UNSUPERVISED HIERARCHICAL MARKOVIAN SEGMENTATION OF 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

This paper is concerned with hierarchical Markov Random Field (MRF) models and with their application to sonar image segmentation. We present a novel unsupervised hierarchical MRF model involving a pyramidal label field and a scale-causal and spatial neighborhood structure. This allows us to more precisely model the local and global characteristics of image content for different scales. Such connections lead to efficiently propagate interactions and this approach seems to be well suited for the segmentation of very noisy sonar images. The MRF prior model parameters are estimated in an accurate and fast way simultaneously to the segmentation process by generalizing the estimation method proposed by Derin *et al.* Experiments with real images indicate that the proposed SCM algorithm (Scale Causal Multigrid) performs better than other hierarchical schemes for sonar image segmentation.

1. INTRODUCTION

In sonar imagery, a low-level step of segmentation in two classes, *i.e. shadow* (corresponding to a lack of acoustic reverberation) and *sea-bottom reverberation*, is necessary to detect and then classify elements located on the sea floor as wrecks, ridges, pebbles, man-made objects ... In this application, contextual information is important to be taken into account to face speckle noise. MRF are appropriate models to specify spatial dependency by means of *a priori* label field distribution [1][2]. Nevertheless, purely spatial MRF models have a limited ability to describe properties on large scale, and may be not sufficient to ensure the regularization process of the set of labels when the sonar image contains strong speckle noise. Such a model can be improved by using a larger neighborhood for each pixel, but this rapidly increases the complexity of the segmentation algorithms and the parameter estimation procedure.

To circumvent this difficulty, a multiresolution scheme has been proposed [1]. A multiresolution image transform is then used to decompose the image data over a pyramidal structure, and reduced similar MRF models are then defined at each resolution level. An other way to ensure an efficient segmentation is to use a hierarchical MRF model. Several approaches to multigrid image modeling have been proposed. Bouman *et al.* use a multiscale MRF model where each scale is causally dependent on the preceding coarser grid field [3]. On one hand, this scheme is not iterative; on the other hand, the neighborhood structure does not completely capture some aspects of local image properties, since a disadvantage of this quadtree model is that spatially adjacent pixels may not have common neighbors at the next coarser scale.

In [4], Kato *et al.* proposed a novel hierarchical model: they introduce a pyramidal structure with a three dimensional (3D) neighborhood system and a 3D label field. However, the system uses a first order neighborhood structure with two cliques and an isotropic model. This model may be well suited for SAR images, but too simple to be used for our application. Scenes are very different in our case: sea floor with pebbles, dunes, ridges, sand, tires or various objects ... Moreover, the segmentation and the estimation with such *a priori* model, using simulated annealing, require a lot of computing time.

Herein, we propose an unsupervised hierarchical segmentation algorithm using an appropriate multiscale model. This model involves a local connection between spatially adjacent sites, and to the father site belonging to the pixel grid at the lower resolution level. We have also developed a computationally efficient method for the unsupervised estimation of the model parameters. This scheme takes into account the variety of the laws in the distribution mixture of a sonar image, estimates the parameters of each noise distribution, and the parameters of the *a priori* model. We use a maximum likelihood estimation technique concerning the noise model parameters. In order to estimate the MRF prior model, our approach consists in extending the method proposed by Derin *et al.* [5]. The MRF prior parameters are estimated jointly to the segmentation process and are optimal in the least squares sense. Finally, we compare the results obtained on real sonar images with respectively, our scheme, the multigrid model defined in [6], the SMAP algorithm [3], and a multiresolution approach [1] .

This paper is organized as follows. In Section 2, we define the notation and the proposed hierarchical model. Section 3 details the parameter *estimation step*, and more precisely, the MRF prior model estimator. Experimental results on real scenes are presented in Section 4.

