HYBRID GENETIC OPTIMIZATION AND STATISTICAL MODEL-BASED APPROACH FOR THE CLASSIFICATION OF SHADOW SHAPES IN SONAR IMAGERY

Person to whom correspondence should be sent:
Name: Christophe Collet
Address: Groupe de Traitement du Signal, Ecole Navale, Lanvéoc-Poullmic,
BP 600 - 29240 Brest-Naval, France.
E-mail: collet@poseidon.ecole-navale.fr
Phone: (33)(0) 298 234 044
Fax: (33)(0) 298 233 857
Name: Mignotte Max
Adresse: Groupe de Traitement du Signal, Ecole Navale, Lanvéoc-Poullmic,
        29240 Brest-Naval, France.
Phone: (33)(0) 298 234 044
Fax: (33)(0) 298 233 857

Name: Patrick Perez
Adresse: IRISA/INRIA, Campus Universitaire de Beaulieu, 35042 Rennes cedex, France.
Email: perez@irisa.fr
Phone: (33)(0) 299 847 273
Fax: (33)(0) 299 847 171

Name: Patrick Bouthemy
Adresse: IRISA/INRIA, Campus Universitaire de Beaulieu, 35042 Rennes cedex, France.
Email: bouthemy@irisa.fr
Phone: (33)(0) 299 847 274
Fax: (33)(0) 299 847 171
HYBRID GENETIC OPTIMIZATION AND STATISTICAL
MODEL-BASED APPROACH FOR THE CLASSIFICATION
OF SHADOW SHAPES IN SONAR IMAGERY

M.Mignotte ‡ C.Collet ‡ P.Pérez • P.Bouthemy •

‡ Groupe de Traitement du Signal, Ecole Navale, BP 600, 29240 Brest-Naval France.
e-mail: collet@poseidon.ecole-navale.fr

• IRISA/INRIA, Campus Universitaire de Beaulieu, 35042 Rennes cedex, France.
e-mail: name@irisa.fr

Acknowledgments: The authors thank GESMA (Groupe d’Étude Sous Marine de
l’Atlantique, Brest, France) for having provided numerous real sonar pictures and DGA
(Direction Générale pour l’Armement, French Ministry of Defense) for partial financial
support of this work (student grant).
ABSTRACT

We present an original statistical classification method using deformable template model to separate natural objects from man-made objects in an image provided by a high resolution sonar. A prior knowledge of the manufactured object shadow shape is captured by a prototype template along with a set of admissible linear transformations to take into account the shape variability. Then, the classification problem is defined as a two-step process. First the detection problem of a region of interest in the input image is stated as the minimization of a cost function. Secondly, the value of this function at convergence allows one to determine whether the desired object is present or not in the sonar image. The energy minimization problem is tackled using relaxation techniques. In this context, we compare the results obtained with a deterministic relaxation technique (a gradient-based algorithm) and two stochastic relaxation methods: Simulated Annealing (SA) and a hybrid Genetic Algorithm (GA). This latter method has been successfully tested on real and synthetic sonar images, yielding very promising results.

Index terms: deformable template, objective function, simulated annealing, gradient-based algorithm, genetic optimization, shape recognition, sonar imagery.
1. INTRODUCTION

Due to their high resolution, existing sonars can detect every object lying on the sea-bed. Each object can be identified efficiently thanks to an analysis of its associated cast shadow (arising from the lack of acoustic reverberation behind the object). However, since the amount and size of acquired images is huge, the exploitation of the collected data has to be achieved with an automatic processing chain. The goal is to be able to detect, in the context of our application (i.e., mine detection and neutralization, underwater wreckage research, submarine rescue...) manufactured objects lying on the sea-bed. On large sea floor areas (typically 1000 meters by 200 meters) the goal is to find an object of size smaller than 6 meters by 3 meters. Immediately after such a detection, others means are deployed in order to confirm the detection and eventually destroyed the threat. It is our applicative context. The proposed method may also be efficiently exploited to detect and to localize cylindrical or spherical objects lying on the sea-floor. Unfortunately, confidentiality reasons prevent us from providing too many informations on the application. The approach presented in this paper allows efficiently to distinguish natural objects from such manufactured objects lying on the sea-bed and the whole paper is presented in this intention.

In computer vision, pattern recognition issues are often cast as a search problem in some appropriate feature space. In that prospect, a classical and generally used recognition approach in sonar imagery consists in using a pre-segmented sonar image and associating for each detected cast shadow shape, a set of appropriate geometrical, spectral or statistical features. The aim is to represent the (shadow) shape with a relevant parameter vector providing maximal information about the extracted shadow. Afterwards, this parameter vector is usually exploited as an input of a supervised classifier (Bayesian classifiers, K-nearest neighbors or multi layer perceptron ...) [1, 2]. The classification problem is reduced to recover from a database the occurrence that is the more similar to the analyzed one. This method is appealing and has been widely used in pattern recognition applications. Nevertheless, it requires that the training set is large enough to ensure
an efficient learning stage. Besides, the efficiency of this technique depends on the feature extraction step for which there are no general rules to determine the optimal number of parameters and the discriminant characteristics of each of them. Note that this method cannot be used efficiently when the extracted shadow shape is partly merged with another object shadow. In our application, it is often the case, cast shadow shapes being altered by the shadow of ridges of sand, dunes or pebbles [3].

