

# Estimation du nombre de sommets avec la méthode du Contour Actif Statistique Polygonal

Olivier RUCH (\*), Philippe RÉFRÉGIER

Institut Fresnel, Équipe Physique et Traitement de l'Image,  
École Nationale Supérieure de Physique de Marseille,

Domaine Universitaire de Saint-Jérôme 13397 Marseille cedex 20, France

(\*) : également avec Thalès Optronique (TOSA), Rue Guynemer, BP 55, 78283 Guyancourt cedex, France

olivier.ruch@fresnel.fr

## Résumé –

Nous proposons d'adapter une technique basée sur le principe de description de longueur minimale (MDL) afin d'estimer la complexité - proportionnelle au nombre de sommets - du polygone employé pour effectuer la segmentation avec des Contours Actifs Statistiques Polygonaux. Nous montrons qu'il est possible d'estimer efficacement ce nombre de sommets, à condition qu'une stratégie d'optimisation en deux étapes soit mise en œuvre. Nous obtenons ainsi une méthode de segmentation rapide reposant sur l'optimisation d'un critère sans paramètre libre.

## Abstract –

We propose to adapt a technique based on the Minimum Description Length principle in order to estimate the complexity (proportional to the number of nodes) of the polygon used for the segmentation with Polygonal Statistical Active Contours. We demonstrate that, provided that a 2-step strategy is implemented, it is possible to estimate efficiently this number of nodes without a priori knowledge and with a fast algorithm leading to a segmentation criterion without free parameters.

## 1 Introduction

Une méthode par contour actif statistique, permettant d'effectuer la segmentation d'un objet dans une image bruitée, a récemment été proposée [1, 2] (méthode dite du Contour Actif Statistique Polygonal). Son principe est le suivant : le contour est décrit par un polygone dont les sommets se déplacent, au cours de la convergence, de façon à optimiser un critère statistique, adapté au modèle de bruit considéré dans l'image. Cette méthode de segmentation est rapide (typiquement, moins de 200 ms sur une image de taille  $256 \times 256$ ) et possède peu de paramètres libres. Néanmoins, le principal obstacle auquel l'on est confronté lors d'une utilisation automatique de cet algorithme est le choix du nombre de sommets décrivant le contour polygonal ; si la forme est simple, peu de sommets suffisent et le fait d'en considérer un nombre trop important peut entraîner une diminution de la qualité de la segmentation. Au contraire, si la forme est complexe il est nécessaire de considérer un nombre de sommets important afin d'obtenir une bonne précision du contour.

Nous proposons d'adapter une technique basée sur le principe de description de longueur minimale (principe MDL) afin d'estimer ce nombre de sommets.

## 2 Rappel sur le contour actif

Pour simplifier les expressions mathématiques, une notation uni-dimensionnelle est adoptée ; les valeurs des niveaux de gris des pixels d'une image quelconque sont

alors représentées sous forme d'un vecteur en utilisant la notation lexicographique. Si l'on suppose que la scène  $\mathbf{s}$  est composée de  $N$  pixels, on écrira alors :  $\mathbf{s} = \{s_i | i \in [1, N]\}$ . Le modèle d'image adopté repose sur l'hypothèse que les niveaux de gris de la région *objet* (notés  $\mathbf{a}$ ) et ceux de la région *fond* (notés  $\mathbf{b}$ ) sont des champs aléatoires correspondant à des réalisations de variables aléatoires indépendantes et qui sont respectivement distribuées avec des lois de probabilité de densités  $P_{\theta_a}(a)$  et  $P_{\theta_b}(b)$ . Un contour polygonal d'ordre  $k$ , c'est-à-dire possédant  $k$  sommets, définit une fonction binaire notée  $\mathbf{w}^k$  ( $\mathbf{w}^k = \{w_i^k | i \in [1, N]\}$ ) de sorte que  $w_i^k$  prenne la valeur 1 à l'intérieur du polygone et 0 à l'extérieur. Ce contour divise ainsi l'image en deux régions : la région  $\Omega_a$  intérieure au contour ( $\Omega_a = \{i \in [1, N] | w_i^k = 1\}$ ) et la région complémentaire  $\Omega_b$  ( $\Omega_b = \{i \in [1, N] | w_i^k = 0\}$ ). Les paramètres  $\theta_a$  et  $\theta_b$  ne sont pas supposés connus; leur valeur peuvent être estimées au sens du maximum de vraisemblance sur les deux régions  $\Omega_a$  et  $\Omega_b$  et sont alors respectivement notées  $\hat{\theta}_a(\mathbf{s}, \mathbf{w}^k)$  et  $\hat{\theta}_b(\mathbf{s}, \mathbf{w}^k)$ . Le Contour Actif Statistique Polygonal consiste ainsi à estimer la forme  $\hat{\mathbf{w}}^k$  de l'objet présent dans la scène en maximisant la pseudo-vraisemblance (quelquefois également dénommée vraisemblance généralisée)  $L[\mathbf{s} | \hat{\mathbf{w}}^k, \hat{\theta}_a(\mathbf{s}, \hat{\mathbf{w}}^k), \hat{\theta}_b(\mathbf{s}, \hat{\mathbf{w}}^k)]$  pour un contour polygonal d'ordre  $k$  donné.