[1] Acknowledgements: The authors thank **GESMA**, 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. SONAR IMAGERY AND HIERARCHICAL MARKOVIAN MODELING

We propose a multiscale model called SCM (Scale Causal Multigrid) which consists of a label field pyramid associated to a single observation level. We consider a scale-causal and spatial neighborhood system. A site s interacts with spatially adjacent sites at a given scale and with its *father* at the upper coarser level (*i.e* a causal neighborhood in scale). Figure 1. depicts the pyramidal structure and the used neighborhood system. The base of our approach is the multiscale model introduced in [6]. The observation field remains at the finest resolution, only the MRF model will be hierarchically defined. Starting from an original full resolution energy model, the energy function is rewritten at each scale as a coarser MRF model.

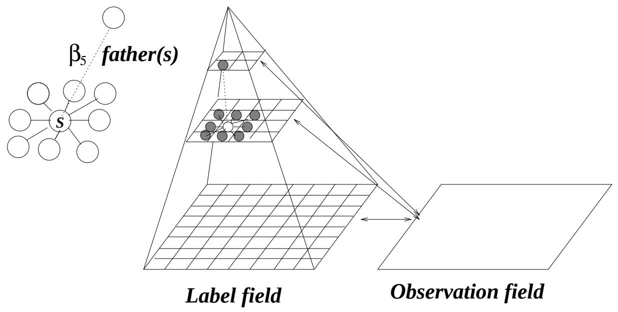

Figure 1: *Neighborhood system and the multiscale structure*

2.1. Monoresolution model

Let $Z = (X, Y)$ be a couple of random fields, with $Y = \{Y_s, s \in S\}$ the field of observations located on a lattice S of N sites s, and $X = \{X_s, s \in S\}$ the label field. Each of the Y_s takes its value in $\Lambda_{obs} = \{0, \ldots, 255\}$ and each X_s in $\{e_0 = \textit{shadow},\ e_1 = \textit{sea bottom reverberation}\}$. The distribution of (X, Y) is defined, firstly, by $P_X(x)$, the distribution of X supposed to be stationary and Markovian, and secondly, by the site-wise likelihoods $P_{Y_s/X_s}(y_s/x_s)$. In this work, these conditional likelihoods depend on the class label x_s. $\mathbf{\Phi}_x$ and $\mathbf{\Phi}_y$ are the parameter vectors associated to the distributions $P_X(x)$ and $P_{Y_s/X_s}(y_s/x_s)$ respectively. The corresponding posterior energy is:

$$U(x,y) = \sum_{<s,t>} \beta_{st}[1 - \delta(x_s, x_t)] + \sum_s \Psi_s(x_s, y_s)$$

where we adopt a 8-connexity spatial neighbourhood and $\beta_{st} = \beta_1, \beta_2, \beta_3$ or β_4 if clique $<s,t>$ is horizontal, vertical, right and left diagonal respectively. $\Psi_s = -\ln P_{Y_s/X_s}(y_s/x_s)$ and δ is the Kronecker delta function.

2.2. Scale causal multigrid model

We now deal with a hierarchy $(x^L, \ldots, x^0)$ of label field, where x^l is defined on grid S^l reduced by 2^l in each direction. We first define:

$$P_{X^0/Y,X^1}(x^0/y, x^1) \propto \exp\{-U^0(x^0, y, x^1)\}$$

with $U^0(x^0, y, x^1) = U(x,y) + \sum_{s \in S} \beta_5[1 - \delta(x^0_s, x^1_{father(s)})]$

where β_5 stands for the inter-level clique parameter. In our approach, the segmentation of sonar images in two classes is stated as a causal statistical labelling problem according to a global Bayesian formulation in which we search for x^l, the labelling at the resolution level l, such as:

$$l = L \ldots 0 \quad \hat{x}^l = \arg\max_{x^l} P_{X^l/Y,\hat{X}^{l+1}}(x^l/y, \hat{x}^{l+1}) \qquad (1)$$

