STATISTICAL SNAKES: APPLICATION TO TRACKING OF BENTHIC CONTOURS

Stefan Rolfes **Maria-João Rendas**

Laboratoire d'Informatique Signaux et Systèmes, de Sophia Antipolis,
CNRS-UNSA,
Sophia-Antipolis, France
{rolfes}{rendas}@i3s.unice.fr

Abstract: This paper presents the formalism of statistical snakes for contour estimation, and applies it to the problem of tracking a natural benthic contour using an AUV equipped with a video camera. To close the control loop that maintains the vehicle on the contour, we need to estimate its location on each frame of the video sequence. We propose a new criterion for contour estimation which is appropriate to the complexity and variability of natural environments. The criterion is based on a non-parametric statistical modeling of the regions adjacent to the tracked boundary, as mixtures of the probability distributions corresponding to the areas on each side of the contour. It is shown on the paper that minimizing the proposed criterion leads to an estimated contour such that the regions on each side have mixture coefficients close to zero and one (meaning that they are "pure" regions). Examples of application of the proposed algorithm to a real underwater image sequence are given.

Keywords: Contour tracking, Vision, Statistical snakes

1. INTRODUCTION

The ability to track benthic contours – that delineate regions of the sea bed occupied by a given type of material or biological species – is not only an efficient observation behaviour in the context of biological studies (analysis of the patchy structure induced in the sea-bed), but also a necessary capability for mapping large scale geometric features of the sea-bed, which provide the most useful and reliable information for robot localisation purposes. This paper addresses tracking of *benthic contours* using *visual information*.

Autonomous contour tracking using vision involves two problems: *(a)* to **estimate** the relative position of the robot with respect to the contour, *(b)* to generate appropriate control actions that **steer** the vehicle such that the contour is kept inside the field of vision of the onboard video camera. This paper focus on problem *(a)*, proposing a fast robust algorithm for *contour tracking accross video frames*.

The solution proposed here improves upon the work presented in [3] where contour estimation involves a complete segmentation of each video frame. The major goal when designing the previous algorithm was to keep the analysis free of *a priori* assumptions about the visual appearance of the observed sea bed regions. The algorithm is grounded on fundamental results from probabilitistic Information Theory, making no assumptions about the characteristics of the probability distributions associated to the distinct regions that may be present in the images. The solution presented here solves the two main drawbacks associated to our previous work: *(1)* its large **computational cost**, by focusing image processing in the vicinity of the tracked contour; *(2)* the **sensitivity with respect to background variations**, by explicitly adapting to the local characteristics of the sea bed. These gains are obtained by resorting the formalism of *deformable contours* (**snakes**), and using the basic measures of Information Theory already involved in the previous solution to define the energy functional that these deformable contours minimize. We call the deformable contours constrained by this energy ***statistical snakes***.

Snakes are curves that evolve in time as a result of the action of a set of active forces, modelled by a partial differential equation (*pde*). In general this equation is obtained from minimization of an energy, even if some authors directly postulate the evolution *pde*. They have been used frequently in the past [9, 8, 4] for segmentation purposes, feature tracking or object recognition.

The snake energy can a combination of *contour* and *region* terms. Contour terms provide an evaluation of the *local* characteristics of the contour, and are usually a superposition of an internal and an external energy. The internal energy is a regularising term, most comonly the superposition of forces modeling elasticity and stiffness. It is balanced by an external energy derived from the image, which drives the snake to regions of high gradient (edges). Contour based snakes rely on the fact that relatively neat and smooth edges can be identified in the image. This is never the case when we track natural features. The problem of clutter that corrupts the feature maps has been addressed, for instance, by restricting the snake to a pre-defined shape space [7], assuming that prior knowledge about the tracked object geometry and dynamic evolutions is available. Again, this is obviously not the case when we consider natural habitats whose small-scale geometry is a priori unknown.

Region terms have been introduced to incorporate measures of the *global* homogeneity of the induced regions, in the context of image segmentationref. Several descriptors of the regions homegeneity have been proposed, in particular statistical indicators such as the mean, variance, and entropy. Several criteria that measure the distance between the histograms of the segmented regions and initial reference histograms have been studied in [2]. The introduction of the region related energy terms efficientlty solves many of the problems associated to contour based snakes. However, application of these approaches to our problem faces several problems: *(1)* they are well defined only for *closed contours,* which is not our case; *(2)* they require the analysis of the *entire image,* imposing high computational load; *(3)* they are suited for segmentation of a *given object* against a *stable background*, which is not the case in our application, since the geometry of the tracked contour and the characteristics of the background bottom can vary as the observation of the sea bottom progresses into new regions.

