

DÉTECTION AUTOMATIQUE DE CIBLES SOUS-RÉSOLUES

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

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

² *Thales Optronique , solenne.thivin@fr.thalesgroup.com , michel.prenat@fr.thalesgroup.com*

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

Résumé. Nous cherchons à mettre en place un algorithme de détection de cibles sous-résolues dans des images de ciel infra-rouges. Nous souhaitons prendre en compte la covariance spatiale de l'image pour effectuer la détection. Nous souhaitons aussi effectuer cette détection tout en garantissant le taux de fausses alarmes. Nous avons donc décidé d'appliquer un test de rapport de vraisemblance généralisé. Pour tenir compte du fait que le modèle est probablement faux , nous n'utilisons pas comme seuil le quantile théorique mais des seuils estimés. Enfin, pour tenir compte des différentes textures, nous avons ajouté une étape de classification avant l'estimation des quantiles empiriques classe par classe. Finalement, lors du test de détection, on a appliqué le seuil correspondant à la classe du patch testé.

Mots-clés. Test du rapport de vraisemblance, Estimation parcimonieuse, Classification ...

Abstract. Our work was motivated by a small target detection problem in passive infrared pictures. The real pictures can be modeled locally by (stationary) Gaussian random field. Our principal constraint was to respect the false alarm rate so we decided to apply the likelihood ratio test. However, since the local model was false, we did not use the theoretical thresholds but empirical thresholds estimated on a test sample composed of real patches. Moreover, the textures in the test sample were very different and before estimating the thresholds, we added a classification step. Using a mixture Gaussian model, we classified the patches of the test sample and we then estimated a threshold in each cluster. Finally, in the detection test, we applied the threshold corresponding to the cluster of the tested patch.

Keywords. likelihood ratio test, Sparse estimation, Clustering

Introduction

Le problème général posé lors de ces travaux est le suivant : on souhaite détecter dans des images réelles des cibles de petite taille. Ces images réelles sont des images de ciel fournies par Thales Optronique (TOSA) et acquises à l'aide de caméras infra-rouge. Elles sont de différentes natures. Elles peuvent contenir du ciel bleu uniforme aussi bien que des nuages de textures très variables.

Dans toute la suite, nous travaillons avec une image I de taille $n \times m$. Chaque pixel est représenté par ses coordonnées spatiales (i, j) .

L'image I est vue comme la réalisation d'un champ spatial $X = (X_{i,j} , i \in [1, n] , j \in [1, m])$. Ce champ X est supposé gaussien et sa loi est caractérisée par un vecteur moyenne m_X et une matrice de covariance Σ_X .

1 Modélisation retenue

1.1 Le vecteur moyenne

Des études empiriques sur données réelles ont montré que le vecteur moyenne était non nul mais que le filtrage du champ X par un filtre gradient permettait de centrer le champ. L'image I est donc filtrée pour obtenir une image J vue comme la réalisation d'un nouveau champ Y :

$$\forall i \leq n \forall j \leq m, Y_{i,j} = X_{i,j} - \frac{1}{4}(X_{i-1,j} + X_{i+1,j} + X_{i,j-1} + X_{i,j+1}).$$

Pour les pixels du bord, c'est à dire ceux pour lesquels l'expression précédente n'est pas bien définie, le filtre ne porte alors que sur les voisins bien définis. Par exemple, pour le pixel $(1, 1)$:

$$Y_{1,1} = X_{1,1} - \frac{1}{2}(X_{1,2} + X_{2,1})$$

Dans toute la suite, nous travaillons alors avec le champ Y gaussien centré et de matrice de covariance toujours notée Σ_Y .

1.2 La matrice de covariance

Des études empiriques sur des images réelles ont montré que la covariance spatiale pouvait être considérée comme nulle pour des couples de pixels qui appartiennent à des textures différentes de l'image (un dans une zone de ciel bleu et un dans un nuage par exemple). Cette propriété est confirmée par une approche physique dans laquelle le ciel bleu et le nuage sont vus comme des systèmes physiques indépendants.

Pour modéliser ce phénomène, nous avons donc supposé qu'il existait une partition de l'image en zones telles que des pixels appartenant à des zones différentes étaient indépendants.

En pratique et pour tenir compte de la faible portée de la covariance du champ Y , on a découpé l'image en patches disjoints de taille 16×16 et on a supposé que les pixels appartenant à des patches différents étaient indépendants. La matrice Σ devient alors une matrice par blocs

$$\Sigma = \begin{pmatrix} P_1 & & & & \\ & P_2 & & 0 & \\ & & P_3 & & \\ & 0 & & P_4 & \\ & & & & \ddots \end{pmatrix}$$

où les matrices P_i sont les matrices de covariances de chacun des patches découpés dans l'image I .