Now let $\Phi^{l,k}$ $(k < l)$ be the "projection" from S^l to S^k, which associates by duplication a blockwise constant configuration on S^k to any configuration on S^l. For $l < L$ we can define: $P(x^l/y, x^{l+1}) \propto \exp\{-U^l(x^l, y, x^{l+1})\}$ with $U^l(x^l, y, x^{l+1}) = U^0(\Phi^{l,0}(x^l), y, \Phi^{l+1,1}(x^{l+1}))$. Simple computations yield [6]:

$$\begin{aligned} U^l(x^l, y, x^{l+1}) &= \underbrace{\sum_{s \in S^l} \Psi^l_s(x^l_s, y_{b^l_s})}_{U^l_1(x^l, y)} \\ &+ \underbrace{\sum_{<s,t>} \alpha^l_{st}\beta_{st} \cdot \left[1 - \delta(x^l_s, x^l_t)\right]}_{U^l_2(x^l)} \\ &+ \underbrace{\sum_{s \in S^l} \alpha^l_5 \beta_5 \cdot \left[1 - \delta(x^l_s, x^{l+1}_{father(s)})\right]}_{U^l_3(x^l, \hat{x}^{l+1})} \end{aligned}$$

$$\text{with} \quad \Psi^l_s(x^l_s, y_{b^l_s}) = \sum_{p \in b^l_s} \Psi_s(x^l_s, y_p)$$

$$\text{and} \quad \alpha^l_{st} = \{\alpha^l_1, \alpha^l_2, \alpha^l_3, \alpha^l_4\}$$

where b^l_s is the $2^l \times 2^l$ block of S corresponding to $p \in S^l$. U^l_1 expresses the adequacy between observations and labels, U^l_2 the energy of the *a priori* model and U^l_3 the energy term expressing the relation with the segmentation at the resolution level $l+1$. $y_{b^l_s}$ stands for the set of all the observations of the block b^l_s associated to site s. The parameters of the Gibbs distribution, $\beta^l_i = \alpha^l_i \beta_i$, are derived over scale in a coherent mathematical way [6]. We easily obtain for our application:

$$\begin{aligned} \alpha^l_i &= \left[2^l + 2(2^l - 1)\right] & i \in \{1, 2\} \\ \alpha^l_i &= 1 & i \in \{3, 4\} \\ \alpha^l_5 &= 4^l \end{aligned}$$

At coarsest level, we define:

$$P_{X^L/Y}(x^L/y) \propto \exp\{-U^L(x^L, y)\}$$

$$\text{with} \quad U^L(x^L, y) = U\big(\Phi^{L,0}(x^L, y)\big)$$

We now have to deal with the coarse-to-fine recursive estimation (1). In order to minimize the energy function associated with the MRF model at each scale, we use the ICM algorithm [7] to get a faster and better estimate of $\hat{x}^l$ $(l < L)$. The final estimate $\hat{x}^{l+1}$ obtained at a given level is interpolated as $\Phi^{l+1,l}(\hat{x}^{l+1})$ to be used as an initialization for the relaxation process at the next finer level (cf. Fig.2).

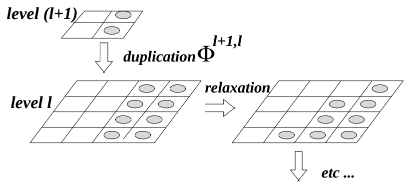

Figure 2: *'coarse-to-fine' minimization strategy*

3. PARAMETER ESTIMATION

Segmenting an image into different classes in a complete unsupervised way is not an easy task. The main difficulty is that the parameter estimation is required for the segmentation, while the segmentation is needed for the parameter estimation. In sonar imagery, no explicit information about the exact properties of the label field X is available since scenes are variable from an image to another. Besides, the noise on the pictures may vary. It is however possible to determine the noise model and the spatial MRF prior model parameters directly from the image data with the procedure presented in [8]. This estimation scheme takes into account the variety of the laws in the distribution mixture of sonar images, and estimates the parameters of each noise distribution. A Gaussian law $\mathcal{N}(\mu, \sigma)$ describes the luminance y within *shadow* regions, and a shifted Rayleigh law $\mathcal{R}(min, \alpha)$ takes into account the speckle noise phenomenon of the *sea bottom reverberation* class [8]. This *Estimation Step* allows us to obtain a preliminary multigrid segmentation [6] from which the SCM model can be estimated. This means that the segmentation process requires two full passes composed of the following steps:

1. Perform an *Estimation Step* (to estimate the noise model parameters $\mathbf{\Phi}_y$ and the spatial MRF prior model parameters with the procedure described in [8]) and compute a first *coarse-to-fine* multigrid segmentation [6] using the estimated parameters.
2. Estimate the SCM model $\mathbf{\Phi}_x$ and re-estimate $\mathbf{\Phi}_y$ based on the multigrid segmentation results.
3. Perform a final *"coarse-to-fine"* segmentation using the estimated SCM model and the strategy described in §2.