Our approach can be considered to fall in between pure contour based and fully region based. The idea is to *associate to each point of the contour* terms that measure the contrast, in statistical terms, of the neighboring regions (see Section 4.2). The region-based characteristics are thus *locally* associated to each point of the contour, instead of being computed for the *entire regions* into which the image is divided by the snake. The variation of the characteristics of the neighboring regions with respect to the contour position, required by the variational approach to energy minimisation, are then obtained by using probabilistic mixture models, whose parameters are dependent on the contour position (see Section 4.4). Processing can thus remain confined to the vicinity of the contour, and accomodation to local characterisics is possible. The energy that defines our snakes is thus defined as a *line integral*, its value in each point depending on the characteristics of the image around that point.

The paper is organised as follows. In section 2 we formulate the visual tracking problem and in section 3 we present the set of assumptions that underly our work. Section 4 presents the defintion of the new (statistically motivated) energy functional, and the necessary expressions for its minimisation using a variational approach. Sections 5 to 6 detail each of the three steps that are executed for each new video frame. Finally, section 8 presents results on a sequence of real underwater images that have been acquired in Villefranche-sur-mer using our platform the Phantom 500SP and section 9 concludes the paper.

2. PROBLEM FORMULATION

Let $S\subset\Re^3$ denote a closed region of the sea bed surface, which we assume to be a smooth 2-dimensional manifold for which $A\subset\Re^2$ is a single chart atlas. We assume that A admits a partition $\{A_i\}_{i=1}^{K}$ that identifies the distinct types of sea bottom habitats in region S, such that

$$p\in A_i \Rightarrow S(p) \text{ is a point of type } C_i.$$

By the previous equation, the partition $\{A_i\}_{i=1}^{K}$ induces a corresponding partition $\{S_i\}_{i=0}^{K-1}$ of S. Let ∂S_i be the boundary of the set S_i, and let $\{\partial S_{in}\}_{n=1}^{N_i}$ be the connected components of ∂S_i . We know from general topoly that, for a fixed i, $\{\partial S_{in}\}_{n=1}^{N_i}$ are *non-intersecting simple closed curves*. However, each ∂S_{in} can intersect several ∂S_{jn} , with $i\neq j$, meaning that the class adjacent to class i changes along that boundary.

Assume that the robot is equipped of a video camera solidary with its body, whose axis is aligned with body vertical axis, and denote by $I(t)$ the video frame acquired by the robot at time t:

$$I_t: \quad R\subset Z^2 \mapsto L \\ (\ell_l,\ell_c)\rightarrow I_t(\ell_l,\ell_c),$$

where R is the (discrete) rectangular support of the image, and L is the value of each pixel (in our case, we use monochromatic images, and $L=[0,255]$). We will occasionally need the set R^c which is the smallest rectangle of $\Re^2$ that contains R.

Assuming that the distance of the robot with respect to the sea bed is large compared to the height variations of the sea bed, such that parallel perspective is a valid approximation. Let $W_t\subset\Re^2$ be the indicator function for the footprint of the video camera at time t, such that

$$p\in W_t \Rightarrow I(t) \text{ contains an image of } S(p)\in S .$$

The set W_t depends on the robot position and attitude as well as on the shape of the sea bed surface, i.e., there is a position dependent image formation operator $T_{X(t),S}[\cdot]$ such that

$$(\ell_l,\ell_c)\in R \Rightarrow \exists p\in W_t \ \therefore\ p=T_{X(t),S}\left[(\ell_l,\ell_c)\right] .$$

To indicate this explicitely, we will sometimes use the heavier notation $W_{X(t)}$.

Defiition **Permanent observing trajectory**

Let X^t be an observing trajectory. We will say that it is a permanent observing trajectory iff

$\forall u \le t \;\; \partial S_{in} \cap S(W_{X(u)}) \neq \emptyset$

i.e., if the contour is present in *every* frame acquired along the trajectory.

With these definitions, we can now formulate the contour tracking problem.

Defiition **Contour Observation Problem**

Find the command actions *u(t), t>0,* that lead the robot through a permament observing trajectory of minimal energy.

