

Segmentation par contour actif probabiliste

Philippe Réfrégier, Olivier Germain, Christophe Chesnaud et Thierry Gaidon

Laboratoire Signal et Image /E.N.S.P.M.

Domaine universitaire de Saint-Jérôme 13 397 Marseille cedex 20 France

email:refregie@enspm009.u-3mrs.fr

RÉSUMÉ

Nous présentons de nouveaux algorithmes optimaux, au sens de la théorie de la décision, pour la segmentation d'un objet d'intensité aléatoire apparaissant sur un fond *non recouvrant* d'intensité aléatoire. Nous étudions les solutions optimales pour des lois Gaussiennes et Gamma.

ABSTRACT

We propose in this paper new optimal snake-based processors to track the shape of a target with random white intensity appearing on a random white spatially disjoint background. We study the optimal solutions for Gaussian and Gamma laws.

1 Introduction

La segmentation d'un objet est un problème classique en traitement d'images. Il apparaît dans de nombreuses applications : poursuite d'objet dans une séquence d'images, reconnaissance de forme ... La technique récente des contours actifs (*snakes*) s'adapte bien à ce type de problèmes. Les premières approches qui ont été étudiées reposaient essentiellement sur les informations des contours de l'image de telle sorte que la *snake* évolue vers les zones de fort gradient d'intensité [1]. D'autres approches ont été proposées par la suite pour pallier les limitations qui apparaissent quand l'objet à segmenter n'est pas simplement délimité de son contexte par des contours nets. C'est en particulier le cas des approches régions [2] [3] pour lesquelles la *snake* évolue afin de séparer deux régions aux propriétés différentes. Récemment, de nouvelles méthodes optimisant des critères statistiques précis ont été proposées [4]. Nous proposons ici de présenter une nouvelle approche [5] permettant d'obtenir des algorithmes optimaux, au sens de la théorie de la décision, pour la segmentation d'un objet d'intensité aléatoire apparaissant sur un fond d'intensité aléatoire. Après une présentation générale de l'approche, nous analyserons les solutions optimales au sens du maximum de vraisemblance lorsque les statistiques en présence suivent des densités de probabilité Gaussiennes et Gamma dont la pertinence sera discutée. Nous terminerons par l'étude d'algorithmes optimaux au sens du maximum a posteriori.

2 Approche générale

Dans la suite nous utiliserons une notation monodimensionnelle, sans perte de généralité. Les vecteurs seront représentés en caractère gras. La scène $\mathbf{s} = \{s_i | i \in [1, N]\}$ est constituée de N pixels. Soit $\boldsymbol{\omega}$, un vecteur de dimension N représentant le support de la cible, de telle sorte que $\omega_i = 1$ lorsque le pixel numéro i est à l'intérieur de la cible

et $\omega_i = 0$ sinon. $\boldsymbol{\omega}$ partitionne donc l'image en deux régions $S_a = \{i | \omega_i = 1\}$ et $S_b = \{i | \omega_i = 0\}$ constituées respectivement de $N_a(\boldsymbol{\omega})$ et $N_b(\boldsymbol{\omega})$ pixels. Les intensités des pixels de la cible a_i (respectivement du fond b_i) sont la réalisation de variables aléatoires, indépendantes entre elles, caractérisées par leur densité de probabilité (pdf) $P_a^{\boldsymbol{\mu}^a}$ (respectivement $P_b^{\boldsymbol{\mu}^b}$). $\boldsymbol{\mu}^u$ est un vecteur dont les composantes, μ^u_j , sont les paramètres de la pdf P_u (pour une loi normale, les composantes du vecteur $\boldsymbol{\mu}$ seront la moyenne et l'écart type).