Il est souvent pratique de considérer un critère  $J^s$  qui est égal, à une constante additive sans intérêt près, à l'opposé de la pseudo vraisemblance. L'expression de ce critère dépend bien entendu de la loi de probabilité considérée. Pour les lois de probabilités gamma et gaussienne,

on a [2] :

$$J_{gamma}^s(\mathbf{s}, \mathbf{w}^k) = N_1 \log \hat{m}_1 + N_2 \log \hat{m}_2, \quad (1)$$

$$J_{gauss}^s(\mathbf{s}, \mathbf{w}^k) = N_1 \log \hat{\sigma}_1^2 + N_2 \log \hat{\sigma}_2^2, \quad (2)$$

où  $\log$  correspond au logarithme népérien. On a noté  $N_l$  le nombre de pixels dans la région  $\Omega_l$  ;  $\hat{m}_l = \frac{1}{N_l} \sum_{i \in \Omega_l} s_i$  et  $\hat{\sigma}_l^2 = \frac{1}{N_l} \sum_{i \in \Omega_l} s_i^2 - \left\{ \frac{1}{N_l} \sum_{i \in \Omega_l} s_i \right\}^2$  ( $l$  prenant les valeurs  $a$  et  $b$ ) correspondent respectivement aux estimées de la moyenne et de la variance de la loi dans chaque région.

Il est possible de réaliser la minimisation du critère  $J^s$  par un processus de déformation itératif et déterministe du polygone  $\mathbf{w}^k$ . Ce processus est très simple et permet d'obtenir de bons résultats avec un faible temps de convergence. Il consiste à considérer les sommets du contour polygonal successivement (par exemple dans le sens trigonométrique), autant de fois que nécessaire. Chaque sommet considéré au cours d'une itération est alors déplacé avec une amplitude fixée (notée  $\gamma$ ) dans l'une des huit directions associées aux huit plus proches voisins. La déformation retenue est celle permettant d'obtenir la valeur du critère  $J^s$  la plus faible.

Afin de réduire la sensibilité de la méthode à la forme initiale du contour, une stratégie multi-résolution a été proposée [3]. Celle-ci consiste à utiliser, dans un premier temps, un contour initial défini avec un nombre restreint de sommets (typiquement 4). Le contour obtenu après convergence, à l'aide de la technique qui vient d'être décrite, est lui-même utilisé comme initialisation d'une nouvelle étape au cours de laquelle le nombre de sommets définissant le contour est augmenté. Pour la suite, nous avons choisi de considérer cette stratégie avec cinq étapes. Elle sera mise en œuvre en diminuant à chaque étape l'amplitude de déformations  $\gamma$  par un facteur 2 et en augmentant le nombre de sommets de sorte que la distance entre deux sommets consécutifs, également appelée distance inter-sommets (notée  $d$ ), soit divisée par 2.