Solving this problem requires the ability of determining the position of the robot with respect to the contour, which in turn imposes the need to localise the contour in each image frame. In [3] tracking is enforced by closing a classical controller acting on the vehicle's heading rate (of PID type) around an error signal equal to the distance of the outgoing point with respect to the the middle of the upper edge of the image, see Fig. 1. To maintain tracking only very crude information about the contour description is required, as the experiments presented in that reference demonstrate. However, to be able to iteratively locate in a stable manner the outgoing point in the current frame we must estimate the entire contour that is observed inside each frame.

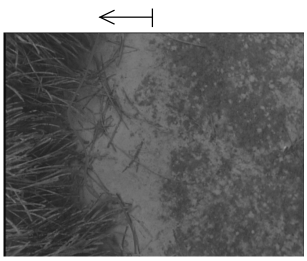


Fig. 1. Control signal for the rate controller.

This paper addresses this last problem:

Defiition **Visual Contour Tracking Problem**

Given a sequence of images $I^t = \{I_u, u \le t\}$, acquired along a permanently observing trajectory, identify the image of the contour S_{in} in frame I_t :

$$c_t = \bigcup_{i=1,\dots,N_t} s_t^i \;,$$

where each s_t^i is a *connected curve* in R^c such that $p \in s_t^i \subset R^c \Rightarrow S\left(T_{X(t),S}[p]\right) \in \partial S_{in}$.

3. ASSUMPTIONS

Our approach to the visual contour tracking problem considers several assumptions, which are summarized in this section.

3.1 Monochromatic camera

We assume that the pixel values are levels of intensity coded in a discrete alphabet: $L = [0,255]$.

3.2 Observation geometry

We consider that during observation the video camera is maintained pointing at the bottom, at a constant altitude above the sea bed, which is assumed to be approximately planar inside the field of view of the camera. We neglect roll and pitch variations, assuming that the camera axis is maitained nearly orthogonal to the sea floor. These assumptions, which impose additional constraints to the observation problem, are motivated by the desire to minimize the variations of the visual appearance of the classes during contour observation.

3.3 Probabilistic model of tracked region

We assume that the pixel intensities corresponding to the homogenous region being tracked, S_0 , are statistically independent realisations of an unknown probability distribution $h_0 \in \mathcal{M}^{|L|}$, where $\mathcal{M}^{|L|}$ is the probabilistic simplex of dimension $|L|$. Note that we assumed that the regions have been numbered such that the tracked region is S_0 . In the sequel we call h_0 the *statistical model for the tracked class*.

In our definition of the the tracking problem, we assume that the users goal is to delineate some class of interest (the **tracked class),** whose visual appearance is stable, while the background classes (on the opposite side of the contour) may change along the boundary.

3.4 Probability model of the background classes

We do not assume homogeneity of the other regions $\{S_i\}_{i=1}^{K-1}$. This is important for natural scenes, as the one shown in Fig. 1, for which the backround scene can contain transition zones between areas occupied by distinct materials or species. In this way, our algorithm will be able to accomodate to complex backgrounds, with characteristics that can vary significantly along the observed contour.

3.5 Motion observation

We assume that (noisy) estimates of the intra-frame displacement and orientation variation may be available from external sensors or processing algorithms. Orientation information is typically available from rate gyros and/or compasses, and displacement can be infered from vision and altimeter information. Availability of these estimates is not mandatory but speeds up processing, allowing the determination of a better initial estimate of the contour inside each new video frame.

4. CONTOUR ESTIMATION

4.1 General approach

As we defined it, the **Visual Contour Tracking Problem** involves tracking of the contour in a sequence of images acquired by a moving obsever. As we said in the Introduction, our algorithm is based on the notion of deformable contours, or *snakes*. A *snake* $s(l,\tau)$ is a one-dimensional curve that evolves in time in order to minimize a given energy $E(s(\cdot,\tau))$, in general defined as the sum of a line integral along the curve and a surface integral over the regions into which the (closed) curve divides the image. From this minimisation problem is derived an Euler-Lagrande *pde* equation that is integrated from some initial condition $s(\cdot,0)$ to obtain a stationary solution $s(\cdot,\infty)$.

We present in this section the steps on which we base determination of the contour for the last received frame I_t . We assume that a valid contour estimate s_{t-1} has been found in the previous frame I_{t-1} and that the observer's motion is such that there is a non-empty intersection of the parts of the contour in successive frames. Let ∂S_{in} be the tracked contour, as before. We assume that $W_t \cap W_{t-1} \cap \partial S_{in} \neq \emptyset$, $\forall t$.