Notre modèle d'images peut donc être décrit par :

$$s_i = a_i \omega_i + b_i(1 - \omega_i)$$

L'objectif de la segmentation est alors l'estimation de la forme $\boldsymbol{\omega}$ de la cible dans la scène. Si nous appelons $H_{\boldsymbol{\omega}}$ l'hypothèse qui consiste à attribuer une certaine forme à la cible, le processeur optimal au sens du *maximum a posteriori* (MAP) est obtenu par la maximisation de la probabilité : $P(H_{\boldsymbol{\omega}}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b | \mathbf{s})$. La loi de Bayes nous permet d'écrire :

$$P(H_{\boldsymbol{\omega}}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b | \mathbf{s}) = P(\mathbf{s} | H_{\boldsymbol{\omega}}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b) \frac{P(\boldsymbol{\mu}^a, \boldsymbol{\mu}^b, H_{\boldsymbol{\omega}})}{P(\mathbf{s})} \quad (1)$$

où $P(\mathbf{s} | H_{\boldsymbol{\omega}}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b)$ est la vraisemblance de \mathbf{s} , c'est-à-dire la probabilité d'observer la scène lorsque la silhouette de la forme est supposé être $\boldsymbol{\omega}$ et que les paramètres sont $\boldsymbol{\mu}^a$ et $\boldsymbol{\mu}^b$. Si nous supposons les hypothèses $H_{\boldsymbol{\omega}}$ équiprobables et indépendantes des vecteurs $\boldsymbol{\mu}^a$ et $\boldsymbol{\mu}^b$ (eux même indépendants entre eux), nous pouvons écrire :

$$P(H_{\boldsymbol{\omega}}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b | \mathbf{s}) = K \times m(\boldsymbol{\omega}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b | \mathbf{s}) \quad (2)$$

où K est une constante et,

$$m(\boldsymbol{\omega}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b | \mathbf{s}) = \prod_{S_a} P_a^{\boldsymbol{\mu}^a}(s_i) P(\boldsymbol{\mu}^a) \times \prod_{S_b} P_b^{\boldsymbol{\mu}^b}(s_i) P(\boldsymbol{\mu}^b).$$

Maximiser la probabilité $P(H_{\boldsymbol{\omega}}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b | \mathbf{s})$ est donc équivalent à maximiser $m(\boldsymbol{\omega}, \boldsymbol{\mu}^a, \boldsymbol{\mu}^b | \mathbf{s})$.

Les divers paramètres, μ^u_j ($u = a$ ou $u = b$) étant inconnus, nous choisissons de les estimer au sens du MAP. Ils doivent alors vérifier l'équation suivante :

$$\frac{\partial m(\omega, \mu^a, \mu^b, s)}{\partial \mu^u_j} \Big|_{\mu^u_j(\omega)} = 0$$

où $\bar{\mu}^a(\omega)$ et $\bar{\mu}^b(\omega)$ sont les valeurs qui maximisent l'équation 2. Le critère à maximiser est alors $m(\omega, \bar{\mu}^a(\omega), \bar{\mu}^b(\omega)|s)$. Dans la suite, nous utiliserons souvent $L(\omega, \mu^a, \mu^b|s)$ qui est le logarithme de $m(\omega, \mu^a, \mu^b)$. L'évolution du *snake* sera régie par la minimisation de :

$$J_{\bar{\mu}^a, \bar{\mu}^b}(\omega, s) = -\ln [m(\omega, \bar{\mu}^a(\omega), \bar{\mu}^b(\omega)|s)].$$

La fonction ω qui minimise ce critère, réalise une segmentation optimale au sens du MAP de la cible.

Nous décrivons ici simplement le contour actif, et donc ω , par la donnée de p nœuds joints entre eux par une interpolation linéaire.

Nous minimisons par la suite le critère précédent à l'aide d'un algorithme stochastique simple (du type Monte Carlo). Il consiste à tirer aléatoirement, suivant une loi uniforme, un nœud, une amplitude et une direction de déformation (limitée par une amplitude de déformation maximale d). Ces tirages aléatoires permettent le déplacement d'un nœud et par conséquent, la définition d'un nouveau support ω' et d'une nouvelle valeur du $J_{\bar{\mu}^a, \bar{\mu}^b}(\omega')$. Ce déplacement sera conservé si cette nouvelle énergie est inférieure à la précédente, il sera rejeté sinon. Cette opération est répétée autant de fois que nécessaire. Nous ne discuterons pas plus la pertinence de cet algorithme dont on peut toutefois souligner que rien n'assure sa convergence vers le minimum global de $J_{\bar{\mu}^a, \bar{\mu}^b}(\omega)$. Un algorithme de type recuit simulé pourrait éventuellement être mieux adapté.