Contrary to the cast shadow of a natural object, the one of a manufactured object has a regular and/or geometrical shape easily identifiable. Bayesian statistical theory is a convenient way of taking this a priori information into consideration. This paradigm in image analysis is quite popular and has been successfully applied for image segmentation [4, 5], image restoration [6] (with a local prior model) or shape matching with deformable template-based methods (with a global prior model) [7, 8]. Laksmanan et al. [9], for example, have used a parametric template model to locate the road boundary in radar images where the two straight, parallel edges of a road are parameterized. Then, the edge detection problem is formulated as a Bayesian optimization problem using a physics-based model of the radar imaging process. A similar approach for shape matching is proposed by Jain et al. [10, 11] which combines in the same manner, both the available knowledge of the shape properties (as prior model) and an observation model (as likelihood model). One such method can be efficiently used for the classification of the cast shadow of each object lying on the sea-bed and more precisely to distinguish natural objects from man-made objects in sonar imagery. In this way, we define a prototype template and a set of admissible linear transformations to take into account the shape variability of the object class to be detected. Also, we define a joint Probability Density Function (PDF) which expresses the dependence between the observed image and the deformed template. Then, the detection problem of an object class is stated as the estimation of the deformation parameters of the template that maximize the posterior PDF. This maximization problem is equivalent to the minimization of a non-concave (usually complex with many local extrema) objective function.
In [10], gradient-based methods are used for the energy minimization of this function. These methods have the disadvantage to require good initial parameter estimates (i.e., a proper initialization of the template), otherwise they will converge toward bad local minima. Stochastic methods based on Simulated Annealing (SA) [12, 13] have the capability of avoiding local minima and no human interaction is required to initialize the model. However, one of the major drawbacks of this procedure is its high computational load. Hereafter, we show that an alternate approach consists in using a genetic exploration of the parameter space.

The main contribution of this paper lies in the use of deformable model to classify objects in sonar imagery. This appears as an interesting alternative to feature-based matching methods, and it is efficient in the case of complex background and/or occlusion phenomena between several cast shadows. We propose an appropriate energy function that differs from previously published works. This energy utilizes the information given by an unsupervised Markovian segmentation of the input sonar image [5] and integrates both region homogeneity and edge information. Finally, we introduce a computationally efficient global optimization method to solve the minimization problem. This method is based on a stochastic search method using a genetic exploration of the parameter search space combined with a steepest ascent procedure and a cooling temperature schedule. In such a context, we will show that the use of Genetic Algorithm (GA) leads to better performance than other stochastic methods (such as simulated annealing) or deterministic relaxation techniques like gradient-based algorithm.

This paper is organized as follows: sections 2 and 3 present the template modeling and the energy-based extraction of template occurrences from data and the classification of extracted shapes. The optimization using gradient-based algorithm, simulated annealing, or GA is described in Section 4. In Section 5, we report some detection/classification results on real and synthetic sonar images. In this context, advantages of GA over the other optimization methods are discussed. Section 6 contains concluding remarks.
2. TEMPLATE REPRESENTATION

Contour-based modeling is a generic tool for shape matching. It can be applied to objects of different shapes by defining different prototype templates.

In sonar imagery, the first goal of the classification step is to automatically separate natural objects from man-made objects. This low-level classification step in two classes is based on the following a priori information: contrary to natural objects, a manufactured object is mainly composed of elements with simple and regular geometrical shape. For this reason, their cast shadow present geometric properties and exhibit straight lines as contours in the case of wrecks, pipe-lines, etc. In particular, the cast shadow of a cylindrical or cubic object is a perfect parallelogram. Therefore, we simply define the corresponding prototype template as a parallelogram characterized by its four vertices. For a different application however, we have to detect spherical objects lying on the seabed. In this case, the associated cast shadow has a typical shape whose representation can easily be defined by a set of \( n \) points manually selected or equally sampled which approximate its outline. A cubic B-spline shape representation involving these \( n \) control points corresponding to “landmarks” is then defined. The outline of the spherical object cast shadow used to build the prototype can either be obtained from a real scene or synthesized (see Fig. 1). The latter way of modeling objects has been widely considered in the object recognition literature, and particularly in the active contour approach [14]. Such a scheme captures the global structure of a shape without specifying a parametric form for each class of shapes.

The prototype template, denoted \( \gamma_0 \), does not describe the possible instances of the shapes to be detected for a class of objects. In order to take into account the variability of the considered object class, we introduce a set of admissible linear transformations on \( \gamma_0 \). Let \( \gamma_\theta \) be a deformed version of the original prototype \( \gamma_0 \) according to a affine transformation with parameters vector \( \theta \). In the case of our first parametric template (used to detect manufactured object), these deformations involve translation, scaling, rotation, stretching and skewing [15] of the template as shown in Figure 2. Corresponding
transformations are given by:

\[
A_s = \begin{pmatrix} s & 0 \\ 0 & s \end{pmatrix} \quad A_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad A_t = \begin{pmatrix} 1 & 0 \\ 0 & t \end{pmatrix} \quad A_u = \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix}
\]

with \( s, t, u \) and \( \alpha \) being respectively the scale, stretch and skew parameters and the angle of rotation.

Due to spherical symmetry, the spherical object cast shadow is symmetric relative to the sonar beam direction. Therefore, the only considered transformations are translation, scaling and stretching for this template (see Fig. 3).

Otherwise, parameters of the translation and the scale transformation are actually constrained by the size of the sub-image. Translation parameter (in the \( x \) or \( y \) axis direction) as well as scale, stretch and skew parameters are constrained such that the template stay, at least partially, within the sub-image.

In order to correctly detect and classify cast shadows that are partly outside the sub-image, we have to partitioned the input sonar image of large size (typically 6000 by 2000 pels) into small overlapping sub-images (256 by 256 pels), whose size is function of the template size we are looking for (typically 30 by 60 pels). The detection-classification method presented in the next part has then to be used for each overlapping sub-image.