The following three steps determine the contour in the present frame, s_t :

- **Prediction:** using available motion estimates, project the part of the contour s_{t-1} inside $W_{t-1} \cap W_t$ onto the new image I_t (see Fig. 2 left and center). Denote by $s_t(l,0)$ the resulting curve.
- **Adjustment:** From initial condition $s_t(l,0)$ use a variational approach to drive the snake towards the contour (see centre of Fig. 2). Let $s_t^*(l,\infty)$ be the resulting curve.
- **Extension:** to provide a complete description of contour inside the observed araea W_t iteratively prolonge $s_t^*(l,\infty)$ of a fixed amount and adjust the new section using a variational approach, until the border of the image is attained (right of Fig. 2). Denote by $s_t = s_t(l,\infty)$ the curve obtained in this way.

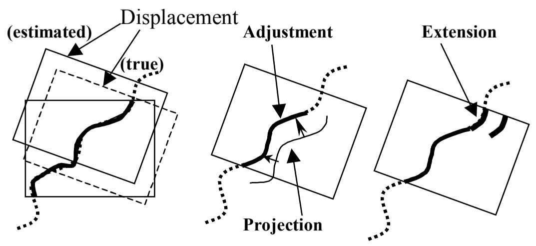


Fig. 2. *Prediction*, *adjustment* and *extension* steps.

4.2 The defining energy

To define our energy we start by defining a potential function over the affine plane. Let $u \in R$ be a point in the image support, and $\vec{v} \in S^1 \subset \Re^2$ be a vector in the unit sphere. Let $R(\vec{v})$ be the orthogonal rotation operator associated to $\vec{v}$, and let $\mathcal{L} = [-dl_x, dl_x] \times [0, dl_y]$. To each point of the affine space $(u,\vec{v})$ associate the sets:

$\mathcal{L}_{in}(u,\vec{v}) = u + R(\vec{v})\mathcal{L}$,

$\mathcal{L}_{out}(u,\vec{v}) = u + R(-\vec{v})\mathcal{L}$,

$\mathcal{L}_{b}(u,\vec{v}) = u + R(-\vec{v})(\mathcal{L} + dl_y)$.

Fig. 3 illustrates the definition of these sets. We define the following scalar positive function on each point of the affine space, $(u,\vec{v})$:

$$P^*: \; R \times S^1 \to R^+$$

$$(u,\vec{v}) \mapsto P^*(u,\vec{v}) = K\left(h_0 \| h_{in}(u,\vec{v})\right) + K\left(h_b(u,\vec{v}) \| h_{out}(u,\vec{v})\right)$$

where

- $K(p\|q) = E_p[\ln p/q]$ is the Kullback-Leibler divergence [1] between distributions p and q .
- $h_{region}(u,\vec{v})$ denotes the histogram of the pixel intensities inside the sets $\mathcal{L}_{region}(u,\vec{v})$, where $region \in \{in, out, b\}$.

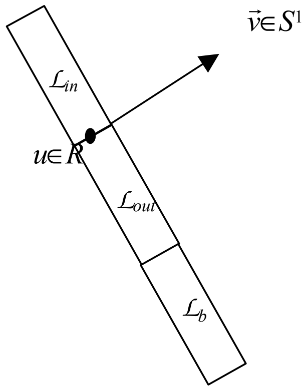


Fig. 3: Definition of the regions $\mathcal{L}_{in}(u,\vec{v})$, $\mathcal{L}_{out}(u,\vec{v})$ and $\mathcal{L}_{back}(u,\vec{v})$.

The first term of P^* measures the similarity (in statistical terms) of the region on the left side of a curve going through point u with orientation $\vec{v}$ to the statistical model of the tracked region, while the second terms measures the similarity of the region on the other side of the curve to the regions further away from the segment in the same direction (the background region, for which no reference statistical model is assumed, a local estimate being computed around each neighborhood of the contour). With this formulation we do not impose that the background has a constant distribution all over the contour. It is easy to check that this energy is equal to zero when the snake point $s_k(l)$ is located on the contour, and the pixel i ntensities in the its vicinity are drawn from the assumed (h_0) and locally estimated (h_b) distributions.

