

# Segmentation markovienne hiérarchique multimodèle d'images sonar haute résolution

## Markovian Hierarchical Multimodel Segmentation for High Resolution Sonar Picture Processing

par Christophe COLLET, Pierre THOUREL, Max MIGNOTTE, Patrick PEREZ\* et Patrick BOUTHEMY\*

Groupe de Traitement du Signal, Ecole Navale,  
F-29240 Brest-Naval  
Tel : 0298 234 044 - Fax : 0298 233 857  
E-mail collet@ecole-navale.fr  
\* IRISA, Campus Universitaire de Beaulieu,  
F-35042 Rennes

### *résumé et mots clés*

Cet article présente une nouvelle approche en segmentation d'images sonar haute résolution, basée sur une modélisation markovienne hiérarchique multimodèle. Le caractère hiérarchique du modèle s'exprime dans la décomposition multirésolution des données, tandis que l'aspect multimodèle intervient dans l'expression du modèle markovien *a priori*. La démarche proposée présente l'intérêt d'être non supervisée : les paramètres des modèles du bruit sont estimés au sens du maximum de vraisemblance. Ils sont réestimés à chaque niveau de résolution de la pyramide des observations. Quant aux paramètres du modèle *a priori*, ils font l'objet d'un calibrage par la méthode des boîtes qualitatives. L'algorithme de relaxation retenu est l'ICM (Iterated Conditional Modes). L'approche multirésolution sur les données permet de s'affranchir en partie du bruit de speckle, très présent sur les images sonar, et permet une bonne initialisation de l'algorithme de relaxation. La convergence vers une configuration énergétique proche du minimum global est alors possible. Cette démarche a été validée avec succès sur de nombreuses images sonar, dont les plus représentatives sont présentées ici.

**Segmentation markovienne, multirésolution, boîtes qualitatives, bruit de speckle, loi de Weibull**

### *abstract and key words*

This paper presents a new approach to high resolution sonar image segmentation : we develop a hierarchical multimodel Markovian modelization. On one hand, the hierarchical approach relies on multiresolution decomposition of the observed data ; on the other hand, multimodel aspects appear in the expression of the Markovian *a priori* model. The proposed Markovian hierarchical multimodel algorithm is unsupervised : noise parameters are estimated at each resolution level according to Maximum Likelihood methods, while the Markovian *a priori* parameters are calibrated using the qualitative box method. The minimization is performed using the Iterated Conditional Modes algorithm. The multiresolution approach on the data allows to eliminate a large part of the speckle noise at coarser resolution levels, providing a good initialization at the finer levels. Fast convergence toward a configuration of the label field near the global minimum is thus possible. This method has been successfully validated on a number of sonar pictures : the most representative results are reported in this paper.

**Markovian segmentation, multiresolution, qualitative box, speckle noise, Weibull law.**

## introduction

Les systèmes sonar haute résolution constituent un moyen efficace pour détecter et classifier les nombreux objets se trouvant sur le fond marin en zones cotières. Depuis quelques années, la résolution de ces sonars augmentant parallèlement à leurs portées, la quantité d'information à traiter a considérablement crû, motivant la recherche de méthodes non supervisées visant à détecter et classifier, entre autre, les objets manufacturés posés sur le fond. Cet article présente donc des résultats intéressants sur la détection d'objets, à partir d'images fournies par un sonar haute résolution<sup>1</sup>, grâce à une modélisation markovienne hiérarchique multimodèle. Nous présenterons dans un premier temps le cadre applicatif associé au traitement des images sonar haute résolution ainsi que les particularités associées à l'analyse de telles images. Nous achèverons cette première partie par un état de l'art des approches markoviennes existantes permettant la segmentation ou la restauration d'images. La seconde partie de ce document présente les techniques utilisées pour la modélisation du bruit et l'estimation des paramètres associés; en particulier nous détaillerons les propriétés des estimateurs retenus. Dans une troisième partie, nous définirons le modèle markovien retenu pour la segmentation à travers l'expression de l'attache aux données et du modèle *a priori*. En particulier, nous présenterons la méthode des boîtes qualitatives que nous avons utilisée pour l'estimation automatique des paramètres du modèle markovien. La quatrième section présente l'approche markovienne hiérarchique multimodèle retenue, tandis qu'une dernière partie valide cette approche sur des images sonar synthétiques et réelles.

# 1. traitement d'images sonar haute résolution

## 1.1. cadre applicatif

Le sonar haute fréquence apporte une réponse efficace aux problèmes liés à la détection de petits objets reposant sur les fonds marins<sup>2</sup>. Après l'émission d'une onde sonore, la détection se fonde sur l'apparition d'un signal réfléchi par l'objet ou d'une interférence générée par la présence de cet objet. Encore faut-il que les dimensions et la position de l'objet soient compatibles avec les capacités de détection du sonar. Nous abordons là les questions qui se posent lors de la conception d'un système sonar et qui conduisent à l'élaboration de sonars spécialisés dans des tâches bien définies :

- le sonar détecteur qui décèle la présence d'un objet;
- le sonar classificateur qui permet l'analyse d'un objet mais avec une portée de détection plus faible;

1. Images fournies par le Groupe d'Etudes Sous-Marines de l'Atlantique, Brest, France.

2. Les auteurs remercient la DGA (Direction Générale de l'Armement) pour le soutien financier partiel apporté à cette étude.

- la caméra acoustique qui conduit à la reconnaissance de l'objet à très courte portée.

Les principales différences entre ces sonars se caractérisent en termes de fréquences d'émission et de réception (donc de taille d'antennes), de mode de visualisation du milieu (vision sectorielle, panoramique ou encore latérale) et d'architecture de système. Les images ainsi acquises par ces différents sonars, permettent de satisfaire les objectifs initialement fixés. Les capteurs utilisés dans une antenne sonar mesurant des variations de pression, il est d'usage d'étudier le module de l'amplitude complexe du signal délivré [10] et non l'intensité du signal comme en imagerie radar [27]. Le principe de formation d'une image est basé sur le découpage d'une région du fond marin en surfaces élémentaires (cellules de résolution). Une voie est formée dans une direction donnée (exploration spatiale de la zone à examiner). Puis, un échantillonnage temporel du signal réverbéré par le fond marin permet d'obtenir l'amplitude de ce signal en fonction de la distance. En formant électroniquement un grand nombre de voies, l'ensemble de la zone est exploré secteur par secteur. Une solution pour explorer toute une zone consiste à ne former qu'un petit nombre de voies très directives et à déplacer le sonar selon une direction perpendiculaire à l'axe d'émission. C'est la solution retenue pour le sonar à balayage latéral remorqué (cf. figure 1a) dont sont issues les images présentées dans cet article.

La mosaïque d'images présentée (cf. figure 2) permet de mesurer la qualité des images délivrées par un sonar à balayage latéral. Ces images représentent différents objets posés sur le fond marin (épave, pipe line, pneu, cylindre, ...). Le signal reçu par les capteurs de l'antenne est essentiellement constitué par le phénomène de réverbération de fond qui correspond à la réflexion par le sol de l'onde émise. La texture du fond joue un rôle prépondérant et se caractérise par un index de réverbération. Cet index exprime une valeur moyenne de la réverbération mais ne prend pas en compte la composition fine (les aspérités dont les dimensions sont petites au regard de la longueur d'onde du signal incident) de cette texture. Or, ce sont ces aspérités, cette rugosité du fond, qui entraînent des fluctuations importantes du signal autour de la valeur moyenne d'une zone de fond. De nombreuses études [9, 11, 16, 32, 40, 51, 53] ont été conduites afin de déterminer les différentes textures de fonds marins. Le critère de classement généralement retenu consiste à comparer la taille moyenne des grains constituant un échantillon prélevé sur le fond. Cette connaissance plus fine des constituants d'une zone permet une approche statistique de la réverbération mesurée par un capteur; c'est l'étude du phénomène de speckle ou de chatoiement.

## 1.2. particularités des images acoustiques des fonds marins

L'observation de petits détails ou de contours présents dans les images est rendue difficile par leur aspect granulaire [1]. Cette mauvaise qualité d'image est inhérente au système sonar et plus généralement à tout système d'émission-réception cohérente. Le bruit de speckle apparaît lorsqu'une onde électromagnétique ou acoustique illumine une surface constituée d'un grand nombre de diffuseurs dont les dimensions sont petites devant la longueur

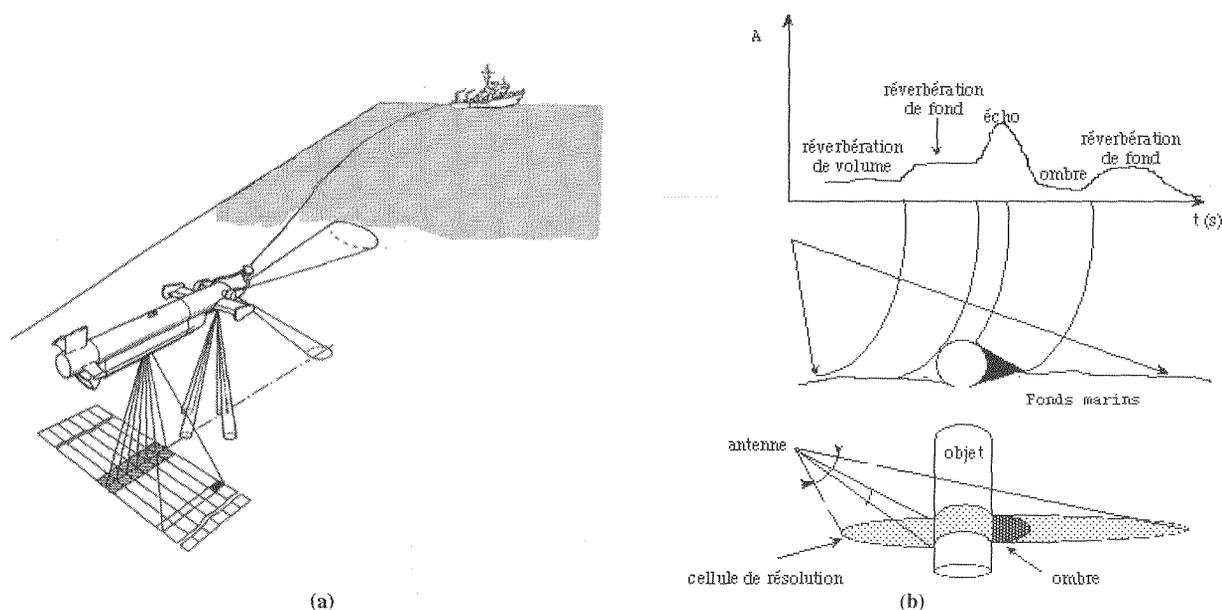


Figure 1. – Principe d'acquisition d'images sonar haute résolution. a) Exploration d'une zone du fond marin à l'aide d'un sonar à balayage latéral remorqué. C'est le déplacement du sonar qui assure le balayage de la zone par juxtaposition d'un petit nombre de voies très directives. b) Principe de la classification sur ombre. La présence d'un objet à l'intérieur de la cellule de résolution entraîne la formation d'une ombre acoustique dont l'analyse permettra de détecter puis de classifier l'objet.

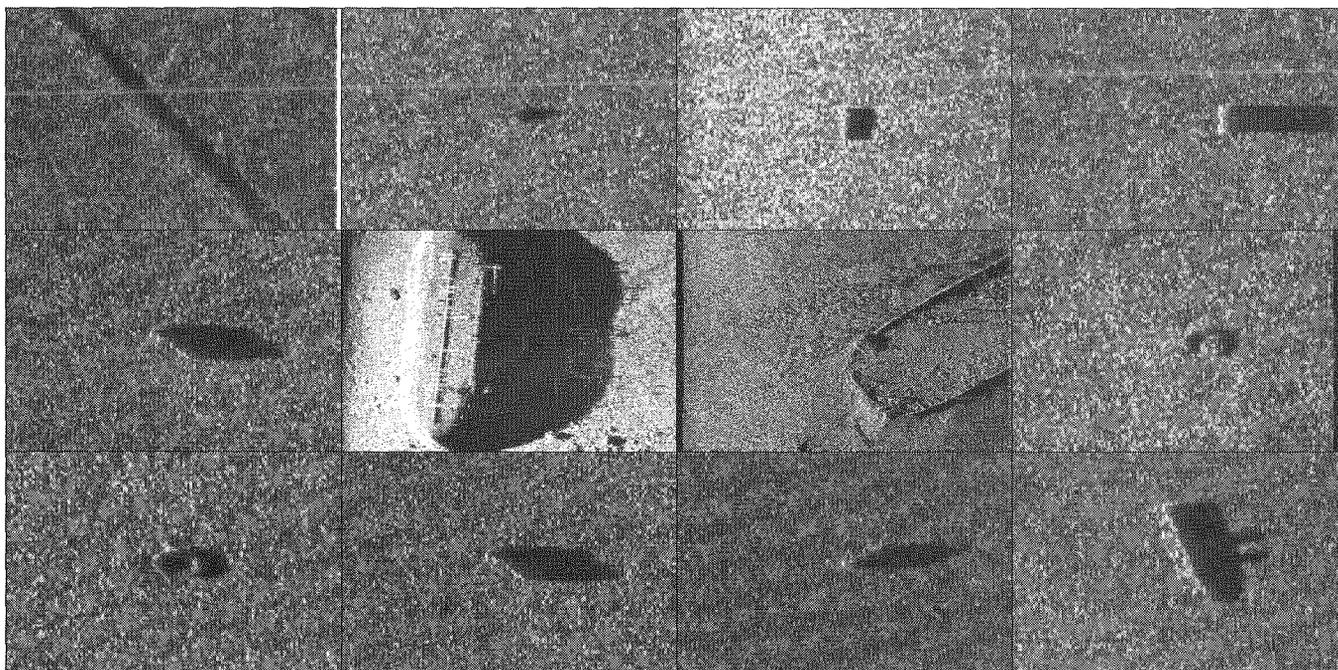


Figure 2. – Mosaïque d'images d'un sonar à balayage latéral. Les différents fonds marins présentés comportent des objets manufacturés divers (roue, épave, pipe-line, cylindre...) reposant sur le fond.

d'onde incidente. Chaque diffuseur se comporte alors comme une source émettant une onde, de même longueur d'onde, mais dont l'amplitude, la phase et la direction sont aléatoires. L'antenne de réception perçoit un signal qui est le résultat de l'interférence de toutes les ondes réfléchies par les diffuseurs présents dans la

cellule de résolution. La modélisation du bruit de speckle dans le cas des images sonar requiert la connaissance de la distribution des ondes réfléchies [10, 22, 33] par le fond marin à l'intérieur d'une cellule de résolution [15, 25, 44] et la prise en compte du principe de formation des images sonar [8].