The parameter vector $\mathbf{\Phi}_x = [\beta_1, \beta_2, \beta_3, \beta_4, \beta_5]$ which defines the **SCM** model is estimated by generalizing the method of estimation proposed by Derin *et al.* [5]. For the spatial model defined in §2.1 (with $\Phi_x = [\beta_1 \ldots \beta_4]$), the method is: let η_i represent a possible label configuration of a generic neighborhood, X_{ν_s} the set of labels assigned to the neighbors of site s, and $P_{X_s, X_{\nu_s}}(e_j, \eta_i)$ the joint distribution of the label $X_s = e_j$ with the neighborhood $X_{\eta_s} = \eta_i$. Assuming they know some segmentation x, Derin *et al.* consider the following overdetermined linear system:

$$\forall \eta_i \;\; [\mathbf{\Theta}(e_0, \eta_i) - \mathbf{\Theta}(e_1, \eta_i)]^t \mathbf{\Phi}_x = \ln \left[\frac{P_{X_s, X_{\nu_s}}(e_1, \eta_i)}{P_{X_s, X_{\nu_s}}(e_0, \eta_i)} \right] \quad (2)$$

where $\mathbf{\Theta}(x_s, \eta_i)^t \; \mathbf{\Phi}_x = \sum_{t \in \nu_s} \beta_{st}[1 - \delta(x_s, x_t)]$. The ratio $P(e_1, \eta_i) \; / P(e_0, \eta_i)$ can be estimated using simple histogramming (by counting the number of 3×3 blocks of type (e_1, η_i) and dividing by the number of blocks of type (e_0, η_i) over the label field). By substituting for each value of η_i in (2), we obtain 256 equations (2^8 possible neighborhood configurations) in four unknowns. We obtain the parameter values by applying the least squares error method to solve this system. The extension of this method to estimate the prior model parameters in Fig.1 is straightforward. The unknown parameter vector to be estimated $\mathbf{\Phi}_x$ can be found by expressing (2) for each scale of the pyramidal structure (cf. Fig. 1). $\forall \eta_i$:

$$\left\{ \begin{array}{rcl} [\mathbf{\Theta}(e_0, \eta_i) - \mathbf{\Theta}(e_1, \eta_i)]^t \underline{\mathbf{M}}^0 \mathbf{\Phi}_x & = \ln \left[\frac{P^0_{X_s, X_{\nu_s}}(e_1, \eta_i)}{P^0_{X_s, X_{\nu_s}}(e_0, \eta_i)} \right] \\ \vdots & \vdots & \vdots \\ [\mathbf{\Theta}(e_0, \eta_i) - \mathbf{\Theta}(e_1, \eta_i)]^t \underline{\mathbf{M}}^{L-1} \mathbf{\Phi}_x & = \ln \left[\frac{P^{L-1}_{X_s, X_{\nu_s}}(e_1, \eta_i)}{P^{L-1}_{X_s, X_{\nu_s}}(e_0, \eta_i)} \right] \end{array} \right.$$

Where $\underline{\mathbf{M}}^l$ is a diagonal matrix defined as:

$$\underline{\mathbf{M}}^l = \text{diag} \; (\alpha_1^l, \alpha_2^l, \ldots, \alpha_5^l)$$

$\underline{\mathbf{M}}^l \cdot \mathbf{\Phi}_x$ stands for the clique potentiel vector of the Gibbs distribution at resolution level l. The ratio $P^l(e_1, \eta_i)$ $/ P^l(e_0, \eta_i)$ of each expression in this system can be estimated by counting the number of blocks of type (e_1, η_i) and dividing by the number of blocks of type (e_0, η_i) over the label field at resolution level l. For each expression in this system, we obtain $512 = 2^9$ equations by considering all possible neighborhood configurations η_i. A special combination (e_j, η_i) may not occur in the label fields. In this case, we cannot obtain a linear equation because of the logarithm. Moreover, $\mathbf{\Theta}(e_0, \eta_i) = \mathbf{\Theta}(e_1, \eta_i)$ implies equations of type $0 = C^{st}$. Therefore we ignore these cases as well. This overdetermined linear system of equations is solved with the least squares method.