We can now define the potential function over all points of any curve $s(l)$ in the image as

$$P(s(l)) = P^*(s(l), \vec{n}(l))$$

where $\vec{n}(l)$ is the unit vector in the direction of the tangent to the curve if the tracked region is on the left, and on the reverse direction otherwise. Using this definition of potential, we finally define the energy associated to a curve in the image as the line integral of the potential along the curve:

$$E(s) = \frac{1}{|L|} \int_s P(s(l))\, dl$$

Our potential function does not explicitly consider regularising terms corresponding to an internal energy of the curve. In [6] it has been shown that the optimal, curvature constrained snake is a cubic spline. In order to avoid the explicit modelling of the internal forces (represented by first and second order derivatives) the authors replaced the snake by a b-snake, introduced in [10] and directly modify the spline parameters.
Mimisation of E will be done under constraints that correspond to a particular assumed parametrisation of the curve, such that in fact $s(l,\tau)=s(l,\alpha(\tau))$, where the exact meaning of the parameters $\alpha(\tau)$ will depend on the step of the general algorithm present before.

4.3 Solving the minimisation problem

We base energy minimisation on an iterative gradient descent by modifying the parameters $\alpha(\tau)$ such that the resulting energy converges to a minima, that us chose $\delta\alpha$, such that $E(s(\cdot,\alpha+\delta\alpha))<E(s(\cdot,\alpha))$. Developping $E(s(\cdot,\alpha+\delta\alpha))$ using a first order Taylor expansion we obtain the following approaximation to the energy variation:

$$\delta E \approx \frac{1}{|L|}\delta a\int\frac{\partial P(l)}{\partial\alpha}dl .$$

By setting

$$\delta\alpha=-\gamma_P\frac{1}{|L|}\int\frac{\partial P}{\partial\alpha}dl \qquad \text{(xxx3)}$$

where γ_α is a positive constant, the energy of the curve will decrease monotonically, and will converge to a stable minimum, since the energy is always positive. Equation (xxx3) can be computed by rewriting equation (xxx2) as

$$\delta\alpha=-\gamma_P\frac{1}{|L|}\int\left(\frac{\partial P(s(l))}{\partial s_x(l,\alpha)}\frac{\partial s_x(l,\alpha)}{\partial\alpha}+\frac{\partial P(s(l))}{\partial s_y(l,\alpha)}\frac{\partial s_y(l,\alpha)}{\partial\alpha}\right)dl$$

where $s(l)=[s_x(l)\,s_y(l)]^T$.

Depending on the specific set of parameters that are adjusted, distinct formulas will be used for the partial derivatives $\partial s_{x/y}/\partial\alpha$, which are presented in Sections 6 and 7. Instead of directly computing the partial derivatives of P with respect to the contour points, $\partial P/\partial s_{x/y}$, we base their evaluation on the assumption of a statistical mixture model for the local histograms h_{in} and h_{out}, as discussed in the next subsection.

4.4 Potential gradient

For an arbitrary location of the snake the distributions $h_{in}(l)$ and $h_{out}(l)$ are, with generality, mixtures of the distribution of the tracked class, h_0 , and of the local background class , $h_b(l)$:

$$h_{in}^{(l)}=\alpha(l)h_0+(1-\alpha(l))h_b^{(l)}$$
$$h_{out}^{(l)}=\beta(l)h_0+(1-\beta(l))h_b^{(l)} . \qquad (4)$$

For this model the energy given by equation (2) is always defined for values of $\alpha(l)$ and $\beta(l)$ in the open interval (0,1) even if h_0 and $h_b(l)$ are not mutually absolutely continuous.

Using this model, we can easily compute the gradient of $P(l)$, $\nabla P(l)=\left[\frac{\partial P(s)}{\partial s_x}\frac{\partial P(s)}{\partial s_y}\right]$ as

$$\nabla P(l)=\nabla_l\alpha(l)\lambda_{in}(l)+\nabla_l\beta(l)\lambda_{out}(l) ,$$

where

$$\lambda_{in}(l)=\sum_i h_0(i)\frac{\left(h_0(i)-h_b^{(l)}(i)\right)}{h_{in}^{(l)}(i)}$$

and

$$\lambda_{out}(l)=\sum_i h_b^{(l)}(i)\frac{\left(h_0(i)-h_b^{(l)}(i)\right)}{h_{out}^{(l)}(i)} .$$

