\hat{p_S}(i) - \sum_j \hat{p_D}(j)\log\hat{p_D}(j) \\
f_5 &= \sum_j \frac{1}{1+j^2}\hat{p_D}(j) \\
f_6 &= \sum_i (i-2\mu)^3 \hat{p_S}(i) \\
f_7 &= \sum_i (i-2\mu)^4 \hat{p_S}(i)
\end{aligned} \tag{10}$$

The symbol $\approx$ indicates that the left hand side can be approximated by the expression on the right with a quantifiable error. Using the expressions above to calculate the Haralick featurs results in a more efficient computation as double summations are replaced by a single ones and memory requirement reduces by a factor $L/4$. The feature extraction algorithm used here is described below:

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) Set up sequences of windows $W_S(n,m)$, $W_D(n,m)$ centred at (n,m) in I_S and I_D respectively.
5) For each one of the pixels (n,m) use the histograms of W_S, W_D to compute features $(f_1,\ldots,f_7)$. Realise features as a vector $(I_1,\ldots,I_7)$ of images extracted from the original image I.

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

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:

$$\frac{\partial C}{\partial t} = \{\frac{1}{n}\sum_{i=1}^{n}\{\lambda_i^{(in)}[I_i - m_i^{(in)}]^2 - \lambda_i^{(out)}[I_i - m_i^{(out)}]^2\} + \mu\kappa\}\vec{\mathbf{N}} \tag{11}$$

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

$$\begin{aligned} \frac{\partial \phi}{\partial t} &= (F_0 + F_1)|\nabla\phi| \qquad (12)\\ F_0 &= \frac{1}{n}\sum_{i=1}^{n}\{\lambda_i^{(in)}[I_i - m_i^{(in)}]^2 - \lambda_i^{(out)}[I_i - m_i^{(out)}]^2\} \\ F_1 &= \mu\kappa = \mu\nabla\cdot\left(\frac{\nabla\phi}{|\nabla\phi|}\right) \end{aligned}$$

with initial condition $\phi(x, y, 0) = sgnd(x, y, C_0)$, the signed distance function from the initial curve C_0.

A first order monotone scheme was used to approximate the term $F_0|\nabla\phi|$ and a first order central difference approximation was used for 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.

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. It was found that in most cases 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, for example images where the object or region in the central part of the image is small relative to the background region, segmentation is 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 [3] 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:

$$\lambda_i^{(in)} = \lambda_i^{(out)} = \frac{|m_i^{(in)} - m_i^{(out)}|}{M} \tag{13}$$

where

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

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

B. Experimental results.

In the results presented 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 (13) 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.

Segmentation results on simulated and real sonar images are presented in figures 1, 2, 3, 4 and 5.

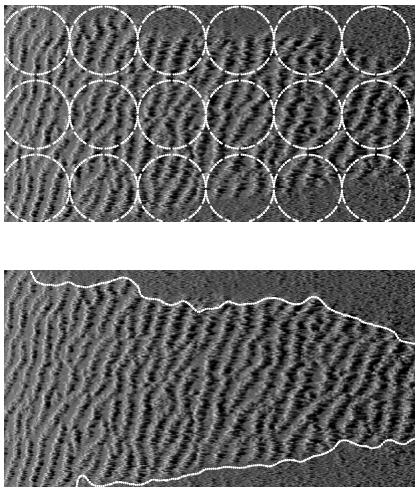

Fig. 1. Initial curve and segmentation result on simulated image containing two types of seabed.

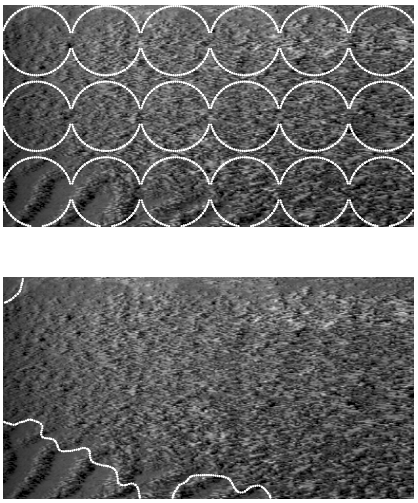

Fig. 2. Initial curve and segmentation result on simulated image containing three types of seabed. The two most similar textures are represented by the top region.