Examinons maintenant les hypothèses que nous sommes amenés à exprimer et qui portent sur le nombre, la répartition et les dimensions des diffuseurs situés à l'intérieur de la cellule de résolution : **1)** A l'intérieur d'une cellule de résolution, le fond marin est constitué d'un grand nombre de diffuseurs indépendants et spatialement répartis de façon aléatoire; **2)** A l'échelle de la cellule de résolution, si le fond marin est constitué de zones homogènes, alors les conditions de réflectivité des ondes sont uniformes; **3)** La distance entre les diffuseurs d'une cellule de résolution et le sonar (quelques dizaines de mètres) est très supérieure à la longueur de l'onde rétrodiffusée (les longueurs d'ondes varient entre  $2 \cdot 10^{-3}$  et  $2 \cdot 10^{-2}$  m). Compte tenu des fréquences utilisées, des dimensions de la cellule de résolution (de quelques  $\text{cm}^2$  à quelques dizaines de  $\text{cm}^2$ ) et de la rugosité des différentes zones de fond marin [16, 29, 44], ces hypothèses sont vérifiées dans de nombreux cas d'utilisation du sonar. L'amplitude du signal reçu par le capteur suit alors une loi de Rayleigh  $\mathcal{R}$  définie par

$$\mathcal{R}(y, \alpha) = \frac{y}{\alpha^2} \exp\left\{-\frac{y^2}{2\alpha^2}\right\} \quad \text{pour tout } y \geq 0 \quad (1)$$

Le paramètre  $\alpha$  caractérisant une loi de Rayleigh, les moments d'ordre 1 et 2 de cette loi s'expriment en fonction de ce paramètre :

$$m_{\mathcal{R}} = E\{y\} = \alpha\sqrt{\frac{\pi}{2}} \text{ et } \sigma_{\mathcal{R}}^2 = \alpha^2\frac{(4-\pi)}{2} \text{ car } E\{y^2\} = 2\alpha^2 \quad (2)$$

Le lien entre la texture du fond marin, les caractéristiques techniques du sonar et la distribution des niveaux de gris des pixels correspondants se modélise donc par une loi de Rayleigh et traduit le bruit de speckle visible sur les images acoustiques.

### 1.3. détection et classification d'objets

Le contexte applicatif visé concerne la détection d'objets posés sur le fond marin et leur classification. La qualité d'un écho ne conduit que très rarement à une classification; aussi est née l'idée d'utiliser la forme de l'ombre produite par un objet posé sur le fond et « éclairé » par une onde acoustique. Par rapport au sonar, la zone située derrière l'objet, n'est pas insonifiée en raison de la présence de cet objet qui intercepte l'onde lors de sa propagation. Cette ombre caractérise donc une absence de réverbération de fond pendant le temps mis par l'onde pour parcourir cette zone. Le signal reçu en provenance de cette zone est constitué du bruit de réverbération de volume, également d'un peu de signal ramené par les lobes secondaires, en particulier sur les bords de la zone d'ombre, et de bruit électronique en provenance de la chaîne de traitement. Si on trace l'évolution du signal reçu en fonction du temps pour une voie donnée (cf. figure 1b), on observe en pratique des images telles que celles présentées en figure 2.

En observant la forme des ombres produites par un objet manufacturé (cylindre, pneu, épave, ...), nous constatons que celle-ci est plutôt régulière et qu'elle peut être décrite avec des attributs géométriques simples (parallélogramme, quadrilatère convexe, polygone dont certains côtés peuvent être approchés par des arcs de cercle ou d'ellipse). Les contours de ces ombres sont eux aussi plutôt réguliers et lisses (segments de droites ou arcs de cercle).

Cette régularité est à opposer à la forme des ombres produites par des rochers de taille similaire. Ces ombres possèdent des formes extrêmement variables, totalement aléatoires et qui n'ont en général, aucune caractéristique géométrique simple. Les contours de ces ombres sont très chahutés et irréguliers. Ces propriétés sont utilisées pour classifier les objets posés sur le fond. Nous avons donc choisi d'améliorer la qualité des ombres détectées grâce à une approche markovienne.

### 1.4. approche markovienne

De nombreuses études comparatives ont été menées afin de déterminer une méthode ou un groupe de méthodes permettant de réduire les effets du bruit de speckle, en particulier dans le domaine de l'imagerie radar [27][52]. Par ailleurs, la segmentation des images sonar s'effectuait jusqu'alors principalement à l'aide d'opérateurs de morphologie mathématique [4] dont l'inconvénient majeur était de travailler sur une image binarisée, après un éventuel filtrage du bruit de speckle. Dans un contexte de segmentation d'image [14][3][26] ou de restauration d'images [13][12], les études menées sur la modélisation markovienne nous sont apparues intéressantes pour diverses raisons : contrairement aux méthodes de type filtrage, les méthodes markoviennes permettent la conservation d'un lien formel de type probabiliste avec les données observées, tout au long du processus d'analyse de l'image; par ailleurs les champs markoviens permettent d'introduire des contraintes génériques qui traduisent les propriétés locales et globales des images modélisées. Les propriétés qui nous intéressent plus particulièrement ici concernent les caractéristiques géométriques des contours d'ombre, qui constituent une information potentiellement classifiante. C'est une approche de type markovien que nous avons développé et que nous allons désormais décrire.

## 2. modélisation du bruit et estimation des paramètres

La modélisation du bruit de speckle par une loi de Rayleigh précédemment justifiée est classique et correspond à de nombreuses situations en imagerie acoustique. Néanmoins, ce modèle n'est pas toujours réaliste expérimentalement lorsque l'on observe :

- une diminution sensible du nombre de diffuseurs par cellule de résolution (la rugosité du fond n'étant plus petite devant la longueur d'onde, le fond peut être modélisé par un petit nombre de facettes réfléchissantes);
- l'apparition de phénomènes d'interférences liés à la présence d'une couche de sédiments dont la taille des grains est importante par rapport à la longueur d'onde, à l'interaction de particules en suspension dans le volume d'eau en particulier dans les zones

peu profondes soumises à des mouvements importants de masses d'eau (marées, embouchures de fleuves, ...).

ou lorsque la dimension des rugosités du fond devient du même ordre de grandeur que les dimensions de la cellule de résolution (en particulier dans le cas du sonar à ouverture synthétique) ou encore lorsque le sonar est exploité dans des eaux très peu profondes [32]. Nous proposons l'utilisation d'une modélisation plus générale, basée sur la loi de Weibull  $\mathcal{W}_b(\alpha, \beta)$  [8] dont l'expression est :

$$\mathcal{W}_b(y, \alpha, \beta) = \frac{\beta}{\alpha} \left(\frac{y}{\alpha}\right)^{\beta-1} \exp\left\{-\left(\frac{y}{\alpha}\right)^\beta\right\} \quad \forall y \geq 0 \quad (3)$$

où  $\alpha$  et  $\beta$  sont respectivement des paramètres d'échelle et de forme, définis strictement positifs.

Largement employée pour modéliser le speckle présent sur les images SAR [20], la modélisation par loi de Weibull offre une grande souplesse. Ainsi pour certaines valeurs de paramètres de la loi de Weibull, nous retrouvons la loi exponentielle ( $\beta = 1$ ) ou la loi de Rayleigh. La loi de Rayleigh est en effet un cas particulier ( $\mathcal{R}(\alpha) = \mathcal{W}_b(\sqrt{2}\alpha, 2)$ ) de la loi de Weibull pour  $\beta = 2$  (cf. figure 3). Il reste maintenant à identifier le type de distribution du bruit sur images réelles, et à estimer les paramètres correspondants.

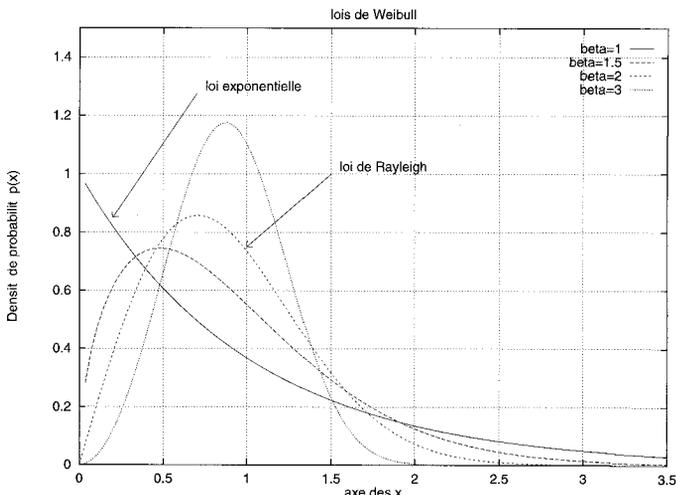


Figure 3. – Lois de Weibull pour différentes valeurs du facteur de forme.

## 2.1. identification des lois de probabilité

Le choix d'une distribution optimale parmi les différentes distributions paramétriques pour modéliser la distribution des niveaux de gris sur l'image s'effectue selon le résultat d'un test d'ajustement. Parmi les nombreux tests existants, nous avons retenu les tests de Kolmogorov-Smirnov et du  $\chi^2$  [37]. Ces deux critères permettent l'identification des distributions correspondant au fond réverbéré et à l'ombre.

Soit une variable aléatoire  $Z$  à valeur dans  $\{z_0, \dots, z_r\}$ . Un nombre  $n$  d'échantillons de  $Z$  sont observés. Parmi ceux-ci,  $k_i$  correspondent à la valeur  $z_i$  ( $i = 0, \dots, r$ ) :  $k_0 + \dots + k_r = n$ .  $Z$  suit une loi modélisable en une loi théorique de formule analytique connue. Ceci signifie que la loi présumée de  $Z$ , ou voisine de celle

de  $Z$ , devrait donner *a priori*  $nP_0$  réalisations ayant valeur  $z_0$  (où  $P_0 = P_Z(z_0)$  en théorie), ...,  $nP_r$  réalisations ayant valeur  $z_r$  (où  $P_r = P_Z(z_r)$  en théorie); pour  $n$  suffisamment grand. On peut chercher à définir l'écart existant entre la loi théorique et la loi empirique observée sur l'image. Pour ce faire, nous avons retenu deux tests :

- **Le test de Kolmogorov** : Il s'agit d'un test non paramétrique d'ajustement à une distribution entièrement spécifiée ([45] pp. 263 et 336). Avec les notations précédemment définies, posons

$$\mathcal{K} = \sum_{i=0}^r |k_i - nP_i|. \text{ Pour choisir la distribution paramétrique qui}$$

s'ajuste le mieux à l'histogramme, nous retenons celle qui minimise la « distance » de Kolmogorov  $\mathcal{K}$ .

- **Le test du  $\chi^2$**  : Le test du  $\chi^2$  mesure l'écart aléatoire entre les effectifs réalisés et les effectifs espérés. Si la loi observée (ou empirique) est bien représentée par la loi théorique, on doit s'attendre à ce que les écarts  $(k_i - nP_i)^2$  soient d'autant plus grands que  $nP_i$  est faible. Aussi, on considèrera plutôt  $\frac{(k_i - nP_i)^2}{nP_i}$  qui, pour un écart donné, décroît quand  $nP_i$  croît. La somme :

$$S = \sum_{i=0}^r \frac{(k_i - nP_i)^2}{nP_i} \text{ est calculée pour mettre en évidence}$$

la « distance » entre la distribution observée et la distribution théorique. C'est le critère du  $\chi^2$ . Ce test sera plus exigeant que le test de Kolmogorov sur les échantillons faiblement représentés ([45] pp. 335).

## 2.2. estimateurs

Les images sonar présentent un offset en amplitude tel que tous les pixels de l'image ont une amplitude supérieure à une valeur *min*. La loi de Weibull  $\mathcal{W}$  étant définie à partir de la valeur nulle, nous sommes amenés à travailler sur une loi décalée de la forme

$$\mathcal{W}_y(\min, \beta, \alpha) = \frac{\beta}{\alpha} \left[\frac{(y - \min)}{\alpha}\right]^{\beta-1} \exp\left[-\left[\frac{(y - \min)}{\alpha}\right]^\beta\right] \quad (4)$$

avec  $y > \min$ ,  $\alpha > 0$  et  $\beta > 0$ .

En supposant la classification des pixels connue, il est possible d'estimer au sens du maximum de vraisemblance les paramètres des deux distributions de probabilité modélisant respectivement les luminances dans chacune des régions (zone d'ombre (étiquette  $e_2$ ) et zone de fond réverbéré (étiquette  $e_1$ )).

Soit  $Y = \{y_i, i = 1 \dots N\}$  le champ des observations dont une réalisation  $y$  est disponible pour effectuer l'estimation des paramètres (i.e.  $y = (y_1, \dots, y_N)$  est l'ensemble des luminances  $y_i \in \mathbb{R}$  observées pour les  $N$  pixels  $i$  appartenant à une région  $e_j, j \in \{1, 2\}$ ). On suppose les  $N$  variables aléatoires indépendantes et identiquement distribuées suivant une loi de Weibull de paramètres  $\Phi_y = (\min, \alpha, \beta)$  inconnus. Pour une réalisation donnée  $Y$ , la vraisemblance associée à  $\Phi_y$  est donnée par :

$$\begin{aligned} \mathcal{L}(\Phi_y) &= P_{Y/\Phi_y}(y/\Phi_y) \\ &= \left[\frac{\beta}{\alpha^\beta}\right]^N \prod_{i=1}^N \left[(y_i - \min)^{\beta-1} \exp\left(-\frac{(y_i - \min)^\beta}{\alpha^\beta}\right)\right] \end{aligned} \quad (5)$$

Le logarithme de l'expression précédente nous donne la fonction de log-vraisemblance :

$$\begin{aligned} \ln \mathcal{L}(\Phi_y) &= \ln [P_{Y/\Phi_y}(y/\phi_y)] \\ &= N \ln \beta + (\beta - 1) \sum_{i=1}^N \ln(y_i - \min) - N\beta \ln \alpha \\ &\quad - \frac{1}{\alpha^\beta} \sum_{i=1}^N (y_i - \min)^\beta \end{aligned} \quad (6)$$

avec  $y > \min$  et  $\alpha > 0$ . Pour déterminer  $\widehat{\min}_{MV}$ ,  $\hat{\alpha}_{MV}$  et  $\hat{\beta}_{MV}$  (valeurs qui maximisent la fonction de vraisemblance), on résoud simultanément les 3 équations suivantes :

$$\frac{\partial \ln \mathcal{L}(\Phi_y)}{\partial \min} = \frac{\partial \ln \mathcal{L}(\Phi_y)}{\partial \alpha} = \frac{\partial \ln \mathcal{L}(\Phi_y)}{\partial \beta} = 0 \quad (7)$$