### 3 Le principe MDL

Le principe MDL [4] (*minimum description length*) est une technique qui a été proposée afin d'estimer l'ordre des modèles statistiques paramétriques. L'estimation de l'ordre d'un modèle est un problème classique en traitement du signal et des images [5, 6]. Il est en effet bien connu que le modèle le plus performant est rarement celui qui utilise le plus de paramètres [6] : au delà d'un certain niveau de complexité, l'ajout de paramètres supplémentaires aboutit à une diminution des performances. Dans notre cas particulier, l'ordre du modèle paramétrique que l'on cherche à estimer correspond au nombre de sommets  $k$  du polygone. Le contexte statistique du contour actif décrit précédemment se prête particulièrement bien à l'application du principe *MDL*. Ce principe a d'ailleurs été appliqué avec succès dans [7] pour l'estimation du nombre de points de contrôle d'une B-spline, dans le cas de formes simples, sans point anguleux. Cependant, comme nous le montrerons dans la suite, dans notre cas l'approche développée dans [7] - appelée processus d'optimisation à

résolution croissante - n'est pas directement applicable. Nous proposons donc un processus plus complexe en deux étapes.

#### 3.1 Application du principe MDL aux contours actifs statistiques polygonaux

Le principe MDL consiste à minimiser le nombre de bits, noté  $\Delta$ , nécessaire pour décrire l'image. L'image étant constituée de trois éléments (la région *objet*, la région *fond* et le contour), cette longueur correspond à la somme de trois termes : la longueur  $\Delta_a$  de la description des niveaux de gris de la région *objet*, la longueur  $\Delta_b$  de la description des niveaux de gris de la région *fond* et la longueur  $\Delta_{w^k}$  de la description du polygone  $w^k$  comportant  $k$  sommets. On a ainsi la relation suivante :

$$\Delta = \Delta_a + \Delta_b + \Delta_{w^k}. \quad (3)$$

La théorie de l'information nous permet de déterminer une approximation de ces trois termes [8, 9]. Commençons par donner une approximation très simple, mais suffisante pour la suite, du terme  $\Delta_{w^k}$ . Chacun des sommets du polygone  $w^k$ , peut être situé sur l'un des  $N$  pixels de l'image. Nous considérerons simplement qu'il y a de l'ordre de  $N^k$  polygones différents comportant  $k$  sommets. Si l'on considère que chacun de ces polygones est équiprobable, le nombre de bits nécessaire pour coder l'un d'entre eux est alors de l'ordre de  $\log_2(N^k)$  (où  $\log_2$  correspond au logarithme en base 2). Déterminons maintenant une approximation des termes  $\Delta_a$  et  $\Delta_b$ . Shannon a montré que le nombre minimal moyen de bits nécessaire pour coder un ensemble d'événements aléatoires correspond à l'entropie de la loi de probabilité qui lui est associée [8]. Le nombre minimal moyen de bits nécessaire pour coder  $N_a$  variables aléatoires distribuées avec une loi de probabilité de densité de probabilité  $P_{\theta_a}$  est ainsi égal au produit de  $N_a$  par l'entropie  $S_a$  de la loi de probabilité. On peut alors écrire  $\Delta_a \simeq N_a S_a$ . L'entropie  $S_a$  est donnée par la relation suivante :  $S_a = - \int P_{\theta_a}(x) \log_2 [P_{\theta_a}(x)] dx - \log_2(q)$  où  $q$  correspond au pas de quantification. Puisque la contribution du terme  $\log_2(q)$  consiste uniquement à ajouter une grandeur constante à la longueur totale de la description, nous ne le considérerons pas dans la suite. De façon similaire, on montre que  $\Delta_b \simeq N_b S_b$ . La longueur totale de la description est alors :

$$\Delta \simeq N_a S_a + N_b S_b + k \log_2(N). \quad (4)$$

En faisant l'approximation classique qui consiste à remplacer la moyenne statistique par la moyenne empirique, on obtient  $N_l S_l \simeq - \sum_{i \in \Omega_l} \log_2 [P_{\theta_l}(a_i)]$  avec  $l = a$  ou  $b$ . Il est alors facile de voir que  $-N_a S_a - N_b S_b$  correspond à la log-vraisemblance (en base 2)  $\ell_2[\mathbf{s} | \mathbf{w}^k, \theta_a, \theta_b]$  de l'hypothèse que le contour polygonal  $\mathbf{w}^k$  corresponde à la silhouette de l'objet présent dans la scène  $\mathbf{s}$ . Dans le cas où les paramètres  $\theta_a$  et  $\theta_b$  sont inconnus, on considèrera une seconde approximation qui consiste à remplacer leur valeur par leur estimée au sens du maximum de vraisemblance. On considèrera donc que :  $\ell_2[\mathbf{s} | \mathbf{w}^k, \hat{\theta}_a(\mathbf{s}, \mathbf{w}^k), \hat{\theta}_b(\mathbf{s}, \mathbf{w}^k)]$  constitue une approximation

