

DÉTECTION DE CIBLES DANS DES TEXTURES COMPLEXES, UNE APPROCHE PAR SEGMENTATION D'IMAGES EN ZONES STATIONNAIRES

Solenne THIVIN ^{1,2} & Erwan LE PENNEC ³ & Michel PRENAT ²

¹ *Laboratoire de mathématiques d'Orsay , solenne.thivin@math.u-psud.fr*

² *Thales Optronique 2 avenue Gay Lussac 78990 Elancourt ,
[solenne.thivin,michel.prenat]@fr.thalesgroup.com*

³ *Département de Mathématiques appliquées, Ecole Polytechnique ,
Erwan.Le-Pennec@polytechnique.edu*

Résumé. On s'intéresse au problème de la détection, dans des images infrarouges passives, d'objets de petite taille apparente sur des fonds complexes (ciel nuageux, mer, sol), tout en maîtrisant le taux de fausses détections. Pour cela, on propose d'approcher localement le contenu des images réelles par des champs gaussiens corrélés et de prendre en compte l'erreur de modélisation à travers les seuils des tests associés. Plus précisément, on a utilisé des champ de Markov aux 4 plus proches voisins. De plus, pour prendre en compte la possible juxtaposition de zones de textures différentes dans l'image (par exemple ciel clair et nuages), on suppose ces champs stationnaires par morceaux. L'algorithme est alors composé de deux étapes. Dans la première, celle de segmentation, on cherche un découpage de l'image en zones stationnaires. On utilise un critère de log-vraisemblance pénalisé pour lequel on a d'ailleurs obtenu une inégalité oracle. On applique alors indépendamment dans chacune des zones un test du rapport de vraisemblance généralisé. Comme l'hypothèse de champ Markovien n'est pas exacte, on tient compte de l'erreur de modélisation dans le choix du seuil de détection en l'apprenant à partir des données réelles.

Mots-clés. Estimation de densité, Détection, taux de fausses alarmes.

Abstract We address the problem of small objects detection in passive infrared pictures with a complex background (cloudy sky, sea, ground). We have to do this ensuring the false alarm probability. We modelize the real pictures by correlated gaussian fields. More precisely, we use Markov random fields with different neighborhoods. To be able to also explain the pictures with different textures (sky without clouds and with clouds for example), we assume that the field is picewise stationary. The algorithm can be separated in two steps. During the first one, we find a partition of the picture in stationary areas. We use a penalized likelihood criterion for which we obtained an oracle inequality. Then, in a second step, we work in each part of the partition independently. To do the detection, we apply a likelihood ratio test based on the fact that the real law of the area is a Markov one. However, the assumption being false, we learn the actual law in order to adapt the threshold used during the test.

Keywords. Density estimation, detection, false alarm rate.

1 Détection et modèles

L'objectif est de déterminer à partir de l'observation d'une image, ou d'une séquence d'images, si la scène contient ou non une cible, ici un point *contrasté* dans l'image. La grande difficulté du problème est qu'il n'y a pas de modélisation simple de l'image sans la cible.

Plus précisément, on dispose d'une image réelle I de taille $M \times M$ que l'on suppose être une réalisation d'un champ aléatoire 2D: $X_{i,j}$, pour $i \in [1, M]$ et $j \in [1, M]$ auquel s'ajoute éventuellement une cible. En vue de bâtir un test statistique, on cherche à modéliser le champ X . Il semble très utopiste d'utiliser un même modèle pour toute l'image. Les modèles que l'on rencontre dans la littérature sont donc des modèles locaux.

Dans la grande majorité, ces modèles locaux autour d'un point (i, j) sont des modèles gaussiens de la forme

$$Z_{i,j} = (X_{i',j'})_{(i',j') \in \mathcal{V}(i,j)} \sim N(m_{i,j}, \Sigma_{i,j}) \quad (1)$$

où $\mathcal{V}(i, j)$ est un voisinage du point (i, j) . Ils diffèrent par le choix de la moyenne et de la variance.

