STATISTICAL MODEL AND GENETIC OPTIMIZATION: APPLICATION TO PATTERN DETECTION IN SONAR IMAGES

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

‡ Groupe de Traitement du Signal, Ecole Navale, Lanvéoc-Poulmic, 29240 Brest-Naval France.
email: name@poseidon.ecole-navale.fr

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

ABSTRACT

We present a new classification method using deformable template model to separate natural objects from man made objects in an image given by a high resolution sonar. A prior knowledge of the manufactured object shadow shape is described by a prototype template and a set of admissible linear transformations to take into account the shape variability. Then, the classification problem is defined as a two step process; firstly the detection problem of a region of interest in the input image is stated in a Bayesian framework and is posed as an equivalent energy minimization problem of an objective function: in this paper, this energy minimization problem is solved by using a hybrid *Genetic Algorithm* (GA). Secondly, the value of this function at convergence allows to determine the presence of the desired object in the sonar image. This method has been successfully tested on real and synthetic sonar images[1], yielding very promissing results.

1. INTRODUCTION

Due to their high resolution, existing sonars allow to detect every object lying on the seabed. Each object can be identified efficiently thanks to an analysis of its associated cast shadow (due to a lack of acoustic reverberation behind the object). However, as data volume has noticeably increased, the exploitation of the collected data has now to be achieved with an automatic processing chain.

Contrary to 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 approach in image analysis has been quite popular and has been sucessfully applied for image segmentation [1] [2] (with a local prior model) or shape matching with deformable template-based methods (with a global prior model). Laksmanan *et al.* [3], for example, have used a parametric template model to locate the road boundary in radar images. Then, the edge detection problem is formulated as a Bayesian estimate using a physics-based model of the radar imaging process. A similar approach for shape matching is proposed by Jain *et al.* [4] 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 to separate 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 object class shape variability 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 in a Bayesian framework and is reduced to the estimation of the deformation parameters of the template that maximize the posterior PDF. This maximization problem is posed as an equivalent energy minimization problem of an objective function.

In [4], gradient based methods are used for energy minimization of this function. These methods have the disadvantage to require good initial parameter estimates, otherwise they will converge toward a local *minima*. Stochastic methods based on *Simulated Annealing* [5] [6] have the capability of avoiding local *minima* and no human interaction is required to initialize the model. But one of the major drawbacks of this procedure is its high computational requirements. Hereafter, we show that an alternate approach consists in using a genetic exploration of the search space.

The main contribution of this paper lies in the use of deformable model and the Bayesian framework to classify objects in sonar imagery. We propose an appropriate energy term that differs from previously published works. This energy term uses the informations given by an unsupervised Markovian segmentation of the input sonar image [2] and integrates both region homogeneity and edge information. Finally, we propose 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 search space combined with a steepest ascent procedure and a cooling temperature schedule.

This paper is organized as follows: §2 and §3 describe the representation of the template and the Bayesian approach for the deformable template matching and the classification problem. The optimization problem using GA is described in §4. In §5, we show some detection/classification results on real and synthetic sonar images.

[1] Acknowledgements: The authors thank **GESMA** (Groupe d'Étude Sous Marine de l'Atlantique), Brest, for having provided numerous real sonar pictures and **DRET** (Direction des Recherches, Etudes et Technique, French Ministry of Defense) for partial financial support of this work (student grant).

2. TEMPLATE REPRESENTATION

This contour based model is appropriate for general shape matching since the same approach can be applied to objects of differents shapes by defining different prototype templates.

Contrary to natural objects, a manufactured object is *a priori* mainly composed of elements with simple geometrical shape. For this reason, their cast shadow is rather regular and shows straight lines and edges. It is the case for wreck, pipe-line, *etc* ... In particular, the cast shadow of a cylindrical or cubic object is a perfect parallelogram. So, we define the corresponding prototype template as a parallelogram representation characterized by its 4 vertices. Fig. 2 illustrates the representation.

For military reasons, 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 spaced which approximate its outline. A cubic B-spline shape representation with these n control points corresponding to the "*landmarks*" is then defined. The outline of the spherical object cast shadow on which we can select this set of n points can be given by a real scene or a synthetic representation of spherical object shadow (Fig. 1).