de  $-N_a S_a - N_b S_b$ . Cette pseudo-vraisemblance correspond à l'opposé du critère de segmentation  $J^s(\mathbf{s}, \mathbf{w}^k)$  qui est minimisé par le Contour Actif Statistique Polygonal. Dans notre cas, le principe MDL consiste donc à minimiser un critère, noté  $J^{MDL}(\mathbf{s}, \mathbf{w}^k, k)$ , correspondant à la longueur  $\Delta$  de la description de l'image et qui s'écrit de la façon suivante :

$$J^{MDL}(\mathbf{s}, \mathbf{w}^k, k) = \Delta(\mathbf{s}, \mathbf{w}^k, k) = J^s(\mathbf{s}, \mathbf{w}^k) + k \log(N). \quad (5)$$

Les expressions du critère  $J^{MDL}$  obtenues avec les lois de probabilité considérées s'écrivent donc :

$$J_{gamma}^{MDL}(\mathbf{s}, \mathbf{w}^k, k) = N_1 \log m_1 + N_2 \log m_2 + k \log N, \quad (6)$$

$$J_{gauss}^{MDL}(\mathbf{s}, \mathbf{w}^k, k) = N_1 \log \sigma_1^2 + N_2 \log \sigma_2^2 + k \log N, \quad (7)$$

où l'on a utilisé les mêmes notations que celles employées dans les expressions 1 et 2. Le contour polygonal  $\widehat{\mathbf{w}}^k$  qui minimise ce critère correspond au contour approximant la silhouette de l'objet avec un nombre de sommets optimal au sens de la description stochastique de l'image [4]. Il sera appelé *contour polygonal de description de longueur minimale* et, par souci de simplification, *contour MDL*.

La minimisation du critère  $J^{MDL}$  est un problème difficile car l'optimisation porte à la fois sur le contour  $\mathbf{w}^k$  et sur le nombre de sommets  $k$  de ce contour. Dans la suite, nous allons montrer que l'utilisation d'un processus simple d'optimisation à résolution croissante analogue à ce qui est réalisé dans [7] ne permet pas d'obtenir une segmentation satisfaisante.

### 3.2 Echec du processus d'optimisation à résolution croissante

Dans le cadre de notre étude, le processus d'optimisation à résolution croissante consiste à approximer la silhouette de l'objet par plusieurs contours polygonaux  $\widehat{\mathbf{w}}^k$ , obtenus chacun après convergence d'un contour actif statistique pour une valeur du nombre de sommets  $k$  fixée. Il est alors possible, pour ces valeurs de  $k$ , d'évaluer le critère  $J^{MDL}(\mathbf{s}, \widehat{\mathbf{w}}^k, k)$  et d'en déduire le contour MDL  $\widehat{\mathbf{w}}^k$  qui minimise ce critère.

Afin d'illustrer que ce processus n'est pas satisfaisant dans le cas du Contour Actif Statistique Polygonal, nous avons appliqué cette méthode sur l'image du bateau présentée sur la figure 1(b). Celle-ci a été synthétisée à partir d'un contour défini par 10 sommets (figure 1(a)), à l'aide de lois de probabilité gamma, selon le modèle d'images présenté précédemment. Avec ce processus direct, on constate sur la figure 1(c) que la segmentation obtenue n'est pas satisfaisante. De plus, le nombre de sommets estimé ( $k = 9$ ) est inférieur à la valeur vraie ( $k_v = 10$ ) et l'on peut observer sur la courbe en trait pointillé de la figure 2 que le minimum associé à cette valeur est peu marqué. Nous avons donc proposé un processus plus complexe permettant d'obtenir une bonne qualité de segmentation.