On peut par exemple, comme dans la thèse d'Emilie Vasquez [Vas11], supposer que la moyenne est un plan et que le bruit est blanc et estimer les paramètres de ce modèle. Ce modèle paramétrique est très pertinent pour des fonds de ciel bleu sans nuages. En revanche, lorsque le patch contient une ou plusieurs zone(s) nuageuse(s), l'algorithme est mis en défaut car le modèle supposé n'est plus vrai. En effet, la "texture" du nuage ne correspond pas à un champ décorréolé mais au contraire à un champ fortement corrélé. Des travaux proposés par Laure Genin [Gen13] lors de sa thèse ont pris en compte cette corrélation. Elle ne fait aucune hypothèses sur $m_{i,j}$ ou $\Sigma_{i,j}$ et les estime de façon non paramétrique.

Une fois les paramètres estimés, on peut alors fabriquer un test de Niemann-Pearson. En effet, ce test est connu pour être le test uniformément le plus puissant si l'on connaît parfaitement le modèle: il garantit le taux de fausses alarmes tout en assurant la plus forte probabilité de détection. Il est donc largement utilisé en pratique.

La garantie de la fausse alarme est une contrainte majeure et nous avons donc décidé d'utiliser nous aussi un test de Niemann-Pearson. En revanche, nous proposons une autre modélisation pour le champ X et nous apprenons les seuils du test à partir de données réelles.

2 Modèle localement stationnaire et partition

Les images réelles sont loin d'être stationnaires donc nous gardons aussi l'idée d'un modèle local. Plus précisément, le modèle que nous proposons est stationnaire par morceaux: il suppose que l'image est constituée de zones indépendantes et dans chacune des zones, on décide de modéliser le contenu des images réelles par un champ de Markov.

Plus particulièrement, on a modélisé l'image comme la somme d'un champ déterministe constant ou plan μ et d'un champ **Auto-Régressif Simultané** [Guy07]. En notant \bar{X} la différence entre X et le champ déterministe μ , on a donc en choisissant un **SAR**(a, σ) isotrope aux 4 plus proches voisins (4 ppvs) de paramètres a et σ :

$$\bar{X}_{i,j} = a(\bar{X}_{i-1,j} + \bar{X}_{i+1,j} + \bar{X}_{i,j-1} + \bar{X}_{i,j+1}) + \sigma\varepsilon_{i,j} \text{ où } \varepsilon_{i,j} \sim \mathcal{N}(0, 1) \quad (2)$$

Concrètement, cela signifie que le comportement du champ recentré \bar{X} au pixel (i,j) est très bien expliqué par une combinaison linéaire de ses 4 plus proches voisins. La figure 1 présente, à gauche, le voisinage utilisé et à droite d'autres tailles de voisinages (8 ou 12 plus proches voisins) envisageables dans la définition du champs **SAR**. Des travaux menés par Nicolas Verzelen mettent en compétition ces différentes tailles de voisinages pour modéliser le champ [Ver09]. On ne prend pas cela en compte ici.

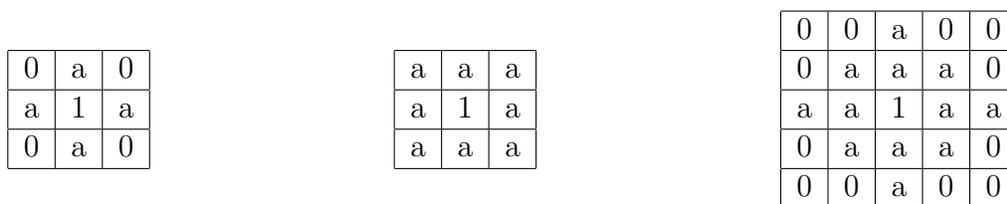


Figure 1: Différents voisinages isotropes

En notant $Z_{i,j} = (X_{i',j'})_{(i',j') \in \mathcal{V}(i,j)}$ le vecteur colonne des valeurs de X au voisinage de (i, j) , on peut vérifier que $Z_{i,j}$ est un vecteur gaussien de moyenne $m_{i,j}$ et de matrice de covariance $\Sigma_{i,j}$ entièrement déterminé par respectivement μ et les paramètres a et σ . Dans le cas des modèles à 4, 8 ou 12 voisins, cette matrice s'exprime simplement comme $\sigma^2 \Sigma(a) = \sigma^2 (Id - aW)^{-2}$ où W est une matrice de voisinage. De manière plus générale, les modèles **SAR** permettent une paramétrisation simple de matrice de covariance à longue portée.