3 Maximum de vraisemblance

Si nous supposons que les densités de probabilité des divers paramètres sont équiprobables, la maximisation de la probabilité *a posteriori* $P(H|\omega, \mu^a, \mu^b|s)$ (équation 1) se ramène à la maximisation de la probabilité *a priori* $P(s|H, \omega, \mu^a, \mu^b)$ (vraisemblance). L'estimation des paramètres se fait alors au sens du maximum de vraisemblance.

Nous allons commencer par déterminer le processeur optimal lorsque les statistiques en présence sont blanches et Gaussiennes. La robustesse de ce processeur sera ensuite illustrée sur une scène réaliste.

Nous nous intéresserons ensuite aux lois Gamma qui sont intéressantes à divers titres. En effet, il est bien connu [6] que les intensités des images speckle sont distribuées suivant une telle densité de probabilité. C'est le cas, en particulier, des images SAR (Synthetic Aperture Radar). De plus, les lois Gamma peuvent être considérées comme une approximation de la statistique de certaines images pré-traitées. En effet, il est classique en traitement d'image d'utiliser la somme des carrés de combinaisons linéaires de pixels. Il est également bien connu, que lorsque le nombre de pixels intervenant dans les combinaisons linéaires augmente, la densité de probabilité de l'image

pré-traitée converge vers une loi Gamma. On peut citer par exemple, l'énergie de certains filtres linéaires comme le Sobel, Prewitt ou le Laplacien.

3.1 Cas Gaussien

Nous supposons ici, que les statistiques en présence sont blanches et Gaussiennes. Les vecteurs paramètres peuvent donc s'écrire : $\mu^a = [m_a, \sigma_a]^T$ et $\mu^b = [m_b, \sigma_b]^T$. La log-vraisemblance s'écrit :

$$\begin{aligned} L^{MV}(\omega, \mu^a, \mu^b, s) &= -N_a(\omega) \ln(\sigma_a) \\ &\quad - \frac{1}{2\sigma_a^2} \sum_{s_a} (s_i - m_a)^2 \\ &\quad - N_b(\omega) \ln(\sigma_b) \\ &\quad - \frac{1}{2\sigma_b^2} \sum_{s_b} (s_i - m_b)^2 \end{aligned}$$

Si nous estimons les paramètres au sens du maximum de vraisemblance, nous obtenons :

$$\bar{m}_u(\omega) = \frac{1}{N_u(\omega)} \sum_{s_u} s_i \quad ; \quad \bar{\sigma}_u^2(\omega) = \frac{1}{N_u(\omega)} \sum_{s_u} [s_i - \bar{m}_u(\omega)]^2$$

En injectant ces estimées dans l'équation précédente et en prenant l'opposé, le critère à minimiser devient :

$$J_\sigma(\omega, s) = N_a(\omega) \ln[\bar{\sigma}_a^2(\omega)] + N_b(\omega) \ln[\bar{\sigma}_b^2(\omega)] \quad (3)$$

Nous pouvons remarquer que ce critère peut être instable. En effet, il diverge lorsque l'estimation de l'une des deux variances est nulle. Dans ce cas nous montrerons dans la section suivante que l'introduction de priors sur les paramètres σ_a et σ_b à estimer permet de pallier simplement cet inconvénient.