### 2.2.1. estimateur de min

L'estimée  $\widehat{\min}_{MV}$  exprimée par l'équation précédente, peut être approchée itérativement et est en fait très légèrement inférieure à  $y_{\min}$  (intensité minimum de l'échantillon  $y$ ). Pour cette raison et pour permettre aux pixels de niveau de gris  $y_{\min}$  d'avoir une probabilité non nulle d'apparition, nous prendrons l'estimateur suivant :

$$\widehat{\min}_{MV} = y_{\min} - 1 \quad (8)$$

### 2.2.2. estimateur de $\alpha$

Posons  $\tilde{y}_i = (y_i - \widehat{\min}_{MV})$ .  $\hat{\alpha}_{MV}$  et  $\hat{\beta}_{MV}$  sont alors solutions du système d'équations suivant :

$$\begin{aligned} \frac{N}{\hat{\beta}} + \sum_{i=1}^N \ln \tilde{y}_i &= N \ln \hat{\alpha} + \sum_{i=1}^N \left(\frac{\tilde{y}_i}{\hat{\alpha}}\right)^{\hat{\beta}} \ln \left(\frac{\tilde{y}_i}{\hat{\alpha}}\right) \\ \text{et} \quad \frac{1}{\hat{\alpha}^{\hat{\beta}}} \sum_{i=1}^N \tilde{y}_i^{\hat{\beta}} &= N \end{aligned} \quad (9)$$

soit

$$\hat{\alpha}_{MV} = \left( \frac{1}{N} \sum_{i=1}^N \tilde{y}_i^{\hat{\beta}_{MV}} \right)^{\frac{1}{\hat{\beta}_{MV}}} \quad (10)$$

On démontre facilement que l'estimateur  $\hat{\alpha}_{MV}$  est non biaisé, convergent et efficace. En effet

$$E[\hat{\alpha}^\beta] = E\left[\frac{1}{N} \sum_{i=1}^N \tilde{y}_i^\beta\right] = \frac{1}{N} \sum_{i=1}^N E[\tilde{y}_i^\beta] = \alpha^\beta \quad (11)$$

L'estimateur est donc non biaisé. Par ailleurs, le calcul de la borne inférieure de Cramer-Rao permet de montrer que cet estimateur est consistant et efficace. Par définition nous avons  $\text{VAR}_{\min}[\hat{\alpha}^\beta] \stackrel{\text{déf}}{=} \frac{1}{I(\alpha^\beta)}$  avec  $I(\alpha^\beta)$  représentant l'information de Fisher que l'on calcule aisément :

$$I(\alpha^\beta) = -E\left[\frac{\partial^2}{\partial \alpha^{2\beta}} \ln P_{Y/\alpha^\beta}(y/\alpha^\beta)\right] = \frac{N}{\alpha^{2\beta}} \quad (12)$$

La borne inférieure de Cramer-Rao est donc donnée par l'expression suivante :  $\text{VAR}_{\min}[\hat{\alpha}^\beta] = \frac{\alpha^{2\beta}}{N}$ . On calcule la variance de l'estimateur et on vérifie aisément que la variance trouvée est bien la borne inférieure de Cramer-Rao : l'estimateur est efficace et consistant.

### 2.2.3. estimateur de $\beta$

L'expression (9) se simplifie un peu et  $\hat{\beta}_{MV}$  peut se mettre sous la forme  $F(\hat{\beta}_{MV}) = \hat{\beta}_{MV}$  :

$$\hat{\beta}_{MV} = \frac{N \sum_{i=1}^N \tilde{y}_i^{\hat{\beta}_{MV}}}{N \sum_{i=1}^N \left(\tilde{y}_i^{\hat{\beta}_{MV}} \ln \tilde{y}_i\right) - \sum_{i=1}^N \ln \tilde{y}_i \sum_{i=1}^N \tilde{y}_i^{\hat{\beta}_{MV}}} \quad (13)$$

Cette équation ne peut se résoudre que numériquement par approximations successives et on ne peut obtenir une forme explicite de l'estimateur de  $\hat{\beta}_{MV}$ . Il est donc difficile de vérifier qu'il possède de bonnes propriétés asymptotiques. Néanmoins, si un estimateur non biaisé et efficace de  $\hat{\beta}_{MV}$  existe, alors il est donné par l'estimation du Maximum de Vraisemblance et donc est solution de l'équation (13), l'inverse n'étant pas vrai. Par ailleurs cette méthode ne sera valable que pour un grand nombre d'échantillons à cause de ses propriétés asymptotiques s'il n'existe pas de statistiques exhaustives ([45] pp.302). On ne pourra donc juger de l'efficacité de cet estimateur qu'expérimentalement. L'équation (13) peut se résoudre par la méthode des suites récurrentes dont nous rappelons ici le principe. Soit la suite récurrente :  $\beta_1, \beta_2 = F(\beta_1), \dots, \beta_n = F(\beta_{n-1})$ . Si cette suite converge vers une limite  $l$ , cette limite est alors solution de l'équation  $l = F(l)$ . Inversement si  $l = F(l)$  possède une solution unique (c'est notre cas puisque  $\hat{\beta}_{MV}$  est la racine unique de l'équation  $F(\hat{\beta}_{MV}) = \hat{\beta}_{MV}$ ), alors cette solution est aussi la valeur de convergence de la suite récurrente  $\hat{\beta}_n = F(\hat{\beta}_{n-1})$ . La valeur initiale  $\hat{\beta}_0$  peut être effectuée par ajustement graphique ([45] pp. 332) pour accélérer la convergence.  $\hat{\beta}_{MV}$  est donc donné par la relation suivante :  $\hat{\beta}_{MV} = \lim_{n \rightarrow \infty} \beta_n$ ; avec  $\beta_n$  défini par la relation de récurrence  $\beta_n = F(\beta_{n-1})$  et  $F$  définie par :

$$F(x) = \frac{N \sum_{i=1}^N \tilde{y}_i^x}{N \sum_{i=1}^N (\tilde{y}_i^x \ln \tilde{y}_i) - \sum_{i=1}^N \ln \tilde{y}_i \sum_{i=1}^N \tilde{y}_i^x} \quad (14)$$

En pratique, cette convergence est très rapide et expérimentalement correcte. En résumé, les estimateurs MV de la loi de Weibull

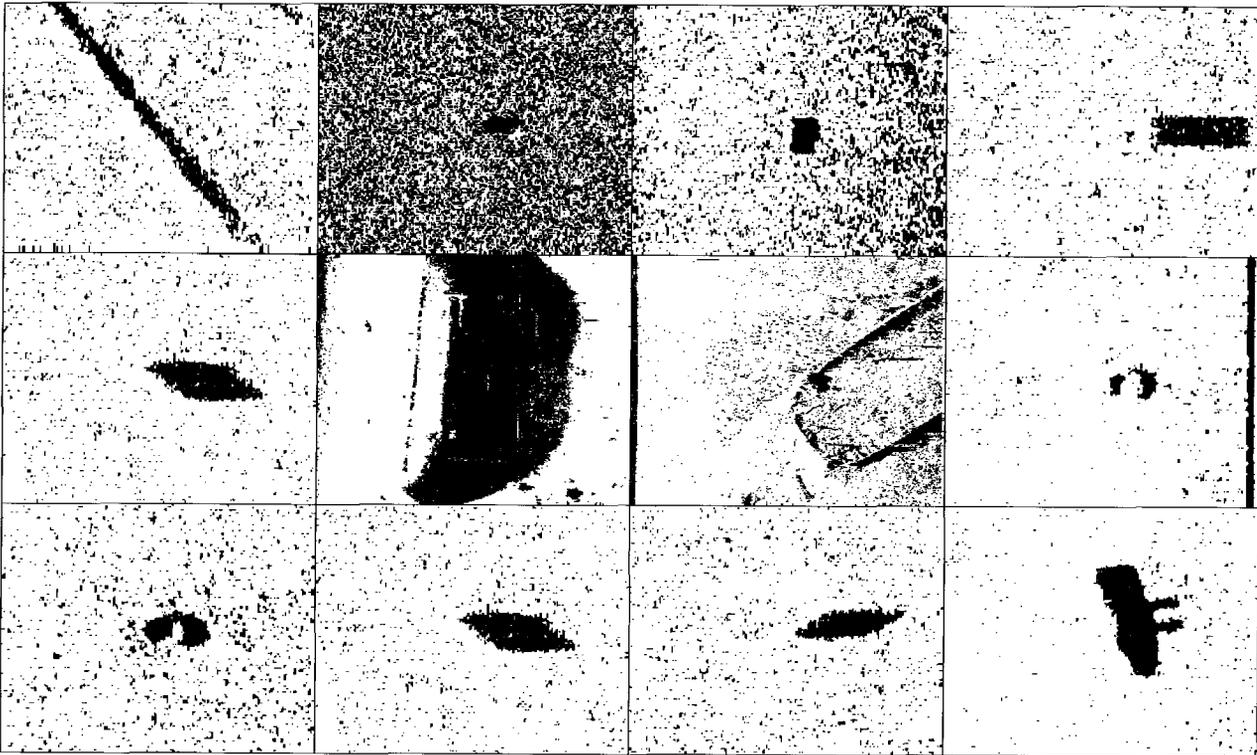


Figure 4. – Mosaïque d’images segmentées au sens du maximum de vraisemblance, avec une modélisation du bruit par lois de Weibull.

sont les suivants :

$$\widehat{min}_{MV} \approx \hat{y}_{min} - 1 \quad (15)$$

$$\hat{\beta}_{MV} = \lim_{n \rightarrow \infty} \beta_n \quad (16)$$

$$\hat{\alpha}_{MV} = \left( \frac{1}{N} \sum_{i=1}^N \tilde{y}_i^{\beta_{MV}} \right)^{\frac{1}{\beta_{MV}}} \quad (17)$$

avec  $\beta_n = F(\beta_{n-1})$ ,  $F(x)$  défini par la relation (14) et  $\beta_0 = 0$ .

Quant aux estimateurs de la loi normale  $\mathcal{N}(\mu, \sigma^2)$ , ils sont classiques :

$$\hat{\mu} = \frac{1}{N} \sum_{i=1}^N y_i \quad (18)$$

$$\hat{\sigma}^2 = \frac{1}{N-1} \sum_{i=1}^N (y_i - \hat{\mu})^2 \quad (19)$$

### 2.3. détection d’ombres par le critère du maximum de vraisemblance

Les résultats obtenus sur la mosaïque d’images (cf. figure 2) à l’aide d’un seuillage par maximum de vraisemblance sont pré-

sentés sur la figure 4. Les pixels des images sont pré-étiquetés « ombre » ou « fond réverbéré » par la méthode des K-moyennes [46][34]. Les résultats obtenus illustrent bien la difficulté associée à une détection précise des zones de fond réverbéré et des zones d’ombre. Une méthode de segmentation basée sur le critère du maximum de vraisemblance n’offre pas une solution suffisamment régularisée : les images segmentées sont clairsemées de pixels malencontreusement classés « ombre » tandis que les contours d’ombre portée des objets sont chahutés. En ce sens la classification sur ombre est alors difficile voire impossible. Nous avons donc adopté une modélisation markovienne qui permettra de régulariser la solution en introduisant explicitement une information *a priori* sur la solution recherchée, comme le montrera la section suivante.

## 3. modélisation markovienne

En associant les techniques d’estimation bayésienne et le formalisme markovien, il devient possible de proposer une approche markovienne de la segmentation. Elle consiste à estimer une image d’étiquettes connaissant l’image des luminances observées. Cette estimation s’effectue en intégrant deux notions distinctes au sein d’une distribution de probabilités :

- la modélisation des effets des systèmes d’acquisition et de formation des images qui permet ainsi de relier les observations

aux étiquettes. C'est le terme d'attache aux données explicitement décrit dans la modélisation;

- les informations *a priori* portant par exemple sur la forme des objets qui conduisent à une régularisation des zones segmentées. Elles constituent le terme contextuel, explicitement décrit dans la modélisation.

### 3.1. modélisation markovienne

Soient :

- $X$  un champ de variables aléatoires (les étiquettes) définies sur l'ensemble des sites  $S$  et à valeurs d'états dans  $\Lambda_{etq} = \{e_1, e_2\}$  ( $e_1$  et  $e_2$  désignent les étiquettes correspondant au fond et à l'ombre).  $x = \{x_s | s \in S\}$  représente une configuration de  $X$  caractérisée par la distribution  $P_X(x)$ ;
- $\Omega_{etq} = \Lambda_{etq}^S$  l'ensemble des configurations du champ des étiquettes;
- $Y$  un champ de variables aléatoires (les observations définies sur l'ensemble des sites  $S$  et à valeurs d'états dans  $\Lambda_{obs}$ .  $y = \{y_s | s \in S\}$  représente une configuration de  $Y$  caractérisée par la distribution  $P_Y(y)$ ;
- $\Omega_{obs} = \Lambda_{obs}^S$  est l'ensemble des configurations du champ des observations.

Dans le cadre de l'estimation bayésienne, il s'agit de modéliser conjointement les étiquettes et les observations à l'aide d'un champ aléatoire couplé  $(X, Y)$  défini par la distribution jointe  $P_{XY}(x, y)$  qu'il faut expliciter. La règle de Bayes indique que

$$P_{XY}(x, y) = P_{Y|X}(y|x) \times P_X(x) \quad (20)$$

La probabilité  $P_{Y|X}(y|x)$  traduit la vraisemblance des observations conditionnellement aux étiquettes, définie par la connaissance des lois de probabilité du bruit précédemment définies, alors que  $P_X(x)$  permet l'introduction de connaissances *a priori* sur les étiquettes. Par hypothèse, les champs étant markoviens, la distribution jointe  $P_{XY}(x, y)$  est une distribution de Gibbs d'énergie  $U(x, y)$  [14, 17] :

$$P_{XY}(x, y) = \frac{1}{Z} \exp\{-U(x, y)\} \quad (21)$$

où l'énergie se décompose en la somme de deux termes :  $U(x, y) = U_1(x, y) + U_2(x)$  avec

- $U_1(x, y)$  issu de la vraisemblance des observations, constitue le terme d'attache aux données qui caractérise l'adéquation des étiquettes aux observations;
- $U_2(x)$  issu de l'*a priori markovien*, constitue le terme contextuel qui introduit un effet régularisant sur le champ des étiquettes.

Le critère qui est le plus fréquemment retenu pour réaliser l'estimation du champ des étiquettes connaissant l'image est celui du MAP (*i.e.*, Maximum A Posteriori). La recherche du champ des étiquettes optimal  $\hat{x}_{opt}$  selon ce critère se traduit donc par la minimisation de la fonction  $U(x, y)$ .