La vraisemblance du modèle local **SAR**(a, σ) est donnée par

$$f_{(a,\sigma,m)}(Z_{i,j}) = \frac{\det(\Sigma(a))^{1/2}}{(2\pi)^{\frac{M^2}{2}}} \exp \left(-\frac{1}{2\sigma^2} {}^t(Z_{i,j} - m) \Sigma(a)^{-1} (Z_{i,j} - m) \right). \quad (3)$$

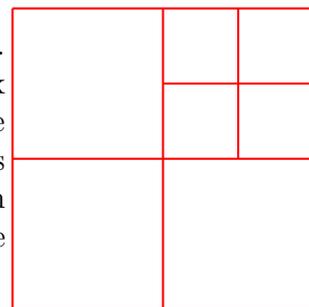
Le principe de maximum de vraisemblance permet d'estimer ces paramètres lorsqu'ils sont inconnus.

À toute partition P de l'image, on peut associer un modèle *par morceaux* (de densité f_P) où l'on suppose que le champ suit un modèle **SAR** sur chaque zone. La dimension de ce modèle est la dimension du modèle local (3 pour un SAR aux 4ppv de moyenne constante) multiplié par le nombre de zones. Pour trouver la *meilleure* partition \hat{P} , on

ne peut se contenter de choisir celle correspondant à la plus grande vraisemblance. Cette stratégie sélectionne toujours des segmentations trop complexes entraînant une grande variance dans l'estimation. On propose pour cela d'utiliser un principe de pénalisation permettant de réaliser automatiquement un bon compromis entre le biais du modèle et la variance de l'estimateur correspondant [Mas03]. Le choix d'une pénalité proportionnelle à la dimension ($\text{pen}(P) = \lambda \dim(P)$) est justifié par une inégalité oracle en terme d'estimation de densité lorsqu'il n'y a pas de cible.

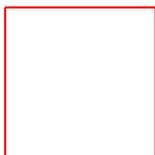
$$E[h^2(\hat{f}_P, f)] \leq C[\inf_P(h^2(f_P, f) + \text{pen}(P))] + K \quad (4)$$

Il est cependant très coûteux de tester toutes les partitions possibles. Ainsi, on décide de restreindre la collection de partitions possibles aux partitions dites "récurives dyadiques" [LPC13]. La figure à droite présente une de ces partitions. Ces partitions sont en fait obtenues de façon réursive (d'où leur nom): partant d'une partition P , on en définit une nouvelle en découpant (ou non) chaque zone de P en quatre nouvelles zones.

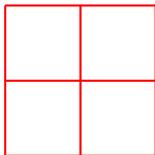


C'est d'ailleurs avec ce même algorithme que l'on va choisir la "meilleure" partition pour une image réelle donnée.

Principe



1. On calcule f_1 la vraisemblance associée à la partition ne contenant qu'une seule zone (figure à gauche).
2. On calcule f_2 la vraisemblance associée à la partition contenant 4 zones rectangulaires (figure à gauche)
3. Si $-\log f_1 + 3\lambda < -\log f_2 + 12\lambda$, on conserve l'hypothèse de stationnarité et l'image n'a donc pas besoin d'être partitionnée. Sinon, l'image n'est pas stationnaire et on relance les étapes 1. et 2. pour chacune des 4 nouvelles zones.



Lorsque l'algorithme s'arrête, on a obtenu la partition associée à l'image.

3 Détection par test du rapport de vraisemblance et calibration

Une fois la segmentation faite et les paramètres dans chaque zone estimés, on propose de construire un test de type rapport de vraisemblance. Comme on cherche à localiser la cible et pas seulement à dire s'il y en a une ou non dans l'image, on fabrique un test local c'est-à-dire un test par pixel où l'on compare les hypothèses

$$\begin{aligned}
H_0 &: "Z_{i,j} \sim \mathcal{N}(m_{i,j}, \sigma^2 \Sigma(a))" \\
H_1 &: "Z_{i,j} \sim \mathcal{N}(m_{i,j} + c\delta_{i,j}, \sigma^2 \Sigma(a))"
\end{aligned}$$

où c représente la puissance de la cible qu'on cherche à détecter et

$$\delta_{i,j}(k, l) = \begin{cases} 1 & \text{si } (k, l) = (i, j) \\ 0 & \text{sinon} \end{cases} \quad (5)$$