It can be shown that $\lambda_{in}(l)\geq 0,\forall\alpha(l)\in[0,1]$ and $\lambda_{out}(l)\leq 0,\forall\beta(l)\in[0,1]$. From equation (4) it is easy to see that for snake points that lie in the **inner** region the mixture term $\alpha(l)$ is equal to one (and $\nabla_l\alpha(l)$ is zero) and, for snake points that lie in the **outer** regions $\beta(l)$ is equal to one (and $\nabla_l\beta(l)$ is zero). One of the two terms of $\nabla P(l)$ is thus always equal to zero.
To compute the gradient, we must thus detect whether the snake point is located in the *inner* or in the *outer* region. This decision can be made if the mixture coefficients are known. We use a test on $\hat\alpha$ and $\hat\beta$, the *maximum likelihood estimates* of $\alpha(l)$ and $\beta(l)$, setting:

$$\nabla_l\alpha(l)=0 \text{ and } \nabla_l\beta(l)=c\vec n, \text{ if } \hat\alpha+\hat\beta>1 ,$$
$$\nabla_l\alpha(l)=c\vec n \text{ and } \nabla_l\beta(l)=0,, \text{ if } \hat\alpha+\hat\beta\leq 1 ,$$

for some constant $c>0$.
To our knowledge, the approach presented in this section to determine the gradient of the energy at each point, based on a mixture model of the regions in the vicinity of the contour, is new. It results in "image gradients" which are not dependent on the intensity of the individual contour pixels (as in [Barluad]) but rather on the collective characteristics of neighboring regions. This explains why in the introduction we claimed that our approach is somehow between pure "contour-absed" and pure "region-based" approaches to deformable contours. We retain the simplicity of contour based approaches (since our energy is indeed a line integral) and it is the determination of the variations of the energy with respect to the location of the contour points that reflects the local (instead of punctual) characteristics of the image.

4.5 Estimation of the mixture coefficients

As explained above, our test of whether the contour point is located inside or outside the tracked region is based on the maximum likelihood estimates of the mixture coefficients $\alpha(l)$ and $\beta(l)$, which are solution of the following minimization problem:

$$\hat\alpha=\arg\min_{\alpha\in[0,1]}\left(K(h_{in}(l)\|\alpha(l)h_0+(1-\alpha(l))h_b(l))\right),$$

with an analogous expression for $\hat{\beta}$.For arbitrary distributions this problem must be solved numerically. However, experimental evidence shows that the minimised criterion is approximately quadratic in its argument,

$$K(h\|\nu h_0+(1-\nu)h_b)\cong c_0+c_1\nu+c_2\nu^2 \ , \qquad (7)$$

enabling determination of a sub-optimal estimate by computing the value of K for three distinct values of ν . Let $K_{0.25}$, $K_{0.5}$ and $K_{0.75}$ denote the values of the Kullback-Leibler divergence for corresponding values of ν. By solving $K=Vc$, where c groups the three coefficients in the previous equation,

$$\begin{bmatrix} K_{0.25} \\ K_{0.5} \\ K_{0.75} \end{bmatrix} = \begin{bmatrix} 1 & 0.25 & 0.25^2 \\ 1 & 0.5 & 0.5^2 \\ 1 & 0.75 & 0.75^2 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix}$$

we can finally compute the estimate of the mixture coefficient as $\hat{\nu}=-\hat{c}_1/(2\hat{c}_2)$.

The equations above for the estimate of the mixture coefficient are only meaningful when the distributions h_0 and $h_b(l)$ are "sufficiently different," otherwise the mixture coefficients are undetermined (the estimation problem is in this case singular). Since the estimation of $h_b(l)$ is done in a small window W_b on the outer region in the vicinity of the snake, it is possible (for large errors) that this window lies entirely inside the tracked region, resulting in a local estimate that resembles the reference histogram: $h_b(l)\approx h_0$. We must thus separately detect this situation, by checking similarity between both histograms. From equation (7) it is clear that the mixture coefficients are only well determined when the second derivative of $K(h\|\nu h_0+(1-\nu)h_b)$ – $2c_2$ – is large. We decide that the snake is inside tracked region when $\max(c_2^{(\alpha)},c_2^{(\beta)})<\varepsilon$,where ε is a user-defined threshold. Automatic choice of this threshold is a point of current research.