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 [3]. 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 [5]. Overall the results obtained in this paper suggest that curve evolution is a valid tool in sonar analysis.

References

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

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

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

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

[5] 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.

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

[7] 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.

[8] 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.

[9] 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.

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

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

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

[13] 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.

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

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

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

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

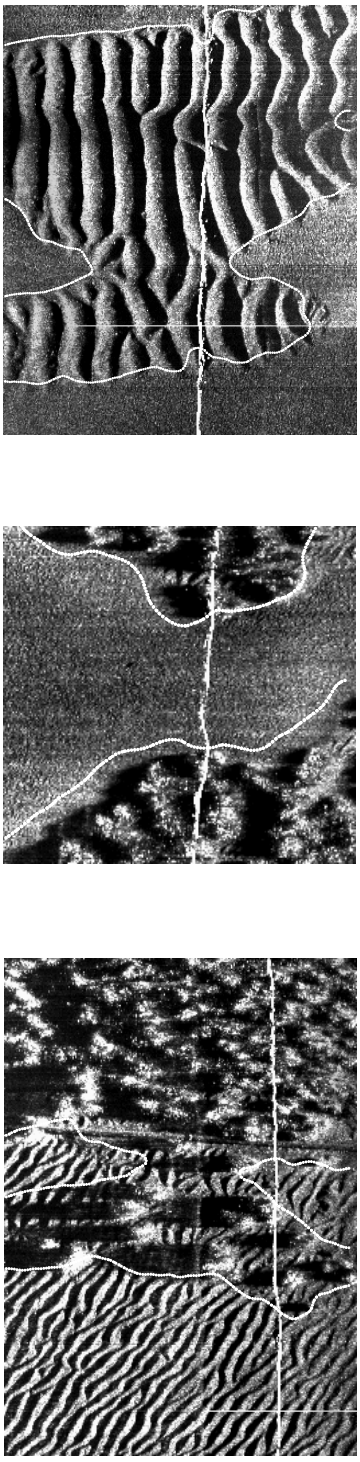

Fig. 3. Segmentation results on real sidescan sonar images.

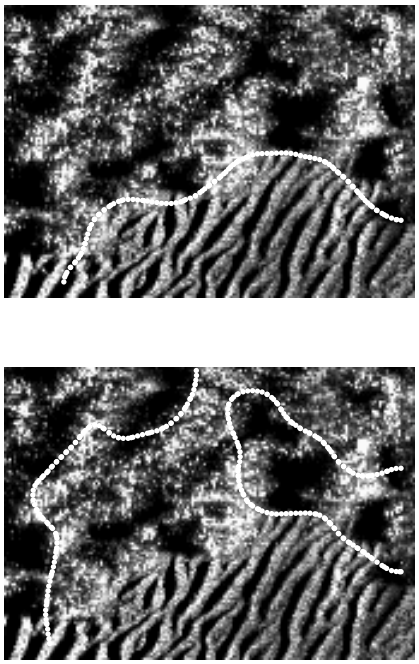

Fig. 4. Segmentation results with and without feature selection.

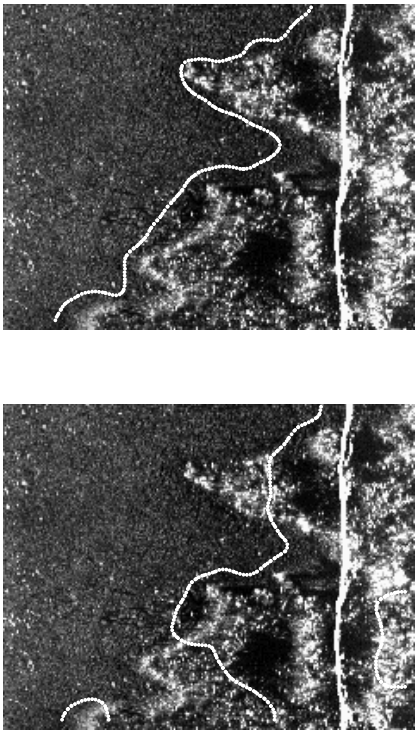

Fig. 5. Segmentation results with and without feature selection.