3.2 Densité de probabilité Gamma

Supposons maintenant que les intensités des pixels de la cible et du fond sont distribuées suivant des densités de probabilité Gamma, définies pour $x > 0$ par :

$$P[x] = \lambda^\beta x^{\beta-1} \exp[-\lambda x] / \Gamma(\beta).$$

$\Gamma(\beta)$ est la fonction gamma et β et λ sont des réels positifs. Les vecteurs paramètres, définis précédemment, se réduisent ici à des scalaires : $\mu^a = \lambda_a$ et $\mu^b = \lambda_b$. En utilisant la même démarche que précédemment, nous obtenons, $\bar{\mu}^a(\omega) = \beta / \bar{m}_a(\omega)$ et $\bar{\mu}^b(\omega) = \beta / \bar{m}_b(\omega)$. Le critère est alors :

$$J_m(\omega, s) = N_a(\omega) \ln[\bar{m}_a(\omega)] + N_b(\omega) \ln[\bar{m}_b(\omega)]$$

Il est intéressant de noter que $J_m(\omega, s)$ ne dépend pas de β . Cette propriété signifie que le même critère est optimal quelque soit l'ordre des densités de probabilité Gamma utilisées.

3.3 Exemples de résultats

La robustesse du modèle Gaussien est illustrée sur une scène réaliste (figure 1). Sur l'image de gauche nous avons superposé à la scène le *snake* initial. Sur celle de droite, nous

montrons le snake final, après minimisation du critère. Le *snake* utilisé comporte $p = 12$ nœuds, l'amplitude maximale des déformations est $d = 2$ et la convergence s'est effectuée en moins de 2000 itérations.

Sur la figure 2 nous avons synthétisé un avion et un fond par des distributions Gaussiennes corrélées de paramètres, $m_a = m_b = 1$ et $\sigma_a = \sigma_b = 0,5$, et de longueurs de corrélation différentes $l_a = 1$ et $l_b = 2$. Nous avons utilisé le processeur optimal dans le cas des lois Gamma qui a été appliqué sur la scène pré-traitée par un filtre Laplacien (de noyau 3×3 dont toutes les valeurs sont égales à 1 sauf la valeur centrale qui est égale à -8). L'état initial et final du *snake* ont été superposés à la scène sur la figure 2. Les paramètres utilisés sont $p = 12$, $d = 2$ et la convergence s'est effectuée en moins de 2000 itérations.

Sur la figure 3 nous montrons les résultats obtenus sur une image réelle SAR de type ERS1 d'une zone agricole. Les *snakes* initial et final sont superposés à la scène ($p = 4$, $d = 3$, convergence effectuée en moins de 1500 itérations).



FIG. 1 — Scène réaliste. *Snake* initial (image de gauche) et *snake* final (image de droite).

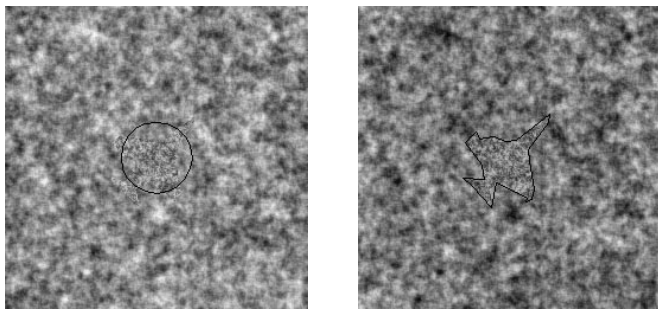


FIG. 2 — La scène est constituée d'un avion et d'un fond modélisés par deux distributions corrélées. Le processeur optimal pour des lois Gamma est appliqué sur l'image pré-traitée par un filtre Laplacien. *Snake* initial (à gauche) et *snake* final après sa convergence sur l'image pré-traitée (à droite).