On cherche à effectuer la détection à taux de fausses alarmes constant donc on décide d'utiliser un test de type Niemann-Pearson : on définit le rapport de vraisemblance [SF94] au pixel (i, j) par

$$\Delta(Z_{i,j}, i, j) = \frac{p_1(Z_{i,j}, c, m_{i,j}, a, \sigma)}{p_0(Z_{i,j}, m_{i,j}, a, \sigma)} \quad (6)$$

Cependant, comme les paramètres sont inconnus, on les remplace par leur valeur estimée sous chacune des hypothèses. Le rapport de vraisemblance généralisé s'écrit alors dans le cas des 4 ppvs

$$\tilde{\Delta}(Z_{i,j}, i, j) := \frac{p_1(Z_{i,j}, \hat{c}, \hat{m}_1, \hat{a}_1, \hat{\sigma}_1)}{p_0(Z_{i,j}, \hat{m}_0, \hat{a}_0, \hat{\sigma}_0)} = \left(\frac{\hat{\sigma}_0}{\hat{\sigma}_1} \right)^{M^2} \frac{\det(I - \hat{a}_1 W)}{\det(I - \hat{a}_0 W)}. \quad (7)$$

On utilise alors ce calcul pour définir un test sous la forme $\tilde{\Delta}(Z_{i,j}, i, j) > \text{seuil}$. Le seuil est à choisir le plus petit possible parmi ceux garantissant la probabilité de fausse alarme. Pour cela, il faut donc connaître la loi de probabilité de $\tilde{\Delta}(X, i, j)$ [SF94]. Cette loi est inconnue et on utilise des méthodes d'apprentissage pour apprendre le seuil. Dans une extension en cours de ce travail, nous étudions comment la prise en compte de covariables dépendantes de l'image dans la zone permet encore d'améliorer les performances de l'algorithme.

4 Résultats

Comme mentionné plus haut, nous avons réussi à démontrer une inégalité oracle pour notre recherche récursive de partitions (4). Cette inégalité garantit que notre procédure sélectionne une partition proche de celle qui serait choisie si on connaissait la vraie loi du champ.

En ce qui concerne la détection, nous avons mis en place un test de détection et nous pouvons montrer que le partitionnement permet de détecter des cibles qui ne sont pas détectées lorsqu'on suppose l'image stationnaire. La figure 2 illustre comment le partitionnement de l'image peut améliorer la détection.

- La figure tout à gauche montre l'image ainsi que la première segmentation qu'on a utilisée pour la détection (1 seule zone).

- La figure du milieu montre le résultat du test : aucune cible détectée
- La figure de droite montre la seconde segmentation que l'on a utilisée (4 zones) ainsi que le résultat du test de détection pour le quadrant contenant la cible : cette fois, elle est détectée.



Figure 2: Exemple d'amélioration due à la segmentation

On voit ainsi que le fait de réduire la zone supposée stationnaire autour du pixel testé permet d'améliorer l'estimation des paramètres et ainsi de mieux faire ressortir l'anomalie due à la présence de la cible.

Ces résultats sont confirmés par des expériences plus extensives sur des données réelles.

References

- [Gen13] L. Genin. *Détection d'objets de petite taille sur des séquences aériennes ou satellitaires*. PhD thesis, ONERA, 2013.
- [Guy07] X. Guyon. Statistiques spatiales. In *Conférence S.A.D.A.*, 2007.
- [LPC13] E. Le Pennec and S. Cohen. Partition-based conditional density estimation. *ESAIM, probability and Statistics*, 17, 2013.
- [Mas03] P. Massart. *Concentration inequalities and model selection*. Springer, 2003.
- [SF94] L. Scharf and B. Freidlander. Matched subspace detector. *IEEE Transaction On signal processing*, 42(8), 1994.
- [Vas11] E. Vasquez. *Techniques statistiques de détection de cibles dans des images infrarouges inhomogènes en milieu maritime*. PhD thesis, Aix Marseille Université, 2011.
- [Ver09] N. Verzelen. Data-driven neighborhood selection of a gaussian field. *Comput. Statist. Data Anal*, 54(5), 2009.