3. MAP DETECTION

A common problem in sonar images are the artefacts caused by the speckle noise effect, which leads to a loss of signal and a very poor quality of the object boundaries [16]. That is why the joint model we propose does not use directly the input image (i.e., the grey levels themselves) or some image gradient measure (as proposed in [10]) in order to recover the contour of each object. In our approach, we use the result of an unsupervised two-class Markovian segmentation of the input sonar image [5]. It provides a binary map \( x = \{ x_s, s \in S \} \) where \( x_s = e_0 \) for shadow areas and \( x_s = e_1 \) for reverberation areas. This allows us to take into account the observed measurements along regions contours but
also the grey-level homogeneity information inside and outside delimited regions. The detection is based on an objective function $\epsilon$ measuring how well a given instance of deformed template $\gamma_\theta$ fits the content of segmented image $x$. From a probabilistic point of view, $\epsilon(\theta, x)$ defines the joint model through the Gibbs distribution:

$$ P_{\Theta, X}(\theta, x) = \frac{1}{Z} \exp\{-\epsilon(\theta, x)\} $$  \hspace{1cm} (1)

where $\Theta$ is the random vector of parameters, and $Z$ a normalizing constant.

3.1. Joint model

The posterior distribution deduced from (1):

$$ P_{\Theta/X}(\theta/x) = \frac{1}{Z_x} \exp\{-\epsilon(\theta, x)\} $$  \hspace{1cm} (2)

provides the probability of a given template given the segmented image. $\epsilon(\theta, x)$ represents a measurement of the similarity between the deformed template and the object present in the image. The energy function $\epsilon(\theta, x)$ is composed of two terms as explained below:

**Edge energy**: Let $S'$ be the set of 1-labeled pixels in a binarized high-pass filtered version of the segmented sonar image $x$ in two classes (i.e., $S'$ represents the set of edges associated to detected cast shadows). The deformable template is attracted and aligned to the salient contours of each object via an edge potential field defined as follows. Each site $s_0 \in S'$ associated to an edge of a detected shadow region in $S'$ creates an elementary potential field $\phi_{s_0}(r)$ such as:

$$ \phi_{s_0}(r) = \frac{1}{r} \exp\left(-\frac{r}{\sigma}\right) $$  \hspace{1cm} (3)

where $r (r \neq 0)$ is the distance to the pixel $s_0$. Figure 4 shows an example of this elementary potential field. The different edges in $S'$ create a potential field $\Phi_{S'}(t)$, given by the total sum of the different elementary potential fields $\phi_{s_0}(r)$:
\[ \Phi_S(t) = \inf \left\{ \sum_{s \in S'} \phi_s(d(s, t)), 1 \right\} \quad \forall t \in S' \tag{4} \]

with \( d(s, t) \) is the distance between pixels \( s \) and \( t \). In fact, this potential field induces a smooth version of the edge image in which a site close to an edge will get a potential value close to 1 as illustrated in Fig. 5. The degree of smoothness can be controlled by parameter \( \sigma \). This edge potential field induces an energy function that relates a deformed template \( \gamma_\theta \) to the edges associated to each detected object in the input image:

\[ \epsilon_c(\theta, x) = -\ln \left\{ \frac{1}{N_{\gamma_\theta}} \sum_{s \in \gamma_\theta} \Phi_S(s) \right\} \geq 0 \tag{5} \]

where the summation is over all the \( N_{\gamma_\theta} \) pixels on the deformed template \( \gamma_\theta \) and the logarithm function is used to increase the range of the energy function \( \epsilon_c \). This function attains its minimum value when the contour of the deformed template coincides exactly with the underlying image edges.

**Region homogeneity energy:** This energy term aims at placing the inside of the deformed template in a region classified as shadow by the segmentation procedure:

\[ \epsilon_r(\theta, x) = -\ln \left\{ \frac{1}{N_{\gamma_\theta}} \sum_{s \in \gamma_\theta} \delta(x_s, \epsilon_0) \right\} \geq 0 \tag{6} \]

where \( \gamma_\theta^* \) and \( N_{\gamma_\theta^*} \) represent the set of pixels and the number of pixels inside the region delimited by \( \gamma_\theta \) and \( \delta(.) \) is the Kronecker delta function. This function attains its minimum value when the set of pixels inside the contour have all been classified as shadow by the segmentation procedure.

Using these two energy terms, the posterior distribution of \( \theta \) given \( x \) is:

\[ P_{\Theta|X}(\theta|x) = \frac{1}{Z_x} \exp \left\{ -\epsilon_c(\theta, x) - \epsilon_r(\theta, x) \right\} \tag{7} \]

\[ = \frac{1}{Z_x} \frac{1}{N_{\gamma_\theta^*} N_{\gamma_\theta}} \left( \sum_{s \in \gamma_\theta} \Phi_S(s) \right) \left( \sum_{s \in \gamma_\theta^*} \delta(x_s, \epsilon_0) \right) \tag{8} \]

where \( Z_x \) is a normalizing constant depending on \( x \) only.
3.2. Detection step

We formulate the detection problem as the search of the Maximum A Posteriori (MAP) estimation of \( \theta \):

\[
\hat{\theta}_{MAP} \in \arg \max_{\theta} \{ P_{\Theta/X}(\theta / x) \} \tag{9}
\]

\[
\in \arg \min_{\theta} \epsilon(\theta, x) \tag{10}
\]

This function, \( \epsilon(\theta, x) \), is minimal when the deformed template exactly coincides with the edges of \( x \) and contains only pixels labeled as shadow.

3.3. Classification step

The resulting value of energy \( \epsilon(\hat{\theta}_{MAP}, x) \) is used to measure the degree of fitness of the template with the region of interest, and then, to determine whether the desired object is present or not. If \( \epsilon(\hat{\theta}_{MAP}, x) \) is lower than a given threshold, then the desired object is assumed to be present and the final configuration of the deformed template reveals the shape and the location of the detected object; otherwise, we decide that the desired object is not present.