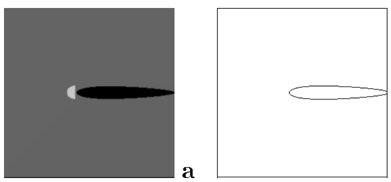

Figure 1: (**a**) *Synthetized shadow and echo of a spherical object given by a ray tracing procedure.* (**b**) *Associated prototype template.*

The prototype template γ_0 describes only one of the possible instances of the shapes to be detected for a class of object. In order to take into account the variability of the considered object class, we introduce a set of admissible linear transformations on γ_0. Let γ_θ be a deformed version of the original prototype according to affine transformation with parameters θ. In the case of our first rigid template (used to detect manufactured object with simple geometrical shape), these deformations involve translation, scaling, rotation, stretching and skewing of the template as is shown in Fig.2.

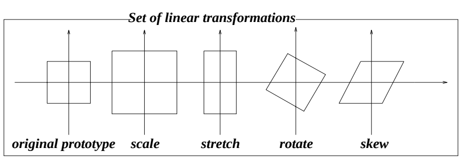

Figure 2: *Considered linear transformations in* $I\!R^2$ *for the original prototype associated to the cast shadow of a simple geometrical shape object.*

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

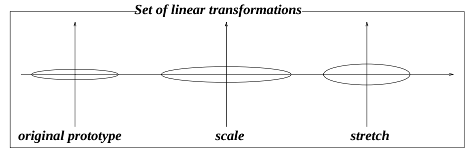

Figure 3: *Considered linear transformations in* $I\!R^2$ *for the original prototype associated to the cast shadow of a spherical shape object.*

3. MAP DETECTION

A common problem in sonar images are artefacts caused by the speckle noise effect which lead to a loss of signal and a very poor quality of the object boundaries [7]. That is why the joint model we propose does not use directly the input image, *i.e.*, the grey-levels themselves or some gradient measure on it in order to detect and use the contour of each object. In our approach, we use the result of an unsupervised two-class (*shadow/reverberation* areas) Markovian segmentation of the input sonar image [2]. This allows us to take into account the observed measurement along the contour but also the grey-level homogeneity information inside and outside the contour. The detection is based on an objective function ϵ measuring how well a given instance of deformed template γ_θ fits the content of segmented image x. From a probabilistic point of view, $\epsilon(\theta, x)$ defines the joint model through the Gibbs Distribution:

$$P_{\Theta,X}(\theta, x) = \frac{1}{Z} \exp\left\{-\epsilon(\theta, x)\right\} \tag{1}$$

where Θ is the random vector of parameters, and Z a normalizing constant.

3.1. Joint model

The posterior distribution deduced from (1):

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

provides the probability of a given template given the segmented image. The joint model $\epsilon(\theta, x)$ specifies the probability of observing the input image, given a deformed template at a specific position, orientation, sketch, stretch and scale. It is 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:

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

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

where r ($r \neq 0$) is the distance to the pixel s_0. The different edges in x' create a potential field $\Phi_{x'}(t)$, given by the total sum of the different elementary potential fields $\phi_{s_0}(r)$:

$$\Phi_{x'}(t) = \inf\Big\{ \sum_{s \in x'} \phi_s\big(d(s,t)\big), 1 \Big\} \qquad \forall \text{ pixel} \tag{4}$$

with $d(s,t)$ is the distance between the pixels s and t. In fact, this potential field induces a smooth version of the edge image x' in which a site close to an edge will get a potential value close to 1 (Fig. 4). The degree of this smoothness can be controled by the parameter σ. This edge potential field induces an energy function that relates a deformed template γ_θ to the edges created by each detected object in the input image:

$$\epsilon_c(\theta, x') = -\ln\Big\{ \frac{1}{N_{\gamma_\theta}} \sum_{s \in \gamma_\theta} \Phi_{x'}(s) \Big\} \tag{5}$$

where the summation is over all the (N_{γ_θ}) pixels on the deformed template γ_θ and the *ln* function is used to increase the dynamics of the energy function ϵ_c.

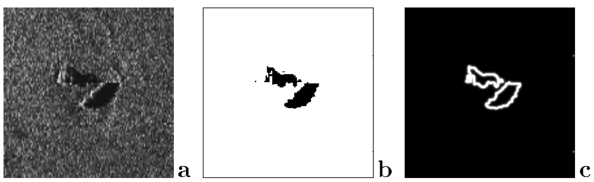

Figure 4: (**a**) *Sonar image.* (**b**) *Associated Markovian two-class segmentation.* (**c**) *Edge potential field with* $\sigma = 1$.