Dans chacun des patches, toujours pour tenir compte de la faible portée de la covariance spatiale, on a décidé de modéliser les matrices P_i par des matrices parcimonieuses avec une structure particulière pour les coefficients non nuls. On suppose que pour chaque pixel, les seuls coefficients non nuls sont ceux correspondant à la covariance avec ses quatre plus proches voisins. Ainsi, chaque ligne de la matrice P_i ne contient au plus que cinq coefficients non nuls.

1.3 Densité du modèle retenu

Une fois les hypothèses sur le modèle fixées, on a pu exprimer sa densité. Pour un champ Y gaussien de matrice de covariance diagonale par blocs avec une structure particulière dans chacun des blocs,

$$f(Y) = \prod_{i=1}^R \frac{1}{\det(P_i)^{\frac{1}{2}}} \exp\left(-\frac{1}{2} {}^t Y_{|i} P_i^{-1} Y_{|i}\right)$$

où $Y_{|i}$ représente la restriction du champ Y au patch $i^{\text{ème}}$ patch.

2 Test du rapport de vraisemblance

Les hypothèses mises en jeu lors de la détection sont les suivantes:

$$\begin{aligned} H_0 &: \text{Absence de cibles dans le pixel } (i, j) \\ H_1 &: \text{Présence de cibles dans le pixel } (i, j). \end{aligned}$$

Tout au long de nos travaux, les cibles sont supposées sous-résolues et additives c'est à dire que dans un pixel contenant une cible on observe la valeur du champ X auquel s'ajoute la puissance de la cible.

2.1 Théorie de Neyman-Pearson

Nous avons souhaité effectuer la détection tout en garantissant le taux de fausses alarmes. La théorie de Neyman-Pearson [4] affirme que le test du rapport de vraisemblance est le test le plus puissant (plus fort taux de détection) parmi tous les tests garantissant le taux de fausses alarmes. C'est donc lui que l'on applique. On note $T_{RV}(i, j)$ le rapport de vraisemblance. Pour un taux de fausses alarmes α souhaité, le test au pixel (i, j) est alors défini par

$$T_{RV}(i, j) = \frac{f_1(Y)}{f_0(Y)} > s \text{ où } P(T_{RV}(i, j) > s | H_0) = \alpha.$$

Comme nous avons choisi de travailler patch par patch, le test du pixel (i, j) ne fait en fait appel qu'à la restriction du champ Y au patch contenant le pixel (i, j) , on la note $Y^{i,j}$. Le test devient alors

$$T_{RV}^{\text{zone}}(i, j) = \frac{f_1(Y^{i,j})}{f_0(Y^{i,j})} > s \text{ où } P(T_{RV}^{\text{zone}}(i, j) > s | H_0) = \alpha.$$

Il nous faut alors connaître la loi du champ $Y^{i,j}$ sous chacune de ces hypothèses afin de construire le test.

En terme de loi du champ X , les hypothèses se traduisent de la façon suivante :

$$\begin{aligned} H_0 &: X^{i,j} \sim N(m, \Sigma^{i,j}) \\ H_1 &: X^{i,j} \sim N(m + c\delta_{i,j}, \Sigma^{i,j}). \end{aligned}$$

où $\delta_{i,j}$ est un champ de même taille que $X^{i,j}$ contenant un 1 en position (i, j) et des 0 partout ailleurs.

Nous travaillons avec le champ Y qui est une version filtrée du champ $X : Y = f(X)$. Les hypothèses pour le champ Y deviennent alors:

$$\begin{aligned} H_0 : Y^{i,j} &\sim N(f(m), P^{i,j}) = N(0, P^{i,j}) \\ H_1 : Y^{i,j} &\sim N(f(m + c\delta_{i,j}), P^{i,j}) = N(cM_{i,j}, P^{i,j}). \end{aligned}$$

où $M_{i,j} = f(\delta_{i,j})$ est un champ de même taille que $\delta_{i,j}$ contenant un 1 en position (i, j) , le coefficient $-\frac{1}{4}$ aux positions $(i-1, j)$, $(i+1, j)$, $(i, j-1)$, $(i, j+1)$ et des 0 partout ailleurs. L'impact du filtre f sur la présence d'une cible au pixel (i, j) est représenté sur la figure 1.