The method can be easily improved in order to retrieve several object cast shadows looking like the prototype template in the input sub-image. In this way, multiple object cast shadows, that are like prototype template, can be localized by using our classification method. Then, we remove the edge in \( \Phi_S \) (the edge potential field) of the detected cast shadow shapes and we repeat the same procedure till the value of \( \epsilon(\theta, x) \) is sufficiently large”. This procedure prevents the algorithm presented in the following section to always “converge” toward the same object. Nevertheless, in our proposed detection-classification method, this abovementioned procedure is not necessary: when a cast shadow shape is detected as being the one belonging to a manufactured object, the associated sonar sub-image is then further analyzed by an expert, in order to confirm or to infirm this detection. Finally, in the mine war context, the number of no manufactured objects (rocks,...) is fortunately much large than manufactured objects!
4. OPTIMIZATION PROBLEM

The objective function to be maximized in Equation 7 is a complex function with several local extrema over the deformation parameter space. A global search is usually impossible due to the size of the configuration space. Instead, we have implemented and compared three different optimization techniques: a gradient-based method as a local minimization technique; the SA and a GA as global minimization techniques, which are able to escape from local energy minima.

4.1. Gradient-based algorithm

This method requires a good initialization of the template, near the true location of the cast shadow shape to be detected. Otherwise, we obtain a sub-optimal solution (local minima) and the value of the resulting objective function $\epsilon(\theta, x)$ cannot be exploited to affirm the presence of the desired object.

In order to avoid this difficulty, a solution, proposed in [10], consists in placing the template at evenly spaced positions and in deforming it according a discretized set of orientation, stretch, skew, values in the input image. These deformed template configurations can then be used to initialize a deterministic gradient descent algorithm. However, the sampling of template positions and transformation parameters must be chosen judiciously. It should be fine enough not to miss the significant local minima of the energy surface, and large enough to avoid high computational requirements.

4.2. Simulated Annealing Algorithm

SA [12] is a stochastic relaxation technique which is based on the analogy to the physical process of annealing a metal: at high temperature, the atoms are randomly distributed. With decreasing temperatures, they tend to arrange themselves in a crystalline state which minimizes the global energy. Using this analogy, the algorithm generates at random new configurations of the template by sampling the probability distribution of the system “at temperature $T$”:
where parameter $T$ slowly decreases to 0. More precisely, starting from a prototype template $\theta^{[0]}$, we construct a sequence of template deformations $\theta^{[1]}, \theta^{[2]}, \ldots, \theta^{[k]}$ such that \( \lim_{k \to \infty} \theta^{[k]} = \arg \min \epsilon(\theta, x) \). In our application, we consider the metropolis algorithm, and for every $k$, we sequentially update each of the component $\theta^{[k]}_i$ of $\theta^{[k]}$ (i.e., each deformation parameter) in the following manner:

- The algorithm generates a new value for $\theta^{[k]}_i$ (called $\theta^{[k]}_{new}$) that differs from the precedent value (that we call $\theta^{[k]}_{old}$) only by $\theta^{[k]}_i$, randomly chosen according to uniform distributions over admissible intervals of values.

- Then $\theta^{[k]}_{new}$ or $\theta^{[k]}_{old}$ is selected with probability $P^{[k]}_{new}$ or $P^{[k]}_{old}$, respectively:

\[
P^{[k]}_{new} = \min \left\{ \frac{\exp \left( -\frac{\epsilon(\theta^{[k]}_{new}, x)}{T_k} \right)}{\exp \left( -\frac{\epsilon(\theta^{[k]}_{old}, x)}{T_k} \right)}, 1 \right\}
\]

and

\[
P^{[k]}_{old} = 1 - P^{[k]}_{new}
\]

Increases of energy can thus be accepted, and the algorithm is able to escape from local energy minima. It has been shown [17] that the algorithm converges to a global energy minimum if the temperature at iteration $k$ is:

\[
T_k \geq \frac{T_0}{\log(1 + k)}
\]

where $T_0$ is a constant depending on the amount of energy which is necessary to escape local minima. Nevertheless, this condition results in an extremely slow procedure. One way to restrict the number of iterations is to use a geometric temperature schedule [12]:

\[
T_k = T_0 \left( \frac{T_f}{T_0} \right)^{\frac{k}{K_{max}}}
\]

where $T_0$ is the starting temperature, $T_f$ is the final temperature and $K_{max}$ is the number of iterations.
4.3. Genetic Algorithm

Genetic Algorithms (GA) are a class of robust stochastic search and global optimization procedures which mimics the evolution of natural systems [17]. The algorithm acts in an iterative way by allowing parallel evolution in a population of $N$ individuals. Each individual represents a point of the search space and is a candidate solution to the optimization problem. It is represented by a string or chromosome, which is composed of a list of $L$ features (corresponding to the $L$ search parameters). The parameters have to be encoded in an appropriate manner. The most common approach is to quantize the parameter values and to binary coded them. The appropriateness of the various individuals (the tentative solution) to the environment is expressed by a fitness function, which (after the characteristics contained in a chromosome have been decoded), gives a performance value to the string.

Genetic search is carried out in a sequence of generations. In each generation, a new population of $N$ chromosomes is created by the genetic operators. These operators mimic the biological phenomena of selection, crossover and mutation. The choice of the solution upon which they are used is dictated by the evolutionary principle of the survival of the fittest. The algorithm begins with an initial population of $N$ chromosomes randomly chosen and terminates when either a specified number of iterations has been performed or a maximally fit individual has emerged.

In our application, let us recall that we have to optimize a $L$ dimensional function ($L = 6$ or $L = 4$ depending on the considered deformable template). Each of the $L$ parameters is quantified on $q$ bits. Therefore, the $i$th chromosome $[\theta]_i$ is a string of $q \cdot L$ bits length:

$$[\theta]_i = (c^i_{11}, c^i_{12}, \ldots, c^i_{1q}; c^i_{21}, c^i_{22}, \ldots, c^i_{2q}; \ldots; c^i_{L1}, c^i_{L2}, \ldots, c^i_{Lq})$$

(16)
Fitness Measure:

We can easily derive a fitness measure $\mathcal{F}$ (to be maximized) directly from equation 9 (the energy function $\epsilon$ to be minimized). To turn $\epsilon(\theta, x)$ into a fitness measure for use in genetic algorithm (i.e., one with range $[0,1]$), one can choose:

$$\mathcal{F}([\theta]_i) = \exp\left\{-\epsilon([\theta]_i, x)\right\} \quad \text{since} \quad \epsilon([\theta]_i, x) \geq 0 \quad (17)$$

The following is the detail of the selection, crossover, and mutation operators. The associated parameters used in our sonar imagery application will be given in subsection 5.3.