Region homogeneity energy: This energy function aims to place the inside of the deformed template in a region classified *shadow* by the segmentation procedure.

$$\epsilon_r(\theta, x) = -\ln\Big\{ \frac{1}{N_{\gamma_\theta^\bullet}} \sum_{s \in \gamma_\theta^\bullet} \delta(x_s - e_0) \Big\} \tag{6}$$

where $\gamma_\theta^\bullet$ and $N_{\gamma_\theta^\bullet}$ represent the set of pixels and the number of pixels inside the contour respectively and δ is the Kronecker delta function.

Using these two energy functions, the posterior distribution of θ given x is:

$$P_{\Theta/X}(\theta/x) = \frac{1}{Z_x} \exp -\Big\{ \underbrace{\epsilon_c(\theta, x') + \epsilon_r(\theta, x)}_{\epsilon(\theta,x)} \Big\} \tag{7}$$

where Z_x is a normalizing constant depending on x only.

3.2. Detection step

We formulate the detection problem as the *Maximum A Posteriori* (MAP) estimation of θ:

$$\begin{aligned} \hat{\theta}_{MAP} &= \arg\max_\theta \big\{ P_{\Theta/X}(\theta/x) \big\} & (8) \\ &= \arg\min_\theta \epsilon(\theta, x) & (9) \end{aligned}$$

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 extracted in x' and then to determine the presence of the desired object. If $\epsilon(\hat{\theta}_{MAP}, x)$ is lower than a given treshold, then the desired object is assumed to be present and the final configuration of the deformed template indicates shape and location of the detected object; otherwise we decide that the desired object is not present.

4. GENETIC OPTIMIZATION

The objective function to be maximized in Eq. 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 a GA-based global optimization technique [8]. We can easily derive a fitness measure $\mathcal{F}$ (to be maximized) directly from Eq. 7 for use in genetic algorithm (*i.e.*, one with range [0,1]):

$$\mathcal{F}(\theta) = \exp\big\{ -\epsilon(\theta, x) \big\} \tag{10}$$

In order to prevent premature convergence [8] and to speed up the convergence rate, we have developed 3 strategies and have combined them:

1.– The first one is an elite-preservation strategy [8]: the individual with the highest fitness always survives to be an individual of the next generation.

2.– The second strategy (called hybrid GA [8]) 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. The fitness function $\mathcal{F}$ at iteration k is defined as follows:

$$\mathcal{F}^k(\theta) = \exp\big\{ -\frac{1}{T_k} \epsilon(\theta, x) \big\} \tag{11}$$

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 uses a smooth version of the energy function ϵ. 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 to sub-optimal solution. For $T_k = 1$ the genetic search is carried out with the real cost function ϵ. At the end of the procedure, $T_k < 1$, the fitness measure falls off rapidly with increasing cost. This allows to maintain a good competition between individuals located near the optimum global and to localize precisely the global *extrema*.

5. EXPERIMENTAL RESULTS

This experiments have been carried out with the template-based detection, the classification scheme, and the genetic optimization described in §4. 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. At each generation, we select 5 % of the best individuals for the hybridation with the local optimization technique and the cooling schedule parameters are: T_0=2, and a=0.99. The optimization procedure takes between 10 and 35 sec (on a IBM P200 workstation) depending on the complexity of $\epsilon(\theta, x)$.

Fig. 5,6,7 show a few examples of classification results from our database. Geometric-shape, *i.e.*, manufactured objects are well detected (Fig. 5) as well as spherical objects (Fig. 6), and this method is efficient even if the cast shadow shape is partially occluded. Fig. 7 shows several natural objects and the associated values of the objective function $\epsilon(\theta, x)$.

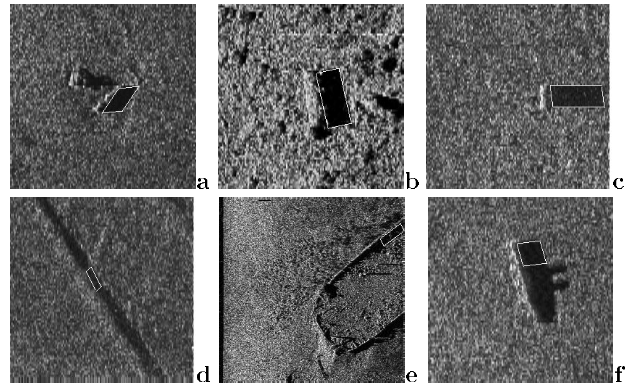

Figure 5: *The used template and the low value obtained for ϵ ($\epsilon<0.2$) allow to classify these shadows as "manufactured object cast shadow."* (**a**-**b**-**c**) *Cylindrical object. ϵ=0.17, ϵ=0.15, and ϵ=0.14 respectively.* (**d**) *pipe-line: ϵ=0.12.* (**e**) *wreck: ϵ=0.15.* (**f**) *trolley: ϵ=0.2. (parallelogram template).*