### 3.3 Processus d'optimisation proposé

Nous proposons de déterminer le contour MDL en considérant un processus de minimisation en deux étapes. La

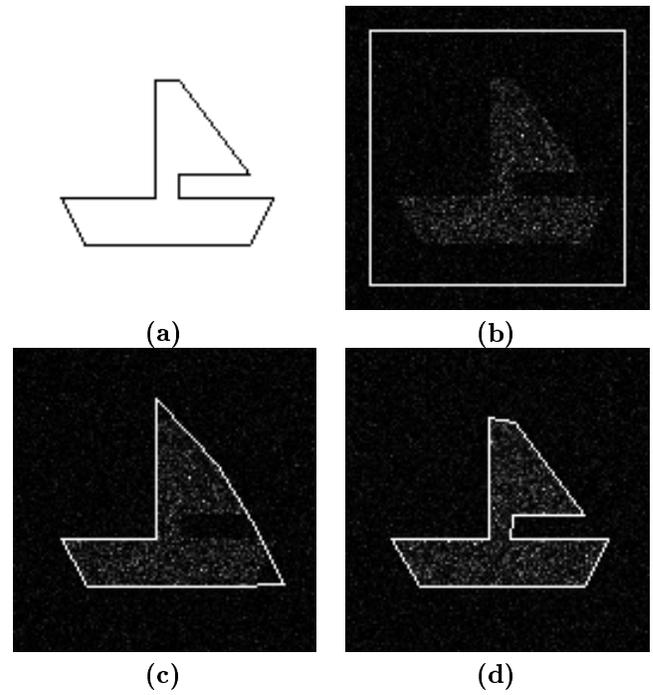


FIG. 1: Echec du processus de minimisation direct avec la méthode du Contour Actif Statistique Polygonal.

(a) : contour polygonal défini par 10 sommets et utilisé pour générer des images synthétiques.

(b) : image synthétique de taille  $128 \times 128$  générée à partir du contour présenté sur la figure (a) avec des lois de probabilités gamma d'ordre 1 dont les moyennes valent 1 dans la région *fond* et 4 dans la région *objet*. En blanc, contour polygonal initial.

(c) : contour MDL obtenu lorsque le nombre de sommets est estimé avec le processus direct ( $k = 9$ ).

(d) : contour MDL obtenu lorsque le nombre de sommets est estimé avec le processus proposé ( $k = 10$ ).

première consiste à extraire le contour de l'objet avec la méthode classique du Contour Actif Statistique Polygonal en considérant un grand nombre de sommets ; on réalise dans un second temps l'*épurage* du contour obtenu en enlevant progressivement des sommets jusqu'à minimiser le critère MDL. Cette seconde étape est réalisée grâce à un algorithme itératif. Pour enlever un sommet du contour, on procède de la façon suivante :

- pour chacun des sommets du contour courant, calculer la valeur du critère MDL obtenue lorsque ce sommet est éliminé ;
- parmi tous ces sommets, éliminer celui qui conduit à la valeur du critère MDL la plus faible ;
- puis, affiner la position des sommets restants en faisant converger le contour obtenu (méthode classique du Contour Actif Statistique Polygonal) ;
- enfin, retenir le contour obtenu après convergence ainsi que la valeur du critère MDL associée.

Le contour MDL obtenu avec ce processus sur l'image de la figure 1(b) est présenté sur la figure 1(d). Il est facile de constater sur la figure 2, que le minimum obtenu dans ce cas est bien marqué et qu'il correspond à la valeur vraie ( $k_v = 10$ ).

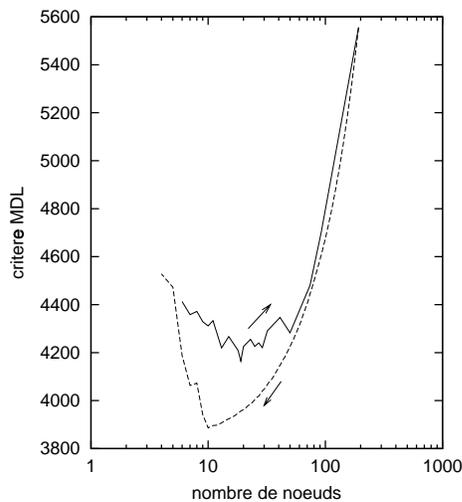


FIG. 2: Trait continu : valeurs du critère MDL obtenues avec le processus d'optimisation croissante. Trait pointillé : valeurs du critère MDL obtenues avec le processus d'optimisation proposé.