Selection:

Individuals with higher fitness survive and individuals with lower fitness die. Let us assume that at iteration $k$, the population of the GA is the set of $N$ chromosomes:

$$\text{POP}^k = \left\{ [\theta]_1^k, \ldots, [\theta]_N^k \right\} \quad (18)$$

Generation of the next population is based on the evaluation of $\mathcal{F}$ for all individuals of $\text{POP}^k$. More precisely, we probabilistically select each chromosome for “reproducing” in the next generation, using their relative fitness:

$$p([\theta]_i^k) = \frac{\mathcal{F}([\theta]_i^k)}{\sum_{j=1}^{N} \mathcal{F}([\theta]_j^k)} \quad (19)$$

Crossover:

A pair of chromosomes is picked up at random and the single-point crossover operator is applied according to a fixed crossover probability. For this operation, a random number in the range of 0 to the length $Lq$ of the string is generated. This is called the crossover point. The portions of two strings lying to the right of the crossover point are interchanged to yield two new strings as shown in Fig. 6.
**Mutation:**

Mutation consists in considering in turn each bit of a given chromosome and changing its value with a predefined low probability called the mutation rate.

In order to prevent premature convergence [17] and to speed up the convergence rate, we have developed three strategies and we have combined them:

1.- The first one is an elite-preservation strategy [17]: the individual with the highest fitness always survives to be an individual of the next generation.

2.- The second strategy (called hybrid GA [17]) consists in associating the genetic search with a local optimization technique. In each generation, a percentage of the best individuals are used to initialize a gradient ascent technique. Therefore, these best individuals explore local neighborhoods in the parameter space to find a point of higher fitness.

3.- In order to improve the results and the robustness of the GA, we propose a third modification reminiscent to SA [18]. The fitness function $\mathcal{F}$ at iteration $k$ is defined as follows:

$$
\mathcal{F}^k([\theta]_i) = \exp\left\{-\frac{1}{T_k} \epsilon([\theta]_i, x)\right\}
$$

with $T_k = T_0 a^k$ and $a < 1$. At the beginning of the genetic search procedure, $T_k > 1$ and the optimization procedure exploits a smooth version of the energy function $\epsilon$. This smooth energy function has fewer spurious local minima, which helps the genetic procedure to maintain a good diversity in the population and to avoid a premature convergence toward a sub-optimal solution (corresponding to an important secondary minima). For $T_k = 1$ the genetic search is carried out with the real cost function $\epsilon$. At the end of the procedure, $T_k < 1$, the fitness measure falls off rapidly with increasing cost. This allows us to maintain a good competition between individuals located near the global optimum and to localize precisely the global extrema. This method is similar to the cooling temperature schedule proposed in the simulated annealing procedure. In our application, these three strategies
are used together in an efficient way. We call this algorithm a hybrid genetic algorithm with an elitist strategy and a cooling temperature schedule.

5. EXPERIMENTAL RESULTS

In order to distinguish natural from man-made objects, we have to search the presence or not of both templates. For the classification step, we use the following decision rules:

- If $\epsilon < 0.2$, we decide that the image contains an object whose cast shadow shape is similar to the template.

- If $0.20 < \epsilon < 0.23$, a doubt exists. An human expert is indispensable to classify these objects.

- If $\epsilon > 0.23$ We decide that the searched object is not present in the input image.

These different threshold values have been chosen empirically after a set of experiments from our database of 100 real sonar images (containing natural or manufactured objects of different nature) and with the optimization procedure ensuring the global minimum (i.e., with the genetic search). As we will show in the following, this decision rule leads, for our application, to a good classification rate.

The proposed deformable template-based detection and the classification scheme has been now applied to several real and synthesized sonar images containing natural or manufactured objects. These experiments have been carried out with the different optimization procedures previously presented. The efficiency of each method is discussed and experimental results are compared.

5.1. Gradient-based optimization

First experiments have been carried out with the gradient-based optimization and the sampling strategy described in subsection 4.1. In our application, we considered a number of possible initialization by selecting five values of each parameter (yielding respectively
5^4 possibilities of the first kind, and 5^2 of the second kind) and different locations evenly spaced within shadow regions.

Figure 7 shows the cast shadow created by a cylindrical object lying on a sandy sea floor and on a pebbly sea floor respectively. Let us note that, for the latter case, the cylinder object cast shadow occludes a number of rock shadows and the background is rather cluttered. For each case, we present the resulting deformed template and we indicate the corresponding energy value. In the first example, we can see that the outline of the cast shadow shape is accurately recovered, and the low value of the energy function allows us to correctly classify this shadow shape as a cylindrical object cast shadow. However, in the second case, the energy function is too complex (i.e., many local minima, due to the complex background and the different occlusions) and/or the initialization given to the gradient-based procedure is not good enough to ensure a correct detection. We obtain a sub-optimal solution (local minima) and the value of the resulting objective function does not allow us to recognize the presence of a manufactured object in this sonar image.

5.2. Simulated Annealing optimization

A second set experiments have been conducted using the SA optimization scheme with a geometrical temperature schedule and the following control parameters: \( T_0 = 2, T_f = 0.005, K_{max} = 15000 \). Tests have shown that this optimization procedure is very sensitive to the parameters of the cooling schedule. In our application, these parameters have been chosen empirically in order to get good convergence in all tested case (around hundred sonar pictures).