4 Maximum a posteriori

Dans cette section, nous allons nous limiter au cas Gaussien. Comme nous l'avons déjà signalé dans la section précédente, la solution optimale au sens du maximum de vraisem-

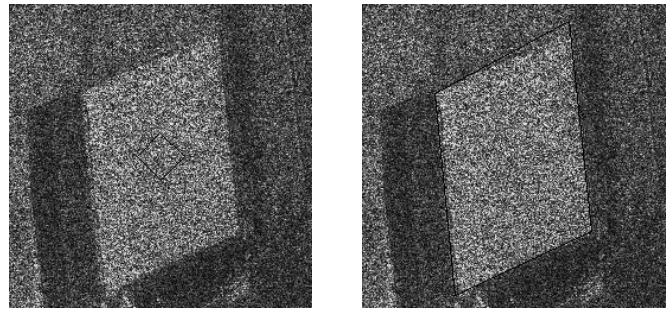


FIG. 3 — Image SAR de type ERS1 d'une zone agricole. *Snake* initial (à gauche) et *snake* final (à droite).

blance peut être instable (équation 3). En effet lorsque l'estimation d'une des deux variances tend vers zéro, le critère diverge. Il peut donc être intéressant d'utiliser un *prior* $P(\sigma)$ qui pénalise les valeurs nulles de l'écart type σ . Cela peut facilement être justifié si on rappelle qu'une image est le résultat d'une quantification. De ce fait, il existe un écart type minimal σ^{min} au dessous duquel l'image numérique est uniforme. C'est pourquoi il est naturel de fortement pénaliser tout les écarts types inférieurs à σ^{min} (via un prior). De même, on peut définir un écart type σ^{max} , écart type maximal que l'on peut obtenir dans une image. La définition de σ^{max} permet seulement d'assurer à $P(\sigma)$ son caractère de densité de probabilité ($\int P(\sigma) d\sigma = 1$).

4.1 Choix d'un prior

Différents choix sont possibles, en particulier les fonctions représentées sur la figure 4. Dans ces divers cas, les solutions

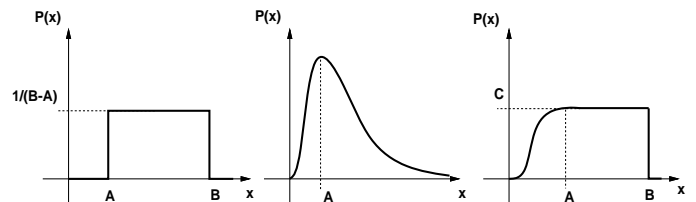


FIG. 4 — Exemples de prior pouvant être utilisé pour l'écart type σ (x sur la figure).

obtenues sont très proches. Nous allons présenter les résultats obtenus lorsque le prior utilisé est la fonction Porte puis lorsqu'il correspond à la fonction Gamma inverse.

4.1.1 Fonction Porte

La fonction porte est représentée à gauche sur la figure 4. Ce prior est défini par : $P(\sigma) = 1/(B - A)$ si $\sigma \in [A, B]$ et $P(\sigma) = 0$ sinon. A et B peuvent correspondre aux écarts types σ^{min} et σ^{max} définis précédemment.

Si on a $\sigma_a, \sigma_b \in [A, B]$ alors $P(H_\omega, \mu_a, \mu_b | s)$ est proportionnelle à la vraisemblance $P(s | H_\omega, \mu)$.

Après calcul, et estimation des divers paramètres au sens du

MAP, nous obtenons le critère :

$$J_\sigma(\omega, \mathbf{s}) = N_a(\omega) \log[(\sigma_a^{MAP})^2] + \theta_a \\ + N_b(\omega) \log[(\sigma_b^{MAP})^2] + \theta_b$$

avec,

$$\begin{cases} (\sigma_u^{MAP})^2 = \bar{\sigma}_u^2 \text{ et } \theta_u = N_u/2 & \text{si } \sigma_u \in [A, B] \\ (\sigma_u^{MAP})^2 = A^2 \text{ et } \theta_u = N_u \bar{\sigma}_u^2 / (2A^2) & \text{sinon} \end{cases}$$

où $\bar{\sigma}_u$ est l'écart type estimé dans la section précédente au sens du maximum de vraisemblance.