5. PREDICTION

Based on the set of assumptions presented in section 3. we assume that the images of the contour inside the intersection $W_{t-1}\cap W_t$ are related by a rigid **affine motion**. We obtain thus an initialisation of the snake for the next imag by

$$s_k(l)=\begin{bmatrix} x(l) \\ y(l) \end{bmatrix}=A_k^0 s_{k-1}(l)+t_k^0 \ . \quad \text{(xxx0)}$$

The affine motion parameters, i.e., matrix A_k^0 (rotation and scaling) and translation t_k^0 are determined from measures of the vehicle's displacement (ΔP_k) rotation ($\Delta\theta_k$) and change of scale (Δs_k) induced by altitude variations:

$$A_k^0=\Delta s_k R_{-\Delta\theta_k}$$

$$t_k^0=-\Delta s_k R_{-\Delta\theta_k}\Delta P_k \ ,$$

where R_θ is the rotation matrix by angle θ . It is clear that this prediction model can be extended to include other perspective deformations induced by strong pitch or roll.

6. ADJUSTMENT

Adjustement proceeds in two consecutive steps. In a first step (registration) we adjust the **affine motion parameters** – preserving the shape of the snake – and in a second step (*shape adjustment*) we release the affine motion constraint to allowing **shape deformations**. Note that in both steps the minimised energy is the same, only the parameters $\alpha(l)$ are distinct, as we have mentioned in section 4.3.

6.1 Registration

In this case we adjust the **affine motion parameters**: $\alpha=(s,\theta,t_x,t_y)$, that define (A_k, t_k) in equation (??). The derivatives with respect to α are easily calculated from the general definition of the affine motion parameters.

The expressions for the gradient of the energy function involve an integration of the gradient of the potential over the contour. We approximate this continuous integral by a summation over a set of discrete sample points along the contour.

6.2 Shape adjustement

During shape adjustment, we directly act upon the set of N control points $\alpha=\{c_1,\ldots,c_N\} c_i=[c_{ix}\ c_{iy}]^T\in R^2$, that determine the contour representation as a B-spline. Each contour point, $s(l,\alpha)$, is a linear combination of the N control points weighted by a set of basis-functions [reference]:

$$s_x(l,\alpha)=\sum_{i=0}^{N-1}c_{ix}b_i(l)\ , \quad s_y(l,\alpha)=\sum_{i=0}^{N-1}c_{iy}b_i(l)\ .$$

The partial derivatives of each contour point with respect to the parameters α (in this case the control points) are thus simply given by:

$$\frac{\partial s_{x/y}(l)}{\partial c_{i(x/y)}}=b_i(l)\ ,$$

which should be used in the expressions of section 4.3. Note that for a constant number L of sampled points, $\{s(0),\ldots,s(L-1)\}$ and of control points the basis functions $b_i(l)$ can be pre-computed, speeding algorithm execution. As in the previous case, the integral that defines the resulting energy gradient is approximated by a finite sum over a discrete set of points,

7. EXTENSION

Assume that the projected snake has converged towards the contour. In the case that the snake's end points are not located in the image borders we extend the snake at both extremes. Assuming local continuity of the contour, the new points are placed at some fixed distance from the current end points along their tangent vectors. Each new point is updated uasing the adjustment step described in section 6.2.

8. EXPERIMENTAL RESULTS

The images that are presented in this section have been acquired at Villefranche-sur-mer, south of France using a ROV Phantom in an autonomous tracking mode at constant altitude. Fig. 3 shows a mosaic of the complete contour that has been tracked. The tracked class is a large patch of Posidonia and the background is mostly occupied by sand. However, as the image shows, there are significant variations of the background along the sequence. Fig. 4 illustrates the initialisation of the snake based on a coarse segmentation of the image [3]. Fig. 5 shows the algorithm result over 8 consecutive frames. The crosses indicate the control points used in the adjustment step. These Figures illustrate the successful propagation of the snake along a contour portion with high curvature.

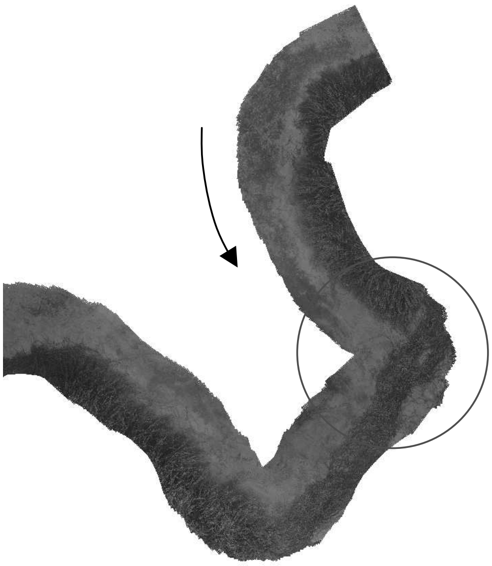


Fig. 3. Mosaic of the tracked contour (390 images) – circle indicates the region shown in the detailed representation in Fig. 5.