Figures 8 illustrate the evolution of the (parallelogram) template during the SA optimization on the same image as in figure 7.b. In the beginning, the temperature is relatively high, then the algorithm explores the space of template parameters even if the corresponding deformed templates do not satisfy the constraints and have a relatively high energy. As the temperature decreases, the constraints become satisfied, the outline of the template gets closer to the edges of the region of interest, and the energy converges...
towards the global minimum. Then, the low value of the energy (\(\epsilon = 0.13\), compared to \(\epsilon = 0.3\)) obtained by gradient-based technique allows to classify this region of interest as the cast shadow of a cylindrical object. In Figure 9, the detected region of interest extracted by the parallelogram template corresponds well to the manufactured object (cylindrical object) and the low value of \(\epsilon\) allows us to classify it as the cast shadow of a cylindrical object.

This procedure ensure a good detection and correct classification rate. Nevertheless, one of the major drawbacks of this optimization procedure is its high computational load. It takes about 7 minutes on a 43P IBM (120 MHz) workstation whereas about 1 minute is necessary for the gradient-based optimization previously described. However, the \(K_{\text{max}}\) parameter should be chosen lesser than 15000 but it seemed that this maximum iteration number was reasonable for such a robustness on a large data base. Moreover, the final temperature is very low (\(T_f = 0.005\)) : this ensures a quasi-deterministic search at the end of the stochastic procedure. This is why it is not necessary to apply a deterministic search at the end of the simulated annealing-based optimization procedure to further improve the optimization result. These two reasons explain the large computing time required for the SA algorithm.

5.3. Genetic optimization

These last experiments have been carried out with the hybrid genetic algorithm using the elitist strategy and the cooling temperature schedule described in subsection 4.3. Tests have shown that this optimization procedure was not very sensitive to the control parameters. In our application, these parameters are the following: population size=100, crossover rate=0.8, mutation rate=0.008, maximum number of generations=150. Some tests have shown that the genetic algorithm-based optimization procedure is not very sensitive to mutation parameter if it is contained within the range \([0.001, 0.01]\), as proposed by Goldberg [17]. At each generation, we select 5% of the best individuals for the hybridization with the local optimization technique, and the cooling schedule parameters
are $T_0=2$, and $a=0.99$.

Our GA takes about 20 – 120 generations to converge to the true solution. In fact, the convergence rate can vary significantly depending on the complexity of the objective function $c$ to be minimized (or the complexity of the input image). The optimization procedure takes between 10 seconds and 35 seconds on a 43P IBM workstation. In all processed cases, we obtain far more quickly the same good results as those provided by the simulated annealing (see Table 1).

Figure 10 (a real sonar image on which a synthetic shadow shape of a cylindrical object have been added) illustrates the best deformed template and the best 5% set of templates before the gradient ascent technique for successive iterations or generations of the genetic search. We can see that this method is efficient even if the cast shadow shape is partially occluded. Figure 11 shows the detection and the classification of a synthetic cast shadow shape associated to a spherical object added on a real sonar image.

Figures 12, 13, 14 show a few examples of classification results from our database. Geometric shape, i.e., manufactured objects are well detected such as the geometric part of a wreck (Fig. 12.e), a section of a pipe-line (Fig. 12.d), a trolley (Fig 12.f) and different cylindrical objects (Fig. 12.a 12.b 12.c). In the same way, spherical objects (Fig. 13) are well detected even if the cast shadow shape is partially occluded. Figure 14 shows several natural objects and the associated values of the objective function $c(\theta, x)$.

6. CONCLUSION

In this paper, we have developed a novel and robust algorithm to distinguish, from sonar images, man-made objects from natural objects lying on the sea-bed. The proposed method enables to distinguish natural objects from manufactured objects lying on the seabed and the whole paper is presented in this intention.\footnote{In a military context, this method may also be used to detect and to localize cylindrical or spherical mines lying on the sea-floor. It is our applicative context. Unfortunately, confidentiality reasons prevent us from providing too many informations on the application.} We have stated the detection
and the classification issues within a statistical setting, and we have shown that this problem can be handled as an equivalent energy minimization problem. The energy function is minimized using a stochastic genetic search combined with a local optimization technique and a cooling temperature schedule. This optimization procedure is fast, robust, simple and well suited to our application compared to other minimization techniques such as gradient-based method or the simulated annealing algorithm. The proposed scheme appears as an original alternative to feature-based matching methods, and remains very efficient in the case of complex cluttered backgrounds and occlusion phenomena between several object cast shadows. This method has been applied to a number of real sonar images; the obtained results demonstrate its efficiency and robustness. This scheme is fast and robust enough to allow an automatic processing of massive amounts of data. Besides, this method can be easily improved in order to retrieve several object cast shadows that are like the prototype template in the input sub-image. In that prospect, multiple object cast shadows looking like the same prototype template can be localized by using our classification method. Then, we remove the edge in $\Phi_{S'}$ of the detected cast shadow shapes and we repeat the same procedure till the value of $c$ is sufficiently large.
References


23


Tables and Figures

<table>
<thead>
<tr>
<th>Optimization</th>
<th>Image</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Gradient-based</td>
<td>$\epsilon = 0.47$</td>
<td>$\epsilon = 0.10$</td>
<td>$\epsilon = 0.31$</td>
</tr>
<tr>
<td>SA</td>
<td>$\epsilon = 0.19$</td>
<td>$\epsilon = 0.13$</td>
<td>$\epsilon = 0.13$</td>
</tr>
<tr>
<td>GA</td>
<td>$\epsilon = 0.17$</td>
<td>$\epsilon = 0.10$</td>
<td>$\epsilon = 0.15$</td>
</tr>
</tbody>
</table>

Table 1: Minima obtained with the different optimization procedures for different sonar image.