## 4 Exemples sur un autre type de bruit

Afin d'illustrer l'efficacité du processus proposé sur un autre type de bruit, nous avons acquis deux images, à l'aide d'une caméra CCD. La première contient deux petits objets alors que la seconde représente un objet plus gros, qui possède un trou. Ces images sont toutes deux de taille  $280 \times 380$  pixels et sont présentées sur la figure 3. Pour

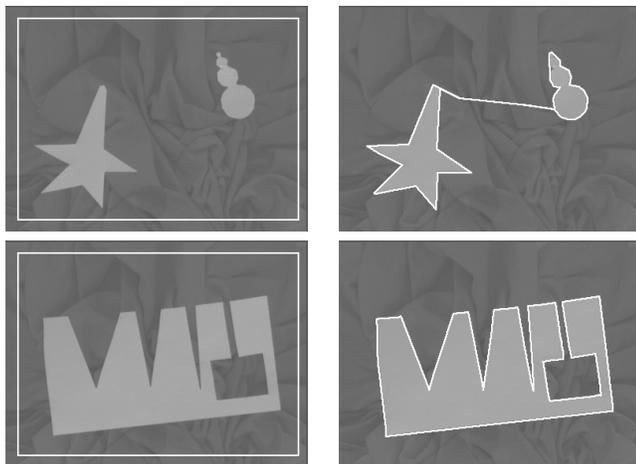


FIG. 3: Segmentation avec le processus proposé de deux images de taille  $280 \times 380$  pixels acquises à l'aide d'une caméra CCD. A gauche : image considérée (contour initial en blanc). A droite : résultat de la segmentation dans le cas du modèle gaussien.

chacune d'elles, nous avons présenté à droite le résultat obtenu avec le processus de minimisation lorsque le modèle gaussien est appliqué (équation 6). On constate que les contours MDL obtenus sont satisfaisants dans les deux cas : dans le premier, les deux objets sont considérés, bien que l'on ait fait l'hypothèse qu'il n'y avait qu'un seul objet

dans l'image et dans le second, le contour rentre dans le trou qu'il y a dans l'objet. Sur ces exemples, les temps de calcul<sup>1</sup> sont respectivement égaux à 560 ms et à 4.1 s.

## 5 Conclusion

En conclusion, nous avons présenté une technique basée sur le principe de description minimale et permettant d'estimer le nombre de sommets nécessaire pour effectuer la segmentation d'un objet dans une image, à l'aide du Contour Actif Statistique Polygonal. Nous avons montré que le processus d'optimisation dit à *résolution croissante* ne permet pas d'obtenir une segmentation satisfaisante et nous avons donc proposé un nouveau processus en deux étapes. Ce processus consiste schématiquement à effectuer, dans un premier temps, une segmentation avec un nombre de sommets très important, puis à éliminer ceux qui sont considérés inutiles. Nous obtenons ainsi une technique de segmentation rapide, reposant sur l'optimisation d'un critère sans paramètre libre.

## Références

- [1] O. Germain and Ph. Réfrégier, "Optimal snake-based segmentation of a random luminance target on a spatially disjoint background," *Opt. Lett.* **21**, 1845–1847 (1996).
- [2] C. Chesnaud, Ph. Réfrégier, and V. Boulet, "Statistical region snake-based segmentation adapted to different physical noise models," *IEEE Trans. on Pattern Analysis and Machine Intelligence* **21**, 1145–1157 (1999).
- [3] C. Chesnaud, V. Pagé, and Ph. Réfrégier, "Robustness improvement of the statistically independent region snake-based segmentation method," *Opt. Lett.* **23**, 488–490 (1998).
- [4] J. Rissanen, *Stochastic Complexity in Statistical Inquiry* (World Scientific, Singapore, 1989).
- [5] G. Schwartz, "Estimating dimension of a model," *Annals of Statistics* **9**, 461–464 (1978).
- [6] C. Olivier, P. Courtellemont, and Y. Lecourtier, "Histogrammes et critères d'information en reconnaissance des formes," in *10ème Congrès RFIA - AFCET, Rennes (France)*, **2**, 1033–1042 (1996).
- [7] Mário A. T. Figueiredo, José M. N. Leitão, and Anil K. Jain, "Unsupervised contour representation and estimation using b-splines and a minimum description length criterion," *IEEE Trans. on Image Processing* **9**, 1075–1087 (2000).
- [8] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.* **27**, 379–423 (1948).
- [9] T. M. Cover and J. A. Thomas, *Elements of Information Theory* (Wiley-interscience, New York, 1991).

<sup>1</sup>évalués sur un PC pentium III, 700 MHz avec 256 Mo de mémoire vive