4. EXPERIMENTAL RESULTS

In this section, we compare the segmentation performance of the proposed SCM algorithm with the multigrid strategy described in [6], the SMAP algorithm on a quadtree proposed by Bouman *et al.* [3], and a classical multiresolution approach (in our case, the data pyramid is built with a low pass filter). All the segmentation results exploit the *Parameter Estimation Step* presented in [8]. This *Estimation Step* is used to estimate the MRF prior model for the multigrid strategy and for the first step of the proposed SCM algorithm. The parameter associated with the prior model of the SMAP algorithm is fixed *a priori* (the probability that the labelling will remain the same from scale $[l+1]$ to $[l]$ is equal to 0.99). Three resolution levels are considered in the multigrid and SCM algorithms and five resolution levels for the SMAP algorithm. Figures 3,4 show the segmentation results obtained with the different methods. The MRF prior model parameters $\mathbf{\Phi}_x$ and the noise model parameters $\mathbf{\Phi}_y$ obtained with our scheme are given in Table 1. Experiments indicate that the SMAP requires less computation than the multigrid or SCM algorithms but the neighborhood structure is not complex enough to describe local image properties. This model often produces blocky segmentations at the boundaries of the shadows. Besides, the multigrid, SMAP and the multiresolution approach do not permit to decrease the number of false alarms (wrong shadow detections) due to the speckle noise effect. SCM performs better, exhibits a good robustness versus speckle noise (false alarms have been eliminated), and allows us to preserve the shadow shapes of little rocks. Manufactured objects or rock shadows are better segmented with our method than with the others (cf. Fig. 3,4).

5. CONCLUSION

In this paper, we have described a novel hierarchical unsupervised segmentation algorithm based on multiscale model, a scale-causal and spatial neighborhood, and a multigrid strategy. The proposed algorithm presents several attractive features compared to other MRF models. The neighborhood structure can more accurately describe image properties and can be estimated efficiently and fastly during the segmentation process. Moreover, this approach seems to be well adapted to sonar images with strong speckle noise. This method has been validated on several real sonar images demonstrating the efficiency and robustness of this scheme. The Markovian segmentation of sonar images in three classes (*shadow*, *sea-bottom reverberation*, and *echo*) will be the topic of our next research.

Fig. 3		**SCM** Algorithm			
$\hat{\Phi}_{y_{(e_0)}}$		$0.07_{(\pi)}$	$4_{(\mu)}$	$23_{(\sigma^2)}$	
$\hat{\Phi}_{y_{(e_1)}}$		$0.93_{(\pi)}$	$12_{(min)}$	$6060_{(\alpha^2)}$	
$\hat{\Phi}_x$	$2.5_{(\beta_1)}$	$2.4_{(\beta_2)}$	$-0.8_{(\beta_3)}$	$-1_{(\beta_4)}$	$0.2_{(\beta_5)}$
Fig. 4		**SCM** Algorithm			
$\hat{\Phi}_{y_{(e_0)}}$		$0.03_{(\pi)}$	$32_{(\mu)}$	$34_{(\sigma^2)}$	
$\hat{\Phi}_{y_{(e_1)}}$		$0.97_{(\pi)}$	$38_{(min)}$	$1230_{(\alpha^2)}$	
$\hat{\Phi}_x$	$2.1_{(\beta_1)}$	$3_{(\beta_2)}$	$-1.2_{(\beta_3)}$	$-1_{(\beta_4)}$	$0.6_{(\beta_5)}$

Table 1: *Estimated parameters. π stands for the proportion of the two classes. μ and σ^2 are the Gaussian parameters (shadow area). min and α are the Rayleigh law parameters (sea floor reverberation) [8]. β_i's are the a priori parameters of the Markovian SCM model.*

6. REFERENCES

[1] P. Thourel, C. Collet, P. Bouthemy, and P. Pérez. Multiresolution analysis and MRF modelling applied to the segmentation of shadows in sonar pictures. In *Proc. 2^{nd} ACCV*, volume 2, pages 81–85, Singapore, Dec. 1996.