### 3.2. l'attache aux données

La modélisation des processus d'acquisition, de formation et d'enregistrement d'une image (*i.e.*, le modèle de dégradation) traduit la relation entre le champ des étiquettes  $X$ , le champ des observations  $Y$  et le bruit. Si à une étiquette  $X_s$  correspond une unique observation  $Y_s$ , et que celle-ci est indépendante des autres  $Y_{t, t \neq s}$  sachant  $X_s$ , alors en posant  $\Psi_s(x_s, y_s) = -\ln P(y_s|x_s)$ , la vraisemblance des observations s'écrit :

$$\begin{aligned} P_{Y|X}(y|x) &= \exp\left\{-\sum_{s \in S} \Psi_s(x_s, y_s)\right\} \\ &= \exp\{-U_1(x, y)\} \end{aligned} \quad (22)$$

et le terme d'attache aux données est donc  $U_1(x, y) = \sum_{s \in S} \Psi_s(x_s, y_s)$ . Dans notre cas, l'expression de la fonction  $\Psi(x_s, y_s)$  est :

$$\begin{cases} \Psi_s(x_s, y_s) = -\ln(W_{y_s}(\min_{x_s}, \beta_{x_s}, \alpha_{x_s})) & \text{si } x_s = e_1 \\ \Psi_s(x_s, y_s) = -\ln(N_{y_s}(\mu_{x_s}, \sigma_{x_s})) & \text{si } x_s = e_2 \end{cases} \quad (23)$$

suivant que l'on choisisse une loi de Weibull décalée (dont la loi de Rayleigh est un cas particulier) ou une loi normale. Nous avons donc :

$$\begin{cases} \Psi_s(x_s, y_s) = -\ln \frac{\beta_{x_s}}{\alpha_{x_s}} - (\beta_{x_s} - 1) \ln \left( \frac{y_s - \min_{x_s}}{\alpha_{x_s}} \right) \\ \quad + \left( \frac{y_s - \min_{x_s}}{\alpha_{x_s}} \right)^{\beta_{x_s}} & \text{si } (y_s - \min_{x_s}) > 0, \\ \quad \text{et } \Psi_s(x_s, y_s) = +\infty \text{ sinon;} \\ \text{ou } \Psi_s(x_s, y_s) = \frac{(y_s - \mu_{x_s})^2}{2\sigma_{x_s}^2} + \ln(\sqrt{2\pi}\sigma_{x_s}) \end{cases} \quad (24)$$

Le type de loi retenue est fixé par le critère de Kolmogorov ou du  $\chi^2$  sur la pré-segmentation obtenue par l'algorithme des K-moyennes [34]. L'estimation des paramètres s'effectue alors grâce aux estimateurs précédemment définis.

### 3.3. le terme contextuel

Les contraintes ou les informations *a priori* que l'on possède sur la solution attendue, vont s'exprimer par l'intermédiaire de termes d'interactions locales. Soit un système de voisinage  $\nu$  sur  $S$  et l'ensemble  $\mathcal{C}$  des cliques correspondantes. Si  $X$  est supposé markovien par rapport à  $\nu$ , alors

$$U_2(x) = \sum_{c \in \mathcal{C}} V_c(x_c) \quad (25)$$

où  $V_c(x)$  est un potentiel d'interaction dépendant uniquement de  $x_c = \{x_s, s \in c\}$ . Nous avons dans un premier temps retenu un modèle *a priori* du type Ising [47] (voisinage  $\nu$  de la 4-connexité) afin d'assurer une bonne régularisation à la fois des zones homogènes correspondant à la rétrodiffusion des fonds marins et des zones d'ombre perturbées par la présence du bruit de speckle :

- Dans le cas des cliques singletons :  $\forall s \in S, V_s(x_s) = \alpha_s$
- Dans le cas des cliques binaires :

$$\forall \{s, t\} \in C, V_{\langle s, t \rangle}(x_s, x_t) = \begin{cases} -\beta_{\langle s, t \rangle} & \text{si } x_s = x_t \text{ où } \beta \geq 0 \\ 0 & \text{sinon} \end{cases}$$

avec  $\alpha_s$  paramètre pouvant dépendre de la valeur d'état de  $x_s$  et  $\beta_{\langle s, t \rangle}$  paramètre spécifique à chaque type de cliques binaires (noté  $\beta_{20}, \beta_{21}, \beta_{22}, \beta_{23}$  pour les cliques respectivement horizontales, verticales, diagonales à  $45^\circ$  et  $-45^\circ$ ). Classiquement, les paramètres  $\alpha_s$  permettent de caractériser chaque classe de régions, tandis que les  $\beta_{\langle s, t \rangle}$  autorisent le contrôle de la taille et de la direction des regroupements de pixels. Dans notre application nous avons choisi  $\alpha_s = 0$ . Or, après une série de segmentations menées sur des images de synthèse, il est apparu que ce modèle ne nous permettait pas de satisfaire simultanément les exigences suivantes : **1**) régularisation des zones homogènes (fond marin ou ombre); **2**) précision dans la segmentation des zones d'ombres de faible largeur; **3**) valorisation de formes géométriques simples; **4**) obtention de zones frontières régulières. Ces exigences sont contradictoires car elles imposent que les paramètres pondérateurs des cliques binaires soient à la fois des grandeurs comparables (pour assurer une régularisation isotrope) et de valeurs nettement différentes (afin de favoriser la segmentation dans une direction privilégiée). Nous avons donc choisi d'enrichir ce modèle en lui adjoignant une fonction de potentiel calculée sur les cliques quaternaires (voisinage de la 8-connexité), avec lesquelles nous allons pouvoir privilégier certaines configurations géométriques particulières. La clique singleton est notée 10 tandis que les configurations possibles des cliques binaires associées au voisinage  $\nu$  sont notées de 20 à 27. Modulo la symétrie par rapport aux étiquettes  $\{e_1, e_2\}$ , huit configurations quaternaires sont possibles (40 à 47) cf. figure 5. L'ensemble des cliques ternaires (également engendrées dans le système de 8-voisinage) n'a pas été retenu car la variété des formes géométriques décrites par ces cliques n'est pas très importante; toutes les cliques ternaires peuvent s'écrire comme un combinaison de configurations de cliques quaternaires. Les cliques quaternaires permettent en outre une discrimination géométrique plus significative; la prise en compte de ces cliques augmente le nombre de paramètres à estimer et allonge sensiblement les temps de calcul des différents algorithmes d'optimisation sans améliorer de façon pertinente le modèle. On notera  $C_2$  l'ensemble des cliques binaires, tandis que  $C_4$  désignera l'ensemble des cliques quaternaires. La fonction de potentiel  $V_c(x)$  associée à une clique quaternaire  $c \in C_4$ , permettra de « favoriser », « défavoriser » ou bien de « rester neutre » vis-à-vis de ces configurations. Cela s'effectue à travers le choix des paramètres  $\beta_{4j}$  (avec  $j = 0, \dots, 7$ ), associés par la fonction de potentiel  $V_c$  aux huit configurations possibles de  $x_c$  :

$$V_c(x) = -\beta_{4j}$$

si les étiquettes de  $x_c$  correspondent à la configuration de type "4j" (cf. figure 5).

Le modèle retenu présente l'avantage de rendre compte aussi bien des formes détaillées et « chahutées » des ombres de rochers que de celles, beaucoup plus géométriques, des objets manufacturés. Cette « fidélité » de la segmentation est fondamentale car elle conditionne le taux de bonne classification de la suite de la chaîne

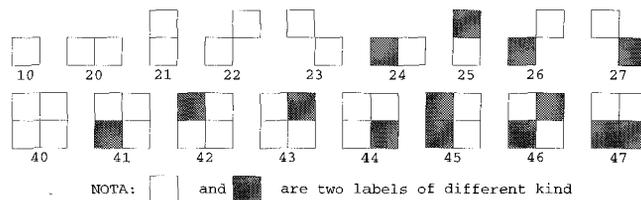


Figure 5. – Configurations possibles des cliques unaires, binaires et quaternaires  $c_{ij} \in C$  pour un système de 8 voisinage, à une symétrie près sur les valeurs des étiquettes.

de traitement, où l'allure du contour de l'ombre est un des critères discriminants. Le nombre de configurations de cliques retenues (douze au total) permet de conserver des temps d'estimation des paramètres raisonnables.

Le terme  $U_2(x)$  d'énergie contextuelle s'écrit donc :

$$\begin{aligned} U_2(x) &= \sum_{c \in C_2} V_c(x_c) + \sum_{c \in C_4} V_c(x_c) \\ &= - \sum_{i=0}^3 \sum_{\langle s, t \rangle \in C_{2i}} \beta_{2i} \delta(x_s, x_t) - \sum_{i=0}^7 \sum_{\substack{c \in C_4: \\ x_c \text{ de type "4j"}}} \beta_{4j} \end{aligned} \quad (26)$$

où  $C_{2i}$  ( $i = 0, \dots, 3$ ) désigne l'ensemble des cliques binaires horizontales, verticales, diagonales à  $45^\circ$  et diagonales à  $-45^\circ$  et  $\delta$  le symbole de Kronecker. Ce terme d'énergie est donc caractérisé par douze paramètres ( $\beta_{2i}$  avec  $i$  variant de 0 à 3 et  $\beta_{4j}$  avec  $j$  variant de 0 à 7)<sup>3</sup> qu'il faut déterminer. Le modèle markovien que nous avons développé pour la segmentation des images sonar [18, 49], s'exprime donc par l'intermédiaire de la fonction d'énergie suivante :

$$U(x, y) = \sum_{s \in S} \Psi_s(x_s, y_s) + \sum_{c \in C_2} V_c(x_c) + \sum_{c \in C_4} V_c(x_c) \quad (27)$$

Les paramètres  $\beta_{ij}$  décrivant *a priori* du modèle markovien doivent désormais être estimés : c'est l'objet du paragraphe suivant.

### 3.4. calibrage par boîtes qualitatives

Cette méthode [2] consiste à définir un certain nombre de couples de configurations locales et particulières d'étiquettes puis, à partir de quelques situations représentatives, à imposer des probabilités minimales ou maximales de transition d'une configuration à l'autre. Grâce au formalisme markovien, ces probabilités de transition s'expriment facilement en terme de variations d'énergie dont les inconnues sont les paramètres  $\beta_{ij}$ . L'ensemble des situations représentatives conduit à la construction d'un système d'inéquations dont la résolution permet de déterminer un intervalle borné de valeurs pour chaque paramètre. La difficulté de cette technique réside dans le choix des situations qui doivent être suffisamment contraignantes pour obtenir des intervalles de

3. Pour alléger la notation, nous emploierons par la suite le terme  $\beta_{ij}$  pour désigner l'ensemble de ces douze paramètres.

valeurs bornés, mais pas trop pour ne pas aboutir à un système impossible.

Soient  $x_{|x_s=e_1}$  et  $x_{|x_s=e_2}$  deux configurations identiques du champ d'étiquettes  $x$  sauf au site  $s$  où elles prennent respectivement les valeurs  $x_s = e_1$  et  $x_s = e_2$ . Lorsque l'on s'intéresse à la probabilité d'occurrence de l'étiquette  $x_s = e_1$  au site  $s$  conditionnellement au champ  $x$ ; on peut, en vertu du formalisme markovien, se restreindre au voisinage  $\nu_s$  et donc à la probabilité  $P(x_s = e_1 | x \in \nu_s, y) = p$ . Si nous souhaitons imposer l'étiquette  $x_s = e_1$  avec une probabilité supérieure à  $p$  alors, la probabilité de transition de la configuration  $x_{|x_s=e_1}$  à  $x_{|x_s=e_2}$ , exprimée en terme d'énergie, s'écrit sous la forme suivante :

$$U(x_{|x_s=e_1}, y) - U(x_{|x_s=e_2}, y) \leq \log\left(\frac{1}{p} - 1\right) \quad (28)$$

Ainsi en fixant des couples de configurations locales et des valeurs pour les probabilités de transition, nous aboutissons à un système d'inéquations dont les inconnues sont les paramètres  $\beta_{ij}$ . Nous allons ainsi traduire trois exemples de situations locales d'étiquettes que nous souhaitons favoriser.

- Nous voulons éviter la détection de points isolés. Si les huit voisins du site  $s$  sont étiquetés « ombre », nous voulons attribuer l'étiquette « ombre » au site  $s$  avec une probabilité  $p = 0.9$ , même si la valeur du pixel observé  $y_s$  privilégie plutôt une étiquette « fond ». C'est-à-dire que la configuration (1) de la figure 6 est plus probable que la configuration (2). L'équation 28 s'écrit donc :

$$\left(-2 \sum_{j=0}^3 \beta_{2j} - 4\beta_{40} + \Psi_s(x_s = e_2, y_s)\right) - \left(-\sum_{k=1}^4 \beta_{4k} + \Psi_s(x_s = e_1, y_s)\right) \leq \eta \quad (29)$$

où  $\eta = \log\left(\frac{1}{p} - 1\right)$ . D'où :

$$\sum_{j=0}^3 \beta_{2j} + 2\beta_{40} - \frac{1}{2} \sum_{k=1}^4 \beta_{4k} \geq \Psi_s(x_s = e_2, y_s) - \Psi_s(x_s = e_1, y_s) - \frac{1}{2}\eta \quad (30)$$

Le terme  $\Psi_s(x_s = e_2, y_s) - \Psi_s(x_s = e_1, y_s)$  est une constante parfaitement déterminée en fonction de la valeur  $y_s$  de l'observation au site  $s$ .

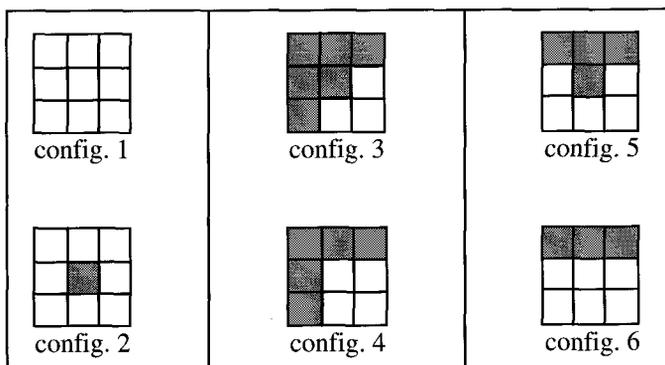


Figure 6. - Trois couples de configurations particulières d'étiquettes utilisés pour la méthode des boîtes qualitatives (1, 2)(3, 4)(5, 6).

- Nous voulons lisser les contours. En présence de cinq voisins de la classe « ombre » et lorsque la valeur de l'observation  $y_s$  privilégie plutôt une étiquette « ombre » nous voulons que le site  $s$  soit effectivement étiqueté « ombre » (configuration (3) de la figure 6) et non pas « fond » (configuration (4) de la figure 6), avec une probabilité  $p = 0.9$ . Après simplification, l'équation 28 s'écrit donc :

$$-2\beta_{22} - \beta_{42} - \beta_{44} + \beta_{45} + \beta_{47} \leq \Psi_s(x_s = e_1, y_s) - \Psi_s(x_s = e_2, y_s) + \eta \quad (31)$$

avec  $\eta = \log\left(\frac{1}{p} - 1\right)$ .

- Pour limiter les effets d'un lissage trop important des contours, nous voulons que la situation (5) de la figure 6 soit plus probable que la situation (6) avec une probabilité de  $p = 0.9$  lorsque la valeur de l'observation  $y_s$  privilégie plutôt une étiquette « ombre ». D'où :

$$2\beta_{20} + 2\beta_{40} + 2\beta_{47} - \sum_{k=1}^4 \beta_{4k} \leq \Psi_s(x_s = e_1, y_s) - \Psi_s(x_s = e_2, y_s) + \eta \quad (32)$$

avec  $\eta = \log\left(\frac{1}{p} - 1\right)$ .