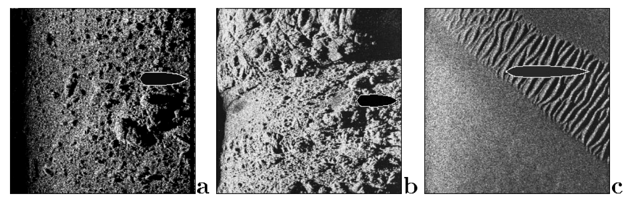

Figure 6: $\epsilon < 0.2$ *class: "spherical object cast shadow."* (**a**) $\epsilon = 0.04$. (**b**) $\epsilon = 0.15$. (**c**) $\epsilon = 0.19$. *(cubic B-spline template).*

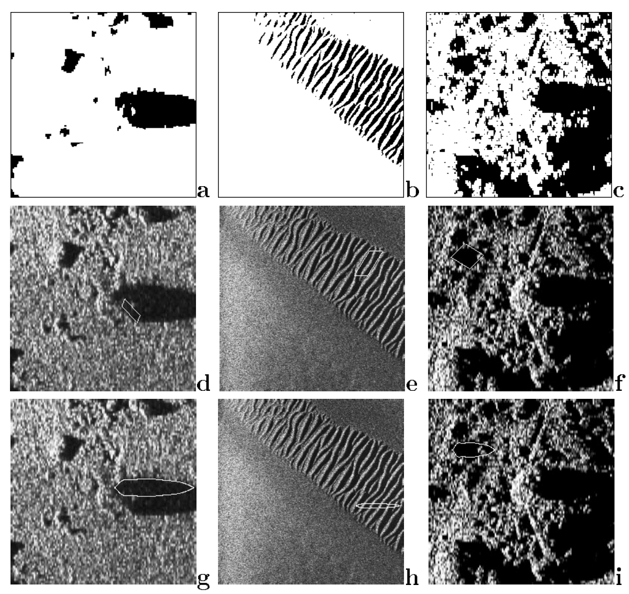

Figure 7: *The value of ϵ (>0.2) is sufficiently large so that we can reject the hypothesis of a manufactured or spherical object present in these images.* [**a**-**c**] *Markovian segmentation of the sonar image presented in* [**d**-**f**]. (**d**) *ϵ=0.40.* (**e**) *ϵ=0.41.* (**f**) *ϵ=0.24.* (**g**) *ϵ=0.55.* (**h**) *ϵ=0.57.* (**i**) *ϵ=0.38.*

6. CONCLUSION

In this paper, we have developped a novel and robust algorithm to distinguish man made objects from natural objects lying on the seabed in sonar images. We have posed the detection and classification problem in a Bayesian setting. This proposed sheme appears as an interesting alternative to feature-based matching methods and remains very efficient in the case of complex background and occlusions of several object cast shadows. This method has been validated on a number of real sonar images demonstrating the efficiency and robustness of this scheme and is compatible with an automatic processing of massive amounts of data.

7. REFERENCES

[1] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, B-48:259–302, 1986.

[2] M. Mignotte, C. Collet, P. Pérez, and P. Bouthemy. Unsupervised Markovian segmentation of sonar images. In *Proc. ICASSP*, volume 4, pages 2781–2785, Munchen, May 1997.

[3] S. Lakshmanan and D. Grimmer. A Deformable Template Approach to Detecting Straight edges in Radar Images. *IEEE Trans. on PAMI*, 18(4):438–443, April 1996.

[4] A.K. Jain, Y. Zhong, and Sridhar Lakshmanan. Object Matching Using Deformable Templates. *IEEE Trans. on PAMI*, 18(3):267–278, March 1996.

[5] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Trans. on PAMI*, 6(6):721–741, November 1984.

[6] C. Kervrann and F. Heitz. Statistical Model-Based Segmentation of Deformable Motion. In *Proc. ICIP*, pages 937–940, Lausanne, 1996.

[7] J. W. Goodman. Some fundamental properties of speckle. *Journal of Optical Society of America*, 66(11):1145–1150, November 1976.

[8] D.E Goldberg. *Genetic Algorithm.* Addison Wesley, 1989.