Figure 1: (a) Cast shadow and echo associated to a spherical object, synthesized with a ray tracing procedure. (b) Associated prototype template.
Figure 2: Considered linear transformations in $\mathbb{R}^2$ for the original prototype associated to the cast shadow of a manufactured object with a regular shape.

Figure 3: Considered linear transformations in $\mathbb{R}^2$ for the original prototype associated to the cast shadow of a spherical shape object.

Figure 4: Elementary potential field created by a site associated to an edge of a detected shadow region in $x$ ($\sigma = 1$).

**List of Tables**

1 Minima obtained with the different optimization procedures for different sonar image. .................................................. 25
Figure 5: (a) Sonar image (image called rock and object). (b) Associated Markovian two-class segmentation. (c) Edge potential field with $\sigma = 1$.

Figure 6: An example of crossover with the partial exchange of information.
Figure 7: Template-based detection and classification with a gradient-based local optimization procedure in case of cast shadow of a cylindrical object on a sandy sea bed (image called object on a sandy sea-bed) (a) and pebbly sea floor (image called object on a pebbly sea-bed) (b). The input two-class segmentation is first shown, then the resulting deformed template is displayed superimposed to original image. (a) Convergence towards the global minimum: $\epsilon = 0.10$. (b) Convergence to a local minima: $\epsilon = 0.31$. 
Figure 8: Evolution of the deformable template during the SA optimization on an image (object on a pebbly sea-bed) involving the cast shadows of a cylindrical object on a pebbly sea-floor. (a) Two-class Markovian segmentation of the sonar image presented in (b-i). (b) Initial position of the deformed template. (c-i) Successive iterations of the simulated annealing procedure. (i) Final segmentation. $\epsilon = 0.13$, leads to classify the selected region as the cast shadow of a cylindrical manufactured object.
Figure 9: Evolution of the deformable template during the SA optimization on an image (rock and object) involving the cast shadow of a rock and a cylindrical object. (a) Two-class Markovian segmentation of the sonar image presented in (b-i). (b) Initial position of the deformed template. (c-i) Successive iterations of the simulated annealing procedure. (i) Final segmentation, $\epsilon = 0.19$, leads to classify the selected region as the cast shadow of a cylindrical manufactured object.
Figure 10: Successive generations of the genetic search. This sonar image contains a synthetic shadow shape of a cylindrical object partially occluded by a dune of sand cast shadow. (a) Two-class Markovian segmentation of the sonar image (b-f) best 5% deformed templates before the gradient ascent (on the left) and the best resulting template (on the right). GA optimization after: (b) 1 (c) 2 (d) 4 (e) 16 (f) 30 generations. Accurate and reliable detection of the cylindrical object cast shadow is obtained ($\epsilon = 0.06$).
Figure 11: Successive generations of the genetic search. This sonar image shows the synthetic shadow shape of a spherical object occluded by several rock cast shadow. (a) Two-class Markovian segmentation of the sonar image (b-f) best 5% deformed templates before the gradient ascent (on the left) and the best resulting template (on the right). GA optimization after: (b) 1 (c) 4 (d) 6 (e) 8 (f) 24 generations. Accurate and reliable detection of the spherical object cast shadow is obtained (ε = 0.10).
Figure 12: The parallelogram template and the low value obtained (with the genetic search) for $\epsilon$ ($\epsilon < 0.2$) allow us to classify these shadows as “manufactured object cast shadow”. (a-b-c) Different cylindrical objects. $\epsilon = 0.17$, $\epsilon = 0.15$, and $\epsilon = 0.14$ respectively. (d) pipe-line: $\epsilon = 0.12$. (e) wreck: $\epsilon = 0.15$. (f) trolley: $\epsilon = 0.2$.

Figure 13: The cubic B-spline template and the low value obtained (with the genetic search) for $\epsilon$ ($\epsilon < 0.2$) allows us to classify these shadows as “spherical object cast shadow”. (a) $\epsilon = 0.04$. (b) $\epsilon = 0.15$. (c) $\epsilon = 0.19$. 

33
Figure 14: The values of $\epsilon$ (>0.2) (obtained with the genetic search) for both templates are sufficiently large so that we can reject the hypothesis of a manufactured or spherical object present in these images. (a-c) Markovian segmentation of the sonar images. (d) $\epsilon=0.40$. (e) $\epsilon=0.41$. (f) $\epsilon=0.24$. (g) $\epsilon=0.55$. (h) $\epsilon=0.57$. (i) $\epsilon=0.38$. 

34
List of Figures

1  (a) Cast shadow and echo associated to a spherical object, synthesized with a ray tracing procedure. (b) Associated prototype template. ........................ 25
2  Considered linear transformations in $\mathbb{R}^2$ for the original prototype associated to the cast shadow of a manufactured object with a regular shape. ........................ 26
3  Considered linear transformations in $\mathbb{R}^2$ for the original prototype associated to the cast shadow of a spherical shape object. ................................. 26
4  Elementary potential field created by a site associated to an edge of a detected shadow region in $x (\sigma = 1)$. ......................................................... 26
5  (a) Sonar image (image called rock and object). (b) Associated Markovian two-class segmentation. (c) Edge potential field with $\sigma = 1$. ............................ 27
6  An example of crossover with the partial exchange of information. ........ 27
7  Template-based detection and classification with a gradient-based local optimization procedure in case of cast shadow of a cylindrical object on a sandy sea bed (image called object on a sandy sea-bed) (a) and pebbly sea floor (image called object on a pebbly sea-bed) (b). The input two-class segmentation is first shown, then the resulting deformed template is displayed superimposed to original image. (a) Convergence towards the global minimum: $\epsilon = 0.10$. (b) Convergence to a local minima: $\epsilon = 0.31$. ........ 28
8  Evolution of the deformable template during the SA optimization on an image (object on a pebbly sea-bed) involving the cast shadows of a cylindrical object on a pebbly sea-floor. (a) Two-class Markovian segmentation of the sonar image presented in (b-i). (b) Initial position of the deformed template. (c-i) Successive iterations of the simulated annealing procedure. (i) Final segmentation. $\epsilon = 0.13$, leads to classify the selected region as the cast shadow of a cylindrical manufactured object. ............................. 29
Evolution of the deformable template during the SA optimization on an image (rock and object) involving the cast shadow of a rock and a cylindrical object. (a) Two-class Markovian segmentation of the sonar image presented in (b-i). (b) Initial position of the deformed template. (c-i) Successive iterations of the simulated annealing procedure. (i) Final segmentation, $\epsilon = 0.19$, leads to classify the selected region as the cast shadow of a cylindrical manufactured object. .................................................. 30