[2] C. Collet, P. Thourel, P. Pérez, and P. Bouthemy. Hierarchical MRF modeling for sonar picture segmentation. In *Proc. 3^{rd} IEEE ICIP*, volume 3, pages 979–982, Lausanne, Sept. 1996.

[3] C.A. Bouman and M. Shapiro. A multiscale random field model for Bayesian image segmentation. *IEEE Trans. on Image Processing*, 3(2):162–177, March 1994.

[4] Z. Kato, M. Berthod, and J. Zerubia. A hierarchical Markov random field model and multitemperature annealing for parallel image classification. volume 58, pages 18–37, 1996.

[5] H. Derin and H. Elliott. Modeling and segmentation of noisy and textured images using Gibbs random fields. *IEEE*, Trans. PAMI-9(1):39–55, Jan. 1987.

[6] F. Heitz, P. Pérez, and P. Bouthemy. Multiscale minimisation of global energy functions in some visual recovery problems. *CVGIP: Image Understanding*, 59(1):125–134, Jan. 1994.

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

[8] 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.

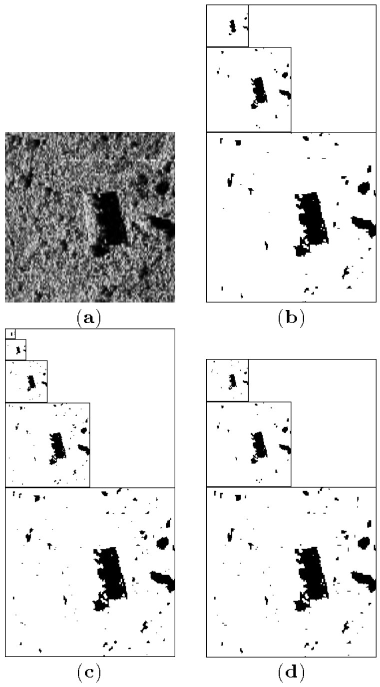

Figure 3: *(**a**) Real sonar image. -Pebbly sea floor with a manufactured object. (**b**) SCM, (**c**) SMAP, (**d**) Multigrid segmentation. The SMAP and the spatial multigrid algorithm do not permit to totally eliminate the speckle noise effect (creating 'shadow' mislabelled isolated pixels).*

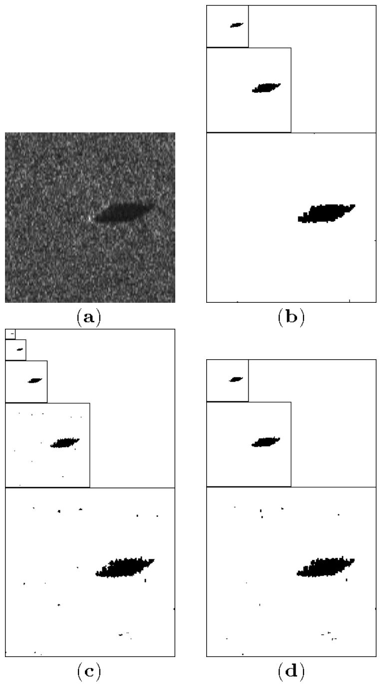

Figure 4: *(**a**) Real sonar image. -Sandy sea floor with the shadow of a man-made object. (**b**) SCM, (**c**) SMAP, (**d**) Multiresolution segmentation. Good results are obtained with the SCM algorithm exhibiting a good robustness against the speckle noise (which induces false small shadow areas for the SMAP or the multiresolution approach), and able to find a regular shadow shape close to the expected one.*