D'autres équations permettant le calibrage des paramètres sont obtenues [48] grâce aux relations entre les segments de droites définissant le quadrilatère formé par l'ombre portée d'un objet cylindrique. Ces équations, jointes aux trois équations déjà décrites, forment un système dont l'ensemble des solutions est situé dans un convexe de  $\mathbb{R}^{12}$ . Pour résoudre ce système, nous avons utilisé le logiciel de programmation linéaire « SIMPLEX » [48] qui minimise une contrainte linéaire sur les paramètres (ici la somme des paramètres). Il peut arriver que les contraintes que nous avons imposées sous la forme des inéquations soient trop contraignantes, rendant le système impossible à résoudre. Dans ce cas, nous choisissons de ne considérer qu'un modèle isotrope portant sur des cliques binaires. L'estimation de l'unique valeur des coefficients pondérateurs  $\beta_{2j}$  se fait en résolvant, par l'intermédiaire du logiciel « SIMPLEX », un système réduit aux deux premières inéquations pour lequel tous les  $\beta_{4k}$  sont nuls.

## 4. approche markovienne hiérarchique multimodèle

Sur la base des premiers modèles markoviens spatiaux [21], diverses approches ont été proposées depuis une dizaine d'années pour augmenter les performances et les potentialités : approches multigrilles, multiéchelles, multirésolutions, hiérarchiques... Une classification des principales approches a été tentée dans [24][23] afin de clarifier ce que chacune d'elle recouvre. On peut globalement considérer les quatre approches suivantes :

- Approche **multirésolution** [6, 13, 35, 36, 49] : elle correspond à la discrétisation spatiale des *observations à différents niveaux de résolution* et aboutit généralement à une représentation pyramidale des données observées. Les techniques utilisées associent opération de décimation spatiale et filtrage (quad-arbre, pyramide gaussienne, transformée en ondelettes...). L'intérêt de cette approche est d'obtenir une représentation plus ou moins grossière des observations permettant de sélectionner la finesse de l'information recherchée en fonction de l'échelle.

- Approche **multiéchelle ou multigrille** [7, 39, 38] : cette approche consiste à résoudre une séquence de problèmes d'optimisation globale sous contrainte, sur les configurations visitées. A chaque échelle, une fonction d'énergie se déduit du modèle initial et peut s'interpréter comme la fonction d'énergie associée à un modèle markovien grossier. Cette approche aboutit à une structure pyramidale de primitives (champ des étiquettes) couplé à un seul niveau d'observation.

- Approche **hiérarchique** [42, 3, 50] : elle définit de manière la plus générale, un modèle global ou un ensemble de modèles interdépendants.

- Approche **multimodèle**, où différents modèles markoviens sont définis sur des graphes quelconques, associés à des topologies pouvant être différentes d'un niveau hiérarchique à l'autre, et mettant en jeu des fonctions potentielles *a priori* quelconques [48][43][28]. On peut évidemment panacher l'approche multirésolution et multimodèle [48].

Nous proposons ici d'associer au modèle markovien, une représentation multirésolution des données, par le biais d'une décomposition pyramidale des images sonar. Par ailleurs, le modèle markovien retenu sera différencié selon l'échelle de représentation des données; d'où le nom d'*approche markovienne hiérarchique multimodèle* retenu.

#### 4.1. approche hiérarchique multimodèle

L'analyse multirésolution des images améliore le RSB [48] dans le cas d'images synthétiques (cf. paragraphe 5.1), ce qui facilite l'opération de segmentation, mais entraîne un lissage des contours des ombres : la certitude de l'existence d'une ombre de taille significative croît tandis que la précision sur la position de ses contours décroît avec le niveau de résolution. Ainsi, aux résolutions grossières, nous ne chercherons pas à modéliser finement les formes des ombres et les zones de transition « ombre-fond » par le biais de configurations d'étiquettes sur des cliques quaternaires.

Pour la construction de la pyramide des observations, nous avons retenu l'algorithme de décomposition multirésolution par ondelettes proposé par S. Mallat [31, 30]. Cet algorithme consiste à projeter l'image sur une base d'ondelettes orthogonales. Nous obtenons alors une série de coefficients d'ondelettes d'« approximation » (ou de tendance) qui contient la majeure partie de l'information issue de l'image (ce sont les basses fréquences de l'image) et une série de coefficients d'ondelettes de « détail » (ou de fluctuation) qui représentent les hautes fréquences de l'image, *i.e.*, les petits détails qui enrichissent l'image. A l'aide des coefficients d'approximation, une image basse-fréquence est

reconstruite à la résolution immédiatement inférieure. Ce processus de décomposition, filtrage et reconstruction, est réitéré jusqu'au niveau de résolution désiré. Une pyramide des observations est ainsi générée, paramétrée par l'indice d'échelle  $l$ ,  $l \in [0, \dots, L]$ , appelé résolution. Pour tous les niveaux de résolution  $l \geq 1$ , nous choisissons de munir les grilles réduites  $S^l$  correspondantes d'un système de 4-voisinage ( $\nu^l$ ) et des ensembles de cliques binaires  $C_2^l$  associés. En revanche, à la pleine résolution, nous adoptons l'ensemble  $C_2 \cup C_4$  des cliques binaires et quaternaires associé au système de 8-voisinage précédemment défini. Chaque modèle markovien est donc caractérisé par un jeu de paramètres  $\theta^l = (\theta_1^l, \theta_2^l)$  qui dépend du niveau de résolution :

- A chaque niveau, la nature et les paramètres  $\theta_1^l$  des termes d'attache aux données sont estimés à partir de l'image de résolution  $l$  correspondante. Le terme d'énergie représentatif de l'attache aux données au niveau  $l$  s'écrira, par convention,  $U_{1,\theta_1^l}^l(x^l, y^l)$ .

- Pour chaque niveau de résolution grossière, nous adoptons un modèle *a priori* isotrope et ne portant que sur des cliques binaires (modèle paramétré par  $\theta_2^l = \{\beta_{2i}^l, i = 0..3\}$ ). Le terme d'énergie, représentatif du modèle *a priori* au niveau  $l \geq 1$  s'écrira, par convention  $U_{2,\theta_2^l}^l(x^l)$ .

- A pleine résolution  $l = 0$ , le modèle *a priori* intègre des informations sur la forme des ombres par le biais des cliques binaires et de configurations particulières d'étiquettes sur les cliques quaternaires (modèle paramétré par  $\theta_2^0 = \{\beta_{2i}^0, i = 0..3; \beta_{4j}^0, j = 0..7\}$ ).

Chaque fonction d'énergie s'exprime de la façon suivante :

$\forall l \in \{1, \dots, L\}$  :

$$\begin{aligned} U_{\theta^l}^l(x^l, y^l) &= U_{1,\theta_1^l}^l(x^l, y^l) + U_{2,\theta_2^l}^l(x^l) \\ &= \sum_{s \in S^l} \Psi_{s,\theta_1^l}(x_s^l, y_s^l) + \sum_{c \in C_2^l} V_{c,\theta_2^l}(x^l) \end{aligned} \quad (33)$$

Pour  $l = 0$  :

$$U_{\theta^0}^0(x, y) = \sum_{s \in S} \Psi_{s,\theta_1^0}(x_s^0, y_s^0) + \sum_{c \in C_2 \cup C_4} V_{c,\theta_2^0}(x^0) \quad (34)$$

où, pour chaque niveau de résolution  $l$ ,  $\Psi_{s,\theta_1^l}(x_s^l, y_s^l)$  modélise le terme d'attache aux données, paramétré par les coefficients  $\theta_1^l$  et  $V_{c,\theta_2^l}(x^l)$  représente les fonctions de potentiel liées à l'*a priori* et pondérées par les paramètres  $\theta_2^l$ .

Les paramètres d'attache aux données  $\theta_1^l$  sont quant à eux identifiés grâce à la méthode présentée en section 2.1 : une première segmentation par l'algorithme des K-moyennes (cf. section 2.3) permet une estimation des paramètres grâce aux estimateurs retenus (cf. section 2.2) au niveau  $l = L$ . Pour les niveaux de résolution  $l < L$ , le champ d'étiquettes obtenu au niveau  $l + 1$  sert à estimer les paramètres d'attache aux données  $\theta_1^l$ .

#### 4.2. algorithme de relaxation

Il importe avant tout de bien segmenter les zones homogènes et de faire disparaître les points isolés mal étiquetés. Un simple modèle markovien isotrope suffit pour répondre à ce besoin et assurer *in fine* une bonne initialisation pour une relaxation à

▲ Algorithme MULTIMODELE ( $L, x_0^L, y, U_{\theta^l}^l, \tau^l, [l = 0, 1, \dots, L]$ )	
$L + 1$	nombre de niveaux de résolution
$x_0^L$	configuration initiale du champ des étiquettes au niveau le plus grossier $L$
$y$	image à pleine résolution
$U_{\theta^l}^l$	fonction d'énergie paramétrée par $\theta^l$ au niveau $l$
$\theta^l$	jeu de paramètres définissant le modèle markovien au niveau de résolution $l$
$\tau^l$	critère d'arrêt au niveau de résolution $l$
<ul style="list-style-type: none"> <li>• <math>l \leftarrow 0</math></li> <li>• <b>Tant que</b> <math>l &lt; L</math> <b>faire</b> :                             <ul style="list-style-type: none"> <li>◇ <math>l \leftarrow l + 1</math></li> <li>◇ construire <math>y^l</math> à partir de <math>y^{l-1}</math></li> </ul> </li> </ul>	
<b>Fin tant que</b>	
<ul style="list-style-type: none"> <li>• <math>l \leftarrow L + 1</math></li> <li>• <b>Tant que</b> <math>l &gt; 0</math> <b>faire</b> :                             <ul style="list-style-type: none"> <li>◇ <math>l \leftarrow l - 1</math></li> <li>◇ <b>si</b> <math>l = L</math> <b>alors</b> <math>x^l \leftarrow x_0^L</math>, <b>sinon</b> <math>x^l \leftarrow \text{Duplication}(x^{l+1})</math></li> <li>◇ estimer <math>\theta^l</math> à partir de <math>y^l</math> et <math>x^l</math></li> <li>◇ <math>x^l \leftarrow \text{ICM}(x^l, y^l, U_{\theta^l}^l, \tau^l)</math></li> </ul> </li> </ul>	
<b>Fin tant que</b>	
<b>Fin Algorithme MULTIMODELE ▲</b>	

Tableau 1. – Algorithme Multimodèle reposant sur une stratégie «descendante» et effectuant à chaque niveau de résolution, une optimisation par ICM. L'algorithme de relaxation ICM a pour paramètre  $\tau^l$  qui désigne le critère d'arrêt de l'algorithme retenu au niveau  $l$ . Le critère retenu en pratique est indépendant de  $l$  et représente le pourcentage d'étiquettes variant entre deux itérations de l'algorithme ICM ( $\tau = 0.1\%$ ); où une itération désigne le balayage de tous les pixels de l'image  $y^l$ .

pleine résolution. Par contre, cette dernière relaxation sera guidée par un autre modèle markovien enrichi d'informations *a priori* portant sur la nature de la forme des ombres, grâce aux cliques quaternaires. C'est ainsi que nous obtenons une segmentation markovienne hiérarchique (multirésolution sur les données) multimodèle (cliques binaires puis quaternaires) utilisant aux différents niveaux de résolution sur les observations, une algorithme de relaxation ICM (Iterated Conditional Modes [5]) : cf. Tableau 1.

L'introduction de  $L + 1$  jeux  $\theta^l$  de paramètres caractérise la fonction d'énergie  $U_{\theta^l}^l(x^l, y^l)$  modélisant les interactions à chaque niveau de résolution, entre les observations et les étiquettes. La chaîne de traitement multimodèle, qui s'articule autour de l'algorithme, permet de propager par l'application d'une stratégie « descendante », une modélisation beaucoup plus précise des lois d'attache aux données, réestimées à chaque niveau de résolution  $l$ , ainsi qu'une initialisation pour la phase de relaxation à pleine résolution qui est de meilleure qualité, comme le présente le paragraphe suivant.

## 5. résultats comparatifs

Cette partie présente les résultats obtenus sur des images de synthèse et sur les images sonar réelles. L'utilisation d'images de synthèse (cf. figures 7 et 8) permet de quantifier l'erreur commise

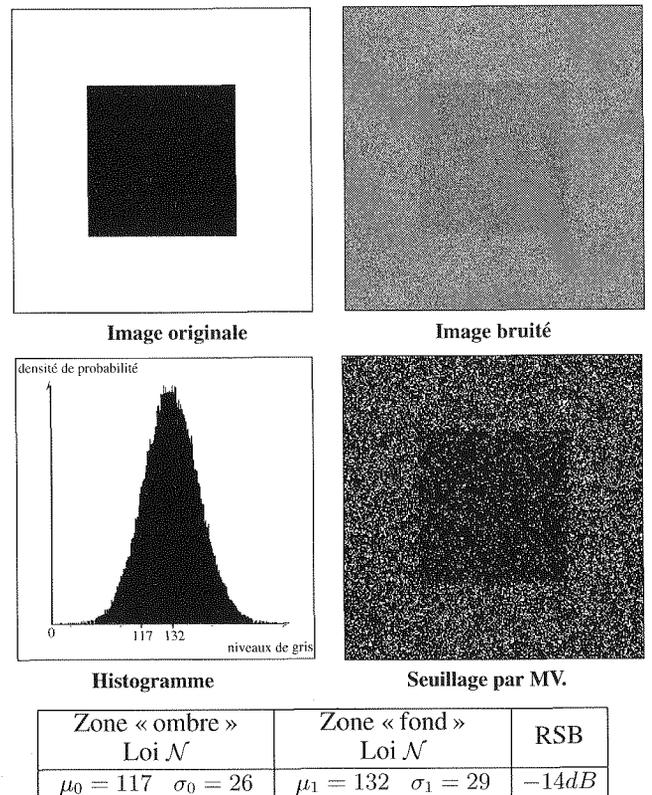
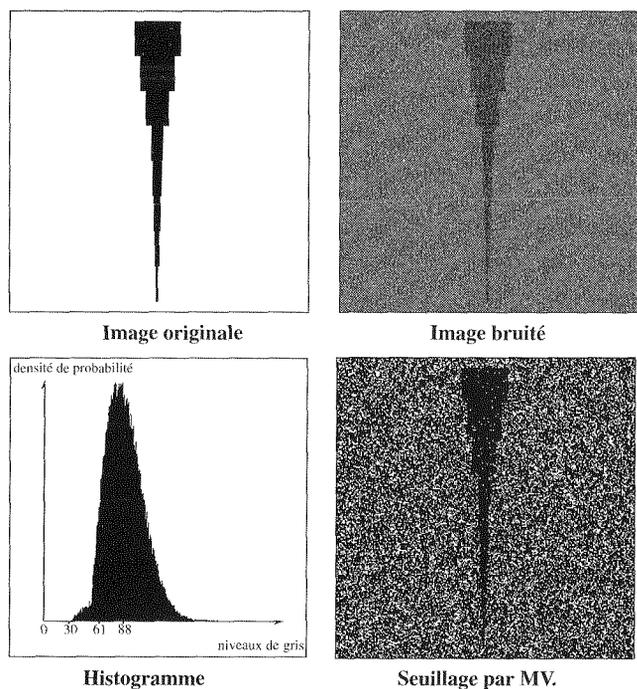


Figure 7. – Image de synthèse représentant un carré sur un fond homogène, chaque zone étant bruitée par une loi normale.