Successive generations of the genetic search. This sonar image contains a synthetic shadow shape of a cylindrical object partially occluded by a dune of sand cast shadow. (a) Two-class Markovian segmentation of the sonar image (b-f) best 5 % deformed templates before the gradient ascent (on the left) and the best resulting template (on the right). GA optimization after: (b) 1 (c) 2 (d) 4 (e) 16 (f) 30 generations. Accurate and reliable detection of the cylindrical object cast shadow is obtained ($\epsilon = 0.06$). ... 31

Successive generations of the genetic search. This sonar image shows the synthetic shadow shape of a spherical object occluded by several rock cast shadow. (a) Two-class Markovian segmentation of the sonar image (b-f) best 5 % deformed templates before the gradient ascent (on the left) and the best resulting template (on the right). GA optimization after: (b) 1 (c) 4 (d) 6 (e) 8 (f) 24 generations. Accurate and reliable detection of the spherical object cast shadow is obtained ($\epsilon = 0.10$). ......................... 32

The parallelogram template and the low value obtained (with the genetic search) for $\epsilon$ ($\epsilon < 0.2$) allow us to classify these shadows as “manufactured object cast shadow”. (a-b-c) Different cylindrical objects. $\epsilon = 0.17$, $\epsilon = 0.15$, and $\epsilon = 0.14$ respectively. (d) pipe-line: $\epsilon = 0.12$. (e) wreck: $\epsilon = 0.15$. (f) trolley: $\epsilon = 0.2$. ................................................................. 33
The cubic B-spline template and the low value obtained (with the genetic search) for $\epsilon$ ($\epsilon < 0.2$) allows us to classify these shadows as “spherical object cast shadow”. (a) $\epsilon = 0.04$. (b) $\epsilon = 0.15$. (c) $\epsilon = 0.19$. 

The values of $\epsilon$ (>0.2) (obtained with the genetic search) for both templates are sufficiently large so that we can reject the hypothesis of a manufactured or spherical object present in these images. (a-c) Markovian segmentation of the sonar images. (d) $\epsilon=0.40$. (e) $\epsilon=0.41$. (f) $\epsilon=0.24$. (g) $\epsilon=0.55$. (h) $\epsilon=0.57$. (i) $\epsilon=0.38$. 

33 34
Max MIGNOTTE received the D.E.A (Postgraduate degree) in Digital Signal, Image and Speech processing of the Institut National Polytechnique of Grenoble (1993) (France) and the Ph.D. degree in Digital signal and image processing from the University of Bretagne Occidentale (France) (1998) (thesis achieved at the laboratory GTS (Groupe de Traitement du Signal) of the French Naval Academy). He is currently an INRIA post-doctoral fellow at Montreal University (DIRO), Canada, Quebec. His current research interests include Bayesian inference, image segmentation (hierarchical markovian models or deformable models), parameters estimation, tracking, classification, deconvolution/restoration and optimization problem in sonar or medical images.
Christophe Collet graduated from the Université Paris-Sud Orsay in 1989 (DEA in Signal Processing) and received the Ph.D. in image analysis from the University of Toulon in 1992. His thesis research was carried out at the Ecole Nationale Supérieure des Télécom. de Bretagne (ENST-Br) and at the French Naval Academy. He is currently "maître de conférences" at the French Naval Academy and chairman of the laboratory GTS "Groupe de Traitement du Signal" since 1994. The line of research of the GTS laboratory concerns signal and image processing for underwater acoustic applications. His main research interests include image modeling, markovian segmentation, bayesian estimation, pattern recognition, texture analysis, statistical optimization technique for image analysis. The main applications deduced from this research framework deal with motion segmentation of infrared image sequences, segmentation of seafloor for man-made pattern classification, multispectral segmentation of SPOT images.

Patrick Pérez was born in 1968. He graduated from “Ecole Centrale Paris”, France, in 1990. He received the Ph.D. degree in Signal Processing and Telecom. from the University of Rennes, France, in 1993. He spent one year as an INRIA post-doctoral fellow at Brown University (Department of Applied Mathematics), USA. He now holds a full-time research position at the INRIA center in Rennes. His research interests include statistical and/or hierarchical models for large inverse problems in image analysis, with a particular focus on motion analysis issues (detection, segmentation, and estimation of the apparent motion, tracking, analysis of elastic and fluid motions, etc.)
Patrick Bouthemy graduated from Ecole Nationale Superieure des Telecommunications, Paris, in 1980, and received the Ph.D degree in Computer Science from the University of Rennes, France, in 1982. From December 1982 until February 1984, he was employed by INRS-Telecommunications, Montreal, P.Q., Canada, in the Department of Visual Communications. Since April 1984, he has been with INRIA, at IRISA in Rennes. He is currently “Directeur de Recherche” Inria and head of Vista project. His major research interests are concerned with image sequence analysis: 2D motion analysis (rigid motion and deformable motion), statistical models for image sequence processing (MRF models, robust estimation, Bayesian estimation), dynamic scene interpretation, active vision, video indexing. He is Associate Editor of the IEEE Transactions on Image Processing.