Figure 1: Transformation de la localisation de la cible après application du filtre gradient

Les densités sous chacune des deux hypothèses sont alors de la forme suivante

$$\begin{aligned} f_1(Y^{i,j}, P^{i,j}, c) &= \frac{1}{\det(P^{i,j})^{\frac{1}{2}}} \exp\left(-\frac{1}{2} {}^t(Y^{i,j} - cM_{i,j})(P^{i,j})^{-1}(Y^{i,j} - cM_{i,j})\right) \\ f_0(Y^{i,j}, P^{i,j}) &= \frac{1}{\det(P^{i,j})^{\frac{1}{2}}} \exp\left(-\frac{1}{2} {}^tY^{i,j}(P^{i,j})^{-1}Y^{i,j}\right) \end{aligned}$$

2.2 Expression du test

Nous sommes alors capables de calculer le test du rapport de vraisemblance. Comme les paramètres c et P sont inconnus, on utilise plutôt le test de rapport de vraisemblance généralisé (dans lequel on a remplacé les paramètres inconnus par leur valeur estimée):

$$T_{RVG} = \frac{f_1(Y^{i,j}, \widehat{P}^{i,j}, \widehat{c})}{f_0(Y, \widehat{P}^{i,j})} \text{ où } (\widehat{P}^{i,j}, \widehat{c}) = \underset{P,c}{\operatorname{argmax}} f_1(Y, P, c)$$

où la recherche de maximum porte sur toutes les matrices P ayant la structure parcimonieuse présentée plus haut et c un réel positif.

D'un point de vue pratique, l'estimation de $P^{i,j}$ n'a pas été évidente. On a appliqué un algorithme de Burn [1] qui permet de construire de manière itérative l'estimateur du maximum de vraisemblance sous la bonne contrainte de la structure.

En utilisant l'expression des paramètres estimés dans la construction du test, on obtient l'expression suivante pour la statistique de test:

$$T_{RVG}(i, j) = \exp\left(\frac{({}^tY^{i,j}\widehat{P}^{i,j}{}^{-1}M_{i,j})^2}{2({}^tM_{i,j}\widehat{P}^{i,j}{}^{-1}M_{i,j})}\right)$$

Comme la fonction $x \rightarrow \exp(\frac{x}{2})$ est strictement croissante, le test que nous avons mis en place en pratique est en fait le suivant :

$$\frac{({}^t Y^{i,j} \widehat{P}^{i,j}{}^{-1} M_{i,j})^2}{{}^t M_{i,j} \widehat{P}^{i,j}{}^{-1} M_{i,j}} > s \text{ où } P\left(\frac{({}^t Y^{i,j} \widehat{P}^{i,j}{}^{-1} M_{i,j})^2}{{}^t M_{i,j} \widehat{P}^{i,j}{}^{-1} M_{i,j}} \geq s \mid H_0\right) = \alpha$$

3 Classification et apprentissage du seuil

Le paramètre de seuillage s est un paramètre difficile à régler. Il correspond normalement à un quantile de la loi de $\frac{({}^t Y^{i,j} \widehat{P}^{i,j}{}^{-1} M_{i,j})^2}{{}^t M_{i,j} \widehat{P}^{i,j}{}^{-1} M_{i,j}}$ sous l'hypothèse H_0 mais celle-ci n'est malheureusement pas connue. Bien qu'elle soit asymptotiquement connue, des expériences numériques ont montré que cette approximation n'était pas suffisante.

3.1 Approche apprentissage

Nous avons donc ajouté à notre algorithme une étape d'apprentissage de seuils. Sur un échantillon test (composés de patches représentatifs des images réelles), nous avons alors estimé le seuil qui permettait de garantir le taux de fausses alarmes du test. Nous avons en fait pris le quantile empirique de cet échantillon. Cependant, ce choix de seuil n'a pas donné les résultats souhaités. La figure 2 permet de comprendre pourquoi. Elle représente les valeurs moyenne du rapport de vraisemblance selon la texture dans le cas d'absence de cibles.