Zone « ombre » Loi $\mathcal{R}$	Zone « fond » Loi $\mathcal{R}$	RSB
$\min_0 = i_0 = 30 \quad \alpha_0 = 25$	$\min_1 = i_1 = 51 \quad \alpha_1 = 30$	-10dB

Figure 8. – Image de synthèse représentant une mire sur un fond homogène, chaque zone étant bruitée par une loi de Rayleigh décalée.

vis-à-vis d'une référence connue (image de synthèse binaire non bruitée), alors que sur images réelles (cf. Figure 11) on ne dispose pas de « vérité terrain » mais seulement d'une « vérité expert ».

## 5.1. images de synthèse

Afin de mesurer de façon pertinente la qualité des résultats des différentes segmentations, nous avons été amenés à générer des images de synthèse dont nous maîtrisons parfaitement les différents paramètres, et pour lesquelles nous disposons d'une « vérité terrain ». Les tests portent sur deux séries d'images binaires de taille 256 par 256 pixels et dont les motifs sont :

- un carré simulant l'ombre (étiquette  $e_2$ );
- une « mire » (i.e., un empilement de rectangles de taille décroissante) simulant également l'ombre.

Pour chaque image simulée, nous présentons l'image originale, l'image bruitée, l'histogramme et le résultat d'un seuillage selon le critère du maximum de vraisemblance (cf. figures 7 et 8). Ces images sont accompagnées d'un tableau résumant toutes les caractéristiques des lois du bruit. Chaque zone de l'image a été bruitée soit par une loi de Rayleigh décalée, soit par une loi normale. Pour quantifier la difficulté des images simulées, nous utilisons le terme de rapport signal-sur-bruit (RSB) généralement

défini par [48] :

$$\text{RSB en dB} = 10 \log \left( \frac{\sigma_{\text{image}}^2}{\sigma_{\text{bruit}}^2} \right) \quad (35)$$

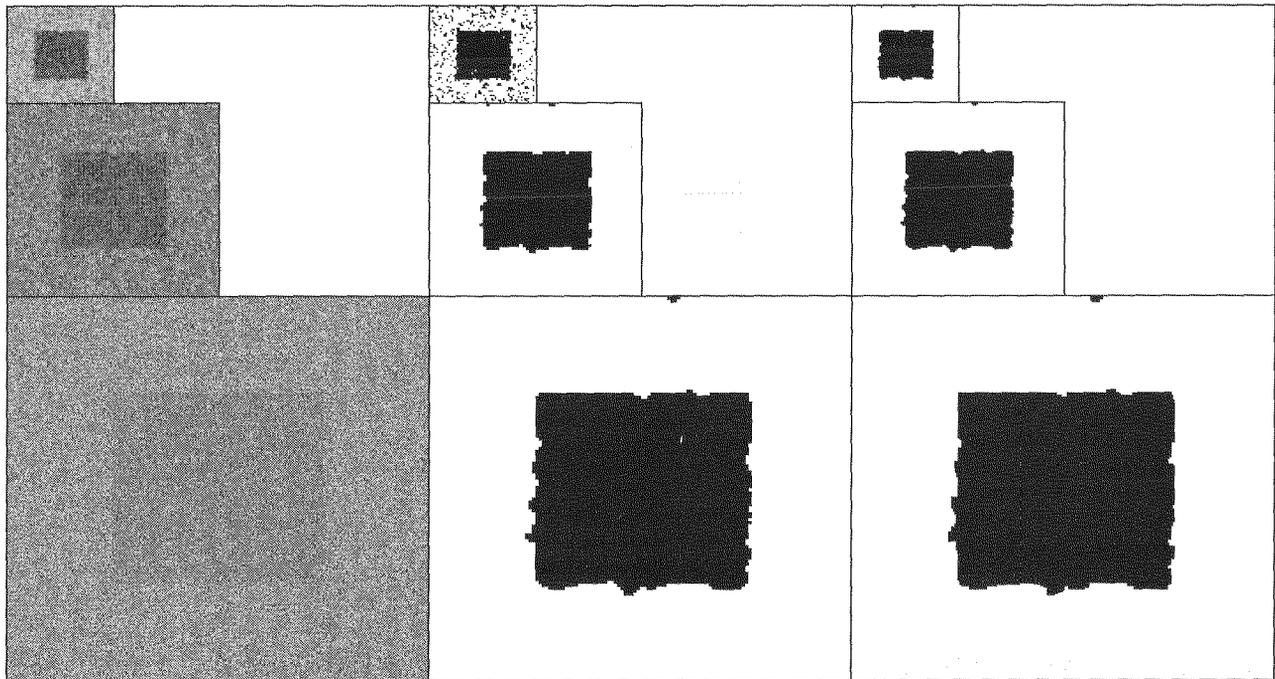
## 5.2. segmentation d'images de synthèse

Dans le cas de l'image du carré bruitée par deux lois normales (cf. figure 7), l'algorithme multimodèle fournit une image segmentée (cf. figure 9) qui offre une très forte probabilité de détection ( $P_{\text{det}} = 99.74\%$ ) et une très faible probabilité de fausse alarme ( $P_{\text{fa}}$ ). La probabilité de détection  $P_{\text{det}}$  désigne le nombre de sites convenablement classés « ombre » sur le nombre total de sites « ombre ». La probabilité de fausse alarme  $P_{\text{fa}}$  désigne le nombre de sites improprement classés « ombre ». Par rapport aux chaînes de traitement « multirésolution-monomodèle » [18] ou multigrille [19], l'amélioration est sensible en terme de diminution du nombre de fausses alarmes. Au niveau le plus grossier ( $L = 2$ ), la valeur de seuil proposée par l'algorithme des K-moyennes n'a pas permis une initialisation de bonne qualité. Néanmoins, le premier modèle markovien isotrope ( $\beta_{2i} = \beta \forall i, \beta_{4j} = 0 \forall j$ ) est suffisamment régularisant pour propager aux niveaux de résolution plus fins un champ initial d'étiquettes de très bonne qualité. A pleine résolution, le modèle markovien adopté permet une régularisation des contours de meilleure qualité que celle obtenue par un modèle isotrope. Les paramètres du bruit sont généralement mieux estimés dans la zone de fond réverbéré car le nombre d'échantillons de réalisation de la loi est supérieur au nombre d'échantillons correspondant aux zones d'ombre. L'image d'une mire bruitée par deux lois de Rayleigh (cf. figure 8) permet de mesurer la robustesse de l'algorithme de segmentation (cf. figure 10) vis-à-vis de fines zones d'ombre. Ces résultats sont à comparer à ceux obtenus par seuillage au sens du maximum de vraisemblance (cf. figures 7 et 8).

Ces résultats obtenus sur images de synthèse où les vérités « terrain » sont connues, ont permis de montrer la robustesse de la méthode. Le paragraphe suivant présente les résultats obtenus sur images réelles, où seule l'analyse d'un expert est disponible.

## 5.3. segmentation d'images sonar haute résolution

Les résultats obtenus avec la chaîne de traitement multimodèle (cf. figure 11) sont particulièrement convaincants dans le cas des images réelles : quelle que soit la taille de l'ombre portée, la segmentation permet de retrouver cette ombre et élimine la majeure partie des étiquettes « ombre » isolées. Les résultats obtenus sont visuellement extrêmement proches du résultat escompté. Les contours de l'ombre des objets manufacturés (pipe-line, cylindres, futs, épaves, pneus, chariot) sont parfaitement détectés. La forme de l'ombre qui est grossièrement estimée à la résolution  $L = 3$ , devient plus précise au fur et à mesure que la résolution devient plus fine. Le premier modèle markovien isotrope fournit au cours des phases de relaxation des niveaux de résolution grossiers, une très bonne allure générale des ombres. A pleine résolution, le

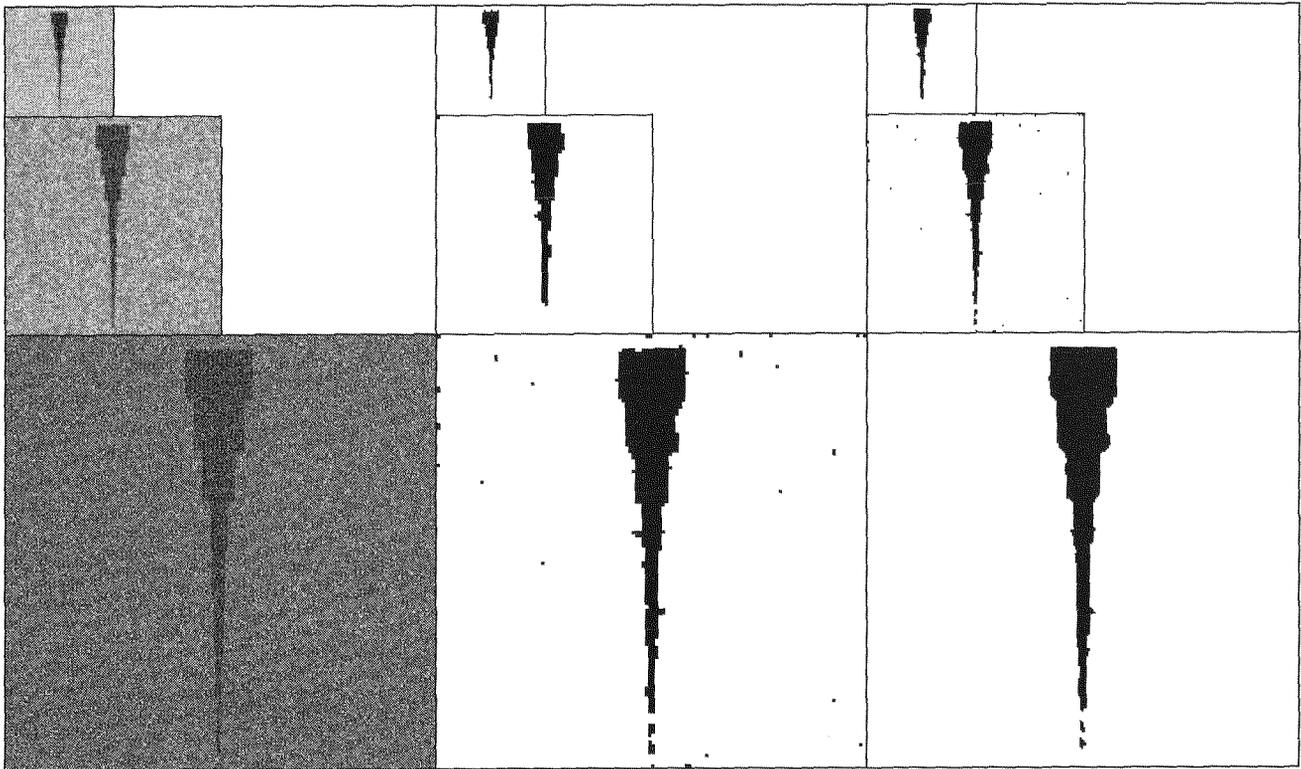


Niveau de résolution	Paramètres du terme d'attache aux données	
	Zone « ombre »	Zone « fond »
2	Gauss : $\mu_0 = 68$ et $\sigma_0 = 84$	Gauss : $\mu_1 = 154$ et $\sigma_1 = 26$
1	Gauss : $\mu_0 = 75$ et $\sigma_0 = 23$	Gauss : $\mu_1 = 114$ et $\sigma_1 = 25$
0	Gauss : $\mu_0 = 118$ et $\sigma_0 = 27$	Gauss : $\mu_1 = 132$ et $\sigma_1 = 29$

Paramètres pondérateurs du terme contextuel											
$\beta_{20}$	$\beta_{21}$	$\beta_{22}$	$\beta_{23}$	$\beta_{40}$	$\beta_{41}$	$\beta_{42}$	$\beta_{43}$	$\beta_{44}$	$\beta_{45}$	$\beta_{46}$	$\beta_{47}$
1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	2	0	0	0	0	0.35	0	0.35

	Nb. pixels « ombre » non détectés	Nb. pixels « ombre » fausses alarmes	$P_{det}$ en %	$P_{fa}$ en %	Valeur Energie finale	Temps en s.
recuit simulé [gts623]	526	134	96.79	0.273	-610759	3131
monorésolution	94	5659	99.43	11.51	-579936	19.8
multirésolution monomodèle [gts383]	10	1997	99.94	4.06	-620215	22.31
multigrille [gts641]	76	1822	99.54	3.71	-600969	31.8
multirésolution Multimodèle	43	885	99.74	1.80	-483001	40.6

Figure 9. – Valeurs estimées des paramètres pour la segmentation de l'image du carré bruitée par deux lois normales (RSB = -14 dB) en fonction du niveau de résolution. Ici l'analyse multirésolution des données s'effectue sur 3 niveaux. La colonne de gauche représente l'image aux différents niveaux de résolution ; la colonne centrale représente le champ d'étiquettes initial utilisé pour initialiser l'algorithme ICM ; la colonne de droite représente le champ d'étiquettes final après relaxation. Les deux tableaux désignent les paramètres estimés du modèle markovien. Enfin, le dernier tableau donne quelques éléments de comparaisons entre différentes approches markoviennes. La probabilité de détection  $P_{det}$  désigne le nombre de sites convenablement classés « ombre » sur le nombre total de sites « ombre ». La probabilité de fausse alarme  $P_{fa}$  désigne le nombre de sites improprement classés « ombre ». Cinq approches sont ici comparées : l'approche classique monorésolution-monomodèle utilisant un algorithme de recuit simulé, l'approche monorésolution-monomodèle utilisant un algorithme de relaxation sous-optimal ICM, l'approche multirésolution-monomodèle, l'approche multigrille et l'approche hiérarchique que nous proposons, baptisée multirésolution-multimodèle.



Niveau de résolution	Paramètres du terme d'attache aux données	
	Zone «ombre»	Zone «fond»
2	Rayleigh : $i_0 = 0$ et $\alpha_0 = 46$	Gauss : $\mu_1 = 179$ et $\sigma_1 = 15$
1	Rayleigh : $i_0 = 0$ et $\alpha_0 = 66$	Gauss : $\mu_1 = 150$ et $\sigma_1 = 15$
0	Rayleigh : $i_0 = 28$ et $\alpha_0 = 29$	Rayleigh : $i_1 = 51$ et $\alpha_1 = 30$

Paramètres pondérateurs du terme contextuel											
$\beta_{20}$	$\beta_{21}$	$\beta_{22}$	$\beta_{23}$	$\beta_{40}$	$\beta_{41}$	$\beta_{42}$	$\beta_{43}$	$\beta_{44}$	$\beta_{45}$	$\beta_{46}$	$\beta_{47}$
1	1	1	1	0	0	0	0	0	0	0	0
1	1	1	1	0	0	0	0	0	0	0	0
1	3	1	1	0.5	0.5	0	0.5	0	1.2	0	0

	Nb. pixels «ombre» non détectés	Nb. pixels «ombre» fausses alarmes	$P_{det}$ en %	$P_{fa}$ en %	Valeur nergie finale	Temps en s.
recuit simulé[48]	113	27	96.86	0.044	-430650	2698
monorésolution	436	36	87.89	0.058	-429295	18.7
multirésolution monomodèle[18]	546	271	84.83	0.44	-375723	22.01
multigrille[19]	194	151	94.61	0.244	-430188	31.3
multirésolution multimodèle	112	308	96.89	0.497	-625095	39.8

Figure 10. – Valeurs estimées des paramètres pour la segmentation de l'image de la mire bruitée par deux lois de Rayleigh (RSB = -10 dB) en fonction du niveau de résolution. La difficulté réside ici dans la finesse de l'ombre portée associée à cet objet conique qui constitue la mire. Les temps de calcul présentés ici ont été obtenus sur station IBM43p-120MHz et sont fournis pour donner un ordre de grandeur du temps nécessaire pour les différentes approches. On constate par exemple que le temps de calcul associé à l'approche proposée (multirésolution- multimodèle) est du même ordre de grandeur que les autres approches, hormis celle utilisant un recuit simulé, et que le résultat obtenu est très fidèle en terme de  $P_{det}$ .