4.1.2 Gamma inverse

La loi Gamma inverse est une pdf qui possède les principales qualités de la pdf de la section précédente et qui permet d'obtenir des résultats simples. Cette loi est définie par l'équation :

$$P(x) = \frac{\alpha^2}{x^2} \exp\left\{-\frac{\alpha^2}{x}\right\}.$$

Si nous supposons que la variance σ^2 a pour loi de probabilité *a priori* (prior), une loi Gamma inverse, on a :

$$P(\sigma) = \frac{2\alpha^2}{\sigma^3} \exp\left\{-\frac{\alpha^2}{\sigma^2}\right\}$$

L'utilisation de ce prior, nous permet alors d'obtenir le critère suivant :

$$J_\sigma(\omega, \mathbf{s}) = N_a(\omega) \ln[\bar{\sigma}_a^2 + 2\alpha_a^2] \\ + N_b(\omega) \ln[\bar{\sigma}_b^2 + 2\alpha_b^2] \quad (4)$$

4.2 Simulations

Afin d'illustrer les résultats obtenus avec l'approche au sens du MAP, nous allons segmenter un objet uniforme sur un fond complexe. Comme nous l'avons déjà fait remarquer, le processeur optimal obtenu dans la précédente section diverge lorsqu'une région est uniforme. Plusieurs initialisations ont été réalisées, et en particulier avec une initialisation interne et une externe à la région uniforme. Dans les deux cas, la convergence ne s'est correctement effectuée qu'avec l'approche du MAP. Nous présentons sur la figure 5 les résultats obtenus, pour une initialisation externe, avec le processeur de l'équation 4 ($\alpha_a = \alpha_b = 10^{-3}$, $p = 12$, $a = 2$, convergence effectuée en moins de 2000 itérations).

5 Conclusion

Nous avons présenté une nouvelle approche pour la segmentation de cibles. La modélisation de la cible et du fond par deux champs aléatoires indépendants nous a permis de décrire des situations difficiles. L'estimation des paramètres au sens du maximum de vraisemblance dans un premier temps, puis du maximum a posteriori dans un second, nous a conduit à des critères à optimiser optimaux dans le cadre de la théorie probabiliste de la décision. La robustesse des algorithmes obtenus a été illustrée sur diverses images.

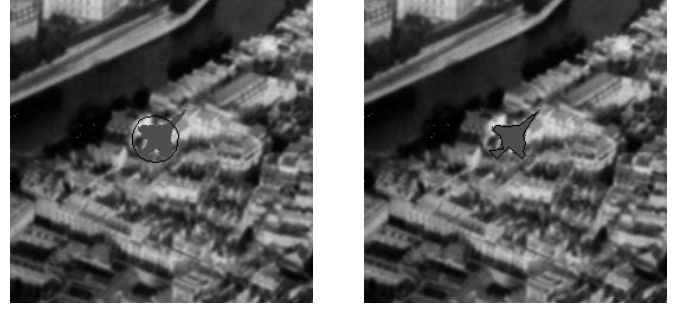


FIG. 5 — Segmentation d'un objet uniforme, grâce à l'utilisation d'un prior.

Remerciements : Les auteurs remercient Madame N. Pourthie et Monsieur D. Massonet du CNES, pour nous avoir fourni et autorisé à publier l'image SAR (figure 3). Ils remercient également V. Pagé, F. Goudail et F. Guérault pour la pertinence de leurs remarques.

Références

- [1] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes : Active contour models," *International Journal of Computer Vision* **1**, 321–331 (1988).
- [2] R. Ronfard, "Region-based strategies for active contour models," *International Journal of Computer Vision* **2**, 229–251 (1994).
- [3] B. Bascle and R. Deriche, "Region tracking through image sequence," *Proc. Conf. Comp. Vision Pattern Recognition (Boston, USA)*, 302–307 (1995).
- [4] C. Kervrann and F. Heitz, "A hierarchical statistical framework for the segmentation of deformable objects in image sequences," *Proc. IEEE Conf. Comp. Vision Pattern Recognition, Seattle*, 724–728 (1994).
- [5] 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).
- [6] J. W. Goodman, *Fourier Optics* (McGraw-Hill Physical and Quantum Electronics Series, San Francisco, 1968).