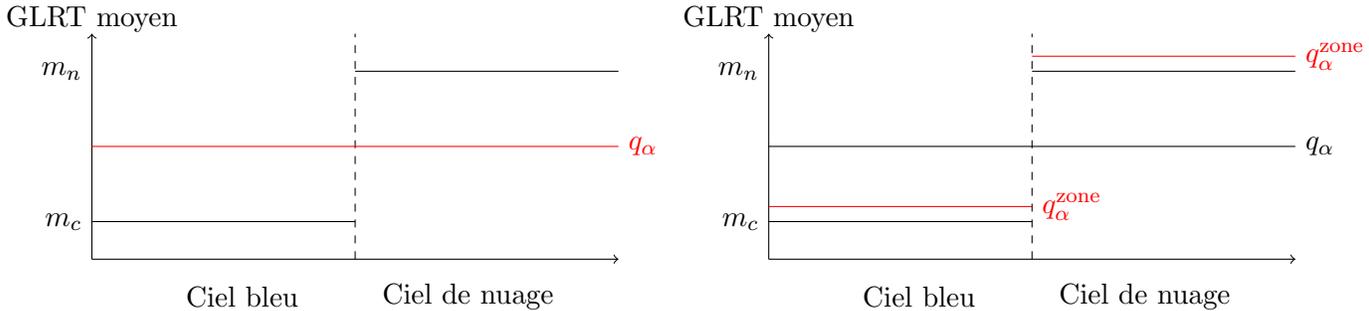


Figure 2: Impact d'une estimation texture par texture

Supposons que l'on se fixe un niveau α et que l'on estime un seuil, noté q_α sur tout l'échantillon. La présence d'une cible va entraîner une augmentation de la valeur du rapport de vraisemblance. Dans la zone de ciel, si cette valeur se trouve entre m_c et q_α elle ne sera pas détectée. En revanche, si on estime un seuil par texture, q_α^{zone} , celui sera plus faible que q_α dans la zone de ciel et plus élevé dans la nuage de nuage. Le taux de détection sera donc augmenté dans la zone de ciel et le taux de fausses alarmes sera diminué dans la zone de nuage.

En pratique, au lieu de faire l'estimation du seuil sur l'échantillon de test tout entier, nous avons d'abord effectué une classification des patches puis nous avons estimé un seuil par classe.

3.2 Approche combinée classification et apprentissage

L'hypothèse sous-jacente est qu'il existe des classes homogènes pour le choix du seuil. Pour les obtenir, nous suivons une approche de type classification non supervisée à l'aide de mélange de gaussiennes. Les descripteurs utilisés ne sont pas les valeurs des pixels mais ceux plus invariants par transformation locale obtenus par Scattering [3].

Une fois les N patches classés, nous avons alors pu estimer un quantile par classe et ainsi obtenir une table seuil contenant autant de ligne que de classes et autant de colonnes que de taux de fausses alarmes souhaités. Cette table de seuils sera une des données interne de notre algorithme.

Conclusion

Le schéma suivant représente le fonctionnement final de l'algorithme proposé.

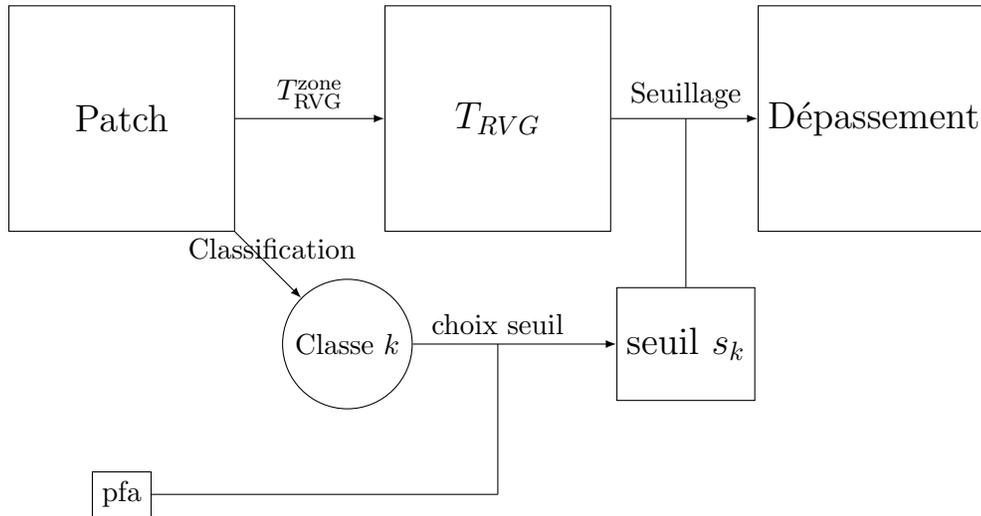


Figure 3: Schéma final de l'algorithme

Bibliographie

- [1] J. Burg, D. Luenberger, D. Wenger (2000), Estimation of structured covariance matrices, Proceedings of the IEEE , 70(9).
- [2] J. Bruna, and S. Mallat (2013), Invariant scattering convolution network, IEEE trans. , 25(8).
- [3] P. Massart. (2003), *Concentration inequalities and model selection*, Springer
- [4] Lévine. (1973), *Fondements théoriques de la radiotechnique statistique*, MIR