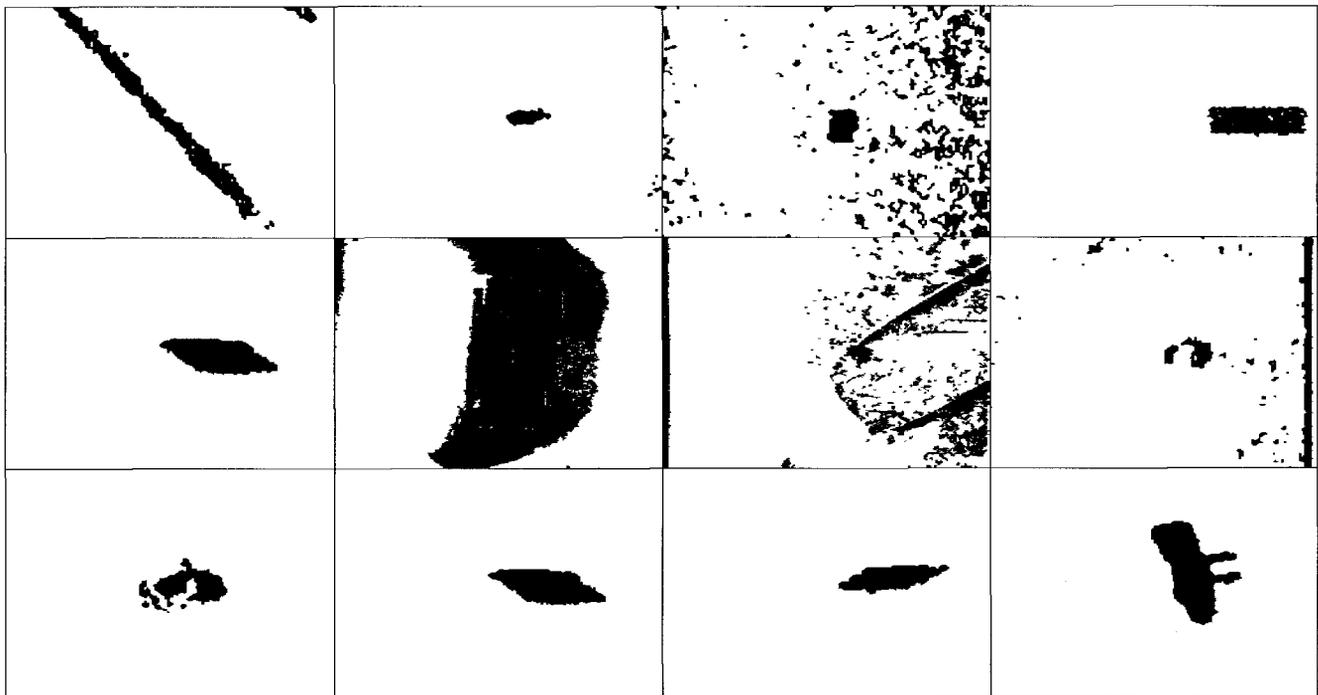


Figure 11. – Résultats obtenus sur la mosaïque d'images, par l'algorithme de segmentation markovien multirésolution-multiéchelle. On observe de bons résultats : le champ des étiquettes est régularisé tant par l'utilisation de cliques quaternaires à pleine résolution, que par la ré-estimation des paramètres du bruit. La distribution retenue à pleine résolution est en règle générale une loi de Weibull, tant pour les zones d'ombre que pour les zones de fond réverbéré. Expérimentalement le choix du critère ( $\chi^2$  ou Kolmogorov) fournit des décisions similaires.

modèle markovien enrichi d'informations portant sur la forme géométrique de l'ombre que l'on souhaite privilégier (introduction des cliques quaternaires), conduit à un résultat très affiné. L'intérêt d'utiliser des configurations particulières d'étiquettes sur des cliques quaternaires apparaît de façon nette sur les contours obliques des ombres qui sont lissés et plus réguliers que ceux obtenus avec un modèle isotrope.

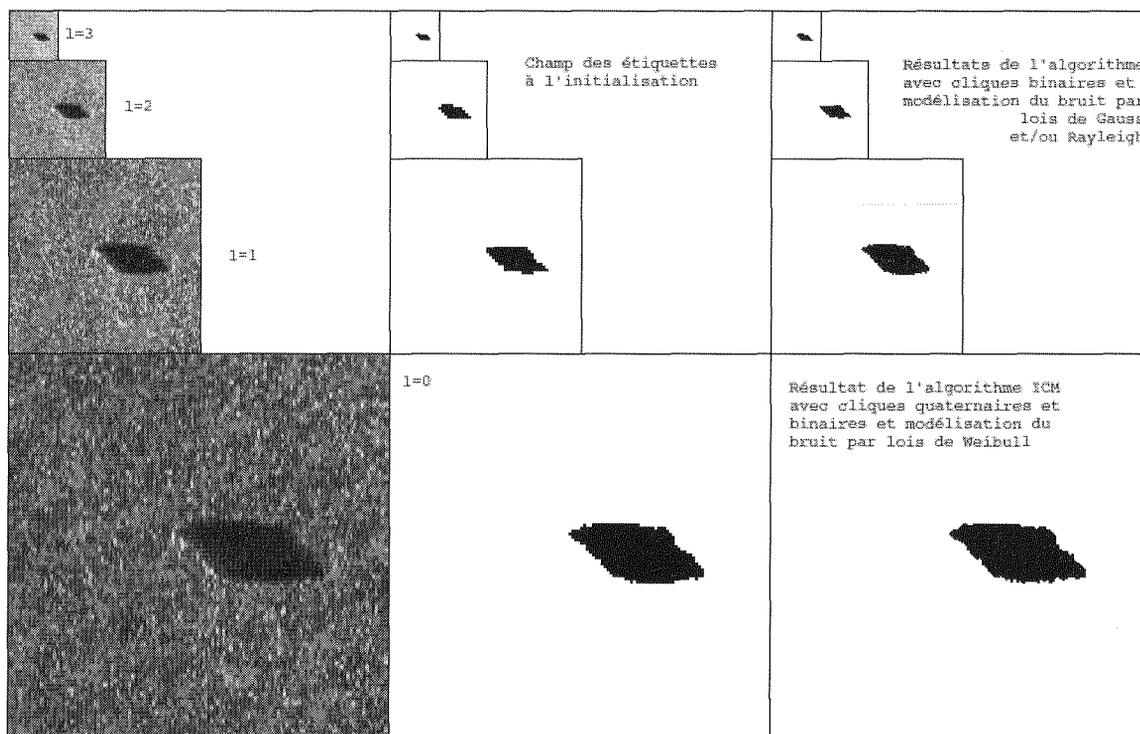
L'enrichissement progressif de la carte des étiquettes au fur et à mesure que la résolution devient plus fine permet de converger vers une carte énergétique des étiquettes proche du minimum global, alors que l'algorithme de relaxation est déterministe. Le paramètre d'attache aux données est estimé au niveau  $l$  grâce au champ d'étiquettes obtenu au niveau  $l - 1$ . Au niveau  $L = 3$ , le champ des étiquettes initial est obtenu par l'algorithme des K-moyennes.

Nous avons choisi de détailler les résultats obtenus pour quatre images.

Ces différents résultats montrent donc l'intérêt de l'approche multirésolution sur les observations, couplée à une modélisation markovienne multimodèle, évoluant en fonction de la résolution. Cette approche, baptisée segmentation markovienne hiérarchique multimodèle permet de rechercher dans des images sonar sévèrement bruitées, l'ombre portée d'un objet géométrique.

## 6. conclusion

En conclusion, l'originalité de la chaîne de segmentation présentée dans cet article repose sur la coopération de deux modèles markoviens associés à une analyse multirésolution des observations. Un premier modèle markovien isotrope, défini pour un système de 4-voisinage, assure la régularisation des zones homogènes des images de résolutions grossières. Avant chaque phase de relaxation par un algorithme ICM, les paramètres de ce modèle sont estimés à partir de l'image multirésolution du niveau considéré, permettant ainsi une meilleure adaptation du modèle aux données. Enfin, à pleine résolution, nous utilisons un deuxième modèle markovien enrichi par la prise en compte de notions de forme géométrique. Initialisé par la projection de l'estimation du niveau précédent, la relaxation fournit alors une image segmentée d'excellente qualité. Le terme d'attache aux données est modélisé par une distribution de Gauss ou de Weibull, cette dernière offrant un cadre plus général que la loi de Rayleigh, dont la justification théorique fut précisée. Le type de loi est déterminé au sens du critère de Kolmogorov ou du  $\chi^2$ ; tandis que l'estimation des paramètres s'effectue au sens du critère de maximum de vraisemblance. La chaîne de segmentation est non supervisée, l'estimation des paramètres du modèle *a priori* s'effectue par la technique des boîtes qualitatives. Des résultats intéressants et nouveaux ont été obtenus et présentés sur des images de synthèse et



Niveau de résolution	Paramètres du terme d'attache aux données	
	Zone « ombre »	Zone « fond »
<b>3</b>	$\mathcal{R} : i_0 = 0 \text{ et } \alpha_0 = 23$	$\mathcal{N} : \mu_1 = 169 \text{ et } \sigma_1 = 23$
<b>2</b>	$\mathcal{R} : i_0 = 19 \text{ et } \alpha_0 = 17$	$\mathcal{N} : \mu_1 = 135 \text{ et } \sigma_1 = 28$
<b>1</b>	$\mathcal{R} : i_0 = 0 \text{ et } \alpha_0 = 15$	$\mathcal{R} : i_1 = 44 \text{ et } \alpha_1 = 108$
<b>0</b>	$\mathcal{W} : i_0 = 17; \alpha_0 = 20.; \beta_0 = 3.5$	$\mathcal{W} : i_1 = 38; \alpha_1 = 5; \beta_1 = 1.8$

Paramètres pondérateurs du terme contextuel												
$\beta_{20}$	$\beta_{21}$	$\beta_{22}$	$\beta_{23}$	$\beta_{40}$	$\beta_{41}$	$\beta_{42}$	$\beta_{43}$	$\beta_{44}$	$\beta_{45}$	$\beta_{46}$	$\beta_{47}$	
1	1	1	1	0	0	0	0	0	0	0	0	
1	1	1	1	0	0	0	0	0	0	0	0	
1	1	1	1	0	0	0	0	0	0	0	0	
2	1.4	1	1	1.4	2	0.26	0.014	0.24	0.04	0	0.4	

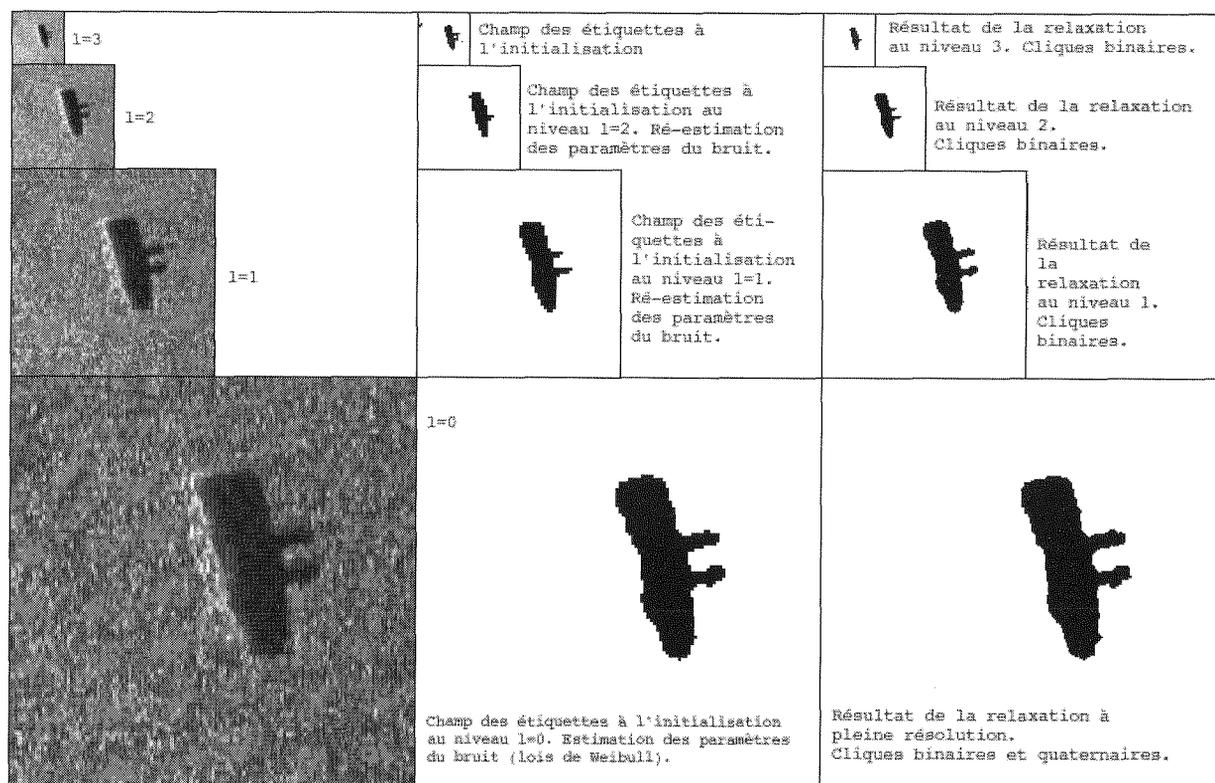
Figure 12. – Valeurs estimées des paramètres pour la segmentation de l'image ci-dessus en fonction du niveau de résolution. La segmentation de cet objet cylindrique s'effectue convenablement depuis les plus faibles résolutions. Les cliques quaternaires interviennent effectivement à la résolution la plus fine. Aucune fausse alarme n'est observée sur les zones de sable avoisinantes. Le choix du type de loi modélisant le bruit de speckle est effectué au sens de la distance de Kolmogorov.

enfin sur des images de sonar haute résolution, permettant ainsi de valider la démarche.