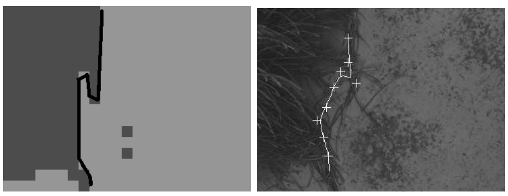
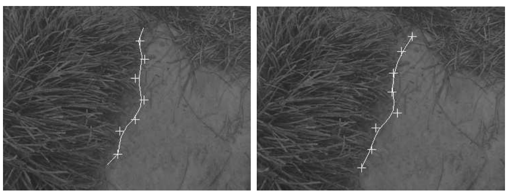


Fig. 4. Initialisation of the snake based on a coarse image segmentation (1st frame).



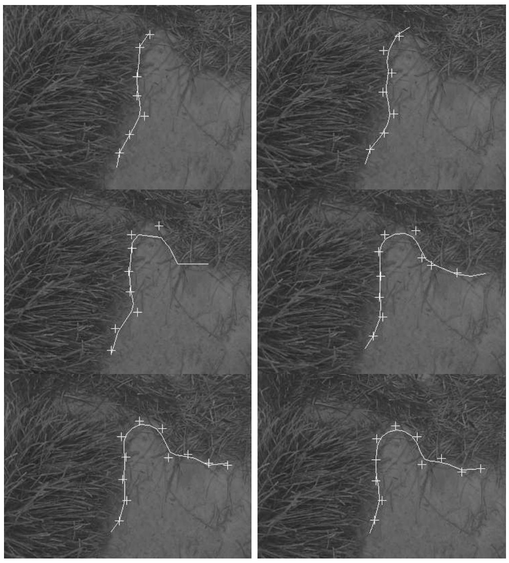


Fig. 5. Frames 123 – 130 of the real video sequence in Fig. 3.

REFERENCES

[1] S. Kullback Information Theory and Statistics, Peter Smith, 1978.

[2] S. Jehan-Besson, Modèles de contours actifs basés régions pour la segmentation d'images et de vidéos, PhD Thesis, Université de Nice-Sophia Antipolis, 2003.

[3] Tenas, A.., M.J. Rendas and J.P. Folcher (2001). Image Segmentation by Unsupervised Adaptive Clustering in the Distribution Space for AUV guidance along sea-bed boundaries using Vision. In: *Proc. Oceans 2001*, Honolulu, Hawaii, USA, Nov. 2001.

[4] McInerney, T. and D. Terzopolous (1995) Topologically Adaptable Snakes. In: *Proc. Of the fifth Int. Conf. On Computer Vision*, Cambridge, MA, USA, June 1995.

[5] Chan, T. and L. Vese (1999). An Active Contour Model without Edges. In : *Scale-Space'99*, pp. 141-151, Springer-Verlag Berlin Heidelberg 1999.

[6] Brigger, P., J. Hoeg and M. Unser (2000). B-Spline Snakes : A Flexible Tool for Parametric Contour Detection. In : *IEEE Transactions on Image Processing*, **Vol. 9**, No. 9, pp. 1484-1496, Sept. 2000.

[7] Blake, A. and M. Isard (1998). *Active Contours*, Springer-Verlag, 1998.

[8] Bascle, B., P. Bouthemy, R. Deriche and F. Meyer (1994). Tracking complex primitives in an image sequence*, Technical Report 2428 INRIA* Soph.-Antipolis France, Dec. 1994.

[9] Toledo, R., X. Oriols, X. Binefa, P. Radeva, J. Vitria and J. Vilanueva (2000). Tracking elongated structures using statistical snakes. In: *Conf. on Computer Vision and Pattern Recognition*, Hilton Head, SC, USA, June 2000, pp 157-162.

[10] Menet, S., P. Saint-Marc and G. Medioni (1990). B-snakes: Implementation and appplication to stereo. In: *Image Understanding Workshop*, Sept 1990, pp. 720-726.