## BIBLIOGRAPHIE

- [1] H. H. Arsenault. Information extraction from images degraded by speckle. In *Proc. of IGARSS'87*, pages 1317–1320, Ann Arbor, 18-21 May 1987.
- [2] R. Azencott. Image analysis and Markov fields. In *Proc. of the Int. Conf. on Industrial and Applied. Math, SIAM*, pages 53–61, Philadelphia, June 1988.
- [3] R. Azencott and C. Graffigne. Non supervised segmentation using multi-level Markov random fields. In *Proc. 11th Intern. Conf. on Pattern Recognition*, pages 201–204, The Hague, The Netherlands, September 1992.
- [4] A. Banerji and J. Goutsias. Detection of minelike targets using grayscale morphological image reconstruction. In *SPIE Detection Technologies for Mines and Minelike Targets*, volume 2496, pages 836–849, Orlando, Florida, 17-21 April 1995.
- [5] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, B-48 : 259–302, 1986.
- [6] L. Blanc-feraud and M. Barlaud. Restauration d'image bruitée par analyse multirésolution et champs de Markov. In *13 ème Colloque sur le Traitement du Signal et des Images (GRETSI)*, pages 829–832, Juan-Les-Pins, septembre 1991.
- [7] C. Bouman and B. Liu. Multiple resolution segmentation of textured images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-13(2) : 99–113, February 1991.

# Segmentation markovienne hiérarchique multimodèle



Niveau de résolution	Paramètres du terme d'attache aux données	
	Zone "ombre"	Zone "fond"
<b>3</b>	$\mathcal{R} : i_0 = 0 \text{ et } \alpha_0 = 21$	$\mathcal{R} : i_1 = 114 \text{ et } \alpha_1 = 30$
<b>2</b>	$\mathcal{R} : i_0 = 0 \text{ et } \alpha_0 = 14$	$\mathcal{N} : \mu_1 = 77 \text{ et } \sigma_1 = 38$
<b>1</b>	$\mathcal{R} : i_0 = 0 \text{ et } \alpha_0 = 14$	$\mathcal{R} : i_1 = 53 \text{ et } \alpha_1 = 42$
<b>0</b>	$\mathcal{W} : i_0 = 10; \alpha_0 = 20.2; \beta_0 = 2.55$	$\mathcal{W} : i_1 = 40; \alpha_1 = 60.2; \beta_1 = 2$

Paramètres pondérateurs du terme contextuel												
$\beta_{20}$	$\beta_{21}$	$\beta_{22}$	$\beta_{23}$	$\beta_{40}$	$\beta_{41}$	$\beta_{42}$	$\beta_{43}$	$\beta_{44}$	$\beta_{45}$	$\beta_{46}$	$\beta_{47}$	
0.5	0.5	0.5	0.5	0	0	0	0	0	0	0	0	
0.5	0.5	0.5	0.5	0	0	0	0	0	0	0	0	
0.5	0.5	0.5	0.5	0	0	0	0	0	0	0	0	
0.5	0.45	0.4	0.4	0.8	0	0	0	0	0.2	0	0.1	

Figure 13. – Valeurs estimées des paramètres pour la segmentation de l'image ci-dessus en fonction du niveau de résolution. L'objet reposant sur le fond marin est un chariot retourné. On observe aux différents niveaux de résolution l'apparition de l'ombre des roues. On observe également le résultat du calibrage par boîtes qualitatives pour les cliques quaternaires à pleine résolution.

[8] M. Bouvet. *Traitements des Signaux pour les Systèmes Sonar*. Collection Technique et Scientifique des Télécommunications. Masson, Paris, 1992.

[9] L. Brekhovskikh and Y. Lysanov. *Fundamentals of Ocean Acoustics*. Springer Series in Electrophysics. Springer-Verlag, 1982.

[10] C. B. Burckhardt. Speckle in ultrasound B-Mode scans. *IEEE Transactions on Sonics and Ultrasonics*, SU-25(1) : pp1-6, January 1978.

[11] W. S. Burdic. *Underwater Acoustic System Analysis*. In Prentice-Hall Signal Processing Series. Prentice-Hall, New Jersey, 1984.

[12] B. Chalmoud. Image restoration using an estimated Markov model. *Signal Processing*, 15(2) : 115-129, September 1988.

[13] P. Charbonnier, L. Blanc-Feraud, and M. Barlaud. Noisy image restoration using multiresolution Markov random fields. *Journal of Visual Communication and Image Representation*, 3(4) : 338-346, December 1992.

[14] R. Chellappa. *Markov Random Fields, theory and application*. Harcourt Brace Jovanovich publishers. Academic Press, Inc, 1993.

- [15] N. P. Choritos. Reflexion and reverberation in normal incidence echosounding. *Journal of Acoustical Society of America*, 96(5) : 2921–2928, November 1994.
- [16] N. P. Chotiros. Ocean bottom acoustic interactions in MCM. In *Proc. of OCEANS'94*, volume 2, pages 250–254, Brest, France, 13-16 September 1994.
- [17] J.P. Cocquerez and S. Philipp. *Analyse d'images : filtrage et segmentation*. Masson - Paris, 1995.
- [18] C. Collet, P. Thourel, P. Bouthemy, and P. Pérez. Détection d'ombres sur images sonar par une analyse multirésolution et une modélisation markovienne. In *15<sup>ème</sup> Colloque sur le Traitement du Signal et des Images (GRETSI)*, pages 569–572, Juan-Les-Pins, septembre 1995.
- [19] C. Collet, P. Thourel, P. Pérez, and P. Bouthemy. Hierarchical MRF modeling for sonar picture segmentation. In *International Conference on Image Processing - ICIP'96*, volume 3, pages 979–982, Lausanne, September 1996.
- [20] Y. Delignon, R. Garelo, and A. Hillion. Etude statistique d'images SAR de la surface de la mer. In *13<sup>ème</sup> Colloque sur le Traitement du Signal et des Images (GRETSI)*, pages 573–575, Juan-Les-Pins, septembre 1991.
- [21] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6) : 721–741, November 1984.
- [22] J. W. Goodman. Some fundamental properties of speckle. *Journal of Optical Society of America*, 66(11) : 1145–1150, November 1976.
- [23] C. Graffigne, F. Heitz, P. Pérez, F. Prêteux, M. Sigelle, and J. Zerubia. Hierarchical Markov random field models applied to image analysis : a review. In *SPIE Neural Morphological and Stochastic Methods in Image and Signal Processing*, volume 2568, pages 2–17, San Diego, 10-11 July 1995.
- [24] F. Heitz, F. Preteux, C. Graffigne, M. Sigelle, and J. Zerubia. Modeles markoviens hiérarchiques pour l'analyse d'images. In *GDR Traitement du signal et des images*, CNRS département SPI - Paris, 1994.
- [25] D. R. Jackson, D. P. Winebrenner, and A. Ishimaru. Application of the composite roughness model to high-frequency bottom backscattering. *Journal of Acoustical Society of America*, 79(5) : 1410–1422, May 1986.
- [26] Z. Kato. *Modélisation markovienne multirésolution en vision par ordinateur. Application à la segmentation d'images SPOT*. Thèse préparée à l'INRIA Sophia Antipolis - version anglaise, Université de Nice - Sophia Antipolis, décembre 1994.
- [27] J. S. Lee and I. Jurkevich. Segmentation of SAR images. volume 27-6, pages 674–680, Proc. of IGARSS'88 - Edinburgh, UK, 12-16 September 1988.
- [28] K. Lim and H. Yang. A new hierarchical mrf modeling for image segmentation and restoration. *ACCV'95*, 1995.
- [29] J. P. Longuemard and A. Goutierre. Diffusion des ondes ultrasonores par le fond marin en fonction de la granulométrie. *Traitement du signal*, 2(1) : 253–258, 1985.
- [30] S. Mallat. A theory for multiresolution signal decomposition : the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7) : 2091–2110, July 1989.
- [31] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(7) : 710–732, July 1992.
- [32] R. L. Martin and R. W. Farwell. High-frequency acoustic modeling. In *Proc. of OCEANS'94*, volume 2, pages 272–277, Brest, France, 13-16 September 1994.
- [33] D. Middleton. A statistical theory of reverberation and similar first-order scattered fields. part 1 : Waveforms and the general process. *IEEE Transaction on Information Theory*, IT-13(3) : 372–392, July 1967.
- [34] 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.
- [35] P. Milanfar, R. R. Tenney, R. B. Washburn, and A. S. Willsky. Modeling and estimation for a class of multiresolution random fields. In *ICIP'94*, pages 397–401, Austin, Texas, 13-16 November 1994.
- [36] K. M. Nam, D. J. Park, and R-H. Park. Multiresolution edge detection in speckle imagery. *Optical Engineering*, 34(3) : 922–932, March 1995.
- [37] Papoulis. *Probability Random, Variables, and Stochastic Processes*, pages 100–105 and 138–149 and 194–197 and 498–503. Mc Graw-Hill - 1991.
- [38] P. Pérez. *Champs markoviens et analyse multirésolution de l'image : application à l'analyse du mouvement*. Thèse, Université de Rennes I, Juillet 1993.
- [39] P. Pérez and F. Heitz. Une approche multiéchelle à l'analyse d'images par champs Markoviens. *Traitement du signal*, 9(6) : 459–472, 1992.
- [40] E. Pouliquen and X. Lurton. Identification de la nature du fond de la mer à l'aide de signaux d'échosondeurs : Modélisation d'échos réverbérés par le fond. *ACTA ACUSTICA*, 2(2) : 113–126, April 1994.
- [41] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes in C. The Art of Scientific Computing*. Cambridge University Press, 2 edition, 1992.
- [42] F. Prêteux and X. Descombes. Synthèse et analyse de textures par coopération de processus multi-échelles. In *8<sup>ème</sup> congrès RFIA*, pages 1015–1023, Lyon-Villeurbanne, novembre 1991.
- [43] C. Regazzoni, F. Arduini, and G. Vernazza. A multilevel gmrf-based approach to image segmentation and restoration. *Signal Processing* 34, pages 43–67, 1993.
- [44] J. O. A. Robertsson and A. Levander. A numerical study of seafloor scattering. *Journal of the Acoustical Society of America*, 97(6) : 3532–3546, June 1995.
- [45] G. Saporta. *Probabilités. Analyse des données et statistique*. Technip - Paris, 1990.
- [46] R.A. Schowengerdt. *Techniques for Image Processing and Classification in Remote Sensing*. Academic Press - New York, 1983.
- [47] M. Sigelle and R. Ronfard. Modèles de Potts et relaxation d'images de labels par champs de Markov. *Traitement du signal*, 9(6) : 449–458, 1992.
- [48] P. Thourel. *Segmentation d'images sonar par modélisation markovienne hiérarchique et analyse multirésolution*. Thèse de doctorat, Université de Bretagne Occidentale, juillet 1996.
- [49] P. Thourel, C. Collet, P. Bouthemy, and P. Pérez. Multiresolution analysis and MRF modeling applied to the segmentation of shadows in sonar pictures. In *Second Asian Conference on Computer Vision ACCV'95*, volume II, pages 81–85, Singapore, December 1995.
- [50] D. Tretter, C. A. Bouman, K. W. Khawaja, and A. A. Maciejewski. A multiscale stochastic image model for automated inspection. *IEEE Transactions on Image Processing*, 4(12) : 1641–1654, December 1995.
- [51] R. J. Urlick. *Principles of Underwater Sound*. Mc Graw-Hill Book Company, 3d edition, 1983.
- [52] M. Zaman and C. Moloney. A comparison of adaptive filters for edge-preserving smoothing of speckle noise. In *Image and Multidimensional Signal Processing Proceedings - ICASSP'93*, volume 5, pages 77–80, Minneapolis, USA, April 1993.
- [53] B. Zerr, E. Maillard, and D. Gueriot. Sea-floor classification by neural hybrid system. In *Proc. of OCEANS'94*, volume 2, pages 239–243, Brest, France, 13-16 September 1994.

Manuscrit reçu le 22 avril 1997.

## Segmentation markovienne hiérarchique multimodèle

### LES AUTEURS

#### Christophe COLLET



Christophe Collet a obtenu son doctorat à l'université de Toulon en 1992. Ses travaux de thèse ont été menés en collaboration avec l'École Nationale Supérieure des Télécom de Bretagne (ENST-Br), au sein du laboratoire Groupe de Traitement du Signal de l'École navale sur les problèmes d'estimation de la cinématique des masses nuageuses en imagerie infrarouge. Il est actuellement maître de conférences à l'Université de Bretagne Occidentale, détaché à l'École navale. Il dirige le laboratoire GTS depuis 1994. Ses centres

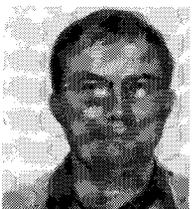
d'intérêt en recherche concernent la modélisation statistique, la segmentation markovienne non supervisée, le traitement des formes, l'analyse de textures en imagerie sonar, les techniques d'optimisation en traitement d'images.

#### Max MIGNOTTE



Max Mignotte est titulaire du DEA Signal-Image-Parole de l'INP de Grenoble (1993). Il poursuit actuellement une thèse de doctorat au Groupe de Traitement du Signal (GTS) de l'École Navale sur la segmentation markovienne hiérarchique non supervisée en imagerie sonar. Ses activités de recherche concernent la segmentation d'images, l'estimation des paramètres, les modèles markoviens hiérarchiques, les modèles statistiques et les algorithmes génétiques.

#### Patrick BOUTHEMY



Patrick Bouthemy est ingénieur Télécom-Paris, Docteur-Ingénieur et Habilité à Diriger des Recherches de l'Université de Rennes 1. Il est actuellement Directeur de Recherche Inria à l'Irisa, et responsable du projet Vista. Ses axes principaux de recherche sont les suivants : analyse du mouvement et des déformations, suivi temporel, modèles statistiques pour l'analyse d'image, vision dynamique et vision active, indexation vidéo.

#### Pierre THOUREL



Pierre Thourel est né à Toulouse en 1960. Diplômé de l'École navale, il est officier de marine depuis 1984. Il est titulaire du DEA d'électronique de l'Université de Bretagne Occidentale (option Traitement du signal - 1993). Il a obtenu son doctorat en 1996 dans le cadre d'un travail de recherche effectué au sein du laboratoire Groupe de Traitement du Signal de l'École navale sur la détection des mines en imagerie sonar. Son travail de thèse porte sur la segmentation

hiérarchique et analyse multirésolution. Ses domaines de recherche comprennent la modélisation markovienne, le traitement multirésolution des images ainsi que l'analyse hiérarchique et multigrille.

#### Patrick PEREZ



Patrick Pérez est né en 1968. Diplômé de l'école Centrale Paris en 1990, il a effectué sa thèse à l'Irisa/Inria-Rennes. Après un an de séjour post-doctoral au Dpt of Applied Mathematics à Brown University (Rhode-Island, USA), il a rejoint l'INRIA en 1994 en tant que chargé de recherche. Ses thèmes de recherches portent sur les modèles statistiques pour le traitement de problèmes inverses de grande dimension, en analyse d'images. Les domaines d'applications privilégiés sont l'analyse du mouvement appa-

rent dans les séquences d'images d'une part, et la fusion multirésolution d'autre part (en télédétection et en imagerie sonar).